On the Conception of Fundamental Time Asymmetries in Physics

Inaugural-Dissertation
zur Erlangung der Doktorwürde
der
Philosophischen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität
zu Bonn
vorgelegt von

Daniel Wohlfarth

aus

Köln

Bonn, 2013

Zusammensetzung der Prüfungskommission:
Prof. Dr. Sturma
(Vorsitzende/Vorsitzender)
Prof. Dr. Bartels
(Betreuerin/Betreuer und Gutachterin/Gutachter)
Prof. Dr. Meißner
(Gutachterin/Gutachter)
Prof. Dr. Brendel
(weiteres prüfungsberechtigtes Mitglied)

Tag der mündlichen Prüfung: 05.02.2013

Danksagung

Herrn Prof. Dr. Andreas Bartels möchte ich überaus herzlich für seine immense Unterstützung in unzähligen Bereichen sowie für die Überlassung und die Erarbeitung des Dissertationsthemas danken.

Herrn Prof. Dr. Ulf-G. Meißner möchte ich für die viele Zeit und Mühe danken welche er als Zweitkorrektor dieser Dissertation investiert hat.

Außerdem möchte ich Herrn Prof. Dr. Jeremy Butterfield danken, welcher mich in der letzte Phase der Dissertation durch zahllose weiterbringende und fruchtbare Diskussionen unterstützt hat.

Für sehr hilfreiche und weiterbringende Diskussionen möchte ich außerdem Dr. Nazim Bouatta, Dr. habil. Cord Friebe, Dr. habil. Jacob Rosenthal, und Dr. Carsten Seck danken.

Für überaus hilfreiche Anmerkungen bezüglich verschiedener Kapitel gilt mein Dank außerdem Dr. Simon Friederich (Bezüglich Kapitel V), Stefan Heidl (Bezüglich Kapitel II), Arianne Shavisi (Bezüglich Kapitel III und V) sowie Dr. Lena Zuchowski (Bezüglich Kapitel IV).

Für Unterstützung in allen erdenklichen Bereichen danke ich ganz herzlich meiner Partnerin Eva Schiffgens sowie meiner Mutter Anna-Maria Wohlfarth und Jost Bornemann.

List of Contents

Chapter I

Introduction and Motivation

I.1. Introduction and Motivation

The question of how to understand the difference between the 'past' and the 'future' (if there is any) has always been a philosophically relevant question. The experience that time appears to flow (bringing the world from a state in the past to a state in the future by passing an ever-changing present) seems to be one of the most basic observations in human life. Thus, attempts have been made to formulate a philosophy of time that suggests that this directed time asymmetry (the flow) is a primitive property of time itself. However, this picture is imbedded in a field with many different issues and unsolved problems, some philosophical and others arising in the physical sciences.

According to the natural sciences, it appears to be a well-motivated view that the most fundamental models of nature are those provided by particular fields of physics. Of course, from a philosophical perspective, this could be rejected, but it need not be. Moreover, if we assume that fundamental physical models describe, even approximately, some properties of nature, these properties can be understood as the most fundamental properties described in scientific theories today. Note that this does not imply that the models of other sciences or other fields in physics are reducible to these fundamental models. My point is only that the view which says that physics describes some of the fundamental properties of nature is well motivated and attractive. The motivation for this investigation is based on this view. However, even if this investigation is motivated by the assumption that the most fundamental properties of nature, which are described in scientific theories, are described in *physical* theories, this does not mean that the investigation's outcome depends on this assumption. The main claim of the investigation is that some time asymmetric structures should be understood as a 'fundamental' property of the physical models currently used to describe nature. Success in proving this claim does not mean that the asymmetries are assumed to be a fundamental property of *nature*; neither does it mean that the asymmetries of time, if they exist, are assumed to be correctly captured by physical theories. The only point is that the asymmetries of time can be seen as a fundamental property of crucial, well-established physical theories and models. This claim is unaffected by any discussion regarding the question: 'Do physical models describe, even approximately, the properties of nature, and are those properties fundamental properties of nature itself?'

Thus, the interesting discussions of this question in the philosophy of science are considered only in some small parts of this investigation. Nevertheless, crucial questions regarding time directions arise mostly in light of the view that at least some fields in physics describe some crucial parts of nature correctly. If this view could be rejected in the first place, it would be attractive, given that the direction of time is a well-observed fact in everyday life, to assume that the direction of time is simply a primitive fact. The experience of the direction of time appears puzzling only if we assume that some crucial and fundamental structures of nature are captured, at least partly, by physics. The nature of the puzzle is revealed by the following observation: A closer look at the laws of fundamental physics shows that they are time-reversal invariant (or CPT-invariant[1]). That is, in the fundamental theories, at least in their standard formulations and interpretations, we find no *fundamental* physical difference between past and future according to the fundamental *laws* of physical theories. This, of course, draws an unsatisfactory picture. The basic issue is that the laws of fundamental physics seem to admit no substantial difference between the past and the future, but the future and the past seem different in everyday experience. Why should that be so?

We find different lines of thought regarding this crucial question. One is based on the fact that it is always possible to argue that fundamental physics captures the fundamental properties of nature incorrectly. If this were so, it would be possible that

a) the 'true' laws of nature (if they exist) are not time reversal invariant, or

b) a non-time reversal invariant formulation or interpretation of known physical laws captures the properties of nature more correctly.

Alternatively, c) fundamental physics cannot capture the real structures of nature, which *may be* time asymmetric.

I now provide a brief motivation for my view that none of these options seems attractive. First consider option c, for which there seems to be two crucial explications:

c1: Fundamental physics cannot capture the real structures of nature, which *may be* time asymmetric. Therefore it is impossible to say whether the time asymmetry of our everyday experience is based on a fundamental property of nature or only on the structure of the human mind or brain.

c2: Fundamental physics cannot capture the real structures of nature, which *may be* time asymmetric. But, the time asymmetry of our everyday life is a basic observation, and we should assume that such basic observations arise from real properties of nature; if fundamental physics cannot explain the origin of this property, it just shows

---

[1]But, the consideration of charge (C) and parity (P) seems unable to solve the problem. I will come back to this point later.

that some crucial properties of nature are not captured by physical theories. Nevertheless, the directedness of time should be assumed to be a fundamental property of nature itself.

Both explications of option c seem disentangled from the aim of this investigation for the following reasons.

Obviously, c1 is a possibility. However, if there is a way to understand the occurring of fundamental time asymmetries on the basis of the theories and models of physics, this understanding, in agreement with modern physics, would be attractive even if this does not guarantee that the fundamental and crucial structures of nature are correctly captured in the physical theories. In fact, this situation is identical regarding almost all properties described in physical theories. Hence, I do not think that this uncertainty is a sufficient reason to deny the fruitfulness of physical explanations for crucial observations in nature. The same, I think, should be assumed (at least prima facie) for the observation of the directedness of time. Hence, the search for an understanding of time asymmetries based on fundamental physics should not be abandoned, even if c1 is taken seriously.

According to c2, everyday life experience is assumed to capture the properties of nature more precisely than fundamental physics does. Although this could be the case, I think it is highly problematic. The problems arise not only because this implies that generations of physicists have spent their lives building sophisticated models of nature without success. More importantly, this view, I think, fails to explain that physics provides new predictions (not only in laboratory experiments but also in observations of the physical environment) and that technologies based on fundamental physical theories turn out to be realizable, at least approximately. From a philosophical point of view, therefore, option c2 seems unattractive. In this context, one motivation for this investigation is to show that an understanding of the crucial time asymmetries based on the fundamental theories of physics is possible.

