

# DILEMMATA

## Jahrbuch der ASFPG

Herausgegeben vom Vorstand der  
Altonaer Stiftung für philosophische Grundlagenforschung

2. Jahrgang  
2007



## Inhaltsverzeichnis

What Ought We to Know About Science and Technology? Or: Philosophy of Science and Science Studies as Science Literacy RAGNAR FJELLAND	1
Commodity Fetishism: Marx's Dialectic of Content and Form VALERIE KERRUISH	19
Violence and Negativity in Hegel's <i>Phenomenology of Spirit</i> TARIK KOCHI	57
Begründung der Metaphysik durch logische Analyse der Selbstbezüglichkeit UWE PETERSEN	77
The Laws of Irresponsibility SCOTT VEITCH	171
Auslobung	189
Offer of a Prize — Call for Entries	190



# What Ought We to Know About Science and Technology? Or: Philosophy of Science and Science Studies as Science Literacy

RAGNAR FJELLAND\*

ABSTRACT. Although the concept of scientific literacy was developed in the nineteen fifties as a goal for science education, it was put on the agenda again in the nineteen nineties. The proponents of scientific literacy argued that the most important problem today is that the level of knowledge of science and technology is far too low in the population at large, and among school children and students in particular. Although some of the alarming reports may be questioned, it is widely accepted that the situation is disturbing. If the general public lacks the knowledge needed to have reasonably well-founded opinions about important scientific and technological issues, it will be a problem for democracy. The paper will argue that promoting scientific literacy is an important aim of philosophy of science. However, philosophers of science should not be just public relations agents for the sciences. On the contrary, it is imperative that they take a critical look at modern science. It will further be argued that a historical perspective is important in pursuing this goal, and that philosophy of science can learn some important lessons from science studies.

## 1. Introduction: scientific literacy

The term “scientific literacy” was introduced in the 1950s, and is regarded as a timeless goal for science education. One description of scientific literacy was given by the American Association for the Advancement of Science in the document *Science for All Americans*. In this document a scientifically literate person is described as one who is aware that science and technology are human enterprises with strengths and limitations, un-

\* Center for the Study of the Sciences and the Humanities (SVT), University of Bergen, Allég. 32, 5020 Bergen, Norway. A previous version of this paper was presented as an invited plenary lecture at the conference *New Perspectives on the relation between history and philosophy of science*, joint meeting of the Division of History of Science and the Division of Logic, Methodology and Philosophy of Science, International Union of History and Philosophy of Science; Paris, 3.10.2002–5.10.2002.

derstand key concepts and principles of science, is familiar with the natural world and recognizes both its diversity and unity, and applies scientific knowledge and skills for individual and social purposes (AAAS, 1990).

Scientific literacy was again put on the public agenda in the 1990s, partly related to the conflict that has come to be known as “the Science Wars”. It started when the biologist Paul R. Gross and the mathematician Norman Levitt published the book: *Higher Superstition. The Academic Left and Its Quarrels with Science* (1994). The book was a fierce attack on certain quarters within the history of science, philosophy of science and sociology of science, such as existentialism, phenomenology, postmodernism, feminism, multiculturalism and so on. The next year, 1995, the book was followed up with a conference in New York given by the New York Academy of Sciences titled *The Flight from Science and Reason*. The conflict gained momentum when the physicist Alan Sokal published the article “Transgressing the Boundaries: Towards a Transformative Hermeneutics of Quantum Gravity” (1996) in a journal for cultural studies, *Social Text*. Soon after the article was published, Sokal revealed that the entire thing had been a hoax. He had intentionally written an article that contained a lot of nonsense, however it was written using fashionably correct terminology with references to a range of “postmodern” thinkers. The hoax gained worldwide publicity, and many of the participants in the debate have claimed that this debate shows that C. P. Snow’s “two cultures” still exist.

The proceedings of the conference *The Flight from Science and Reason* were published in an anthology with the same title as the conference. In his introduction Paul R. Gross sums up the main concerns of the organizers:

We believe that there is today in the West, among professors and others who are paid, in principle, to think and teach, a new and most systemic flight from science and reason. It is given endless and contradictory justifications; but its imperialism — for example under the banner of “science studies” — and the high esteem in which it holds the trendiest irrationalisms, are undeniable. This has brought with it, from that unexpected academic quarter, a truculent defense in the name of “democracy” of New Age and traditional forms of sophistry and charlatanism. Younger programs of anti-logic and anti-science are both diffuse (oppositional

movements being by their nature fractious) and angrier than the siege, already a few decades old, of “objectivity” in the social sciences and humanities (Gross, Levitt, Lewis 1997, p. 2).

This concern was not without foundation, and it is still a cause for concern. Sometimes we receive alarming reports on the low level of scientific literacy in the population at large. “Anti-science” and “alternative science” movements, like creationism, astrology and healing, to take a few examples, are so widespread that they cannot be neglected. Some of the most alarming reports are probably exaggerated, but according to a conservative assessment of the situation there is no doubt that the knowledge of and interest in science and technology do not increase at the same rate as the significance of science and technology in society.

This must be a cause for concern for the authorities, and therefore, “scientific and technological literacy” (STL) and “public understanding of science and technology” (PUST) have been put at the top of the political agenda. In USA president Bill Clinton made it one of the main political issues in his second term in office.<sup>1</sup> Although there is sometimes a considerable discrepancy between word and action, all politicians in industrialized countries (and most other countries as well) maintain that scientific and technological literacy has the highest priority. However, the motivations may be different. Most people can agree that scientific and technological literacy in the population is a prerequisite for keeping up productivity and as a guarantee for a democratic development. If the population by and large lacks the knowledge required for making well-founded decisions on important issues related to science and technology, it will undermine the democratic control of the development.

Some of the advocates of scientific and technological literacy go further, though. They maintain that all public scepticism and resistance against controversial science and technology, like nuclear power and genetically modified food, is based on superstition and ignorance. Therefore, more information and better knowledge will change the negative public opinion. *The Flight from Science and Reason* is one example. The articles follow up the introduction quoted previously with attacks on disciplines that are perceived as representing irrational tendencies in academic life, such as existentialism, phenomenology, feminist epistemology and deep-ecology. The authors probe deeply and also find targets to attack deep

<sup>1</sup> Needless to say, the priorities of his successor have been different.

within the established sciences. That Ilya Prigogine would be criticized might be expected, but Werner Heisenberg and Niels Bohr also come under criticism. It is difficult to avoid having the impression that the proponents maintain a rigid and orthodox notion of science, where everything that does not fit into a narrow framework is considered to be “anti-science” and “irrational”.

## 2. Ronald Giere on scientific literacy

If we look at the American Association for the Advancement of Science’s description of scientific literacy, the interesting thing is that it does not concentrate so much on scientific facts, but rather emphasizes higher levels of cognition, such as critical thinking.

When scientific literacy is taken in this sense, no doubt philosophers of science have promoted scientific literacy. Obvious examples are the logical positivists and Karl Popper. An important goal for their activity was the dissemination of the scientific attitude to all areas of society.

At least one introductory textbook in philosophy of science has the explicit aim of promoting scientific literacy. The book is Ronald Giere, *Understanding Scientific Reasoning*. This is a widely used introductory text at the undergraduate level. The book is supposed to give the basics of scientific reasoning:

The primary answer to the question, “why study scientific reasoning?” is that it will help you to be better able to understand and evaluate scientific information in both your personal life and your work (Giere 1991, p. 4).

It will not enable the reader to do science in the laboratory, but to evaluate scientific information in a more critical way:

For the purposes of this text, then, learning to understand scientific reasoning is a matter of learning how to understand and evaluate reports of scientific findings of the type one would find in a popular magazine, a national newspaper, or a news magazine. This requires only a very general idea of what goes on in laboratories. And it does not require the skills that are necessary to do laboratory research (*ibid.*, p. 5).



The book consists of three parts, dealing with models, statistics and theory of decisions respectively.

The virtue of the book is that it is closer to real science than many other introductory textbooks in the philosophy of science. For example, it deals with randomized, prospective and retrospective trials, and the notion of cause as used in epidemiology. The reader is supposed to learn what it means that, say, smoking causes cancer. It uses examples from real-life science, and has many good examples. There are no theoretical discussions, though.

I said that the explicit aim of the book is to promote scientific literacy. And one aspect of the book places it firmly in this tradition: sources of error, ignorance and irrationality are all placed outside science. The two last editions of the book have a chapter with the title “Marginal science”. In this chapter psychoanalysis, astrology, clairvoyance, von Däniken and New Age phenomena are dealt with. There is nothing wrong with this. But it is a striking fact that the book does not address the problem of errors within science, or the abuse of science. If the aim of the book is to promote scientific literacy, this is a serious omission. It is a much more serious possibility that the students who read the book will one day be in the position to abuse science than that they will be the victims of, say, astrology.

It is interesting to notice that in the description given by the American Association of the Advancement of Science quoted previously, it is emphasized that the scientifically literate person is aware that science and technology are human enterprises with strengths and limitations. Learning about the limitations of science is as much a part of scientific literacy as is learning about the strengths of science. Two books which might serve this purpose are Stephen Jay Gould’s *The Mismeasure of Man* and Kristin Shrader-Frechette’s *Risk and Rationality*.

### 3. A different approach to scientific literacy

A quite different book, which also has the explicit aim of promoting scientific literacy, is Harry Collins and Trevor Pinch: *The Golem: what everyone should know about science*:

The point is that, for citizens who want to take part in the democratic process of a technological society, all the science they need to know about is controversial . . . (Collins and Pinch 1993, p. 3).

Therefore, the main task of the book is to remove science from its “Platonic heaven”. The main strategy is to show that scientific results are the product of disagreement, uncertainty, compromise, power struggle and decisions. In other words: it is a social construction. The problem of traditional history of science (and here they follow Kuhn) is that this is mostly left out, leaving the false impression that scientific “truths” prevail with the inevitable force of logical necessity.

In the book seven historical cases are examined. One is the theories of relativity. Another is cold fusion. After having dealt with the cases the authors draw the following conclusion:

. . . we have shown that scientists at the research front cannot settle their disagreements through better experimentation, more knowledge, more advanced theories, or clearer thinking. It is ridiculous to expect the general public to do better.

We agree with the public understanders that the citizen needs to be informed enough to vote on technical issues, but the information needed is not about the content of science; it is about the relationship of experts to politicians, to the media, and to the rest of us (Collins and Pinch 1993, p. 145).

However, Collins and Pinch go further:

We have no reason to think that relativity is anything but the truth — and a very beautiful, delightful and astonishing truth it is — but it is a truth which came into being as a result of decisions. [...] . . . it was a truth brought about by agreement to agree about new things. It was not a truth forced on us by the inexorable logic of a set of crucial experiments (ibid., p. 54).

I have already mentioned that two of the cases in the book are the theories of relativity (the special as well as the general) and cold fusion: one success story and one failure. However, according to the book it looks as if the only difference between the success of the theory of relativity and cold fusion was that the first was accepted whereas the second was rejected by the scientific community.

Collins and Pinch need not go that far, though. The important thing is to show that the results of science are not inevitable, they might have been different. I think that Ian Hacking is right when he gives the following characterization of social constructivists. Social constructivists about X tend to hold that:

- (1) X need not have existed, or need not be at all as it is. X, or X as it is at present, is not determined by the nature of things; it is not inevitable.

Very often they go further, and urge that

- (2) X is quite bad as it is.
- (3) We would be much better off if X were done away with, or at least radically transformed (Hacking 1999, p. 6).

For X we substitute “science” (or “science and technology”).

#### 4. The importance of idealization

Collins and Pinch have later modified their position. However, I shall use their original position to show why I cannot follow them all the way, and I shall start with an example from their book.

The example is an exercise to teach elementary school pupils to measure the boiling point of water. The pupils are told to put their thermometers into a beaker of water and read the temperature when the water boils. Hardly any of the pupils obtain the result 100 °C if they do not already know the answer. In the example used in the book the results are like this: Skip gets 102 °C, Tania gets 105 °C, Johnny gets 99,5 °C, Mary gets 100,2 °C, Zonker gets 54 °C, whereas Brian does not obtain any result. Smudger boils his beaker dry, and bursts his thermometer. Ten minutes before the end of the lesson the teacher gathers all the pupils and starts the “social engineering” process: Skip held his thermometer in a bubble of superheated steam when he made his reading, Tania had impurities in her water, Johnny did not wait until the water boiled, Mary’s result demonstrates the effect of slightly higher air pressure, and Zonker, Brian and Smudger have not yet acquired the required competence. After this lesson all the pupils are convinced that they have demonstrated that the boiling point of water is 100 °C, or they would have demonstrated it if

there had not been a few local problems. According to Collins and Pinch this simple exercise demonstrates the essence of science:

In the end, however, it is the scientific community (the head teacher?) who brings order to this chaos, transmuting the clumsy antics of the collective Golem Science into a neat and tidy scientific myth. There is nothing wrong with this; the only sin is not knowing that it is always thus (Collins and Pinch 1993, p. 151).

However, if we look closer at the example, it is worth noticing that something is missing: In each case a specific explanation is given for why the pupil did not obtain the exact value  $100^{\circ}\text{C}$ . For example, it is said that Skip got  $102^{\circ}\text{C}$  because he put the thermometer in a bubble of superheated vapor. But the crucial question is: Are these explanations correct? Would all the pupils have obtained  $100^{\circ}\text{C}$  if they had carried out their measurements correctly under ideal conditions, or is this something that their teacher tries to make them believe? We know that the answer to that question is yes, that the pupils would indeed have obtained the result  $100^{\circ}\text{C}$ .

The lesson of this example is not that science is a social construction, but that the result can only be obtained under special, idealized conditions. To understand an important aspect of modern science we have to understand the importance of idealization. In fact, we learned this from Alexandre Koyré, who pointed to what he called the Platonism of modern science. In his *Galileo Studies* he draws a line from Pythagoras and Plato, via Archimedes and to Galileo. We might extend this line to Einstein and Hawking.

The basic problem which Platonism tries to solve, is the question if and how certain knowledge can be obtained. The Platonist answer is: through mathematics. Plato's theory of knowledge was inspired by geometry as the paradigm of knowledge, and according to Galileo "the book of nature" is written in the language of mathematics. However, there is an important difference between Plato on the one side and Galileo and modern science on the other. Whereas Plato's reality was immaterial, Galileo's reality was material. Galileo called objective reality "primary sense qualities". The essential property of matter is that it can be described mathematically.

Galileo recognized that a mathematical description requires measurements, and that measurements require controlled laboratory experiments.

The aim of the controlled laboratory experiment is to keep all or most factors constant. Only one or a few factors are varied at a time. These ideal conditions increase *certainty*. According to the traditional view, controlled experiments are merely simplification and purification of natural situations. We have to leave out some factors to make the problems manageable. Afterwards we “add back” the factors that were left out, and in this way we come closer to natural situations.

However, we do not only remove complicating factors. We *impose* artificial conditions on the object as well, because the ideal conditions are normally not realized in everyday life. Therefore, “adding back” may not be an easy task. There is an alternative, though. We may realize the ideal conditions through technology. From this point of view technology is a way of reducing uncertainty. It is interesting to note that Galileo was aware of the intimate relationship between the ideal conditions required to carry out experiments, and technology. In *Dialogue Concerning Two New Sciences* he pointed out that his own results had been proved in the abstract, and when applied to concrete cases they would yield false results. The horizontal motion would not be uniform, a freely falling body would not move according to the law, and the path of a projectile would not be a parabola. However, speaking of the difficulties arising by these limitations, he immediately adds:

... in order to handle this matter in a scientific way, it is necessary to cut loose from these difficulties; and having discovered and demonstrated the theorems, in the case of no resistance, to use them and apply them with such limitations as experience will teach. And the advantage of this method will not be small; for the material and shape of the projectile may be chosen, as dense and round as possible, so that it will encounter the least resistance in the medium (Galileo 1638/1954, p. 251).

To put it simply: technology to a large extent realizes the ideal conditions of the laboratory at a larger scale. We can see how this works in biotechnology by looking at the so-called “green revolution” of the 1960s as an example. It involved the development and introduction of new plant species that gave larger yields per acre than did the traditional species. This first happened with wheat in Mexico, and later with wheat and rice in Asia. The new “high-yielding varieties” could make better use of fertilizer with far higher concentrations than traditional varieties, and they

had noticeably faster maturation rates. One crucial factor was that the plants receive the correct amount of watering at the correct time. They were also less resistant against a number of diseases and parasites. To summarize: if the new variants were to give higher yields, a “*technological package*”, in the form of the correct amounts of fertilizer, water, pesticides and plant protection products was needed. If these things are not in place, the process can go wrong. The desired effects can in other words only be achieved if one also has control over the environment. What was needed to carry out the “green revolution”, was to realize *controlled laboratory conditions in agriculture*.<sup>2</sup>

## 5. Recognizing uncertainty

Although we can control parts of nature in this way, the inescapable problem is, however, that there will always be something outside the system we control. A factory is a typical example of a controlled system. However, the control is normally far from perfect. First, the production process itself is full of risks, for example the risk of explosions and chemical hazards. Second, there is the area around the factory. Traditionally this was heavily polluted. Although regulations have reduced local pollution, the problem is often moved to other places. In particular, heavily polluting production is often moved from the rich countries to third world countries, where regulations are absent or less strict. Third, we have the uses of the products and the disposal of the worn-out products, and so on. Therefore, when an area is subject to technical control, there is always a large area which escapes control.

In what follows I shall use the term “natural conditions” in contrast to the simplified and idealized conditions of the laboratory and the factory. However, the word “natural” does not imply that the conditions are prior to human intervention.

Then we return to the problem of “adding back” from the simplified and idealized conditions of the laboratory to natural conditions. This

<sup>2</sup> Cf. the following quotation from Ian Hacking: “In fact, few things that work in the laboratory work very well in a thoroughly unmodified world — in a world which has not been bent toward the laboratory” (Hacking 1992, p. 59). Hacking refers to Latour 1987.

problem has been recognized by ecologists. Therefore, laboratory experiments have limited value in ecology because the artificial conditions sometimes prevent important natural effects from appearing and may magnify incidental and trivial effects. To quote an ecologist:

Laboratory studies are effective in isolating a response to a factor but the response may not be ecologically relevant and the number of potential factors that could be investigated is so large that the study of any isolated factors may be futile (Peters 1991, p. 138).

If laboratory experiments fail, field experiments might do the job. They are something between a laboratory experiment and the natural system. Because they are closer to the natural systems, they are popular in ecology. However, it looks as if we get nothing for free. There is a trade-off between control of the conditions on the one hand and relevance to natural situations on the other: The better the field experiments, the less relevant they are.

The problem of uncertainty may also be formulated in the language of risk assessment. We must (at least) distinguish between two different situations: uncertainty and ignorance. When we have uncertainty, it means that we know what can go wrong. (When we also know the probabilities, we are talking about risk.) However, there are often situations where we have no idea of what can go wrong. These situations are characterized by *ignorance*. In risk assessment it is desirable to reduce uncertainty to risk, because it enables the application of the mathematical methods of risk analysis (probability theory, statistics and the like). This requires simplification and idealization, either in the form of experiments as described earlier or by applying mathematical models. However, we have a similar problem as in the case of ecology: the reduction of uncertainty may increase ignorance (Wynne 1992, p. 114).

Mathematically speaking the problem is nonlinearity. The mathematical sciences have since the time of Galileo largely concentrated on linear or approximately linear systems. One reason is that the analytical tools of mathematics can be used. However, when the interactions between the parts of a system or the factors determining a process are nonlinear, the situation is changed. This was first observed in chaos. Chaos is characterized by *sensitive dependence on initial conditions* (the “Butterfly effect”): small uncertainties in the determination of the initial conditions of a system may increase exponentially until they are the same magnitude as the

parameters of the system. In that case “adding back” does not work, and predictions are limited.

However, organic nature is in general not chaotic, but complex. Complexity and chaos are not the same. It is sometimes said that complexity arises at “the edge of chaos”. But they do have nonlinearity in common, rendering impossible both “adding back” and exact predictions.

Is uncertainty genuinely new in science? The answer is simply “no”. From the very beginning of philosophy and science there were alternative schools of thought that emphasized uncertainty. The contemporaries of Plato, the Sophists, and even Plato’s own teacher, Socrates, stressed both uncertainty and ignorance. Therefore, we can draw an alternative line, from Socrates, via the Renaissance Humanism of Erasmus and Montaigne, to the present situation (cf. Toulmin 1990).

In a certain sense the two aspects, certainty and uncertainty, are combined in the theories of probability and statistics. Furthermore, the recognition of uncertainty is the very foundation of one of the most basic and influential theories of contemporary science, quantum mechanics.

However, what is genuinely new today is the recognition that uncertainty cannot be tamed or ignored. Previously, unintended side-effects of industrial production that were outside our control could to a large extent be ignored. However, the global character of some environmental problems has shown that there is no “outside”: the biosphere is finite. Therefore, scientists and technologists have in many ways come into a new situation. The Chernobyl accident is a dramatic example, however problems such as a possible global warming, a possible reduction of the ozone layer, and so on are all of the same type. These encompass totally different problems than scientists and technologists are traditionally trained to deal with (Funtowicz and Ravetz 1991, p. 85).

## 6. Some consequences for scientific literacy

Recognizing the importance of idealization in science, the complexity of nature and the irreducibility of uncertainty has consequences for scientific literacy. I shall restrict myself to dealing with one problem, the uses of idealized models.

Experts trained in a field have a tendency to apply the kinds of models that conform with their field. The following example has been taken from



Brian Wynne's "Uncertainty and Environmental Learning": In May of 1986 a cloud of radioactive material from the Chernobyl accident passed over Cumbria in North Wales. Heavy rains caused a large amount of radioactive cesium to fall over an area used to raise sheep. The authorities in charge assured everyone that there was no cause for concern, but in spite of this, six weeks after the rains a ban against selling meat from sheep that had grazed in the area was imposed because of the high levels of radioactivity found in the meat. Experts claimed however that the radioactivity would rapidly decrease, and that the ban would be lifted in a few weeks. Yet, even after six years the level of radioactivity was so high in some of the affected areas that restrictions had to be upheld.

How could the experts be so wrong? Their predictions were based on extrapolations from the behavior of cesium in alkaline clay soil to the acid peat soil of Cumbria. Measurements showed that the dispersion of cesium in these types of soil was fairly similar, and on that basis they assumed that cesium would sink so far down into the ground that after a short period of time there would be no problem. This was based on the assumption that the radiation would come from the cesium in the soil and would be absorbed by people or animals who happened to be in the area. Under this assumption it was the physical transport of cesium in the soil that was important. However, this assumption was wrong. The sheep got cesium in their bodies through the grass they ate. The important question was therefore not how the cesium was dispersed throughout the soil but if it was absorbed into the vegetation. Here there proved to be a significant difference between alkaline clay soil and acid peat. In alkaline clay soil cesium adsorbs onto aluminum silicate molecules so that it does not get absorbed into the vegetation, whereas in peat it remains chemically mobile and can therefore be taken up into the vegetation. The experts did not consider these possibilities, and that was the cause of their mistaken predictions (Wynne 1992, p. 121).

Should not a model that took into consideration for example chemical properties, have been used at the onset? The answer is, of course, yes. But to understand why the experts made such an apparently elementary error we have to take into consideration that they had been trained as physicists. Physicists are used to think in terms of physical transportation and radiation. Chemists are trained to think in terms of chemical reactions and chemical mobility. The problem is that it is not a part of professional

training to learn about the limits of the models and methods of a field. This is the neglected part of scientific literacy.

In particular, situations involving complex systems are new to most researchers. In the mathematical sciences one is trained to deal with idealized situations and use simple models. The physicist Per Bak tells a story to demonstrate how inadequate this way of thinking may be:

The obsession among physicists to construct simplified models is well illustrated by the story about the theoretical physicist asked to help a farmer raise cows that would produce more milk. For a long time, nobody heard from him, but eventually he emerged from hiding, in a very excited state. "I now have figured it all out," he says, and proceeds to the blackboard with a piece of chalk and draws a circle. "Consider a spherical cow . . ." Here, unfortunately, it appears that universality does not apply. We have to deal with the real cow (Bak 1997, p. 45).

"Extended peer communities" implies an extension of the traditional scientific community to include non-experts as well. However, this does not mean that non-experts should invade the research laboratories and carry out research. It does mean, though, that non-experts should take part in discussions of priorities, evaluation of results and policy debates.

The arguments in favor of extended peer communities are similar to Paul Feyerabend's arguments for a democratization of science.<sup>3</sup> I regard it as a continuation of an important element in the Socratic tradition. We know that it was part of Socrates' strategy to pretend that he was more ignorant than he actually was. By asking apparently naive questions to an expert one may reveal tacit assumptions which the expert himself is not aware of. Many scientists are sceptical of public debates about controversial scientific and technological questions, like nuclear power and genetically modified food, and allege that public opinion is often based on prejudices and lack of information. No doubt this is sometimes the case. But there are at least two reasons for not keeping these kinds of questions away from the public. First, non-experts may be wrong because they are prejudiced or lack the required information. But experts may also be wrong. Some of their errors may be corrected by bringing in non-experts. To put it simply: The public may be wrong because it is too far away from

<sup>3</sup>Cf. "Laymen can and must supervise Science" (Feyerabend 1978, p. 96).

the technical problems, whereas experts may be wrong because they are too close. The “tunnel” vision of experts is at least as great a problem as the ignorance of non-experts. The second reason is that common people are affected by the decisions which are made. The questions of global warming, the ozone layer, radioactive waste and genetically modified food concerns everybody, experts as well as non-experts. These questions are too important to be left only to the experts.

It can be argued that these consequences do not influence science and “the scientific method”, but only science policy. In a certain sense this is true, but it depends on what is meant by “the scientific method”. What is affected, is not science per se, but a dominating ideal of what science should be, emphasizing measurements, mathematics, idealized models, laboratory experiments, exact predictions and reductionism. When it is recognized that this scientific ideal is too narrow even for the mathematical (or “exact”) sciences, the toxicologist or ecologist should have few reservations against doing the same.

However, one might take one step further and argue that the root of uncertainty is complexity. Therefore, to come to terms with the new situation, a new science of complexity is required. An increasing number of authors argue in this way (for a small selection, see Nicolis and Prigogine 1989, Bak 1997 and Auyang 1998). This is an important question, but it goes beyond the scope of this paper.

## References

- Auyang, Sunny Y.: *Foundations of Complex-System Theories*, Cambridge: Cambridge University Press, 1998.
- Bak, Per: *How Nature Works. The Science of Self-Organized Criticality*, Oxford: Oxford University Press, 1997.
- Collins, Harry and Trevor Pinch: *The Golem: what everyone should know about science*, Cambridge: Cambridge University Press, 1993.
- Feyerabend, Paul: *Science in a Free Society*, London: NLB, 1978.
- Funtowicz, Silvio and Jerome Ravetz: *Uncertainty and Quality in Science for Policy*, Dordrecht: Kluwer, 1990.
- Funtowicz, Silvio and Jerome Ravetz: *The Emergence of Post-Normal*

Science, C. Rossi and E. Tiezzi (eds.): *Ecological Physical Chemistry, Proceeding of an international Workshop*, Amsterdam: Elsevier Science Publishers, 1991.

Funtowicz, Silvio and Jerome Ravetz: *Science for the Post-Normal Age, Futures*, September 1993: 739–755.

Galilei, Galileo: *Dialogue Concerning Two New Sciences* (1638), translated by Henry Crew & Alfonso de Salvio (1914), New York: Dover Publications, 1954.

Giere, Ronald N.: *Understanding Scientific Reasoning*, Fort Worth: Harcourt Brace Jovanovich, third edition 1991.

Gross, Paul R. and Norman Levitt: *Higher Superstition. The Academic Left and Its Quarrels with Science*, Baltimore, MD: Johns Hopkins, 1994.

Gross, Paul R., Norman Levitt and Martin W. Lewis: *The Flight from Science and Reason*, New York Academy of Sciences/Johns Hopkins University Press, 1997.

Hacking, Ian: The Self-Vindication of the Laboratory Sciences, in Andrew Pickering (ed.): *Science as Practice and Culture*, Chicago and London: The University of Chicago Press, 1992, pp. 29–64.

Koyré, Alexandre: Galileo and Plato, in A. Koyré: *Metaphysics and Measurement* (1943), Baltimore and London: Johns Hopkins Press, 1968.

Kuhn, Thomas: Postscript–1969, *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press, 1970.

Latour, Bruno: *Science in Action*, Milton Keynes: Open University Press, 1987.

Nicolis, Grégoire and Ilya Prigogine: *Exploring Complexity*, New York: Freeman and Company, 1989.

Peters, Robert Henry: *A Critique for Ecology*, Cambridge: Cambridge University Press, 1991.

Shrader-Frechette, Kristin: *Risk and Rationality*, Berkeley: University of California Press, 1991.

Sokal, Alan: Transgressing the Boundaries: Towards a Transformative Hermeneutics of Quantum Gravity, *Social Text*, Spring/Summer 1996.

Toulmin, Stephen: *Cosmopolis. The Hidden Agenda of Modernity*, New York: The Free Press, 1990.

Wynne, Brian: Uncertainty and environmental learning, *Global Environmental Change*, June 1992, pp. 111–127.



# Commodity Fetishism: Marx's Dialectic of Content and Form

VALERIE KERRUISH

## 1. Overview

Thought flies free of the mundane world. It flies free and constructs figures and fantasies, concepts and illusions. It leads and it misleads. It is in excess of the material, habitual, time-bound world of everyday life. This to my mind is why Marx adds the section on the fetish character of the commodity and its secret to his chapter on the theory of value. The thesis of the chapter, read as a whole, is that the double character of the commodity (the commodity form of the product of labour or the value-form of the commodity) gives rise to a substitution of this for that, as actual as it is unreasonable, which is detectable by asking after the historical and material conditions of the validity of formulae and expressions of value.

Marx saw *Capital* as a work of science and so do I. Whether good science or bad science; whether a world historical event that established a new scientific paradigm or fundamentally confused are judgements I leave to those who see some point in making them. What I think is significant in this context is that science, for Marx, is as much a product of social conditions as wealth or religion. Classical political economy, in his view, had already attained the status of a science. Social relations of production in industrialising Europe had evolved to that point at which the concepts necessary to put the study of political economy on a scientific basis had been grasped. Given however the class structure of these societies, the role of class struggle in processes of social change, and the location of mental labour within a division of labour separating mental and manual labour, Marx's persuasion is that the further advance of this science should take the form of its critique from a standpoint that is for the working class.

For my part returning to Marx on commodity fetishism is returning to a classical thinker of the social realm whose idea of fetishism has long struck me as implicated in the various ways in which the promises of legal

rights fail those most in need of them.<sup>1</sup> But having been persuaded in the meantime by Hegel's idea of thought's dialectical and speculative logical foundation it is a return in disagreement with Marx's idea that Hegel's logical dialectic stood wrong way up and around. Certainly Marx thought that scientific knowledge was conceptually mediated and structured but he maintained that what, as independent of such knowledge finds its objectivity, is not the logical forms which structure and contain human thinking, but the very nature of things which Hegel suspends in the artifice of constructing a purely logical realm.<sup>2</sup>

Quite apart however from the point that Hegel makes no notable contribution to political economy, he too is not to be taken whole. If it no longer makes epistemological sense to me to posit a material reality that is independent of human knowledge of it, I regard Hegel's enthronement of philosophical consciousness at the very summit of human self-awareness, as a conceit. As regards Hegel's idea of a logic that would replace metaphysics and the means and methods that might accomplish or realise the practical-theoretical task of so doing, it is a conceit to be set aside by recourse to mathematical methods and reasoning, that is to mathematical logic. From the premise then that Hegel's idea of thought's dialectical and speculative logical foundation, rid of its philosophical conceit, may gain a basis in constructive mathematical practice, my question became where Marx's thinking of the fetish character of the commodity, should be revised. My revisionary intent is to dissolve the opposition he supposes between ideal and material foundations of thought. The revision of Hegel's idea removes what I consider to be a justified objection to his thought that may well have motivated Marx's turn from it. My motivation for persisting with Marx's idea of the fetish character of the commodity is the belief that it is a profoundly speculative and dialectical thinking of the social realm. My sense, that in its flawed brilliance this idea remains, in some way, critical to thinking the social and cultural dimensions of law, stayed stubbornly in place. But in what way? And how might that be winked out of its lair in the complex fabric of Marx's desire, commitments and thought?

<sup>1</sup> See generally [Kerruish 1991].

<sup>2</sup> Referring here to [Hegel 1969] at 36; [Hegel 1812] at 26–7. Here, as with citations to Marx, I cite the English translation used with its date of publication followed by the German edition used with the original date of publication.



This paper documents a re-reading of the first chapter of *Capital* vol. I that has been made with these questions and aims in view. The following section ‘Preliminaries’ deals selectively with the first three sections of that chapter relating them back to observations made in the Postface to the second edition particularly as they concern Marx’s relation to Hegel. Marx’s engagement with commodity fetishism in the final section of his chapter is too often read in isolation from these earlier sections. Such readings miss the shift from analysis of what the double character of the commodity *is* to what it *does*. The third section of the paper, titled ‘The Fetish Character of the Commodity and its Secret’ after the final section of Marx’s chapter, moves off from this observation, reading Marx as re-traversing the abstract theory of the first three sections from the perspective of action: the action, or if you will agency of forms of value in perception and representation, in human efforts at deciphering the effects of their own incompletely intentional practices.

The final section of the paper, ‘Religion, Science and Ideology’ draws on the reading and commentary of the previous sections to argue that the famed analogy with religion, the flight to its misty realm to which Marx sees himself obliged, is a disastrous consequence of, on the one hand, his rejection of Hegel’s idea of thought’s logical foundation and, on the other a narrative of European social, scientific and cultural progress toward freedom which Marx inherited from Hegel and to which he trusted the emancipatory force of socialist revolution.

It will be said, going back the *Paris Manuscripts* and the use made there of a notion of fetishism; or going even further back to his Bonn days and the acquaintance then made with the term,<sup>3</sup> that the analogy with religion was always the leading idea: that the forced flight to that misty realm is all guise and that fetishism here as there is about objectification as externalisation and alienation of a human essence. That is the normative or ideological reading and I do not argue that Marx is not motivated by a mighty passion to change what he sees as changing and changeable in the interest of realising, in this world, his ideal of the social as being for others. What I will argue is that the theoretical character of the insight that connects the fetish character of the commodity to its double character

<sup>3</sup> Marx came across de Brosses’ *Du culte des dieux fetiches* in the course of intensive studies in the history of art and religion in Bonn in 1842. According to Pietz the term ‘Fetischismus’ was coined by de Brosses [Pietz 1993]; [Böhme 2001].

is lost to this move, effaced by what Marx assumes is unchanging and unchangeable: laws of thought that, as some people are still inclined to say, cannot be questioned because any (sane/reasonable/rational) questioning of them would have to use them.

My thesis is that in rejecting Hegel's logical dialectic, Marx rejects a notion of objectivity, the objectivity of illusion, which he then re-instates as fetish phenomena. These he names fetishism, via the analogy with religion and here, I shall argue, he loses an idea, an equivocation which, if left as an instantiation in the social realm of the objectivity of illusion, is both relevant and valuable for legal theory. Rejecting Hegel's logical dialectic Marx is disarmed in the face of his own perception of that very excess of thought that haunts logics of the understanding. What, it seems, Marx does not suspect or envisage, is that Hegel's questioning of the universality of so called laws of thought is an idea which will find its time in and through a new science of mathematical logic.<sup>4</sup>

This may seem to leave out of account the role played in Marx's thinking by laws of nature and the contrasting laws of social change which he imagines himself to have discovered. In part that is the fault of the attempted brevity of an overview. There is no doubt that Marx's engagement in the early chapters of *Capital* with the logic of a discourse, classical political economy, is marked as a work of nineteenth century science. I hope that will come out with the force proper to it in my reading. But there is a more substantial reason for taking the issue back to the diversity of formal logical systems which have emerged from the mathematical approach to logic. Whereas I cannot imagine that Marx, counterfactually projected into the twentieth century, would have rejected the new physics, I doubt that he would have moved from classical logic and set theory.<sup>5</sup> Classical logic is the logic of realist thought, Platonist or materialist. That cannot, I think, be disassociated from Marx's rejection, early and late, of Hegel's logical dialectic. Logically, Marx insists, particulars, whether objects or forms, must hold the grammatical place of a subject to which predicates ascribe properties. Indeed it is thus that Marx sees

<sup>4</sup> Such questioning, it is true, rarely connects to Hegel. It is however part of my enterprise to make that connection.

<sup>5</sup> While Alain Badiou in his *Being and Event* makes no gesture to this aspect of Marx's thinking, his re-writing of Zermelo–Fraenkel set theory as ontology, with its defence of classical logic, strikes me as consistent with Marx's logical classicism.

in expressions of value, the mystifying inversion that he attributes to Hegel's logic. Thus as regards the new logic, my hunch is that he would have made company with Zermelo and approved putting set theory on an axiomatic basis to block the antinomies that emerged within it and which, in their logical or property theoretic form, breathe life into Hegel's idea for a dialectical and speculative logic.

Revisions are always fraught with the risk and subject to the charge of having cut out just that moment in a thinking that makes it memorable. Where, as here, I am tampering with Marx's idea of material foundations of consciousness, a justification seems to be called for. Marx *cannot* accept that thought itself is at odds with itself because 'thought itself', for him, is thought that forgets its absolute dependence on the production and reproduction of material life. This absolute is just that which is equivocated, leadingly-misleadingly by fetish phenomena only then to be re-instated as 'fetishism' by the analogy with religion.

Bitter fruits! I take the view that the history of ideas is a comedy of errors with a dark side of crime that calls theory to account for the privileges that condition and enable its productions and performances. That takes me into the ironies, the conceits, the outrageous happenstances of place and time in that history, yet with the aim of recuperating objective insights and notions of objectivity from the present foundational malaise.

## 2. Preliminaries

The task Marx set himself in the section of his chapter on value immediately preceding that on commodity fetishism, "a task never even attempted by bourgeois economics", is to show the origin of the universal form of value in the simple, isolated or accidental form of value. The universal form has the entire universe of commodities on the left hand side of an equation and a universal equivalent on the right. The money-form of value is not formally different from it. It is preceded by the "total or expanded form of value" ( $z$  commodity A =  $u$  commodity B or =  $v$  commodity C or =  $w$  commodity D or =  $x$  commodity E or + etc) which, in turn is preceded by the simple form, 20 yards of linen = 1 coat; or  $5x = 2y$ . "The whole secret of the form of value lies hidden" in this simple form, Marx writes.<sup>6</sup> Its solution is within our reach, once we grasp the

<sup>6</sup> [Marx 1976] at 139; [Marx 1873] at 63.

point that human labour power “creates value, but is not itself value”. It *becomes* value in its congealed state in objective form (*gegenständlicher Form*), and that value can only be expressed as an ‘objectivity’ (*Gegenständlichkeit*): “a thing which is materially different to the linen itself and yet common to the linen and all other commodities”.<sup>7</sup>

The analysis of the simple form distinguishes its two ‘poles’, the relative form of value and the equivalent form. Discussing various peculiarities of the latter, Marx endorses and praises Aristotle’s early investigations of the value-form.<sup>8</sup> Aristotle’s genius, according to Marx, is to have seen the equality relation in the value expression of commodities. In particular, he accepts the idea that the equation of two different commodities in the simple value form (‘5 beds = 1 house’), since it involves things that are distinct to the senses, requires a qualitative equation and that this would not be possible in the lack of an essential identity. “There can be no exchange” he [Aristotle] says “without equality, and no equality without commensurability”.<sup>9</sup>

The Aristotelian assumption of a ‘third’ a common substance or identical essence which enables the commensuration of qualitatively unlike things is not inconsistent with identifying the act of commensuration in the practice of exchange or counting as equal. These are aspects of inquiry that answer different questions. The latter asks after a social practice, asks after what it is that people do and the conditions of that doing which at some stage or another results in a concept of value entering economic discourse. That is the perspective that Marx defers to the final section of the chapter. The former seeks truth conditions of expressions of equality. That is the issue here. In this discussion Marx, accepting the need for an essential identity as a condition of a valid equation, explains why Aristotle did not get far with his analysis of value: he could not grasp human labour power as the condition of commensurability in a society founded on the labour of slaves and concluded that the equation could “only be something foreign to the true nature of things . . . ‘a makeshift for practical purposes’”.<sup>10</sup> He lacked, so Marx, *a concept of value*, not for

<sup>7</sup> Ibid at 142; 65–6.

<sup>8</sup> Ibid at 151–2; 73–4.

<sup>9</sup> Ibid at 151; 73–4.

<sup>10</sup> Ibid.

want of genius, but because the concept of human equality was not yet fixed in popular opinion (*Volksvorurteil*: literally, ‘popular prejudice’).

The concept of value, according to Marx, answers the question “What is the same (*das Gleiche*), that is the common substance which, within the expression of value of the chairs [5 chairs = 1 house] the house represents for the chairs?”<sup>11</sup> It is labour, abstractly conceived as human labour; that which is ‘really’ or ‘in truth’ equal in the relation expressed. It is thus that the “power of abstraction” replaces microscopes and chemical reagents in the analysis of economic forms,<sup>12</sup> if only when social conditions permit. On the other and prior hand,

[l]abour ... as the creator of use-values, as useful labour, is a condition of human existence which is independent of all forms of society; it is an eternal natural necessity which mediates the metabolism between man and nature, and therefore human life itself.<sup>13</sup>

Useful labour, or better labouring, this concrete, sensuous, practical activity, will later in *Capital* in the context of the labour process, be characterised as an unrest (*Unruhe*).<sup>14</sup> It gives the sense in which concrete useful labour takes the place in Marx’s thinking of the unrest of thought at odds with itself in Hegel’s: of the restlessness of the negative, in Nancy’s phrase.

This is the sense in which Marx’s theory of value is a *labour* theory of value and a labour *theory* of value. He assumes, in addition to the metaphysics of the equality relation and the conditioning principle of concrete, useful labour, the achieved status of classical political economy as a science which investigates the “internal framework” of bourgeois production relations<sup>15</sup> and undertakes its critique. This assumption locates him

<sup>11</sup> Ibid at 151; 74.

<sup>12</sup> Ibid at 90; 12.

<sup>13</sup> Ibid at 133; 57.

<sup>14</sup> “Labour has become bound up in its object: labour has been objectified, the object has been worked on. What on the side of the worker appeared in the form of an unrest now appears, on the side of the product in the form of being [*Sein*] as a fixed immobile characteristic. The worker has spun and the product is a spinning” (ibid at 287: 195). The German ‘Er hat gesponnen, und das Produkt ist ein Gespinnst’ plays on figurative senses of ‘Gespinnst’ that take in lying and fabricating. These senses might better be captured for English readers by translating ‘Gespinnst’ as ‘spun yarn’.

<sup>15</sup> Ibid at 174 n.34; 95 n.32.

within an established discourse, with its conventionally established categories. He writes, in his Postface to the second edition of a more rigorous “derivation of value by analysis of the equations in which every exchange-value is expressed” making clear the discursive context of that analysis.<sup>16</sup> The aim (and accomplishment) of this opening chapter of *Capital* is to revise the categorial framework, value, use-value and exchange-value, thus their determinations and mutual relations, bringing his approach and methods with their implicit claims, theoretical and practical/critical, to that task.

Notoriously, in that same Postface, Marx writes a striking homage to Hegel, after forcefully declaring his method “in its foundations not only different from the Hegelian but exactly opposite to it”.<sup>17</sup> In general it seems to me that Marx has the idea of identifying the fetish character of the commodity in order to dispel a mystifying logic of the concept that is repeated in the value-expression. While the passage that most explicitly articulates this point appears only in an Appendix to the first edition which was subsequently dropped, the idea of the ‘inversion’ (*Verdrehung*) of the value-expression is kept.<sup>18</sup> But the basis of that undertaking is the analysis of the first three sections and that in turn proceeds from his decision to take the individual commodity as elementary or cell form

<sup>16</sup> Ibid at 94; 18.

<sup>17</sup> Ibid at 102; 27.

<sup>18</sup> Ibid at 150, 72 and see further below at p.31. The first edition passage reads as follows: “This *inversion* whereby the sensibly-concrete counts only as appearance-form of the abstractly-universal, and it is not to the contrary that the abstractly-universal counts as property of the concrete — this inversion characterizes the value-expression. At the same time it renders difficult its comprehension. If I say: Roman Law and German Law are both law, that is obvious. But if I say, on the other hand, *the Law* (this abstract entity) *realizes itself* in Roman Law and German Law (these concrete laws), then the connection becomes mystical” ([Marx 1867] at 56–57). Compare a younger Marx’s Feuerbachian objection to Hegel’s ‘subject-predicate inversion’ in his *Critique of Hegel’s ‘Philosophy of Right’*. “Had Hegel started with the real subjects as the bases of the state it would not have been necessary for him to let the state become subjectified in a mystical way. ‘However the truth of subjectivity’ says Hegel, ‘is attained only in a subject, and the truth of personality only in a person’. This too is a mystification. Subjectivity is a characteristic of subjects and personality is characteristic of the person. Instead of considering them to be predicates of their subjects, Hegel makes the predicates independent and then lets them be subsequently and mysteriously converted into their subjects” ([Marx 1970] at 23). Further references to this repeated criticism are given by O’Malley in his Introduction at xxxiii.

of the appearance of the wealth of capitalist societies as his ‘subject’ or object of analysis. It is a decision, I would say, which again reflecting the comments on Hegel in the Postface, answers the question, ‘With what should a critical science of political economy begin’.

We should go back to the remarkable opening sentence of *Capital*.

The wealth of societies in which the capitalist mode of production prevails appears as a ‘monstrously immense collection of commodities (*ungeheure Warensammlung*)’; the individual commodity appears as its elementary form.<sup>19</sup>

Marx quotes from the opening sentence of his 1859 *Zur Kritik der politischen Ökonomie*, where he gives the individual commodity as “*elementarisches Dasein*”. His decision to begin with the commodity is already reached, presumably as a result of the investigations (1857–8) comprising the *Grundrisse: Foundations of the Critique of Political Economy (Rough Draft)*, but the shift from *Dasein* (existence) to form attests further deliberation and the role of a form–content dichotomy in organising the text.<sup>20</sup> The individual commodity, singled out by theoretical artifice from the multiplicity, the “monstrously immense collection of commodities” in which the wealth of capitalist societies appears, and given as elementary or cell-form of that wealth knots quite some threads together. It is the content of a decision or judgement reached on the question of beginning. Described as “the simplest economic *concretum*”,<sup>21</sup> it is (relatively) concrete, a particular, in comparison with the universals of its various attributes use-value, value and exchange value. As such, and holding the place in the expressions of value of the (grammatical) subject, it is the basis of a dialectic that is ‘right way up and around’. And, subject(-matter) of a work written from a standpoint of viewing “the development of the economic formation of society . . . as a process of natural history”<sup>22</sup> it holds this formal place within a narrative pursuant on Marx’s rejection

<sup>19</sup> [Marx 1976] at 125; [Marx 1873] at 49.

<sup>20</sup> Gyorgy Markus argues a partial shift between the critical theory of *Grundrisse* and *Capital* ([Markus 1986] at 126–145, esp. 139f.). While sharing what he terms an Aristotelian-Hegelian content and form dichotomy, he sees the course of exposition in the earlier work organised “according to the principle of ascendance from the abstract to the concrete” whereas in *Capital* it goes according to that of essence and appearance (142).

<sup>21</sup> [Marx 1879] at 215.

<sup>22</sup> [Marx 1976] at 92; [Marx 1873] at 16.

of the generic difference between logical and historical dialectics affirmed by Hegel.<sup>23</sup> The chapter is on the theory of value and involves a critique of the theory of value in classical political economy.<sup>24</sup> But it is intended at, is about, ‘commodity’ in a phenomenological sense. ‘Form’ here is not then just a metaphysical category, determined only by place within a set of abstract structures. It also is more specific: an appearance form, an elementary or cell-form, a natural form, a value form, a form of expression and more. Content depends on the type of form at issue.

The individual commodity will turn out to be contradictory, a thing with a double character of use-thing and value-thing. As ‘elementary form’ it has this contradiction within itself, but as ‘appearance-form’ it is constituted in its relation to another commodity and the expression of that relation in the ‘value-form’. Thus, later, after consideration of its dual character and after a section on the dual character of the labour embodied in commodities, at the beginning of the section ‘The value-form, or exchange-value’ we get to “the form of commodities”, the way they appear “in so far as they possess a double form i.e. a natural form and a value form”.<sup>25</sup> But “first of all”, as a first step in the analysis of this subject, “the commodity . . . is an external object, a thing which through its qualities satisfies human needs of whatever kind”. Tied into (conditioned) by the physical properties of commodities, use-values “in the form of society to be considered here are also material bearers (*Träger*) of — exchange-value”.<sup>26</sup>

<sup>23</sup> Marx’s acerbic insistence in his ‘Marginal Notes on Wagner’ that he does “not proceed on the basis of ‘concepts’ hence also not from the ‘value-concept’ ” but from “the simplest social form in which the product of labour in contemporary society manifests itself, and this is as ‘commodity’ ” ([Marx 1879] at 214), as the statement there that his approach does not set “ ‘logical’ and ‘historical’ concepts in contrariety” (ibid at 221) underlines these points.

<sup>24</sup> “As regards value in general, classical political economy in fact nowhere distinguishes explicitly and with a clear awareness between labour as it appears in the value of a product, and the same labour as it appears in the product’s use-value. Of course the distinction is made in practice, since labour is treated sometimes from its quantitative aspect, and at other times qualitatively. But it does not occur to the economists that a purely quantitative distinction between the kinds of labour presupposes their qualitative unity or equality, and therefore their reduction to abstract human labour” ([Marx 1976] at 173 n.33; [Marx 1873] at 94 n.31).

<sup>25</sup> [Marx 1976] at 138; [Marx 1873] at 62.

<sup>26</sup> Ibid at 126; 50.



‘Matter’ — ‘stuff’, ‘body’, ‘substratum’ — nature’s provision to the wealth of societies with its ‘coarsely sensuous existence’ tends to be under-emphasised in accounts of Marx’s thinking which, quite properly emphasise the practical character of his materialism: the role of social practices, the situatedness of social personae and things within historically developed relations of production, and the integration of political commitments, aims and ideals in theoretical standpoints. Individual commodities as use-values and bearers of exchange-value are physical things with properties such as existence in space and time, weight, chemical composition that are investigated in mathematical and natural sciences. The value of a commodity, in Marx’s analysis, *cannot be* a “geometrical, physical, chemical or other natural property” of the commodity. These properties are relevant only to its use-value. “On the other hand, it is just the abstraction from their use-value, which evidently (*augenscheinlich*) characterises the exchange relation of commodities.<sup>27</sup> Value pertains to a social relation between two different commodities — ‘social’ in the sense that it is a product of human practice.

The practical nature of Marx’s materialism with its consequential doctrine of the conceptually and historically mediated character of scientific knowledge of these objects may be its saving grace. But this chapter is incomprehensible without the theoretical sense of that ‘cannot be’. The sense in which use-value is the contrary of value and its appearance form in bourgeois society, exchange-value, depends on it, as also the distinction Marx makes between value and exchange-value. Political economy as a science, like mathematical and natural sciences, seeks knowledge of laws and the ‘objectivity’ of its knowledge rests in laws, but as laws of social change these are “special laws that regulate the origin existence, development and death of a given social organism and its replacement by another higher one”.<sup>28</sup> The distinction made here between laws of nature and laws of political economy is critical to Marx’s theory of value in its divergence from classical political economy.<sup>29</sup> The independent existence of matter in nature posited — wrongly to my mind — by Marx against the independence

<sup>27</sup> Ibid at 127; at 51–2.

<sup>28</sup> Ibid at 102; 27, quoting a description of which Marx approved.

<sup>29</sup> Rosa Luxemburg considers that a rejection of an absolute universality inhering in a natural law perspective on science and its laws is constitutive of Marx’s critical approach ([Luxemburg 1913] at 67f).

that Hegel vests in the form of the logical Notion,<sup>30</sup> distinguishes their ideas of ‘foundation’. The various secrets, puzzles and mysteries that accompany “the value-form of the commodity or the commodity form of the product of labour” are foundational questions, questions basically of the ‘objectivity’ of the value of commodities.

By undertaking the task with which I began these preliminaries, that of showing the simple value-form to be the origin of the money-form, Marx confronts himself with a problem of equality relations: a problem of the difference or sameness of equality and identity which, in the work of Gottlob Frege, will write its own chapter in the history of symbolic logic and the foundations of arithmetic. But the emergence of that problem lies in the future and indeed in a purely formal science. The problem as Marx states it at the beginning of the section on the value-form, concerns the objectivity of the value of commodities (*Wertgegenständlichkeit der Waren*). It “differs from Dame Quickly ‘in the sense that one does not know where to have it’”.<sup>31</sup> Not an “atom of matter” enters it and while, as Marx bids his reader recall, commodities possess this objectivity only in so far as they are expressions of abstract human labour so that, being purely social in character, it can only appear in a social relation between commodities, this is merely a re-statement of the problem: a summary statement which calls for analysis of the value-form and its secret.

This secret, that which is hidden in the value-relation between two commodities, is fairly quickly aired, first through Marx’s analysis but then with evident paradox by the commodity, the value of which is being expressed. The linen in association with the coat

reveals its thoughts in a language with which it alone is familiar, the language of commodities. In order to tell us that labour creates its own value in its abstract quality of being human labour, it says that the coat, in so far as it counts as its equal, i.e. is value, consists of the same labour as it does itself. In order to inform us that its sublime objectivity as a value differs from its stiff and starchy existence as a body, it says that value has the appearance of a coat, and therefore that in so far as the linen itself is a value-thing, it and the coat are as like as two peas.<sup>32</sup>

<sup>30</sup> [Hegel 1969] at 586; [Hegel 1816] at 18; and see [Kerruish 2006] at 41.

<sup>31</sup> Ibid at 138; 62.

<sup>32</sup> Ibid at 66–7; 143–4.

This is a foretaste of things to come and the thread of continuity between the first three and the last sections which, not without ground, is so regularly severed. For if Marx will have it that, if one is investigating the production of value, there is “nothing mysterious” in having to consider the human labour power expended in tailoring and in weaving in terms of their general characteristic of being human labour, the ‘inversion’ (*Verdrehung*) that prompts him to begin the section on commodity fetishism with the mystical, magical dance of a table, is stated in more or less the same breath: “in the value expression of the commodity, the matter is stood on its head (*wird die Sache verdreht*)”.<sup>33</sup> It is at this point in his text that the discussion of Aristotle is called in aid.

Two secrets or one? And to what extent is it or are they phantoms of Marx’s persuasion that the sought after objectivity exists? Or if not that, are they products of a set of metaphysical persuasions that will not leave such objectivity where it is, in a form of expression, in language or in the thought expressed in the formula? Or, as I am suggesting, does the secret have a parallel in the enigmatic behaviour of signs which “suddenly display their own selves when they are combined by means of the sign for identity of content.”<sup>34</sup> The first question should not detain us. Resolving this problem of the objectivity of the value of commodities goes through the breach of an ‘inversion’. It is to Althusser’s credit that he inveighed against the ambiguities of this term and too hasty investments in the “rapidly written lines” of the Postface to the second edition of *Capital*.<sup>35</sup> I would redescribe to say that Marx confronts a certain ‘objectivity of illusion’.<sup>36</sup> It is an objectivity that is constituted in the flash of an eye: in

<sup>33</sup> Ibid at 150, 72.

<sup>34</sup> [Frege 1879] at 20.

<sup>35</sup> [Althusser 1967] at 17. I am also sympathetic to his suggestion that, as Marx used it, ‘inversion’ is a word lacking a concept. But I part company with Althusser early, with his claim that the term holds the place of an absent question of the specific difference between Marxist and Hegelian dialectic, not because I suppose that Marx was fully aware of what he was doing theoretically, rather more because I find the terms ‘Marxist’, ‘Hegelian’ and ‘dialectic’ too fuzzy to yield a ‘specific difference’ with substantial theoretical content (cf. [Althusser and Balibar 1970] at 32f.).

<sup>36</sup> The phrase is taken from Hegel’s homage to Kant’s recuperation of dialectic as a “necessary function of reason” in the Introduction to his *Science of Logic*. Hegel writes of the “*objectivity of the illusion* and the *necessity of the contradiction* which belongs to the nature of thought determinations” ([Hegel 1969] at 56; [Hegel 1812] at 54). See further [Kerruish 2006] at 30.

just that abstraction from their use-value, which “evidently (*augenscheinlich*) characterises the exchange relation of commodities.”<sup>37</sup> Its secret, to my mind, has gained a laboratory in the very strange universe of mathematical logic, a universe so strange as to be devoid of and applicable to that stuff, body, substratum which Marx will have it is being abstracted from, which he encases and encloses in ‘nature’ and, differently, in the concept of value.