Option b) also seems problematic. There are time asymmetric formulations or interpretations of some fundamental physical theories, but they do not seem to be motivated independently from the issue. In fact, they seem motivated by the issue at hand, which means that they were constructed to achieve the goal of formulating a time-asymmetric formalism in fundamental physics. In this situation, it is not clear which formulation (the time symmetric one or time asymmetric one) captures the properties of nature (more) correctly (if any). But, I shall argue that even without time asymmetric non-standard formulations or interpretations, an understanding of time asymmetries, based on fundamental physical theories, is possible. Hence, the issues that arise from taking the view b, so I will show in this investigation, can be avoided.

Moreover, we find that some time asymmetric formalisms (for example, the rigged Hilbert space approach) in fundamental physics are usually constructed by the following considerations.

1. In the considered standard formalisms of fundamental physics, time evolutions in both time directions are describable and allowed (symmetrically). So it becomes possible to 'cut out via hand' the possible evolution in the past direction.

2. Then the task is 'only' to find a coherent mathematical formulation of the remaining possible temporal evolution.

3. The result is a time asymmetric formalism in fundamental physics.

However, such formalisms seem ad hoc and unmotivated by independent reasons. Thus, in cases like the rigged Hilbert space approach, option b appears unattractive, at least from a philosophical perspective. Nevertheless, I shall argue in chapter VI that some time asymmetric formulations can be motivated independently by the analysis presented later. Specifically, in the rigged Hilbert space approach the time asymmetric formulation can be seen as motivated by a physical analysis if the right conceptual framework is used. Here, however, I will simply conclude that option b seems unattractive, as long as (as this investigation claims) there is a way of understanding time asymmetries as fundamental properties of the standard formulation and interpretation of physical theories.

Option a is always a possibility. However, it shifts the question only to a not-yet-formulated (or never formulated) physical theory. Therefore, this option should not be taken if other options are available to solve the issue, especially in the context of currently formulated theories of physics.

But, one different and prominent suggestion is that time asymmetries are surely *not* a fundamental property of the physical theories [see, for example, Price (1996)]. In this view, physics is taken seriously, and the experience of time in everyday life is assumed to capture only some other properties, which could be provided by the biological structure of the human brain. This option, I think, looks more attractive than the three views discussed above. Note that this option is slightly distinguished from option c1. The distinction arises from the fact that, according to c1 it is impossible to say if time asymmetries are understandable as a natural property or not, whereas e.g. Price (1996) argues that there are good physical arguments to assume that the experience of time directions in everyday life is provided from other structures and in particular *not* from physics.

Nevertheless, by recognizing how deeply imbedded the notion of 'past' and 'future' (and also the notion of a fundamental difference between past and further) is in our everyday experience, it would be a more satisfying option to base the time directions on fundamental

properties of the physical theories used to describe nature. This investigation claims to support this view even if the standard formulation and interpretation of fundamental physics is assumed and even if, in this formulation and interpretation, all fundamental physical laws are time reversal invariant (or CPT invariant). In chapter II, I propose a possible understanding of 'fundamentality' of time asymmetries, which is based not on time-reversal variance of fundamental physical laws but on the structure of their solution sets. However, before I come to this proposal in chapter II, I discuss some other aspects of the investigation.

In different physical models, we deal with different time parameters. In general and special relativity, we deal with proper time, which is the fundamental time coordinate in relativistic physics. In the Newtonian limit, those proper time coordinates become approximately the Newtonian background time from non-relativistic physics, which seems to describe most experiences in everyday life. According to cosmology, however, there is also cosmic time, which is an important time coordinate. In cosmological models, the cosmic time parameter, if it is definable in a particular spacetime (which seems to be the case in our actual universe), plays a similar (although not identical) role to Newtonian time, and it is connected to the fundamental proper times of different world lines (and not only by a non-relativistic approximation). Thus, according to physics, we find at least two interesting time parameters, which could be directed or not. It is also noteworthy that, according to physics, in a non-time orientable spacetime, which is allowed according to the Einstein equation, the proper times, even of parallel world lines, can have opposite directions. Thus, regarding proper times, we find that every world line, and thus every elementary physical system, can have its own time direction, which would be valid only in a local environment of a given spacetime point on one particular world line. This notion of a 'fundamental' time asymmetry (fundamental, because this notion is based on proper times) surely cannot capture some intuitive requirements for a 'fundamental' time asymmetry. One of these requirements, which I think is reasonable, is that a fundamental time asymmetry should be valid for at least a spacetime region that captures most parts of our environment in our particular universe without switching the alignment. I shall argue, following Castagnino, Lara and Lombardi (2003) and Castagnino and Lombardi (2009), that the asymmetries of cosmic time are *fundamentally* imbedded in the models of cosmology (chapters III and V). Also, I will argue that asymmetries of proper times arise in many physical contexts and can be seen as consequences of a time asymmetric energy flux in spacetimes similar to ours. But, in contrast to Castagnino and Lombardi (2009), I shall argue that the crucial question as to whether the proper time asymmetries can be seen as fundamental too, will remain unsolved in this investigation. However it will be shown that the time asymmetries of proper times, understood as consequences from a time asymmetric energy flux, produce an understanding of many properties of many prominent time arrows, even if the crucial question of fundamentality will not be solved but only revealed and formulated in this investigation.

Thus, this investigation may propose a counterintuitive picture, as follows:

The fundamental time asymmetry appears in cosmic time. In contrast, the time asymmetry of the fundamental time coordinates appears as non-fundamental and only understood as consequences of fundamental time asymmetries if some crucial questions regarding the connections between the cosmic and proper time asymmetries could be solved, as we will see. I will show this in greater detail in chapters III, IV and VI.

I.1.1. Arrows of Time in Physical Models

In discussions regarding the formulation of time arrows in physical models, some prominent examples are often discussed in physics and in the philosophy of physics and time. These examples are the time arrow in cosmology, the arrow of radiation, the arrow of time in thermodynamics and the arrow of time in quantum mechanics. In chapters III, IV, V and VI, I develop an alternative understanding of those time arrows motivated by the suggestion in chapter II (the suggestion of a new understanding of the term 'fundamentality' in the context of time asymmetries). The alternative understanding of these time arrows is distinguished from the traditional understanding in many ways. From the philosophical point of view, this differentiation occurs most importantly by

> a) the possibility that the time asymmetries (arrows) regarding cosmic time (in classical cosmology, quantum cosmology and thermodynamics) can be seen as fundamental properties or products of fundamental properties of the physical theories used to describe nature.

> b) the possibility to understand the origin of proper time asymmetries by considering a time asymmetric energy flux in spacetimes similar to ours. The question as to whether those asymmetries can be understood in a fundamental way, like the asymmetries with respect to cosmic time, will depend on an unsolved question regarding the connections between the alignments of proper and cosmic time asymmetries.

In this chapter, I shall only give a short overview of the different aspects and issues regarding these prominent time arrows. Because this will be discussed in greater detail in the following chapters, here I only sketch the traditional understanding of these arrows. Moreover, I will add some brief thoughts which motivate the claim that a new understanding of the different arrows is required in order to understand the *origin* of time asymmetries on the basis of fundamental physics (if the asymmetry is one of cosmic time) or a time asymmetric energy flux in spacetimes similar to ours (for proper time asymmetries).

The arrow of time in cosmology

Traditionally, the cosmological time arrow is defined in terms of the evolution and expansion of the three-dimensional universe and it is defined for cosmic time coordinates. The past, in this definition, is identified as the cosmic time direction in which the three-dimensional universe has a lower three-volume, and the cosmic future is defined as that in which the three-dimensional universe has a larger three-volume. In modern cosmological models, the universe expands. Thus, as long as this expansion holds, the time arrow will not change direction.

Regarding this traditional definition of the cosmological time arrow, many objections can be made to demonstrate that this cannot provide a fundamental understanding of cosmological time asymmetries. I deal with some of them in chapter III; at this point I only sketch one crucial objection, which shows that the cosmological time arrow, in this simple traditional notion, cannot describe a fundamental time asymmetry. This is because our universe can be described as a particular solution of the Einstein equation. Also, the Einstein equation shows that, even if the definability of cosmic time is assumed, there is no reason to rule out a closed spacetime in general. Even the discovery that our particular universe shows an accelerated expansion does not mean that closed spacetimes are ruled out as possible spacetimes (see also chapter III). Thus, even if our particular universe is an ever expanding universe, the cosmological time arrow defined by this expansion is not necessarily a fundamental property of physics because closed spacetimes are also possible according to the fundamental physical laws. Therefore, I think the cosmological time asymmetry in this traditional understanding cannot be seen as fundamental. I shall focus on some prominent and more sophisticated accounts of the cosmological time arrow in chapter III. Nevertheless, I will argue (chapter III) that they give rise to essentially the same problems as the traditional notion outlined here. Thus, in my suggestions in chapters III and V, I shall propose an understanding of cosmological time asymmetries, which can be seen as explications of a fundamental time asymmetry in the solution set of the fundamental dynamical equations of cosmology.

The arrow of radiation

It seems hard to determine the *traditional* understanding or characterisation of the arrow of radiation. The most traditional view seems to be the standard characterisation from physics [see, for example, Jackson (1999), Frisch (2000) or Rohrlich (2005)]. According to this characterisation, the arrow of radiation arises from the empirical fact that fully advanced radiation (of a specific type) is not observable in nature. Both the fully advanced and the

fully retarded solutions of Maxwell equations are time-mirrored pictures of each other. Thus, it seems that in nature, one time direction, identified with the fully retarded solution, is favoured.

I think that most attempts to understand this fact are problematic; I shall discuss this in much greater detail in chapter IV. Various accounts attempt to explain or describe the origin of the arrow of radiation, but, as I shall argue in chapter IV, they are unsuccessful in explaining its origin in physics. Thus, I argue, partly on the basis of philosophical suggestions from Frisch (2000), Castagnino and Lombardi (2009) and the physical analysis of Castagnino, Lara and Lombardi (2003), that there are crucial structures, at least in spacetimes similar to ours, that forbid the occurrence of fully advanced radiation of a specific but crucial kind. This will provide the retardation condition of Frisch (2000) on the basis of physics. Additionally, I shall draw attention to the question of the connection between this time arrow of proper times and the fundamental time asymmetry in spacetimes similar to ours (see chapter IV).

The arrow of time in thermodynamics

Perhaps the most prominent arrow of time is that in thermodynamics. This arrow is traditionally defined in terms of the behaviour of entropy in closed systems. According to the second law of thermodynamics, apart from fluctuations, the entropy of a closed system will increase with time up to a maximum value.

In discussions of this arrow of time in physics as well as in the philosophy of physics and time, crucial objections can be made to show that the thermodynamic time arrow, based on this definition, is not caused by fundamental physical reasons. I shall consider some of them in chapter V, but in this chapter I will present some more introductory thoughts on the understanding of this prominent time arrow.

Statistical mechanics predicts that, as the time coordinate decreases, most entropy values, apart from fluctuations, would also increase, as they do for an increasing time coordinate (as long as no initial conditions are used). Thus, according to the descriptions of statistical physics, thermodynamics does not include an entropic time asymmetry at a more fundamental level as the initial condition. Instead, some crucial initial conditions must be applied to provide the entropic arrow of time in thermodynamics. One crucial condition is that, in the systems 'past', the entropy value was low (which, then, defines 'past'); the second law of thermodynamics provides a time asymmetry only under this condition. Of course, the set of possible initial conditions also includes other initial conditions (and in fact more likely ones according to statistical physics) that cannot yield a time arrow in thermodynamics. In fact, the special choice of a particular initial condition seems to be motivated not by intrinsic structures of the theory in question (thermodynamics) but by

empirical data or anthropic considerations. Both approaches seem prima facie unable to provide an understanding of entropic time asymmetries based on fundamental physics but only based on boundary conditions or anthropic considerations. Thus, I think, the time arrow in thermodynamics, in this traditional understanding and based on crucial initial conditions, cannot constitute an understanding of the time direction based on the fundamental properties of physical theories.

In chapter V, I show that, according to specific entropy definitions in quantum thermodynamics and motivated by the physical analysis of Castagnino and Laciana (2002), the thermodynamic time arrow can be understood as a necessarily occurring by-product of a more fundamental cosmological time asymmetry if some crucial conditions are fulfilled (which seems to be the case in our particular universe). Thus, I argue that the arrow of time in quantum thermodynamics cannot be understood as fundamental itself but as a necessarily occurring by-product of a more fundamental cosmological time asymmetry. I shall argue that the behaviour of some specific entropy values in cosmic time will be intrinsically asymmetric in our particular (and similar) spacetime(s); hence, with decreasing cosmic time, the entropy value will also decrease, and with increasing cosmic time, the entropy value will increase (apart from fluctuations). Additionally, and independent of an epistemic or ontic interpretation of entropy itself, I will show that the *origin* of the time asymmetry in the behaviour of entropy is physically effective, whether or not entropy itself is understood as a purely epistemic content. Thus, I think the analysis of the thermodynamic time arrow in chapter V provides new views and advantages for the understanding of this prominent time arrow and the second law of thermodynamics itself.

Quantum mechanics

According to the standard formulation and interpretation of ordinary quantum mechanics, the time arrow in this field is mostly understood as a result of quantum measurements. Without considering any attempts to resolve the measurement problem, the most traditional understanding of this time arrow seems to be the following: The time evolution of a quantum system is described according to the time reversal invariant (according to the view that an complex conjugated equation is physical equivalent to the original equation) Schrödinger equation (or the Klein–Gordon equation and the Dirac equation regarding relativistic quantum mechanics, but they do not add anything to this discussion, so I focus on ordinary non-relativistic quantum mechanics here). Thus, it does not favour one particular time direction. However, the Schrödinger dynamics breaks down when a measurement or an analogous physical process is performed on a quantum system. In this case, we observe classical states. In the traditional formulation of quantum mechanics, this fact is attributed to the collapse of the wave function, which is time asymmetric, at least according to the

traditional formulation and interpretation. This means that time evolution, if begun in a classical state, will guide the measured classical state of the system slowly to one in which quantum effects become stronger. However, if a measurement is performed on a quantum state, the quantum state collapses into one classical state at the moment when the measurement is performed, and not slowly but on a very short (or infinitely short) time scale.

I shall argue in chapter VI, motivated partly by the physical analysis of Castagnino, Lara and Lombardi (2003), that, because it seems totally outside the scope of this investigation to provide a solution to the measurement problem, I divide the fields of applications of quantum physics into three levels. The first consists of laboratory experiments, in which we can deal only with measured physical entities. At this level, we can treat a quantum measurement as a black box in order to avoid the measurement problem and to define subsystems. The second level covers quantum measurements or analogous physical processes, and the third level is that of a pure von Neumann–Schrödinger quantum dynamics where no measurement is assumed. In the third level, the measurement problem is avoided by simply ignoring the possibility of quantum measurements or analogous physical processes at all. Using this distinction between 'levels' of quantum physical descriptions, I show in chapter VI that at the levels of laboratory experiments, as well as at the level of a pure von Neumann–Schrödinger dynamics, we find an arrow of time as an intrinsic property of the physical processes themselves. These arrows will be understandable as by-products of an energetic time asymmetry whereby the connection between those local asymmetries and cosmic time asymmetries will, again, be unsolved but revealed as the crucial point to understand the arrows in a fundamental sense. Moreover, the time arrow in the pure von Neumann–Schrödinger quantum dynamics provides strong arguments and motivations for a particular time asymmetric formulation of the rigged Hilbert space approach (see chapter VI).