As I have said, for Marx, the solution to the secret of the simple form of value is within reach once it is grasped that human labour power creates value but is not itself value. Does the puzzle or riddle of money (*Geldrätsel*) disappear then, as promised, when the trace connecting the simple form to the “dazzling” money form has been detected and the forms of expression of value analysed?<sup>38</sup> Not entirely: the analysis shows, so Marx, that the money form of value is not formally different to the universal form. Gold takes the place of linen as universal equivalent. “The advance consists only in that the form of direct and universal exchangeability, in other words, the universal equivalent form, has now by social custom finally become entwined with the specific natural form of the commodity gold.”<sup>39</sup> The significant transformations are between the simple, total or expanded, and universal forms.<sup>40</sup> But a formal analysis of what the double form of the commodity *is*, in a social theory yet lacks an account of what it *does* and why it does it. These transformations and the double character of the commodity and of the labour embodied in it will now be re-traversed, re-covered, from the perspective of its ‘activity’: the perspective of practice, based for Marx on what people do, how they perceive the products of that doing and, if they are political economists, represent it.

<sup>37</sup> [Marx 1976] at 127; [Marx 1873] at 51–2.

<sup>38</sup> *Ibid* at 139; 62. The Fowkes’ translation of ‘puzzle’ or ‘riddle’ as of the ‘secret’ to be considered shortly as ‘mystery’ doesn’t allow for the sense of this question.

<sup>39</sup> *Ibid* at 162: 84.

<sup>40</sup> The last, the transformation to the universal form brings, for Marx, a simplification and unification (*ibid* at 79: 157). Constructed by a totalisation over a potentially infinite series of value relations, constitutive of the universe of commodities through the exclusion of one of their party to serve as their universal equivalent, it represents yet another mode of abstraction: one that seems lately to be taken as some kind of original sin.

### 3. ‘The Fetish Character of the Commodity and its Secret’

The value-form of the commodity or the commodity form of the product of labour: Marx treats them as synonymous in his Preface to the first edition of *Capital*. The difference is modal, a shift from analysis of what the double character of the commodity is and its origin in the simple value-form to what it does. It is a difference neither in the subject of the chapter (the commodity) nor its topic (the theory of value), only an order of exposition that could have been otherwise. Still, it should come as no surprise, that the latter perspective will provide the occasion for a greater emphasis on the practical dimensions of Marx’s thinking.

The section is written in a single piece, without subdivision, but it also, more or less systematically, shifts in endeavour. It first describes and names the phenomenon, then explains it, then seeks to dispel it, then addresses itself to bourgeois political economy in the form of a critique of its theory of value. Value and values are bewilders, incomprehensible, without paying attention to that which, through their forms they themselves conceal.

Which is not to say that this text is not itself bewildering! I think it is clear that Marx regards the fetish character of the commodity as an actual social phenomenon of commodity producing societies within which capitalist production is dominant. It shall disappear with the disappearance of capitalist social relations. His explanation of why this is so is also relatively clear: it is a phenomenon of social relations of commodity production within which producers labour in private, enclosed from and independently of other producers.

I say relatively clear because the notions of ‘private’ and ‘social’ are very thin or abstract here. What, as a matter of the method of political economy has been presupposed, “the subject, society”,<sup>41</sup> is given a few

<sup>41</sup> As stated in the 1857 Introduction to the *Grundrisse*, the method of political economy presupposes “the subject, society” ([Marx 1973] at 101–2; [Marx 1857–1858] at 21–2). This follows an account of the “reproduction of the concrete by way of thought” that is accomplished by Hegelian dialectic and the illusion as regards the nature of reality into which that very accomplishment, so Marx, lead Hegel. Marx appears to have had second thoughts as to the methodological reflections of this Introduction and probably intended the (1859) Preface to *Zur Kritik der politischen Ökonomie* to take their place ([Marx 1859] at LIII). Certainly it is this later formulation of his foundation that he repeats in *Capital* ([Marx 1976] at 175 n.35; [Marx 1873] at 96 n.33). I don’t see that either this or the partial shift between *Grundrisse* and

bare determinants here: just a few initial pen strokes. It is characterised by a “spontaneously arisen” or naturally grown (*naturwüchsig*)<sup>42</sup> social division of labour which assigns to individual useful labours a character as heterogenous types of useful labour. The ‘private’ is sketched only in terms of mutual independence and isolation of the acts of labour. Within these sparse determinants, the contention is that the social character of production, while in itself unproblematic and present “as soon as humans work for each other in any way”,<sup>43</sup> is increasingly disguised in the course of history, reaching an apogee in capitalist social formations.

It is also clear that for Marx the mystical character of the commodity comes neither from its use-value nor from the “nature of the determinants of value”: useful work, time and social relations between humans working for each other. It comes from the form of the commodity with its double character as use-value and value. Within this form the equality of human labour, time (duration) as the measure of the expenditure of human labour, and social relations between producers all receive doubles. The problems already begin in trying to decipher the double of the equality of human labour.

Die Gleichheit der menschlichen Arbeiten erhält die sachliche  
Form der gleichen Wertgegenständlichkeit der Arbeitsprodukte  
...<sup>44</sup>

The thought here, better captured in the Moore and Aveling translation, is to set up the equality of products of labour in their all being, indiscriminately, values, as the double. Words have a hard time getting at *Doppelgänger!*

The mysterious character of the commodity form consists simply therein, that it mirrors back to humans the social character of their own labour as the objective character of their very own

*Capital* discharges the presupposition in question, although perhaps it should not be nominated as ‘subject’.

<sup>42</sup> [Marx 1976] at 166; [Marx 1873] at 87.

<sup>43</sup> Ibid at 164; 86.

<sup>44</sup> [Marx 1873] at 86. Rendered by Fowkes as: “The equality of the kinds of human labour takes on a physical form in the equal objectivity of the products of labour as values” (at 164); and by Moore and Aveling as “The equality of all sorts of human labour is expressed objectively by their products all being equally values” ([Marx 1889] at 42); I would translate “The equality of human labours gains the factual form of the products of work being all equally values.”

products of labour, as social, natural-properties of these things, and thereby also the social relation of the producers to the sum total of labour as a social relation between objects, existing externally to them. Through this substitution (*Quidproquo*) products of work become commodities, sensuous supersensuous or social things<sup>45</sup>

So, a substitution, a constitutive substitution, worked by the commodity form, exchanges the social character of that which produces value with the character of that which is produced, values. The substitution is located in perception. Thereby the products of labour become commodities: things of a strange kind. The social character of labour, that is its character as labouring for others becomes the social character of things, their being for each other. It would be like the mechanical optics of vision were one dealing with physical phenomena, but that is not the case. The commodity-form and the value-relation of products of labour within which it appears, has absolutely nothing to do with the physical character of those products of labour and their relations as things. (Here is that flash of the eye, the abstraction which ‘evidently’ characterises the exchange relation of commodities.) In its very duplicity, the commodity form plays up, plays tricks, so that “definite social relations between humans themselves . . . assumes here, for them, the bizarre (*phantasmagorisch*) form of a relation between things”.<sup>46</sup>

In order therefore to find an analogy we must take flight into the misty realm of religion. There the products of the human brain appear as autonomous figures endowed with a life of their own, which enter into relations with each other and with the human race. So it is in the world of commodities with the product of men’s hands. I call this the fetishism which attaches itself to the products of labour as soon as they are produced as commodities and is therefore inseparable from the production of commodities.<sup>47</sup>

What Marx does here is name the phenomenon with a term coined in colonising Europe to characterise the superstitions of primitive peoples. I suppose he is using the analogy and name to tell against the conceit

<sup>45</sup> [Marx 1976] at 164–5; [Marx 1873] at 86.

<sup>46</sup> *Ibid.*

<sup>47</sup> *Ibid.*

of enlightened bourgeois Europe, to hoist it on its own petard. But the hoist is the petard: a doctrine of historical progression of increasingly civilised religious beliefs which go hand in hand with economic, scientific and political progress. Marx adheres to and propagates this narrative, locating the promise of socialist revolution within it.

It is a move that fits that narrative and it is a narrative that was and is used to justify crime. Is it further imposed by the narrative? Here there are other questions that press. Why the shift to analogy? Whence the “therefore” that instructs the flight? Because the optical analogy founders on the absence of light being transmitted from one physical object to another? But this must be a feint! Or, if this bizarre appearance is the trammel of exchange-value as “merely appearance-*form*, not its own *content*”,<sup>48</sup> what or where is the, or is there a form that *is* its own content?

I take these questions up against a pretty obvious sense, not unrelated to the polemical intention mentioned, in which the analogy *is* imposed by the narrative. The shift to analogy is forced by Marx’s location of the substitution in perception and the metaphor of mirroring covering the function of the commodity form. Marx sees himself as confronted by that inversion (*Verdrehung*) in the forms of expression of value, as facing a breach of *sense* worked by these forms which he relates to an ‘inverted’ and mystical logic. The sight analogy fits the location (partially, but then it is only an analogy) but, given the claim that the abstraction involved works a complete divorce from the physicality of labour products, it is bound to fail. It functions only to instruct the flight to the misty realms of religion, to ground the ‘therefore’ and then on the supposition that there are just two possible realms where an analogy may be found: physical/neurological and spiritual. And metaphysical/logical?<sup>49</sup> The form that is its own content is Marx’s concept of value and that there might be questions of concept formation going to Aristotle’s approach to equality relations or indeed to a singular canon of classical logic consisting in some so called ineffable ‘laws of thought’ is excluded.<sup>50</sup>

<sup>48</sup> [Marx 1879] at 215.

<sup>49</sup> I do not suggest this is the only other possibility. The psychoanalytic appropriation of fetishism is an obvious further candidate. My suggestion is directed by my interest and endeavour.

<sup>50</sup> “There is nothing in any of Marx’s writings to justify burdening him with the absurd doctrine that the law of non-contradiction is invalid” ([Callinicos 1983] at 54). Alas and that on many fronts!



It is not that all is lost from this point. The paragraph ends and the text shifts into a second phase, explaining what has been identified or at least named, by taking up, not the double form of the commodity but that of the labour embodied in commodities. It recurs back to the second section of the chapter, to the point which Marx has introduced into political economy. On the one hand the use-values of commodities come from diverse, heterogeneous concrete labour embodied in them. On the other, as values, their substance is the same: equal, or abstract human labour; labour reduced by a social process “that goes on behind the backs of the producers”<sup>51</sup> to the expenditure of simple labour power.

The fetish attaching to commodities “arises from the peculiar (*eigentümlichen*) social character of the labour that produces them”. The significance of practices of exchange as that point at which producers labouring in private come into contact with each other is brought into relief. The social character of their private labours only appears within the context of that contact.

To the producers, therefore, the social relations between their private labours appear as what they are, i.e. they do not appear as direct social relations between persons in their work, but rather as material [*dinglich*] relations between persons and social relations between things.<sup>52</sup>

No illusions then? The bizarre appearance turns out, to use an Hegelian phrasing, to be the truth of social relations of production under conditions of private ownership of the means of production? I leave this run. Hegel is no political economist.

The text continues into a play of forms and processes of abstraction through which Marx conceptualises and represents how this state of affairs comes into being: exchange of commodities, its deepening and extension over a period of time to a point where useful things are produced for the purpose of exchange, and the effect of this “moment” in history on the labour of the individual producer. “From this moment on the labour of the individual producer acquires a twofold social character.” As an element in an originally natural but increasingly complex social division of labour, that labour must hold its place within this totality, by being concrete useful labour that satisfies definite social needs. But this it can

<sup>51</sup> Ibid at 135; 59.

<sup>52</sup> Ibid at 165–66; 87.

do, “only insofar as every particular kind of useful, private labour can be exchanged with i.e. counts as the equal of every other kind of useful private labour.”<sup>53</sup> Then on through the abstraction that constitutes the equality of completely different labours, over the role of social practices in the constitution of ideas<sup>54</sup> and back to a somewhat more determined representation of social processes going on behind the backs of intending individuals. The paragraph renders the transformation of the expanded to the universal form of value in less formal terms, renders the formulae of the previous section into a language of social theory. Its immediate destination is a critical point.

Men do not therefore bring the products of their labour into relation with each other as values because they see these objects merely as the material integuments of homogeneous human labour. The reverse is true; by equating their different products to each other in exchange as values they equate their different kinds of labour as human labour. They do not know it, but they do it.<sup>55</sup>

Intentionality and its incompleteness, the gap between actual and intended results of human action, between purposive human practices and their social effects enters the text.<sup>56</sup> It is the occasion for the social theorist or scientist to intervene and say what has happened.

Value, therefore, does not have its description branded on its forehead; but rather transforms every product of labour into a social hieroglyphic. Later on, humans try to decipher the hieroglyphic,

<sup>53</sup> Ibid.

<sup>54</sup> I am condensing an already dense text: “Equality of completely different labours, can only consist (*bestehen*) in an abstraction from their real inequality, in the reduction to the common characteristic which they have as the expenditure of human labour power, human labour in the abstract. The private producer’s brain reflects this twofold social character of his labour only in the forms which appear in practical intercourse, in the exchange of products” (ibid). “Gleichheit toto coelo verschiedener Arbeiten kann nur in einer Abstraktion von ihrer wirklichen Ungleichheit bestehn . . .”. It is not clear to me whether “toto coelo”, meaning fully or completely, qualifies ‘equality’ as in the Fowkes translation, or ‘different labours’ as I have rendered it. “Equality in the full sense” would indeed be identity.

<sup>55</sup> Ibid at 166; 88.

<sup>56</sup> Cf. [Markus 1986] at 6f, coupling this with objectivation (*Vergegenständlichung*) as basic to Marx’s idea of social theory.

to get behind the secret of their own social product: for the characteristic which objects of utility have of being values is as much their social product as is their language. The belated scientific discovery that the products of labour, in so far as they are values, are merely the material expressions of the human labour expended to produce them, marks an epoch in the history of mankind's development, but by no means banishes the semblance of objectivity (*gegenständlichen Schein*) [objectual appearance] possessed by the social characteristics of labour.<sup>57</sup>

What term, thing or object does have its description branded on its forehead? What doesn't have the character of hieroglyph? That must be a question. The 'therefore' however relates the thought here to the incomplete intentionality of human practices of production and exchange of commodities and the difficulties, after the event, after use-objects have become values, of deciphering what unbeknown to themselves they have done. Certainly, for Marx, a labour theory of value is indispensable to that task. As itself a product of definite social conditions, it tells of a progressively developing awareness. But theory does not change what it comprehends merely in comprehending it. The continuation of the paragraph now gets back to the specificity of value implied at the beginning. While the objectual appearance of social labour, like the chemical composition of the earth's atmosphere is not changed by its valid scientific analysis, it should not be forgotten that this analysis, unlike chemical analysis, is valid only for the particular social formation within which labour products become commodities. It is again a critical point. Marx is distinguishing between the universality of the valid statements of geometry, physics, chemistry etc. and the particularity to a given social formation of those of political economy.

For all that is problematic here, I wish to keep a focus on the objectual appearance that Marx is (and I am) struggling with. It is not the 'objectivity of illusion' (*Objektivität des Scheins*). It has the sense, as far as I can grasp it, of 'appearing to have the properties of a sensible object'. It is an appearance that goes through, happens on account of the *Quid-pro-quo*. The 'objectivity' of scientific knowledge on the other hand, has a different sense, one that refers back to the mode of scientific comprehension and is intended in 'the objectivity of illusion'. It has a connotation of

<sup>57</sup> [Marx 1976] at 167; [Marx 1873] at 88.

value or validity tied to its being (taken as) a form of knowledge of laws governing the realm of inquiry or, as I would prefer, of inescapability. The objectivity of the value of commodities plays between these senses: one does not know where to have it.

Following a paragraph covering a shift from quantitative to qualitative stabilisation of value in broadly socio-historical terms which, en route to locating the objectivity of the value of commodities in the socially necessary labour time for their production “asserting itself as a regulative law of nature”,<sup>58</sup> explicitly articulates Marx’s accusation of things controlling people rather than people controlling things, the text shifts again from the phenomenon and its explanation to representation of the phenomenon in political economy and theories of value.

Political economy, according to Marx, can avoid being misled in the theory of economic value even while the social conditions giving rise to the fetish prevail. Objects of utility become values within a social and historical process. But,

[r]eflection on the forms of human life, hence also the scientific analysis of those forms, takes a course directly opposite to their real development. Reflection begins *post festum* and therefore with the results of the process of development ready to hand. The forms which stamp products as commodities and which are therefore the preliminary requirements for the circulation of commodities already possess the fixed quality of natural forms of social life before man seeks to give an account not of their historical character, for in his eyes they are immutable, but of their content.<sup>59</sup>

It will shortly become clear that the plea that Marx makes in this passage for an historical approach as antidote to forgetting historical processes of becoming,<sup>60</sup> also observes the synchronic organisation of the categories of a science. The observation on forms stamping products of labour as commodities misrecognised as having “the fixed quality of natural forms of social life” is this forgetting. It anticipates the specific criticism of

<sup>58</sup> Ibid at 168; 89.

<sup>59</sup> Ibid at 168; 89–90.

<sup>60</sup> Cf. [Marx 1973] at 85; [Marx 1857–1858] at 7: “The whole profundity of those modern economists who demonstrate the eternity and harmoniousness of existing social relations lies in this forgetting.”

classical political economy, namely that it assumes capitalist production to be a natural, normal form of social organisation.

“The whole mystery of the form of value”, Marx has said, lies hidden in the simple form of value. But it is the finished (*fertige*) form, he now adds, the money form, which both gives commodities the character of values and does the work of concealment of the social character of private labour and the social relations between producers, “by making them appear as relations between things, instead of revealing them plainly”.<sup>61</sup> And yet, as noted, a few pages back, in the context of his explanation of the phenomenon, Marx has said, “[t]o the producers, . . . the social relations between private labours *appear as what they are* (my emphasis).”

This is the question of illusion (or not) that I let run. I do not think that Marx is contradicting himself but there is slippage. The money-form is that which conceals and that *it* should have the capacity to reveal things plainly is counterfactually intended. What I think Marx wants to say is that the money-form *qua* universal form of value, constitutes the character of commodities as values and *merely* on account of gold or silver rather than coats or boots serving as the universal equivalent — a contingent matter without formal theoretical significance — conceals the social character of the labour and the social relations between producers. Marx draws here on his earlier analysis. He also assumes that the money-form bedazzles. The absurdity of saying that coats or boots are the universal incarnation of abstract labour, he claims, is self-evident. It is the dazzling money-form that conceals this absurdity. But there are several layers to ‘this absurdity’ and it is here that there is slippage. That this absurdity is no error of judgement on the part of those to whom it appears is the earlier point. The situation is one in which producers are ruled by their product. It is both actual and unreasonable. That it gives rise to appearance forms which substitute the properties of products of labour for the character of the labour process which produces them is also actual. These forms are themselves products of the relations of production that obtain. Is the substitution itself unreasonable? As the fetish character of commodities it is phenomenal: the earlier point. As *fetishism* it is pejoratively connoted as unreasonable, perverse. Something further is involved: a practice or representation that, at least, fails to grasp, is

<sup>61</sup> [Marx 1976] at 169; [Marx 1873] at 89.

tricked by, falls prey to the dynamic of essence and appearance involved *and thus* constitutes a force against a possible, reasonable future.

That the material conditions of all this are alterable; that indeed a certain completion of a phase of historical development reached with the emergence of the universal form of value, both satisfies a condition for social revolution *and* obscures the possibility of revolution, are all now in issue. The normativity indicated stems from Marx's judgement that it is unreasonable that producers be ruled by their product. It gains its point from his conviction that this can be changed. It makes a demand on theory to distinguish what is actual from what is reasonable and thus to contribute to making what is reasonable, actual. An explanation is also emerging: political economy is caught, retarded in failure to see that the forms which constitute objects as commodities are relative to a particular social formation.

Such is the character attached to the categories of political economy. That is the generalisation to which Marx moves.<sup>62</sup> Later, he will specify the question which the political economists failed to ask. My reservations here go to the money form as the "finished form" of the world of commodities. When is a form — a category, an institution, a shape of life — finished or completed? One can only say, it seems to me, *post festum*, with the experience of collapse, contradiction, sudden insight or unpredicted change in conditions that engenders uproar, crisis, new horizons or circumstances. Yet there is also a great insight. What is new to capitalist society is that the disguise of universality has fallen over the particularity of capitalist production relations themselves. So and even so, it is just with the emergence of universal forms of production relations that political economy comes onto the scene as a science. The categories or concepts constituting it as such become thinkable. They are true to ("socially valid and therefore objective (*objektiv*) thought-forms of")<sup>63</sup> the relations of production of this place and time: appropriate to the scientific analysis of capitalist but not all social formations. They appear as universal, unlimited in their applicability, but they are not.

The disguise of universality is only the more difficult to see through in the thought it enables. In order to dispel its "magic and necromancy" (*Zauber und Spuk*) Marx moves to a representation of four modes of pro-

<sup>62</sup> Ibid, at 169; 90.

<sup>63</sup> Ibid.

duction.<sup>64</sup> In a following paragraph, within a narrative of progression from unfreedom to freedom and from a supposed portent of increasingly powerful means and enriching modes of production, Marx assigns various “religious reflections of the real world” to their corresponding social formations and concludes:

The shape of the social life-process, that is, of the material process of production, will only shed its mystical veil of mist, when, as the product of freely associated humans, it stands under their conscious and planned control. That, however, requires a material foundation of the society or a series of material conditions of existence, which are themselves, again, the naturally grown product of a long and tormented historical development.<sup>65</sup>

Placed into the grammatical subject position, the shape (*die Gestalt*) of the material process of production is given a self-revealing potentiality, the actualisation or realisation of which is deferred to the activity of (some) humans, producers who will constitute themselves as revolutionary subjects. In an envisioned new sovereignty, the incomplete intentionality of human action should disappear from the realm of necessity.

The text moves back at this point to political economy and the question that separates it from Marx’s critique. Political economy, he writes,

has uncovered the content concealed in these forms [of value and its magnitude]. But it has never once asked the question why this content has assumed that particular form, that is to say, why labour is expressed in value, and why the measurement of labour by its duration is expressed in the magnitude of the value of the product. These formulas, which bear the unmistakable stamp of belonging to social formations in which the process of production has mastery over man, instead of the opposite, appear to the political economists’ bourgeois consciousness to be as much a

<sup>64</sup> Crusonade, feudal, communal peasant and communist. They have distinctive intended essences: the classical liberal individualism of the Crusonade narrative, the hierarchy of feudal relations of personal dependence, the patriarchy of the peasant family and the transparent relations of production and distribution in an imagined “association (*Verein*) of free humans, working with the means of production held in common, and expending their many different forms of labour power in full self-awareness as one single social labour force” (ibid at 171; 92).

<sup>65</sup> Ibid at 173; 94.

self-evident and nature-imposed necessity as productive labour itself.<sup>66</sup>

Marx's question is a demand on theory to grasp content revealed and concealed in expressions. What stands before or behind it is the way in which Marx has sought to meet that demand in the preceding body of the chapter. What is presented is an answer to the question of why political economy, even at its best, has *not* asked Marx's question. No doubt this is a move that is characteristic of critique. Nor should it be forgotten that, although the analysis of the discourse of political economy that might support Marx's contention is not before the reader, Marx has been working the archives of political economy for over twenty years: a labour that may well be described as a labour of love. Three long footnotes appended to the paragraph may serve as such a reminder. Perhaps a residing unease with this answer should be put to the limitations of critique. Or perhaps it is the category or notion within which it is formulated, namely self-evident, nature-imposed necessity that provokes dissent.

Withal, finally, remarkably, in still another attempt to communicate his idea, Marx has commodities re-take the stage to speak their own secret. They take the stage as self-aware and articulate subjects, gaily poking fun at the economists they mislead. Not use-values but exchange values are what they are about, how they relate to each other, what they intend.

Our use-value may interest men, but it does not belong to us as things (*Dingen*). What belongs to us as things however is our value. Our own intercourse as commodities proves it. We relate to each other merely as exchange-values.<sup>67</sup>

#### 4. Religion, Science, Ideology

It may not seem like much of a secret. It may, for those for whom there is no stuff and matter only only signs and symbols, tell only of a Marx caught in his own illusion. The secret concerns things and anything, everything is or can be taken as a 'thing'.<sup>68</sup> But then the particular, peculiar thingness of commodities is what is at stake. *That* is what is constituted by the

<sup>66</sup> Ibid at 174–5; 94–5.

<sup>67</sup> Ibid at 176–7; 97.

<sup>68</sup> Cp. [Nancy 1993] at 167–88.



*Quidproquo* even as commodities as use-values that are bearers of value are products of humans labouring in bourgeois society.

What commodities, these sensuous-supersensuous or social things, would say of themselves, were it the case that they had that other peculiarly social characteristic of humans, namely language, is what Marx has them say. His critique is of political economy, but the metaphysical and theological subtleties of the thing-ness of commodities concern substance and subject in a thinking that is also a thinking of necessity and freedom.

“[E]verything turns on grasping the True not only as *Substance* but also as *Subject*”, Hegel wrote in the Preface to his *Phenomenology of Spirit*.<sup>69</sup> For Marx, substance become subject is ensouled commodities. Such a portrayal works a *reductio ad absurdum* of Hegel’s standpoint, but not as an argument that dangles in mid-air. Ensouled commodities speaking out their own character or identity is the inscenation of an idea it is written against and in that sense could not have been written without.

It is by all means a remarkable scene and its continuation, the invitation to compare what commodities would say of themselves with what the economists cited as “speaking out of the soul of the commodity” say of them, is remarkably ambiguous, indeed paradoxical.

Value (exchange value) is a property of things, riches (use-value) of man. Value, in this sense, necessarily implies exchanges, riches do not.

And:

Riches (use-value) are the attribute of man, value an attribute of commodities. A man or a community is rich, a pearl or a diamond is valuable . . . . A pearl or a diamond is valuable as a pearl or diamond.<sup>70</sup>

The scene plays out within the spell of the fetish (or as a play within a play) and it does give rise to the suspicion, which Derrida voices, that

<sup>69</sup> [Hegel 1977] at 9–10; [Hegel 1807] at 22.

<sup>70</sup> [Marx 1976] at 177; [Marx 1873] at 97. The economists cited reject Ricardo’s theory of value. They are particular cases, for Marx, of getting things wrong as regards value, use-value and exchange value. Ricardo’s followers, Marx notes, are unable to make a convincing response to the criticism of him made by the (unnamed) author first cited and S. Bailey (the author of the second citation) for the familiar reason that they do not find in Ricardo’s work any elucidation of the internal connection between value and the form of value, or exchange value (ibid n.38; n.36).

Marx puts words into the mouths of commodities only then to claim that the economist naively reproduces this fictive speech.<sup>71</sup> I am not clear what Derrida is aiming at, but he seems to by-pass just that which I wish to dwell on: the problematic ‘objectivity’ of fetish phenomena. What these commodities say is that use-value “does not belong to us as things” (*Er kommt uns nicht als Dingen zu*) and the thing-ness of commodities is just that which their form, their double character both constitutes and dissembles. That this thing-ness, this ‘objectivity’ is *not* the ‘objectivity’ of value, that it plays between two senses of ‘objectivity’ and in this mischievous play, plays up with these senses is the point. The hyperbole of their speech, the extravagance of linen in association with a coat revealing “its thoughts in language with which it alone is familiar” positions Marx, outside the ‘world of commodities’, as de-cipherer of a code written by sensuous, practical human activity in the space of its own incomplete intentionality. That it, this code, this speech, is thus written; that it is not of Marx’s sole authoring; that there is an ‘objectivity’ in all this which Marx wants to get at and bind into a transformable present is the point I wish to hold on to. It is *an* objectivity of illusion; an instantiation of such objectivity in the social realm. It is not *the* objectivity of illusion in the sense it which Hegel takes such objectivity, product of contradictions existing in *all* ideas, concepts, and categories, as a *logos*, a principle of intelligibility from which a science of logic becomes possible.<sup>72</sup>

The claim that the code is not, trivially, of Marx’s sole authoring is, in one dimension, the same implausible claim that Hegel makes for his derivation of the categories of metaphysics. It supposes a method that is internal to the task in hand which is to trace a contradiction driven *and constitutive/productive* dynamic, and that within an epistemology (or counter-epistemology)<sup>73</sup> which admits subject-object interaction in the production of knowledge. *Capital* is written from a standpoint that is for the working class, the class which, in Marx’s political thought, is the ‘subject of history’. These first chapters of *Capital* however have the

<sup>71</sup> Cp. [Derrida 1994] at 157f.

<sup>72</sup> That totalisation in Hegel’s case, undermining as it does, literal readings of his absolute idea, holds his logic open; makes his logical foundation, unlike Kant’s, a foundation that does not justify and so attempt to conserve, existing knowledge that is presumed complete and perfect. Cf. [Nancy 2002] esp. at 19f and 66f.

<sup>73</sup> For a sketch of the epistemological dimension of Hegel’s *Phenomenology* ‘standpoint of absolute knowing’ and its relation to his *Logic* see [Kerruish 2006] at 33.

significant similarity to Hegel's *Logic* of being subjectless in *that* sense of subject. In making the commodity the subject of his chapter, Marx here no less than Hegel there, is engaged with the agency of forms in knowing and knowledge: logical forms for Hegel; forms that have emerged as the conventionally established categories of classical political economy (value, use-value, exchange-value) for Marx. And just as Hegel's *Logic* is intended to be an organon or tool for the production of objective insights, Marx's theory of value is intended as a tool of economic analysis.<sup>74</sup> In both cases then the task at hand is practical-theoretical, reliant on a method (as well as presupposing an epistemic standpoint), although not justified by it.<sup>75</sup>

In reading Hegel's Idea as a demiurgos outside the system Marx seems not to recognise this dimension in Hegel's approach, even as his own approach, relative to the different contexts of the two endeavours and a corresponding difference in methods employed, has that same dimension. Given foundational claims made in both enterprises, this perhaps gives the sense of inversion which Marx intends. That is to say, for Marx, Hegel's logical foundation is subject to *his*, Marx's foundational claim: that *all* forms of human consciousness, science, art, philosophy, law, religion are founded on the mode of production of material life. Such totalisations or universal quantifications are notoriously tricky. But *in principle* Marx's foundational claim is no less open than Hegel's (minus of course the construal of the absolute idea as a demiurgos). 'In principle': that is if Marx, unlike his Marxist followers, is *not* read as excluding his own thought from the range of applicability of his foundational thesis. That is why, at the outset, I suggested that the nineteenth century paradigm of science that Marx is thinking within, has only so much force. It is then, rather the suspicion of something like a *tertium non datur* being applied to a question of material or ideal foundations that is hard to dispel. The satire of ensouled commodities, read as a *reductio*, is a rejection of Hegel's logical dialectic. To that extent I agree with Althusser's arguments for a radical break between Marx and Hegel.<sup>76</sup> It takes shape here, as Marx makes his own abstract beginning with a logic of capital, and indeed with

<sup>74</sup> Cf. [Burkitt 1984] at 54f.

<sup>75</sup> Cf. [Rose 1981] at 24f for an account of neo-Kantian influences in the emergence of Marxist sociology; on the role of method in Hegel's *Logic* on which however I do not follow her, see *ibid* at 201.

<sup>76</sup> I am unconvinced by Althusser's substitution of Spinoza for Hegel as Marx's "only direct ancestor" ([Althusser and Balibar 1970] at 102). Whether one is better off

the insistence that the commodity is ‘first of all’ to be taken as an external, physical object, although Marx knows well that a ‘cell-form’ is a concept reached by abstraction.

I want then to lead that break back into different ideas of abstraction or concept formation. From early to late Marx rejects Hegel’s *substantivierte Ausageform*, the abstraction or objectification whereby a predicate is turned into a substantive which can then take the place of a grammatical subject.<sup>77</sup> It is Hegel’s way of bracketing out ‘the nature of things’ and constructing a purely logical realm, within which thought confronts only itself.<sup>78</sup> Set firmly against this move, Marx refuses a distinction between logical and historical concepts. It is a rejection of the value or validity in scientific discourse of the artificed construction of a realm of pure thought. The very abstract beginning of *Capital* with its consideration of the agency of forms is, as a logic, the logic of a discourse (political economy) which is itself, for Marx, the product of historically changing social relations of production. Yet formal logic is also in play, willy nilly, whatever Marx thought of it as a science; whether he saw his rejection of Hegel’s logic as an affirmation of the classical paradigm or whether he thought formal logic largely irrelevant to his science. It is in play, regulating reasoning and, at least by default, the ‘it’ is classical logic with its authorisations and permissions regarding assumptions and its incontrovertible laws of thought. The dynamical dimension of Marx’s thought comes from his standpoint of viewing the economic formation of

with a Spinozistic Marx rather than an Hegelian Marx when it comes to the ethics and politics of Marxism is not my question.

<sup>77</sup> He actually uses this device in the section on commodity fetishism, where following the analogy with religion, the text moves into explanatory mode. The following very literal translation which does not follow the verb change in the English translation from “acts practically” to “appears in practice” brings it out. “Only within their exchange, [do] the products of work receive a socially equal objectivity with regard to value which is separated from their perceptibly different objectivity with regard to use. *This splitting of the product of work into a useful thing and a value-thing only comes to act practically (betätigt sich nur praktisch)* when exchange has already attained sufficient extension and importance, so that useful things are produced for exchange, [and] the value-character of the things itself therefore already comes into consideration with their production” (my emphasis; [Marx 1873] at 87; cp. [Marx 1976] at 166). I have puzzled about that but, unable to determine whether it is a conscious play or a standard use of the German language of the times, have had to let it be.

<sup>78</sup> See further [Kerruish 2006] at 35.

society as a process of natural history and that in turn is embedded in his narrative. There is then nothing to liberate the concepts of necessity and freedom from the cultural narrative Marx thinks within, nothing *except* Marx's very own analytic endeavour and it is just this that gives out when he takes flight to the misty realm of religion.

One will lose Marx's whole analytic endeavour with its 'contradiction' between private production and social exchange together with the modal shift from analysis of what the double form of the commodity is to what it does to the disaster of the analogy with religion and its narrative unless these *formal* logical considerations are aired.<sup>79</sup> They must however be complemented by asking what the analytic endeavour gives out *to* if one is not merely to shift the application of a dubious use of *tertium non datur* from ontology to formal logic. It cannot be, Marx has said, in statement or affirmation of his presupposition, natural properties of commodities that constitute their value. It cannot be, I would say, that the relation between classical and non-classical logics is determined by a classical logical principle; not at least if one is attempting to reason otherwise.<sup>80</sup> Certainly, it is here, on a question of formal logic that I revise Marx. It seems to me that he does not come to this question because he is barred from it by his break with Hegel. That is the sense in which I have said that Marx *cannot* accept thought itself as being at odds with itself. It is, I think, equally clear that it is to the political and (anti-)religious normativity of his thought, that the analytic endeavour gives out.

As I have said, all is not lost thereby so far as Marx's scientific work is concerned. The costs will lie elsewhere. The text shifts to explanation, keeping its grip on the actuality of the *Quidproquo*, holding the passionate belief in the unreasonableness of this actuality in the keeping of saying how and why it comes to this actuality. I do not think Marx errs in leading that back into social practices of exchange in political economies with the venerable institution of private ownership of the means of production.

<sup>79</sup> The notion of 'formal logical' used here differs from that cognate to the 'formal logic' which Hegel considered a heap of dead bones. *That* logic is pretty much dead and gone though its place as sanctified canon of reason is held today by first order, classical logic. Taught in standard introductions to logic as the logic of 'correct inference', often enough to enhance or as a course in 'critical thinking' it seems to me to be not the least of academic institutions' contribution to stifling both interest in and critical thinking about formal logic.

<sup>80</sup> For considerations going to this claim see [Kerruish and Petersen 2006] at 79f.

I think rather that a justified focus of critique based on the perception of structural tension between social being as being for others and private production as production in isolation from other producers has been made hostage to a shoddy analogy and a prejudiced naming. That this takes place on the basis of an historical narrative Marx shares, even inherits from Hegel is less ironic than bitter. I shall come back to that and to the thinking of freedom and necessity associated with it very shortly. I want first to add a few words in defence of my claim that all is not lost thereby.

Marx would not be Marx had he not rejected Hegel's logical dialectic and seen in the expressions of exchange-value a similarly mystified expression. The fetish character of the commodity as an instantiation of the objectivity of illusion in the social realm would not have been written up. The insight of the form of universality falling over production relations themselves and in enabling a science also concealing its limitations would not have been articulated. The engagement with the logic of a discourse<sup>81</sup> (or system, situation, field, text, institution) which is Marx's engagement in these early chapters of *Capital* would be lacking.

There is a gap between an imagined or constructed realm of pure thought and the social realm which Marx presupposes. No doubt that gap is as imaginary or as artficed as the realm of pure thought itself and some say that common sense and experimental science get along perfectly well without burdening themselves with such a realm and the gap it opens. Marx to my mind is not of this persuasion and for that reason a most enigmatic materialist. He imagines a new sovereignty and an infinite freedom based on human mastery of nature, on fully rational control of the realm of necessity.<sup>82</sup> I have wondered at that, wondered how on earth Marx could have imagined that humans might escape the incomplete intentionality of their actions, most of all as regards their own needs and desires. But any answer to that question which would not, foolishly, treat Marx as a fool, would need to take up the subjective side

<sup>81</sup> The usage communicates a theoretical sense of what must or does hold or happen in the designated location on account of the nature or structure of that location. It takes in the fetish-character of the commodity as an objective, in the sense of necessary phenomenon of capitalist social relations consistently with the play between the thing-sense and law-sense of 'objectivity' that Marx is confronting.

<sup>82</sup> Sketched in the cited passage from the section on commodity fetishism, this is somewhat expanded towards the end of volume III of *Capital* ([Marx 1971] at 818f; [Marx 1894] at 826f.

of an interpenetrated subjective–objective structure of human knowing and being. I have taken up the side of objectivity: an objectivity that plays between that said to be founded on laws of one type and another and an objectivity that is tied to external objects that are what they are independent of human intercourse with them in thinking and knowing.

That is evidently one-sided and partial. Its motivation, and my own suspension of the political dimensions of Marx's thought is itself political. The narrative that Marx brings to the social realm is the more easily denounced than displaced and I see no point adding more denunciations in the face of that. I rather follow Marx (against Marx) in saying that this narrative and the consciousness it speaks out is not sheer, mere fiction without its own conditions of production and reproduction in social relations and practices. And further and still in a way, following Marx: what it does, its justificatory function as applied to the crimes of colonialism is inseparable from their commission and repetition. Evidently I reject the analogy with religion. As fitting the narrative it is implicated in it. As imposed by the narrative on Marx's formal conceptualisation of the social realm, it removes the sense of fetish phenomena from the double character of the commodity. But it is not that I do not have my own normative point. I am hoping to persuade my reader that, in addition to counter-narratives, critiques and deconstructive arguments, plays and performances, theory needs to pursue logical/metaphysical issues against the justificatory function of a narrative that does not go away. Science, a discourse or form of knowing that gets the name of objective knowledge because it is a powerful instrument of change is part of that hope. A science of logic Hegelian in spirit, mathematical in reasoning and tools, that investigates the very notion of 'objectivity' that, to my mind, imbues and redeems the idea of the fetish character of the commodity, opens a realm of investigation that takes the place of Marx's analogy.

In sum: my argument is that in rejecting Hegel's logical dialectic, Marx wittingly/unwittingly disarms himself in the face of his own apprehension of that very excess of thought that haunts logics of the understanding. He disarms himself of the question Hegel asked, however obscurely, of those laws of thought that were once thought, by logicians, philosophers and mathematicians alike to be unquestionable. It rests on the broader contention that the substitution or *Quidproquo* that Marx names fetishism is a particular instance of a dialectic of form and content

which shows up in the antinomies, incompleteness and undecidability results of twentieth century higher order logic and set theory. The perfectly innocent thing — sign, number, word, table — used daily without ever giving anybody any trouble, is in no way innocent. It escapes the control of ordinary use and ordinary language, defies conventions, ignores bans, plays up, plays havoc, not with common sense, that withstands such inordinacy, but with the unquestionability of ‘laws of thought’. It has another character, a *Doppelgänger*. It is itself and it stands for something else. It is used to calculate and to code. It revels in ambiguity. It is the solid thing at which we take our meals and it encodes distinctions of consumption and taste. One can celebrate and make fine plays of all this, particularly in the realm of the symbolic. One can insist on control and try banning use–mention and other ‘confusions’. One can take refuge in what some seem to think is the earthly saviour of social theory, common sense. Or, and here one can, in principle, be with both Marx and Hegel, seek to fashion from it a tool, a theory or a logic as a tool of objective insights.

How would that look as concerns legal theory? I have no programmatic answer to this question. I can only say that I would revise Marx’s thesis to contend that fetish phenomena are actual and reasonable/unreasonable, deferring its equivocation into law’s claim to authoritative determinations of right, into a question of law’s reasonableness, and developing that to an idea of the wrong of law.

## References

- [Althusser 1967] Louis Althusser. Contradiction and overdetermination. *New Left Review*, 41:15–35, 1967.
- [Althusser and Balibar 1970] Louis Althusser and Etienne Balibar. *Reading Capital*, tr. Ben Brewster. Verso, London, 1979.
- [Apter and Pietz 1993] Emily Apter and William Pietz, editors. *Fetishism as Cultural Discourse*. Cornell University Press, Ithaca and London, 1993.
- [Böhme 2001] Hartmut Böhme. Das Fetischismus-Konzept von Marx und sein Kontext. In Volker Gerhardt, editor, *Marxismus. Versuch einer Bilanz*, pages 289–318. Scriptorum Verlag, Magdeburg, 2001.



- [Burkitt 1984] Brian Burkitt. *Radical Political Economy: an Introduction to the Alternative Economics*. Wheatsheaf Books Ltd, Brighton, Sussex, 1984.
- [Callinicos 1983] Alex Callinicos. *Marxism and Philosophy*. Oxford University Press, Oxford, 1983.
- [Derrida 1994] Jacques Derrida. *Specters of Marx*, tr. Peggy Kamuf. Routledge, New York and London, 1994.
- [Dragstedt 1976] Albert Dragstedt, editor. *Value: Studies by Karl Marx*. New Park Publications, London, 1976.
- [Frege 1879] G. Frege. Begriffsschrift, a formula language, modeled upon that of arithmetic, for pure thought. In [van Heijenoort 1967], pages 5–82. 1879.
- [Glockner 1927/30] Hermann Glockner, editor. *Hegel Sämtliche Werke. Jubiläumsausgabe in zwanzig Bänden*. Friedrich Frommann Verlag (Günther Holzboog), Stuttgart-Bad Cannstatt, 1961–1968.
- [Hegel 1807] Georg Wilhelm Friedrich Hegel. *Phänomenologie des Geistes*, volume 2 of [Glockner 1927/30].
- [Hegel 1812] Georg Wilhelm Friedrich Hegel. *Wissenschaft der Logik. Erster Teil. Die objektive Logik*, volume 4 of [Glockner 1927/30].
- [Hegel 1816] Georg Wilhelm Friedrich Hegel. *Wissenschaft der Logik. Zweiter Teil. Die subjektive Logik oder Lehre vom Begriff*, volume 5 of [Glockner 1927/30].
- [Hegel 1969] Georg Wilhelm Friedrich Hegel. *Hegel's Science of Logic*, tr. A. V. Miller. George Allen & Unwin Ltd, London, 1969. Reprinted by Humanities Press, New York, 1976.
- [Hegel 1977] Georg Wilhelm Friedrich Hegel. *Hegel's Phenomenology of Spirit*, tr. A. V. Miller. Oxford University Press, Oxford, 1977.
- [Kerruish 1991] Valerie Kerruish. *Jurisprudence as Ideology*. Routledge, London, 1991.
- [Kerruish 2006] Valerie Kerruish. On re-staging the universal: Butler, Hegel and contesting the closure of logic. *Dilemmata: Jahrbuch der ASFPG*, 1:23–60, 2006.
- [Kerruish and Petersen 2006] Valerie Kerruish and Uwe Petersen. Philosophical sanity, mysteries of the understanding and dialectical logic. *Dilemmata: Jahrbuch der ASFPG*, 1:61–91, 2006.
- [Luxemburg 1913] Rosa Luxemburg. *The Accumulation of Capital*, tr. Agnes Schwarzschild. Routledge and Kegan Paul Ltd, London, 1951.

- [Markus 1986] Gyorgy Markus. *Language and Production: A Critique of the Paradigms*. D. Reidel, Dordrecht, 1986.
- [Marx 1970] Karl Marx. *Critique of Hegel's 'Philosophy of Right'*, tr. Annette Jolin and Joseph O'Malley. Cambridge University Press, Cambridge, 1970.
- [Marx 1857–1858] Karl Marx. *Grundrisse der Kritik der politischen Ökonomie. Rohentwurf*. Dietz Verlag, Berlin, 1974.
- [Marx 1859] Karl Marx. *Zur Kritik der politischen Ökonomie*. Verlag von J. H. W. Dietz, Stuttgart, 1907.
- [Marx 1867] Karl Marx. The Form of Value. In [Dragstedt 1976], pages 49–70, 1867.
- [Marx 1873] Karl Marx. *Das Kapital: Kritik der politischen Ökonomie I. Karl Marx Friedrich Engels Werke*, volume 23. Dietz Verlag, Berlin, 1972.
- [Marx 1879] Karl Marx. Marginal Notes on Wagner. In [Dragstedt 1976], pages 195–229, 1879.
- [Marx 1894] Karl Marx. *Das Kapital: Kritik der politischen Ökonomie III. Karl Marx Friedrich Engels Werke*, volume 25. Dietz Verlag, Berlin, 1973.
- [Marx 1971] Karl Marx. *Capital: A Critique of Political Economy III*. Progress Publishers, Moscow, 1971.
- [Marx 1973] Karl Marx. *Grundrisse: Foundations of the Critique of Political Economy (Rough Draft)*, tr. Martin Nicolaus. Penguin Books, Harmondsworth, 1973.
- [Marx 1976] Karl Marx. *Capital: A Critique of Political Economy I*, tr. B. Fowkes. Penguin Books, London, 1976.
- [Marx 1889] Karl Marx. *Capital: A Critical Analysis of Capitalist Production I*, tr. T. S. Moore and E. Aveling. George Allen and Unwin, London, 1938.
- [Nancy 1993] Jean-Luc Nancy. The heart of things. In *The Birth to Presence*, tr. Brian Holmes and others, pages 167–188. Stanford University Press, Stanford, California, 1993.
- [Nancy 2002] Jean-Luc Nancy. *Hegel: the Restlessness of the Negative*, tr. Jason Smith and Steven Miller. University of Minnesota Press, Minneapolis and London, 2002.
- [Pietz 1993] William Pietz. Fetishism and materialism: The limits of theory in Marx. In [Apter and Pietz 1993], pages 119–151, 1993.

- [Rose 1981] Gillian Rose. *Hegel contra Sociology*. Athlone Press, London, 1981.
- [van Heijenoort 1967] van Heijenoort, editor. *From Frege to Gödel: a Source Book in Mathematical Logic 1879–1931*. Harvard University Press, Cambridge, MA, 1967.



# Violence and Negativity in Hegel's *Phenomenology of Spirit*

TARIK KOCHI\*

“In the beginning was the deed”.<sup>1</sup>

## Introduction

This paper offers a reading of the opening chapters of G.W.F. Hegel's *Phenomenology of Spirit* (1807).<sup>2</sup> Hegel, inheriting the notion of the spontaneity of thinking from Immanuel Kant, and the notion of the self's positing of itself from Johann Gottlieb Fichte, develops a radical and original account of the metaphysics of subjectivity. Central to Hegel's account, yet often overlooked, is an approach which begins to think seriously about the relation between thinking and action, and in particular, the relationship between negativity and negation within thinking and violence within the world. Hegel's account remains relevant to contemporary attempts to develop a philosophy of the subject and a philosophy of ethics. Each of these attempts would do well to consider the role and consequences of negativity and violence within subjectivity.

\* Lecturer, School of Law, Queen's University Belfast, Northern Ireland. Email: T.Kochi@qub.ac.uk. Thanks to the Altonaer Stiftung für philosophische Grundlagenforschung (ASFPG) for its financial support and hospitality. Thanks also to Valerie Kerruish for her comments and to both Valerie and Uwe Petersen for the many conversations in Altona, Hamburg, which have contributed to this paper.

<sup>1</sup> Goethe, J.W. *Faust*, (München: Deutscher Taschenbuch Verlag, 1997), p. 67. Goethe's original German reads: "... *im Anfang war die That!*"

<sup>2</sup> Hegel, G.W.F. *Phenomenology of Spirit*, Miller, A.V. tr. (Oxford: Oxford University Press, 1977).

## I

In the *Critique of Pure Reason* (1781/1788) Kant argues that cognition arises from two sources in the mind, the first involves the reception of representations (intuition), the second involves an active faculty of cognising an object by means of these representations. Kant argues that the latter notion involves a certain ‘spontaneity’ of cognition and refers to it in a general sense as the faculty of the understanding (*Verstand*).<sup>3</sup> This spontaneous activity of thinking is crucial for Kant. It occurs as a fundamental act of self-emanation without which the cognition of sensible representations would not be possible. Kant writes:

The *I think* must be able to accompany all my representations; for otherwise something would be represented in me that could not be thought at all, which is as much as to say that the representation would either be impossible or else at least would be nothing for me. That representation that can be given prior to all thinking is called *intuition*. Thus the manifold of all intuition has a necessary relation to the *I think* in the same subject in which this manifold is to be encountered. But this representation is an act [*Actus*] of *spontaneity*, i.e. it cannot be regarded as belonging to sensibility. I call it the *pure apperception*, in order to distinguish it from the *empirical* one, or also the *original apperception*, since it is that self-consciousness which, because it produces the representation *I think*, which must be able to accompany all others and which in all consciousness is one and the same, cannot be accompanied by any further representation. I also call its unity the *transcendental* unity of self-consciousness in order to designate the possibility of *a priori* cognition from it. For the manifold representations that are given in a certain intuition would not altogether be *my* representations if they did not altogether belong to a self-consciousness; i.e., as my representations (even if I am not conscious of them as such) they must yet necessarily be in accord with the condition under which alone they *can* stand together in a universal self-consciousness, because otherwise they

<sup>3</sup> Kant, I. *Critique of Pure Reason*, Guyer, P. & Wood, A. eds. tr. (Cambridge: Cambridge University Press, 1998), p. 193, A51/B75.

would not throughout belong to me. From this original combination much may be inferred.<sup>4</sup>

For Kant the 'I think' expresses the act of the thinking self's determination of its existence. It is through the *spontaneous act of thinking* that there is such a thing as intelligence.<sup>5</sup> This activity underlying human intelligence bears degrees of similarity with an intellectual activity that is claimed by Kant as the basis of transcendental freedom within his resolution of the 'third antinomy of pure reason'.<sup>6</sup> While each activity should not be reduced to the other, for Kant, there seems to be something significant that is shared between these two activities — a spontaneous activity of the 'I' residing at the basis of both human intelligence and human freedom.

An account of the spontaneity of thinking as an *act* or *deed* is inherited and given a radical emphasis by Fichte and his broad notion of

<sup>4</sup> Ibid, pp. 246–247, B132–B133.

<sup>5</sup> Ibid, p. 260, B158, in footnote.

<sup>6</sup> Ibid, p. 533, A533/B561. Kant argues:

By freedom in the cosmological sense, on the contrary, I understand the faculty of beginning a state *from itself*, the causality of which does not in turn stand under another cause determining it in time in accordance with the law of nature. Freedom in this signification is a pure transcendental idea, which, first, contains nothing borrowed from experience, and second, the object of which also cannot be given determinately in any experience, because it is a universal law — even of the possibility of all experience — that everything that happens must have a cause, and hence that the causality of the cause, as *itself having happened* or arisen, must in turn have a cause; through this law, then, the entire field of experience, however far it may reach, is transformed into the sum total of mere nature. But since in such a way no absolute totality of conditions in causal relations is forthcoming, reason creates the idea of spontaneity, which could start to act from itself, without needing to be preceded by any other cause that in turn determines it to action according to the law of causal connection.

It is especially noteworthy that it is this *transcendental* idea of *freedom* on which the practical concept of freedom is grounded, and the former constitutes the real moment of the difficulties in the latter, which have long surrounded the question of its possibility. *Freedom in the practical sense* is the independence of the power of choice from *necessitation* by impulses of sensibility. For a power of choice is *sensible* insofar as it is *pathologically affected* (through moving-causes of sensibility); it is called *animal* power of choice (*arbitrium brutum*) if it can be *pathologically necessitated*. The human power of choice is indeed an *arbitrium sensitivum*, yet not *brutum* but *liberum*, because sensibility does not render its action necessary, but in the human being there is a faculty of determining oneself from oneself, independently of necessitation by sensible impulses.

thinking as ‘*Tathandlung*’. In the *Science of Knowledge* (1794)<sup>7</sup> Fichte develops an account of the thinking self that is not originally reliant upon a sensible or empirical world but which creates itself or emanates from itself as an act of *self-positing*. For Fichte, while the activity of thinking or the intellect is spontaneous, it is not chaotic or lawless. Fichte argues that the intellect acts, but owing to its nature it can only act in a certain fashion. In this sense, for Fichte, there is a certain “law of action” of the intellect and there are “necessary laws of the intellect”.<sup>8</sup>

Taking the place of, or re-interpreting, Kant’s notion of the ‘transcendental unity of apperception’, Fichte introduces a conception of the activity of the thinking intellect as the act of *positing* (*setzen*). Such a move is seen as necessary so that the categories of the understanding (within the framework of Kant’s architectonic) can actually be derived. For Fichte the intellect or the self is not posited by something else but is posited by itself, it is an act of *self-positing*. He writes:

The self’s own positing of itself is thus its own pure activity. The *self posits itself*, and by virtue of this mere self-assertion it *exists*; and conversely, the self *exists* and *posits* its own existence by the virtue of merely existing. It is at once the agent and the product of action; the active, and what the activity brings about; action and deed are one and the same, and hence the ‘I am’ expresses an Act, and the only one possible, as will inevitably appear from the Science of Knowledge as a whole.<sup>9</sup>

For Fichte this conception of thinking being an act which posits itself cannot be proved empirically.<sup>10</sup> Rather, it arises as a form of intellectual intuition (*intellektuelle Anschauung*) which Fichte compares to the form

<sup>7</sup> Fichte, J.G. *The Science of Knowledge*, Heath, P. and Lachs, J. ed. tr. (Cambridge: Cambridge University Press, 1982).

<sup>8</sup> *Ibid*, p. 21.

<sup>9</sup> Fichte, J.G. *Science of Knowledge*, p. 97.

<sup>10</sup> *Ibid*, Fichte, at p. 93 argues:

Our task is to *discover* the primordial, absolutely unconditioned first principle of all human knowledge. This can be neither *proved* nor *defined*, if it is to be an absolutely primary principle.

It is intended to express that *Act* which does not and cannot appear among the empirical states of our consciousness, but rather lies at the basis of all consciousness and alone makes it possible.



of sensible intuition (*sinnliche Anschauung*) of space and time presupposed by Kant.<sup>11</sup> On this notion of intellectual intuition Fichte argues:

This intuiting of himself that is required of the philosopher, in performing the act whereby the self arises for him, I refer to as *intellectual intuition*. It is the immediate consciousness that I act, and what I enact: it is that whereby I know something because I do it. We cannot prove from concepts that this power of intellectual intuition exists, nor evolve from them what it may be. Everyone must discover it immediately in himself, or he will never make its acquaintance. The demand to have it proved for one by reasoning is vastly more extraordinary than would be the demand of a person born blind to have it explained to him what colours are, without his needing to see.

Everyone, to be sure, can be shown, in his own admitted experience, that this intellectual intuition occurs at every moment of his consciousness. I cannot take a step, move a hand or foot, without an intellectual intuition of my self-consciousness in these acts; only so do I know that *I* do it, only so do I distinguish my action, and myself therein, from the object of action before me. Whosoever ascribes an activity to himself, appeals to this intuition. The source of life is contained therein, and without it there is death.<sup>12</sup>

Fichte attempts to outline the basic operation of the self-positing of the self (*das Ich*). He begins with the basic proposition 'A is A' or 'A = A' and its suggestion that 'if A exists, then A exists'.<sup>13</sup> For Fichte, the necessary connection between 'if' and 'then' (named 'X') expresses a form of positing, in the sense that 'A' is posited, therefore 'A' is.<sup>14</sup> Fichte extends this conception to include a universal notion of self abstracted from empirical content, expressed as 'I = I' or 'I am I'.<sup>15</sup> Thus:

But the proposition 'I am I' has a meaning wholly different from that of 'A is A'. For the latter has content only under a certain condition. *If* A is posited, it is naturally posited *as* A, as having

<sup>11</sup> Ibid, pp. 49–51.

<sup>12</sup> Ibid, p. 38.

<sup>13</sup> Ibid, p. 94.

<sup>14</sup> Ibid, p. 95.

<sup>15</sup> Ibid, p. 96.

the predicate A. But this proposition still tells us nothing as to *whether* it actually is posited, and hence whether it is posited with any particular predicate. Yet the proposition ‘I am I’ is unconditionally and absolutely valid, since it is equivalent to the proposition X; it is valid not merely in form but also in content. In it the I is posited, not conditionally, but absolutely, with the predicate of equivalence to itself; hence it really *is* posited, and the proposition can be expressed as *I am*.<sup>16</sup>

This ‘I’ for Fichte, as the spontaneous act of the intellect, is grounded upon its own activity, it is “absolutely posited”.<sup>17</sup> For Fichte the ‘I’ refers to the “absolute subject”: “That whose being or essence consists simply in the fact that it posits itself as existing”.<sup>18</sup> Following the notion of the positing of identity Fichte introduces the notion of the positing of opposition, through the proposition ‘ $\sim A$  is not equal to A’ ( $\sim A \neq A$ ; which involves:  $\sim A = \sim A$ ).<sup>19</sup> Fichte argues that if any  $\sim A$  were posited this would presuppose an A, and in this sense the  $\sim A$  would need to be understood as materially conditioned. For Fichte, in the positing of an  $\sim A$ , an A must also be posited.<sup>20</sup> Fichte writes:

As surely as the absolute certainty of the proposition ‘ $\sim A$  is not equal to A’ is unconditionally admitted among the facts of em-

<sup>16</sup> Ibid.

<sup>17</sup> Ibid, p. 97.

<sup>18</sup> Ibid, p. 98. Fichte at p. 99 states: “The self begins by an absolute positing of its own existence”. At pp. 100–1 Fichte argues further:

That our proposition is the absolutely basic principle of all knowledge, was pointed out by *Kant*, in his deduction of the categories; but he never laid it down specifically *as* the basic principle. *Descartes*, before him, put forward a similar proposition: *cogito, ergo sum* — which need not have been merely the minor premise and conclusion of a syllogism, with the major premise: *quodcumque cogitat est*; for he may very well have regarded it as an immediate datum of consciousness. It would then amount to *cogitans sum, ergo sum* (or as we should say, *sum, ergo sum*). But in that case the addition of *cogitans* is entirely superfluous; we do not necessarily think when we exist, but we necessarily exist whenever we think. Thinking is by no means the essence, but merely a specific determination of existence; and our existence has many other determinations besides this.

<sup>19</sup> Ibid, p. 103.

<sup>20</sup> Ibid.

pirical consciousness, *so surely is a not-self opposed absolutely to the self.*<sup>21</sup>

From this basis Fichte puts forward the notion that the positing of every A involves the implicit positing of the  $\sim A$  and vice versa. Fichte extends this to the notion that the positing of the self (I) involves the implicit positing of the not-self ( $\sim I$ ) and that the self is posited through the simultaneous positing of both itself and what is other than itself, the not-self. For Fichte when the self posits itself through its opposite it is positing its limit.<sup>22</sup> He argues that: "To *limit* something is to abolish its reality, not *wholly* but in *part* only, by negation".<sup>23</sup> Such limitation through negation points to the notion of divisibility and the notion that the self and not-self are posited as divisible.<sup>24</sup> Fichte argues that: "In the self I oppose a divisible not self to the divisible self".<sup>25</sup>

The notion of the spontaneous activity of thinking which posits both itself and its not-self or 'other' is inherited by Hegel.<sup>26</sup> Further, the central idea of thinking as a deed is taken up by Hegel and given a radical re-interpretation, especially in relation to the operation of negation. For Kant and Fichte the deed of thinking, its self-positing and negation of its self and its other, took place purely in the realm of thinking. In the *Science of Logic* (1812–16)<sup>27</sup> Hegel is similarly concerned with the realm

<sup>21</sup> Ibid, p. 104.

<sup>22</sup> Ibid, p. 108. Fichte states:

Hence, as before, we must make an experiment and ask: How can A and  $\sim A$ , being and nonbeing, reality and negation, be thought together without mutual elimination and destruction? ... We need not expect anyone to answer the question other than as follows: They will mutually *limit* one another.

<sup>23</sup> Ibid.

<sup>24</sup> Ibid.

<sup>25</sup> Ibid, p. 110.

<sup>26</sup> Significant in the movement between Fichte and Hegel is the philosophy of Schelling. See: Schelling, F.W.J. *System of Transcendental Idealism*, Heath, P. tr. (Charlottesville: University Press of Virginia, 2001). For more general accounts of the philosophical movements between Kant and Hegel see: Henrich, D. *Between Kant and Hegel*, Pacini, D.S. ed. (Cambridge, Mass.: Harvard University Press, 2003); Di Giovanni, G. *Between Kant and Hegel: Texts in the Development of Post-Kantian Idealism*, (Indianapolis: Hackett, 2000); Kolakowski, L. *Main Currents of Marxism*, vol. 1. Falla, P.S. tr. (Oxford: Oxford University Press, 1981).

<sup>27</sup> In the first half of the *Logic* Hegel attempts to derive the categories of the understanding. He shows how contradiction is contained at the heart, not just at the peripheries of pure reason. See: Hegel, G.W.F. *Hegel's Science of Logic*, Miller, A.V. tr.

of pure reason and demonstrates, contra Kant, that contradictions are not confined to merely the four ‘cosmological antinomies of pure reason’ but are contained within every act of thinking. For Hegel pure reason, as with all human thinking, is *antinomical*, it is characterised by *contradiction* and hence, transcendental philosophy needs to comprehend the role, operation and activity of contradiction within pure reason.<sup>28</sup> While Hegel in his *Logic* is concerned with the realm of pure reason, the relation between his *Phenomenology of Spirit* and the *Logic* is such as to make his undertaking consistent with the activity of actual thinking within the living, social world in the realm of *spirit* (*Geist*). In the *Phenomenology of Spirit* Hegel shows how the deeds of thinking (drawn into view by Kant and Fichte) occur also as physical, social deeds within an inter-subjective and inter-institutional human world. By making such a move Hegel forges an original and crucial link between the tradition of transcendental philosophy, negativity and the operation of human violence. The link has consequences for any account of morality or ethics that presupposes or develops historically from Kant’s notion of transcendental freedom.

## II

In the *Phenomenology of Spirit* Hegel inherits the necessary relationship occurring in the positing of the self between self and not-self, or between self and other, developed by Fichte. Hegel takes this relationship, which for Fichte is implicit in both the notion and existence of the ‘I’, and places it back within a living world of not only sense experience and perception but, also a world containing physical materiality and living, breathing other beings. A notion of ‘world’ is not added to Fichte’s account of the

(London: Allen and Unwin, 1969); Hegel, G.W.F. *The Encyclopaedia Logic: Part I of the Encyclopaedia of Philosophical Sciences with the Zusätze*, Geraets, T.F., Suchting, W.A. and Harris, H.S. tr. (Indianapolis: Hackett, 1991).

<sup>28</sup> Hegel, G.W.F., *Encyclopaedia Logic*, §48, states:

The main point that has to be made is that antinomy is found not only in the four particular objects taken from cosmology, but rather in *all* objects of all kinds, in *all* representations, concepts, and ideas. To know this, and to be cognizant of this property of objects, belongs to what is essential in philosophical study; this is the property that constitutes what will determine itself in due course as the *dialectical* moment of logical thinking.

self, it is already implicit in the notion of the not-self, which must include everything *other* than the self. Hegel simply makes this more explicit and asks the question: how does the self come to be itself as distinct from its not-self? Hegel's account asks further: how does the self precisely come to draw the limits of itself from its other and thus *become conscious of itself* in a way that it might be *certain* that what it *intuits* as itself (the 'I') is actually itself and not, in fact, something other?

For Hegel if we might have an immediate sense or intellectual intuition of a particular 'I' that simply *is*, or that exists, then there arises something of a philosophical problem of how this specific, immediate 'I' can be distinguished from something else. He claims that if we try to identify the immediate 'I' that we *mean*, then it is taken as something different from something else, it is a 'this' rather than a 'that'.<sup>29</sup> Hegel argues that there is a difference between the 'this' or the object that we might *mean* and the 'this' or the object that is expressed through words. Hegel argues:

It is as a universal too that we *utter* what the sensuous [content] is. What we say is: 'This', i.e. the *universal* This; or 'it is', i.e. *Being in general*. Of course, we do not *envisage* the universal This or Being in general, but we *utter* the universal; in other words, we do not strictly say what in this sense-certainty we *mean* to say. But language, as we see, is the more truthful; in it, we ourselves directly refute what we *mean* to say, and since the universal is the true [content] of sense-certainty and language expresses this true [content] alone, it is just not possible for us ever to say, or express in words, a sensuous being that we *mean*.<sup>30</sup>

In one sense Hegel can be understood as arguing against a tradition that tries to ground transcendental philosophy upon a notion of *intellectual intuition* and demands that if an account of self-consciousness is to be a subject of philosophy (rather than a subject of aesthetics) then it needs to be explained through a philosophical language. In another sense, by suggesting the 'truthfulness' of language Hegel draws out a speculative theme latent in Fichte but not fully brought to bear in philosophy. The theme involves the notion of *mediation* in general which is most apparent in the use of symbols and words within language which have meaning as points

<sup>29</sup> Hegel, *Phenomenology of Spirit*, p. 58.

<sup>30</sup> *Ibid*, p. 60.

of designation and universality only in some form of relation to their others. The relation between self and not-self occurs as one basic speculative form, each term has meaning only in its relation of distinction from the other, and further, when one is *meant*, a sense of the distinction from the other, a reference to this other, and thus, some aspect or representation of the other, is contained within this meaning. For Hegel, comprehending an object concretely, rather than abstractly, involves grasping how its content, or the object itself, is mediated through such a relation.

From Hegel's viewpoint Kant and Fichte have only presented an abstract account of the 'I' or self-consciousness. They have not yet grasped a concrete account of self-consciousness which would come about through paying attention to how self-consciousness is mediated. Part of the structure of Hegel's *Phenomenology of Spirit* presents a speculative account of self-consciousness by moving from the abstract to the more concrete. Hegel can be seen to do this by drawing attention to the operation of mediation (*Vermittlung*), negation and the manner in which the knowledge of an object gains a greater speculative or determinate content when viewed in terms of its necessary inter-relations with its others. In relation to the Fichtean account of the self, this involves comprehending its self-positing in relation to its not-self, this being a world of physical objects and other intelligent human subjects.

For Hegel, the Fichtean self-consciousness exists in a world where it has some immediate sense of itself, the 'I', and where it is confronted by something that is not itself, a world of appearance. In its existence self-consciousness posits both itself and its not-self, however, it is not completely certain of the boundaries or limits of itself. That is, it cannot be certain of the limits of the self, as determined by the not-self, because this boundary is marked (for Kant and Fichte) by the uncertain and often erroneous world of appearance. Hegel argues that the self wants to become fully conscious of what it is, that is, it wants to become self-conscious and wants to know itself with certainty rather than vagueness.<sup>31</sup> Hegel calls this demand, need or drive that the self know itself and know itself as an identity or unity that is distinct from its not-self, 'desire' (*Begierde*) in general.<sup>32</sup>

<sup>31</sup> Ibid, p. 105.

<sup>32</sup> Ibid.

By placing the abstract 'I' in the world Hegel positions the self within a broader category of 'life'.<sup>33</sup> Positioned within a world of life and driven by a desire to know itself with certainty, the self, in the manner of Kant, cannot withdraw from a world of appearance and turn back into itself. If the self withdrew from the world of appearance into itself it would never be able to know with any certainty the limits or boundaries of itself — the self would not *know* where *it* ends and where the shadowy world of appearance begins. Rather, for Hegel, the only means by which a self might come to know what *it is*, is through the experience of learning the limits of itself through becoming aware of the independence of its objects, that is, the independence of what *is* the not-self.<sup>34</sup> For Hegel it is through the outward mediated *experience* of the world and not through some form of purely *internal* knowing or intellectual intuition that a consciousness becomes a self-consciousness. This is not to say that Hegel rejects entirely any immediate knowledge of self. Rather, he draws attention the fact that any form of immediate knowledge is *always* mediated through the self's experience of its implicit relation with the external world.<sup>35</sup>

One aspect of the speculative account Hegel presents involves the sense that through the self's experience of the not-self there occurs a radical change in both the self's knowledge of itself and the self's knowledge of the world. By experiencing a certain independence of an object in the world, for example, in the way that it needs to be moved, or that it can be eaten or chopped down, the self comes to a firmer, more detailed, and more concrete knowledge of its objects. Further, through this greater knowledge of the objects of the not-self, the self comes to a firmer, more concrete and more detailed knowledge of itself: a knowledge of what it *is* through reflecting upon the character, qualities and forms contained within the relations and mediations between itself and what is other than itself. What was initially a very abstract concept of the self's being, the simple assertion 'I am' has now taken on a more concrete content. What being is for the 'I' is radically different after the process of experiencing and its own reflection upon this process. By drawing attention to how the self's knowledge of itself is *constituted* through its experience of

<sup>33</sup> Ibid, p. 106.

<sup>34</sup> Ibid.

<sup>35</sup> On the relation between immediacy and mediation in Hegel's philosophy see the discussion in: Henrich, D. *Hegel im Kontext*, (Frankfurt am Main: Suhrkamp, 1971).

mediation with the not-self and its reflection upon this experience, Hegel draws attention to the *speculative nature* of knowing and the production of knowledge. In this both the knowing *subject* and the known *object* are constituted through each other. For Hegel, grasping the *subject-object* character of all knowledge is central to the task of any critical, or now, speculative philosophy.

For Fichte, the self's relation to its not-self involved the operation of negation. The self distinguished or determined itself and maintained its self-identity through the negation of what was other than it. In Fichte this self-inscription of the unity of the 'I' through negation took place purely at the level of thinking. Negation was an act of thinking and as taking place within the realm of thinking, when the not-self was 'negated', no physical harm was done. For Hegel, if we are to consider what self-consciousness in fact *is* then the self needs to be understood as positioned firmly within the physical and social world, within *life*, where the self's demand for unity and certainty of self-identity operates as desire. Placed within the realm of life the self's negation of the object occurs no longer purely as an act of thinking. Rather, this act of thinking occurs also as a physical act whereby the negation of the object by the self involves harm, destruction, killing, and violence. For the self which has not yet understood the speculative connection between itself and the other this negation is easy, it sees the object merely as the negative of itself (without a positive content of its own) and treats it accordingly. Hegel argues:

Certain of the nothingness of this other, it explicitly affirms that this nothingness is *for it* the truth of the other; it destroys the independent object and thereby gives itself the certainty of itself as a *true* certainty, a certainty which has become explicit for self-consciousness itself *in an objective manner*.<sup>36</sup>

For Hegel the desire for a self to fully become certain of itself, to become conscious of itself, does not gain satisfaction in the self's relation to the bare world of objects. Part of this rests in the fact that these

<sup>36</sup> Hegel, *Phenomenology of Spirit*, p. 109. A certain sense of this violence is picked up by Derrida in the essay 'Violence and Metaphysics'. Derrida's interpretation of Hegel, however, appropriates Hegel's philosophy for his own purposes and presents an account of Hegel's philosophy as inadequate for contemporary concerns. See: Derrida, J. 'Violence and Metaphysics: An Essay on the Thought of Emmanuel Levinas' in *Writing and Difference*, Bass, A. tr. (Chicago: University of Chicago Press, 1978).



objects have only a limited independence in relation to the self, and hence they can be negated relatively easily in the self's assertion of its self-identity. Where the self does however gain satisfaction, argues Hegel, where it becomes self-conscious, is when it confronts another intelligent self: when the opposing not-self or, other, is an intelligent self. Hegel argues: "Self-consciousness achieves its satisfaction only in another self-consciousness".<sup>37</sup> For Hegel:

A self-consciousness exists *for a self-consciousness*. Only so is it in fact self-consciousness; for only in this way does the unity of itself in its otherness become explicit for it. The 'I' which is the object of its Notion is in fact not 'object'; the object of Desire, however, is only independent, for it is the universal indestructible substance, the fluid self-identical essence. A self-consciousness, in being an object, is just as much 'I' as 'object'. With this, we already have before us the Notion of *Spirit*. What still lies ahead for consciousness is the experience of what Spirit is — this absolute substance which is the unity of the different independent self-consciousnesses which, in their opposition, enjoy perfect freedom and independence; 'I' that is 'We' and 'We' that is 'I' [*ich, das Wir, und Wir, das Ich ist*]. It is in self-consciousness, in the Notion of Spirit, that consciousness first finds its turning point, where it leaves behind it the colourful show of the sensuous here and now and the nightlike void of the supersensible beyond, and steps out into the spiritual daylight of the present.<sup>38</sup>

For Hegel, the self only finds some form of certainty in its pursuit of self-knowledge when it begins to comprehend itself in relation to another intelligent self that stands before it as the not-self. The self may only comprehend its unity or self-identity through a comprehension of how it is *constituted* through a relation of *mediation* with the other. The relation of mediation between self and other constitutes the limits of the self; the boundaries of the self are drawn or produced through the active relation of mediation between self and other, and through the self's reflection upon this relation as the operation of self-inscription and self-positing.

In this respect the self is not something that exists before coming into contact with the world but is itself *produced* through its relation of

<sup>37</sup> Hegel, G.W.F. *Phenomenology of Spirit*, p. 110.

<sup>38</sup> *Ibid*, pp. 110–1.

mediation with its others. On this side of thinking and self-reflection the creative, constitutive, spontaneous act of thinking occurs *as mediation itself* — it is mediation (*Vermittlung*). In a sense, for Hegel, the self or 'I' is always mediated, its act of thinking is always an act of mediation with its self and with its other. The 'I' is a middle term, it is a middle and process of mediation between the self and the other. From such a standpoint the 'I' needs to be understood as a 'We'. The self cannot be thought as the merely atomistic individual but is always a process of mediation such that the concept of 'self' now includes within it the process of co-constitution between intelligent beings. For Hegel, self-consciousness is a relation between immediacy and mediation, a relation between the drive of the immediate 'I' for its self-identity and the reflection upon the constitution of selves through their others as a 'We': a relation between the notions of the *subject as spirit* and *spirit as subject*.

For Hegel, understood as a relation between immediacy and mediation, the attempt of the self to know itself concurrently involves the effort to see-into or know the other. Only through knowing what the other is, might the self more firmly comprehend itself and come to understand the content and limits of itself. Hegel calls the process of attempting to comprehend the not-self, or the other, the process of recognition (*Anerkennung*).<sup>39</sup> Hegel sketches two forms or outcomes of this process which are not mutually exclusive.<sup>40</sup> The first involves a successful moment of

<sup>39</sup> The term 'recognition' is taken by Hegel from Fichte, though Hegel's philosophical employment of the term is far more radical. See: Fichte, J.G. *Foundations of Natural Right*, Neuhouser, F. ed. Baur, M. tr. (Cambridge: Cambridge University Press, 2000).

<sup>40</sup> For an account which emphasizes these two moments of recognition see: Williams, R.R. *Hegel's Ethics of Recognition*, (Berkeley: University of California Press, 1997). Contrast this to the account given in Kojève, A. *Introduction to the Reading of Hegel*, Nichols, J.H. tr. (Ithaca, N.Y: Cornell University Press, 1980). While Kojève has done much to draw attention to the importance of Hegel's theory of recognition, his account of the theory is limited. Kojève's account only gives emphasis to the moment of struggle and does not give due regard to moments of mutual recognition present within the *Realphilosophie*, the *Phenomenology of Spirit* and the *Philosophy of Right*. Another important Marxist account of Hegel's theory of recognition is the interpretation and extrapolation given by Slavoj Žižek. See: Žižek, S. *The Sublime Object of Ideology*, (London: Verso, 1989); and Žižek, S. *Tarrying With the Negative: Kant, Hegel and the Critique of Ideology*, (Durham: Duke University Press, 1993). While Žižek's use of recognition is a novel and powerful critique of modern society and culture, his interpretation of Hegel through Lacan has a tendency to reduce or structure recognition

recognition in which each self recognises itself as mutually recognising the other.<sup>41</sup> The second involves an unsuccessful moment of recognition in which one self is recognised and the other is recognising.<sup>42</sup> This unsuccessful moment involves the operation of *misrecognition*.<sup>43</sup>

In the successful moment, or moment of mutual recognition, each intelligent self is faced by the fact that their not-self is also an intelligent being. In this moment what was formerly simply something other and defined only negatively in relation to the self, is now seen to have a content, as intelligence, which is not completely alien to the self's comprehension of itself. The self's recognition of the other does involve a moment of identification with the other but does not involve the reduction of the other to the same. Rather, in the process of recognition the self comes to know and appreciate the position and status of the not-self within its own self. The self comes to recognise and appreciate *the other within itself*. For Hegel, in recognising the other within itself, the self affirms this otherness and therefore does not try to negate the other, but lets the other go free.<sup>44</sup> What is recognised is not simply the other, but the constitutive role of the other in the self, the necessity of the other for the self to be itself. Recognition involves an affirmation by the self that a necessary and essential element of itself resides in the other and in the relation of mediation with it.

For Hegel, the moment of successful, mutual recognition is bound up with the moment of unsuccessful recognition, or misrecognition. In

around the Lacanian 'void'. For a critique of Žižek's reading of Hegel see: Dews, P. *The Limits of Disenchantment: Essays on Contemporary European Philosophy*, (London: Verso, 1995). For an account of the French reception of Hegel see: Butler, J.P. *Subjects of Desire: Hegelian Reflections in Twentieth-Century France*, (New York: Columbia University Press, 1987).

<sup>41</sup> Hegel, G.W.F. *Phenomenology of Spirit*, p. 112.

<sup>42</sup> *Ibid*, 113.

<sup>43</sup> The term is taken from Gillian Rose, see: Rose, G. *Hegel Contra Sociology*, (London: Athlone, 1981); Rose, G. *Mourning Becomes the Law*, (Cambridge: Cambridge University Press, 1996). For more general accounts of the *Phenomenology of Spirit* see: Pippin, R.B. *Hegel's Idealism: The Satisfactions of Self-Consciousness*, (Cambridge: Cambridge University Press, 1989); Pinkard, T. *Hegel's Phenomenology: The Sociality of Reason*, (Cambridge: Cambridge University Press, 1994). See also: Williams, R.R. *Recognition: Fichte and Hegel on the Other*, (Albany: S.U.N.Y. Press, 1992); Honneth, A. *The Struggle for Recognition: The Moral Grammar of Social Conflicts*, Anderson, J. tr. (Cambridge: Polity Press, 1995).

<sup>44</sup> Hegel, G.W.F. *Phenomenology of Spirit*, p. 112.

the operation of misrecognition the self proceeds in relation to its now intelligent other in the same way it proceeded in relation to the world at large. That is, the self in its drive towards self-identity and self-certainty affirms itself by negating its other. In the moment of misrecognition the self does not recognise the other within itself or the relation of mutual co-constitution between self and other, it merely sees its immediate view of self as significant and everything else as insignificant.

Further, misrecognition can be interpreted as involving the operation of *error* within the moment of self-reflection. The *mis* of misrecognition refers to a moment of cognitive error, mistake, and slippage or a moment where self-consciousness cannot get past elements of its own immediacy to comprehend its mediated nature. The *mis* of misrecognition also involves the operation of finitude within human cognition in its efforts to comprehend the infinite. In a sense, the finite self can only comprehend its infinite mediated character, by grasping differing parts that make up a whole. In attempting to grasp itself as a unity or a whole the self can only grasp fragile, weak, limited moments of itself in its relations to its many others.

For Hegel, the recurrent moment of error does not condemn the self's attempt to know itself and the world to failure. Rather, the process of recognition describes the effort of a spontaneous, creative human intelligence attempting to come to terms with its own errors and limitations. This involves a certain restlessness of thinking — thinking's desire — which is always pushed by itself to reach beyond its own limitations.<sup>45</sup> In this respect, the self's recognition of the other does not just happen once, it is instead an act of thinking that needs to be repeated again and again. As a process every cognition needs to be continually re-recognised as the self will only come to some form of self-certainty through ongoing acts, efforts and errors. Further, recognition draws attention to the creative aspect of human thinking whereby it only begins to develop by coming to terms with its own errors. The capacity of thinking to come to terms with error, limitation, and contradiction more generally is given significant emphasis by Hegel.

<sup>45</sup> On the 'restlessness' contained within thinking as presented by Hegel, see: Nancy, J.-L. *The Speculative Remark: One of Hegel's Bon Mots*, Suprenant, C. tr. (Stanford: Stanford University Press, 2001); Nancy, J.-L. *Hegel: The Restlessness of the Negative*, Smith, J. and Miller, S. tr. (Minneapolis: University of Minnesota Press, 2002).

If operating purely within the realm of thinking misrecognition, error, mistake and the assertion of a naive sense of self-certainty against the other would not have violent consequences. Hegel, however, places the recognising self within a living, physical world of intelligent human subjects. In this context the moment of misrecognition of the other, the refusal to recognise the other, the act of self-positing at the expense of the other, and the negation of the other, all have physical, material consequences. Hegel quite clearly describes this as *violence*: the seeking of the death of the other and the staking of the self's own life so that two selves engage in a life and death struggle.<sup>46</sup> In this struggle each self learns not only that the other is independent, but, also, that as a living being the self may also be treated as an object that can be negated. In a sense, through being negated by the other's misrecognition, the self is faced with its own limits and its extreme finitude within the possibility of death. For Hegel, through relative levels of power and the fear of death, one self succumbs to the other, recognises the other as master (*Herr*) and assumes the subservient position of slave (*Knecht*).<sup>47</sup>

Through this account Hegel can be seen to be presenting quite a stark and original theory which draws attention to the relation between the act of thinking and the deeds that occur in the natural, physical world. For Hegel the act of self-positing and the act of negation have physical consequences which impact upon the self's attempt to gain knowledge of itself. In one sense, Hegel's account of recognition draws attention to a certain form of violence, or at least the unaccounted for possibility of violence, residing within the transcendental philosophy of Kant and Fichte. In abstracting from the world, this transcendental philosophy abstracts from violence. However, to the extent that Kant and Fichte speak of a self-consciousness which exists in the world, then their account participates in a form of violence which they remain unaware of or hide. Such an unawareness impacts upon their attempts to construct a practical, moral philosophy. In another sense, Hegel begins to point to one significant theoretical form of comprehending human, social violence based upon the operation of positing and negation within human thinking. Through Hegel the *deeds* of thinking become closely related to the physical deeds of human life: a certain violence of positing and negativity.

<sup>46</sup> Hegel, G.W.F. *Phenomenology of Spirit*, pp. 113–4.

<sup>47</sup> *Ibid.*, p. 115.

### III

So what does Hegel in these passages present to philosophy? Amongst other things Hegel presents a radical and original account of the metaphysics of subjectivity. For the first time in the history of Western philosophy Hegel draws together the relationships between thinking and action and between negativity and violence and shows their central role in the production of the human subject and its construction (and destruction) of its world.

Hegel radicalises the Fichtean idea of thoughts as deeds. What for Fichte was an act only at a transcendental level is shown by Hegel to be also an act within the physical, material and social world. Positing and negation become the activities that link the transcendental subject to its world. In this link, or relation of mediation, it is the subject's activity that is shown to be necessary and essential. The active, creative and destructive subject thinks both itself and its world, it is physically shaped and moulded by its material environment, and it is represented and constructed by the thoughts and acts of other thinking subjects. These acts of self-positing, inter-subjective positing and negation sit in a causal relationship with both the comprehension of the transcendental subject (as it is portrayed by differing historical modes of philosophical reflection) and the living human subject within politics, ethics and the social world. Hegel shows that the thinking subject, the 'I', whose primacy is given to modern philosophy by Descartes and Kant and which stands as the central figure of Enlightenment rationality and liberalism, is fundamentally connected to history, inter-subjectivity and the social-material world. The question of where the 'not-I' ends and the 'I' begins (and *visa versa*), framed by Hegel as *the question* of modern subjectivity, continues to chastise modern philosophy, even if many philosophers and theorists are unaware that they are being chastised by a modern question arranged (at least in part) by Hegel.

By drawing out conceptually how thoughts operate as deeds Hegel's account grasps the world creating function of modern subjectivity. Because the subject's positing of itself and its account of the world (i.e. of how the world *is* and *should* be in relation to the self) is also a physical

positing through individual and (later<sup>48</sup>) through institutional force, the subject's acts of positing and negation, as a form of labour, create and shape the social-material world. A modern world is created not simply by the subject's learning of its ability to shape and create objects and master the physical world, but also at an institutional level through the positing of laws, through the positing of economic, ethical and political modes of thought that arrange and order social life, and finally through positing as war.

Such a creative character of modern subjectivity is also profoundly destructive. Every act of positing contains a moment of negation and this moment of negation bears out differing degrees of immediate and institutional violence. The negativity and violence within the subject cannot be separated from the self-positing of the subject at any level. Every act of positing will involve a moment of differentiation and negation which, if it does not amount to violence, institutes a potentiality of violence — a *not-yet*<sup>49</sup> violence, that may be enacted or realised at some point in the future. Hegel's account draws out the links between transcendental (*a priori*) negativity, the desire for self-identity, negation, violence, mastery and domination. For Hegel this is not the whole of subjectivity but it is an important part which creeps into and upsets all knowledge, action and ethical effort. The religious aspect of Hegel's thought gives this negativity a certain role, purpose and moment of redemption.<sup>50</sup> For those of us who might not share Hegel's speculative account of religion and who look back on the violent train wreck of modern, Enlightenment enthusiasm that was the nineteenth and twentieth centuries (inclusive of romanticism, nationalism, colonialism and imperialism), then we may view the negativity and violence within subjectivity somewhat differently. The violence of our recent history offers us a lesson. That is, the status of contemporary attempts to outline both a philosophy of the subject and a philosophy

<sup>48</sup> For example, Hegel's account of the operation of institutional positing as negation during the French Terror. See: Hegel, G.W.F. *Phenomenology of Spirit*, p. 355–363.

<sup>49</sup> To use a term taken from Ernst Bloch. See: Bloch, E. *The Spirit of Utopia*, Nassar, A.A. tr. (Stanford: Stanford University Press, 2000).

<sup>50</sup> I.e. negativity, violence and death played a different role to Hegel as a Christian philosopher than it does to more secular thinkers today. For an account of the religious elements of Hegel's thinking see: Küng, H. *The Incarnation of God: An Introduction to Hegel's Theological Thought as a Prolegomena to a Future Christianity*, Stephenson, J.R. tr. (New York: Crossword, 1987).

of ethics will be inadequate if they do not pay enough attention to the negativity and violence imbedded within human subjectivity.

If we take seriously the Hegelian account of the operation of positing and negation within subjectivity, or at least, if we give it some degree of credence, then it would be naive to attempt to ground an ethics or politics upon the notion of a 'non-violent ethics' or a 'non-violent relationship to the other'.<sup>51</sup> Rather, a better approach would be to attempt to develop an idea of ethics and politics which understands the negativity and violence contained within human subjectivity, and knowing this, sets up practices and institutions that might predict and minimise resultant harm. One aspect of a future ethics and politics built along these lines could be grounded upon a *praxis* of recognition — one which paid attention to and planned for the *mis* of misrecognition. While a *praxis* of recognition cannot be the whole of ethics or politics it does offer a way of characterising ethical decision and political action in a manner that is consistent with the violence and negativity that resides at the heart of the modern subject.

<sup>51</sup> See: Cornell, D. *The Philosophy of the Limit*, (New York: Routledge, 1992).



# Begründung der Metaphysik durch logische Analyse der Selbstbezüglichkeit

UWE PETERSEN

ZUSAMMENFASSUNG. Denken, das auszog, sich seiner eigenen Grundlagen zu versichern, fand sich wieder in einem heillosen Widerstreit seiner Einsichten. So erklärte es sich selbst zu diesem Behufe ungeeignet, nicht ahnend, daß es eben in diesem Widerstreit auf sich selbst und seine ureigene konstitutive Kraft gestoßen war. Dieser Situation der Selbstbezüglichkeit soll hier durch eine nicht-klassische Logik Rechnung getragen werden, um für eine Form der spekulativen Erkenntnis Raum zu schaffen: ein Wissen, das seinen Grund nicht in der äußeren Welt hat, sondern in einer kognitiven Rückwendung auf sich selbst, auf seine eigene Begrifflichkeit. Dazu werden Techniken und Methoden aus der mathematischen Logik und Grundlagenforschung herangezogen, wo Selbstbezüglichkeit schon seit langem eine wichtige Rolle spielt.

Was der Geist will, ist,  
seinen eigenen Begriff zu erreichen,  
aber er selbst verdeckt sich denselben . . . <sup>1</sup>

It's the self-reference, stupid.<sup>2</sup>

## 1. Einleitung

Vor rund 75 Jahren hat Rudolf Carnap in einem Aufsatz mit dem Titel „Widerlegung der Metaphysik durch logische Analyse der Sprache“<sup>3</sup> zu einem Schlag ausgeholt, der die Metaphysik ein für allemal aus dem Kreis der Wissenschaften verbannen sollte. Die damals noch relativ junge Disziplin der symbolischen Logik war auserkoren, dafür die Mittel zu liefern:

<sup>1</sup> [28], S. 90 f. Es geht weiter: „ist stolz und voll von Genuß, in dieser Entfremdung seiner selbst.“

<sup>2</sup> Ich, dieser Band, S. 77; frei nach Bill Clinton.

<sup>3</sup> Kann man im 21. Jahrhundert noch davon ausgehen, daß Carnaps Aufsatz [6], zumindest dem Titel nach, bekannt ist? Wenn nicht, so hat wohl der Titel dieses Aufsatzes seinen Sinn verfehlt.

Durch die Entwicklung der *modernen Logik* ist es möglich geworden, auf die Frage nach der Gültigkeit und Berechtigung der Metaphysik eine neue und schärfere Antwort zu geben. Die Untersuchungen der „angewandten Logik“ oder „Erkenntnistheorie“, die sich die Aufgabe stellen, durch logische Analyse den Erkenntnisgehalt der wissenschaftlichen Sätze und damit die Bedeutung der in den Sätzen auftretenden Wörter („Begriffe“) klarzustellen, führen zu einem positiven und zu einem negativen Ergebnis. Das positive Ergebnis wird auf dem Gebiet der empirischen Wissenschaft erarbeitet; die einzelnen Begriffe der verschiedenen Wissenschaftszweige werden geklärt; ihr formal-logischer und erkenntnistheoretischer Zusammenhang wird aufgewiesen. Auf dem Gebiet der *Metaphysik* (einschließlich aller Wertphilosophie und Normwissenschaft) führt die logische Analyse zu dem Ergebnis, daß *die vorgeblichen Sätze dieses Gebietes gänzlich sinnlos sind*. Damit ist eine radikale Überwindung der Metaphysik erreicht, die von den früheren antimetaphysischen Standpunkten aus noch nicht möglich war.<sup>4</sup>

„Sinnlos“ wird dabei von Carnap folgendermaßen näher bestimmt:

Im strengen Sinn *sinnlos* ist [. . .] eine Wortreihe, die innerhalb einer bestimmten, vorgegebenen Sprache gar keinen Satz bildet.<sup>5</sup>

Damit läßt sich nun der Titel von Carnaps Aufsatz wie folgt modifizieren: „Überwindung der Metaphysik durch logische Analyse ihrer Sätze (Wortreihen) innerhalb einer bestimmten vorgegebenen Sprache“.

Aber dann ergibt sich sofort die Frage: Was ist mit anderen, nicht vorgegebenen Sprachen? Und weiter: Wer gibt überhaupt die Sprache vor, und aufgrund welcher Autorität? Reicht es aus, eine bestimmte Sprache zu finden, in der metaphysische Sätze ungrammatikalisch sind? Ist es nicht nötig, den Reichtum der möglichen Sprachen in Betracht zu ziehen?

Das wird uns noch etwas weiter unten beschäftigen. Vorher aber will ich auf die Sprache eingehen, die Carnap bei seinem Kriterium der Sinnlosigkeit im Auge hatte.

Carnap greift auf eine machtvolle Entwicklung der Logik seit der Mitte des 19. Jahrhunderts zurück, die außer ihrer intendierten Anwendung

<sup>4</sup> [6], S. 219 f.

<sup>5</sup> Ebd., S. 220.

auch ein bemerkenswertes, und für die herkömmliche Auffassung keineswegs willkommenes, Resultat hatte. Nachdem Frege die Logik zu einem funktionstüchtigen Werkzeug ausgebaut und für eine logische Grundlegung der Arithmetik eingesetzt hatte, entdeckte Russell, daß sich damit nicht nur arithmetische Wahrheiten beweisen lassen, sondern auch Widersprüche. Sein Rezept zur Vermeidung solcher beweisbaren Widersprüche war eine Unterscheidung von Typen,<sup>6</sup> die dann als „einfache Typentheorie“ in die Geschichte eingegangen ist. In der Sprache dieser Typentheorie hat man Zeichen für Variablen verschiedener Typen zur Verfügung, und das Zeichen für die Elementschäftsrelation kann nur zwischen Zeichen für Variablen bestimmter Typen stehen, wobei insbesondere ausgeschlossen ist, daß Zeichen für Variablen desselben Typs auf beiden Seiten der Elementschäftsrelation zu stehen kommen. Mit anderen Worten, die Sprache ist so aufgebaut, daß unmittelbar an der Syntax erkennbar ist, ob eine Zeichenreihe ein Satz dieser Sprache ist oder nicht.

Eine solche Typenunterscheidung dient Carnap als Vorbild für seine Auffassung von sinnlosen Sätzen. In vager Analogie dazu führt er erst „Sphärenunterscheidungen“ ein, was ihm dann erlaubt, unliebsame Sätze als „Sphärenvermengungen“ zu brandmarken:

Die Nichtbeachtung des Unterschiedes sphärenfremder Begriffe bezeichnen wir als „Sphärenvermengung“.<sup>7</sup>

Sphärenvermengungen sind nach Carnap die Ursache sinnloser Sätze. Unter der Überschrift „Literatur“ führt Carnap dann weiter aus:

Eine ausdrückliche Beachtung hat die genannte Art der Mehrdeutigkeit bisher in der Logik nicht gefunden. Sie hat aber eine gewisse Verwandtschaft mit der Vielheit der „Suppositionen“ eines Wortes, die die Scholastiker zu unterscheiden pflegten [ . . . ]. Enger hängt sie jedoch mit der Typentheorie zusammen, die Russell zur Überwindung der logischen Paradoxien aufgestellt und in seinem System der Logistik zur Anwendung gebracht hat [ . . . ]. Russell hat jedoch diese Theorie nur auf formal-logische Gebilde angewendet, nicht auf ein System der konkreten Begriffe (d. h. genauer: nur auf Variable und logische Konstanten, nicht auf nichtlogische Konstanten). Unsere „Gegenstandssphären“ sind

<sup>6</sup> Vgl. [46], Appendix B. Unten auf S. 107 findet sich ein kurzes Zitat zur Charakterisierung der Typentheorie in Russells eigenen Worten.

<sup>7</sup> [5], S. 40.

die Russellschen „Typen“, angewendet auf nichtlogische Begriffe. Somit liegt auch die Rechtfertigung unserer Unterscheidung der Gegenstandssphären [...] in der Typentheorie[.]<sup>8</sup>

Das erlaubt nun eine weitere Modifizierung des Titels von Carnaps Aufsatz: „Überwindung der Metaphysik durch logische Analyse ihrer Sätze (Wortreihen) innerhalb einer bestimmten vorgegebenen Sprache welche eine Sphärenvermischung ausschließt“.

Was hat man sich nun unter einer „Sphärenvermischung“ konkret vorzustellen? Anders gesagt, was sind die Russellschen „Typen“, „angewendet auf nichtlogische Begriffe“? Carnap gibt folgendes Beispiel:

[[D]ie Wortreihe [„Caesar ist eine Primzahl“ ist] syntaxgemäß gebildet, [...] ist aber trotzdem sinnlos. „Primzahl“ ist eine Eigenschaft von Zahlen; sie kann einer Person weder zu- noch abgesprochen werden. Da [[diese Wortreihe] aussieht wie ein Satz, aber kein Satz ist, nichts besagt, weder einen bestehenden noch einen nicht bestehenden Sachverhalt zum Ausdruck bringt, so nennen wir diese Wortreihe einen „Scheinsatz“.<sup>9</sup>

Aus Carnaps Sicht

wird hier ein Prädikat zwar als Prädikat verwendet, aber als Prädikat einer anderen „Sphäre“: es liegt eine Verletzung der Regeln der sog. „Typentheorie“ vor. [...] Personennamen und Zahlwörter gehören zu verschiedenen logischen Sphären, und daher auch Personenprädikate (z. B. „Feldherr“) und Zahlenprädikate („Primzahl“).<sup>10</sup>

Und wie erkennt man, daß „Caesar“ und „Primzahl“ zu verschiedenen logischen Sphären gehören? An der Syntax erkennt man es jedenfalls nicht, und daß ein Satz wie „Caesar ist keine Primzahl“ zu Widersprüchen führen würde, zeigt Carnap auch nicht.<sup>11</sup> Da muß man eben das richtige Sprach-

<sup>8</sup> Ebd.

<sup>9</sup> [6], S. 227 f.

<sup>10</sup> Ebd., S. 235.

<sup>11</sup> In seinen *Grundlagen der Arithmetik* (§56) formuliert es Frege als Mangel einer Festlegung nach dem Muster, „einem Begriffe kommt die Zahl 0 zu, wenn kein Gegenstand unter ihn fällt“, daß wir durch sie nie entscheiden können, „ob einem Begriffe die Zahl Julius Caesar zukomme, ob dieser bekannte Eroberer Galliens eine Zahl ist oder nicht.“ Deshalb hat Frege Begriffsumfänge eingeführt: Sie erlauben es ihm, Zahlen *dingfest* zu machen. Von „Sinnlosigkeit“ ist bei Frege jedoch keine Rede.

verständnis haben. Carnap hat das natürlich, bzw. bildet sich ein es zu haben, so wie es sich für einen Philosophen gehört. Es ist das inhaltliche Wissen, das die Philosophie prägt und sie so überlegen macht. Aber mit der Russellschen Typentheorie hat das, über eine vage Analogie hinaus, nichts mehr zu tun. Deshalb werde ich Carnaps Idee einer Sphärenunterscheidung keine weitere Beachtung schenken und durch das Konzept der Typenunterscheidung ersetzen, das Carnap ohnehin als Rechtfertigung für seine Unterscheidung der Gegenstandssphären anführt. Wenn ich also im folgenden Carnap zitiere, so sollte „Scheinsatz“ im Sinne von „Wortreihe mit Typenvermengungen“ gelesen werden, um die stärkste mögliche Lesart zu erhalten.

Wenn unsere These, daß die Sätze der Metaphysik Scheinsätze sind, zu recht besteht, so würde also in einer logisch korrekt aufgebauten Sprache die Metaphysik gar nicht ausgedrückt werden können. Daraus ergibt sich die große philosophische Bedeutsamkeit der Aufgabe des Aufbaus einer logischen Syntax[.]<sup>11</sup>

Worauf es mir hier vor allem ankommt, ist eine gewisse Unklarheit in der Formulierung: „[[I]n einer logisch korrekt aufgebauten Sprache [[kann] die Metaphysik gar nicht ausgedrückt werden“. Das kann auf zwei Arten gelesen werden:

- Es gibt eine logisch korrekt aufgebaute Sprache, in der die Metaphysik nicht ausgedrückt werden kann.
- Es gibt keine logisch korrekt aufgebaute Sprache, in der die Metaphysik ausgedrückt werden kann.

Es kommt darauf an, mit welcher Betonung man den Satz liest. Was Carnap tatsächlich — ansatzweise zumindest — in Angriff nimmt, ist die erste Variante, der Aufbau einer logischen Syntax. Es scheint mir aber klar, daß Carnaps Anspruch auf die zweite Variante abzielt, die auch einfach folgendermaßen formuliert werden kann:

- Wenn eine Sprache logisch korrekt aufgebaut ist, dann kann die Metaphysik nicht in ihr ausgedrückt werden.

Dazu muß aber präzisiert werden, was überhaupt als „logisch korrekt aufgebaute Sprache“ gelten kann. Die Aufgabe einer solchen Präzisierung formuliert Carnap jedoch nicht. Das scheint für Carnap nicht zum Anliegen einer *logischen Analyse* zu gehören. Hier eröffnet sich aber ein grundsätzliches Problem für jeden Versuch einer *Überwindung*, ein Problem, das sich

für den Versuch einer *Begründung* nicht stellt. Eine Überwindung der Metaphysik nach dem Carnapschen Muster erfordert ein Quantifizieren über alle logisch korrekt aufgebauten Sprachen.<sup>12</sup> Für eine *Begründung* würde es ausreichen, eine nach allgemeinen Maßstäben *logisch korrekt aufgebaute Sprache* zu finden, bzw. aufzubauen, in der sich Metaphysik betreiben läßt, um Carnaps Schlag abgewehrt zu haben. Das ist es, worum es hier geht.

Insoweit Carnap glaubt, in metaphysischen Aussagen eine Sphärenvermischung entdecken zu können, mag er ja durchaus etwas Richtiges getroffen haben. Aber eine Sphärenvermischung auf der Grundlage der Typentheorie als ungrammatikalisch zu verwerfen, ist solange zu billig, als die Typentheorie nicht als einzig sinnvolle Logik nachgewiesen worden ist. Das war wohl auch Carnap klar, und so finden wir folgende Bemerkung:

Die Typentheorie ist zwar nicht allgemein anerkannt, aber es ist bisher keinem ihrer Gegner gelungen, ein logisches System aufzuweisen, das ohne Verwendung einer Typentheorie imstande wäre, die Widersprüche (die sog. „Paradoxien“), an denen die alte Logik krankt, zu vermeiden.<sup>13</sup>

Es mag ja sein, daß 1928 niemand in der Lage war, der Typentheorie ein logisches System entgegenzusetzen, das ohne Typenunterscheidung imstande war, die Paradoxien zu vermeiden. Aber schon zwei Jahre später hat Łukasiewicz seinen unendlichwertigen Aussagenkalkül vorgestellt, in dem die logischen Paradoxien zwar formuliert werden können, aber nicht mehr zu beweisbaren kontradiktorischen Widersprüchen führen.<sup>14</sup> Sollte Carnap wirklich geglaubt haben, die (einfache) Typentheorie stelle das letzte

<sup>12</sup> Das läßt sich durchaus mit dem Problem der Entscheidbarkeit (von Theorien) vergleichen: der Begriff des Entscheidungsverfahrens (Algorithmus) muß präzise gefaßt sein, um beweisen zu können, daß eine Theorie unentscheidbar ist. Das erfordert ein Quantifizieren über alle Entscheidungsverfahren. Im Falle der Entscheidbarkeit von Theorien wird das bekanntlich über den Begriff der allgemein-rekursiven Funktion (oder einen äquivalenten Begriff) und die Church-Turing-These erreicht. Das bedeutet jedoch nicht, daß man schon die Church-Turing-These braucht, um spezielle Algorithmen durch rekursive Funktionen zu erfassen.

<sup>13</sup> [5], S. 40.

<sup>14</sup> Vgl. [33]. Weitere Systeme dieser Art sind im Laufe der vierziger und fünfziger Jahre des letzten Jahrhunderts von Fitch, Ackermann und Schütte entwickelt worden. Literaturhinweise finden sich unter 78.15 in [37], S. 1099.

Wort in Sachen Paradoxienvermeidung dar? Wie auch immer, Carnap hat seinen Versuch einer Überwindung der Metaphysik von der Notwendigkeit einer Typenunterscheidung à la Russell zur Vermeidung der logischen Paradoxien abhängig gemacht und damit implizit eine Richtung gewiesen, in der Metaphysik sinnvoll betrieben werden kann: als Theorie jener Begriffsbildungen, die unter klassischen Bedingungen zu beweisbaren Widersprüchen führen. Die Widersprüche, an denen, in Carnaps Worten, „die alte Logik krankt“, stellen ein tief wurzelndes Problem begrifflichen Denkens dar, das nicht nur der Fragestellung der Metaphysik zugrundeliegt, sondern dem Carnap mit seinem Versuch einer Überwindung der Metaphysik selbst nicht entrinnen konnte.

Eine Form dieses Problems zeigt sich im sogenannten *Ad-hominem*-Argument, das die Praxis des Sprechers an dessen eigenen Ansprüchen mißt.<sup>15</sup> Carnap scheut sich nicht, diese Frage zu stellen:

But now it may perhaps be objected: “How about your own propositions? In consequence of your view your own writings, including this book, would be without sense, for they are neither mathematical nor empirical, that is verifiable by experience.” What answer can be given to this objection? This question is decisive for the consistency of the view which has been explained here.<sup>16</sup>

Das ist tatsächlich eine Frage, von deren Beantwortung die Glaubwürdigkeit von Carnaps Ansatz zur Überwindung der Metaphysik abhängt. Zunächst einmal beantwortet Carnap diese Frage nicht selbst, sondern schiebt Wittgenstein vor:

An answer to the objection is given by Wittgenstein in his book *Tractatus Logico-Philosophicus*. This author has developed most radically the view that the propositions of metaphysics are shown by logical analysis to be without sense. How does he reply to the criticism that in that case his own propositions are also without sense? He replies by agreeing with it. He writes: “*The result of philosophy is not a number of ‘philosophical propositions,’ but to make propositions clear*” (S. 77). “My propositions are elucidatory in this way: he who understands me finally recog-

<sup>15</sup> Vgl. [12], S. 68.

<sup>16</sup> [7], S. 36.

nizes them as senseless, when he has climbed out through them, on them, over them. (He must so to speak throw away the ladder after he has climbed up on it.) He must surmount these propositions; then he sees the world rightly. Whereof one cannot speak, thereof one must be silent." (S. 189).<sup>17</sup>

Worauf es mir hier ankommt, ist der Dreh mit dem „Resultat“: Die Philosophie hat ein Resultat, und zu dem kommt man nur, wenn man vergißt, wie man dahin gekommen ist. Mit anderen Worten, was Wittgenstein als Philosophie propagiert, ist eine Form der Einsicht, die mit ihrem eigenen Werdegang auf Kriegsfuß steht. Was macht nun Carnap daraus?

I, as well as my friends in the Vienna Circle, owe much to Wittgenstein, especially as to the analysis of metaphysics. But on the point just mentioned I cannot agree with him. In the first place he seems to me to be inconsistent in what he does. He tells us that one cannot state philosophical propositions and that whereof one cannot speak, thereof one must be silent; and then instead of keeping silent, he writes a whole philosophical book. Secondly, I do not agree with his assertion that all his propositions are quite as much without sense as metaphysical propositions are. My opinion is that a great number of his propositions (unfortunately not all of them) have in fact sense; and that the same is true for all propositions of logical analysis.<sup>18</sup>

Haben wir eine Antwort erhalten, oder sind wir nur an der Nase herumgeführt worden? Carnap scheint sehr wohl zu sehen, daß Wittgenstein sich in einer Art von „performative inconsistency“ verfangen hat, doch erfahren wir lediglich was er davon hält — ohne irgendeine logische Analyse. Dafür verspricht er

<sup>17</sup> Ebd., S. 37. Carnaps Seitenangaben beziehen sich auf die erste englische Übersetzung von Wittgensteins Tractatus von 1922. Das erste Wittgenstein-Zitat lautet im (deutschen) Original: „Das Resultat der Philosophie sind nicht ‚philosophische Sätze‘, sondern das Klarwerden von Sätzen.“ ([63], S. 48); das zweite: „Meine Sätze erläutern dadurch, daß sie der, welcher mich versteht, am Ende als unsinnig erkennt, wenn er durch sie — auf ihnen — über sie hinausgestiegen ist. (Er muß sozusagen die Leiter wegwerfen, nachdem er auf ihr hinaufgestiegen ist.) Er muß diese Sätze überwinden, dann sieht er die Welt richtig. Wovon man nicht sprechen kann, darüber muß man schweigen.“ ([63], S. 150).

<sup>18</sup> [7], S. 37 f.



to show a way of formulating the results of logical analysis, a way not exposed to the objection mentioned, and thus to exhibit an *exact method of philosophy*.

Im Rückblick wissen wir, daß Carnap dieses Versprechen nicht hat einlösen können. Aber das ist eine andere Geschichte. Was mich hier interessiert, ist eine gewisse Unverträglichkeit von Reden und Tun, Rhetorik und Praxis, wenn man so will. Deshalb füge ich hier noch ein Zitat aus Russells Vorwort zur englischen Übersetzung von Wittgensteins *Tractatus* an, in dem Russell ebenfalls auf Wittgensteins Inkonsequenz zu sprechen kommt:

Mr Wittgenstein's attitude towards the mystical [...] grows naturally out of his doctrine in pure logic, according to which the logical proposition is a picture (true or false) of the fact, and has in common with the fact a certain structure. It is this common structure which makes it capable of being a picture of the fact, but the structure cannot itself be put into words, since it is a structure *of* words, as well as of the facts to which they refer. Everything, therefore, which is involved in the very idea of the expressiveness of language must remain incapable of being expressed in language, and is, therefore, inexpressible in a perfectly precise sense. This inexpressible contains, according to Mr Wittgenstein, the whole of logic and philosophy. The right method of teaching philosophy, he says, would be to confine oneself to propositions of the sciences, stated with all possible clearness and exactness, leaving philosophical assertions to the learner, and proving to him, whenever he made them, that they are meaningless. It is true that the fate of Socrates might befall a man who attempted this method of teaching, but we are not to be deterred by that fear, if it is the only right method. It is not this that causes some hesitation in accepting Mr Wittgenstein's position, in spite of the very powerful arguments which he brings to its support. What causes hesitation is the fact that, after all, Mr Wittgenstein manages to say a good deal about what cannot be said, thus suggesting to the sceptical reader that possibly there may be some loophole through a hierarchy of languages, or by some other exit.<sup>19</sup>

<sup>19</sup> [48], p. xx f.

In alldem zeigt sich ein Problem — ein Problem mit der Sprache, angesichts eines Inhaltes, der sich nur um den Preis der Inkonsistenz ausdrücken läßt.

Für den Fall der Russellschen Typentheorie hat Fitch das folgendermaßen formuliert:

a ramified theory of types could not even be stated. Its sweeping restrictions against self-reference would apply to every theory, including itself, and so would be self-referential in violation of its own edicts. A similar criticism can be made even against the more moderate simplified theory of types, if regarded as universally applicable.<sup>20</sup>

Anders gesagt, beim Versuch, ein Kriterium der Sinnlosigkeit von metaphysischen Sätzen aufzustellen, ist Carnap ein Opfer jener Grammatikfehler geworden, die er ausschließen wollte.

Mit einem einfachen Slogan könnte man das folgendermaßen fassen:

*Die Formulierung eines Kriteriums zur Überwindung der Metaphysik ist ein Teil derselben.*

Das wird Metaphysikgegner jedoch nicht davon abbringen, weiterhin die Unmöglichkeit jeglicher Metaphysik zu behaupten, denn

in Wirklichkeit liegt die Sache so, daß es keine sinnvollen metaphysischen Sätze geben kann. Das folgt aus der Aufgabe, die die Metaphysik sich stellt: sie will eine Erkenntnis finden und darstellen, die der empirischen Wissenschaft nicht zugänglich ist.<sup>21</sup>

Wenn es natürlich „in Wirklichkeit“ so ist, dann kann dagegen auch ein logisches Argument nichts ausrichten.<sup>22</sup> Deshalb beeile ich mich zu erklären, daß es mir hier auch nicht darum geht, Metaphysikgegner von der Unhaltbarkeit ihrer Argumente zu *überzeugen*, sondern diese Unhaltbarkeit als einen Einstieg in die Metaphysik zu nutzen. Mein Versuch einer Begründung setzt da an, wo Wittgenstein die Leiter wegwirft, bzw. sich ins Mystische flüchtet, als wüßte er, daß es etwas zu vertuschen gibt. Das, was für Wittgenstein nicht sagbar ist, ist das Aufschlußreiche, gerade weil es sich der Sprache des Schulmeisters entzieht.