Also, I will show that a crucial time asymmetry in one prominent process (often associated with quantum measurements), the decoherence process, is understandable in ways similar to the quantum mechanical time arrows from the other levels of description. Additionally, I mention that I do not go into much detail in the interesting discussion regarding the decoherence account of the measurement problem itself. Thus, I do not discuss all the arguments in favour of or against the view that the decoherence account can 'solve' the measurement problem. Here my motivation is only to show that, according to one important quantum process connected with quantum measurements, the description of the decoherence process is time asymmetric, at least in spacetimes similar to ours.

In summary, this subsection has sketched my main motivations and my main claims, which will be discussed in much greater detail in the following chapters. But before I present an

understanding of 'fundamentality' in the context of time directions in chapter II, I shall briefly mention another general assumption of this investigation.

I.1.2. The Role of Quantum Gravity

It could appear puzzling that this investigation claims to understand time directions on the basis of fundamental properties of the physical theories used to describe nature without taking into account the diverse formulations of quantum gravity. The fundamental time asymmetry, which I will define as a structural property of the solution set of crucial dynamic equations in cosmology, appears only in classical cosmology (chapter III) and in ordinary 'semi-classical' cosmology (with no attempt to quantize gravity; chapter V). Thus, the fundamental theory for all the models considered in this investigation is general relativity without a quantisation of gravity. Hence, my claims could be seen as a bit inconsistent. In fact, according to physics, it seems reasonable to assume that the most fundamental physical theory we can think of today is a theory of quantum gravity unified with the quantum dynamical description of the other three fundamental interactions. So, a fundamental time asymmetry should be based on the properties of such a fundamental unified theory of quantum gravity.

I have much sympathy for this view, but, according to physics, at least as far as I know, there is no well-established theory of quantum gravity that fulfils all the physically motivated requirements for such a theory. For example, as far as I know, there is no formulated theory of quantum gravity that has a well-defined classical limit or that explains the processes of supersymmetry breaking (SUSY) (if SUSY were imbedded in such a theory). Thus, because of the great range of speculations about the formulation of a unified theory of quantum gravity, for this investigation it seemed plausible to consider only the well-established physical theories of general relativity and quantum field theory (QFT). This restriction seems even more attractive considering that most formulations of quantum gravity have serious problems in defining even some sort of time coordinate on which a *fundamental* time asymmetry could be based. Thus, given the situation that we find in fundamental physics and in the scientific attempts to formulate quantum gravity, it seems well-motivated to focus on well-established physical theories. Moreover, I think an independent motivation for avoiding the field of quantum gravity in this investigation is the hope that, if a unified theory of quantum gravity is formulated, at least in their limit solutions, it should provide ordinary QFT and the traditional theory of general relativity as approximations. So, the fundamental time asymmetry in classical and 'semi-classical' cosmology (which I show more precisely in chapters III and V) could be a property of the approximation of the fundamental unified theory of quantum gravity. However, this is of course a question that can be investigated only after a well-established theory of unified quantum gravity is formulated.

However, even if physics should someday show that all the theories used in our current physical descriptions are inadequate and must be exchanged for others, I think the analysis in chapter II, which shows that a fundamental time asymmetry should not be based on the property of time-reversal variance of a fundamental physical law, could still be valid as long as the mathematical language of the theories change not too drastically. In fact, the considerations from chapter II are independent of specific physical theories and are basically of a philosophical kind. Thus, the suggested understanding of 'fundamentality' in the context of time asymmetries could still be an adequate notion of fundamentality even if the entire field of modern fundamental physics were *basically* inadequate (even in approximation).

I.2. The Main Claims

i) I will begin this investigation by motivating and defining a new notion of 'fundamentality' in the context of time asymmetries. This new notion is not based on the time-reversal variance of fundamental physical laws but instead on the structure of the solution set of a fundamental dynamical equation. This proposal is based and motivated on philosophical, physical and mathematical work in the context of asymmetries in general and time asymmetries in particular. I shall take the opportunity to mention some crucial research that is important for the motivation for my own suggestion: Boltzmann (1897), Castagnino, Gaioli and Gunzig (1996), Castagnino and Gunzig (1997), Castagnino and Laura (1997), Castagnino (1998), Castagnino and Gunzig (1999), Castagnino, Gueron and Ordonez (2002), Castagnino and Laciana (2002), Castagnino, Catren and Ferraro (2002) and very crucially Castagnino and Lombardi (2009) as well as Feynman (1964).

ii a) I will consider some accounts regarding the cosmological time arrow [see, for example, Price (1996), Price (2002) as well as Ćirković and Miloševic-Zdjelar (2004)], and I shall show that none of them can provide a fundamental understanding of even a cosmological time asymmetry, nor can they rule out the possibility of such an fundamental understanding (see chapter III).

ii b) Therefore, I will develop my proposal: that the solution set of the crucial dynamical equation in cosmology (analysed similar to Castagnino and Lombardi (2009)) provides a situation that is captured and described by my definition of fundamental time asymmetries from chapter II (provided some crucial requirements are fulfilled, see chapters II and III).

iii) Additionally, in chapter IV I will consider the actual discussion of the arrow of radiation. I will concentrate on crucial attempts to understand its origin [see, for example, Rohrlich

(2005), Frisch (2000), Price (1994), Price (1996), Price (2006), Zeh (1989), Zeh (1999)]. In chapter IV, my claim will be:

> a) that the traditional characterisation of the arrow of radiation is well-motivated for physical reasons, where the characterisation is given by the empirical fact that no fully advanced radiation of a specific kind seems to occur in our particular spacetime region. Moreover, I shall show

> b) [Motivated in part by the philosophical suggestions of Frisch (2000)] that the arrow of radiation, in the suggested characterisation, can be understood as a simple consequence from time asymmetric energy flows in spacetimes similar to ours (following parts of Castagnino and Lombardi (2009)). Thereby, the question of fundamentality of the radiation arrow will depend on the connection of the energy flow (in spacetimes similar to ours) and the cosmological time asymmetry investigated in chapter III.

iv a) Because the fundamental time asymmetry in cosmology (described in chapter III) was explicated only in the context of classical cosmology, one of the claims of chapter V is that the same type of explication is also given in 'semi-classical' quantum cosmology (omitting quantum gravity). I shall show that this is the case by considering the Einstein equations in some quantum cosmological models. Here the analysis is motivated in part by the cosmological investigation of Castagnino and Laciana (2002) (see chapters III and V).

iv b) In chapter V, I also concentrate on the understanding of the arrow of time in quantum thermodynamics. At the beginning of chapter V, I shall discuss an account of this particular time arrow that seems to be guided by artificial definitions and some crucial approximation methods [see Allahverdyan and Gurzadyan (2002)]. I investigate this account in order to show that my suggestions will not use similar kinds of arguments. Thus, in the second part of chapter V, I shall argue that, according to some prominent entropy definitions from Landau and Lifshitz (1970) as well as from Glansdorff and Prigogine (1971), time asymmetric behaviours of entropy values appear as a necessarily occurring by-product of the fundamental time asymmetry in cosmology.

v) In chapter VI, I will focus on the time arrows in non-relativistic quantum mechanics. Here the distinctions (mentioned above) among three different levels of quantum physics are crucial to deal with the measurement problem. The different levels are given by

> a) the level of laboratory descriptions, where a quantum measurement as well as analogue physical processes can be handled as black boxes,

> b) the level of the measurement itself, where only the prominent decoherence approach is considered, and

c) the level of the pure von Neumann–Schrödinger quantum mechanics, where no measurement or analogous physical process is assumed.

Adopting parts of the physical analysis of Castagnino, Lara and Lombardi (2003), I show that at all levels, a time arrow can be found and understood as a consequence of time asymmetric energy flows, but, in contrast to Castagnino, Lara and Lombardi (2003), not depending on the definition of subsystems in general. Thereby, as well as by the radiation arrow, the fundamentality of those time asymmetries depend on the connection to the fundamental time asymmetry of cosmology (which is defined in chapters III and V). Moreover, I argue (see also Bishop (2004)) that, regarding the pure quantum mechanical level of description, a particular time asymmetric formulation of ordinary quantum mechanics, the rigged Hilbert space formulation in a particular form, is strongly supported.

To summarize the aims of this investigation,

First: I will show that an understanding of time asymmetries based on the properties of fundamental physics is possible without time-reversal variant laws in the fundamental theories of physics.

Second: I will show that the fundamentality of time asymmetries can be based on the structure of the solution set of time-reversal invariant physical equations.

Third: I will show that this understanding of fundamentality is applicable to physics; in particular, to classical cosmology and 'semi-classical' cosmology, which will provide time arrows in QFT and in quantum thermodynamics. Other local time asymmetries and arrows can be understood as time asymmetric consequences, for example in classical electrodynamics and ordinary quantum mechanics, if a particular kind of connection between the cosmological time asymmetry and other local processes is assumed. This point will be clarified in the chapters IV and VI.

Figure 1.1 shows a schematic diagram of the coarse structure of the investigation as sketched in this chapter. At the beginning of each chapter, I shall return to this diagram to clarify which part of the investigation that particular chapter covers. The aim of this investigation, which is that the understanding of a difference between two time directions, labelled as 'past' and 'future', can be based on fundamental considerations in physics, is illustrated in the diagram.

Cosmic time asymmetries	Proper time asymmetries

Motivation and Definition of 'Fundamentality' in the Context of Time Asymmetries (Chapter II)

Investigating the Solution Set of Crucial Dynamic Equations in Cosmology (Chapter III and V)

Fundamental Time Asymmetry in the Solution Set of Crucial Dynamic Equations in Cosmology (Chapter III)

Time Asymmetric Behaviour of the Expectation Value of the Particle Number Operator in Hyperbolic Curved Spacetimes (Chapter V)

The Entropic Time Arrow Understood as a Consequence of the Fundamental Time Asymmetry in Cosmology (Chapter V)

Time Asymmetric Behaviour of the Relativistic Energy Flux in Spacetimes similar to ours (Chapter IV)

The Arrow of Radiation Understood as a Consequence of Time Asymmetric Behaviour of the Relativistic Energy Flux (Chapter IV)

The Traditional Arrow of Time in Quantum Mechanics Understood as a Consequence of the Time Asymmetric Behaviour of the Relativistic Energy Flux (Chapter VI)

Motivating the Rigged Hilbert Space approach to non-Relativistic Quantum Mechanics (Chapter VI)

Time Asymmetric Decoherence Processes; Understood as a Consequence of the Time Asymmetric Behaviour of the Relativistic Energy Flux (Chapter VI)

15

Moreover, I think the proposed understanding of some time arrows investigated here could provide new arguments in the discussions of

a) an epistemic or ontic understanding of entropy and the second law of thermodynamics and

b) causality and the asymmetry between causes and effects in fundamental physics as well as in non-fundamental physics and specific sciences.

I shall return to these points briefly in chapter VII, where the conclusions of the entire investigation are summarized.

Chapter II

'Fundamental' Time Asymmetries

This very short but crucial chapter argues for a new understanding of fundamentality with regard to time asymmetries. My assertion is that the fundamentality of time asymmetries, even when it is based on fundamental physics, should not be based on the time-reversal invariance of physical laws. I argue that it is possible to construct a notion of fundamentality regarding time asymmetries that is applicable to physics, even if all the fundamental laws were time-reversal invariant. I shall argue that the fundamentality of a time asymmetries can be based on the set of solutions of a time reversal invariant law (TRIL). If this law is part of fundamental physics, the structure of the associated solution set is also fundamental. I show that this structure can yield a fundamental difference between time directions. In addition, it appears that this is possible without any speculation about a non time-reversal invariant formulation of quantum physics.

On the diagram below, we see that this chapter aims only to clarify and motivate the new understanding of 'fundamentality' in the context of time asymmetries. The parts of the structure that is considered in the following chapter are marked with a light violet colour.

'Fundamental' Time Asymmetries

Cosmic time asymmetries

Proper time asymmetries

Motivation and Definition of 'Fundamentality' in the Context of Time Asymmetries (Chapter II)

Investigating the Solution Set of Crucial Dynamic Equations in Cosmology (Chapter III and V)

Fundamental Time Asymmetry in the Solution Set of Crucial Dynamic Equations in Cosmology (Chapter III)

Time Asymmetric Behavior of the Expectation Value of the Particle Number Operator in Hyperbolic Curved Spacetimes (Chapter V)

Time Asymmetric Behavior of the Relativistic Energy Flux in Spacetimes similar to ours (Chapter IV)

The Arrow of Radiation; Understood as a Consequence of Time Asymmetric Behavior of the Relativistic Energy Flux (Chapter IV)

The Traditional Arrow of Time in Quantum Mechanics; Understood as a Consequence of the Time Asymmetric Behavior of the Relativistic Energy Flux (Chapter VI)

The Entropic Time Arrow; Understood as a Consequence of the Fundamental Time Asymmetry in Cosmology (Chapter V)

Motivating the Rigged Hilbert Space approach to non-Relativistic Quantum Mechanics (Chapter VI)

18

Time Asymmetric Decoherence Processes; Understood as a Consequence of the Time Asymmetric Behavior of the Relativistic Energy Flux (Chapter VI)

As the diagram shows, this chapter is central to the entire analysis. In fact, the notion of fundamentality developed in this chapter is the basic motivation for the different considerations in chapters III–VI. This situation is reflected in the diagram by the fact that the box representing the new notion of fundamentality has only outgoing arrows but no incoming ones. Because the diagram above does not provide useful information about how this chapter is organized, I hope the following graphic clarifies the chapter's overall structure.

'Fundamental' Time Asymmetries

Two possibilities for defining time asymmetries
a) Time-reversal variants of physical laws
b) Time asymmetric solutions of physical laws
Hence, four combinations are possible

Time-reversal invariant laws
&
Only time symmetric solutions

Time-reversal invariant laws
&
Also time asymmetric solutions

Time-reversal variant laws
&
Only time symmetric solutions

Time-reversal variant laws
&
Also time asymmetric solutions

No time asymmetric properties

Mixed properties. Only useful if the solution set favours the time asymmetric solutions.

Mixed properties; But, no predictable or describable physical process would be time asymmetric. Hence: not useful.

All properties seem time asymmetric. But, the combination seems only adequate to describe nonstandard formulations or interpretations of physics. (The CPT-Theorem dose not shed new light on the situation; see chapter III)

A 'fundamental' favouring of time asymmetric solutions seem possible if the set of time symmetric solutions is a subset of measure zero (according to an ordinary measure).

But, no favouring between two time mirrored and asymmetric solutions seem possible. Hence, other properties of the solutions set are necessary to base a notion of 'fundamentality' on the structure of the solution set. Sufficient and necessary seems that two time mirrored (and intrinsically asymmetric solutions) are physically identically, even if formal distinguishable. This produces a fundamental time asymmetry in the whole solution set.