<sup>20</sup> [12], p. 72.

<sup>21</sup> [6], S. 236.

<sup>22</sup> Und wenn es sich tatsächlich um Metaphysik handelt, so um *gute Metaphysik*, eben deshalb, weil sie einer *schlechten Metaphysik* entgegengesetzt ist.

Hier will ich ansetzen. Ich beginne damit, daß ich meine Alternative den zwei Lesarten von Carnaps Behauptung, die ich auf S. 81 aufgeführt habe, gegenüberstelle. Das sieht dann folgendermaßen aus:

1. Es gibt *keine* logisch korrekt aufgebaute Sprache, in der irgendein metaphysischer Satze sinnvoll wäre.
2. Es gibt logisch korrekt aufgebaute Sprachen, in der alle metaphysischen Sätze sinnlos sind.
3. Es gibt logisch korrekt aufgebaute Sprachen, in der (gewisse) metaphysische Sätze sinnvoll sind.

Auch bei einer extrem wohlwollenden Haltung gegenüber Carnaps Unterfangen kann man wohl nicht ernsthaft behaupten, daß Carnap die Richtigkeit der ersten Alternative gezeigt hat. Die zweite Alternative erscheint wenig beeindruckend. Man kann immer eine Sprache vorgeben, die so rudimentär ist, daß sich kaum etwas darin ausdrücken läßt. Die dritte Behauptung ist die, deren Richtigkeit ich hier zeigen will. Mit anderen Worten, es gibt eine (nach den Maßstäben der mathematischen Logik korrekt aufgebaute) Sprache  $\Sigma$  (die man vorgeben kann), in der metaphysische Begriffsbildungen von der Art, die Carnap als sinnlos bezeichnet, widerspruchsfrei durchgeführt werden können. Inwieweit dabei eine *sinnvolle Wissenschaft* herauskommt, soll hier der Erfahrung überlassen werden, nicht — wie für die Empiristen — einer Spekulation über die Natur der menschlichen Erkenntnis.

Die Entwicklung der modernen Logik hat nicht nur die Voraussetzung für Carnaps Behauptung geschaffen, daß metaphysische Sätze (aufgrund inhärenter Sphärenvermengungen) in gewissen Sprachen sinnlos sind, sondern auch die Mittel, alternative Systeme der Logik zu entwerfen. Diese Möglichkeit hat Carnap nicht berücksichtigt, und das ist es, was seinem Versuch einer Überwindung der Metaphysik den penetranten Beigeschmack des Wunschenkens verleiht.

Die symbolische Logik zu Anfang des 20. Jahrhunderts, mit dem Versprechen, alle Mathematik als analytisch nachzuweisen, und einer durch die Paradoxien der (höheren) Logik und Mengenlehre erzwungenen Typenunterscheidung, war für Carnap ein verheißungsvoller Ausgangspunkt, seinen puristischen Ideen eine Grundlage zu verschaffen. Aber die Zutaten paßten nicht zusammen: die Typenunterscheidung machte den Traum von der rein logischen Begründung der Mathematik zunichte. Das logizistische Programm in der Gestalt der *Principia Mathematica* erfüllte die

Hoffnungen des logischen Empirismus nicht. Spätestens mit der Einführung eines Unendlichkeitsaxioms ist der Anspruch auf eine rein logische Begründung verwirkt.

Es war aber nicht einfach nur ein Versuch der Reduktion von Mathematik auf Logik gescheitert; mit Gödels Unentscheidbarkeitsergebnis war klar, daß eine vollständige Rückführung der Arithmetik auf Logik grundsätzlich nicht gelingen kann. Der logische Empirismus hatte zu seinem Ausgangspunkt einen Ansatz gewählt, der zu diesem Zeitpunkt schon gescheitert war. Aber Carnap und Wittgenstein haben sich für eine Vogel-Strauß-Politik entschieden und schlichtweg nicht zur Kenntnis genommen, was ihrer Grundüberzeugung zuwiderlief. Wittgenstein hat seine Haltung auf einen einfachen Nenner gebracht: „Meine Aufgabe ist es nicht, über den Gödelschen Beweis, z. B., zu reden; sondern an ihm vorbeizureden.“<sup>23</sup>

Die symbolische Logik als Verbündete des Empirismus war ein Traum, der an der Komplexität formaler Systeme zerbrochen ist. In der Philosophie, nicht nur der empiristischen, gibt es eine gewisse Vorliebe für simplifizierte Modelle, und der logische Empirismus ist ein Opfer dieser Vorliebe geworden. Jahrzehnte später schreibt Stegmüller unter der Überschrift „Falsche Orientierung am großen Bruder?“:

Die Entwicklung der Philosophie der Mathematik zu einer exakten Wissenschaft, genannt *Metamathematik*, ist durch die mathematische Grundlagenkrise hervorgerufen worden. Da diese Krise durch die Entdeckung der mengentheoretischen Antinomien ausgelöst wurde, wird sie oft so dargestellt, als habe es sich dabei um ein *tragisches Ereignis in der modernen Mathematik* gehandelt.

Betrachtet man diesen Vorgang unter dem Aspekt der Wirkung, so gelangt man eher zu der gegenteiligen Beurteilung: Die Entdeckung von Antinomien war ein höchst *glückliches Ereignis*; denn sie bewirkte den Zwang zur Formalisierung und Präzisierung des Erkenntnisgegenstandes der Philosophie der Mathematik. Intuitive Vorstellungen vom mathematischen Denken wurden durch genau beschreibbare Objekte ersetzt und die Philosophie der Mathematik entwickelte sich zur mathematischen Grundlagenforschung, die in allen ihren Verzweigungen zu Disziplinen führte, die der Mathematik an Präzision nicht nachstanden und die heute selbst als Teile der Mathematik angesehen werden.

<sup>23</sup> [64], S. 383.

In den empirischen Wissenschaften ist es *leider* zu keiner Grundlagenkrise dieser Art gekommen. Und daher ergab sich auch kein Zwang, die Begriffs- und Theorienbildungen in diesen Disziplinen in Objekte zu verwandeln, die sich für exakte metatheoretische Studien eignen. Man tat nur so, *als ob* solche Objekte vorlägen. [. . .]

Somit blieb [. . .] nur eine *Philosophie des Als-Ob*, die sich am großen Bruder *Metamathematik* orientierte: Was sich dort als so außerordentlich fruchtbar erwiesen habe, das *müsse* sich auch hier als fruchtbar erweisen.<sup>24</sup>

In den empirischen Wissenschaften mag es zu keiner Grundlagenkrise dieser Art gekommen sein, aber für die Metaphysik sieht das ganz anders aus. Kant hat, in einer großangelegten Studie mit dem Titel „Kritik der reinen Vernunft“, versucht zu zeigen, daß es in der Natur der metaphysischen Begriffsbildung liegt, zu antinomischen Ergebnissen zu führen. Seine eigene Reaktion darauf war klassisch in dem Sinne, daß er diese Begriffsbildung beschränken wollte. Hegel hat demgegenüber auf die konstitutive Rolle dieser Begriffsbildung verwiesen und versucht, die Widersprüche in eine Ableitung von Denkbestimmungen zu integrieren.<sup>25</sup> In dieser Situation, in der es tatsächlich auf exakte metatheoretische Untersuchungen der Begriffs- und Theorienbildung ankommt, sollten die Methoden der Metamathematik von außerordentlichem Nutzen sein.

Im folgenden werde ich, im Anschluß an Hegels Idee, ein nicht-klassisches System der (höheren) Logik vorstellen, das es erlaubt, auf unverfängliche Weise die wildesten Typenvermengungen vorzunehmen, um dann mit den daraus resultierenden, aber nun durch eine nicht-klassische Logik entschärften, Widersprüchen viele schöne Dinge zu tun, die aus

<sup>24</sup> [56], S. 1 f.

<sup>25</sup> Man vergleiche [29], S. III:

Hegel [hat] der Erscheinung des Widerspruchs einen systematischen Ort im Denken zugewiesen und hat das Schema des Widerspruchs im Geschehen der Wirklichkeit nachzuweisen versucht. Aber er und auch die folgenden Denker haben sich nicht bemüht, das Phänomen deskriptiv zu erfassen. Unterdessen sind die Möglichkeiten einer deskriptiven Bearbeitung der Widersprüche gewachsen. Um nur eine zu nennen: in einem Zweig der Mathematik, der Mengentheorie, sind Widersprüche aufgetreten.

Dieses Zitat stammt aus derselben Zeit wie Carnaps Versuch einer Überwindung der Metaphysik.

Carnaps Sicht zu sinnlosen Sätzen führen müßten.<sup>26</sup> Tatsächlich wird der klassische Standpunkt, mit seinen tradierten Vorstellungen von Sinn, Bedeutung und Wahrheit, einen solchen Eingriff nicht unbeschadet überstehen; aber das ist es gerade, was für die Begründung der Metaphysik eine neue Möglichkeit schafft — jenseits aller Haken und Ösen eines semantischen Korsetts.

Zuerst aber will ich versuchen zu klären, worum es mir in der Metaphysik eigentlich geht.

## 2. Was heißt hier „Metaphysik“?

Die frühesten uns überlieferten Zeugnisse eines Denkens, das ich als metaphysisch begreife und für das ich hier einen Begründungsvorschlag ausarbeiten will, stammen aus dem fünften Jahrhundert vor unserer Zeitrechnung und werden Parmenides zugeschrieben:<sup>27</sup>

οὔτε γὰρ ἄν γνῶις τό γε μὴ ἔδῃ (οὐ γὰρ ἀνυστόν)  
οὔτε φράσαις.<sup>28</sup>

... τὸ γὰρ αὐτὸ νοεῖν ἐστὶν τε καὶ εἶναι.<sup>29</sup>

χρῆ τὸ λέγειν τε νοεῖν τ' ἐὼν ἔμμεναι· ἔστι γὰρ εἶναι,  
μηδὲν δ' οὐκ ἐστὶ.<sup>30</sup>

ταύτων δ' ἐστὶ νοεῖν τε καὶ οὔνεκεν ἔστι νόημα.  
οὐ γὰρ ἄνευ τοῦ ἐόντος, ἐν ᾧ πεφρατισμένον ἐστὶν,

<sup>26</sup> Die Definition eines Begriff der Sinnlosigkeit wird jedoch nicht dazu gehören, da es sich dabei, aus meiner Sicht jedenfalls, um einen wirklich sinnlosen Begriff handelt.

<sup>27</sup> Natürlich gibt es Gelehrte, die gegen diese Einordnung des Parmenides in den Kontext einer logisch-metaphysischen Tradition Einspruch erheben, doch, frei nach Quine ([45], S. 345), das ist nun einmal die Strafe, die darauf steht, einem berühmten Philosophen irgend etwas zuzuschreiben.

<sup>28</sup> [10], S. 231, Parmenides Frag. 2, 7; übersetzt als „denn weder erkennen könntest du das Nichtseiende (das ist ja unausführbar) noch aussprechen“. Vgl. [8], S. 53, fr. 3.

<sup>29</sup> Ebd., Parmenides Frag. 3; übersetzt als „denn dasselbe ist Denken und Sein“. Vgl. auch [8], S. 55, fr. 4.

<sup>30</sup> Ebd., S. 232, Parmenides Frag. 6; übersetzt als „Nötig ist zu sagen und zu denken, daß *nur* das Seiende ist; denn Sein ist, ein Nichts dagegen ist nicht“.

εὐρήσεις τὸ νοεῖν· οὐδὲν γὰρ ⟨ἦ⟩ ἔστιν ἦ ἔσται  
 ἄλλο πάρεξ τοῦ ἐόντος [.]<sup>31</sup>

Soweit Parmenides. Aristoteles schreibt den folgenden Ausspruch Herakleitos zu:

πάντα εἶναι καὶ μὴ εἶναι,<sup>32</sup>

und kommentiert, daß diese Lehre „alles wahr zu machen“ scheint.

Leukippos und Demokritos wird, ebenfalls von Aristoteles, folgender Ausspruch zugeschrieben:

οὐδὲν μᾶλλον τὸ ὄν τοῦ μὴ ὄντος εἶναι [.]<sup>33</sup>

Worauf es mir bei diesen überlieferten Bruchstücken ankommt, ist die zentrale Stellung von Sein und Nichts (oder Nichtsein). Das sind hier und später die immer wiederkehrenden Themen, und das vor dem Hintergrund ihrer Beziehung zum Denken: Wie kann man das, was nicht ist, widerspruchsfrei denken oder sagen?

In seinen Vorlesungen zur Geschichte der Philosophie sagt Hegel über Parmenides:

Mit Parmenides hat das eigentliche Philosophieren angefangen, die Erhebung in das Reich des Ideellen ist hierin zu sehen. Ein Mensch macht sich frei von allen Vorstellungen und Meinungen, spricht ihnen alle Wahrheit ab und sagt, nur die Nothwendigkeit, das Seyn ist das Wahre. Dieser Anfang ist freilich noch trübe und unbestimmt; es ist nicht weiter zu erklären, was darin liegt; aber gerade dieß Erklären ist die Ausbildung der Philosophie selbst, die hier noch nicht vorhanden ist. Damit verband sich die Dialektik, daß das Veränderliche keine Wahrheit habe; denn wenn man diese Bestimmungen nimmt, wie sie gelten: so kommt man auf Widersprüche.<sup>34</sup>

<sup>31</sup> Ebd., S. 238, Parmenides Frag. 8, 34 ff; übersetzt als „Dasselbe ist Denken und der Gedanke, daß IST *ist*; (35) denn nicht ohne das Seiende, in dem es als Ausgesprochenes ist, kannst du das Denken antreffen. Es ist ja nichts und wird nichts anderes sein außerhalb des Seienden“.

<sup>32</sup> [21], S. 202; Aristoteles, Metaphysik IV, 7, 1012 a 25. „Alles ist und ist nicht“.

<sup>33</sup> Ebd., S. 30; Aristoteles, Metaphysik I, 4, 985 b 9. In Hegel [27], S. 348 (Fußnote \*\*), unter Berufung auf Aristoteles irrtümlich Herakleitos zugeschrieben, übersetzt als: „Das Seyn ist nicht mehr als das Nichtseyn“.

<sup>34</sup> [27], S. 312 f.

Das ist im großen und ganzen eine Sicht, die ich mir zu eigen gemacht habe und die meiner Auffassung der Metaphysik zugrunde liegt.

Zu sehen, daß sich im Begriff des Seins überhaupt ein philosophisches Problem auftut, ist schon eine begriffliche Leistung, im wesentlichen die Leistung, die Carnap und seine Nachfolger so gerne diskreditiert gesehen hätten.

Bei Aristoteles scheint sich schon ein Verlust des Problembewußtseins abzuzeichnen:

τὸ [[δὲ]] ὄν λέγεται πολλαχῶς μὲν, ἀλλ' ἅπαν πρὸς μίαν ἀρχήν· τὰ μὲν γὰρ ὅτι οὐσίαι ὄντα λέγεται, τὰ δ' ὅτι πάθη οὐσίας, τὰ δ' ὅτι ὁδὸς εἰς οὐσίαν, ἢ φθορὰ ἢ στερήσεις ἢ ποιότητες ἢ ποιητικὰ ἢ γεννητικὰ οὐσίας, ἢ τῶν πρὸς τὴν οὐσίαν λεγομένων, ἢ τούτων τινὸς ἀποφάσεις ἢ οὐσίας· διὸ καὶ τὸ μὴ ὄν εἶναι μὴ ὄν φαμέν.<sup>35</sup>

Auch wenn Aristoteles das Problem des Parmenides nicht gesehen zu haben scheint, so war er es doch, der das Vorhaben der Wissenschaft, die wir heute als *Metaphysik* bezeichnen, formuliert hat:

Ἔστιν ἐπιστήμη τις ἣ θεωρεῖ τὸ ὄν ἢ ὄν καὶ τὰ τοῦτω ὑπάρχοντα καθ' αὐτό. αὕτη δ' ἐστὶν οὐδεμίᾳ τῶν ἐν μέρει λεγομένων ἢ αὐτῇ· οὐδεμία γὰρ τῶν ἄλλων ἐπισκοπεῖ καθόλου περὶ τοῦ ὄντος ἢ ὄν, ἀλλὰ μέρος αὐτοῦ τι ἀποτεμόμεναι περὶ τούτου θεωροῦσι τὸ συμ-βεβηχός, οἷον αἱ μαθηματικὰ τῶν ἐπιστημῶν.<sup>36</sup>

<sup>35</sup> [21], S. 146 & 148; Aristoteles, *Metaphysik* IV, 2, 1003 b 5–11; in [52], S. 83 übersetzt als:

[[Man]] spricht [...] zwar in vielfacher Bedeutung vom Seienden, doch stets im Hinblick auf *ein* Prinzip. Man spricht nämlich vom Seienden, weil es entweder ein Wesen ist oder eine Affektion eines Wesens oder weil es ein Weg zum Wesen oder weil es Vergehen, Privation, Qualität, Bewirkendes oder Erzeugendes eines Wesens ist oder von etwas, das in Beziehung auf das Wesen ausgesagt wird, oder gar, weil es Verneinung von etwas derartigem oder eines Wesens ist. Daher sagen wir ja auch, ein Nichtseiendes *sei* nichtseiend.

<sup>36</sup> [21], S. 146; Aristoteles, *Metaphysik* IV, 1, 1003 a 19–27. In [52], S. 82, übersetzt als:

Es gibt eine Wissenschaft, die das Seiende, insofern es seiend ist, betrachtet und das, was ihm an sich zukommt. Diese ist aber mit keiner der sogenannten Einzelwissenschaften identisch; denn keine der anderen Wissenschaften betrachtet allgemein das Seiende, insofern es seiend ist, sondern indem sie sich einen Teil vom Seienden herauschneiden, betrachten sie diesen hinsichtlich seines Akzidens, wie das etwa die mathematischen Wissenschaften tun.



Da ist noch keine Rede von letzten Ursachen, höchsten Wesen oder sonstigen Gebilden aus dem Reich des Absoluten. Erste Philosophie, πρώτη φιλοσοφία, betrachtet Seiendes, insofern es seiend ist. Dann geht es aber weiter:

ἐπει δὲ τὰς ἀρχὰς καὶ τὰς ἀκροτάτας αἰτίας ζητοῦμεν, δῆλον ὡς φύσεώς τινος αὐτὰς ἀναγκαῖον εἶναι καθ' αὐτήν. εἰ οὖν καὶ οἱ τὰ στοιχεῖα τῶν ὄντων ζητοῦντες ταύτας τὰς ἀρχὰς ἐζήτουν, ἀνάγκη καὶ τὰ στοιχεῖα τοῦ ὄντος εἶναι μὴ κατὰ συμβεβηχός, ἀλλ' ἢ ὄν· διὸ καὶ ἡμῖν τοῦ ὄντος ἢ ὄν τὰς πρώτας αἰτίας ληπτέον.<sup>37</sup>

Mit einem letzten Zitat von Aristoteles will ich den Kreis schließen, dessen Ausgangspunkt Parmenides (Denken und Sein sind dasselbe) war:

αὐτὸν δὲ νοεῖ ὁ νοῦς κατὰ μετάληψιν τοῦ νοητοῦ· νοητὸς γὰρ γίγνεται θιγγάνων καὶ νοῶν, ὥστε αὐτὸν νοῦς καὶ νοητόν.<sup>38</sup>

Das klingt für mich nicht nur wie eine (teilweise) Rückkehr zu Parmenides (αὐτὸν νοῦς καὶ νοητόν), sondern darüber hinaus auch wie eine Betonung der Selbstbezüglichkeit im Denken: die Vernunft hat am Gedachten Anteil. Was hier angesprochen wird, ist eine Verquickung dessen, was Carnap auseinanderzuhalten versuchte.

Damit bin ich bei dem, was für mich metaphysisches Denken ausmacht: die Ausrichtung der Aufmerksamkeit auf das Denken als Aktivität, das mit seinem Objekt in einer Weise verbunden ist, daß es konstitutiv für die Erkenntnis seines Gegenstandes wird, d.h. aus dem Wissen nicht weggedacht werden kann. Aber in der weiteren Geschichte der Metaphysik scheint dieser Faden erst einmal verloren gegangen zu sein. Die Frage nach dem Sein als solchem (τὸ ὄν ἢ ὄν) wurde lange Zeit durch Fragen nach bestimmten „höchsten“ Entitäten verdrängt. Da es die Hegelsche

<sup>37</sup> Ebd. 27–33. In [52], S. 82, übersetzt als:

Da wir aber die Prinzipien und die höchsten Ursachen suchen, so ist es klar, daß diese Ursachen einer an sich existierenden Natur sein müssen. Wenn nun diejenigen, die die Elemente der Dinge suchten, eben diese Prinzipien suchten, so waren notwendigerweise auch die Elemente nicht Elemente des Seienden in akzidentellem Sinne, [30] sondern des Seienden, insofern es seiend ist. Also müssen auch wir die ersten Ursachen des Seienden, insofern es seiend ist, erfassen.

<sup>38</sup> [22], S. 148; Aristoteles, Metaphysik XII, 1072 b 20–22. In [52], S. 314, übersetzt als:

Sich selbst denkt die Vernunft, indem sie am Gedachten Anteil hat. Gedacht nämlich wird sie selbst, indem sie Gedachtes berührt und denkt, so daß die Vernunft und Gedachtes dasselbe sind.

Aufarbeitung der Metaphysik ist, in deren Tradition mein Begründungsvorschlag steht, soll Hegel selbst bezüglich dieser Metaphysik zu Worte kommen:

Ihre Gegenstände waren zwar Totalitäten, welche an und für sich der Vernunft angehören; — Seele, Welt, Gott — aber die Metaphysik nahm sie aus der Vorstellung auf, legte sie als fertige gegebene Subjecte, bey der Anwendung der Verstandesbestimmungen darauf, zu Grunde, und hatte nur an jener Vorstellung den Maßstab, ob die Prädicate passend und genügend seyen oder nicht.<sup>39</sup>

Weiterhin, so führt Hegel aus,

wurde sie dadurch Dogmatismus, weil sie nach der Natur der endlichen Bestimmungen annehmen mußte, daß von zwey entgegengesetzten Behauptungen, dergleichen jene Sätze waren, die eine wahr, die andere aber falsch seyn müsse.<sup>39</sup>

Diese Metaphysik ist durch Kants *Kritik der reinen Vernunft* gewaltig aufgemischt worden. Kant entwickelt darin ein Argument, das zeigen soll, daß wir niemals die *Dinge an sich* erkennen können, sondern nur so, wie sie uns erscheinen, d.h. im wesentlichen durch die Formen der Anschauung (Raum und Zeit), und die Kategorien des Verstandes, die in den vier Gruppen Quantität, Qualität, Relation, und Modalität angeordnet sind.

Kants Argument basiert wesentlich auf der Konstruktion von gewissen beweisbaren Widersprüchen (Antinomien) in der reinen Vernunft (Widerstreit der transzendentalen Ideen). In einem Abschnitt mit dem Titel „Antithetik der reinen Vernunft“ entwickelt Kant vier Antinomien, betreffend Anfang und Ende des Raums und der Zeit, (un)endliche Teilbarkeit der Materie, Kausalität und Freiheit, und Existenz eines höchsten Wesens. Hegels Idee einer Metaphysik als Logik setzt an Kants Interpretation und Auflösung dieser Antinomien der reinen Vernunft an:

Die Hauptsache, die zu bemerken ist, ist, daß nicht nur in den vier besonders aus der Kosmologie genommenen Gegenständen die Antinomie sich befindet, sondern vielmehr in allen Vorstellungen, Begriffen und Ideen. Dieß zu wissen und die Gegenstände in dieser Eigenschaft zu erkennen, gehört zum Wesentlichen der philosophischen Betrachtung; diese Eigenschaft macht das aus,

<sup>39</sup> [24], S. 39.

was weiterhin sich als das dialektische Moment des Logischen bestimmt.<sup>40</sup>

In einem darauffolgenden Zusatz heißt es dann:

Auf dem Standpunkt der alten Metaphysik wurde angenommen, daß wenn das Erkennen in Widersprüche gerathe, so sey dieses nur eine zufällige Verirrung und beruhe auf einem subjectiven Fehler im Schließen und Raisoniren. Nach Kant hingegen liegt es in der Natur des Denkens selbst in Widersprüche (Antinomien) zu verfallen, wenn dasselbe das Unendliche erkennen will. Ob nun schon [. . .] das Aufzeigen der Antinomien insofern als eine sehr wichtige Förderung der philosophischen Erkenntniß zu betrachten ist, als dadurch der starre Dogmatismus der Verstandesmetaphysik beseitigt und auf die dialektische Bewegung des Denkens hingewiesen worden ist, so muß doch dabei zugleich bemerkt werden, daß Kant auch hier bei dem blos negativen Resultat der Nichterkennbarkeit des Ansich der Dinge stehen geblieben und nicht zur Erkenntniß der wahren und positiven Bedeutung der Antinomien hindurch gedrungen ist. Die wahre und positive Bedeutung der Antinomien besteht nun überhaupt darin, daß alles Wirkliche entgegengesetzte Bestimmungen in sich erhält und daß somit das Erkennen und näher das Begreifen eines Gegenstandes eben nur so viel heißt, sich dessen als einer konkreten Einheit entgegengesetzter Bestimmungen bewußt zu werden.<sup>41</sup>

Hier setzt nun Hegels logische Wende an.

Die Logik, in der wesentlichen Bedeutung speculativer Philosophie, tritt an die Stelle dessen, was sonst Metaphysik genannt und als eine von ihr abgesonderte Wissenschaft abgehandelt wurde.<sup>42</sup>

Hegel selbst hat schon — allerdings mit äußerst zweifelhaftem Erfolg — in seiner *Wissenschaft der Logik* versucht, die Idee der Metaphysik als Logik umzusetzen. Ich werde nicht versuchen, seinen Argumentationen einen Sinn abzugewinnen. Die Widersprüche von Kant und Hegel sind nicht schlüssig, und Generationen von fleißigen Exegeten haben daran

<sup>40</sup> [25], S. 141.

<sup>41</sup> Ebd., S. 141 f.

<sup>42</sup> [24], S. 38.

nichts ändern können. Statt dessen werde ich mir die Widersprüche der uneingeschränkten Abstraktion zunutze machen. Metaphysik nach Hegelschem Muster zielt dann auf die Widersprüche ab, die Carnap ausschließen will.

Damit wird Metaphysik als Logik zu einer Theorie der Begriffe, etwa in der Tradition von Freges Grundgesetzen, die auf zweierlei abzielt:

1. auf eine Definition in rein logischen Termen von grundlegenden Denkbestimmungen, wie zum Beispiel der Modalität, Quantität, Kausalität, Zeit und Raum (um nur die nächstliegenden zu nennen), und
2. auf eine Ableitung ihrer charakteristischen Gesetze durch rein logische Mittel.

Dafür entlehne ich von Hegel die Bezeichnung „spekulative Logik“.<sup>43</sup> Das darf jedoch nicht zu der Annahme verleiten, ich wolle Hegels Logik in irgendeiner Weise zur Darstellung bringen, sei es durch „Formalisierung“, „Rekonstruktion“, oder sonst einer Form der Auseinandersetzung mit Hegels Text. In den 175 Jahren seit Hegels Tod hat die Hegelinterpretation ihre Unfähigkeit genügend dokumentiert, und ich sehe nicht, daß daran etwas geändert werden kann. Weder Hegel noch seine Exegeten können einen *Inhalt* vorweisen, an dem sich die Ansprüche der spekulativen Logik und Philosophie auf eine besondere Art der Erkenntnis, die sich dem klassischen Denken entzieht, messen lassen. Um einen solchen Inhalt geht es aber im folgenden. Dabei muß klar sein, daß auch wenn die ursprüngliche Idee für einen solchen Inhalt aus der Philosophie Hegels stammt, der hier vorgestellte Inhalt erst durch die Methoden und Resultate der mathematischen Logik zutage getreten ist und sich auch nur durch sie erschließen läßt.

<sup>43</sup> Dieser Ausdruck geht zwar auf Hegel zurück, wurde aber von ihm nur sehr sporadisch gebraucht. Meines Wissens nur in [25], S. 53: „die spekulative Logik enthält die vorige Logik und Metaphysik ...“ und S. 195/196, weiterhin [24], S. 36: „In der spekulativen Logik ist die bloße Verstandes-Logik enthalten ...“, und S. 275, daß es „gerade die Philosophie, und zwar zunächst die spekulative Logik, ist, welche die Nichtigkeit der ... bloßen Verstandesbestimmungen aufzeigt“, sowie [26], S. 57: „die speculative Logik ... zeigt, dass alle jene auf die Seele angewandten Bestimmungen ... in ihr Gegenteil umschlagen“. Hegel verwendet ihn in einem Kontext, der ihn in unmittelbare Nähe zum Ausdruck „spekulative Philosophie“ stellt, wie etwa in [23], S. 171 (§14): „Die Logik ist ... an sich selbst spekulative Philosophie“.

### 3. Was heißt hier „Selbstbezüglichkeit“?

Es stünde schlecht um das Projekt einer Zurückführung der charakteristischen Eigenschaften von Denkbestimmungen auf Eigenschaften widersprüchlicher Begriffe, gäbe es nur Kants Antinomien und die Widersprüche, die Hegel selbst entdeckt zu haben glaubte. Deshalb eignet sich Carnaps Versuch einer Überwindung der Metaphysik so gut als Hintergrund für meinen Versuch einer Begründung der Metaphysik. Die (einfache) Typentheorie, die Carnap bemüht hat, um metaphysische Satzkonstruktionen als sinnlos zu brandmarken, war von Russell entworfen worden, um Paradoxien in den Grundlagen der (höheren) Logik und Mengenlehre zu vermeiden. Der Effekt der Typenunterscheidung war die Vermeidung einer Form von Selbstbezüglichkeit, wie sie insbesondere der Russellschen Antinomie zugrunde liegt. Mit anderen Worten, wir brauchen uns um Widersprüche nicht zu bemühen — sie sind längst da.

Daß den Widersprüchen der Logik eine erkenntnistheoretische Bedeutung zukommen könnte, die über die rein negative Dimension der Widerlegung einer Grundannahme hinausgeht, wird Carnap wohl nie in den Sinn gekommen sein.<sup>44</sup> Aber für eine Philosophie, die in Widersprüchen den Schlüssel zu den Kategorien sucht, sollten sie wie ein Geschenk des Himmels kommen.<sup>45</sup>

Natürlich ist damit noch nicht gesagt, daß es sich bei den Paradoxien der Logik tatsächlich um Widersprüche handelt, die sich für eine Grundlegung von Denkbestimmungen eignen. Aber es muß klar sein, daß dies eine Frage ist, die ich zu einer Angelegenheit der logischen Forschung — und nicht des philosophischen Wunschdenkens — machen will.

<sup>44</sup> Das ist keine Besonderheit logischer Empiristen. Man vergleiche in dieser Hinsicht [4], p. 26:

Der Widerspruch ist nicht eine unumgängliche Station auf dem Wege der Wahrheit, sondern ein Zeichen dafür, daß man von diesem Wege abgekommen ist. Er ist im Prinzip vermeidbar und hat keine spekulative Kraft.

Da bleibt nur noch die Frage, ob die Verkündung solcher Einsichten von der hohen Warte der Philosophie für oder gegen die Philosophie spricht.

<sup>45</sup> In diesem Sinne kann man auch folgende Bemerkung in [11], S. 204, verstehen:

Hegel hat die sprachlichen wie die ontologischen Antinomien der modernen Logik nicht gekannt. Es darf als gesichert gelten, daß er an ihnen seine helle Freude gehabt hätte.

Der Gedanke spricht mich an, scheint mir aber doch zu optimistisch. Davon abgesehen, wenn es so klar ist, daß Hegel die modernen logischen Paradoxien zu würdigen gewußt hätte, warum sind dann seine Adepten so zimperlich?

Vorher aber will ich auf die Frage eingehen, wie überhaupt Selbstbezüglichkeit entsteht.

Selbstbezüglichkeit muß hergestellt werden — und das ist nicht immer einfach. Dabei kann einiges schiefgehen, sei es durch Unverstand oder widrige Umstände. Das folgendes Beispiel aus der anglophonen Hegel-Literatur soll das illustrieren:

Consider a phrase like, “The phrase on page 5.” Note that it applies to itself. Note that if I wrote that phrase on another page, it would not apply to itself, but to the phrase on page 5. Now consider something like, “This very phrase.” It refers to itself irrespective of which page it is written on. It refers to itself not contingently, as does the phrase, “The phrase on page 5” but essentially[.]<sup>46</sup>

Diese Passage steht auf S. 7; als ob ein übelgesinnter Setzer sich über vermeintlichen philosophischen Tiefsinn lustig machen wollte. Oder vielleicht auch nur vom Unsinn ablenken, denn das ist nicht, was an diesem Beispiel so grundsätzlich schiefgegangen ist; es kann leicht richtig gestellt werden, indem man „5“ durch „7“ ersetzt. Das Problem ist, daß auf S. 7 (wie auch auf S. 5) mehr als nur ein einziger Ausdruck steht, und daher der bestimmte Artikel keinen der Ausdrücke auf S. 7 bezeichnet.<sup>47</sup> Sicher kann man auch diesen Mangel beheben, indem man nur den einen Ausdruck „der Ausdruck auf S. 7“ auf S. 7 stehen läßt; aber die Frage bleibt: Sollte es im Jahre 1989 tatsächlich noch Philosophen gegeben haben, denen nicht klar war, worum es beim bestimmten Artikel geht? Und was ist mit der Unterscheidung von „kontingenter“ und „wesentlicher“ Selbstbezüglichkeit? Kann ein Demonstrativpronomen wie „dieser“ die Bürde der „Wesentlichkeit“ tragen, auch wenn es (im Englischen) durch ein „very“ verstärkt wird? Im Deutschen würde man „eben“ oder „genau“ sagen. Und genau darum scheint es mir in dem folgenden Zitat zu gehen:

»Dies« bezeichnet, was immer im Augenblick, wo das Wort gebraucht wird, das Aufmerksamkeitszentrum in Anspruch nimmt. Bei Wörtern, die nicht egozentrisch sind, ist der unveränderliche Gehalt etwas, was den Gegenstand betrifft, auf den sie sich beziehen, aber »dies« bedeutet in jedem Fall des Gebrauchs etwas an-

<sup>46</sup> [40], S. 7.

<sup>47</sup> Pinkard spricht von „application“, aber ich denke, man kann das durch „denoting“ ersetzen, ohne Pinkard Unrecht zu tun.

deres. Was hier unveränderlich ist, ist nicht der Gegenstand, auf den sich das Wort bezieht, sondern dessen Beziehung zu dem besonderen Gebrauch des Wortes. So oft das Wort gebraucht wird, hat die Person, die es gebraucht, auf irgend etwas ihre Aufmerksamkeit gerichtet, und dieses irgend etwas bezeichnet das Wort. Wenn ein Wort nicht egozentrisch ist, so braucht es zwischen den verschiedenen Anlässen nicht unterschieden zu werden, bei denen es benutzt worden ist; wohl aber müssen wir diesen Unterschied bei egozentrischen Wörtern machen, weil das, was sie anzeigen, etwas ist, was eine gegebene Beziehung zu dem besonderen Gebrauchsfall des Wortes hat.<sup>48</sup>

Ich hoffe das reicht, um die Schwierigkeiten beim Gebrauch eines Demonstrativpronomens wie „dieser“ anzudeuten.

Was dem Autor des oben zitierten Versuchs zur Herstellung von Selbstbezüglichkeit vorgeschwebt haben mag, war vielleicht etwas wie die folgende Darstellung der Antinomie des Lügners bei Tarski:

[[B]]etrachten wir die Aussage:

Die Aussage 149 Z 20 dieses Aufsatzes ist nicht wahr.

Der Kürze wegen wollen wir diese Aussage durch  $\rangle s \langle$  ersetzen. Gemäß unserer Vereinbarung bezüglich des angemessenen Gebrauchs des Terms  $\rangle \text{wahr} \langle$  behaupten wir nun die Äquivalenz der Form (T):

(1)  $\rangle s \langle$  ist genau dann wahr, wenn die Aussage 149 Z 20 dieses Aufsatzes nicht wahr ist.

Wenn wir uns an den Sinn von  $\rangle s \langle$  erinnern, dann stellen wir empirisch fest:

(2)  $\rangle s \langle$  ist mit der Aussage 149 Z 20 dieses Aufsatzes identisch.

Nun folgt nach einem bekannten Gesetz der Theorie der Identität (Leibniz) aus (2), daß wir in (1) den Ausdruck  $\rangle \text{die Aussage 149 Z 20 dieses Aufsatzes} \langle$  durch  $\rangle s \langle$  ersetzen können. Dergestalt erhalten wir:

(3)  $\rangle s \langle$  ist wahr genau dann, wenn  $\rangle s \langle$  nicht wahr ist.

<sup>48</sup> [49], S. 100.

Wir kommen auf diese Weise also zu einem offenkundigen Widerspruch.<sup>49</sup>

Merkwürdigerweise gebraucht Tarski beide Formen der Selbstbezüglichkeit, die indirekte Art der eindeutigen Beschreibung des Satzes und das Demonstrativpronomen „dieser“. Die Wendung „dieser Aufsatz“ kann vermieden werden durch einen Verweis auf den Band der Zeitschrift, in welchem der Aufsatz zuerst erschienen ist. Warum Tarski diese Möglichkeit nicht nutzt und stattdessen auf das Demonstrativpronomen „dieser“ zurückgreift, ist nicht klar, denn worum es ihm geht, ist zu zeigen, daß das empirische Element in der Herstellung der Selbstbezüglichkeit nicht wesentlich ist. Es hätte darüber hinaus den Vorteil, daß die Formulierung auch im Zitat noch funktionieren würde,<sup>50</sup> wie beispielsweise im gegenwärtigen Fall, wo „dieser Aufsatz“ irreführend sein kann.

Ich will das nicht weiter vertiefen, weil es einfache Möglichkeiten gibt, Selbstbezüglichkeit herzustellen, die keinen Gebrauch von Demonstrativpronomina machen. Was mir hier wichtiger erscheint, ist die Möglichkeit, Paradoxien der Selbstbezüglichkeit zu konstruieren, die unabhängig von einem empirischen Auffinden sind. Als Beispiel einer solchen Paradoxie führt Tarski die folgende Formulierung an:

*S* sei eine Aussage die mit den Wörtern ›Jede Aussage‹ beginnt. Wir bringen mit *S* die neue Aussage *S*\* in Wechselbeziehung, indem wir *S* den folgenden beiden Modifikationen unterwerfen: wir ersetzen in *S* das erste Wort ›Jede‹ durch ›Die‹ und fügen an das zweite Wort ›Aussage‹ die ganze Aussage *S* in Anführungszeichen. Wir wollen vereinbaren, die Aussage *S* ›(auf sich selbst) anwendbar‹ oder ›nicht (auf sich selbst) anwendbar‹ zu nennen, und zwar in Abhängigkeit davon, ob die in Wechselwirkung zu ihr stehende Aussage *S*\* wahr oder falsch ist. Betrachten wir nun die Aussage: ›Jede Aussage ist nichtanwendbar‹. Es kann leicht gezeigt werden, daß die oben aufgestellte Aussage anwendbar und nichtanwendbar sein muß, also eine Kontradiktion.<sup>51</sup>

<sup>49</sup> [57], S. 149. Der Vollständigkeit halber will ich darauf hinweisen, daß die Zahlenangaben tatsächlich richtig sind in dem Sinn, daß der fragliche Satz tatsächlich der einzige Satz in Zeile 20 auf S. 149 des angegebenen Aufsatzes ist.

<sup>50</sup> Sozusagen eine „portable paradox formulation“, kurz „ppf“, im Unterschied zu „pdf“.

<sup>51</sup> [57], S. 183 (Anmerkung 12).



Im Vergleich zu dieser Konstruktion von Selbstbezüglichkeit zur Erzeugung einer semantischen Paradoxie erscheint das mengentheoretische Vorgehen zur Herstellung von Selbstbezüglichkeit fast trivial: Grundlage ist die Aussageform  $*_1 \in *_1$ .<sup>52</sup> Am Beispiel der Russellschen Antinomie sieht das dann so aus:  $R := \{x : x \notin x\}$ . Nun nimmt man sich  $R \in R$  vor und erhält:

$$R \in \{x : x \notin x\} \leftrightarrow R \notin R,$$

das heißt:

$$R \in R \leftrightarrow R \notin R.$$

Mehr braucht es nicht, um klassische Logik ins Chaos zu stürzen. *Tertium non datur*  $A \vee \neg A$  erledigt den Rest, indem es daraus die beiden Formeln  $R \in R$  und  $R \notin R$ , also einen Widerspruch ableitet.

Gegen die hier verwendete Typenvermischung von der einfachen Gestalt ( $x \in x$ ) richten sich sowohl die (einfache) Typentheorie als auch die Systeme der axiomatischen Mengenlehre. Sie sind erfolgreich, insoweit es darum geht, die Paradoxien auszuschließen,<sup>53</sup> aber nicht, insoweit es darum geht, Selbstbezüglichkeit grundsätzlich zu vereiteln. Gödel hat gezeigt, wie man in der Arithmetik eine Form der Selbstbezüglichkeit erzeugen kann, die — wie die Anführung von Tarski — mit einem Wahrheitsprädikat zu Widersprüchen führt. Das Verfahren ist aufwendig,<sup>54</sup> ich will aber versuchen, die wesentlichen Stationen zu skizzieren.

Was ich als bekannt voraussetze, ist die Möglichkeit, jedem Grundzeichen einer formalen Sprache in eindeutiger Weise eine natürliche Zahl zuzuordnen.<sup>55</sup>

<sup>52</sup> Aussageform in dem Sinne, daß eine Aussage entsteht, wenn alle Vorkommnisse des Zeichen  $*_1$  durch ein- und demselben Objektterm ersetzt werden.

<sup>53</sup> Ein Widerspruchsfreiheitsbeweis für die einfache Typentheorie mit Unendlichkeitsaxiom existiert in der Form der Schnitteliminierbarkeit der einfachen Typentheorie ohne Unendlichkeitsaxiom (Beweis von „Takeuti's Vermutung“, vgl. [16], S. 176 ff). Ein Widerspruchsfreiheitsbeweis für die „allgemeine Mengenlehre“, i.e. die Zermelo-Fraenkel Mengenlehre ohne Unendlichkeitsaxiom, ist von Ackermann 1937 geliefert worden.

<sup>54</sup> Gödel führte es 1931 in seinem berühmten Aufsatz 'Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I' ein; [18] im Literaturverzeichnis.

<sup>55</sup> Da eine formale Sprache normalerweise unendlich viele Variablen unter ihren Grundzeichen hat, ist das nicht unbedingt völlig trivial. Man behilft sich mit Primzahlen, von denen es nachweislich unendlich viele gibt, und die, sozusagen, *immun* sind gegenüber der Möglichkeit einer Zerlegung in arithmetische Bestandteile.

Der zweite Schritt ist der, Reihen von natürlichen Zahlen wiederum eine natürliche Zahl so zuzuordnen, daß aus dieser Zahl die Reihe wieder rekonstruierbar ist. Das geschieht unter Heranziehung des *Gaußschen Fundamentalsatzes der Arithmetik*,<sup>56</sup> nach dem jede natürliche Zahl größer als 1 in bis auf die Reihenfolge eindeutiger Weise als ein Produkt von Primzahlen geschrieben werden kann oder selbst eine Primzahl ist. Auf diese Weise kann jedem Ausdruck  $X$  der formalen Sprache auf umkehrbar eindeutige Weise eine Ziffer zugeordnet werden, die sogenannte Gödelzahl  $\ulcorner X \urcorner$ . Dabei zeigen die kleinen Ecken  $\ulcorner \urcorner$  die numerische Kodifizierung eines Ausdrucks an.

Der nächste Schritt ist die Definition einer Substitutionsfunktion in der Arithmetik. Diese Funktion ist in — vergleichsweise — elementarer Weise berechenbar (primitiv-rekursiv) und kann damit als arithmetische Funktion dargestellt werden, d.h. die folgende Gleichung ist für jede Aussageform  $\mathfrak{A}$  und jede natürliche Zahl  $n$  in der Arithmetik beweisbar:

$$\text{sub}(\ulcorner \mathfrak{A}[x] \urcorner, n) = \ulcorner \mathfrak{A}[n] \urcorner.$$

Umgangssprachlich würde das etwa folgendermaßen lauten:  $\text{sub}(\ulcorner \mathfrak{A}[x] \urcorner, n)$  ist gleich der Gödelzahl des Ausdrucks  $\mathfrak{A}[n]$ , d.h. des Ausdrucks der aus  $\mathfrak{A}[x]$  entsteht, wenn jedes Vorkommen von  $x$  durch  $n$  ersetzt wird.

Nun geht man folgendermaßen vor: man wählt für beide Argumente dieser Substitutionsfunktion die Gödelzahl des Ausdrucks  $\mathfrak{A}[\text{sub}(x, x)]$ ,<sup>57</sup> d.h.  $\ulcorner \mathfrak{A}[\text{sub}(x, x)] \urcorner$ . Wir nehmen nun  $k_{\mathfrak{A}}$  als abkürzende Schreibweise für  $\ulcorner \mathfrak{A}[\text{sub}(x, x)] \urcorner$  und erhalten so die folgende („indirekte“) „Fixpunkteigenschaft“:<sup>58</sup>

$$\text{sub}(k_{\mathfrak{A}}, k_{\mathfrak{A}}) = \ulcorner \mathfrak{A}[\text{sub}(k_{\mathfrak{A}}, k_{\mathfrak{A}})] \urcorner.$$

Weshalb das „Fixpunkteigenschaft“ heißt, sollte ausreichend klar werden, wenn wir die Abkürzung  $f_{\mathfrak{A}}$  für  $\text{sub}(k_{\mathfrak{A}}, k_{\mathfrak{A}})$  hernehmen:

$$f_{\mathfrak{A}} = \ulcorner \mathfrak{A}[f_{\mathfrak{A}}] \urcorner.$$

<sup>56</sup> So genannt, weil Gauß ihn als erster bewiesen hat.

<sup>57</sup> Die Selbstbezüglichkeit wird hier durch  $\text{sub}(x, x)$  hergestellt, das offensichtlich die Rolle von  $(x \in x)$  übernimmt.

<sup>58</sup> Es muß klar sein, daß eine „abkürzende Schreibweise“ nur ein Mittel zur einfacheren Darstellung ist, kein Teil der formalen Sprache. Sie kann also jederzeit wieder eliminiert werden.

Dann ist  $f_{\mathfrak{A}}$  ein Fixpunkt bezüglich  $\mathfrak{A}$  in folgendem Sinn: Wenn man  $\mathfrak{A}$  als Aussagefunktion betrachtet, dann ist ihr Wert für das Argument  $f_{\mathfrak{A}}$  wiederum  $f_{\mathfrak{A}}$  selbst.

Damit kann nun ein Fixpunktergebnis der folgenden Gestalt bewiesen werden:

$$F \leftrightarrow \mathfrak{F}[\ulcorner \mathfrak{C}[F] \urcorner].$$

Eine derartige Fixpunkteigenschaft verursacht Schwierigkeiten für die Ausdrückbarkeit von grundlegenden semantischen Begriffen auf der Ebene der formalisierten Theorie selbst (d.h. als ein arithmetisches Prädikat im Falle der Arithmetik, wie z.B. das Prädikat, eine Primzahl zu sein), allen voran das Wahrheitsprädikat, also des Prädikats *tru*, das die folgende Wahrheitsbedingung erfüllt:

$$tru(\ulcorner A \urcorner) \leftrightarrow A.$$

Mit einem Fixpunkt der Gestalt

$$\ulcorner \neg tru(f) \urcorner = f$$

führt das zu

$$tru(f) \leftrightarrow tru(\ulcorner \neg tru[f] \urcorner) \leftrightarrow \neg tru(f),$$

was wiederum mit der klassischen Logik unverträglich ist.

Damit haben wir zwei Arten zur Herstellung von Selbstbezüglichkeit kennen gelernt, die — unabhängig von Demonstrativpronomina, Beschreibung und empirischer Auffindung oder semantischen Methoden wie Anführung — rein syntaktisch funktionieren: mengentheoretisch (höhere Logik) durch Komprehension (Abstraktion) und arithmetisch durch Kodierung. Das läßt sich folgendermaßen in einer Tabelle gegenüberstellen:

mengentheoretisch	arithmetisch
$x \in x$	$sub(x, x)$
$R_{\mathfrak{A}} := \{x : \mathfrak{A}[x \in x]\}$	$k_{\mathfrak{A}} := \ulcorner \mathfrak{A}[sub(x, x)] \urcorner$
$F_{\mathfrak{A}} := R_{\mathfrak{A}} \in R_{\mathfrak{A}}$	$f_{\mathfrak{A}} := sub(k_{\mathfrak{A}}, k_{\mathfrak{A}})$
$\mathfrak{A}[F_{\mathfrak{A}}] \leftrightarrow F_{\mathfrak{A}}$	$\ulcorner \mathfrak{A}[f_{\mathfrak{A}}] \urcorner = f_{\mathfrak{A}}$

Als Beispiel für die linke Spalte habe ich Russells Antinomie angeführt, wo  $\mathfrak{A}$  im Wesentlichen die Negation ist:  $\neg F \leftrightarrow F$ . Als Beispiel für die zweite Spalte habe ich Tarskis Wahrheitsantinomie genannt.

Die Typentheorie wurde entworfen, um die erste (mengentheoretische) Form der Selbstbezüglichkeit zu unterbinden. Die zweite (arithmetische) Form der Selbstbezüglichkeit läßt sich nicht unterbinden, sobald ein gewisses Maß an Arithmetik zur Verfügung steht.

Es ist die indirekte Form der Selbstbezüglichkeit, die philosophisch aufschlußreich ist. Es kann passieren, daß eine Aussage über einen Sachverhalt auf der Gegenstandsebene gleichzeitig als eine (kodierte) Aussage über einen Sachverhalt auf der Theorieebene interpretiert werden kann, und zwar so, daß die beiden unvereinbar sind. Das klassische Beispiel ist die Konstruktion von Gödels unentscheidbarem Satz. Eine formalisierte Theorie, die die Formulierung aller primitiv-rekursiven Funktionen erlaubt, wie zum Beispiel die Theorie **PRA** der primitiv-rekursiven Arithmetik, kann so in den Objektbereich abgebildet werden, daß es möglich wird, in dieser Theorie Aussagen — in kodierter Form — über dieselbe zu machen.

Aussagenebene		Objektebene
$\mathfrak{A}[sub(x, x)]$	$\longleftarrow$ $\searrow$	$sub(x, x)$ $k_{\mathfrak{A}} := \ulcorner \mathfrak{A}[sub(x, x)] \urcorner$
$f_{\mathfrak{A}} = \ulcorner \mathfrak{A}[f_{\mathfrak{A}}] \urcorner$	$\longleftarrow$	$\downarrow$ $f_{\mathfrak{A}} := sub(k_{\mathfrak{A}}, k_{\mathfrak{A}})$
$\downarrow$		
$G_{\mathfrak{A}} := \mathfrak{A}[f_{\mathfrak{A}}]$		
$\downarrow$		
$G_{\mathfrak{A}} \leftrightarrow \mathfrak{A}[\ulcorner G_{\mathfrak{A}} \urcorner]$		

Man mache sich die Gleichung

$$sub(\ulcorner \mathfrak{A}[sub(x, x)] \urcorner, k_{\mathfrak{A}}) = \ulcorner \mathfrak{A}[sub(k_{\mathfrak{A}}, k_{\mathfrak{A}})] \urcorner$$

klar, die den letzten Übergang von der Objektebene zur Aussagenebene kennzeichnet. Sie ist das Herz der (indirekten) Fixpunkteigenschaft.

Das Hin- und Her zwischen den beiden Ebenen im obigen Diagramm ist charakteristisch für das Herstellen von Selbstbezüglichkeit in Theorien, die ausreichend Arithmetik enthalten:

*self-reference is never direct: it comes from a controlled confusion of two levels of meaning for integers, which are seen both as numbers and as names for formulas.*<sup>59</sup>

Das ist es, was Wittgenstein nicht sehen konnte oder wollte; oder, um mit Adorno zu reden, was „den Positivisten wie Musik in den Ohren mißtönt“.<sup>60</sup> Aber, um gleich Adorno und Co. auf Distanz zu halten, dieses Moment des „Unauflöslchen“ ist keine Angelegenheit sprachlicher Verrenkungen und Kunststückchen, wie sie uns von einer gewissen philosophischen Seite so gerne vorgeführt werden, sondern ein Ergebnis der formalisierten Arithmetik, das wiederum zu einem integralen Bestandteil der Informatik geworden ist.

Mit anderen Worten, Selbstbezüglichkeit erzeugt Doppeldeutigkeit („Ambiguität“), die sich im Falle der Arithmetik nicht vermeiden läßt. Das hat eine erkenntnistheoretische Bedeutung. Ambiguität ist eine unvermeidliche „Realität“ unserer Beschreibung, die den Puristen nicht gelegen kommen mag, die aber Raum für eine Grundlegung von Denkbestimmungen nach Hegelschem Muster schafft: Bestimmungen, die ihren Ursprung in dieser Ambiguität haben, können zwar nicht als objektiv in dem Sinne gelten, daß sie den Objekten *an sich* zuzuschreiben wären, aber dennoch sind sie intersubjektiv überprüfbar, d.h. es kommt ihnen eine subjektunabhängige Gültigkeit zu.

Im Unterschied zur arithmetischen Form der Selbstbezüglichkeit, die, wie Odifreddi bemerkt, nie direkt ist, verträgt sich die mengentheoretische Selbstbezüglichkeit, wie sie z.B. in Russells Antinomie verwendet wurde, nicht mit der klassischen Logik. Das weist auf einen grundsätzlichen Unterschied zwischen den beiden Formen der Selbstbezüglichkeit hin. Dann stellt sich aber die Frage, ob bzw. wie weit die klassische Logik überhaupt für die indirekte Form der Doppeldeutigkeit angemessen ist. Es ist durchaus denkbar, daß die klassische Logik bei der Anwendung auf die indirekte Fixpunkteigenschaft zu Ergebnissen führt, die in gewisser Weise *falsch* sind. Diesen Gedanken will ich aber hier nicht weiter verfolgen, da ich ohnehin nur mit der direkten Fixpunkteigenschaft arbeiten werde, und da führt an einer Einschränkung der Logik kein Weg vorbei.

<sup>59</sup> [35], S. 165; kursiv im Original.

<sup>60</sup> [34], S. 45.

#### 4. Reaktionen auf die Paradoxien der Selbstbezüglichkeit

Was ich oben (S. 92) hinsichtlich des Begriffs des Seins gesagt habe, gilt auch für die Paradoxien der Selbstbezüglichkeit: Zu sehen, daß sich hier überhaupt ein philosophisches Problem auftut, ist schon eine begriffliche Leistung. Daß es sich dabei nicht um eine Selbstverständlichkeit handelt, wird beispielsweise durch folgendes Zitat nahegelegt:

Sätze wie [der Lügner], die ihre eigene Falschheit aussagen, müssen offenbar vermieden werden, und wir vermeiden sie auch faktisch in der natürlichen Sprache aus ebenso prinzipiellen Gründen wie den Widerspruch überhaupt.<sup>61</sup>

Wie gut für die natürliche Sprache und die Philosophie, die sich ihrer bedient. Damit kann das Problem der Paradoxien aus der Philosophie verabschiedet und in den Bereich deduktiver Systeme abgeschoben werden.

Für Frege ließ Russells Antinomie keinen Zweifel an der Unverträglichkeit seines Grundgesetzes V mit dem Gesetz vom ausgeschlossenen Dritten. In einem Nachwort zu [15] erwägt Frege Möglichkeiten, dieser Situation zu begegnen. Zwei Fragen ragen heraus:<sup>62</sup>

1. Sollen wir annehmen, das Gesetz vom ausgeschlossenen Dritten gelte von den Klassen nicht?<sup>63</sup>
2. Oder sollen wir annehmen, es gebe Fälle, wo einem unanfechtbaren Begriffe keine Klasse entspreche, die sein Umfang wäre?