Even if not all the details of the diagram are comprehensible now, during this chapter its structure should become clear. On this point, the only purpose of the diagram is to present the claims and describe how the chapter is organized in order to demonstrate these claims. At the end of this short chapter, I will show this diagram again in order to demonstrate that all the argumentation 'arrows' are considered in this chapter and that the suggested new notion of 'fundamentality' for time asymmetries is well motivated. I will continue with this procedure in all chapters. So that also the structure of the different following chapters can easily be ascertained.

II.1. Preliminary Considerations

In 1927, Eddington coined the phrase 'arrow of time' to describe the asymmetry of time directions according to physical phenomena. At least since then, the validity and fundamentality of different time arrows and asymmetries have been discussed. However, in the standard formulation and interpretation, the fundamental laws seem time-reversal invariant.[2] Thus, the fundamental laws in physical models do not seem to provide a favoured time direction or any kind of fundamental time asymmetry.

Many attempts have been made to formulate a time asymmetry according to proper times or cosmic time in modern physics as well as in the philosophy of science. For example, Kanekar, Sahni and Shtanov (2001) tried to ground time directions in the cosmology of oscillating spacetimes in string cosmological models, and Allahverdyan and Gurzadyan (2002) grounded the time arrow on the chaotic properties of the background radiation. Rohrlich (2005) tried to define time directions in classical electrodynamics on the basis a causality principle; in contrast to Rohrlich, Frisch (2000) stipulated a new time asymmetric ad hoc law in classical electrodynamics to deduce a time arrow. Price (1996) and (2006) tried to argue that, according to fundamental physics, there are no favoured time directions; empirical time directions such as the arrow of radiation or the thermodynamic time asymmetry would be provided by boundary conditions and thus should not be understood as fundamental time directions. In contrast to Price, Bohm, Gadella and Wickramasekara (1999) [see also Bishop (2004)] developed a time asymmetric formulation of non-relativistic quantum physics based on the rigged Hilbert space approach in order to ground a time direction in fundamental properties of quantum physics.

---

[2] Or they are CPT-invariant, which does not change the issue (see chapter III).

So the modern discussion of the validity and fundamentality of time directions has achieved an enormous variety. In addition to the suggestions described above, many authors have, at least implicitly, touched the question of the fundamentality of time directions in physical models and especially in cosmology, even at the end of the last century. See, for example: Davies (1994), Earman (1974), Schulman (1999), Penrose (1979), Sachs (1987), Price (1996), Grünbaum (1973), Matthews (1979), Reichenbach (1956) or Hawking and Ellis (1973).

However, the main problem in understanding time directions as a fundamental property of physics arises from the fact that, according to the standard formulation and interpretation of the fundamental physical theories, the fundamental laws are time-reversal invariant.[2] Thus, it seems that they cannot provide a favoured time direction.

Thus, this chapter should show that it is unnecessary, and moreover insufficient, to base the fundamentality of time asymmetries on the time-reversal variance of dynamical laws. Instead, a definition of fundamentality is suggested that is based on the structure of the solution set of a physical dynamical equation. The solutions are associated with physical processes or models of nature and, as I will show, the entire set of solutions can include time asymmetry as a structural property, even if the law is time-reversal invariant. Moreover, if the considered equation is a fundamental equation, the structure in its solution set is also fundamentally given. Thus, I argue that this structure in the solution space of fundamental equations can provide a fundamental time asymmetry. The idea is motivated by the cosmological investigation of Castagnino, Lara and Lombardi (2003) and Castagninio and Lombardi (2009), in which they show that a cosmological time arrow based on the solution space of the dynamical equations in cosmology may occur even if the fundamental Einstein equations are time reversal invariant. I argue that this idea should be generalized in order to provide a new definition of 'fundamental' time asymmetries.

II.2. A Proper Notion of 'Fundamental' Time Asymmetries

The first step in building my argument is well known and has been analysed by many writers on both philosophy and physics. It arises from the distinction between the property of time-reversal (in)variance, as a property of dynamical equations, and the general property of time (a)symmetry, applicable to solutions of fundamental dynamical equations. In this investigation, 'time-reversal invariance' is used only to describe law-like dynamical equations, and 'time (a)symmetry' is understood as a property of the solutions of those equations. More precise, a dynamical equation is understood as time-reversal invariant iff the transformation $t \rightarrow -t$ dose not change the form of the equation. Time symmetry of a solution $f(t)$ instead is given if there is at least one symmetry point t_0 such that $f(t_0+t)=f(t_0-t)$

for every scalar time coordinate t.[3] Moreover, we can associate dynamical equations with physical laws and solutions to those equations with physical models that satisfy such laws. Therefore, we can combine these properties in four ways:

a) Time-reversal invariance and only time symmetric solutions,
b) Time-reversal invariance and some time asymmetric solutions,
c) No time-reversal invariance and only time symmetric solutions and
d) No time-reversal invariance and some time asymmetric solutions.

It is easy to find physical examples of combination a). However, this combination is not useful if we are interested in time asymmetries.

Combination b) looks interesting because it shows that TRILs could have time asymmetric solutions. Nevertheless, such traditional asymmetries are usually not understood as fundamental time asymmetries because they occur only in some special models of the TRIL, and the occurrence of such asymmetries is provided by boundary conditions. Thus, they are not 'fundamental'.

The applicability of combination c) or d) in fundamental physics is at least problematic. It seems that time-reversal variant[2] laws cannot be found within the laws of fundamental physics in the standard interpretation. Hence those combinations seem applicable only in some special formulation of e.g. quantum laws, for example in some formulations of the rigged Hilbert space approach (see, for example, Bohm, Gadella and Wickramasekara (1999); Bishop (2004), Castagnino, Gadella and Lombardi (2005) or Castagnino, Gadella and Lombardi (2006)). However, it will be shown below that combination c) or d) need not be used to understand time asymmetries in a fundamental manner.

I suggest that combination b) indicates another plausible way for defining fundamentality and explaining the occurrence of a fundamental time asymmetry. As we will see, this suggestion does not require special interpretations or formulations of quantum physics. However, a definition of fundamentality in terms of combination b) as it stands above does not seem plausible. Instead, we can add some conditions to combination b).

Definition I:

Suppose L is a fundamental linear TRIL, and S(L) is the solution space with dim(S(L)) = n. I will call a time asymmetry 'fundamental' if and only if:

---

[3] The fact that time reversal (in)variants is understandable as a special type of time (a)symmetry is not from crucial interest here and it will changes the considerations in no way.

i) There is no more than a countable collection $S_i(L)$ of subspaces of dimensions $m_i < n$ and no more than an uncountable collection $S_i'(L)$ of subspaces of dimension $m_i' + 1 < n$, such that if $f(t) \in S(L)$ is time symmetric, then $f(t) \in S_i(L)$ or $f(t) \in S_i'(L)$ for some i, and if $f(t) \in S(L)$ is time asymmetric, then $f(t) \notin S_i(L)$ and $f(t) \notin S_i'(L)$ for all i .

ii) For time asymmetric solutions $f(t) \in S(L)$, the solution $f(-t) \in S(L)$ refers to the same physical world as $f(t)$ does.

Condition ii) is important because if $f(t) \in S(L)$ and $f(-t) \in S(L)$ describe physically different models, we would have to explain why only one direction (+ or − sign) occurs in nature. Then, this argument would construct the time asymmetry. However, if condition ii) holds, no such additional argument is needed. Such a situation is possible, for example, if the sign of t refers to a non-physical structure. For example, according to general relativity, time directions, which are physical time directions, are independent of the sign of a hypothetical Newtonian background time parameter, which can be the sign of t in definition I. I show this in more detail in chapters III and V in the context of classical and quantum cosmology. I shall show that the sign of the crucial parameter distinguishes the solutions only in a mathematical and formal, and not in a physical, way. But, on this point I will only reflect on the definition, independently from particular physical theories.

Moreover, if condition i) is fulfilled, the solutions of a particular fundamental law are almost all asymmetric. If condition i) is fulfilled, then all time symmetric solutions are contained in one or many [see condition i)] subspaces of measure zero.[4] This situation is illustrated in Figure II.1.
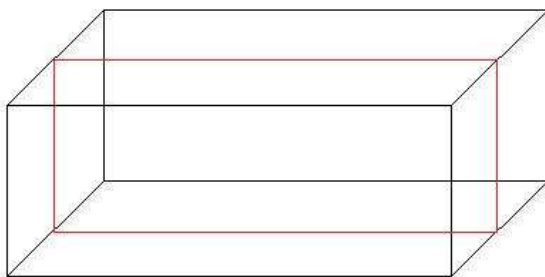

Fig. II.1 The cuboid illustrates a part of a three-dimensional solution space that includes both time symmetric and time asymmetric solutions. All time symmetric solutions are located on

---

[4] If an ordinary measure such as the Lebesgue measure is used

the surface indicated by the red rectangle. This fragmentation is possible only because the solution space of time symmetric solutions has only two dimensions (where the entire solution space has three), and this fact results (as a possible explication) from the first part of definition I. Therefore, almost all solutions of the dynamical law include an intrinsic time asymmetry.

Thus, a law with a set of solutions that satisfies conditions i) and ii) of definition I will always (except for some subspaces with dimension $m_i < n$) rise a time asymmetry in the described or predicted processes. Therefore, it seems clear that a time asymmetry in physics that satisfies definition I can (and perhaps should) be called fundamental (because if the law is fundamental, the structure of the solution space is given for fundamental reasons).
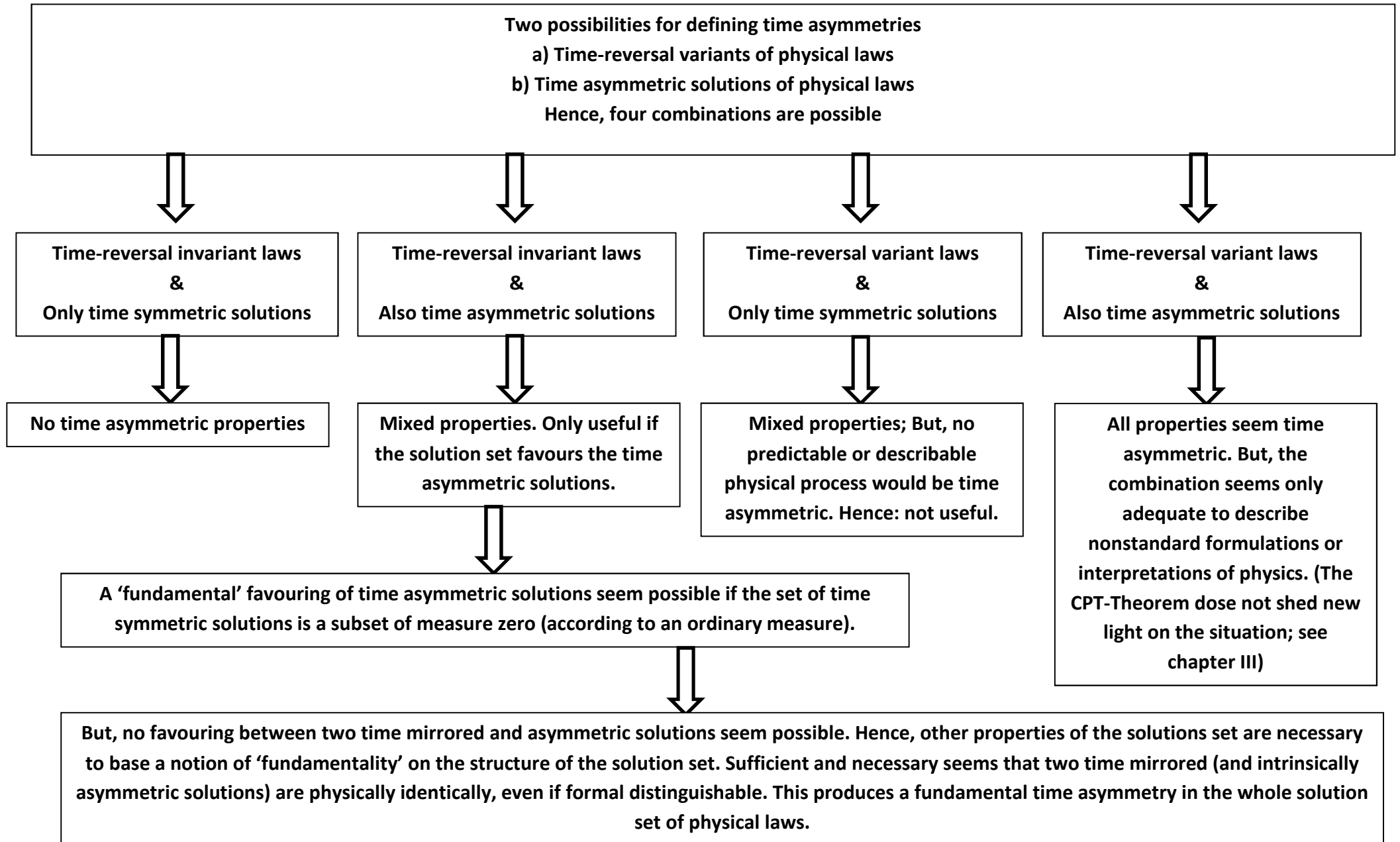
To examine this idea in a bit more detail, note that the structure of a solution space of a fundamental law is clearly a fundamental property. In fact, it is based only on the fundamental law. Thus, it seems quite reasonable to say that time asymmetries are fundamentally imbedded in a theory if the solution space of the fundamental equations fulfils condition i), which means that the law provides that almost all of its solutions are time asymmetric. However, both conditions i) and ii) must be fulfilled in order to define a fundamental time asymmetry because the set of time asymmetric solutions can be time *symmetric* as a whole, if only i) is fulfilled. Thus, the entire set will *not* show a time asymmetry if only condition i) is fulfilled. To define a fundamental time asymmetry, it is necessary to refer to condition ii) as well. If condition ii) is also fulfilled, it follows that a time asymmetry is fundamentally described by the structure of the solution space of the fundamental law under consideration. The reason is that if i) and ii) are fulfilled, the intrinsic time asymmetries in almost all solutions [see condition i)] cannot be used to construct a time *symmetric* solution set [see condition ii)].

II.3. Conclusion

The aim of this short but crucial chapter was to show that fundamental time asymmetries, as properties of fundamental physics, should not be understood as arising from time-reversal variant laws. Instead, definition I show another way of defining fundamentality in this context. For a time asymmetry based on definition I, the time reversal (in)variance of the fundamental law is irrelevant. Nevertheless, the time asymmetry, which is given by the

solution space of such a law, can be understood as a fundamental time asymmetry, because if the law is fundamental, then the structure of the solution space should also be fundamental.

To show that all the relevant arguments that motivated my suggestion were made in this chapter, I again present the diagram illustrating the structure of this chapter. This should demonstrate in a simplified and compressed way that, at least as a working hypothesis, the new concept of 'fundamentality' in the context of time asymmetries is well motivated.