Das läuft auf folgende Alternative hinaus: Einschränkung der Logik oder Einschränkung der Begriffsbildung. Der Weg einer Einschränkung der Begriffsbildung wird typischerweise von der einfachen Typentheorie und Zermelos Mengenlehre beschrritten. Auch Frege selbst hatte noch im Nach-

<sup>61</sup> [60], p. 197. Die „prinzipiellen Gründe“ würden mich interessieren, aber darüber von einem Philosophen detaillierten Aufschluß zu erhalten, ist wohl zu viel verlangt.

<sup>62</sup> [15], S. 254.

<sup>63</sup> In [15], auf S. 69, sagt Frege:

Eine Definition eines Begriffes (möglichen Prädikates) muss vollständig sein, sie muss für jeden Gegenstand unzweideutig bestimmen, ob er unter den Begriff falle (ob das Prädikat mit Wahrheit von ihm ausgesagt werden könne) oder nicht.

Eine Aufweichung dieses Grundsatzes scheint mir Frege hier als Möglichkeit in Betracht zu ziehen, ohne jedoch grundsätzlich das logische Gesetz vom ausgeschlossenen Dritten anzuzweifeln.

wort zum zweiten Band der Grundgesetze einen Vorschlag gemacht, wie Umfänge einzuschränken seien,<sup>64</sup> der sich aber als unzureichend herausstellte.

Etwa zur gleichen Zeit machte sich Russell Gedanken, wie er der nach ihm benannten Antinomie begegnen sollte. Hier ist seine erste, noch vorläufige, Formulierung der (einfachen) Typentheorie:

It will [ . . . ] be necessary to distinguish (1) terms, (2) classes, (3) classes of classes, and so on *ad infinitum*; we shall have to hold that no member of one set is a member of any other set, and that  $x \in u$  requires that  $x$  should be of a set of a degree lower by one than the set to which  $u$  belongs. Thus  $x \in x$  will become a meaningless proposition; and in this way the contradiction is avoided.<sup>65</sup>

Das läuft auf eine Einschränkung der Sprache hinaus, in dem Sinne, daß  $x \in x$  gar nicht erst gebildet werden kann. Hier kündigt sich Carnaps Sphärenunterscheidung an.

Es war Russell jedoch klar, daß eine solche (einfache) Typenunterscheidung gegen die sogenannten *semantischen* Paradoxien nichts ausrichten konnte, da letztere nicht von Mengen und Elementen handeln.

Erst fünf Jahre später formulierte Russell sein berühmtes *vicious circle principle*, das eine zusätzliche Schichtung innerhalb der Typen erforderte, wodurch auch die semantischen Paradoxien ausgeschlossen wurden.

These fallacies [ . . . ] are to be avoided by what may be called the ‘vicious-circle principle’; i.e., ‘no totality can contain members defined in terms of itself’. This principle, in our technical language, becomes: ‘Whatever contains an apparent variable must not be a possible value of that variable’. Thus whatever contains an apparent variable must be of a different type from the possible values of that variable; we will say that it is of a *higher* type.

<sup>64</sup> In [15], auf S. 262. Die uneingeschränkte Abstraktion hatte Frege nicht unter die Grundgesetze aufgenommen, sie kann aber unschwer aus Grundgesetz V abgeleitet werden. Frege setzte sich ausführlich mit dem Problem der Begriffsbildung auseinander, insbesondere wie Begriffe als Objekte der Sprache zur Verfügung stehen können. Seine philosophische Analyse ist in seinem Aufsatz „Begriff und Gegenstand“ enthalten, wo er den Begriff „Werthverlauf“ einführt. Frege hat auch daran gedacht, zu beweisen, daß alle Terme eine Bedeutung haben. Dabei ist ihm aber ein Fehler unterlaufen, wie Bartlett [2], S. 72 f, aufgezeigt hat.

<sup>65</sup> [46], S. 517.

Thus the apparent variables contained in an expression are what determines its type. This is the guiding principle[.]<sup>66</sup>

Zu dieser Zeit legte auch Zermelo seine Untersuchungen zur Grundlegung der Mengenlehre vor. Sein Konzept hat in der Form der Zermelo-Fraenkel-Skolem-Mengenlehre bis heute Bestand.

In der Mengenlehre hat man sich darauf zurückgezogen, daß der ursprüngliche (naive) Mengenbegriff nicht von einer Klassifikation ausging, sondern von dem Begriff der Kollektion vorhandener Elemente („iteratives Mengenkonzept“ oder „kumulative Hierarchie der Mengen“). Damit scheint so etwas wie  $(x \in x)$  tatsächlich ausgeschlossen zu sein, was bedeuten würde, daß die Mengenlehre nicht von den gängigen Paradoxien der Selbstbezüglichkeit betroffen ist. So hat Gödel schon in [19] argumentiert. Aber gleichzeitig hat er auch immer wieder darauf hingewiesen, daß die intensionalen Paradoxien, also die Paradoxien der Logik, nach wie vor (1964) nicht gelöst sind, womit sich die ganze Problematik einfach nur auf die höhere Logik verlagert.

Anfang der dreißiger Jahre des letzten Jahrhunderts hat Tarski einen Vorschlag gemacht, wie die semantischen Paradoxien zu umgehen sind, der weitgehende Zustimmung gefunden hat.

Wenn wir [.] die [wesentlichen] Voraussetzungen analysieren, die zur Antinomie des Lügners führen, so stellen wir das Folgende fest:

(I) Wir haben unausgesprochen vorausgesetzt, daß die Sprache, in der die Antinomie konstruiert worden ist, neben deren Ausdrücken auch die Namen derselben enthält, ferner semantische Terme wie ›wahr‹ in bezug auf Aussagen dieser Sprache. Wir haben auch vorausgesetzt, daß alle Aussagen, die den angemessenen Gebrauch dieses Terms festlegen, in der Sprache behauptet werden können. Eine Sprache mit diesen Eigenschaften wird als ›semantisch geschlossen‹ bezeichnet.

(II) Wir haben vorausgesetzt, daß in dieser Sprache die üblichen Gesetze der Logik gelten.

[.] Da jede Sprache, die diese beiden Voraussetzungen erfüllt, inkonsistent ist, müssen wir wenigstens eine von ihnen verwerfen.<sup>67</sup>

<sup>66</sup> [47], S. 75.

<sup>67</sup> [57], S. 150 f.



Bekanntlich entschließt sich Tarski, nur die Möglichkeit einer Preisgabe der Annahme (I) zu erwägen, d.h. Preisgabe semantisch geschlossener Sprachen.

Damit haben wir die wichtigsten Strategien zur Vermeidung von Paradoxien erwähnt: die einfache Typentheorie und das iterative Mengenkonzept, sowie die Sprachstufen als Strategie zur Vermeidung von semantischen Paradoxien.

Strategien, die klassische Logik einzuschränken, haben es schwerer gehabt sich durchzusetzen. Schlimmer noch, es hat immer wieder Stimmen, auch von logisch nicht ganz Unbedarften gegeben, die die Möglichkeit einer solchen Einschränkung, oder zumindest ihren Sinn, bezweifelt haben. Der Grundtenor ist von Tarski vorgegeben worden.<sup>68</sup>

Es wäre überflüssig, hier die Konsequenzen der Verwerfung von Voraussetzung (II) hervorzuheben, das heißt, unsere Logik (vorausgesetzt, daß das möglich ist) in ihren elementarerem und fundamentaleren Teilen zu verändern.<sup>69</sup>

Warum es überflüssig wäre, verrät Tarski nicht, und wahrscheinlich tut er gut daran. So wie die Behauptung dasteht, ist sie ein wohlfeiles und höchst effektives Mittel, naive Gemüter zu beeindrucken. Entsprechend ist Tarskis Auffassung endlos wiederholt und variiert worden, bis hin zur phantasievollen Ausschmückung:

Wir stehen also vor der Alternative: Preisgabe der semantischen Geschlossenheit der Sprache oder Ersetzung logischer Grundregeln durch neue. Das letztere würde eine wissenschaftliche Katastrophe darstellen; denn es ist nicht einzusehen, wie bei Verwerfung jener einfachen logischen Prinzipien und Deduktionsregeln, die bei der Konstruktion der semantischen Antinomien verwendet wurden, auch nur ein geringer Bestandteil des als „Wissenschaft“ bezeichneten Forschungsbetriebes aufrechterhalten werden könnte. Es bleibt daher nur der erste Ausweg[.]<sup>70</sup>

<sup>68</sup> Das ist umso bemerkenswerter, als Tarski als Koautor des in Fußnote 14 erwähnten Aufsatzes erscheint, in dem eine unendlich-wertige Aussagenlogik vorgestellt wird, die sehr wohl eine uneingeschränkte Abstraktion für die Aussagenlogik zuläßt (vgl. [54]) — und bei einer geeigneten Erweiterung auch für die Prädikatenlogik (vgl. [62]).

<sup>69</sup> [57], S. 151.

<sup>70</sup> [55], S. 38 f.

Für Schulmeisternaturen sind die Paradoxien offensichtlich ein willkommenen Anlaß, der Sprache Einschränkungen aufzuerlegen und zu bestimmen, was gesagt werden darf bzw. sich sinnvoll sagen läßt. Carnap war da kein Einzelfall mit seinem Verbot von Sphärenvermengungen. Auch runde vierzig Jahre später lebt der Traum vom Verbiehen von Ausdrücken fort, wie das nächste Zitat zeigt:

[[D]as Auftreten der von Russell gefundenen Antinomie zeigt, daß ein sinnloser Ausdruck verwendet worden ist und diese Verwendung keineswegs harmlos war. Damit wird man die Frage stellen dürfen, weshalb wir denn überhaupt sinnlose Ausdrücke zulassen sollen, statt Freges Forderungen an die korrekte Bildung von Ausdrücken zu verschärfen und sinnlose Ausdrücke ein für allemal zu verbieten.<sup>71</sup>

Warum sollten wir sinnlose Ausdrücke zulassen? Gegenfrage: Wer bestimmt, was sinnlose Ausdrücke sind? Auch der von Russell benutzte Term kann in einer entsprechenden Mengenlehre sehr sinnvoll eingesetzt werden: nämlich um zu zeigen, daß durch ihn keine Menge bezeichnet wird. Auch wenn ein Kriterium der Sinnlosigkeit es schaffen sollte, die Klippe der Ideologie zu umschiffen, so bleibt doch immer noch die Frage, ob man sinnlose Ausdrücke ausschließen kann, ohne möglicherweise auch sinnvolle Ausdrücke auszuschließen. Ausdrücke sind nicht mit Etiketten versehen, die sie als sinnlos oder sinnvoll ausweisen.

Frege war sich der Schwierigkeit bewußt, wie folgendes Zitat aus dem Nachlaß nahelegt:

Es ist schwer, einen allgemein üblichen Ausdruck zu vermeiden, wenn man die Fehler, die daraus entspringen können, noch nicht kennengelernt hat. Es ist gar schwer, vielleicht unmöglich, jeden Ausdruck, den uns die Sprache darbietet, auf seine logische Unverfänglichkeit zu prüfen. So besteht denn ein grosser Teil der Arbeit des Philosophen — oder sollte wenigstens bestehen — in einem Kampfe mit der Sprache. Aber vielleicht sind nur wenige sich dieser Notwendigkeit bewusst.<sup>72</sup>

Ich fürchte, Freges Kampf mit der Sprache ist, insbesondere von denen, die sich so eifertig auf ihn berufen, nicht als permanente Auseinander-

<sup>71</sup> [58], S. 42. Und der Rohrstock schwingt immer mit.

<sup>72</sup> [30], S. 289.

setzung mit der Sprache verstanden worden, sondern als Einzwängen, ein Knebeln der Sprache. Es wird sich von selbst verstehen, daß ich diesem Beispiel nicht zu folgen beabsichtige. Was ich anstrebe, ist eine Einschränkung der Logik zugunsten einer uneingeschränkten Abstraktion, ohne daß kontradiktorische Widersprüche, also Widersprüche der Gestalt  $A$  und nicht- $A$ , beweisbar werden. Auf diese Weise will ich versuchen, einen Bereich für die Vernunft zu erschließen, den die klassische (Verstandes-) Logik mit ihren rigiden Prinzipien verbaut.

## 5. Eine Einschränkung der klassischen Logik

Um das Privileg einer uneingeschränkten Abstraktion (Begriffsbildung) in Anspruch nehmen zu können, muß ich einen Weg finden, die klassische Logik so einzuschränken, daß die Paradoxien der uneingeschränkten Abstraktion keinen Schaden anrichten können.

Es liegt nahe, beim Satz vom ausgeschlossenen Dritten (*Tertium non datur*) anzusetzen. Hegel hat ihn als einen Satz des bestimmten Verstandes charakterisiert:

Der Satz des ausgeschlossenen Dritten ist der Satz des bestimmten Verstandes, der den Widerspruch von sich abhalten will, und indem er dies thut, denselben begeht.<sup>73</sup>

Von daher könnte man erwarten, daß er ihn, als ungeeignet für das spekulative Denken, aus der spekulativen Logik ausschließt. Dagegen scheint jedoch folgende Bemerkung zu stehen:

In der spekulativen Logik ist die bloße Verstandes-Logik enthalten und kann aus jener sogleich gemacht werden; es bedarf dazu nichts, als daraus das Dialektische und Vernünftige wegzulassen[.]<sup>74</sup>

Das kann als verwirrend empfunden werden. Wenn die Verstandeslogik in der spekulativen enthalten ist, enthält dann die spekulative Logik auch den Satz vom ausgeschlossenen Dritten? Das scheint mir keinen rechten Sinn zu ergeben.

Ich will Hegels Worten nicht zuviel Bedeutung beilegen, aber tatsächlich zeigt sich hier ein Problem allgemeiner Art. So ist zum Beispiel die

<sup>73</sup> [25], S. 276. Vgl. oben S 94: Hegel über den Dogmatismus in der Metaphysik.

<sup>74</sup> [25], S. 195 f.

intuitionistische Logik auf der einen Seite ein Teilsystem der klassischen, auf der anderen läßt sich aber die klassische Logik innerhalb der intuitionistischen interpretieren, d.h. man kann eine Vorschrift für eine Übersetzung von Formeln derart angeben, daß die Übersetzung einer beweisbaren Formel der klassischen Logik wieder in der intuitionistischen Logik beweisbar ist. Auf diese Weise ergibt sich ein logischer Raum sozusagen *unterhalb* der klassischen Logik. Damit ist die intuitionistische Logik in dem Sinne *reicher* als die klassische Logik, als sie über eine stärkere Aussagekraft verfügt, die es zum Beispiel erlaubt, die Church-Turing-These auf der Objektebene als Axiom widerspruchsfrei zu formulieren.

Um die uneingeschränkte Abstraktion zulassen zu können, ohne daß jede Formel beweisbar wird, muß von den Axiomen der klassischen Aussagenlogik etwas weggelassen werden. In diesem Sinne geht das, was ich weiter unten als „dialektische Logik“ bezeichnen werde, aus der Verstandeslogik, d.h. der klassischen Logik (mit dem Satz vom ausgeschlossenen Dritten), hervor, indem alle Sätze des Verstandes weggelassen werden, die eine strikte Zweiwertigkeit postulieren. Die Frage ist natürlich, was alles zu diesen Sätzen des Verstandes gehört, und insbesondere, ob es reicht, den Satz vom ausgeschlossenen Dritten aus der Liste der Axiome zu streichen.

Nach dem Gesagten kann man die Sache aber auch so sehen, daß hinsichtlich dessen, was als *wahr* erachtet wird, in der klassischen Logik alles *Vernünftige* ausgeschlossen wird, insbesondere natürlich die Widersprüche der unbeschränkten Begriffsbildung. In diesem Sinne wäre dann die klassische Logik die *ärmere* Logik. Im folgenden werde ich aus axiomatischer Sicht sagen, daß gewisse Axiome oder Schlußregeln aufgegeben werden müssen, um in den Genuß der uneingeschränkten Abstraktion zu gelangen; was man auch einfach so formulieren kann: Weniger ist mehr bzw. weniger Axiome erlauben (potentiell) mehr Differenzierung, und damit Inhalt.

Eine Einschränkung der klassischen Logik kann auf verschiedene Weisen unternommen werden, wie beispielsweise durch ein Hinzufügen von nicht-klassischen *Wahrheitswerten* zu den beiden klassischen Werten „wahr“ und „falsch“, oder eine Preisgabe klassischer Axiome. Derartige Einschränkungen taugen nicht immer für eine widerspruchsfreie Zulassung der uneingeschränkten Abstraktion. Entsprechend gibt es Autoren, die glauben, daß es nicht ausreicht, das *Tertium non datur* aufzugeben,

um zu einer Logik zu gelangen, die eine widerspruchsfreie Zulassung der uneingeschränkten Abstraktion erlaubt. Als Begründung wird mitunter auf die intuitionistische Logik verwiesen, in der das *Tertium non datur* nicht gilt, die uneingeschränkte Abstraktion aber trotzdem zur Inkonsistenz führt und dazu, daß jede Formel beweisbar ist. Das Problem dabei ist, was man unter Preisgabe des *Tertium non datur* versteht. Es reicht nicht, irgendein Axiomensystem der klassischen Logik herzunehmen, in welchem die Formel  $A \vee \neg A$  als Axiom auftritt, und das dann aus der Liste zu streichen.<sup>75</sup> Etwas pointierter könnte man sagen: Die Preisgabe eines logischen Gesetzes ist Glücksache, und es gibt eine Menge Pechvögel.

Die Auffassung, daß es praktisch unmöglich sei, oder zumindest sehr schwierig, die Logik so einzuschränken, daß uneingeschränkte Abstraktion zugelassen werden kann, ist umso erstaunlicher, als es eine lächerlich einfache Möglichkeit gibt, zumindest aus technischer Sicht, wenn man nur über ein bißchen Hintergrundwissen zu Gentzens Sequenzenkalkül verfügt. Die Grundeinsicht dieser Strategie lautet folgendermaßen:

*Schnitteliminierung ohne Zusammenziehung erfordert keinen Rückgriff auf die Komplexität der Schnittformel.*

Anders gesagt, die Schnittregel im sogenannten *Idealkalkül*<sup>76</sup> ohne Zusammenziehungen läßt sich leicht als überflüssig nachweisen, womit man einen Konsistenzbeweis erhält, und zwar mit einer einfachen Induktion. Diese Tatsache stellt auch eine Verbindung zur theoretischen Informatik her: Girards Ansatz zu einer *logic of polytime* basiert auf eben dieser Einsicht in die Komplexität der Schnitteliminierbarkeit der Logik ohne Zusammenziehungen.<sup>77</sup>

Es wird jetzt nötig, eine formale Sprache und einen formalen Beweisbegriff als Bezugsrahmen für die weiteren Überlegungen festzulegen.<sup>78</sup>

Als Grundzeichen haben wir

1. freie Variablen, mitgeteilt durch  $a, b, a_1, \dots$  ;
2. gebundene Variablen, mitgeteilt durch  $x, y, z, x_1, \dots$  ;

<sup>75</sup> Vgl. meine Bemerkungen in [37], S. 1558 ff, 113.16.

<sup>76</sup> Was die Geschichte der Bezeichnung „Idealkalkül“ angeht, so verweisen Fraenkel et al. [13], S. 155, auf den Artikel von Hermes und Scholz [31].

<sup>77</sup> Vgl. meine Bemerkungen in [38], S. 671.

<sup>78</sup> Notgedrungen sind die folgenden Angaben knapp. Im wesentlichen folge ich der Terminologie von [37], § 12.

3. als Hilfszeichen runde Klammern;
4. als logische Zeichen die Konstanten  $\subseteq$  (Inklusion),  $\in$  (Element) und den  $\lambda$ -Operator.

Wie üblich wird ein Term bzw. eine Formel, die keine freien Variablen enthält, als *geschlossen* bezeichnet.

Was die logischen Konstanten angeht, so ist dies ein sparsames Fundament. Abgesehen davon, daß dadurch beweistheoretische Untersuchungen erleichtert werden, soll auch der Ausdrucksfähigkeit der uneingeschränkten Abstraktion Raum zur Entfaltung gelassen werden.

Terme und Formeln werden nun auf übliche Weise simultan definiert:

1. Jede freie Variable ist ein Term.
2. Sind  $s$  und  $t$  Terme, so sind  $(s \subseteq t)$  und  $s \in t$  Formeln.
3. Wenn  $t$  ein Term ist,  $\mathfrak{F}[t]$  eine Formel und  $x$  eine gebundene Variable, die nicht in  $\mathfrak{F}$  vorkommt, dann ist  $\lambda x \mathfrak{F}[x]$  ein Term.

Ich werde  $r, s, t$  als Mitteilungszeichen für Terme gebrauchen,  $A, B, C$  für Formeln, große griechische Buchstaben  $\Gamma, \Pi$  und  $\Theta$  für Folgen von Formeln, und  $\Gamma[A]$ , zum Beispiel, um *ein bestimmtes* Vorkommen einer Formel  $A$  in einer Folge von Formeln anzuzeigen, und  $\Gamma[\ ]$ , um anzuzeigen, daß ein bestimmtes Vorkommen einer Formel gestrichen worden ist. *Nennformen* sind wie in [51], S. 1, definiert; sie werden durch  $\mathfrak{F}, \mathfrak{C}, \mathfrak{F}_1$ , usw. mitgeteilt. Alle Mitteilungszeichen können auch indiziert gebraucht werden. Eine Nennform  $\mathfrak{F}$  wird als *erster Stufe* bezeichnet, wenn das Nennzeichen  $*_1$  nirgends in  $\mathfrak{F}$  auf der linken Seite von  $\in$  vorkommt, d.h. die Folge  $*_1 \in$  kommt nirgends in  $\mathfrak{F}$  vor.

Der Einfachheit halber werde ich  $\lambda A$  für  $\lambda x A$  schreiben, wenn  $x$  nicht in  $A$  vorkommt.  $\lambda A$  wird dann auch als *leere Abstraktion* bezeichnet.

Wenn  $\Gamma$  eine Folge von Formeln ist und  $A$  eine Formel, dann heißt  $\Gamma \Rightarrow A$  eine *Sequenz*. Ich schreibe  $\Gamma, [A]^n, \Pi \Rightarrow C$  für  $\Gamma, A, \dots, A, \Pi \Rightarrow C$ , wenn  $A, \dots, A$  eine Folge von  $n$  Vorkommnissen der Formel  $A$  ist.

Die Schlußregeln in Sequenzenform für  $\in$  sind:<sup>79</sup>

$$\text{links : } \frac{\mathfrak{F}[t], \Gamma \Rightarrow C}{t \in \lambda x \mathfrak{F}[x], \Gamma \Rightarrow C} \qquad \text{rechts : } \frac{\Gamma \Rightarrow \mathfrak{F}[t]}{\Gamma \Rightarrow t \in \lambda x \mathfrak{F}[x]}$$

und die für  $\subseteq$ :

<sup>79</sup>  $\in$ -Schlüsse werden oft auch „ $\lambda$ -Schlüsse“ genannt.

$$\text{links : } \frac{\Gamma \Rightarrow r \in s \quad r \in t, \Pi \Rightarrow C}{s \subseteq t, \Gamma, \Pi \Rightarrow C} \quad \text{rechts : } \frac{a \in s, \Gamma \Rightarrow a \in t}{\Gamma \Rightarrow s \subseteq t}$$

mit der Variablenbedingung: die freie Variable  $a$  (*Eigenvariable*) darf nicht in der Untersequenz auftreten.

Schlüsse gemäß dieser Regeln werden als *logische Zeichen Grundschlüsse* bezeichnet.

Ein *Schnitt* ist ein Schluß nach dem Schema

$$\frac{\Gamma \Rightarrow C \quad \Pi[C] \Rightarrow B}{\Gamma, \Pi[] \Rightarrow B},$$

wobei  $C$  als *Schnittformel* bezeichnet wird.

*Strukturschlußregeln* sind Verdünnungen, Vertauschungen und Zusammenziehungen:

$$\frac{\Gamma \Rightarrow C}{A, \Gamma \Rightarrow C} \quad \text{und} \quad \frac{\Gamma, A, B, \Pi \Rightarrow C}{\Gamma, B, A, \Pi \Rightarrow C} \quad \text{und} \quad \frac{A, A, \Gamma \Rightarrow C}{A, \Gamma \Rightarrow C}.$$

Die formale Theorie, die aus den vorgenannten Regeln mit den üblichen Grundsequenzen ( $A \Rightarrow A$ ) besteht, heißt **LJ $_{\lambda}$** . **LJ $_{\lambda}$**  ist die intuitionistische Version des sogenannten *Idealkalküls* und erlaubt, genauso wie der volle klassische Kalkül, die Beweisbarkeit jeder beliebigen Formel erlaubt (*Trivialität*).

Zum Beweis führe ich die Implikation als eine definierte logische Konstante ein:

$$(5.1) \quad A \rightarrow B := \lambda A \subseteq \lambda B.$$

Schlüsse gemäß folgender Schemata können nun als **LJ $_{\lambda}$** -herleitbar nachgewiesen werden:

$$\text{links : } \frac{\Gamma \Rightarrow A \quad B, \Pi \Rightarrow C}{A \rightarrow B, \Gamma, \Pi \Rightarrow C} \quad \text{rechts : } \frac{A, \Gamma \Rightarrow B}{\Gamma \Rightarrow A \rightarrow B}.$$

Im Fall der linken Regel geht das folgendermaßen:

$$\frac{\frac{\Gamma \Rightarrow A}{\Gamma \Rightarrow r \in \lambda A} \quad \frac{B, \Pi \Rightarrow C}{r \in \lambda B, \Pi \Rightarrow C}}{\lambda A \subseteq \lambda B, \Gamma, \Pi \Rightarrow C}.$$

Damit läßt sich nun ein äußerst sparsamer Beweis der Paradoxie von Curry in **LJ $_{\lambda}$**  führen. Es sei  $C := \lambda x(x \in x \rightarrow A)$ , wobei  $A$  eine beliebige Formel sein darf:

$$\begin{array}{c}
 \frac{C \in C \Rightarrow C \in C \quad A \Rightarrow A}{C \in C \rightarrow A, C \in C \Rightarrow A} \rightarrow \text{-links} \\
 \frac{\quad}{C \in C, C \in C \Rightarrow A} \in \text{-links} \\
 \frac{\quad}{C \in C \Rightarrow A} \text{Zusammenziehung}
 \end{array}$$

und weiter:

$$\begin{array}{c}
 \frac{C \in C \Rightarrow A}{\Rightarrow C \in C \rightarrow A} \rightarrow \text{-rechts} \\
 \frac{\quad}{\Rightarrow C \in C} \in \text{-rechts} .
 \end{array}$$

Durch einen Schnitt erhält man  $\Rightarrow A$ , wobei  $A$  nach Definition von  $C$  eine beliebige Formel sein kann. Damit haben wir die Behauptung bewiesen, daß  $\mathbf{LJ}_\lambda$  trivial ist.

Bei einem derart einfachen Beweis bleiben nicht viele Möglichkeiten, die Logik einzuschränken: außer der Abstraktion sind nur Implikation, Zusammenziehung und Schnitt beteiligt.

Die Möglichkeit einer Einschränkung der Abstraktion wird von den modernen axiomatischen Mengenlehren genutzt. Ein solches Vorgehen verträgt sich aber nicht mit der Idee einer Begründung der Metaphysik als Logik.<sup>80</sup> Die Implikation ist eine logische Konstante, hier noch dazu auf der Grundlage von  $\subseteq$  *definiert*, so daß eine Strategie, die hier ansetzen würde, eine ganze Reihe von weiteren Einschränkungen nach sich ziehen würde. Zusammenziehung und Schnitt sind sogenannte Strukturschlußregeln. Es liegt in der Natur der Schnittregel, daß sie von einem Widerspruch zur Trivialität führt. Ohne eine Form der Schnittregel gibt es also keine Trivialität. Das machen sich die sogenannten *parakonsistenten* Logiken zunutze: Widersprüche können geduldet werden, weil sie nicht zur Trivialität führen.

Offensichtlich spielt die Zusammenziehung in dem obigen Beweis eine wichtige Rolle. So stellt sich die Frage, ob eine Preisgabe der Zusammenziehungsregel ausreichen würde, um die uneingeschränkte Abstraktion widerspruchsfrei zulassen zu können? Tatsächlich folgt aus der oben genannten Grundeinsicht, daß Schnitteliminierung ohne Zusammenziehungen keinen Rückgriff auf die Komplexität der Schnittformel erfordert, ziemlich unmittelbar: Ohne Zusammenziehungen ist die Schnittregel auch

<sup>80</sup> [1] scheint eine derartige Linie zu verfolgen, aber es ist mir nicht gelungen, in seinem Werk mehr als nur eine idiosynkratische Formulierung der Ergebnisse der Zermelo-Fraenkel-Mengenlehre zu finden.



bei uneingeschränkter Abstraktion eliminierbar. Die formale Theorie, die aus  $\mathbf{LJ}_\lambda$  entsteht, wenn als Strukturschlüsse keine Zusammenziehungen mehr, sondern nur noch Verdünnungen und Vertauschungen zugelassen werden, heißt  $\mathbf{LJD}_\lambda$ .<sup>81</sup> Dann ist zu zeigen: Was immer mithilfe von Schnitten in  $\mathbf{LJD}_\lambda$  beweisbar ist, kann auch ohne Schnitt bewiesen werden.

Die Bedeutung der Schnitteliminierbarkeit liegt in der Konsequenz für die Widerspruchsfreiheit. Man kann sich leicht überzeugen, daß die leere Sequenz  $\Rightarrow$  nicht ohne Schnitt  $\mathbf{LJD}_\lambda$ -beweisbar sein kann: Sie ist weder eine Anfangssequenz noch die Untersequenz eines Grundschlusses. Falls ein Widerspruch herleitbar wäre, so wäre mit Schnitt auch die leere Sequenz herleitbar.<sup>82</sup>

$$\frac{\Rightarrow A \quad \neg A \Rightarrow}{\Rightarrow}$$

Auf den eigentlichen Beweis der Schnitteliminierbarkeit kann hier nicht eingegangen werden, da er den Rahmen sprengen würde, aber der wesentliche Punkt soll zumindest angedeutet werden.

Klassischerweise erfordert ein Schnitteliminierungsbeweis eine Induktion nach der Länge der Schnittformel. Uneingeschränkte Abstraktion steht dem im Weg: die Länge einer Formel  $\mathfrak{F}[t]$  kann länger sein als die von  $t \in \lambda x \mathfrak{F}[x]$ , obwohl es auf den ersten Blick nicht so aussehen mag. Man nehme z.B. Russells Klasse:  $\mathfrak{F}[t]$  entspricht dann der Formel  $\lambda x (x \notin x) \notin \lambda x (x \notin x)$  und  $t \in \lambda x \mathfrak{F}[x]$  der Formel  $\lambda x (x \notin x) \in \lambda x (x \notin x)$ . Der relevante  $\epsilon$ -Schluß hat dann beispielsweise folgende Gestalt:

$$\frac{\lambda x (x \notin x) \notin \lambda x (x \notin x), \Gamma \Rightarrow C}{\lambda x (x \notin x) \in \lambda x (x \notin x), \Gamma \Rightarrow C}$$

Deshalb darf die übliche Strategie des „Nach-oben-Verschiebens-des-Schnittes“ nicht von der Verkürzung der Schnittformel abhängen. Das

<sup>81</sup> Natürlich gibt es auch äquivalente Formulierungen in anderen Kalkülarten, beispielsweise im sogenannten Hilbert-Frege-Stil (axiomatisch), auf die ich hier aber nicht eingehen will. Eine Behandlung der dazu benötigten Techniken findet sich in [37], §§ 16 und 17.

<sup>82</sup> Da in der Formulierung von  $\mathbf{LJD}_\lambda$  ein Negationszeichen fehlt, sieht das in der Praxis etwas komplizierter aus: Anstelle der leeren Sequenz muß man eine Sequenz minimaler Länge hernehmen, wie etwa  $\Rightarrow a \in a$ , die nachweislich nicht ohne Schnitt beweisbar sein kann.

ist aber durch die Preisgabe der Zusammenziehungen erreicht worden: Eine Induktion nach der Anzahl der logische Zeichen Schlüsse reicht aus.

Ich will versuchen, das etwas für den Fall zu verdeutlichen, daß die Schnittformel  $C$  die Form  $t \in \lambda x \mathfrak{F}[x]$  hat. In diesem Fall sieht der Schnitt folgendermaßen aus:

$$\frac{\frac{\Gamma \Rightarrow \mathfrak{F}[t]}{\Gamma \Rightarrow t \in \lambda x \mathfrak{F}[x]} \quad \frac{\mathfrak{F}[t], \Pi \Rightarrow C}{t \in \lambda x \mathfrak{F}[x], \Pi \Rightarrow C}}{\Gamma, \Pi \Rightarrow C} \text{Schnitt.}$$

Nun läßt sich ein Schnitt schon mit den Obersequenzen ausführen, wodurch wir eine gewisse Reduktion erreicht haben:

$$\frac{\begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \Gamma \Rightarrow \mathfrak{F}[t] \end{array} \quad \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \mathfrak{F}[t], \Pi \Rightarrow C \end{array}}{\Gamma, \Pi \Rightarrow C} \text{Schnitt.}$$

Was hier reduziert worden ist, ist die gesamte Länge der Herleitung und nicht notwendigerweise der Schnittformel; letztere kann länger geworden sein — das kommt auf die Form von  $\mathfrak{F}$  an.

Diese Strategie reicht jedoch nicht mehr aus, wenn Zusammenziehungen erlaubt sind. Dann kann folgende Situation entstehen:

$$\frac{\frac{\Gamma \Rightarrow \mathfrak{F}[t]}{\Gamma \Rightarrow t \in \lambda x \mathfrak{F}[x]} \quad \frac{\frac{\mathfrak{F}[t], t \in \lambda x \mathfrak{F}[x], \Pi \Rightarrow C}{t \in \lambda x \mathfrak{F}[x], t \in \lambda x \mathfrak{F}[x], \Pi \Rightarrow C}}{t \in \lambda x \mathfrak{F}[x], \Pi \Rightarrow C} \text{Zusammenziehung}}{\Gamma, \Pi \Rightarrow C} \text{Schnitt.}$$

Um einen derartigen Schnitt zu reduzieren, wird üblicherweise folgende Strategie benutzt:

$$\frac{\begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \Gamma \Rightarrow \mathfrak{F}[t] \end{array} \quad \frac{\frac{\Gamma \Rightarrow \mathfrak{F}[t]}{\Gamma \Rightarrow t \in \lambda x \mathfrak{F}[x]} \quad \begin{array}{c} \text{---} \\ \text{---} \\ \text{---} \\ \mathfrak{F}[t], t \in \lambda x \mathfrak{F}[x], \Pi \Rightarrow C \end{array}}{\mathfrak{F}[t], t \in \lambda x \mathfrak{F}[x], \Pi \Rightarrow C} \text{Schnitt}}{\mathfrak{F}[t], \Gamma, \Pi \Rightarrow C} \text{Schnitt}}{\Gamma, \Gamma, \Pi \Rightarrow C} \text{Schnitt}}{\text{evtl. Zusammenziehungen}}{\Gamma, \Pi \Rightarrow C}$$

Hier hat, wie vorher, der erste Schnitt eine geringere Komplexität der Herleitung. Für den zweiten Schnitt muß jedoch ein Teil der Herleitung ( $\Gamma \Rightarrow \mathfrak{F}[t]$ ) ein zweites Mal benutzt werden, und so vergrößert sich die Komplexität des zweiten Schnitts (und der gesamten Herleitung). Für den zweiten Schnitt bräuchten wir daher ein Komplexitätsmaß, das von dem der Herleitung unabhängig ist, wie eben die Länge der Schnittformel. Aber letztere ist im vorliegenden Fall nicht unbedingt reduziert worden, und das macht die ganze Aktion wertlos, was uns allerdings nicht überraschen sollte, da wir ja wissen, daß die Schnitteliminierung im vorliegenden Fall unmöglich ist.

Mit anderen Worten, wenn es keine Zusammenziehungsregeln (auch nicht versteckt, d.h. integriert in die logische Zeichen Schlußregeln!) gibt, kann die Schnitteliminierbarkeit (bei entsprechender Form der Grundschlußregeln) ohne Rückgriff auf die Länge der zur Diskussion stehenden Schnittformel bewiesen werden. Daran liegt es, daß Logik ohne Zusammenziehungen so sicher vor allen Antinomien ist, die durch uneingeschränkte Abstraktion erzeugt werden.<sup>83</sup>

Das reicht aus, die üblichen (intuitionistischen Versionen der) Schlußregeln aller gängigen logischen Konstanten zu erhalten. Hier ist eine Liste von Definitionen, auf die ich gelegentlich zurückgreifen werde:

$$\begin{aligned} \mathcal{V} &::= \lambda x(x \subseteq x) \quad (\text{Allmenge}), \\ \bigwedge x \mathfrak{F}[x] &::= \mathcal{V} \subseteq \lambda x \mathfrak{F}[x] \quad (\text{Allquantor}), \\ \perp &::= \bigwedge x(\mathcal{V} \subseteq x) \quad (\text{falsum}), \\ \neg A &::= A \rightarrow \perp \quad (\text{Negation}), \\ \emptyset &::= \lambda \perp \quad (\text{Leermenge}), \\ \top &::= \neg \perp \quad (\text{verum}), \\ \bigvee x \mathfrak{F}[x] &::= \bigwedge y(\bigwedge x(\mathfrak{F}[x] \rightarrow \lambda \top \subseteq y) \rightarrow \lambda \top \subseteq y) \quad (\text{Existenz}), \\ s \equiv t &::= \bigwedge y(s \in y \rightarrow t \in y) \quad (\text{Identität}), \\ A \square B &::= \bigwedge x((A \rightarrow (B \rightarrow \lambda \top \subseteq x)) \rightarrow \lambda \top \subseteq x), \quad (\text{a-Konjunktion}), \\ A \wedge B &::= \bigwedge x(\bigwedge y(\lambda A \in y \rightarrow (\lambda B \in y \rightarrow x \in y)) \rightarrow \lambda \top \subseteq x) \\ & \hspace{15em} (\text{m-Konjunktion}), \end{aligned}$$

<sup>83</sup> In Abschnitt 9 werden wir sehen, daß sich das nicht auf die klassischen Beschreibungoperatoren wie  $\varepsilon$ ,  $\mu$  und  $\iota$  ausdehnen läßt.

$$A \vee B := \bigwedge y((A \rightarrow \lambda\top \subseteq y) \wedge (B \rightarrow \lambda\top \subseteq y) \rightarrow \lambda\top \subseteq x)$$

(m-Disjunktion),

$$\{s\} := \lambda x(s \equiv x) \quad (\text{exklusive Einermenge}),$$

$$\{s, t\} := \lambda x(x \equiv s \vee x \equiv t) \quad (\text{exklusive Paarmenge}),$$

$$A \leftrightarrow B := (A \rightarrow B) \wedge (B \rightarrow A) \quad (\text{logische Äquivalenz}),$$

$$s = t := \bigwedge x(x \in s \leftrightarrow x \in t) \quad (\text{Gleichheit}),$$

$$\{s\} := \lambda x(x = s) \quad (\text{inklusive Einermenge}),$$

$$\{s, t\} := \lambda x(x = s \vee x = t) \quad (\text{inklusive Paarmenge}),$$

$$\langle s, t \rangle := \{\{\{s\}, \emptyset\}, \{\{t\}\}\} \quad (\text{inklusives geordnetes Paar}),$$

$$\langle s, t \rangle := \{\{\{s\}, \emptyset\}, \{\{t\}\}\} \quad (\text{exklusives geordnetes Paar}),$$

$$\lambda xy \mathfrak{F}[x, y] := \lambda z \forall x \forall y (z \equiv \langle x, y \rangle \square \mathfrak{F}[x, y]) \quad (\text{zweistellige Abstraktion}),$$

$$\mathbb{C} := \lambda xy (y = \lambda z (z \notin x)) \quad (\text{Komplement}),$$

$$st := \lambda x (\langle t, x \rangle \in s) \quad (\text{Nachbereich eines Arguments}),$$

$$s[t] := \lambda x \bigwedge y (\langle t, y \rangle \in s \rightarrow x \in y) \quad (\text{Anwendung}),$$

$$s \cup t := \lambda x (x \in s \vee x \in t) \quad (\text{m-Vereinigung}),$$

$$0 := \emptyset \quad (\text{Null}),$$

$$s' := s \cup \{s\} \quad (\text{Nachfolger}),$$

$$s \sqcap t := \lambda x (x \in s \square x \in t) \quad (\text{a-Durchschnitt}).$$

Ein Großteil dieser Definitionen wird mehr oder weniger vertraut sein.<sup>84</sup> Eine Unterscheidung zwischen „a“- und „m“-Versionen von Konjunktion usw. empfiehlt sich angesichts der unterschiedlichen Behandlung der Annahmen: a-Konjunktion akkumuliert die Annahmen, m-Konjunktion bildet ihren kleinsten gemeinsamen Nenner, sozusagen. Eine Unterscheidung zwischen *exklusiven* und *inklusive*n Versionen der Einermenge usw. wird durch die Unterscheidung von Gleichheit und Identität motiviert. Besondere Erwähnung verdienen vielleicht noch die Operationen *Nachbereich eines Arguments* und *Anwendung*. *st* gibt den Wertebereich von *s* für

<sup>84</sup> Die Symbole  $\bigwedge$  für den Allquantor und  $\bigvee$  für den Existenzquantor scheinen aus der Mode gekommen zu sein, aber sie haben den Vorteil einer Nähe zu den Symbolen  $\wedge$  und  $\vee$ , eine Nähe, die ich im Hinblick auf eine gewisse Parallelität im Fall der alternativen Konjunktion  $\square$  und des Notwendigkeitsoperators  $\square$ , die weiter unten eingeführt werden, erhalten möchte.

das Argument  $t$ . Für den Fall, daß  $s$  eine Funktion ist, gibt  $st$  die Eimermenge des Funktionswertes für das Argument  $t$  und  $s[[t]]$  den Funktionswert selbst. Ansonsten erhält man, beispielsweise,  $\emptyset[[s]] = \mathcal{V}$  und  $\mathcal{V}[[s]] = \emptyset$ , hingegen  $\emptyset s = \emptyset$  und  $\mathcal{V} s = \mathcal{V}$ . Was diese beiden Operationen jedoch gemeinsam haben, ist, daß sie so hübsch den typentheoretischen Vorstellungen von einer korrekten Syntax der Sprache zuwiderlaufen. Deshalb wird es auch nicht überraschen, daß sie sich für die Erzeugung von Paradoxien der Selbstbezüglichkeit eignen. Das sieht man schön an folgendem Term:

$$\phi := \lambda xy (xx = 0).$$

In einem ersten Schritt läßt sich dafür zeigen:

$$\begin{array}{l} \phi\phi = 0 \Rightarrow \phi\phi = 0 \\ \hline \phi\phi = 0 \Rightarrow \langle \phi, a \rangle \in \phi \\ \hline \phi\phi = 0 \Rightarrow a \in \phi\phi \qquad a \in 0 \Rightarrow \\ \hline a \in \phi\phi \rightarrow a \in 0, \phi\phi = 0 \Rightarrow \\ \hline a \in \phi\phi \leftrightarrow a \in 0, \phi\phi = 0 \Rightarrow \\ \hline \phi\phi = 0, \phi\phi = 0 \Rightarrow \\ \hline \phi\phi = 0 \Rightarrow \text{Zusammenziehung.} \end{array}$$

Damit erhält man weiter:

$$\begin{array}{l} \phi\phi = 0 \Rightarrow \\ \hline \langle \phi, a \rangle \in \phi \Rightarrow a \in 0 \\ \hline a \in \phi\phi \Rightarrow a \in 0 \\ \hline \Rightarrow a \in \phi\phi \rightarrow a \in 0 \qquad \Rightarrow a \in 0 \rightarrow a \in \phi\phi \\ \hline \Rightarrow a \in \phi\phi \leftrightarrow a \in 0 \\ \hline \Rightarrow \phi\phi = 0 \end{array},$$

somit einen Widerspruch.

### 6. Von der Selbstbezüglichkeit zur Fixpunkteigenschaft

Mit den Mitteln, die im vorigen Abschnitt zur Verfügung gestellt worden sind, läßt sich das Thema der Selbstbezüglichkeit zur Fixpunkteigenschaft ausdehnen. In Abschnitt 3 habe ich Russells Antinomie als Beispiel für

Selbstbezüglichkeit angeführt. Aber Selbstbezüglichkeit ist damit gerade erst berührt; das Beispiel der Russellschen Antinomie kann dahingehend verallgemeinert werden, daß es zu jeder einstelligigen Aussageform  $\mathfrak{F}$  einen *Fixpunkt* gibt, d.h. eine Formel derart, daß:  $R_{\mathfrak{F}} := \{x : \mathfrak{F}[x \in x]\}$  und man erhält

$$R_{\mathfrak{F}} \in \{x : \mathfrak{F}[x \in x]\} \leftrightarrow \mathfrak{F}[R_{\mathfrak{F}} \in R_{\mathfrak{F}}],$$

d.h., wenn man  $F_{\mathfrak{F}}$  für  $R_{\mathfrak{F}} \in R_{\mathfrak{F}}$  schreibt:

$$(6.1) \quad F_{\mathfrak{F}} \leftrightarrow \mathfrak{F}[F_{\mathfrak{F}}],$$

ein hübsches Fixpunktergebnis.

Es ist aber noch eine weitere Verallgemeinerung der Fixpunkteigenschaft möglich, und zwar für Terme.<sup>85</sup> Selbstbezüglichkeit kann man jetzt auf der Grundlage des Terms  $\lambda z (\exists t, z \in s)$  herstellen, für den oben die Abkürzung  $st$  eingeführt wurde. Es sei nun  $g := \lambda xy \mathfrak{F}[xx, y]$  und  $f_g := gg$ :<sup>86</sup>

$$\frac{\frac{\mathfrak{F}[gg, a] \leftrightarrow \mathfrak{F}[f_g, a]}{\exists g, a \in \lambda xy \mathfrak{F}[xx, y] \leftrightarrow \mathfrak{F}[f, a]}}{a \in \lambda z (\exists g, z \in g) \leftrightarrow a \in \lambda x \mathfrak{F}[f_g, x]}}{\Rightarrow a \in f_g \leftrightarrow a \in \lambda x \mathfrak{F}[f_g, x]}}{\Rightarrow f_g = \lambda x \mathfrak{F}[f_g, x]}.$$

Das heißt, wir haben eine Fixpunkteigenschaft für Termformen:

$$(6.2) \quad f_g = \lambda x \mathfrak{F}[f_g, x].$$

Als Beispiel sei hier ein Fixpunkt für die Einermenge angeführt:<sup>87</sup>

$$\lambda xy (xx \equiv y) \lambda xy (xx \equiv y).$$

<sup>85</sup> Vgl. [36], S. 382 f.

<sup>86</sup> In [36] habe ich eine Formulierung mithilfe der Anwendungsoperation  $s[t]$  benutzt, die etwas komplizierter ist, im Endeffekt aber auf dasselbe hinausläuft.

<sup>87</sup> Koch in [32], S. 3, bemüht Aczels Anti-Fundierungsaxiom, um die Existenz eines Fixpunktes für die Einermenge, d.h. eines Objekts  $f$ , für das gilt  $\{f\} = f$ , als Rechtfertigung für die Möglichkeit eines Fixpunktes für die Negation heranzuziehen, und meint dann bezüglich der Negation: „Es sollte sich ohne weiteres ein Fixpunkt denken lassen, in dem Argument und Wert, Eingabe und Ausgabe der Operation zusammenfallen.“ Angesichts der hier bewiesenen Fixpunkteigenschaft ist das ein unnötiger Umweg: Die uneingeschränkte Abstraktion liefert Fixpunkte für alle Aussage- und Termformen.

Es sei  $g := \lambda xy (xx \equiv y)$  und  $f := gg$ :

$$\frac{\frac{\frac{a \equiv gg \Rightarrow a \equiv f}{\{g, a\} \in g \Rightarrow a \equiv f}}{a \in \lambda x (\{g, x\} \in g) \Rightarrow a \equiv f}}{a \in gg \Rightarrow a \in \{f\}} \Rightarrow a \in gg \rightarrow a \in \{f\} \quad \Rightarrow a \in \{f\} \rightarrow a \in gg$$

$$\frac{\Rightarrow a \in gg \leftrightarrow a \in \{f\}}{\Rightarrow gg = \{f\}}.$$

Selbstbezüglichkeit kann nun in der Form der Fixpunkteigenschaft die Gestalt einer einfachen geschlossenen Formel gegeben werden, oder eigentlich zwei, je nachdem, welche der beiden Operationen, Bildbereich oder Anwendung, man dazu hernimmt:

$$\bigwedge x \bigvee y (xy = y) \quad \text{und} \quad \bigwedge x \bigvee y (x[[y]] = y).$$

Als ziemlich unmittelbare Konsequenz der Fixpunkteigenschaft erhält man:

$$\neg \bigvee y \bigwedge x (x \neq yx) \quad \text{und} \quad \neg \bigvee y \bigwedge x (x \neq y[[x]]).$$

Das ist ein bemerkenswertes Ergebnis, das ich nicht zuletzt deshalb ausdrücklich erwähne, weil es der klassischen Auffassung so unversöhnlich entgegengesetzt ist. Nach klassischer Auffassung ist beispielsweise die Komplementbildung  $\mathbb{C}$  eine Operation, für die

$$\bigwedge x (x \neq \mathbb{C}[[x]])$$

erfüllt ist.

In diesem Zusammenhang will ich auch auf das hinweisen, was Hao Wang „Gödel’s paradox“ genannt hat.<sup>88</sup> Nach Hao Wang wollte Gödel damit zeigen, daß es keine einfache Lösung für die *intensionalen Paradoxien* gibt. Gödel benutzt eine Funktion  $E$ , die folgendermaßen durch Fallunterscheidung definiert wird:

$$E(x) = \begin{cases} 1, & \text{falls } x = 0; \\ 0, & \text{falls } x \neq 0. \end{cases}$$

<sup>88</sup> Vgl. 8.6.24–8.6.26 auf S. 279 in [61].

Nun behauptet Gödel, man könne unmittelbar sehen, daß  $E(x) \neq x$  für alle  $x$  gilt. In Anbetracht der Fixpunkteigenschaft für  $\mathbf{I}\mathbf{D}_\lambda$  kann das jedoch so nicht zutreffen. Ohne eine Form des *Tertium non datur* oder ein äquivalentes klassisches Gesetz kann es nicht gelingen,  $E(x) \neq x$  für beliebige  $x$  zu beweisen.

Zunächst einmal will ich zeigen, wie man das Resultat  $\bigwedge x(E(x) \neq x)$  auf einfache Weise erhalten kann (ohne jedoch zu behaupten, daß es nicht einfacher ginge):

$$\begin{array}{c} a \neq 0 \Rightarrow E(a) = 0 \quad 0 = a \Rightarrow a = 0 \\ \hline a \neq 0, E(a) = a \Rightarrow a = 0 \\ \hline a \neq 0, a \neq 0, E(a) = a \Rightarrow \\ \hline a = 0, E(a) = a \Rightarrow \quad a \neq 0, E(a) = a \Rightarrow \\ \hline \Rightarrow a = 0 \vee a \neq 0 \quad a = 0 \vee a \neq 0, E(a) = a \Rightarrow \\ \hline E(a) = a \Rightarrow \\ \hline \Rightarrow E(a) \neq a \end{array} .$$

Offensichtlich wird das *Tertium non datur* darin gebraucht und eine Zusammenziehung.

Es mag einfacher gehen, aber sicher nicht ohne eine gewisse Form des *Tertium non datur* oder der Zusammenziehung. Das folgt aus der Definierbarkeit der folgenden Funktion  $nc$  in  $\mathbf{I}\mathbf{D}_\lambda$ , die ich im Prinzip schon in [36], S. 383, betrachtet habe (allerdings mit einer komplizierteren Definition):

$$nc := \lambda xy((x = 0 \wedge y = 1) \vee (x \neq 0 \wedge y = 0)).$$

Dafür erhält man ohne Schwierigkeiten:

$$(6.3) \quad t = 0 \Rightarrow nc[t] = 1, \quad \text{und}$$

$$(6.4) \quad t \neq 0 \Rightarrow nc[t] = 0,$$

d.h. die Definitionsbedingungen von  $E$  sind erfüllt.

Aufgrund der Fixpunkteigenschaft gilt weiterhin:

$$\Rightarrow nc[f] = f .$$

Damit haben wir folgende Situation:

- In  $\mathbf{I}\mathbf{D}_\lambda$  kann eine Funktion  $nc$  definiert werden, welche die Definitionsbedingungen von  $E$  erfüllt;



- zu jeder einstelligen Termform  $\mathfrak{Z}$  gibt es einen Term  $f_{\mathfrak{Z}}$  (Fixpunkt bezüglich  $\mathfrak{Z}$ ) derart, daß  $\mathfrak{Z}[f] = f$   $\dot{\mathbf{I}}\mathbf{D}_{\lambda}$ -beweisbar ist, d.h. insbesondere für  $nc$ :  $\dot{\mathbf{I}}\mathbf{D}_{\lambda} \vdash nc[[f]] = f$ ; d.h.  $\dot{\mathbf{I}}\mathbf{D}_{\lambda} \vdash \neg \bigwedge x(nc[x] \neq x)$ ;
- $\dot{\mathbf{I}}\mathbf{D}_{\lambda}$  ist widerspruchsfrei: die Schnitteliminierbarkeit in  $\dot{\mathbf{I}}\mathbf{D}_{\lambda}$  kann durch einen einfachen Induktionsbeweis (d.h. bis  $\omega$ ) geliefert werden.

Mit anderen Worten,  $E(x) \neq x$  für alle  $x$  kann keine unmittelbare Folgerung der Definition von  $E$  sein.

Was passiert dann aber mit dem Fixpunkt  $f$ ? Aus 6.4 erhält man zuerst

$$\frac{\Rightarrow nc[[f]] = f \quad f \neq 0 \Rightarrow nc[[f]] = 0}{f \neq 0 \Rightarrow f = 0}$$

und dann

$$\frac{\Rightarrow nc[[f]] = f \quad f = 0 \Rightarrow nc[[f]] = 1}{\frac{f = 0 \Rightarrow f = 1 \quad 1 = 0 \Rightarrow \perp}{f = 0, f = 0 \Rightarrow \perp}}$$

Offensichtlich würden Zusammenziehungen hier Unheil anrichten. Man könnte dann nämlich wie folgt schließen:

$$\frac{\frac{f = 0, f = 0 \Rightarrow \perp}{f = 0 \Rightarrow \perp}}{\Rightarrow f \neq 0} \quad \frac{f \neq 0 \Rightarrow f = 0}{\Rightarrow f = 0} \quad \Rightarrow f = 1$$

$$\frac{\Rightarrow f = 0 \quad \Rightarrow f = 1}{\Rightarrow 1 = 0}$$

Hier ist wohl der Punkt, wo die Differenz zwischen klassischer und dialektischer Position am deutlichsten zutage tritt:

- in der klassischen Theorie gibt es eine Abbildung  $g$  derart, daß  $g[x] \neq x$  für alle  $x$  gilt;
- in einer dialektischen (höheren) Logik gibt es *keine* Abbildung  $g$ , für die immer  $g[x] \neq x$  gilt.

Der Punkt ist höchst metaphysisch: Ist es grundsätzlich möglich, die *Welt* (oder das *Universum*) in zwei elementfremde Teile aufzuteilen, deren Vereinigung die gesamte *Welt* ist. Die klassische Position kann dahingehend charakterisiert werden, daß sie diese Frage mit Ja beantwortet (*Tertium non datur*). Die Position, die ich mit meinem Versuch einer Begründung

der Metaphysik durch logische Analyse der Selbstbezüglichkeit verfolge, beantwortet die Frage mit Nein. Aufgrund der Fixpunkteigenschaft gibt es einen Term  $f_{cp}$  derart, daß das Komplement  $\lambda x(x \notin f_{cp})$  von  $f_{cp}$  gleich  $f_{cp}$  ist. Anders gesagt, der Durchschnitt zwischen dem Fixpunkt  $f_{cp}$  und seinem Komplement  $\lambda x(x \notin f_{cp})$  ist gleich  $f_{cp}$  selbst:  $\lambda x(x \notin f_{cp}) \cap f_{cp} = f_{cp}$ . Von diesem Fixpunkt kann man nicht zeigen, daß er leer ist, obwohl man auch nicht zeigen kann, daß etwas unter ihn fällt. Die folgenden beiden Beweisfiguren mögen das vermitteln:

$$\frac{\frac{\frac{\frac{a \in f_{cp} \Rightarrow a \in f_{cp} \quad a \in \emptyset \Rightarrow \perp}{a \in f_{cp} \rightarrow a \in \emptyset, a \in f_{cp} \Rightarrow \perp}}{f_{cp} = \emptyset, a \in f_{cp} \Rightarrow \perp}}{f_{cp} = \emptyset \Rightarrow a \notin f_{cp}}}{\Rightarrow \lambda x(x \notin f_{cp}) = f_{cp} \quad f_{cp} = \emptyset \Rightarrow a \in \lambda x(x \notin f_{cp})}}{\frac{f_{cp} = \emptyset \Rightarrow a \in f_{cp} \quad a \in \emptyset \Rightarrow \perp}{a \in f_{cp} \rightarrow a \in \emptyset, f_{cp} = \emptyset \Rightarrow \perp}}{\frac{f_{cp} = \emptyset, f_{cp} = \emptyset \Rightarrow \perp}{f_{cp} = \emptyset, f_{cp} = \emptyset \Rightarrow \perp}},$$

und:

$$\frac{\frac{\frac{a \in f_{cp} \Rightarrow a \in f_{cp}}{a \in f_{cp}, a \notin f_{cp} \Rightarrow \perp}}{\Rightarrow \lambda x(x \notin f_{cp}) = f_{cp} \quad a \in f_{cp}, a \in \lambda x(x \notin f_{cp}) \Rightarrow \perp}}{a \in f_{cp}, a \in f_{cp} \Rightarrow \perp}.$$

## 7. Fixpunkte I: Definition rekursiver Funktionen

Die Fixpunkteigenschaft legitimiert Definitionspraktiken, die nach herkömmlicher Auffassung logisch unzulässig sind; dazu gehören insbesondere Zirkeldefinitionen. Sie ermöglichen eine Darstellung aller rekursiven Funktionen in der Theorie  $\mathbf{L}^{\mathbf{D}}\lambda$ . Wie das im einzelnen geht, kann in [39], auf den Seiten 101–122, nachgelesen werden. Hier will ich nur auf die verbotene Art der Definition hinweisen.<sup>89</sup>

<sup>89</sup> Die Theorie der partiell-rekursiven Funktionen erlaubt über den Rekursionsatz ein analoges Vorgehen, das sogar mit der klassischen Logik vereinbar ist — allerdings

Als Beispiel für die Definition einer Funktion als Fixpunkt bietet sich folgendes Schema zur Definition der Addition an:

$$add = \lambda x_1 x_2 x_3 ((x_2 \equiv 0 \square x_3 \equiv x_1) \diamond \bigvee y \bigvee z (x_2 \equiv y' \square x_3 \equiv z' \square \langle\langle x_1, y \rangle\rangle, z \rangle \in add)).$$

Das ist nur als Schema gedacht, um einen Eindruck zu vermitteln. Logische Konstanten wie  $\equiv$  und  $'$  müssen eventuell variiert werden, um den spezifischen Gegebenheiten eines typenfreien Systems Rechnung zu tragen.

Damit lassen sich dann leicht relevante Eigenschaften der folgenden Art zeigen:

$$\begin{aligned} &\Rightarrow \langle\langle s, 0 \rangle\rangle, s \rangle \in add, \\ &\Rightarrow \langle\langle s, t' \rangle\rangle, r' \rangle \in add, \\ &\langle\langle s, 0 \rangle\rangle, t \rangle \in add \Rightarrow t \equiv s, \\ &\bigwedge x (\langle\langle s, n \rangle\rangle, x \rangle \in add \rightarrow x \equiv p), \langle\langle s, n' \rangle\rangle, t \rangle \in add \Rightarrow t \equiv p'. \end{aligned}$$

Dann läßt sich die Addition in vertrauter Form als  $s + t \equiv add[s, t]$  einführen, und ihre Rekursionsgleichungen

$$\begin{aligned} &\Rightarrow m + 0 = s, \\ &\Rightarrow m + n' = (m + n)', \end{aligned}$$

können ziffernweise durch metatheoretische Induktion in  $\mathbf{LID}_\lambda^Z$  bewiesen werden.

Tatsächlich lassen sich auf diese Weise alle primitiv-rekursiven Funktionen ziffernweise repräsentieren, und darüber hinaus auch alle k-rekursiven Funktionen.

Um die allgemein-rekursiven Funktionen ziffernweise repräsentieren zu können, wird noch ein Suchoperator benötigt, der die kleinste Nullstelle einer Funktion  $f$  liefert, sofern eine solche existiert, und  $f$  für alle kleineren Argumente definiert ist, d.h. Werte in den natürlichen Zahlen besitzt. Dazu wird die Gesamtheit der natürlichen Zahlen nach folgendem Schema als Fixpunkt eingeführt:

nur um den Preis der Zulassung partieller, d.h. nicht überall definierter, Funktionen. Es ist mir nicht klar, wie weit diese Analogie reicht.

$$\mathbf{N} = \lambda x (x \equiv 0 \diamond \bigvee y (y \in \mathbf{N} \square x \equiv y')).$$

Das liefert zwar keine vollständige Induktion, aber genug, um die relevanten Eigenschaften eines solchen Suchoperators beweisen zu können.

Auf diese Weise können alle rekursiven Funktionen in  $\mathbf{L}^{\mathbf{D}}\lambda$  ziffernweise repräsentiert werden, und damit erhält man auch die Unentscheidbarkeit von  $\mathbf{L}^{\mathbf{D}}\lambda$ . Um die Totalität der primitiv-rekursiven Funktionen zu zeigen, braucht man aber schon eine Form der vollständigen Induktion, von der man aufgrund einfacher ordinaler Überlegungen sagen kann,<sup>90</sup> daß sie in  $\mathbf{L}^{\mathbf{D}}\lambda$  nicht zu haben ist. Das wird uns weiter unten beschäftigen.

In der theoretischen Informatik, und allgemein in den mathematischen Disziplinen, ist man weniger dogmatisch eingestellt als in der Philosophie, und so mag es nicht weiter verwundern, daß die Logik, die ich hier  $\mathbf{L}^{\mathbf{D}}\lambda$  genannt habe, heute in der theoretischen Informatik ein beliebtes Forschungsobjekt ist. Ein Grund, weshalb diese Logik für die theoretische Informatik eine besondere Rolle spielt, besteht darin, daß alle in ihr als total beweisbaren Funktionen in polynomialer Zeit berechenbar sind. Das mag für Metaphysiker nicht sonderlich interessant sein, aber es wirft noch einmal ein interessantes Licht auf den Versuch, hier eine „wissenschaftliche Katastrophe“ herbeizureden.

## 8. Fixpunkte II: die Unverträglichkeit der Extensionalität

Die Fixpunkteigenschaft erlaubt eine Darstellung der allgemein-rekursiven Funktionen auf eher unorthodoxe Weise. Sie hat aber auch andere unorthodoxe Auswirkungen, die zunächst einmal überraschen können. Als erstes will ich zeigen, daß eine einfache Substitutionseigenschaft nicht einmal mehr für Nennformen der ersten Stufe allgemein gilt:

$$(8.1) \quad s = t, \mathfrak{F}[s] \Rightarrow \mathfrak{F}[t].$$

Um das zu sehen, nehme man  $\mathfrak{F} := s = *_{1} \square s = *_{1}$ :

$$s = t, s = s \square s = s \Rightarrow s = t \square s = t.$$

Da  $s = s \square s = s$  ziemlich trivialerweise gilt, erhält man

$$s = t \Rightarrow s = t \square s = t,$$

<sup>90</sup> Die Widerspruchsfreiheit von  $\mathbf{L}^{\mathbf{D}}\lambda$  kann durch eine einfache Induktion gezeigt werden. Nach Gödels zweitem Satz kann damit die Induktion in  $\mathbf{L}^{\mathbf{D}}\lambda$  nicht beweisbar sein.

d.h. ein Schema der Zusammenziehung für Gleichheitsformeln. Mit

$$\lambda A = \lambda \top \leftrightarrow A$$

erhält man daraus sofort ein Schema der Zusammenziehung für beliebige Formeln:

$$A \Rightarrow A \square A.$$

Damit ist klar, daß 8.1 auch für Nennformen der ersten Stufe nicht mit uneingeschränkter Abstraktion verträglich ist.

Was jedoch gezeigt werden kann: Wenn  $\mathfrak{F}$  eine Nennform der ersten Stufe ist, dann gibt es eine natürliche Zahl  $n$  derart, daß

$$(8.2) \quad [s = t]^n, \mathfrak{F}[s] \Rightarrow \mathfrak{F}[t]$$

$\mathbf{LD}_\lambda$ -beweisbar ist.

Wie sieht nun aber die Situation aus, wenn wir die Beschränkung auf Nennformen der ersten Stufe fallen lassen? Die Antwort lautet einfach: Die Fixpunkteigenschaft für Aussageformen macht kurzen Prozeß mit jeder Gleichsetzung von Gleichheit und Identität, auch in der Form einer Regel wie etwa

$$(8.3) \quad \frac{\Rightarrow \lambda F = \emptyset}{\Rightarrow \lambda F \equiv \emptyset}.$$

Aufgrund von 6.1 wissen wir, daß es einen Fixpunkt gibt, der

$$F \leftrightarrow \lambda F \equiv \emptyset$$

erfüllt. Nun läßt sich zeigen, daß

$$\begin{aligned} \lambda F \equiv \emptyset &\Rightarrow \perp, \quad \text{und} \\ \Rightarrow \lambda F = \emptyset & \end{aligned}$$

beide in  $\mathbf{LD}_\lambda$  beweisbar sind,<sup>91</sup> d.h. es gibt einen Term  $\lambda F$  derart, daß die

<sup>91</sup> Der Beweis ist ausgeführt in [36], S. 376, sowie [37], S. 1734. Dort habe ich den Term  $R^\circ := \lambda x(\lambda(x \in x) \equiv \emptyset)$  dafür herangezogen. Dieser Term kann als Variation von  $R := \lambda x(x \notin x)$  gesehen werden, wenn man sich folgendes klar macht: Der Russellsche Term ist logisch äquivalent zu  $\lambda x(\lambda(x \in x) = \emptyset)$ . Man braucht also nur das Zeichen = durch  $\equiv$  zu ersetzen. Das wird unmittelbar deutlich, wenn man statt der Selbstelementenschaft  $x \in x$  die Selbstanwendung  $xx$  betrachtet und anstelle des Terms  $\phi := \lambda xy(xy = \emptyset)$  von S. 121 den Term  $\phi^\circ := \lambda xy(xy \equiv \emptyset)$  hernimmt: der Beweis von S. 121 geht dann über in einen Beweis von  $\phi^\circ \phi^\circ \equiv 0 \Rightarrow$  und  $\Rightarrow \phi^\circ \phi^\circ = 0$ , und das *ohne* eine einzige Zusammenziehung.

Zulassung einer Schlußregel der Gestalt 8.3 unmittelbar zur Inkonsistenz führen würde.

Mithilfe der Fixpunkteigenschaft 6.2 für Termformen kann dieses Ergebnis noch dahingehend verschärft werden, daß jedes Objekt, das sich überhaupt von irgendeinem anderen Objekt unterscheiden läßt, einen Doppelgänger hat, d.h. ein Objekt, das ihm gleich ist, von dem es sich aber auch unterscheiden läßt.<sup>92</sup>

Was also Quine als „confusion“ brandmarkt, ist für die höhere Logik unabdingbar: eine Trennung von Gleichheit und Identität. Als Faustregel kann man davon ausgehen, daß in dem, was analytische Empiristen verteuflern, eine Quelle spekulativer Einsicht verborgen liegt. In diesem Fall: die Morgenröte der Intensionen, die manchen analytischen Philosophen ein Grauen ist.<sup>93</sup>

## 9. Fixpunkte III: das Versagen der Beschreibung

Die Unverträglichkeit der Extensionalität mit der uneingeschränkten Abstraktion kann schon als Anzeichen dafür gewertet werden, daß auch beim Bezeichnen mit Schwierigkeiten zu rechnen ist, wenn uneingeschränkte Abstraktion zugelassen wird. Die Fixpunkteigenschaft macht auch tatsächlich kurzen Prozeß mit Beschreibungsoperatoren in ihrer klassischen Gestalt. Dies ist am augenfälligsten bei der sogenannten unbestimmten Beschreibung, die sich umgangssprachlich durch die Wendung „was die Eigenschaft A hat“ wiedergeben läßt. In der Mathematik, wie auch sonst, ist folgende Art der Argumentation üblich: Angenommen es gibt eine natürliche Zahl mit der Eigenschaft  $\mathfrak{A}$ ; es sei  $p$  eine solche Zahl, dann gilt offensichtlich  $\mathfrak{A}[p]$ . In [39], S. 123, habe ich die Unverträglichkeit der unbestimmten Beschreibung ( $\varepsilon$ -Operator) mit  $\mathbf{L}^1\mathbf{D}_\lambda$  mit einem lächerlich einfachen Fixpunkt gezeigt.

<sup>92</sup> Cf. [36], S. 385, theorem 8.1.

<sup>93</sup> Falls Quines Neigungen in Vergessenheit geraten sein sollten, will ich hier daran erinnern:

Intensions are creatures of darkness, and I shall rejoice with the reader when they are exorcised ([44], p. 180.)

Ich schätze mich glücklich, in einer Zeit zu leben, in der sich christliche Sorge um das Seelenheil des Nächsten nicht mehr in der Form von Inquisition, Folter und Scheiterhaufen ausdrückt, bin mir aber nicht sicher, ob ich dem Frieden trauen kann.

Ich will hier nicht viel Aufhebens um den  $\varepsilon$ -Operator oder seinen nahen Verwandten, den  $\mu$ -Operator, machen, sondern mich vor allem auf die bestimmte Beschreibung, d.h. den  $\iota$ -Operator konzentrieren, und das nicht zuletzt, weil sich in den Beweis in [39], S. 126 f, ein hartnäckiger Druckfehler eingeschlichen hat, den ich hier korrigieren will.

In [39] habe ich u.a. die folgenden Unverträglichkeiten bewiesen:

$$\mathbf{L'D}_\lambda \cup \{\forall x(\mathfrak{F}[x] \wedge \wedge y(\mathfrak{F}[y] \rightarrow x = y)) \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]]\} \vdash \perp, \text{ und}$$

$$\mathbf{L'D}_\lambda \cup \{\forall x \mathfrak{F}[x], \wedge z_1 \wedge z_2(\mathfrak{F}[z_1] \square \mathfrak{F}[z_2] \rightarrow z_1 = z_2) \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]]\} \vdash \perp.$$

Ich folge dem Vorgehen in [39], S. 126, und verbessere nur die Druckfehler.

Man nehme die Nennform

$$\mathfrak{F} := *_1 \in \{0, 1\} \square \phi \neq *_1 \square \wedge y(y \in \{0, 1\} \square y < *_1 \rightarrow \neg(\phi \neq y)),^{94}$$

wobei  $\phi$  wie in [39] der Fixpunkt  $\phi = \iota x \mathfrak{F}[x]$  ist. Durch  $\square$ -Beseitigungen ergibt sich daraus sofort  $\mathfrak{F}[\iota x \mathfrak{F}[x]] \Rightarrow \phi \neq \iota x \mathfrak{F}[x]$ . Damit erhält man:

$$\forall x(x \in \{0, 1\} \square \phi \neq x \square \wedge y(y \in \{0, 1\} \square y < x \rightarrow \neg(\phi \neq y)) \Rightarrow \phi \neq \iota x \mathfrak{F}[x].$$

Das weitere Vorgehen ergibt sich aus [39], S. 127, indem zuerst einmal die oberste Formel auf dieser Seite (Zeile 4) durch

$$0 \in \{0, 1\}, \phi \neq 0, \wedge y(y \in \{0, 1\} \square y < 0 \rightarrow \neg(\phi \neq y)) \Rightarrow$$

ersetzt wird. Wie im Beweis von Proposition 3.4 in [39], S. 124, erhält man  $\phi \neq 0 \Rightarrow$  und damit  $\phi \neq 1 \Rightarrow$ . Das wiederum ergibt die Ersetzung der zweiten Formel auf S. 127 (Zeile 6) durch

$$\Rightarrow \wedge y(y \in \{0, 1\} \square y < 1 \rightarrow \neg(\phi \neq y)).$$

Mit  $\Rightarrow 1 \in \{0, 1\}$  und

$$1 \in \{0, 1\}, \phi \neq 1, \wedge y(y \in \{0, 1\} \square y < 1 \rightarrow \neg(\phi \neq y)) \Rightarrow$$

erhält man daraus durch einen Schnitt  $\phi \neq 1 \Rightarrow$ , und damit einen Widerspruch.

Damit wird erneut die Frage relevant, was durch den bestimmten Artikel eigentlich bezeichnet wird, Wenn wir zum Beispiel von der leeren Menge sprechen, sprechen wir von  $\emptyset$ , d.h.  $\lambda x(x \neq x)$ , und/oder auch von  $\lambda x(\lambda(x \in x) \neq \emptyset)$ ? Die beiden sind gleich, aber nicht identisch.<sup>95</sup> Eine

<sup>94</sup> Das letzte  $y$  ersetzt das Nennzeichen  $*_1$  in [39], S. 126.

<sup>95</sup> Cf. [36], S. 376.

derartige Konstellation wird in den herkömmlichen Ansätzen zum Problem des Bezeichnens gar nicht berücksichtigt. Aber die Situation hat eine auffallende Ähnlichkeit mit derjenigen, die Hegel als seinen Ausgangspunkt in der *Wissenschaft der Logik* mit den Begriffen *Sein* und *Nichts* gewählt hat. Hegel beginnt mit der extensionalen Identität und intensionalen Nichtidentität von zwei Begriffen — Sein und Nichts — und nutzt, von hier ausgehend, diesen intensionalen Unterschied für eine Grundlegung von Kategorien aus. Kein Zweifel, Hegels Argumentation ist hoffnungslos, aber die Idee, daß die uneingeschränkte Begriffsbildung („substantivierte Aussageform“) es offensichtlich unmöglich macht, Objekte eindeutig zu bezeichnen, und daß wir so dazu getrieben werden, weitere Verfeinerungen der Bestimmungen vorzunehmen, die aus intensionalen Unterschieden gespeist werden, ist eine herausfordernde Konzeption, die bisher wohl kaum eine ernsthafte Beachtung erfahren hat. Die Unvereinbarkeit von klassischen logischen Prinzipien (*Tertium non datur*, Zusammenziehung), Extensionalität und Beschreibung mit uneingeschränkter Begriffsbildung mag dieser Idee ausreichenden Nachdruck verleihen, um eine Forschung in Richtung intensionaler Bestimmungen zu motivieren. Die Einführung der Modalitäten in [36] war ein bewußter Versuch in diese Richtung. Allgemeiner formuliert, ich hoffe, daß es möglich sein wird, die Gesetze intensionaler Entitäten wie der Modalität, Kausalität, usw. von den Strukturschlußregeln der Gentzenschen Logik gewissermaßen „ablesen“ zu können. Aber bis dahin ist noch ein weiter Weg.

## 10. Die Unverträglichkeit der klassischen Induktion

Anders als bei den bisherigen Unverträglichkeitsresultaten lege ich hier kein direktes Fixpunktergebnis vor, sondern nur einen Beweis der linken Zusammenziehungsregel mithilfe der uneingeschränkten Abstraktion und des klassischen Induktionsschemas. Im Endergebnis läuft das aber auf dasselbe hinaus: Das klassische Induktionsschema ist nicht mit der uneingeschränkten Abstraktion verträglich.<sup>96</sup>

<sup>96</sup> Dieses Ergebnis sollte vor dem Hintergrund einer Bemerkung wie der folgenden gesehen werden:

Es gibt [...] Beweisverfahren, die ganz unabhängig von irgendwelchen Systemen, denen sie einverleibt werden können, als gültig einsehbar sind. Zu ihnen



Zuerst will ich erklären, was mit dem Ausdruck „klassische Induktion“ gemeint ist. Wenn es einen Term  $\mathbf{X}$  in der Sprache einer Theorie  $\Sigma$  (in Sequenzenformulierung) der Mengenlehre (oder höheren Logik) gibt, so daß die folgenden (als *Peano-Axiome* bekannten) Bedingungen

- (X1)  $\Rightarrow 0 \in \mathbf{X}$ ,
- (X2)  $s \in \mathbf{X} \Rightarrow s' \in \mathbf{X}$ ,
- (X3)  $s \in \mathbf{X}, 0 = s' \Rightarrow$ ,
- (X4)  $s \in \mathbf{X}, t \in \mathbf{X}, s' = t' \Rightarrow s = t$ ,
- (X5)  $\mathfrak{F}[0], \wedge x (\mathfrak{F}[x] \rightarrow \mathfrak{F}[x']), s \in \mathbf{X} \Rightarrow \mathfrak{F}[s]$

$\Sigma$ -beweisbar sind, so sage ich, daß  $\Sigma$  über *klassische Induktion* verfügt.<sup>97</sup>

Mit X5 erhält man sofort das *klassische Induktionsschema*

$$(X5') \quad \frac{\Gamma \Rightarrow \mathfrak{F}[0] \quad \Gamma, \mathfrak{F}[a] \Rightarrow \mathfrak{F}[a']}{s \in \mathbf{X}, \Gamma \Rightarrow \mathfrak{F}[s]}$$

mit der üblichen Variablenbedingung: Die Eigenvariable  $a$  darf nicht in der Untersequenz auftreten.

Zusammenziehungen für  $s \in \mathbf{X}$ -Formeln ergeben sich als unmittelbare Konsequenz aus X1, X2, und X5:

$$\frac{\begin{array}{c} \Rightarrow 0 \in \mathbf{X} \quad \Rightarrow 0 \in \mathbf{X} \\ \hline \Rightarrow 0 \in \mathbf{X} \square 0 \in \mathbf{X} \end{array}}{\begin{array}{c} a \in \mathbf{X} \Rightarrow a' \in \mathbf{X} \quad a \in \mathbf{X} \Rightarrow a' \in \mathbf{X} \\ \hline a \in \mathbf{X}, a \in \mathbf{X} \Rightarrow a' \in \mathbf{X} \square a' \in \mathbf{X} \\ \hline a \in \mathbf{X} \square a \in \mathbf{X} \Rightarrow a' \in \mathbf{X} \square a' \in \mathbf{X} \\ \hline s \in \mathbf{X} \Rightarrow s \in \mathbf{X} \square s \in \mathbf{X} \end{array}}$$

Der Bequemlichkeit halber führe ich eine abgekürzte Schreibweise für relativierte Quantoren ein:

$$\begin{aligned} \wedge^x \mathfrak{F}[x] &::= \wedge x (x \in \mathbf{X} \rightarrow \mathfrak{F}[x]), \quad \text{und} \\ \vee^x \mathfrak{F}[x] &::= \vee x (x \in \mathbf{X} \square \mathfrak{F}[x]). \end{aligned}$$

Weiterhin wird folgender Term benötigt:

$$\mathbf{B}^\circ ::= \lambda x (x \equiv 0 \vee \vee^x y (x \equiv y')).$$

gehört etwa die vollständige Induktion[.] ([59], S 102.)

<sup>97</sup> Man beachte, daß X1-X4 für  $\mathbf{X} \equiv \mathbf{N}^\circ$  in  $\mathbf{L}^1\mathbf{D}_\lambda^Z$  beweisbar sind. Mit anderen Worten, es geht eigentlich nur um X5.

Folgende Sequenzen lassen sich nun in  $\mathbf{UD}_\lambda$  beweisen:

- (B1)  $\Rightarrow 0 \in \mathbf{B}^\circ$ ,  
 (B2)  $s \in \mathbf{X} \Rightarrow s' \in \mathbf{B}^\circ$ ,  
 (B3)  $s \in \mathbf{B}^\circ \Rightarrow s \in \mathbf{X}$ ,  
 (B4)  $s \in \mathbf{B}^\circ \Rightarrow s' \in \mathbf{B}^\circ$ ,  
 (B5)  $s \in \mathbf{B}^\circ \Rightarrow s \in \mathbf{B}^\circ \square s \in \mathbf{B}^\circ$ .

Die Beweise machen keine Schwierigkeiten. Ich zeige nur B3 und B5.

$$\frac{\frac{\frac{s \in \mathbf{X} \Rightarrow s \in \mathbf{X} \quad \Rightarrow s' \equiv s'}{s \in \mathbf{X} \Rightarrow s \in \mathbf{X} \square s' \equiv s'}}{s \in \mathbf{X} \Rightarrow \bigvee^x y (s' \equiv y')}}{s \in \mathbf{X} \Rightarrow s' \equiv 0 \vee \bigvee^x y (s' \equiv y')}}{s \in \mathbf{X} \Rightarrow s' \in \mathbf{B}^\circ}.$$

Unter Einsatz von B3 kann man folgendermaßen fortfahren:

$$\frac{\frac{\frac{\Rightarrow 0 \in \mathbf{B}^\circ \quad \Rightarrow 0 \in \mathbf{B}^\circ}{\Rightarrow 0 \in \mathbf{B}^\circ \square 0 \in \mathbf{B}^\circ}}{s \equiv 0 \Rightarrow s \in \mathbf{B}^\circ \square s \in \mathbf{B}^\circ}}{\frac{\frac{\frac{b \in \mathbf{X} \Rightarrow b' \in \mathbf{B}^\circ \quad b \in \mathbf{X} \Rightarrow b' \in \mathbf{B}^\circ}{b \in \mathbf{X}, b \in \mathbf{X} \Rightarrow b' \in \mathbf{B}^\circ \square b' \in \mathbf{B}^\circ}}{b \in \mathbf{X}, b \in \mathbf{X}, s \equiv b' \Rightarrow s \in \mathbf{B}^\circ \square s \in \mathbf{B}^\circ}}{b \in \mathbf{X}, s \equiv b' \Rightarrow s \in \mathbf{B}^\circ \square s \in \mathbf{B}^\circ}}{\frac{\bigvee^x y (s \equiv y') \Rightarrow s \in \mathbf{B}^\circ \square s \in \mathbf{B}^\circ}}{x \equiv 0 \vee \bigvee^x y (s \equiv y') \Rightarrow s \in \mathbf{B}^\circ \square s \in \mathbf{B}^\circ}}{s \in \mathbf{B}^\circ \Rightarrow s \in \mathbf{B}^\circ \square s \in \mathbf{B}^\circ}.$$

Um nun zeigen zu können, wie klassische Induktion mit uneingeschränkter Abstraktion eingesetzt werden kann, um Zusammenziehungen zu erhalten, führe ich den folgenden Term ein (wobei  $A$  diejenige Formel ist, für die die Zusammenziehung gezeigt werden soll):

$$\rho_A \equiv \lambda x_1 x_2 \bigwedge y (\langle 0, \lambda A \rangle \in y \square \bigwedge^x z (\langle z', \lambda A \square \lambda A \rangle \in y) \rightarrow \langle x_1, x_2 \rangle \in y),$$

wobei eine *inklusive* (d.i. gegenüber Ersetzungen offene) Version  $\rho_A$  wie folgt definiert ist:

$$\rho_A \equiv \lambda x_1 x_2 \bigvee^x y \bigvee z (y = x_1 \square z = x_2 \square \langle y, z \rangle \in \rho_A).$$

Damit erhalten wir unmittelbar:

$$\begin{aligned} &\Rightarrow \langle 0, \lambda A \rangle \in \rho_A, \quad \text{und} \\ &\Rightarrow \langle s', \lambda A \sqcap \lambda A \rangle \in \rho_A. \end{aligned}$$

Als einfache Folgerung davon und den Eigenschaften der Anwendung erhalten wir:

$$\begin{aligned} &\Rightarrow \rho_A \llbracket 0 \rrbracket \subseteq \lambda A, \quad \text{und} \\ &\Rightarrow \rho_A \llbracket s' \rrbracket \subseteq \lambda A \sqcap \lambda A. \end{aligned}$$

Ich zeige nur die erste:

$$\begin{aligned} &\Rightarrow \langle 0, \lambda A \rangle \in \rho_A \quad a \in \lambda A \Rightarrow a \in \lambda A \\ &\hline &\langle 0, \lambda A \rangle \in \rho_A \rightarrow a \in \lambda A \Rightarrow a \in \lambda A \\ &\hline &\bigwedge y (\langle 0, y \rangle \in \rho_A \rightarrow a \in y) \Rightarrow a \in \lambda A \\ &\hline &\frac{a \in \rho_A \llbracket 0 \rrbracket \Rightarrow a \in \lambda A}{\Rightarrow a \in \rho_A \llbracket 0 \rrbracket \rightarrow a \in \lambda A} \\ &\hline &\Rightarrow \rho_A \llbracket 0 \rrbracket \subseteq \lambda A \end{aligned}$$

Die andere Richtung erfordert eine gewisse „Eindeutigkeit“ (Rechts-eindeutigkeit), die eine wohlbekannte Konsequenz der Art und Weise ist, wie der Term  $\rho_A^\circ$  definiert ist:

$$\begin{aligned} &\Rightarrow \lambda A \subseteq \rho_A \llbracket 0 \rrbracket, \quad \text{und} \\ &\Rightarrow \lambda A \sqcap \lambda A \subseteq \rho_A \llbracket s' \rrbracket. \end{aligned}$$

Ich behandle nur den ersten Fall:

$$\begin{aligned} &\Rightarrow 0 = 0 \quad \Rightarrow \lambda A = \lambda A \\ &\hline &\Rightarrow 0 = 0 \sqcap \lambda A = \lambda A \\ &\hline &\Rightarrow (0 = 0 \sqcap \lambda A = \lambda A) \vee \forall^x y (0 = y' \sqcap \lambda A = \lambda A \sqcap \lambda A) \\ &\hline &\Rightarrow \langle 0, \lambda A \rangle \in \lambda x_1 x_2 ((x_1 = 0 \sqcap x_2 = \lambda A) \vee \forall^x y (x_1 = y' \sqcap x_2 = \lambda A \sqcap \lambda A)) \end{aligned}$$

Der einfacheren Darstellbarkeit halber schreibe ich in den beiden folgenden Beweisfiguren  $\lambda A^2$  anstelle von  $\lambda A \sqcap \lambda A$ :

$$\begin{array}{c}
c \in \mathbf{X} \Rightarrow c' = c' \quad \Rightarrow c' = c' \\
\hline
c \in \mathbf{X} \Rightarrow c \in \mathbf{X} \square c' = c' \quad \Rightarrow \lambda A^2 = \lambda A^2 \\
\hline
c \in \mathbf{X} \Rightarrow c \in \mathbf{X} \square c' = c' \square \lambda A^2 = \lambda A^2 \\
\hline
c \in \mathbf{X} \Rightarrow \bigvee^x y (c' = y' \square \lambda A^2 = \lambda A^2) \\
\hline
c \in \mathbf{X} \Rightarrow (c' = 0 \square \lambda A^2 = \lambda A) \vee \bigvee^x y (c' = y' \square \lambda A^2 = \lambda A^2) \\
\hline
c \in \mathbf{X} \Rightarrow \langle c', \lambda A^2 \rangle \in \lambda x_1 x_2 ((x_1 = 0 \square x_2 = \lambda A) \vee \bigvee^x y (x_1 = y' \square x_2 = \lambda A^2)) \\
\hline
\Rightarrow \bigwedge^x z (\langle z', \lambda A^2 \rangle \in \lambda x_1 x_2 ((x_1 = 0 \square x_2 = \lambda A) \vee \bigvee^x y (x_1 = y' \square x_2 = \lambda A^2)))
\end{array}$$

Unter Einsatz von X3 lässt sich dann zeigen:

$$\begin{array}{c}
a \in \mathbf{X}, 0 = a' \Rightarrow b = \lambda A \\
\hline
b = \lambda A \Rightarrow b = \lambda A \quad a \in \mathbf{X}, 0 = a', b = \lambda A^2 \Rightarrow b = \lambda A \\
\hline
0 = 0, b = \lambda A \Rightarrow b = \lambda A \quad a \in \mathbf{X} \square 0 = a' \square b = \lambda A^2 \Rightarrow b = \lambda A \\
\hline
0 = 0 \square b = \lambda A \Rightarrow b = \lambda A \quad \bigvee^x y (0 = y' \square b = \lambda A^2) \Rightarrow b = \lambda A \\
\hline
(0 = 0 \square b = \lambda A) \vee \bigvee^x y (0 = y' \square b = \lambda A^2) \Rightarrow b = \lambda A \\
\hline
\langle 0, b \rangle \in \lambda x_1 x_2 ((x_1 = 0 \square x_2 = \lambda A) \vee \bigvee^x y (x_1 = y' \square x_2 = \lambda A^2)) \Rightarrow b = \lambda A
\end{array}$$

Jetzt werden diese Ergebnisse in der folgenden Beweisfigur verwendet. Um die Darstellung zu erleichtern, wird im folgenden  $\xi$  an die Stelle von  $\lambda x_1 x_2 ((x_1 = 0 \square x_2 = \lambda A) \vee \bigvee^x y (x_1 = y' \square x_2 = \lambda A \square \lambda A))$  treten:

$$\begin{array}{c}
\langle 0, b \rangle \in \xi \Rightarrow b = \lambda A \quad a \in \lambda A, b = \lambda A \Rightarrow a \in b \\
\hline
\Rightarrow \langle 0, \lambda A \rangle \in \xi \square \bigwedge^x z (\langle z', \lambda A \square \lambda A \rangle \in \xi) \quad a \in \lambda A, \langle 0, b \rangle \in \xi \Rightarrow a \in b \\
\hline
a \in \lambda A, \langle 0, \lambda A \rangle \in \xi \square \bigwedge^x z (\langle z', \lambda A \square \lambda A \rangle \in \xi) \rightarrow \langle 0, b \rangle \in \xi \Rightarrow a \in b \\
\hline
a \in \lambda A, \bigwedge y (\langle 0, \lambda A \rangle \in y \square \bigwedge^x z (\langle z', \lambda A \square \lambda A \rangle \in y)) \rightarrow \langle 0, b \rangle \in y \Rightarrow a \in b \\
\hline
a \in \lambda A, \langle 0, b \rangle \in \rho_A \Rightarrow a \in b \\
\hline
a \in \lambda A \Rightarrow \langle 0, b \rangle \in \rho_A \rightarrow a \in b \\
\hline
a \in \lambda A \Rightarrow \bigwedge y (\langle 0, y \rangle \in \rho_A \rightarrow a \in y) \\
\hline
a \in \lambda A \Rightarrow a \in \rho_A[0] \\
\hline
\Rightarrow a \in \lambda A \rightarrow a \in \rho_A[0] \\
\hline
\Rightarrow \lambda A \subseteq \rho_A[0]
\end{array}$$

Damit erhält man unmittelbar:

$$(10.1) \quad \Rightarrow \rho_A[[0]] = \lambda A, \text{ und}$$

$$(10.2) \quad \Rightarrow \rho_A[[s']] = \lambda A \sqcap \lambda A.$$

Aus 10.2 wiederum erhält man sofort

$$(10.3) \quad \rho_A[[s']] = \lambda \top \Rightarrow A \sqcap A$$

auf folgende Weise:

$$\frac{\Rightarrow \rho_A[[s']] = \lambda A \sqcap \lambda A}{\rho_A[[s']] = \lambda \top \Rightarrow \lambda A \sqcap \lambda A = \lambda \top \quad \lambda A \sqcap \lambda A = \lambda \top \Rightarrow A \sqcap A} \cdot \frac{\rho_A[[s']] = \lambda \top \Rightarrow A \sqcap A}{\rho_A[[s']] = \lambda \top \Rightarrow A \sqcap A}.$$

Das wird etwas weiter unten zur Anwendung kommen.

Jetzt kann ich zeigen, daß aus 10.1 und 10.2 mithilfe von X5' Zusammenziehungen folgen. Zuerst die „Induktionsbasis“:

$$\frac{\Rightarrow \rho_A[[0]] = \lambda A \quad A \Rightarrow \lambda A = \lambda \top}{\Rightarrow 0 \in \mathbf{B}^\circ \quad A \Rightarrow \rho_A[[0]] = \lambda \top} \cdot \frac{A \Rightarrow 0 \in \mathbf{B}^\circ \sqcap \rho_A[[0]] = \lambda \top}{A \Rightarrow 0 \in \mathbf{B}^\circ \sqcap \rho_A[[0]] = \lambda \top}.$$

Als nächstes kommt der „Induktionsschritt“. In einem ersten Schritt erhalten wir  $A, s \in \mathbf{B}^\circ, \rho[[s]] = \lambda \top \Rightarrow \rho[[s']] = \lambda \top$  (um Platz zu sparen, lasse ich hier das indizierte  $A$  weg und schreibe einfach nur  $\rho$ ):

$$\frac{\Rightarrow \rho[[0]] = \lambda A \quad A, \lambda A = \lambda \top \Rightarrow \rho[[0']] = \lambda \top \quad \frac{\rho[[b']] = \lambda \top \Rightarrow \rho[[b'']] = \lambda \top}{s \equiv b', \rho[[s]] = \lambda \top \Rightarrow \rho[[s']] = \lambda \top}}{\frac{A, \rho[[0]] = \lambda \top \Rightarrow \rho[[0']] = \lambda \top \quad \frac{\bigvee^x y (s \equiv y'), \rho[[s]] = \lambda \top \Rightarrow \rho[[s']] = \lambda \top}{A, \bigvee^x y (s \equiv y'), \rho[[s]] = \lambda \top \Rightarrow \rho[[s']] = \lambda \top}}{A, s \in \mathbf{B}^\circ, \rho[[s]] = \lambda \top \Rightarrow \rho[[s']] = \lambda \top}.$$

Und weiter:

$$\frac{s \in \mathbf{B}^\circ \Rightarrow s' \in \mathbf{B}^\circ \quad A, s \in \mathbf{B}^\circ, \rho_A[[s]] = \lambda \top \Rightarrow \rho_A[[s']] = \lambda \top}{\frac{A, s \in \mathbf{B}^\circ, s \in \mathbf{B}^\circ, \rho_A[[s]] = \lambda \top \Rightarrow s' \in \mathbf{B}^\circ \sqcap \rho_A[[s']] = \lambda \top}{A, s \in \mathbf{B}^\circ, \rho_A[[s]] = \lambda \top \Rightarrow s' \in \mathbf{B}^\circ \sqcap \rho_A[[s']] = \lambda \top}} \cdot \frac{A, s \in \mathbf{B}^\circ \sqcap \rho_A[[s]] = \lambda \top \Rightarrow s' \in \mathbf{B}^\circ \sqcap \rho_A[[s']] = \lambda \top}{A, s \in \mathbf{B}^\circ \sqcap \rho_A[[s]] = \lambda \top \Rightarrow s' \in \mathbf{B}^\circ \sqcap \rho_A[[s']] = \lambda \top}.$$

Zusammen, durch Anwendung von X5', ergibt das:

$$\frac{A \Rightarrow 0 \in \mathbf{B}^\circ \square \rho_A[[0]] = \lambda\top \quad A, a \in \mathbf{B}^\circ \square \rho_A[[a]] = \lambda\top \Rightarrow a' \in \mathbf{B}^\circ \square \rho_A[[a']] = \lambda\top}{\Rightarrow 0' \in \mathbf{X} \quad 0' \in \mathbf{X}, A \Rightarrow 0' \in \mathbf{B}^\circ \square \rho_A[[0']] = \lambda\top}$$

$$\frac{A \Rightarrow 0' \in \mathbf{B}^\circ \square \rho_A[[0']] = \lambda\top}{A \Rightarrow \rho_A[[0']] = \lambda\top}.$$

Schließlich kommt ein Schnitt mit 10.3 zur Anwendung:

$$\frac{A \Rightarrow \rho_A[[0']] = \lambda\top \quad \rho_A[[0']] = \lambda\top \Rightarrow A \square A}{A \Rightarrow A \square A}.$$

Das ergibt dann Zusammenziehungen auf die übliche Weise:

$$\frac{A, A, \Gamma \Rightarrow C}{A \Rightarrow A \square A \quad A \square A, \Gamma \Rightarrow C}$$

$$\frac{A, \Gamma \Rightarrow C}{A, \Gamma \Rightarrow C}.$$

Damit ist gezeigt, daß auch die klassische Induktion nicht mit uneingeschränkter Abstraktion verträglich ist.

## 11. Die Unverträglichkeit des Auswahlaxioms

Die Unverträglichkeit des  $\varepsilon$ -Operators (mit dem klassischen  $\varepsilon$ -Axiom) von Abschnitt 9 kann schon als Hinweis darauf gewertet werden, daß auch das Auswahlaxiom durch die uneingeschränkte Abstraktion in Mitleidenschaft gezogen wird.

Ich habe keinen Fixpunkt, der in unmittelbarer Weise das Auswahlaxiom zu Fall bringt, aber es ist bekannt, daß das Auswahlaxiom auf der Basis einer intuitionistischen Mengenlehre zur Beweisbarkeit des *Tertium non datur* führt.<sup>98</sup> Dieses Argument kann der Situation der zusammenziehungsfreien Logik angepaßt werden.<sup>99</sup>

<sup>98</sup> Vgl. [3], S. 163, wo auf [9] verwiesen wird. Mir scheint [20] jedoch der Sache näher zu kommen.

<sup>99</sup> Ein entsprechendes Argument geht auch für den Auswahloperator  $\varepsilon$  in der zusammenziehungsfreien Logik durch, aber das interessiert hier kaum, angesichts des Fixpunktarguments in Abschnitt 9.

Die Form des Auswahlaxioms, die im intuitionistischen Fall zur Anwendung kommt, ist

$$\bigwedge x (\mathfrak{A}[x] \rightarrow \bigvee y \mathfrak{B}[x, y]) \rightarrow \bigvee y \bigwedge x (\mathfrak{A}[x] \rightarrow \mathfrak{B}[x, y[x]]) .$$

Man kann leicht zeigen, daß das eine klassische Konsequenz des Auswahlaxioms in der Formulierung

$$(11.1) \quad \bigwedge x \bigvee y (\mathfrak{A}[x] \rightarrow \mathfrak{B}[x, y]) \rightarrow \bigvee y \bigwedge x (\mathfrak{A}[x] \rightarrow \mathfrak{B}[x, y[x]])$$

ist:

$$\begin{array}{c} \mathfrak{B}[a, b], \mathfrak{A}[a] \Rightarrow \mathfrak{B}[a, b] \\ \hline \mathfrak{B}[a, b], \mathfrak{A}[a] \Rightarrow \mathfrak{B}[a, b] \\ \hline \mathfrak{B}[a, b] \Rightarrow \mathfrak{A}[a] \rightarrow \mathfrak{B}[a, b] \\ \hline \mathfrak{B}[a, b] \Rightarrow \bigvee y (\mathfrak{A}[a] \rightarrow \mathfrak{B}[a, y]) \\ \hline \mathfrak{A}[a] \Rightarrow \mathfrak{A}[a] \quad \bigvee y \mathfrak{B}[a, y] \Rightarrow \bigvee y (\mathfrak{A}[a] \rightarrow \mathfrak{B}[a, y]) \\ \hline \mathfrak{A}[a] \rightarrow \bigvee y \mathfrak{B}[a, y], \mathfrak{A}[a] \Rightarrow \bigvee y (\mathfrak{A}[a] \rightarrow \mathfrak{B}[a, y]) \\ \hline \mathfrak{A}[a] \rightarrow \bigvee y \mathfrak{B}[a, y], \mathfrak{A}[a] \Rightarrow \mathfrak{B}[a, b], \bigvee y (\mathfrak{A}[a] \rightarrow \mathfrak{B}[a, y]) \\ \hline \mathfrak{A}[a] \rightarrow \bigvee y \mathfrak{B}[a, y] \Rightarrow \mathfrak{A}[a] \rightarrow \mathfrak{B}[a, b], \bigvee y (\mathfrak{A}[a] \rightarrow \mathfrak{B}[a, y]) \\ \hline \mathfrak{A}[a] \rightarrow \bigvee y \mathfrak{B}[a, y] \Rightarrow \bigvee y (\mathfrak{A}[a] \rightarrow \mathfrak{B}[a, y]), \bigvee y (\mathfrak{A}[a] \rightarrow \mathfrak{B}[a, y]) \\ \hline \mathfrak{A}[a] \rightarrow \bigvee y \mathfrak{B}[a, y] \Rightarrow \bigvee y (\mathfrak{A}[a] \rightarrow \mathfrak{B}[a, y]) \\ \hline \bigwedge x (\mathfrak{A}[x] \rightarrow \bigvee y \mathfrak{B}[x, y]) \Rightarrow \bigvee y (\mathfrak{A}[a] \rightarrow \mathfrak{B}[a, y]) \\ \hline \bigwedge x (\mathfrak{A}[x] \rightarrow \bigvee y \mathfrak{B}[x, y]) \Rightarrow \bigwedge x \bigvee y (\mathfrak{A}[x] \rightarrow \mathfrak{B}[x, y]) \end{array} .$$

Das ist zwar kein intuitionistisch gültiger Beweis, aber es geht ja nur darum, eine Anwendung des klassischen Auswahlaxioms zu finden, aus dem sich intuitionistisch das *Tertium non datur* beweisen läßt.

Für den Fall der zusammenziehungsfreien Logik mit uneingeschränkter Abstraktion muß das noch weiter abgewandelt werden. Die Form des Auswahlaxioms, die hier zur Anwendung kommt, sieht folgendermaßen aus:

$$(11.2) \quad [\bigwedge x (\mathfrak{A}[x] \rightarrow \bigvee y \mathfrak{B}[x, y])]^2 \rightarrow \bigvee y [\bigwedge x (\mathfrak{A}[x] \rightarrow \mathfrak{B}[x, y[x]^*])]^2 ,$$

wobei  $s[[t]]^*$  eine *inklusive* (gegenüber Ersetzungen offene) Version der Anwendung ist, die folgendermaßen definiert ist:

$$s[[t]]^* := \lambda x \bigwedge y (\exists t, y \in s^* \rightarrow x \in y) ,$$





Zu 11.5: Mithilfe von 11.4 und dem Auswahlaxiom in der Form

$$\bigwedge x(x = \theta \vee x = \eta \rightarrow \bigvee y(y \in x)) \rightarrow \bigvee y \bigwedge x(x = \theta \vee x = \eta \rightarrow y[x]^* \in x)$$

erhält man:

$$\bigvee y \bigwedge x(x = \theta \vee x = \eta \rightarrow y[x]^* \in x).$$

Damit sind weiterhin die folgenden Sequenzen beweisbar:

$$(11.6) \quad A \Rightarrow \theta = \eta,$$

$$(11.7) \quad a[\theta]^* \neq a[\eta]^* \Rightarrow \neg A,$$

$$(11.8) \quad a[\theta]^* \in \theta, a[\eta]^* \in \eta \Rightarrow a[\theta]^* \neq a[\eta]^* \vee A,$$

$$(11.9) \quad a[\theta]^* \in \theta, a[\eta]^* \in \eta \Rightarrow A \vee \neg A,$$

$$(11.10) \quad \bigvee y \bigwedge x(x = \theta \vee x = \eta \rightarrow y[x]^* \in x) \Rightarrow A \vee \neg A,$$

$$(11.11) \quad \Rightarrow A \vee \neg A.$$

Zum Beweis betrachte man der Reihe nach:

Zu 11.6. Das Ganze ist symmetrisch und ich zeige nur die eine Richtung:

$$\frac{\frac{A \Rightarrow A}{A, a = 0 \Rightarrow a = 1 \vee A} \quad \frac{A \Rightarrow A}{A, A \Rightarrow a = 1 \vee A}}{A, a = 0 \vee A \Rightarrow a = 1 \vee A} \frac{A, a \in \theta \Rightarrow a \in \eta}{A \Rightarrow a \in \theta \rightarrow a \in \eta}.$$

Zu 11.7. Man nehme 11.6 und 11.3, d.h. die Möglichkeit, gleiche Terme durch einander zu ersetzen:

$$\frac{A \Rightarrow \theta = \eta \quad \theta = \eta \Rightarrow a[\theta]^* = a[\eta]}{A \Rightarrow a[\theta]^* = a[\eta]} \frac{a[\theta]^* \neq a[\eta]^* \Rightarrow \neg A}{a[\theta]^* \neq a[\eta]^* \Rightarrow \neg A}.$$

Zu 11.8.

$$\begin{array}{c}
 0 = 1 \Rightarrow \\
 \hline
 a[\theta]^* = 0, a[\eta]^* = 1, a[\theta]^* = a[\eta]^* \Rightarrow \\
 \hline
 a[\theta]^* = 0, a[\eta]^* = 1 \Rightarrow a[\theta]^* \neq [\eta] \\
 \hline
 a[\theta]^* = 0, a[\eta]^* = 1 \Rightarrow a[\theta]^* \neq a[\eta]^* \vee A \quad A \Rightarrow a[\theta]^* \neq a[\eta]^* \vee A \\
 \hline
 a[\theta]^* = 0 \vee A, a[\eta]^* = 1 \vee A \Rightarrow a[\theta]^* \neq a[\eta]^* \vee A \\
 \hline
 a[\theta]^* \in \theta, a[\eta]^* \in \eta \Rightarrow a[\theta]^* \neq a[\eta]^* \vee A
 \end{array}$$

Zu 11.9. Man nehme 11.8 und 11.7:

$$\begin{array}{c}
 a[\theta]^* \neq a[\eta]^* \Rightarrow \neg A \\
 \hline
 a[\theta]^* \neq a[\eta]^* \Rightarrow A \vee \neg A \quad A \Rightarrow A \vee \neg A \\
 \hline
 a[\theta]^* \in \theta, a[\eta]^* \in \eta \Rightarrow a[\theta]^* \neq a[\eta]^* \vee A \quad a[\theta]^* \neq a[\eta]^* \vee A \Rightarrow A \vee \neg A \\
 \hline
 a[\theta]^* \in \theta, a[\eta]^* \in \eta \Rightarrow A \vee \neg A
 \end{array}$$

Zu 11.10. Man nehme 11.9:

$$\begin{array}{c}
 \Rightarrow \theta = \theta \vee \theta = \eta \quad a[\theta]^* \in \theta, a[\eta]^* \in \eta \Rightarrow A \vee \neg A \\
 \hline
 \theta = \theta \vee \theta = \eta \rightarrow a[\theta]^* \in \theta, \theta = \theta \vee \theta = \eta \rightarrow a[\eta]^* \in \eta \Rightarrow A \vee \neg A \\
 \hline
 [\wedge x(x = \theta \vee x = \eta \rightarrow a[x]^* \in x)]^2 \Rightarrow A \vee \neg A \\
 \hline
 \forall y[\wedge x(x = \theta \vee x = \eta \rightarrow y[x]^* \in x)]^2 \Rightarrow A \vee \neg A
 \end{array}$$

11.11 ist nun eine unmittelbare Folgerung aus 11.5 und 11.10 mithilfe eines Schnitts.

Da uneingeschränkte Abstraktion nicht mit dem Gesetz des *Tertium non datur* verträglich ist, ist damit auch die Unverträglichkeit (einer Form) des (klassischen) Auswahlaxioms mit der uneingeschränkten Abstraktion gezeigt.

## 12. Selbstbezüglichkeit im Einsatz: die Definition von Z

Wie die letzten vier Abschnitte zeigen, fordert die Zulassung der uneingeschränkten Abstraktion massive Einschränkungen an klassischen Prinzipien, die grundsätzlich über die Einschränkung der Logik hinausgehen, nämlich:

1. Unterscheidung von Gleichheit und Identität;
2. Preisgabe klassischer Formen der Beschreibung;
3. Preisgabe der klassischen Form der Induktion;
4. Preisgabe des Auswahlaxioms.

All diese Unverträglichkeitsresultate zeigen aus meiner Sicht etwas an, nämlich, daß mit der klassischen Lehrmeinung etwas nicht stimmt. Die klassische Lehrmeinung suggeriert, wir könnten so schließen, „als ob“ wir vollständiges Wissen hätten („Gottähnlichkeit“ oder „divinisimilitude“).

Wir schließen mit realen Annahmen, sprachlichen Gebilden, die ihrerseits auf der Objektebene (evtl. kodiert) in Erscheinung treten können und auf diese Weise Doppeldeutigkeit erzeugen. Wie tragen wir einer solchen Rolle der Annahmen in Regeln für ein logisches Schließen Rechnung? Das heißt im Falle der Zusammenziehungsregel, daß wir in gewisser Weise Buchhaltung führen über unseren Gebrauch der Zusammenziehungsregeln.<sup>100</sup> Mit anderen Worten, ob wir in einem Beweis eine Annahme ein- oder zweimal gebrauchen, soll im Endresultat erkennbar sein.

Das Problem: Wie kann man über Annahmen gleicher Gestalt quantifizieren? D.h., wie kann man ausdrücken, daß eine Formel  $A$  mehrmals gebraucht wird, um zu  $B$  zu gelangen, ohne sich dabei genau festlegen zu müssen, wie oft  $A$  tatsächlich gebraucht wird?

Man betrachte die Reihe:

$$(A \rightarrow B) \vee (A \square A \rightarrow B) \vee (A \square A \square A \rightarrow B) \vee \dots,$$

wobei  $\square$  eine nicht-kontrahierende Konjunktion ist, d.h. es gilt nicht  $A \rightarrow A \square A$ . Gesucht ist eine Möglichkeit,  $A \square^n A$  für eine gewisse Form von natürlichen Zahlen  $n$  so ausdrücken zu können, daß

$$A \square^0 A \leftrightarrow A$$

$$A \square^1 A \leftrightarrow A \square A$$

$$A \square^2 A \leftrightarrow A \square A \square A$$

usw. gilt, um schließlich so etwas wie

$$(12.1) \quad (x)(x \in \mathbf{N} \rightarrow (A \square^x A))$$

oder

$$(12.2) \quad (Ex)(x \in \mathbf{N} \square (A \square^x A \rightarrow B))$$

<sup>100</sup> Wer sich an die lineare Logik aus der theoretischen Informatik erinnert fühlt, möge sich durch Fußnote 100 auf S. 143 bestätigt sehen.

formulieren zu können. Die intuitive Symbolik  $\Box^x$  muß präzise gefaßt werden, deshalb haben auch die Quantoren hier eine intuitive Gestalt ( $x$ ) und ( $Ex$ ), damit sie nicht mit den Symbolen eines formalen Systems verwechselt werden.

Dafür müssen eine paar neue Definitionen eingeführt werden. Es geht darum, ausdrücken zu können, daß eine Formel noch einmal zur Verfügung steht. Dazu muß man aber in der Lage sein, über die Gültigkeit von Formeln reden zu können, und dazu, wiederum, braucht man so etwas wie eine *Wahrheitsdefinition*. In einer typenfreien Logik erhält man das leicht:

$$\lambda A \in \{\mathcal{V}\} \leftrightarrow A.$$

Natürlich hat diese Wahrheitsdefinition nichts mit Wahrheit zu tun.<sup>101</sup> Alles was ich damit vorhabe, beruht auf den folgenden Festsetzungen:

$$I := \{\mathcal{V}\}, \text{ und} \\ s^I := s \Box I.$$

Sie versetzen mich in die Lage, so etwas auszudrücken wie:

$$\lambda A \in I \leftrightarrow A, \\ \lambda A \in I^I \leftrightarrow A \Box A, \\ \lambda A \in I^{II} \leftrightarrow A \Box A \Box A,$$

usw. Das bringt mich meinem Ziel näher, wenn es mir gelingt, den Ausdruck  $A \in I^{IIII}$ , zum Beispiel, mit „5 mal  $A$ “ zu verknüpfen.

Zu diesem Zweck definiere ich zunächst auf der Metaebene induktiv die Menge  $\Psi$ :

$I$  ist ein Element von  $\Psi$ .

Wenn  $x$  ein Element von  $\Psi$  ist, dann ist auch  $x^I$  ein Element von  $\Psi$ .

Dann stellt eine weitere induktive Definition eine Korrespondenz her zwischen natürlichen Zahlen und Elementen von  $\Psi$ :

<sup>101</sup> Ich will ja gar nicht behaupten, daß Tarskis sogenannte *Wahrheitsdefinition* ein Etikettenschwindel ist, aber es muß klar sein, daß ich nicht viel Aufhebens mache, wenn es darum geht, auf der formalen Ebene eine allgemeine Äquivalenz zwischen einer Aussage  $A$  und einer Aussage der Form  $s_A \in t$  herzustellen, wobei  $t$  ein fester Term ist und die Gestalt des Terms  $s_A$  von  $A$  abhängt. Es ist jedoch zu beachten, daß Tarskis Vorgehen rekursiv ist; ansonsten aber ist es so unphilosophisch, wie das, was ich hier eingeführt habe.

$I$  ist das korrespondierende  $\Psi$ -Element zu 1.

Wenn  $\tilde{n}$  das korrespondierende  $\Psi$ -Element zu der natürlichen Zahl  $n$  ist, dann ist  $\tilde{n}^I$  das korrespondierende  $\Psi$ -Element zu der natürlichen Zahl  $n'$ .

Jetzt geht es natürlich darum, diese Menge  $\Psi$  auf der Objektebene repräsentieren zu können — und da kommt das Problem mit der Unendlichkeit herein.

Wenn ich sage, „da kommt das Problem mit der Unendlichkeit herein“, so bedeutet das nicht, daß es ein Problem wäre, einen Term zu formulieren, der alle Elements von  $\Psi$  als Elemente enthält. Das Problem der Unendlichkeit besteht darin, zeigen zu können, daß ein Term nur diese Objekte als Elemente enthält.

Die grundsätzliche Idee für meinen Ansatz geht zurück auf den Beweis der aussagenlogischen Unvollständigkeit in Schüttes [50], S. 238. Schütte zeigt auf einfache Weise, daß in einer schnittfreien Logik mit den üblichen Regeln für die uneingeschränkte Abstraktion die Formel

$$\lambda x(x \in x) \in \lambda x(x \in x)$$

sofort als unentscheidbar eingesehen werden kann. Wegen der Schnitteliminierbarkeit kann diese Formel nur durch einen  $\in$ -Schluß links oder rechts gewonnen worden sein. In beiden Fällen aber führt diese Formel nur auf sich selbst zurück:

$$\frac{\lambda x(x \in x) \in \lambda x(x \in x) \Rightarrow}{\lambda x(x \in x) \in \lambda x(x \in x) \Rightarrow} \quad \frac{\Rightarrow \lambda x(x \in x) \in \lambda x(x \in x)}{\Rightarrow \lambda x(x \in x) \in \lambda x(x \in x)} .$$

Eine kleine Variation dieser Überlegung führt nun zu meinem Ansatz. Und zwar bediene ich mich der folgenden Formel:

$$\lambda x((x \in x) \square A) \in \lambda x((x \in x) \square A) .$$

Diese Formel absorbiert, sozusagen, Antezedentformeln (also Annahmen) der Gestalt  $A$ :

$$\frac{\lambda x((x \in x) \square A) \in \lambda x((x \in x) \square A), A \Rightarrow}{(\lambda x((x \in x) \square A) \in \lambda x((x \in x) \square A)) \square A \Rightarrow} \quad \frac{\lambda x((x \in x) \square A) \in \lambda x((x \in x) \square A) \Rightarrow}{\lambda x((x \in x) \square A) \in \lambda x((x \in x) \square A) \Rightarrow} .$$

Auf der rechten Seite

$$\begin{array}{l} \Rightarrow \lambda x((x \in x) \sqsupset A) \in \lambda x((x \in x) \sqsupset A) \quad \Rightarrow A \\ \hline \Rightarrow (\lambda x((x \in x) \sqsupset A) \in \lambda x((x \in x) \sqsupset A)) \sqsupset A \\ \hline \Rightarrow \lambda x((x \in x) \sqsupset A) \in \lambda x((x \in x) \sqsupset A) \end{array}$$

ist die Situation im wesentlichen wie im Fall von Schüttes Formel: Man muß die Formel schon bewiesen haben, um sie beweisen zu können (immer Schnitteliminierbarkeit vorausgesetzt).

Um eine einfache, gängige Schreibweise zu haben, definiere ich:

$$\check{\gamma}[A] := \lambda x((x \in x) \sqsupset A) \in \lambda x((x \in x) \sqsupset A).$$

Dann läßt sich die Formel

$$(\lambda x((x \in x) \sqsupset A) \in \lambda x((x \in x) \sqsupset A)) \leftrightarrow (\lambda x((x \in x) \sqsupset A) \in \lambda x((x \in x) \sqsupset A)) \sqsupset A$$

einfach als

$$\check{\gamma}[A] \leftrightarrow \check{\gamma}[A] \sqsupset A$$

schreiben und man erkennt deutlich den Fixpunktcharakter. Ein Schlußschema der Gestalt

$$\frac{\check{\gamma}[A], A, \dots, A, \Gamma \Rightarrow}{\check{\gamma}[A], \Gamma \Rightarrow}$$

läßt sich leicht (durch eine Induktion nach der Anzahl der Vorkommnisse von  $A$ ) als  $\mathbf{L}^1\mathbf{D}_\lambda$ -herleitbar nachweisen.

Nun kann man einen Term formulieren, der nachweislich alle natürlichen Zahlen als Elemente enthält:

$$N_\gamma := \lambda x \wedge y(0 \in y \sqsupset \check{\gamma}[\wedge z(z \in y \rightarrow z' \in y)] \rightarrow x \in y).$$

Das reicht aber nicht, um ein Induktionsschema der folgenden Art

$$\frac{\Rightarrow \mathfrak{F}[0] \quad \mathfrak{F}[a] \Rightarrow \mathfrak{F}[a']}{s \in N_\gamma \Rightarrow \mathfrak{F}[s]}$$

beweisen zu können, aus dem einfachen Grund, weil die rechte Obersequenz in der folgenden Schlußfigur keine Aussicht hat, jemals bewiesen werden zu können:

$$\frac{\Rightarrow 0 \in \lambda x \mathfrak{F}[x] \quad \Rightarrow \check{\gamma}[\wedge z(z \in \lambda x \mathfrak{F}[x] \rightarrow z' \in \lambda x \mathfrak{F}[x])]}{s \in \lambda x \wedge y(0 \in y \sqsupset \check{\gamma}[\wedge z(z \in y \rightarrow z' \in y)] \rightarrow x \in y) \Rightarrow \mathfrak{F}[s]}.$$

In Anbetracht der obigen Überlegungen besteht keine Aussicht, jemals in der Lage zu kommen,  $\check{\gamma}[A]$  (in  $\mathbf{L}^1\mathbf{D}_\lambda$  oder einer sinnvollen Erweiterung davon) für irgendeine Formel  $A$  beweisen zu können. Aber die Situation

ist nicht ganz hoffnungslos. Es mag nicht möglich sein, die Gesamtheit der natürlichen Zahlen auf diese Weise direkt so zu definieren, daß gewisse Induktionsregeln bewiesen werden können, aber was möglich ist, ist, das  $N_\gamma$  zugrundeliegende Schema für die Definition eines Zwischenterms heranzuziehen, welcher dazu dient eine Implikation auszudrücken, in der das Vorderglied beliebig oft gebraucht werden kann, etwa in der Form

$$A \square \dots \square A \rightarrow B,$$

d.h. es gibt eine Anzahl von Anwendungen von  $A$  als Annahme, die  $B$  impliziert, also eine *schwache Implikation*.

Anders gesagt, mit  $\check{\gamma}[A]$  hat man zwar noch keine Induktion, aber einen wesentlichen Baustein dazu: die Möglichkeit auszudrücken, daß der Induktionsschritt beliebig oft zur Verfügung steht, d.h. einen Ersatz für Zusammenziehungen einer bestimmten Form, erzeugt mithilfe der Selbstbezüglichkeit. Das ist aber nicht, wie im Fall der natürlichen Zahlen, ein Term, sondern eine Formel. Zu diesem Zweck definiere ich den Term

$$\mathbf{Z} := \lambda x \wedge y (I \in y \square \check{\gamma}[\wedge z (z \in y \rightarrow z^I \in y)] \rightarrow x \in y).$$

Für  $\mathbf{Z}$  läßt sich nun zeigen, daß

$$\begin{aligned} &\Rightarrow I \in \mathbf{Z}, \text{ und} \\ &s \in \mathbf{Z} \Rightarrow s^I \in \mathbf{Z} \end{aligned}$$

in  $\mathbf{L}^I \mathbf{D}_\lambda$  beweisbar sind.<sup>102</sup> Somit gilt: wenn  $s$  ein Element von  $\Psi$ , so ist  $s \in \mathbf{Z}$  in  $\mathbf{L}^I \mathbf{D}_\lambda$  beweisbar.

Was aber von größerer Bedeutung ist, ist die umgekehrte Richtung: Wenn  $\Rightarrow s \in \mathbf{Z}$  in  $\mathbf{L}^I \mathbf{D}_\lambda$  beweisbar ist, dann ist  $s$  ein Element von  $\Psi$ .

Das läßt sich durch folgende Reihe von Überlegungen wahrscheinlich machen.<sup>103</sup>

Wenn  $\Rightarrow s \in \mathbf{Z}$  in  $\mathbf{L}^I \mathbf{D}_\lambda$  beweisbar ist, so ist es sicher auch

$$I \in b, \check{\gamma}[\wedge z (z \in b \rightarrow z^I \in b)] \Rightarrow s \in b,$$

und zwar aufgrund einer Eliminierung von logischen Konstanten. Wegen der Schnitteliminierbarkeit muß damit auch

$$I \in b, \wedge z (z \in b \rightarrow z^I \in b), \check{\gamma}[\wedge z (z \in b \rightarrow z^I \in b)] \Rightarrow s \in b$$

<sup>102</sup> Vgl. [36], S. 388.

<sup>103</sup> Eine etwas ausführlichere Beweisskizze findet sich [36], S. 391. Ich will aber nicht verhehlen, daß der Beweis, auch wie er in [37], Abschnitt 132a, dargelegt ist, Lücken aufweist.

$\mathbf{L}^1\mathbf{D}_\lambda$ -beweisbar sein. Das läßt sich dahingehend verallgemeinern, daß es eine natürliche Zahl  $n$  derart geben muß, daß

$$I \in b, [\bigwedge z (z \in b \rightarrow z^I \in b)]^n \Rightarrow s \in b$$

$\mathbf{L}^1\mathbf{D}_\lambda$ -beweisbar ist, was wiederum darauf reduziert werden kann, daß es eine Reihe von Termen  $\xi_1, \dots, \xi_n$  derart gibt, daß

$$I \in b, \xi_1 \in b \rightarrow \xi_1^I \in b, \dots, \xi_n \in b \rightarrow \xi_n^I \in b \Rightarrow s \in b$$

in  $\mathbf{L}^1\mathbf{D}_\lambda$  beweisbar ist. Aufgrund der Irreduzibilität von Formeln der Gestalt  $s \in b$  ( $b$  ist eine freie Variable) läßt sich sagen, daß jedes der  $\xi_i$  ( $i \in \{1, \dots, n\}$ ) mit einem Element aus  $\{I, \dots, \tilde{n}\}$  identisch sein muß. Damit erhält man dann, daß  $s \in \{I, II, \dots, \tilde{n}\}$  sein muß, d.h.  $s \in \Psi$ .

### 13. Z-Schlüsse, Notwendigkeit und schwache Implikation

Wir haben festgestellt, daß  $\Rightarrow s \in \mathbf{Z}$  nur dann  $\mathbf{L}^1\mathbf{D}_\lambda$ -beweisbar ist, wenn  $s \in \Psi$ . Wenn wir nun  $[A/s]$  für  $\lambda A \in s$  schreiben, folgt daraus, daß die Schlußregel

$$\frac{\Rightarrow s \in \mathbf{Z} \quad \Rightarrow A}{\Rightarrow [A/s]}$$

in  $\mathbf{L}^1\mathbf{D}_\lambda$  zulässig ist, d.h. zu den Grundslußregeln hinzugenommen werden kann, ohne daß neue Formeln beweisbar werden. Das sieht man folgendermaßen. Wenn  $\Rightarrow s \in \mathbf{Z}$  in  $\mathbf{L}^1\mathbf{D}_\lambda$  beweisbar ist, dann ist  $s$  ein Element von  $\Psi$ , d.h.  $s$  ist entweder  $I$  oder es gibt ein Element  $t$  von  $\Psi$  derart, daß  $s \equiv t^I$ . Die Behauptung folgt durch eine meta-theoretische Induktion:

Induktionsbasis:

$$\frac{\Rightarrow A}{\Rightarrow [A/I]}$$

Induktionsschritt: die Induktionsvoraussetzung ist

$$\frac{\Rightarrow t \in \mathbf{Z} \quad \Rightarrow A}{\Rightarrow [A/t]},$$

d.h. man erhält  $\Rightarrow [A/t]$ . Da außerdem

$$\frac{\Rightarrow [A/t] \quad \Rightarrow A}{\Rightarrow [A/t^I]}$$

gilt, erhält man  $\Rightarrow [A/s]$  mit  $s \equiv t^I$ .



Jetzt kommt der spekulative Schritt.

Was wir festgestellt haben, ist, daß aus der Beweisbarkeit von  $s \in \mathbf{Z}$  und  $A$  auf die Beweisbarkeit von  $\lambda A \in s$  geschlossen werden kann. Was ist aber, wenn  $s \in \mathbf{Z}$  nur unter Annahmen gezeigt werden kann, d.h. wenn nur  $\Gamma \Rightarrow s \in \mathbf{Z}$  (für nichtleeres  $\Gamma$ ) beweisbar ist? Kann man dann das Schema

$$(Z) \quad \frac{\Gamma \Rightarrow s \in \mathbf{Z} \quad \Rightarrow A}{\Gamma \Rightarrow [A/s]}$$

widerspruchsfrei als eine neue Grundschlußregel zu  $\mathbf{I}^1\mathbf{D}_\lambda$  hinzunehmen? Mit dieser Frage begeben wir uns in das eigentliche Gebiet der spekulativen Logik: synthetisch-apriorische Erkenntniserweiterung.

Das System, das aus  $\mathbf{I}^1\mathbf{D}_\lambda$  durch Hinzunahme von  $\mathbf{Z}$  als Grundschlußregel entsteht, soll  $\mathbf{I}^1\mathbf{D}_\lambda^Z$  heißen. Was nun gezeigt werden kann, ist, daß alle Anwendungen der  $\mathbf{Z}$ -Schlußregel in  $\mathbf{I}^1\mathbf{D}_\lambda^Z$ -Beweisen von Endsequenzen der Form  $\Rightarrow s \in \mathbf{Z}$  eliminiert werden können, d.h., wenn  $\Rightarrow s \in \mathbf{Z}$  mit  $\mathbf{Z}$ -Schlüssen  $\mathbf{I}^1\mathbf{D}_\lambda^Z$ -beweisbar ist, so ist es auch ohne  $\mathbf{Z}$ -Schlüsse, d.h.  $\mathbf{I}^1\mathbf{D}_\lambda$ -beweisbar.<sup>104</sup> Mittels Kontraposition erhalten wir dann: Wenn die Sequenz  $\Rightarrow s \in \mathbf{Z}$  nicht in  $\mathbf{I}^1\mathbf{D}_\lambda$  (also ohne  $\mathbf{Z}$ -Schlüsse) beweisbar ist, dann ist sie es auch mit  $\mathbf{Z}$ -Schlüssen nicht. Da die Sequenz  $\Rightarrow 0 \in \mathbf{Z}$  in  $\mathbf{I}^1\mathbf{D}_\lambda$  nicht beweisbar ist, ist sie es also auch nicht mit  $\mathbf{Z}$ -Schlüssen. Damit ist gezeigt, daß  $\mathbf{Z}$ -Schlüsse widerspruchsfrei zu  $\mathbf{I}^1\mathbf{D}_\lambda$  hinzugenommen werden können. Es ist eine rein meta-logische Einsicht, die hier zur Anwendung kommt.

Nun können wir das Problem der Formulierung eines Terms angehen, der uns eine Form der Induktion liefert. Wir ziehen  $\mathbf{Z}$  zur Definition des folgenden Terms heran:

$$\check{\mathbf{I}}^\circ := \lambda x (x \in \mathbf{Z} \square \wedge y ([I \in y \wedge \wedge z (z \in y \rightarrow z^I \in y) / x] \rightarrow x \in y)).$$

Offensichtlich gilt wie für  $\mathbf{Z}$ , daß

$$\begin{aligned} \Rightarrow I \in \check{\mathbf{I}}^\circ, \quad \text{und} \\ s \in \check{\mathbf{I}}^\circ \Rightarrow s^I \in \check{\mathbf{I}}^\circ \end{aligned}$$

$\mathbf{I}^1\mathbf{D}_\lambda$ -beweisbar sind.  $\check{\mathbf{I}}^\circ$  erlaubt nun eine Form der Induktion, deren Gegenstand allerdings nicht Zahlen sind, sondern Elemente von  $\mathbf{Z}$ , d.h.  $I, I^I, I^{II}, \dots$ :

<sup>104</sup> Der Beweis ist detaillierter ausgeführt in [37], Abschnitt 133b. Ich fürchte aber, daß der Beweis auch dort noch unzureichend ist.

$$(13.1) \quad \frac{\Gamma \Rightarrow \mathfrak{F}[I] \quad \mathfrak{F}[a] \Rightarrow \mathfrak{F}[a^I]}{s \in \check{\mathbf{\Pi}}^\circ, \Gamma \Rightarrow \mathfrak{F}[s]} .$$

Es sei in der folgenden Beweisfigur  $\xi := \lambda x_1 (C \rightarrow \mathfrak{F}[x_1])$ :

$$\frac{\frac{\frac{\frac{\Gamma \Rightarrow \mathfrak{F}[I]}{\underline{\underline{C \Rightarrow \mathfrak{F}[I]}}} \Rightarrow C \rightarrow \mathfrak{F}[I]}{\Rightarrow I \in \xi} \quad \frac{\frac{\frac{C \Rightarrow C \quad \mathfrak{F}[a] \Rightarrow \mathfrak{F}[a^I]}{C \rightarrow \mathfrak{F}[a], C \Rightarrow \mathfrak{F}[a^I]} \quad C \rightarrow \mathfrak{F}[a] \Rightarrow C \rightarrow \mathfrak{F}[a^I]}{\underline{\underline{a \in \xi \Rightarrow a^I \in \xi}}} \Rightarrow a \in \xi \rightarrow a^I \in \xi}{\Rightarrow \bigwedge z (z \in \xi \rightarrow z^I \in \xi)} \quad \frac{\Gamma \Rightarrow C \quad \mathfrak{F}[s] \Rightarrow \mathfrak{F}[s]}{C \rightarrow \mathfrak{F}[s], \Gamma \Rightarrow \mathfrak{F}[s]}}{\frac{\Rightarrow I \in \xi \wedge \bigwedge z (z \in \xi \rightarrow z^I \in \xi)}{s \in \mathbf{Z} \Rightarrow [I \in \xi \wedge \bigwedge z (z \in \xi \rightarrow z^I \in \xi) / s]} \quad \frac{\Gamma \Rightarrow \mathfrak{F}[s]}{s \in \xi, \Gamma \Rightarrow \mathfrak{F}[s]}}{\frac{s \in \mathbf{Z}, [I \in \xi \wedge \bigwedge z (z \in \xi \rightarrow z^I \in \xi) / s] \rightarrow s \in \xi, \Gamma \Rightarrow \mathfrak{F}[s]}{s \in \mathbf{Z}, \bigwedge y ([I \in y \wedge \bigwedge z (z \in y \rightarrow z^I \in y) / s] \rightarrow s \in y), \Gamma \Rightarrow \mathfrak{F}[s]} \quad \frac{\Gamma \Rightarrow \mathfrak{F}[s]}{s \in \mathbf{Z} \square \bigwedge y ([I \in y \wedge \bigwedge z (z \in y \rightarrow z^I \in y) / s] \rightarrow s \in y), \Gamma \Rightarrow \mathfrak{F}[s]}}{s \in \check{\mathbf{\Pi}}^\circ, \Gamma \Rightarrow \mathfrak{F}[s]} .$$

Diese Induktion erlaubt noch keine Seitenformeln, aber sie reicht aus, Zusammenziehungen für Formeln der Gestalt  $s \in \check{\mathbf{\Pi}}^\circ$  zu beweisen:

$$(13.2) \quad \frac{\Rightarrow I \in \check{\mathbf{\Pi}}^\circ \square I \in \check{\mathbf{\Pi}}^\circ \quad a \in \check{\mathbf{\Pi}}^\circ \square a \in \check{\mathbf{\Pi}}^\circ \Rightarrow a^I \in \check{\mathbf{\Pi}}^\circ \square a^I \in \check{\mathbf{\Pi}}^\circ}{s \in \check{\mathbf{\Pi}}^\circ \Rightarrow s \in \check{\mathbf{\Pi}}^\circ \square s \in \check{\mathbf{\Pi}}^\circ} .$$

Das erlaubt zumindest eine kleine Verstärkung von 13.1:

$$(13.3) \quad \frac{\Gamma \Rightarrow \mathfrak{F}[I] \quad \mathfrak{F}[a], a \in \check{\mathbf{\Pi}}^\circ \Rightarrow \mathfrak{F}[a^I]}{s \in \check{\mathbf{\Pi}}^\circ, \Gamma \Rightarrow \mathfrak{F}[s]} .$$

Mit dieser Form der Induktion, die ich auch *Protoinduktion* genannt habe,<sup>105</sup> lassen sich nun die intuitiven Formulierungen 12.1 und 12.2 vom vorigen Abschnitt formal fassen. Ich beginne mit einem Begriff der Notwendigkeit.

<sup>105</sup> Z.B. in [36], S. 392.

Die Ausdrucksfähigkeit der uneingeschränkten Abstraktion gestattet die Formulierung eines Begriffs der Notwendigkeit, der die Intuition von 12.1 symbolisch erfaßt durch

$$(13.4) \quad \Box A := \bigwedge x(x \in \mathbf{Z} \rightarrow [A/x]),$$

in Worten etwa:  $\lambda A \in x$  für jedes Element  $x$  von  $\mathbf{Z}$ ; d.h. es sieht fast so aus, als hätte ich erreicht, was ich oben auf S. 143 als Ziel erklärt habe: eine Möglichkeit, über Annahmen gleicher Gestalt zu quantifizieren.

Tatsächlich läßt sich einiges, aber nicht alles, Wünschenswerte zeigen. Zunächst einmal kann  $\Box A \Rightarrow A$  schon ohne  $\mathbf{Z}$ -Schlüsse bewiesen werden:

$$\frac{\frac{\frac{A \Rightarrow A}{[A/I] \Rightarrow A}}{\Rightarrow I \in \mathbf{Z}}}{I \in \mathbf{Z} \rightarrow [A/I] \Rightarrow A} \cdot \frac{}{\bigwedge x(x \in \mathbf{Z} \rightarrow [A/x]) \Rightarrow A}.$$

Tatsächlich läßt sich für jede natürliche Zahl  $\Box A \Rightarrow A \square^n A$  in  $\mathbf{I}^1 \mathbf{D}_\lambda$  beweisen.

Mit dem Einsatz von  $\mathbf{Z}$ -Schlüssen läßt sich weiterhin zeigen, daß Schlüsse nach dem Schema

$$\frac{\Rightarrow A}{\Rightarrow \Box A}$$

herleitbar sind:

$$\frac{\frac{\frac{\Rightarrow A}{a \in \mathbf{Z} \Rightarrow [A/a]}}{\Rightarrow a \in \mathbf{Z} \rightarrow [A/a]}}{\Rightarrow \bigwedge x(x \in \mathbf{Z} \rightarrow [A/x])}.$$

Dann wird aber deutlich, daß wir mit der Definition 13.4 spätestens dann nicht weiter kommen, wenn es um das sogenannte K-Axiom geht:  $\Box(A \rightarrow B), \Box A \Rightarrow \Box B$ . Das sollte aus folgender Reduktion klar werden:

$$\frac{\frac{\frac{a \in \mathbf{Z} \rightarrow [A \rightarrow B/a], a \in \mathbf{Z} \rightarrow [A/a], a \in \mathbf{Z} \Rightarrow [B/a]}{\bigwedge x(x \in \mathbf{Z} \rightarrow [A \rightarrow B/x]), \bigwedge x(x \in \mathbf{Z} \rightarrow [A/x]), a \in \mathbf{Z} \Rightarrow [B/a]}}{\bigwedge x(x \in \mathbf{Z} \rightarrow [A \rightarrow B/x]), \bigwedge x(x \in \mathbf{Z} \rightarrow [A/x]) \Rightarrow a \in \mathbf{Z} \rightarrow [B/a]}}{\bigwedge x(x \in \mathbf{Z} \rightarrow [A \rightarrow B/x]), \bigwedge x(x \in \mathbf{Z} \rightarrow [A/x]) \Rightarrow \bigwedge x(x \in \mathbf{Z} \rightarrow [B/x])}.$$

Eine Herleitung der obersten Sequenz kann  $\mathbf{Z}$  nicht bieten. Das Problem ist, daß  $\mathbf{Z}$  noch keine Induktion gestattet und auch keine Zusammenziehungen für Formeln der Gestalt  $s \in \mathbf{Z}$ , wie es in 13.2 für Formeln der Gestalt  $s \in \check{\mathbf{\Pi}}^\circ$  formuliert wurde. Deshalb wende ich mich folgender Definition zu:

$$\square A := \bigwedge x (x \in \check{\mathbf{\Pi}}^\circ \rightarrow [A/x]),$$

die vielversprechender ist, weil sie anstelle von  $\mathbf{Z}$  den Term  $\check{\mathbf{\Pi}}^\circ$  hat, der, wie wir gesehen haben, durchaus Induktionen ermöglicht.

Damit können wir nun tatsächlich das K-Axiom beweisen. Anstelle von

$$a \in \mathbf{Z} \rightarrow [A \rightarrow B/a], a \in \mathbf{Z} \rightarrow [A/a], a \in \mathbf{Z} \Rightarrow [B/a]$$

in der vorgehenden Reduktion erhalten wir nun

$$a \in \check{\mathbf{\Pi}}^\circ \rightarrow [A \rightarrow B/a], a \in \check{\mathbf{\Pi}}^\circ \rightarrow [A/a], a \in \check{\mathbf{\Pi}}^\circ \Rightarrow [B/a]$$

und können folgendermaßen weiter reduzieren:

$$\frac{\frac{a \in \check{\mathbf{\Pi}}^\circ \Rightarrow a \in \check{\mathbf{\Pi}}^\circ \quad [A \rightarrow B/a], [A/a], a \in \check{\mathbf{\Pi}}^\circ \Rightarrow [B/a]}{a \in \check{\mathbf{\Pi}}^\circ \rightarrow [A \rightarrow B/a], a \in \check{\mathbf{\Pi}}^\circ \rightarrow [A/a], a \in \check{\mathbf{\Pi}}^\circ, a \in \check{\mathbf{\Pi}}^\circ, a \in \check{\mathbf{\Pi}}^\circ \Rightarrow [B/a]}}{a \in \check{\mathbf{\Pi}}^\circ \rightarrow [A \rightarrow B/a], a \in \check{\mathbf{\Pi}}^\circ \rightarrow [A/a], a \in \check{\mathbf{\Pi}}^\circ \Rightarrow [B/a]}.$$

Was bleibt, ist nun eine Protoinduktion mit dem linken Ast:

$$\frac{\frac{A \rightarrow B, A \Rightarrow B}{[A \rightarrow B/I], [A/I] \Rightarrow [B/I]}}{[A \rightarrow B/I] \square [A/I] \Rightarrow [B/I]} \Rightarrow [A \rightarrow B/I] \square [A/I] \rightarrow [B/I]$$

und dem rechten Ast:

$$\frac{\frac{[B/a] \Rightarrow [B/a] \quad A \rightarrow B, A \Rightarrow B}{[B/a], A \rightarrow B, A \Rightarrow [B/a] \square B}}{[A \rightarrow B/a], [A/a] \Rightarrow [A \rightarrow B/a] \square [A/a] \quad [B/a], A \rightarrow B, A \Rightarrow [B/a^I]}}{[A \rightarrow B/a] \square [A/a] \rightarrow [B/a], [A \rightarrow B/a], A \rightarrow B, [A/a], A \Rightarrow [B/a^I]}}{\frac{[A \rightarrow B/a] \square [A/a] \rightarrow [B/a], [A \rightarrow B/a^I], [A/a^I] \Rightarrow [B/a^I]}{[A \rightarrow B/a] \square [A/a] \rightarrow [B/a], [A \rightarrow B/a^I] \square [A/a^I] \Rightarrow [B/a^I]}}{[A \rightarrow B/a] \square [A/a] \rightarrow [B/a] \Rightarrow [A \rightarrow B/a^I] \square [A/a^I] \rightarrow [B/a^I]}.$$

Mit anderen Worten, wir sind bei einem Begriff der T-Modalität angelangt — und das ohne irgendein Gerede von „möglichen Welten“, wie es in der analytischen Philosophie üblich ist, auch wenn es sich dabei nur um algebraische Modelle handelt. Bei dem Begriff der Modalität, den ich hier eingeführt habe, werden keinerlei Modelle zuhulfe genommen; es handelt sich um eine rein syntaktisch eingeführte theoretische Konstante, wie es auch die  $\Box$ -Konjunktion selbst ist, als deren Verallgemeinerung die  $\Box$ -Notwendigkeit aufgefaßt werden, so wie der  $\bigwedge$ -Quantor eine Verallgemeinerung der  $\wedge$ -Konjunktion darstellt.

Das Verhältnis des Notwendigkeitsoperators zur klassischen Logik wird durch die  $\mathbf{LD}_\lambda^Z$ -Beweisbarkeit der beiden folgenden Sequenzen treffend charakterisiert:

$$\begin{aligned} \Box(A \vee \neg A), A &\Rightarrow \Box A, \\ \Box(A \vee \neg A), \Box A \rightarrow B &\Rightarrow A \rightarrow B. \end{aligned}$$

D.h. modale Unterscheidungen brechen für klassisch beweisbare Formeln zusammen.

Auf entsprechende Weise läßt sich nun auch 12.2 (schwache Implikation) auf der formalen Ebene erfassen:

$$A \supset B \equiv \forall x (x \in \check{\mathbf{I}}^\circ \Box ([A/x] \rightarrow B)).$$

Damit erhält man dann beispielsweise:

$$\begin{aligned} A \rightarrow B &\Rightarrow A \supset B, \\ A \supset B, A \supset (B \rightarrow C) &\Rightarrow A \supset C, \\ \Box(A \vee \neg A), A \supset B &\Rightarrow A \rightarrow B, \\ A \supset B, \Box((A \rightarrow B) \rightarrow B) &\Rightarrow B, \\ A \supset B, \Box((A \rightarrow B) \rightarrow (C \rightarrow B)) &\Rightarrow C \supset B, \end{aligned}$$

und Schlüsse nach folgenden Schemata sind direkt ableitbar:

$$\frac{A, \dots, A, \Gamma \Rightarrow B}{\Gamma \Rightarrow A \supset B}, \quad \frac{\Gamma \Rightarrow A \quad \Pi \Rightarrow A \supset B}{\Box \Gamma, \Pi \Rightarrow B}, \quad \frac{\Gamma \Rightarrow A \quad B, \Pi \Rightarrow C}{A \supset B, \Box \Gamma, \Pi \Rightarrow C},$$

$$\frac{(A \rightarrow B) \Rightarrow (C_1 \rightarrow (\dots \rightarrow (C_n \rightarrow B) \dots))}{(A \supset B) \Rightarrow (C_1 \supset (\dots \supset (C_n \supset B) \dots))}.$$

## 14. Beschreibung neu gefaßt

Die Fixpunkteigenschaft für Termformen ist unverträglich mit dem bestimmten Artikel in seiner klassischen axiomatischen Form. Das ist die eine Seite. Die andere ist, daß die Ausdruckskraft der mengentheoretischen Komprehension bzw. der Abstraktion im Rahmen einer klassischen Mengenlehre die Formulierung einer ganzen Reihe von Termen erlaubt, die für den Beweis des  $\iota$ -Axioms ausreichen.<sup>106</sup> Angesichts der Widerspruchsfreiheit von  $\mathbf{L}\mathbf{D}_\lambda$  wird klar sein, daß da irgendwo eine Zusammenziehung versteckt sein muß, die die völlige Übertragung des klassischen Ansatzes vereitelt. Damit ergibt sich aber die Frage, wo genau geht ein Beweis des  $\iota$ -Axioms mit diesen Definitionen nicht mehr durch und ist es möglich, für die anfallenden Zusammenziehungen einen Ersatz zu finden bzw. sie in anderer Weise zu integrieren, um ein modifiziertes  $\iota$ -Axiom zu erhalten?

Natürlich ist es möglich, daß eine Antwort auf diese Frage von der besonderen Gestalt des Terms abhängt, der zur Darstellung des bestimmten Artikels in der höheren Logik hergenommen wird. Quine in [43], S. 147, gibt dem Term

$$\lambda x \forall y (x \in y \square \bigwedge z (z = y \leftrightarrow \mathfrak{F}[z])),$$

der die „waste cases“ einheitlich mit  $\emptyset$  gleichsetzt, den Vorzug. Es ist mir nicht gelungen, 14.5 bei Zugrundelegung dieses Terms ohne den Gebrauch einer Form der Extensionalität oder Zusammenziehung zu beweisen. Deshalb wäre es zumindest nötig, auf einen anderen Term wie etwa

$$\lambda x \forall y (x \in y \square (\mathfrak{F}[y] \wedge \bigwedge z (\mathfrak{F}[z] \rightarrow y = z)))$$

auszuweichen, der mit Sicherheit weder Extensionalität noch Zusammenziehung erfordert und ebenfalls die „waste cases“ einheitlich mit  $\emptyset$  gleichsetzt. Da ich aber nicht an einer einheitlichen Behandlung der „waste cases“ interessiert bin, zumindest jetzt noch nicht, bin ich besser dran, wenn ich die *Eindeutigkeitsbedingung*  $\bigwedge z (\mathfrak{F}[z] \rightarrow y = z)$  ganz fallen lasse und die folgende einfache Definition heranziehe, die auch schon in Quine [43], S. 149, zu finden ist:

$$\iota x \mathfrak{F}[x] := \lambda x \forall y (x \in y \square \mathfrak{F}[y]).$$

Unter klassischen Bedingungen, d.h. bei Verfügbarkeit von Zusammenziehungen, erhält man mit dieser Definition das  $\iota$ -Axiom auf einfache

<sup>106</sup> Cf. [43], S. 147 ff. Der Unterschied zwischen den verschiedenen Formulierungen ist die Behandlung der „waste cases“.

Weise. Aber auch ohne Zusammenziehungen kommt man noch ein ganzes Stück weit. Ich führe eine Reihe wichtiger Schritte auf.

$$(14.1) \quad s \in \iota x \mathfrak{F}[x], \mathfrak{F}[t], \wedge z_1 \wedge z_2 (\mathfrak{F}[z_1] \square \mathfrak{F}[z_2] \rightarrow z_1 = z_2) \Rightarrow s \in t,$$

$$(14.2) \quad s \in t, \mathfrak{F}[t] \Rightarrow s \in \iota x \mathfrak{F}[x],$$

$$(14.3) \quad \mathfrak{F}[t], \wedge z_1 \wedge z_2 (\mathfrak{F}[z_1] \square \mathfrak{F}[z_2] \rightarrow z_1 = z_2) \Rightarrow \iota x \mathfrak{F}[x] = t,$$

$$(14.4) \quad s \in \iota x \mathfrak{F}[x], \wedge y (\mathfrak{F}[y] \rightarrow t = y) \Rightarrow s \in t,$$

$$(14.5) \quad \mathfrak{F}[t], \wedge y (\mathfrak{F}[y] \rightarrow t = y) \Rightarrow \iota x \mathfrak{F}[x] = t,$$

$$(14.6) \quad s \in t, \wedge y (\mathfrak{F}[y] \leftrightarrow t = y) \Rightarrow s \in \iota x \mathfrak{F}[x],$$

$$(14.7) \quad \wedge y (\mathfrak{F}[y] \leftrightarrow t = y) \Rightarrow \iota x \mathfrak{F}[x] = t.$$

Jede dieser Sequenzen kann leicht bewiesen werden. Aber da Zusammenziehungen tückisch sein können, will ich hier die beiden wichtigsten, 14.2 und 14.4 beweisen.

Zu 14.2.

$$\frac{\frac{s \in t \Rightarrow s \in t \quad \mathfrak{F}[t] \Rightarrow \mathfrak{F}[t]}{s \in t, \mathfrak{F}[t] \Rightarrow s \in t \square \mathfrak{F}[t]}}{s \in t, \mathfrak{F}[t] \Rightarrow \bigvee y (s \in y \square \mathfrak{F}[y])}}{s \in t, \mathfrak{F}[t] \Rightarrow s \in \iota x \mathfrak{F}[x]}.$$

Zu 14.4.

$$\frac{\frac{\frac{\frac{\mathfrak{F}[b] \Rightarrow \mathfrak{F}[b] \quad s \in b, t = b \Rightarrow s \in t}{s \in b, \mathfrak{F}[b], \mathfrak{F}[b] \rightarrow t = b \Rightarrow s \in t}}{s \in b, \mathfrak{F}[b], \wedge y (\mathfrak{F}[y] \rightarrow t = y) \Rightarrow s \in t}}{s \in b \square \mathfrak{F}[b], \wedge y (\mathfrak{F}[y] \rightarrow t = y) \Rightarrow s \in t}}{\bigvee y (s \in y \square \mathfrak{F}[y]), \wedge y (\mathfrak{F}[y] \rightarrow t = y) \Rightarrow s \in t}}{s \in \iota x \mathfrak{F}[x], \wedge y (\mathfrak{F}[y] \rightarrow t = y) \Rightarrow s \in t}.$$

Bis hierher gibt es noch nichts, was für die Widersprüche von Abschnitt 9 verantwortlich gemacht werden könnte, d.h. es sind noch keine Zusammenziehungen aufgetreten. Es ist der letzte Schritt auf dem Weg von 14.5 zu

$$\bigvee x (\mathfrak{F}[x] \square \wedge y (\mathfrak{F}[y] \rightarrow x = y)) \Rightarrow \mathfrak{F}[\iota x \mathfrak{C}[x]]$$

mithilfe eines Schnittes mit einer Substitutionsformel der Gestalt 8.2, wo Zusammenziehungen erforderlich werden, zumindest eine, immer vorausgesetzt, daß  $\mathfrak{F}$  überhaupt eine Formel der ersten Stufe ist:

$$\frac{\mathfrak{F}[a], \wedge y(\mathfrak{F}[y] \rightarrow a = y) \Rightarrow \iota x \mathfrak{F}[x] = a \quad [\iota x \mathfrak{F}[x] = a]^n, \mathfrak{F}[a] \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]]}{\text{n Schnitte, im wesentlichen derselben Gestalt}} \\ \frac{[\mathfrak{F}[a]]^n, [\wedge y(\mathfrak{F}[y] \rightarrow a = y)]^n, \mathfrak{F}[a] \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]]}{\text{Vertauschung(en) und Zusammenziehung(en)}} \\ \frac{\mathfrak{F}[a], \wedge y(\mathfrak{F}[y] \rightarrow a = y) \Rightarrow \mathfrak{F}[\iota x \mathfrak{C}[x]]}{\mathfrak{F}[a] \square \wedge y(\mathfrak{F}[y] \rightarrow a = y) \Rightarrow \mathfrak{F}[\iota x \mathfrak{C}[x]]} \\ \frac{\mathfrak{F}[a] \square \wedge y(\mathfrak{F}[y] \rightarrow a = y) \Rightarrow \mathfrak{F}[\iota x \mathfrak{C}[x]]}{\vee x(\mathfrak{F}[x] \square \wedge y(\mathfrak{F}[y] \rightarrow x = y)) \Rightarrow \mathfrak{F}[\iota x \mathfrak{C}[x]}}$$

Da dies die einzigen Zusammenziehungen im Beweis des  $\iota$ -Axioms sind, müssen sie als Ursache der Unverträglichkeit des  $\iota$ -Axioms mit der uneingeschränkten Abstraktion angesehen werden. Es ist der einzige Punkt, an dem ein klassisches Prinzip zur Anwendung kommt. In Anbetracht von 8.2 bedeutet das, daß die Anzahl der Substitutionen eine Rolle spielt, d.h. der Leitgedanke des *Ressourcenbewußtseins* reicht bis hierher.<sup>107</sup>

Dann ist aber die Frage: Wie kann dem in einer zusammenziehungsfreien Logik Rechnung getragen werden?

Ich knüpfe an die obige Beweisfigur an, bevor die Zusammensetzungen zur Anwendung kommen. An die Stelle der Zusammensetzungen treten folgende Schlüsse:

$$\frac{[\mathfrak{F}[a]]^n, [\wedge y(\mathfrak{F}[y] \rightarrow a = y)]^n, \mathfrak{F}[a] \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]]}{[\mathfrak{F}[a]]^{n+1}, [\wedge y(\mathfrak{F}[y] \rightarrow a = y)]^n \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]]} \\ \frac{\mathfrak{F}[a] \square^n \mathfrak{F}[a], [\wedge y(\mathfrak{F}[y] \rightarrow a = y)]^n \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]]}{\vee x(\mathfrak{F}[x] \square^n \mathfrak{F}[x]), [\wedge y(\mathfrak{F}[y] \rightarrow a = y)]^n \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]}}$$

D.h. zu jeder Nennform  $\mathfrak{F}$  der ersten Stufe gibt es eine natürliche Zahl  $n$  derart, daß die vorgehende Sequenz  $\mathbf{LD}_\lambda$ -beweisbar ist.

<sup>107</sup> Dieser Punkt ist wichtig im Hinblick auf die Möglichkeit einer Interpretation des  $\lambda$ -Kalküls in einer zusammenziehungsfreien Logik wie etwa in Abschnitt 4 in [39].



Aufgrund der Möglichkeit, einen Begriff der Notwendigkeit  $\Box$  als eine unendliche  $\Box$ -Konjunktion zu formulieren, der es zumindest gestattet, eine Akkumulation von gleichgestaltigen Antezedentformeln auszudrücken, kann der vorgehende Ersatz weiterentwickelt werden.

Es sei  $\mathfrak{F}$  wie oben. Dann sind Sequenzen der folgenden Gestalt  $\mathbf{L}^i\mathbf{D}_\lambda$ -beweisbar:

$$\begin{aligned} & \forall x \Box \mathfrak{F}[x], \Box \wedge z_1 \wedge z_2 (\mathfrak{F}[z_1] \Box \mathfrak{F}[z_2] \rightarrow z_1 = z_2) \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]], \\ & \forall x \Box (\mathfrak{F}[x] \Box \wedge y (\mathfrak{F}[y] \rightarrow x = y)) \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]], \\ & \forall x \Box \wedge y (\mathfrak{F}[y] \leftrightarrow x = y) \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]]. \end{aligned}$$

Das Vorgehen ist für jede der drei Sequenzen im wesentlichen dasselbe; ich beschränke meine Aufmerksamkeit auf die zweite. Ich setze mit der zweiten Zeile der vorigen Beweisfigur an:

$$\begin{array}{c} \frac{\mathfrak{F}[a]^{(n+1)}, [\wedge y (\mathfrak{F}[y] \rightarrow a = y)]^n \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]]}{\mathfrak{F}[a]^{(n+1)}, [\wedge y (\mathfrak{F}[y] \rightarrow a = y)]^{(n+1)} \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]}} \\ \hline (n+1) \Box\text{-Einführungen, möglicherweise mit Vertauschungen} \\ \hline \frac{\mathfrak{F}[a] \Box \wedge y (\mathfrak{F}[y] \rightarrow a = y) \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]]}{\Box (\mathfrak{F}[a] \Box \wedge y (\mathfrak{F}[y] \rightarrow a = y)) \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]}} \dagger \\ \hline \forall x \Box (\mathfrak{F}[a] \Box \wedge y (\mathfrak{F}[y] \rightarrow a = y)) \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]] \end{array}$$

In dem mit  $\dagger$  gekennzeichneten Schluß kommt 1i aus [38], S. 670, zur Anwendung.

Tatsächlich lassen sich die obigen Schemata schon für die  $\mathbf{L}^i\mathbf{D}_\lambda$ -Beweisbarkeit zeigen. Schema 1i aus [38], S. 670, das eine Art von  $\Box$ -Einführung links darstellt, kann ohne den Gebrauch von  $\mathbf{Z}$ -Schlüssen gezeigt werden, d.h. ist schon  $\mathbf{L}^i\mathbf{D}_\lambda$ -beweisbar; und mehr wird von  $\Box$  in dem obigen Beweis nicht gebraucht. Das bringt allerdings wenig, da es in  $\mathbf{L}^i\mathbf{D}_\lambda$  unmöglich ist, eine Formel der Gestalt  $\Box A$  zu beweisen.<sup>108</sup> Das ist eine entscheidende Änderung für  $\Box$  in  $\mathbf{L}^i\mathbf{D}_\lambda^Z$ .

<sup>108</sup> Da die Schnittregel in  $\mathbf{L}^i\mathbf{D}_\lambda$  eliminierbar ist, kann die  $\mathbf{L}^i\mathbf{D}_\lambda$ -Beweisbarkeit von  $\Box A$  auf die einer Formel der Gestalt  $\check{\mathfrak{Y}}[A]$  zurückgeführt werden. Letztere kann jedoch niemals  $\mathbf{L}^i\mathbf{D}_\lambda$ -beweisbar sein, wie aus Proposition 132.21 in [37] hervorgeht.

Damit ergeben sich die folgenden Schlußschemata als  $\mathbf{L}^1\mathbf{D}_\lambda^Z$ -herleitbar, sofern  $\mathfrak{F}$  eine Nennform der ersten Stufe ist:

$$\frac{\Gamma \Rightarrow \forall x \square \mathfrak{F}[x] \quad \mathfrak{F}[a], \mathfrak{F}[b], \Pi \Rightarrow a = b}{\Gamma, \square \Pi \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]]},$$

$$\frac{\Gamma \Rightarrow \mathfrak{F}[s] \quad \mathfrak{F}[a], \Pi \Rightarrow s = a}{\square \Gamma, \square \Pi \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]]}.$$

Auf diese Weise erhält man einen durchaus brauchbaren Ersatz für den bestimmten Artikel.

Angesichts des Zwischenergebnisses

$$[\mathfrak{F}[a]]^{(n+1)}, [\bigwedge y (\mathfrak{F}[y \rightarrow a = y])]^{(n+1)} \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]]$$

bietet sich jedoch noch eine andere Version an, die den Vorteil einer vertrauten Form hat. In der klassischen Formulierung wird die  $\rightarrow$ -Implikation durch die schwache Implikation  $\supset$  ersetzt:

$$\begin{aligned} & \forall x \mathfrak{F}[x] \square \bigwedge z_1 \bigwedge z_2 (\mathfrak{F}[z_1] \square \mathfrak{F}[z_2] \rightarrow z_1 = z_2) \supset \mathfrak{F}[\iota x \mathfrak{F}[x]], \\ & \forall x (\mathfrak{F}[x] \square \bigwedge y (\mathfrak{F}[y] \rightarrow x = y)) \supset \mathfrak{F}[\iota x \mathfrak{F}[x]], \\ & \forall x \bigwedge y (\mathfrak{F}[y] \leftrightarrow x = y) \supset \mathfrak{F}[\iota x \mathfrak{F}[x]]. \end{aligned}$$

Das ergibt sich leicht aus folgender Variation der letzten Beweisfigur:

$$\begin{aligned} & \frac{[\mathfrak{F}[a] \square \bigwedge y (\mathfrak{F}[y \rightarrow a = y])]^{(n+1)} \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]]}{[\mathfrak{F}[a] \square \bigwedge y (\mathfrak{F}[y \rightarrow a = y])/\tilde{n}^I] \Rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]]} \\ \Rightarrow \tilde{n}^I \in \check{\mathbf{II}} & \Rightarrow \frac{[\mathfrak{F}[a] \square \bigwedge y (\mathfrak{F}[y \rightarrow a = y])/\tilde{n}^I] \rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]]}{\Rightarrow \tilde{n}^I \in \check{\mathbf{II}} \square ([\mathfrak{F}[a] \square \bigwedge y (\mathfrak{F}[y \rightarrow a = y])/\tilde{n}^I] \rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]])} \\ & \Rightarrow \frac{\Rightarrow \tilde{n}^I \in \check{\mathbf{II}} \square ([\mathfrak{F}[a] \square \bigwedge y (\mathfrak{F}[y \rightarrow a = y])/\tilde{n}^I] \rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]])}{\Rightarrow \forall x (x \in \check{\mathbf{II}} \square ([\mathfrak{F}[a] \square \bigwedge y (\mathfrak{F}[y \rightarrow a = y])/\tilde{n}^I] \rightarrow \mathfrak{F}[\iota x \mathfrak{F}[x]])}, \end{aligned}$$

wobei natürlich, wie immer,  $\mathfrak{F}$  eine Nennform der ersten Stufe sein muß.

## 15. Induktion neu gefaßt

Die schwache Implikation versetzt uns auch in die Lage, eine Definition der natürlichen Zahlen anzugeben:

$$\mathbf{N}^\circ := \lambda x \bigwedge y (0 \in y \wedge \bigwedge z (z \in y \rightarrow z' \in y) \supset x \in y),$$

wobei  $^\circ$  anzeigen soll, daß es sich hier um einen *exklusiven* Term in dem Sinne handelt, daß die Elemente von  $\mathbf{N}^\circ$  wirklich nur  $0, 0', \dots$  sind, und nicht auch ihnen gleiche Terme anderer Gestalt.

$\mathbf{N}^\circ$  erlaubt die Formulierung eines Induktionsschemas.<sup>109</sup> Schlüsse nach folgenden Schemata sind  $\mathbf{LD}_\lambda^Z$ -herleitbar, sofern die freie Variable  $a$  nicht in der Untersequenz auftritt:

$$\frac{\Gamma \Rightarrow \mathfrak{F}[0] \quad \mathfrak{F}[a], a \in \mathbf{N}^\circ, \text{ n } \Gamma \Rightarrow \mathfrak{F}[a']}{s \in \mathbf{N}^\circ, \square \Gamma \Rightarrow \mathfrak{F}[s]} .$$

Das erlaubt auch eine doppelte Induktion in der Form:

$$\frac{a \in \mathbf{N}^\circ, \Gamma \Rightarrow \mathfrak{F}[a, 0] \quad \bigwedge x \mathfrak{F}[x, b], a \in \mathbf{N}^\circ, b \in \mathbf{N}^\circ, \Gamma \Rightarrow \mathfrak{F}[a, b']}{s \in \mathbf{N}^\circ, t \in \mathbf{N}^\circ, \square \Gamma \Rightarrow \mathfrak{F}[s, t]} .$$

In Abschnitt 7 hatten wir gesehen, daß alle rekursiven Funktionen in  $\mathbf{LD}_\lambda$  dargestellt werden können. Was dort jedoch nicht möglich war, war zu zeigen, daß sie auch *total* sind, d.h. für alle Argumente aus den natürlichen Zahlen auch Werte in den natürlichen Zahlen haben. Jetzt können wir zumindest für alle primitiv-rekursiven Funktionen die Totalität nachweisen.

Das reicht, um die primitiv-rekursive Arithmetik **PRA** in  $\mathbf{LD}_\lambda^Z$  zu interpretieren. Ein entsprechendes Ergebnis kann in [37] oder auch [38] nachgelesen werden.

Es reicht aber nicht, wie im klassischen Fall, für eine verschränkte Induktion, d.h. eine Induktion der Form

$$\frac{\Gamma \Rightarrow \mathfrak{F}[0, b] \quad \mathfrak{F}[a, t_1], \Pi \Rightarrow \mathfrak{F}[a', 0] \quad \mathfrak{F}[a, t_2], \mathfrak{F}[a', b], \Xi \Rightarrow \mathfrak{F}[a', b']}{\Gamma, \Pi, \Xi \Rightarrow \mathfrak{F}[s, t]} .$$

Das kann man sich an Hand folgender Beweisfiguren deutlich machen. Mit dem klassischen Schema der Induktion erhält man:

<sup>109</sup> Das ist der ganze Sinn und Zweck des exklusiven Terms.

$$\frac{\frac{\frac{\mathfrak{F}[a, t_1] \Rightarrow \mathfrak{F}[a', 0]}{\wedge y \mathfrak{F}[a, y] \Rightarrow \mathfrak{F}[a', 0]} \quad \frac{\mathfrak{F}[a, t_2], \mathfrak{F}[a', b] \Rightarrow \mathfrak{F}[a', b']}{\wedge y \mathfrak{F}[a, y], \mathfrak{F}[a', b] \Rightarrow \mathfrak{F}[a', b']}}{\wedge y \mathfrak{F}[a, y] \Rightarrow \mathfrak{F}[a', c]}}{\wedge y \mathfrak{F}[a, y] \Rightarrow \wedge y \mathfrak{F}[a', y]}}{\frac{\Gamma \Rightarrow \mathfrak{F}[0, b]}{\Gamma \Rightarrow \wedge y \mathfrak{F}[0, b]}} \quad \frac{\Gamma \Rightarrow \wedge y \mathfrak{F}[s, y]}{\Gamma \Rightarrow \mathfrak{F}[s, t]}.$$

Das Problem ist die Induktion im rechten Ast. Alles was man in  $\mathbf{LD}_\lambda^Z$  erhält, ist:

$$\frac{\frac{\frac{\mathfrak{F}[a, t_1] \Rightarrow \mathfrak{F}[a', 0]}{\wedge y \mathfrak{F}[a, y] \Rightarrow \mathfrak{F}[a', 0]} \quad \frac{\mathfrak{F}[a, t_2], \mathfrak{F}[a', b] \Rightarrow \mathfrak{F}[a', b']}{\wedge y \mathfrak{F}[a, y], \mathfrak{F}[a', b] \Rightarrow \mathfrak{F}[a', b']}}{\square \wedge y \mathfrak{F}[a, y] \Rightarrow \mathfrak{F}[a', c]}}{\square \wedge y \mathfrak{F}[a, y] \Rightarrow \wedge y \mathfrak{F}[a', y]}.$$

Das eignet sich aber nicht mehr für die zweite Induktion. Was hier fehlt, ist ein verstärkter Notwendigkeitsoperator  $\square$ , der folgenden Schlußregeln gehorcht:

$$\frac{\square^n A, \Gamma \Rightarrow C}{\square A, \Gamma \Rightarrow C} \quad \text{und} \quad \frac{\square^n \Gamma \Rightarrow C}{\square \Gamma \Rightarrow \square C}.$$

Damit ließe sich eine verschränkte Induktion folgendermaßen auf einfache Induktionen zurückzuführen:

$$\frac{\frac{\frac{\frac{b \in \mathbf{N}^\circ \Rightarrow \exists [b] \in \mathbf{N}^\circ \quad \mathfrak{F}[a, \exists [b]], \mathfrak{F}[a', b] \Rightarrow \mathfrak{F}[a', b']}{b \in \mathbf{N}^\circ, \exists [b] \in \mathbf{N}^\circ \rightarrow \mathfrak{F}[a, \exists [b]], \mathfrak{F}[a', b] \Rightarrow \mathfrak{F}[a', b']}}{\Rightarrow \mathfrak{F}[a', 0]} \quad \frac{b \in \mathbf{N}^\circ, \wedge^\circ y \mathfrak{F}[a, y], \mathfrak{F}[a', b] \Rightarrow \mathfrak{F}[a', b']}{c \in \mathbf{N}^\circ, \square \wedge^\circ y \mathfrak{F}[a, y] \Rightarrow \mathfrak{F}[a', c]}}{\square \wedge^\circ y \mathfrak{F}[a, y] \Rightarrow \mathfrak{F}[a', c]}}{\Rightarrow \wedge^\circ y \mathfrak{F}[0, y]} \quad \frac{\square \wedge^\circ y \mathfrak{F}[a, y] \Rightarrow \wedge^\circ y \mathfrak{F}[a', y]}{\square \wedge^\circ y \mathfrak{F}[a, y] \Rightarrow \square \wedge^\circ y \mathfrak{F}[a', y]}}{\Rightarrow \square \wedge^\circ y \mathfrak{F}[0, y]} \quad \frac{s \in \mathbf{N} \Rightarrow \square \wedge^\circ y \mathfrak{F}[s, y]}{s \in \mathbf{N} \Rightarrow \wedge^\circ y \mathfrak{F}[s, y]} \quad \frac{s \in \mathbf{N}, t \in \mathbf{N} \Rightarrow \mathfrak{F}[s, t]}{s \in \mathbf{N}, t \in \mathbf{N} \Rightarrow \mathfrak{F}[s, t]}.$$

In [39] habe ich auf den Seiten 136–159 eine Möglichkeit vorgestellt, einen solchen verstärkten Notwendigkeitsoperator zu definieren, die im wesentlichen auf dem Verfahren aufbaut, das ich schon für die **Z**-Erweiterung eingesetzt habe. Damit lassen sich dann auch 2-Rekursionen darstellen.

Die gleiche Schwierigkeit taucht aber wieder auf, wenn es um eine dreifach verschränkte Induktion geht. Das Verfahren, das ich in [39] beschrieben habe, läßt sich jedoch entsprechend erweitern, was dann zu einer Hierarchie der (mathematischen) Induktionen führt, die ihrerseits wieder in eine entsprechende Hierarchie der Rekursionen mündet. Es steht zu erwarten, daß ein Widerspruchsfreiheitsbeweis für die  $n$ -fach verschränkte Induktion eine Induktion bis  $\omega_n$  erfordert, d.h. man erhält folgende Hierarchie für Rekursionen:

- 1-Rekursion:  $\omega^\omega$  oder, anders gesagt,  $\omega_1$
- 2-Rekursion:  $\omega^{\omega^\omega}$  oder, anders gesagt,  $\omega_2$
- ⋮

Dann stellt sich aber die Frage nach einem Abschluß dieser Hierarchie. Nicht zuletzt deshalb habe ich einen Ersatz für die Beschreibung eingeführt, um mit dem  $\mu$ -Operator in der Lage zu sein, die Definitionsstruktur für allgemein rekursive Funktionen nachvollziehen zu können. Die entscheidende Frage ist, ob sich damit der Begriff der allgemein rekursiven Funktion für die zusammenziehungsfreie Logik einführen läßt. Das ist eine offene Frage, aber es scheint mir, insbesondere angesichts der inversen Church-Turing-These, eine Frage von eminenter metaphysischer Bedeutung zu sein.

## 16. Ausklang

Metaphysik im Anschluß an Kant und Hegel ist ein Projekt das sich als *Ableitung der Denkbestimmungen* aus einer Fakultät des Logischen kennzeichnen läßt. Das bedeutet, daß Denkbestimmungen als logische Entitäten definiert und ihre Eigenschaften innerhalb einer (höheren) Logik bewiesen werden. Das ist es, was Metaphysik zur Logik macht, und zwar zu einer spekulativen Logik, d.h. zu einer Logik, die immer wieder — durch Totalisierungen, die auf ihren eigenen Bereich abzielen — über sich selbst hinausgeht.

Das zentrale Problem einer Metaphysik als einer spekulativen Logik ist, mehr noch als in Freges Unterfangen einer logischen Grundlegung des Begriffs der Zahl, die Klärung der Frage: „Wie kommen die leeren Formen der Logik dazu, aus sich heraus solchen Inhalt zu gewinnen?“<sup>110</sup>

Mein grundsätzlicher Ansatz zur Beantwortung dieser Frage und gleichzeitig der Ausgangspunkt für meinen Versuch einer Begründung der Metaphysik durch logische Analyse der Selbstbezüglichkeit ist durch das Auftreten der Paradoxien in den Grundlagen der Logik, Semantik und Mengenlehre gegeben. Sie haben Freges ursprüngliches Unterfangen einer logischen Grundlegung der Arithmetik zunichte gemacht, aber gleichzeitig einen neuen Weg für eine logische Grundlegung der Dialektik und spekulativen Philosophie gewiesen, indem sie zeigen, daß schon die elementaren Bausteine der (höheren) Logik, Abstraktion und Prädikation, Widersprüche erzeugen. Auch wenn Poincaré dem nichts Positives abgewinnen konnte, er hat den Nagel auf den Kopf getroffen: „[L]a Logistique n'est plus stérile, elle engendre l'antinomie.“<sup>111</sup> Was auf der Grundlage der klassischen Logik nur als Antinomie in Erscheinung treten kann, wird in der spekulativen Logik zum Gegenstand einer Untersuchung gemacht. Die Beziehungen widersprüchlicher Begriffe untereinander, die Möglichkeit, auf der Grundlage selbstbezüglicher Begriffsbildungen Hierarchien zu definieren und über sie zu totalisieren, wie das etwa im Fall von **Z** geschieht, stellen eine Quelle möglicher Inhalte für logische Einsicht dar. Damit werden wahrheitskonservierende Erweiterungsschlüsse ohne Rückgriff auf äußere Erfahrung möglich.

Das bedeutet, daß die uneingeschränkte Abstraktion, wie sie erstmals von Frege in die Logik eingeführt wurde, einen der beiden Eckpfeiler der Metaphysik als einer spekulativen Logik darstellt. Der andere Eckpfeiler ist eine Einschränkung der klassischen Logik, die nötig wird, wenn man in konsistenter Weise mit uneingeschränkter Abstraktion, und mit der daraus resultierenden Fixpunkteigenschaft und ihren Widersprüchen, arbeiten will. Darüber hinaus ist die spekulative Logik durch eine extreme Sparsamkeit ihrer theoretischen Konstanten gekennzeichnet. Idealerweise sollte neben der uneingeschränkten Abstraktion nur eine Form der Inklusion eingeführt werden. Höchstens noch die Identität, oder, alternativ, die Prädikation soll zu den Grundzeichen gehören.

<sup>110</sup> [14], S. 22.

<sup>111</sup> [41], S. 316.

In diesem Sinne habe ich auf den vorhergehenden Seiten versucht, einen Einblick in die Möglichkeiten zu geben, die sich eröffnen, wenn man die dogmatische Haltung des logischen Empirismus und seiner Tradition aufgibt und eine unbeschränkte Begriffsbildung mit ihren Typenvermengungen zuläßt. Die Einschränkung der Logik, die das ermöglicht, zielt auf den Umgang mit Annahmen beim Schließen ab. Klassisches Schließen erlaubt u.a. das Zusammenziehen von Annahmen derselben Gestalt, d.h. eine Annahme  $A$ , die im Verlauf eines Beweises mehrmals gebraucht wird, kann so behandelt werden, als ob sie nur ein einziges Mal gebraucht wird. Das verträgt sich nicht mit Typenvermengungen. Mit Typenvermengungen kommt Ambiguität, und mit der Ambiguität verlieren wir das, was gemeinhin als Voraussetzung für sinnvolles Schließen gilt: Stabilität der Aussagen.<sup>112</sup> Sinnvolles Schließen ohne Stabilität der Aussagen ist jedoch möglich, wenn Zusammenziehungen aufgegeben werden. Hier kündigt sich an, wie wichtig für das spekulative Denken die Rolle von Annahmen im Schließen ist. Es ist nicht eine besondere Form des Satzes, etwa der *spekulative Satz*, wie Hegel meinte, wodurch sich spekulatives Denken vom klassischen abhebt, sondern eine besondere Form des Schließens. Spekulative Logik kann geradezu als Wissenschaft der Dynamik von Annahmen im logischen Schließen gelten. Die Aufgabe für die Metaphysik als einer spekulativen Logik besteht dann darin, in dieser Dynamik die charakteristischen Gesetze von Kategorien dingfest zu machen.<sup>113</sup> Wohlgermerkt, die *Gesetze*, nicht die Kategorien selbst, sollen dingfest gemacht werden. Das heißt, auf der formalen Ebene sollen Objekte definiert werden, die gerade jene Gesetzmäßigkeiten erfüllen, die von Kategorien erwartet werden; so wie etwa die Folge  $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \dots$  der von Neumann'schen Definition der natürlichen Zahlen die Gesetze der Arithmetik erfüllt. Das läßt sich natürlich nicht so ohne weiteres auf außer-mathematische Bereiche übertragen: Axiomatische Formulierungen kategorialer Begriffe sind rar. Da trifft es sich gut, daß die Axiomensysteme der modalen Logik eine Ausnahme darstellen: Mit der Erweiterung von  $\mathbf{LD}_\lambda$  durch  $\mathbf{Z}$ -Schlüsse werden nämlich Gesetze ableitbar, die

<sup>112</sup> Vgl. [42], S. 462, zum Beispiel.

<sup>113</sup> Manch ein Philosoph wird hier Einspruch erheben wollen: Es darf nicht als ausgemacht gelten, daß Kategorien überhaupt Gesetzen gehorchen. Es werden wohl jene Philosophen sein, die sich in ihrem angestammten Bereich bedroht fühlen, sobald sie verbindlich angeben sollen, welchen Ursprung ihre Einsichten haben. Wenn es nach ihnen ginge, gäbe es weder eine moderne Physik noch eine moderne Logik.

dadurch als Gesetze der T-Modalität und einer schwachen Implikation identifiziert werden können. Diese Gesetze ermöglichen es auch, die Definitionsgleichungen der primitiv-rekursiven Funktionen zu beweisen; damit wird der Schritt von einer ziffernweisen Repräsentierbarkeit der primitiv-rekursiven (d.i. 1-rekursiven) Funktionen zu einer formalen Beweisbarkeit ihrer Definitionsgleichungen auf der Ebene der Objekttheorie vollzogen. Dieser Ansatz läßt sich schrittweise so erweitern, daß die Definitionsgleichungen der  $k$ -rekursiven Funktionen für jedes  $k > 1$  ( $k \in \mathbb{N}$ ) beweisbar werden.<sup>114</sup> Auf diese Weise ist die Metaphysik als spekulative Logik mit der Grundlegung der Mathematik aufs engste verknüpft.

Was bleibt, ist die Frage, ob und wie dieser Ansatz fortgeführt werden kann, um mehr als nur modale Kategorien zu gewinnen. Die Fixpunktkonstruktionen, die bisher zur Anwendung gekommen sind, zielen darauf ab, eine Akkumulation von Formeln derselben Gestalt ausdrücken zu können. Es steht nicht zu erwarten, daß sich auf diese Weise jemals etwas anderes als ein Notwendigkeitsbegriff ergibt. Der Ausweg wird sein, sich nicht nur auf den Ausschluß von Zusammenziehungen zu beschränken, sondern den einmal eingeschlagenen Weg auch auf weitere Strukturschlußregeln, insbesondere etwa Vertauschungen, auszudehnen. Dazu wird man aber Paradoxien betrachten müssen, die neben Zusammenziehungen auch auf signifikante Weise Vertauschungen involvieren.

Bei alledem muß klar sein, daß es sich hier um eine Aufgabe handelt, die sich mit den üblichen Mitteln der philosophischen Argumentation nicht bewältigen läßt. Deshalb habe ich der formal-logischen Behandlung so viel Raum gegeben. Es braucht mehr als nur schöne Worte, um aus der Hegelschen Idee der Metaphysik als einer spekulativen Logik etwas zu machen, das die Bezeichnung „Theorie“ verdient.

## Literatur

- [1] BADIOU, Alain: *L'être et l'événement*. Paris: Éditions du Seuil, 1988.
- [2] BARTLETT, James M.: *Funktion und Gegenstand. Eine Untersuchung in der Logik von Gottlob Frege*, München: Diss. Philosophie, 1961.
- [3] BEESON, Michael J.: *Foundations of constructive mathematics*. Berlin, Heidelberg, New York: Springer-Verlag, 1985 (Ergebnisse der Mathematik und ihrer Grenzgebiete, Folge 3, Band 6).

<sup>114</sup> Für den Fall  $k = 2$  habe ich das in [39] auf den Seiten 143–159 gezeigt.



- [4] BRÖCKER, Walter: *Formale, transzendente und spekulative Logik*. Frankfurt am Main: Vittorio Klostermann, 1962 (Wissenschaft und Gegenwart Heft 23).
- [5] CARNAP, Rudolf: *Der logische Aufbau der Welt*. Hamburg: Felix Meiner, 1974 (vierte Auflage). – Erste Auflage 1928.
- [6] CARNAP, Rudolf: Überwindung der Metaphysik durch logische Analyse der Sprache. In: *Erkenntnis* 2 (1931), S. 219–241.
- [7] CARNAP, Rudolf: *Philosophy and logical syntax*. London: Kegan Paul, Trench, Trübner & Co. Ltd., 1935 (Psyche Miniatures, General Series 70). – Three lectures delivered at the University of London.
- [8] COXON, A. H.: *The Fragments of Parmenides. A critical text with introduction, translation, the ancient testimonia and a commentary*. Assen/Maastricht: Van Gorcum, 1986.
- [9] DIACONESCU, Radu: Axiom of choice and complementation. In: *Proceedings of the American Mathematical Society* 51 (1975), S. 176–178.
- [10] DIELS, Hermann; KRANZ, Walther: *Die Fragmente der Vorsokratiker*. Erster Band. Berlin: Weidmannsche Verlagsbuchhandlung, 1951 (sechste Auflage).
- [11] ESSLER, Wilhelm K.: Zur Topologie der Arten dialektischer Logik bei Hegel. In: HENRICH, Dieter (Hrsg.): *Hegels Wissenschaft der Logik: Formation und Rekonstruktion*. Stuttgart: Klett-Cotta, 1986, S. 198–208.
- [12] FITCH, Frederic B.: Self-Reference in Philosophy. In: *Mind* LV (1946), S. 64–73.
- [13] FRAENKEL, Abraham A.; BAR-HILLEL, Yehoshua; LEVY, Azriel: *Foundations of Set Theory*. Amsterdam: North-Holland Publishing Company, 1973 (zweite Auflage). – Erste Auflage 1958.
- [14] FREGE, Gottlob: *Grundlagen der Arithmetik*. Breslau: Marcus, 1934 (zweite Auflage). – Erste Auflage: Wilhelm Koebner, Breslau 1884.
- [15] FREGE, Gottlob: *Grundgesetze der Arithmetik, begriffsschriftlich abgeleitet*. Bd. II. Hildesheim: Georg Olms Verlagsbuchhandlung, 1966. – Orig.: Verlag von Hermann Pohle, Jena 1903.
- [16] GIRARD, Jean-Yves: *Studies in Proof Theory*. Bd. I: *Proof Theory and Logical Complexity*. Napoli: Bibliopolis, 1987.
- [17] GLOCKNER, Hermann (Hrsg.): *Hegel Sämtliche Werke. Jubiläumsausgabe in zwanzig Bänden*. Stuttgart-Bad Cannstatt: Friedrich

- Frommann Verlag (Günther Holzboog), 1961–68 (vierte Auflage). – Erste Auflage 1927–30.
- [18] GÖDEL, Kurt: Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme. In: *Monatshefte für Mathematik und Physik* 38 (1931), S. 173–198.
- [19] GÖDEL, Kurt: Russell's Mathematical Logic. In: SCHILPP, Paul A. (Hrsg.): *Library of living philosophers*. Bd. 5: *The Philosophy of Bertrand Russell*. La Salle, Illinois: Open Court Publishing Company, 1944, S. 123–153.
- [20] GOODMAN, Nicolas D.; MYHILL, John: Choice implies excluded middle. In: *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 24 (1978), S. 461.
- [21] GOOLD, George P. (Hrsg.): *Aristotle. The Metaphysics. Books I–IX*. Cambridge MA und London: Harvard University Press, 1989 (The Loeb Classical Library 271). – Erstauflage 1933.
- [22] GOOLD, George P. (Hrsg.): *Aristotle. The Metaphysics: Books X–XIV, Oeconomica, Magna Moralia*. Cambridge MA und London: Harvard University Press, 1990, (The Loeb Classical Library 287). – Erstauflage 1935.
- [23] HEGEL, Georg Wilhelm F.: *Philosophische Propädeutik*. [17], Bd. 3.
- [24] HEGEL, Georg Wilhelm F.: *Enzyklopädie der philosophischen Wissenschaften im Grundrisse*. [17], Bd. 6.
- [25] HEGEL, Georg Wilhelm F.: *System der Philosophie. Erster Teil. Die Wissenschaft der Logik*. [17], Bd. 8.
- [26] HEGEL, Georg Wilhelm F.: *System der Philosophie. Dritter Teil. Die Philosophie des Geistes*. [17], Bd. 10.
- [27] HEGEL, Georg Wilhelm F.: *Vorlesungen über die Geschichte der Philosophie. Erster Band*. [17], Bd. 17.
- [28] HEGEL, Georg Wilhelm F.: *Vorlesungen über die Philosophie der Geschichte*. [17], Bd. 11.
- [29] HEISS, Robert: *Logik des Widerspruchs. Eine Untersuchung zur Methode der Philosophie und zur Gültigkeit der formalen Logik*. Berlin und Leipzig: Walter de Gruyter & Co., 1932.
- [30] HERMES, Hans; KAMBARTEL, Friedrich; KAULBACH, Friedrich (Hrsg.): *Gottlob Frege. Nachgelassene Schriften und wissenschaftlicher Briefwechsel I*. Hamburg: Felix Meiner, 1969.

- [31] HERMES, Hans; SCHOLZ, Heinrich: *Mathematische Logik*. Leipzig: Teubner, 1952 (Enzyklopädie der mathematischen Wissenschaften, Band I, 1. Teil, Heft 1, Teil I).
- [32] KOCH, Anton F.: Die Selbstbeziehung der Negation in Hegels Logik. In: *Zeitschrift für philosophische Forschung* 53 (1999), S. 1–29.
- [33] ŁUKASIEWICZ, Jan; TARSKI, Alfred: Untersuchungen über den Aussagenkalkül. *Comptes rendus des séances de la Société des Sciences et des Lettres de Varsovie*, vol. 23, 1930, cl. iii, 30–50.
- [34] MAUS, Heinz; FÜRSTENBERG, Friedrich (Hrsg.): *Der Positivismusstreit in der deutschen Soziologie*. Neuwied und Berlin: Luchterhand, 1971 (Soziologische Texte 58); (dritte Auflage).
- [35] ODIFREDDI, Piergiorgio: *Studies in Logic and the Foundations of Mathematics*. Bd. 125: *Classical Recursion Theory — The Theory of Functions and Sets of Natural Numbers*. Amsterdam und anderswo: Elsevier, 1989.
- [36] PETERSEN, Uwe: Logic Without Contraction as Based on Inclusion and Unrestricted Abstraction. In: *Studia Logica* 64 (2000), S. 365–403.
- [37] PETERSEN, Uwe: *Diagonal Method and Dialectical Logic. Tools, Materials, and Groundworks for a Logical Foundation of Dialectic and Speculative Philosophy*. Osnabrück: Der Andere Verlag, 2002.
- [38] PETERSEN, Uwe:  $\mathbf{LD}_\lambda^Z$  as a Basis for  $\mathbf{PRA}$ . In: *Archive for Mathematical Logic* 42 (2003), S. 665–694.
- [39] PETERSEN, Uwe: Some Additions and Corrections to *Diagonal Method and Dialectical Logic*. In: *Dilemmata. Jahrbuch der ASFPG* 1 (2006), S. 93–173.
- [40] PINKARD, Terry P.: The Categorical Satisfaction of Self-Reflexive Reason. In: *Bulletin of the Hegel Society of Great Britain* 19 (1989), S. 5–17.
- [41] POINCARÉ, Henri: Les mathématiques et la logique. In: *Revue de métaphysique et de la morale* 14 (1906), S. 294–317.
- [42] POINCARÉ, Henri: La logique de l'infini. In: *Revue de métaphysique et de la morale* 17 (1909), S. 461–482.
- [43] QUINE, Willard Van O.: *Mathematical Logic*. Cambridge MA: Harvard University Press, 1951 (zweite Auflage). Erste Auflage 1940.
- [44] QUINE, Willard Van O.: Quantifiers and Propositional Attitudes. In: *The Journal of Philosophy* 53 (1956), S. 177–187.

- [45] QUINE, Willard Van O.: *Wort und Gegenstand*. Stuttgart: Philipp Reclam jun., 1980.
- [46] RUSSELL, Bertrand: *The Principles of Mathematics*. London: Allen & Unwin, 1937 (third reprint of second edition). – Orig.: Cambridge: Cambridge University Press, 1903.
- [47] RUSSELL, Bertrand: Mathematical Logic as Based on the Theory of Types. In: *Logic and Knowledge, Essays 1901–1950* (R. C. Marsh, Hrsg.). London: George Allen & Unwin Ltd, 1956, S. 59–102. – Orig.: *American Journal of Mathematics* 30 (1908), 222–262.
- [48] RUSSELL, Bertrand: Introduction. In: [63], 1922, S. ix–xxii.
- [49] RUSSELL, Bertrand: *Das menschliche Wissen*. Darmstadt: Holle Verlag, 1951.
- [50] SCHÜTTE, Kurt: *Beweistheorie*. Berlin, Göttingen, Heidelberg: Springer-Verlag, 1960.
- [51] SCHÜTTE, Kurt: *Vollständige Systeme modaler und intuitionistischer Logik*. Berlin, Heidelberg, New York: Springer-Verlag, 1968.
- [52] SCHWARZ, Franz F.: *Aristoteles. Metaphysik*. Stuttgart: Philipp Reclam Jun., 1993. Erste Auflage 1970.
- [53] SKIRBEKK, Gunnar (Hrsg.): *Wahrheitstheorien. Eine Auswahl aus den Diskussionen über Wahrheit im 20. Jahrhundert*. Frankfurt am Main: Suhrkamp, 1977 (suhrkamp taschenbuch wissenschaft 210).
- [54] SKOLEM, Thoralf: Bemerkungen zum Komprehensionsaxiom. In: *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* 3 (1957), S. 1–17.
- [55] STEGMÜLLER, Wolfgang: *Das Wahrheitsproblem und die Idee der Semantik*. Wien und New York: Springer-Verlag, 1970 (zweite Auflage). – Erste Auflage 1957.
- [56] STEGMÜLLER, Wolfgang: *Theorie und Erfahrung, Zweiter Halbband: Theorienstrukturen und Theoriendynamik*. Berlin, Heidelberg, New York: Springer-Verlag, 1973 (Probleme und Resultate der Wissenschaftstheorie und Analytischen Philosophie, Band II).
- [57] TARSKI, Alfred: Die semantische Konzeption der Wahrheit und die Grundlagen der Semantik. In: [53], S. 140–188.
- [58] THIEL, Christian: Gottlob Frege: Die Abstraktion. In: SPECK, Josef (Hrsg.): *Grundprobleme der großen Philosophen, Philosophie der Gegenwart*, 1. Göttingen: Vandenhoeck & Ruprecht, 1972, S. 9–44.

- [59] THIEL, Christian: Das Begründungsproblem der Mathematik und die Philosophie. In: KAMBARTEL, Friedrich; MITTELSTRASS, Jürgen (Hrsg.): *Zum normativen Fundament der Wissenschaft*. Frankfurt am Main: Athenäum-Verlag, 1973, S. 91–114.
- [60] TUGENDHAT, Ernst: Tarskis semantische Definition der Wahrheit und ihre Stellung innerhalb der Geschichte des Wahrheitsproblems im logischen Positivismus. In: [53], S. 189–223. – Orig.: *Philosophische Rundschau* 8 (1960), 131–159.
- [61] WANG, Hao: *A logical journey: from Gödel to philosophy*. Cambridge MA und London: The MIT Press, 1996.
- [62] WHITE, Richard B.: The Consistency of the Axiom of Comprehension in the Infinite-Valued Predicate Logic of Łukasiewicz. In: *Journal of Philosophical Logic* 8 (1979), S. 509–534.
- [63] WITTGENSTEIN, Ludwig: *Tractatus logico-philosophicus*. London: Routledge & Kegan Paul, 1969 (vierte Auflage).
- [64] WITTGENSTEIN, Ludwig: *Bemerkungen über die Grundlagen der Mathematik* (G. H. von Wright, R. Rhees, und G. E. M. Anscombe, Hrsg.). Ludwig Wittgenstein Werkausgabe Band 6. Frankfurt am Main: Suhrkamp Verlag, 1984 (suhrkamp taschenbuch wissenschaft 506).



# The Laws of Irresponsibility

SCOTT VEITCH\*

ABSTRACT. This paper considers the ways in which the legal organisation of responsibility is able to legitimate the disappearance of responsibility, that is, to organise irresponsibility. It assesses the importance of aspects of legality under contemporary social conditions as a means of constituting impunities that facilitate the legitimation of suffering as a matter of routine practice.

## Problems with responsibility

According to Young (2004) “The most common model of assigning responsibility derives from legal reasoning to find guilt or fault for a harm. Under the fault model, one assigns responsibility to particular agents whose actions can be shown as causally connected to the circumstances for which responsibility is sought” (368). Young recognises a number of problems with this model including the following: the identification of the agent; the establishing of cause and effect; the relevant status of the actions (intentional, negligent etc.); and, finally, that “Assigning responsibility to some agents, on this model, usually also has the function of absolving other agents who might have been candidates for fault” (ibid).

There is a sense in which this last observation seems like a truism. Where responsibility is assigned to some person or group of persons, then those who are not that person or within that group are, almost by definition, not responsible. They are, to use Young’s term, absolved from responsibility. But Young’s point is intended to go deeper than that. It is concerned with a genuine attempt to understand and engage with the complex and varied sources of actions and events within which harms

\* School of Law, University of Glasgow. This paper draws on work that was first discussed at a workshop on Complicity held at the Altonaer Stiftung für philosophische Grundlagenforschung in Hamburg in October 2006. I am immensely grateful for the opportunity afforded by the Stiftung to present that work in progress, and in particular to Valerie Kerruish and Uwe Petersen for the invitation to do so. I am grateful to them and Emiliios Christodoulidis, Tarik Kochi, and Stewart Motha for their detailed comments on that occasion.

are caused, an attempt that many moral and political approaches simply do not make. (Young's primary concern in this article is with trying to re-think, let us say at the least, the *involvement* of consumers in the propagation of harms — whether human or environmental — through their consumer activities in the context of global labour and consumption practices.) Hence in reviewing an analysis offered by Onora O'Neill, Young claims that where 'we' benefit from, or at least contribute to by our actions, processes that are inevitably complex, such as the global division of labour, then we nonetheless have responsibilities, or are in a relation of implication, with respect to being a part of these processes themselves. But where the fault liability model would limit, or, what is more likely exclude responsibility entirely, and hence work its absolution of 'us', Young proposes instead not abandoning the model, but paying more attention to the alternative notion of *political* responsibility. According to this, and drawing, she says, on the work of Hannah Arendt, individuals are responsible "precisely for things they themselves have *not* done. The reason to assume political responsibility involves not individual fault, but derives from 'my membership in a group (a collective) which no voluntary act of mine can dissolve ...'" (375: quoting Arendt).

What appears radical already from such a short description is the dual sense of an actor's *implication* in harms that is not, on the one hand, related to intentional or negligent action *by* the actor; and on the other, the claim that any excuse from responsibility, based on the non-voluntary nature of the membership of a group, is delegitimated. Underpinning both these claims is the attention Young says is required to query the 'normality' which is assumed, or taken for granted, with respect to the more conventional understandings of the fault model. That is, as Young notes, in the legal liability model, "what counts as a wrong ... is generally conceived as a deviation from a baseline. Implicitly we assume a normal background situation that is morally acceptable, if not ideal" (377). Yet, in the alternative, she suggests, "A concept of political responsibility in relation to structural injustices ... evaluates not a harm that deviates from the normal and acceptable, but rather often brings into question precisely the background conditions that ascriptions of blame or fault assume as normal." Such a critique will be justified in so far as it understands that the assessment of what is 'normal' operates *routinely* — that is why it is *normal* — to make invisible certain forms and experiences



of harms by setting them as background conditions where the matter of ‘deviation’ or breach does not arise with respect to them. Hence at this point, Young explains, “When we judge that structural injustice exists, we are saying precisely that at least some of the normal and accepted background conditions are not morally acceptable” (378).

But what can this mean? If I have sympathy with Young’s analysis as I have briefly reported it here, then that sympathy is suddenly given critical provocation by a stark question: What is the significance of that word ‘*morally*’? In the first instance, what meaning, if any, is lost if that word is dropped? Alternatively what meaning, if any, is gained by it being there? Let me make a few comments with respect to this issue which suggest profound problems with the role of ‘morality’ in situations of the type that Young is referring to.

First consider the way in which ‘morality’ seeks to carve out a specific difference within the possibilities of what might otherwise be candidates for ‘unacceptability’. The practices being referred to here are those which form the ‘background conditions’ within which, it is claimed, ‘structural injustices’ exist. But if these conditions are themselves complex, then what Young has ably discerned is that part of that very complexity involves observing an inability of ‘morality’ to discern unacceptability based on the postulated differences between ‘normal’ and ‘exceptional’. The issue here is a general one. Young argues that instead of concentrating on, for example, finding causal agents, we should concentrate on seeing that “those participating in the production and reproduction of the structures should recognize that their actions contribute along with those of others to this injustice, and take responsibility for altering the processes to avoid or reduce injustice” (379). The problem, however, when put in these terms, is precisely this: that those responsible are frequently and most likely to be those who don’t see the injustice *as their* problem; or even more acutely put, see it as *not being a case of injustice* at all, because it is an instance of some other form of acting the operating conditions of which are in fact thought to be *just*. (For example, the EU may want to respond to ‘moral’ claims to incorporate labour or human rights standards in its external relations, but it still, famously, gives more subsidy to its cows than it does to the impoverished of the third world.) Most commonly, and certainly in the types of case Young is referring to, these conditions and forms of acting are nothing less than those of a functioning capitalist

market, one of the key attributes of which is that the relations and rules of engagement take place and are understood by the actors themselves in economic and not political terms. For this reason *political* responsibility, where it is given any consideration at all, will be either denied or, if it is articulated, have already been invoked as just.

If ‘political responsibility’ faces this severe difficulty then it is also clear that *moral* disapprobation (whatever that means) may tinker at the edges, or assume whatever mantle suits best — charitable obligation, noblesse oblige, the language of regret — but it will in no way open to critique the kind of massive, and detailed, structural transformations required properly to address these injustices. Hence when Young writes that, “We share responsibility to fashion organized means of changing how the processes work so they will issue in less injustice” (380), she fails to acknowledge that many of the major players here — if we are to use such an actor-based notion, then these are major in terms of capital clout rather than anything so quaint as people or peoples — explicitly *denounce* such ‘sharing’, or broad participation, as anathema to the proper functioning of the economy (as observable, for example, with respect to the constituency of the summits of the World Economic Forum at Davos) or, when they do not so denounce, do not in any way open for negotiation the ‘background’ conditions that the very political instruments — state, EU, or WTO say — are themselves constitutively disallowed from changing *as* background conditions.

And this, ironically, returns us to the legal liability model: deviation from the norm there was correctly understood as the model for assessing responsibility, but with political responsibility, the normal action still has to be seen to be a deviation in order to recognise it as *injustice*; and yet to the extent that the structures and the political institutions implicated in this themselves use ‘normal’ mechanisms to consolidate their position — whatever form these may take (and they are many: the ‘baseline’ constitution of global capital; guaranteed property/profit rights such as TRIPs; bi-lateral investment treaties, etc.) — then those actors most prominently involved in fact consider that *these are* the politically responsible ways of acting. And what is perhaps most powerful of all about the securing of this ‘normality’ — and we will return to this shortly — is not merely, and not only, that they are produced and reproduced as economic, and thus

essentially background and non-political, but that they are so commonly inaugurated, sustained and legitimated as *legal*: as *the law*.

If we are sympathetic to the cause Young espouses — to reduce the suffering consequent on the workings of a global capitalist economy — we may nonetheless wonder whether, despite its endeavours, a lack of sociological and political depth may work to limit the initial impetus of her analysis. The analysis would, one is tempted to say, remain super-structural. Perhaps this fact explains why, despite arguing that the share of political responsibility is not equal amongst actors, Young rejects the idea that we should concentrate — as even Thomas Pogge suggests — on those who design the global economic order; for this would mean, she suggests, that it lets too many of ‘us’ off the hook of our responsibilities. This is so even where it is the case, Young claims, referring to Liam Murphy’s work, that “so long as the society in which we live is far from the ideal of justice, moral demands on individuals will be rather stringent and perhaps difficult to meet” (383–4).

Of course, we encounter here exactly the same difficulty noted a moment ago, though expressed now in a different way. For the great giveaway with the line just quoted is to be observed in those apparently tentative or innocuous terms ‘rather’ and ‘perhaps’. They tell us a great deal because they are as far from tentative and innocuous as is it is possible to be; on the contrary, they are virulent. What is ‘rather stringent’ about these ‘moral demands’ if it is not an open admission of a readily defeasible obligation? What is ‘perhaps difficult’ to meet about them if not the fact that they will just not be met? Instead, by ‘rather’ and ‘perhaps’ are best understood the reality that the ‘moral’ part of the ‘moral demands’ on individuals are not at all stringent, nor intended to be, nor *could* they be. For they play out in entirely the wrong register, where wrongs will not register except as regret, as options, as perhaps, perhaps, perhaps. The nature of the ‘moral’ enterprise here is (necessarily) to mis-understand the nature of the experience of suffering from the perspective of which so many ‘perhaps’s are in reality just so many blank stares, ‘blind eyes turned’, so many determined, doggedly-defended impunities.

Young acknowledges that often “the problem is not that people are failing in the performance of their tasks; on the contrary, they are doing their jobs rather well. The problem is the way that the institutions are defined, their power, purposes and interactions with one another, as well

as how they define tasks to fulfil those purposes" (384–5). This seems true in an important sense, that is, one which will seek to grasp an understanding of the lived conditions of pernicious exploitation. But will an overhaul of these conditions come from the result of us — ‘us’ again — coming to ‘our’ senses *morally*? (Are ‘we’ really all just so dumb?) Young claims that the transformation of institutions, “is everyone’s task and no one’s in particular” (385), and yet it is precisely this that the ‘shrewd jargon of morality’ — to borrow Philip Roth’s coruscating formulation — does and will negotiate so brilliantly, turning ‘demand’ into ‘absolution’ via ‘perhaps’. Hence despite it being said that different persons “stand in differing positions in structures that produce unjust outcomes”, differences Young considers according to “connection, power, and privilege”, there is once again a failure, for much the same reasons as noted before, to see any necessary responses as being irrelevant in so far as they are ‘moral’; something that can be recognised clearly from the conclusion: “Persons who benefit relatively from structural inequalities have special moral responsibilities to contribute to organized efforts to correct them, not because they are to blame for them, but because they have more resources and are able to adapt to changed circumstances without suffering serious deprivation” (387).

The sad inadequacy, the empty elitism of that ‘special’ moral obligation, is again a tell-tale symptom, although this time one in which the ‘rather’s and ‘perhaps’s become formulated more powerfully in as much as they must be attentive to the care that must be taken not to suggest — god forbid — that those who benefit are in any way ‘to blame’, or that in seeking change their resources would in any way be threatened such as to cause them to suffer ‘serious deprivation’.

What is admirable about, indeed the reason for considering, Young’s analysis, is that it does seek to understand that a primary attentiveness be given to the experience of suffering, particularly that associated with extensive poverty, exploitation and immiseration in a global context of economic structures that simultaneously produce unprecedented wealth for the few. But what I have suggested, albeit briefly, is that despite attempts to move and think beyond the academic propensity to keep their pristine concepts well-removed from the diseased scourges of suffering, the language of political, and certainly moral, responsibility fails adequately to comprehend the constitution of forces that maintain this suffering.

More than that even: that this language remains *complicit* in its reproduction. But does this mean that the very language of responsibility, no matter what content it may take on, is suspect, at least with respect to suffering consequent on this kind of scale of organisation? No doubt one of the strongest statements of a rejection of the notion that a search for responsibility is important with respect to understanding and changing capitalist modes of exploitation is given by Marx in the preface to *Capital* when he writes:

Individuals are dealt with here only in so far as they are the personifications of economic categories, the bearers of particular class-relations and interests. My standpoint, from which the development of the economic formation of society is viewed as a process of natural history, can less than any other make the individual responsible for relations whose creatures he remains, socially speaking, however much he may subjectively raise himself above them (Marx, 1990: 92).

Or alternatively, does the language of responsibility — at the very least in its etymologically minimal sense of being a call for a *response* — not mark the existence of an intuition the rejection of which would be to ignore or negate potentially worthwhile forms of motivation and action in the face of such suffering?

I want to come at these questions through a route that will not line them up for adjudication, but rather see how there are important sociological and conceptual factors that need first to be understood in a way that might in fact turn out to undermine the very questions. Specifically, I want to pursue the ways in which *legal* modes of organisation have come to play a central role in organising responsibilities whilst simultaneously providing means of negating responsibilities. But this process, I suggest, is more profound than Young's sense of absolution through law, because it concerns not only modern law's conceptual abilities, but its social pervasiveness and dominance. And it is this that explains further why a shift of emphasis towards moral or political responsibilities as a suggested means of redress, is misguided.

## Law and organised irresponsibility

For C W Mills, with whom, so far as I can tell the term ‘organised irresponsibility’ originates, it is “Within the corporate worlds of business, war-making and politics, [that] the private conscience is attenuated — and the higher immorality is institutionalized” (Mills 1956: 343). Mills calls this “structural immorality” (ibid), the “mindlessness of the powerful that is the true higher immorality of our time; for, with it, there is associated the organized irresponsibility that is today the most important characteristic of the American system of corporate power” (342). Consistent with the problems just raised with Young’s analysis, Mills had observed that, “In economic and political institutions the corporate rich now wield enormous power, but they have never had to win the moral consent of those over whom they hold this power” (344). Instead, “recurrent economic and military crises spread fears, hesitations, and anxieties which give new urgency to the busy search for moral justifications and decorous excuses.” If crisis is an overused, devalued term, he suggests, it is largely because it has been used by people in power “to cover up their extraordinary policies and deeds. For genuine crises involve situations in which men at large are presented with genuine alternatives.” The “higher immorality”, the organized irresponsibility “have not involved any public crises; on the contrary, they have been matters of a creeping indifference and a silent hollowing out” (345).

As I suggested earlier, it may be in no small measure due to the legitimization function offered by legal norms and institutions that crises are not seen for what they are, at least with respect to those whose suffering is not acknowledged, and that the very legality of actions is a key mode of facilitating the fact that responses are commonly those, as Mills so tellingly puts it, of ‘creeping indifference’. In order to explore this a little more, let me draw briefly on the work of Ulrich Beck and his analysis of risk society and its environmental dangers, to show one prominent way in which this may occur. Beck has taken up these themes in a way that explicitly connects the role of law to the organisation of irresponsibility. This is, importantly, a far more devastating account of law’s role than that of a mere inadequacy associated with the fact that a fault liability model dominates law’s modes of cognition, action and organisation. Instead it suggests that law and legal institutions are themselves complicit in the ongoing production and normalization of global risks, and that this

works consistently — that is, in an organised manner — to favour certain interests to the exclusion of others. If so, then we are confronted not only by the possibility of law's inability to confront global hazards, but by the fact that law and legal institutions operate to promote them.

If modern industrial society was centrally concerned with the distribution of goods, argues Beck, then risk society is concerned with the distribution of 'bads'. But such distribution takes place according to certain dynamics, including importantly that of legal regulation. Yet in this instance, instead of law operating, as would more commonly be understood, to regulate responsible behaviour and impose consequences for irresponsible behaviour — that is, to set standards and impose sanctions consequent on their breach — the actual operation of law turns out to be a way of organizing *irresponsible* behaviour. In other words dangerous actors — be they states or corporations — carry out activities that contribute to global risk, but these harmful activities are legalized through a regime that claims still to be there to protect people and the environment from such harms. The reason this happens is because the principles law has for organizing responsibility — in particular those regarding causation, individual liability, and proof — are rooted in an earlier form of modernity and have not caught up with the nature and extent of the dangers now being faced. Yet their operation is *not* simply inadequate but contributes decisively to the problem. This, says Beck, is the “concealed gap of the century”, and in the context of ecological harms, leads to nothing short of what he labels “legalized universal pollution”:

dangers worldwide make it harder to prove that a single substance is the cause; the international production of harmful substances works against proving the culpability of a single company or perpetrator; the individual character of criminal law contradicts the collective danger; and the global character of the danger has abolished “causes” as our industrial forefathers understood them (Beck 1995: 131–132).

By assigning responsibilities only under certain well-established legal categories, what modern law and legal institutions do then is organise a system of *non*-liability, a regime of unaccountability: environmental harms continue at an alarming rate, and yet either no-one at all can be or is held responsible for them, (or everyone and thus no-one is responsible for causing them), or responsibility is massively asymmetrical to the harms

caused (small fines, say, for massive damage caused). Here then is how Beck sums up the overall effect:

If one wanted to think up a system for turning guilt into innocence, one could take this collaboration between justice, universal culpability, acquittal and pollution as one's model. Nothing criminal is happening here, nothing demonstrably criminal anyway. Its undemonstrability is guaranteed precisely by compliance with, and strict application of, the fundamental rule of justice — the principle of individual culpability, whereby both pollution and non-pollution, justice and (coughing) injustice, are guaranteed (ibid 135).

It is in exactly this way that modern law is not simply inadequate to the task of combating the dangers of risk society, but is rather also directly involved in *perpetuating* their ongoing development. This, says Beck, shows how law in risk society engineers a global system of *organized irresponsibility*.

Drawing on this graphic example of legal complicity in the production and legitimation of harms, we are able to work, I suggest, towards a fuller understanding of the role of law with respect to the legitimation of suffering more generally. As I have argued more extensively elsewhere (Veitch, 2007), the legal organisation of irresponsibility in this respect relies on three very conventional features of law under contemporary social conditions, the effects of which are no less significant for these being conventional or normal — indeed that is precisely the point.

These features are: law's claim to correctness, its force, and its social priority. According to Alexy's analysis (2004) legal normativity involves a socially effective institutionalised force, and the claim that this force is right or just (the claim to correctness). And it is because of the former (socially efficacious coercion), that the latter (correctness) necessarily extends its reach beyond matters of merely internal legal validity and makes its claim to legitimacy effective in the terrain of social relations more generally. It is this then that constitutes the third feature of law in contemporary society. That is, given law's claim to correctness and its enforceability, law attains a level of priority and prominence in social life and its normative hierarchies even when its effects may in fact be claimed to be unjust.



If we consider a further sociological dimension to this, then where in contemporary society there exists an increase in formal law, that is, where under conditions of *juridification* there is an expansion into and densification of legal normativity within the fabric of social relations (Habermas 1987), then that double structure of coercion and right increasingly imbricates those social relations with the force and correctness of law. A necessary effect of this is that to an increasing degree the juridical form permeates, structures and organises the available range of normative understandings, expectations and responses in society generally. Consequently, as Habermas himself has observed in his later work, it has come to be the case in modern society that law increasingly becomes “the only medium in which it is possible reliably to establish morally obligated relationships of mutual respect even amongst strangers” (Habermas 1996: 460), with the result that positive law “*relieves* the judging and acting person of the considerable cognitive, motivational and . . . organisational demands of a morality centred on the individual’s conscience” (452, emphasis in original). It is this that sees law’s measure increasingly predominate in matters concerning controversial aspects of morality and politics including even the basic elements of the reproduction of social life — or their denial. It is this that forces us to think not of the influence of morality or politics on law, but rather the reverse: the influence of, indeed the determination by, law because of its social effectiveness, on other conventional sources of norms. For it is with this observation, moreover — to return to our earlier critical remarks — that we can now see more fully why the claim about ‘moral’ and ‘political’ responsibility, as supposedly independent sources of motivation, is deeply implausible.

Given this we should understand that the reach and force of law’s normativity is necessarily a most powerful and extensive means of organizing the very material practices of responsibility and irresponsibility, and so must be understood as being implicated in, rather than merely indifferent to, the production of suffering. In other words, law’s measure — its normative organising, force, and claim to correctness — stands in a relation of *implication* to the commission of mass suffering and its normalization, whether that is with respect to environmental devastation or suffering brought about through legalized practices of economic exploitation. Moreover, and returning to Beck’s insights, law and legal institutions are implicated in the very disavowals that establish legitimate impunities.

Hence to the extent — and it is indeed a large extent — that law does in fact dominate the field of responsibility in a socially effective and legitimate manner, then it is in precisely these more profound respects that we should understand the limitations of discourses of responsibility. For it is the failure to grasp the constitutive role of law here which in large part accounts for the sense in which claims about moral and political responsibility under conditions of global capitalist exploitation appear super-structural: they are not sufficiently able to engage with the possibility that it is the very ‘normality’ of what Young earlier termed the ‘background conditions’ that poses the greatest threat to human and environmental well-being. But this is *not* because they are, in some vague sense, ‘morally unacceptable’; rather something far more troubling is signified. Put at its most extreme (for it is the extremity, that is, of the ‘normal’), and returning for a final time to Beck, it involves realising that

This threat to all life does not come from outside; it does not begin with the exceptional case, war, the enemy. It emerges within, enduringly, as the reverse side of “progress”, peace and normality. Those responsible for safety and rationality, and those who threaten it most, are no longer isolated from one another by national or group boundaries, polarized by the roles they play, *but are potentially one and the same* (Beck 1995: 163, emphasis added).

In case one is tempted to think that this expresses something specific to the environment, as opposed to the harms of legally organised economic activity, we need only be reminded of some of the realities of what Eduardo Galeano (2000) has called our ‘upside down’ world to see just how normal the madness has become: where “The countries that sell the world the most weapons are the same ones in charge of world peace” (113) (“fortunately for them” Galeano adds, “the threat of world peace is receding”); where in Latin America, he writes, but it is not of course only there, “the streets and avenues tend to bear the names of those who stole the land and looted the public purse” (201); where “the most successful companies in the world are the ones that do the most to murder it” and which, “in the name of freedom make the planet sick and then sell it medicine and consolation” (215); where debt is “something even those who have nothing have” (247–8); and where, generally, “Never have so many suffered so much for so few” (236). To maintain this ‘upside down world’

requires an enormous bureaucratic, legal and administrative effort, both at an ideological and material level. It requires nothing short, in other words, of the unerring force that only normatively organised coercion, that is, that only *law*, can offer.

In the global context, which is, as Young rightly discerns, the context within which it is necessary to think of these issues, Lindqvist makes this point with great articulacy: “More and more people”, he writes “are being born into poverty, ignorance, and hunger. More and more people are born superfluous, worthless to the interdependent economy, yet still vulnerable to its effects. More and more are born for whom violence is the only way out.” But that violence is and has historically been met, not by reason, nor morality, nor by political responsibility; *except*, that is, to the extent that all these are implicated *in* the response, a response which is made in the only register with which it is meaningfully commensurate, namely normatively organised violence itself:

Throughout this century, it has been clear that the standard of living enjoyed in industrial countries cannot be extended to the world’s population. We have created a way of life that must always be limited to a few. These few make up the broad middle class in a few countries and a small upper class in the rest. The members know each other by their buying power. They have a common interest in preserving their privileges, by force if necessary. They, too, are born into violence . . . Global violence is the hard core of our existence (Lindqvist 2001: sections 397–398).

And the measure of its success is the extent to which it is organised legally.

Here is a simple question: is it not unlikely that a legal system and its concepts which have come to prominence in modernity, which have co-existed with slavery, colonial brutality, and economic immiseration, and which continue to exist — despite claims of equality — with most of these, and which make up a prominent and indeed valued institution in such societies, is it not unlikely that this institution is not in some way involved in these very harms? To say that it is not so implicated historically is a denial of the historical reality, and few would sensibly hold it to be true. But to say that it is not implicated now, given the immiseration and violence of the contemporary global order is surely equally implausible. And yet, the dominant mode of thought and action does not and will not accept this designation, and it does not and will not do so, I have suggested,

in large part because of the legitimacy that legality bestows. This legitimacy is the embodiment of what Stanley Cohen has called ‘magical legalism’ which is, he said, “a method to ‘prove’ that an allegation could not possibly be correct because the action is illegal.” For example, a state may have signed up to torture conventions which make torture illegal, and so — the “magical syllogism” — whatever horrific practices the state is engaged in cannot be torture. These kinds of “legalistic moves”, says Cohen, “are wonderfully plausible as long as common sense is suspended” (Cohen 2001: 108).

Such suspension of disbelief is one that also infects an optimistic view of our times, one that would seem to prefer ignorance of the fact that, despite ours being supposedly the age of equality and human rights, “Our era has witnessed more violations of their principles than any previous, less ‘enlightened’ one. Ours is the epoch of massacre, genocide, ethnic cleansing, the age of the Holocaust . . . No degree of progress allows us to ignore that never before in absolute figures, have so many men, women, and children been subjugated, starved, or exterminated on earth” (Douzinas 2002: 20).

Yet to the extent that so many of the harms of the past and present century were, and continue to be, *not* legal violations, that is, continue to be perfectly legal, then this is why, it seems to me at any rate, it is so important to understand that the kind of harms to which we need to pay far more attention are those in which the law is *itself* implicated. That is, not crimes and delicts, but legal behaviour itself. This involves paying attention to law’s *successes*, not its impotence. It involves paying attention to the fact that law is both culpable in the sense of engaged in the active commission of harms and not culpable at exactly the same time. For this is precisely how the disappearance of responsibility for extensive harms succeeds, and it is not through law’s failure but amongst the successful practices of legal responsibility. This is how an asymmetry between suffering and establishing responsibility for it operates in a legalised manner, and it is, finally, how legally *caused* harms, not responses to them, should best be understood: *not* as impotence, but as the system *working*.

To accept this requires a fair degree of upsetting of conventional viewpoints, for it requires an acknowledgment that that which is commonly believed can, and sometimes does, promote social goods and benefits, also

and at the same time is capable of producing extensive harms and covering these up. But this is exactly how the ‘laws of irresponsibility’ work.

### Motivating responsibility?

Generals’ strategies, it is often said, are based on preparations designed to fight the last war. There is something of this which is reminiscent when it comes to the grand legal codes of the modern era. These codes — whether of the Enlightenment revolutions or of the European or United Nations declarations of human rights — have all emerged from amid the rubble of human devastation and the broken bodies of human violence and degradation, and even then, representing only a segment of its total. These codes, so commonly now universal — that is, applicable across all times and spaces — are the juridical embodiments of the innumerable litanies of *Never again*, the legal responses to the so many solemn declarations of *Nunca mas!* and *Lest we forget*.

But spare this a thought: if these laws and legal codes are really like that, if they really do embody the highest aspirational ideals of universal right and humanity; if they really do provide us with the *most* progressive standards of human decency, then what, after the *next* full scale horror *will the set of legal codes have to embody?* The response to this is likely only to be visceral, for one dare not begin to imagine. Yet against such potentially apocalyptic scenarios, is not responsibility also still a worthwhile response? I have suggested that the field of responsibility is marked so heavily by legal forms and expectations that it is inevitably a suspect notion in terms of addressing the existence of suffering consequent on humanly organised action. In particular this is the case with respect to the unexamined motivations attributed to moral and political responsibility. But, to the extent that an understanding of responsibility practices — or indeed of legally regulated practices of irresponsibility — is necessary to see how such organisation works, it is, nonetheless, still an important, perhaps even essential, area of enquiry with respect to the assessment of the nature, and hence possible alleviation, of harms.

Beyond that, perhaps one will take inspiration from all kinds of other sources, but we should not pretend that they are ‘moral’, or ‘political’, as if that meant some form of autonomous or independent reason or motivation. Instead, perhaps one can turn one’s face away, if only momentarily,

to other rhythms; times and timings that Orwell captured with such enduring eloquence in his ‘Some Thoughts on the Common Toad’:

At any rate, spring is here, even in London N.1, and they can’t stop you enjoying it. This is a satisfying reflection. How many a time have I stood watching the toads mating, or a pair of hares having a boxing match in the young corn, and thought of all the important persons who would stop me enjoying this if they could. But luckily they can’t. So long as you are not actually ill, hungry, frightened or immured in a prison or a holiday camp, spring is still spring. The atom bombs are piling up in the factories, the police are prowling through the cities, the lies are streaming from the loudspeakers, but the earth is still going round the sun, and neither the dictators nor the bureaucrats, deeply as they disapprove of the process, are able to prevent it.

True enough, no doubt; but as Orwell knew, one cannot avert the gaze for too long.

## Bibliography

- Alexy, R. (2004), ‘The nature of legal philosophy’, *Ratio Juris* 17: 156–167.
- Beck, U. (1995), *Ecological Politics in an Age of Risk*, trans. A. Weisz, Cambridge: Polity.
- Cohen, S. (2001), *States of Denial: Knowing about Atrocities and Suffering*, Cambridge: Polity.
- Douzinas, C. (2002), ‘Postmodern just wars: Kosovo, Afghanistan and the new world order’ in J. Strawson (ed) *Law after ground zero*, London: Glasshouse, 20–36.
- Galeano, E. (2000), *Upside Down: A primer for the looking glass world*, trans. M. Fried, New York: Picador.
- Habermas, J. (1987), *The Theory of Communicative Action, Vol. 2 Life-world and System*, trans. Thomas McCarthy, Cambridge: Polity.
- Habermas, J. (1996), *Between Facts and Norms*, trans. W. Rehg, Cambridge, Mass.: MIT Press.

Lindqvist, S. (2001), *A History of Bombing*, London: Granta.

Marx, K. (1990), *Capital Vol. 1* 'Preface to the First Edition', Harmondsworth: Penguin.

Mills, C. W. (1956), *The Power Elite*, New York: OUP.

Veitch, S. (2007), *Law and Irresponsibility: On the legitimation of human suffering*, London: Routledge-Cavendish.

Young, I. M. (2004), 'Responsibility and Global Labour Justice' *Journal of Political Philosophy* 12: 365–388.





# Auslobung

Im Rahmen ihres Entwicklungsprojekts einer spekulativen Logik setzt die Altonaer Stiftung für philosophische Grundlagenforschung ein Preisgeld von € 5.000,- für einen

Beweis der Widerspruchsfreiheit von  $\mathbf{LD}_\lambda^Z$

aus.  $\mathbf{LD}_\lambda^Z$  ist dabei die formale Theorie, die in

PETERSEN, Uwe: Logic Without Contraction as Based on Inclusion and Unrestricted Abstraction. *Studia Logica* 64 (2000), 365–403,

auf S. 392 definiert wird. In dieser Arbeit wird auch ein möglicher Ansatz für einen Widerspruchsfreiheitsbeweis durch partielle Schnittelimierung skizziert. Es besteht keine Verpflichtung, diesem Ansatz zu folgen, aber es wird erwartet, daß aus einem vorgelegten Widerspruchsfreiheitsbeweis eine Abschätzung der ordinalen Stärke von  $\mathbf{LD}_\lambda^Z$  hervorgeht.

Teilnahmeberechtigt sind natürliche Personen sowie deren Zusammenschlüsse.

Eine Frist zur Einsendung von Beiträgen besteht nicht. Einsendungen werden in der Reihenfolge ihres Eingangs bei der Altonaer Stiftung für philosophische Grundlagenforschung durch eine vom Stiftungsvorstand berufene Jury begutachtet. Der Preis wird an die erste Einsendung vergeben, die nach Ermessen der Jury die Anforderungen erfüllt. Der Rechtsweg ist ausgeschlossen.

Mitteilungen über den Stand von Einsendungen sowie die Vergabe des Preisgeldes werden auf [www.asfpg.de](http://www.asfpg.de) unter „Projekte — Entwicklungszentrum für spekulative Logik — Auslobung“ veröffentlicht.

Beiträge können in deutscher oder englischer Sprache abgefaßt sein und auf postalischem oder elektronischem Wege eingereicht werden.

Postadresse:

Altonaer Stiftung für philosophische Grundlagenforschung  
Ehrenbergstr. 27  
22767 Hamburg  
Germany

E-mail-Adresse: [preis@asfpg.de](mailto:preis@asfpg.de)

Alle Rechte von Einsendungen bleiben bei den jeweiligen Autoren.

## Offer of a Prize — Call for Entries

Within the project of the development of a speculative logic, the Altonaer Stiftung für philosophische Grundlagenforschung (ASFPG) offers a prize of € 5,000.– for a

Consistency proof of  $\mathbf{LD}_\lambda^Z$ .

The system  $\mathbf{LD}_\lambda^Z$  is defined in

PETERSEN, Uwe: Logic Without Contraction as Based on Inclusion and Unrestricted Abstraction. *Studia Logica* 64 (2000), 365–403,

on p. 392. In this paper a possible approach to a consistency proof by way of a partial cut elimination is sketched. There is no obligation to follow this approach, but it is expected that any consistency proof submitted will provide an estimation of the ordinal strength of  $\mathbf{LD}_\lambda^Z$ .

All natural persons and groups of natural persons are eligible for participation.

There is no deadline for entries. Entries will be judged in the order of their receipt at the Altonaer Stiftung für philosophische Grundlagenforschung by a jury to be appointed by the Board of the Stiftung. The first entry giving a proof judged to be correct by the jury will be awarded the prize. Recourse to legal action is excluded.

Information on the number of entries under consideration and on the award of the prize will be given at [www.asfpg.de](http://www.asfpg.de) under “Projects — Centre for speculative logic — Prize”.

Entries can be made in German or English. They can be submitted by post or by electronic mail.

Postal address:

Altonaer Stiftung für philosophische Grundlagenforschung  
Ehrenbergstr. 27  
22767 Hamburg  
Germany

Email: [preis@asfpg.de](mailto:preis@asfpg.de)

All rights in submissions remain with the authors.