'Fundamental' Time Asymmetries

Two possibilities for defining time asymmetries
a) Time-reversal variants of physical laws
b) Time asymmetric solutions of physical laws
Hence, four combinations are possible

Time-reversal invariant laws
&
Only time symmetric solutions

Time-reversal invariant laws
&
Also time asymmetric solutions

Time-reversal variant laws
&
Only time symmetric solutions

Time-reversal variant laws
&
Also time asymmetric solutions

No time asymmetric properties

Mixed properties. Only useful if the solution set favours the time asymmetric solutions.

Mixed properties; But, no predictable or describable physical process would be time asymmetric. Hence: not useful.

All properties seem time asymmetric. But, the combination seems only adequate to describe nonstandard formulations or interpretations of physics. (The CPT-Theorem dose not shed new light on the situation; see chapter III)

A 'fundamental' favouring of time asymmetric solutions seem possible if the set of time symmetric solutions is a subset of measure zero (according to an ordinary measure).

But, no favouring between two time mirrored and asymmetric solutions seem possible. Hence, other properties of the solutions set are necessary to base a notion of 'fundamentality' on the structure of the solution set. Sufficient and necessary seems that two time mirrored (and intrinsically asymmetric solutions) are physically identically, even if formal distinguishable. This produces a fundamental time asymmetry in the whole solution set of physical laws.

The advantage of this concept of fundamentality is that some crucial problems for fundamental time directions can be avoided. The well-known problems mentioned above that arise in discussions of time directions are based on the crucial point that it seems highly problematic to search for fundamental time directions if the fundamental laws are time-reversal invariant. In this case, the time direction is traditionally based on boundary conditions that are not fundamentally given. With the notion of fundamentality defined here, this is no longer problematic because even a TRIL can lead to a solution space that includes almost exclusively time asymmetric solutions. Moreover, according to condition ii) in definition I, those asymmetries must lead to a fundamental time asymmetry if *f(t)* and *f(-t)* describe the same physical model. In this case the sign of *t* refers only to a mathematical construct such as an absolute background time (e.g. a Newtonian time parameter).

I shall demonstrate in the next chapter that the proposed conception of fundamental time asymmetries is applicable to physical theories. I will demonstrate this in the context of cosmology and I will show precisely that, under some motivated conditions, the crucial dynamical equations in cosmology lead to a solution set which fulfils definition I.

Chapter III

Time's Arrow in Cosmology

The aim of this chapter is to show that the new conception of fundamentality (chapter II) can be applied successfully in physical theories.

In the introduction of this chapter (section III.1), I shall refer to various views from the literature. I begin by reviewing the work of Price (1996), which suggested that nature has no arrow of time as a fundamental property. Thus, in subsection III.1.1 of this chapter, I provide reasons for disagreeing with the arguments in Price's work.

Next, in subsection III.1.2, I refer to suggestions for defining the arrow of time in cosmology via the behaviour of entropy. In addition, I argue that these various types of approaches cannot explain the occurrence of a *fundamental* arrow of time in cosmology. Moreover, in subsection III.1.3, I broaden my scope to approaches in the context of inflation theories as well.

To verify the applicability of this new conception of fundamentality, I present an example to show that it can be applied in physical theories in sections III.2.1–III.2.4. The example is classical cosmology, in which the theory of general relativity (as well as empirical equivalent spacetime theories) is treated as the fundamental physical theory. In sections III.2.1–III.2.4, I define a fundamental time asymmetry in cosmology (under some well-motivated conditions) by combining the outcomes from by Castagnino, Lara and Lombardi (2003a) and Castagnino and Lombardi (2009) with the proposed understanding of fundamentality (chapter II); I address a possible objection in section III.2.5. Finally, in section III.3, I conclude by discussing the effects of adding the new understanding of fundamentality to classical cosmology.
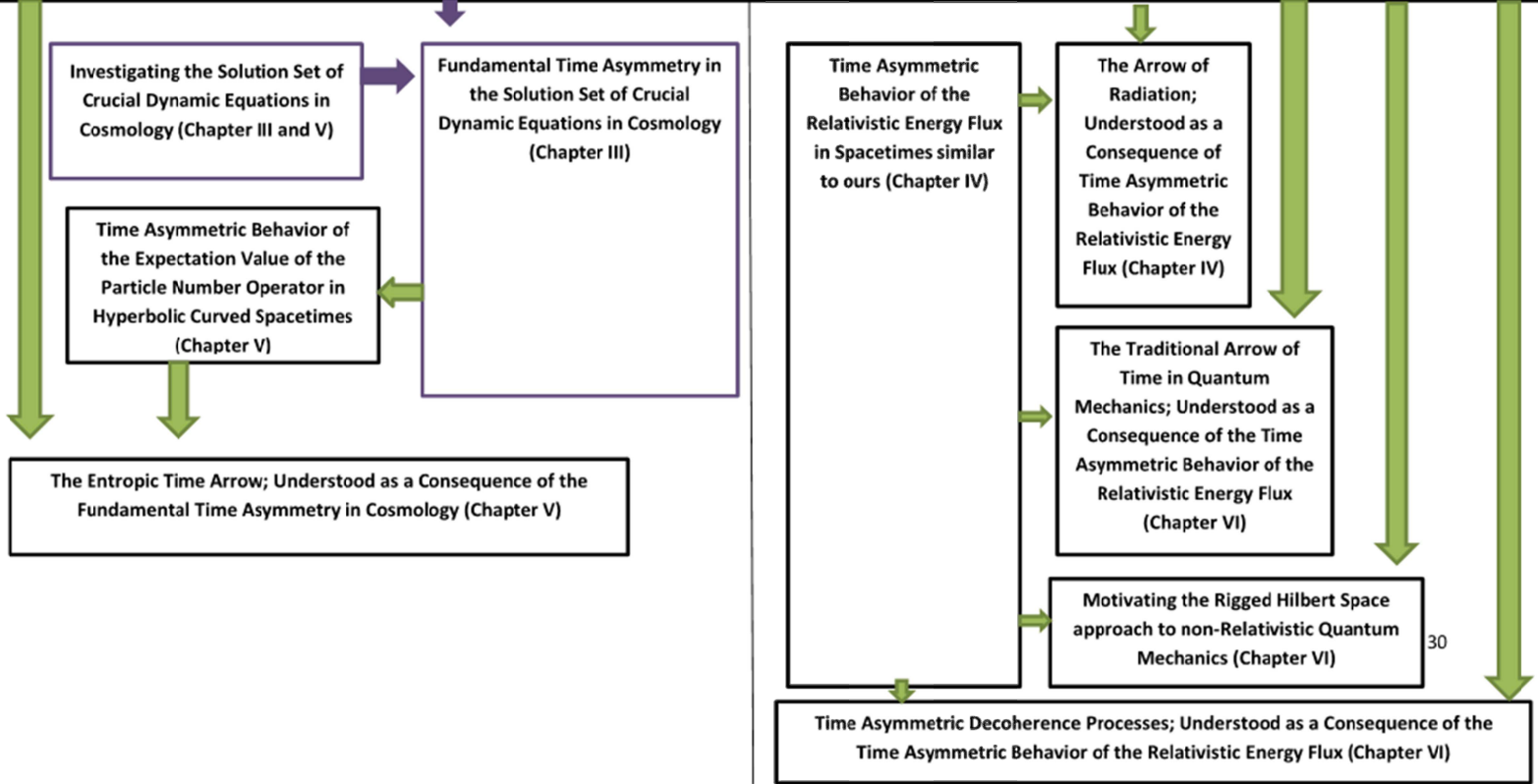
Regarding this chapter's relationship to the entire investigation, the following diagram, as usual, illustrates which parts of the analysis are considered in this chapter.
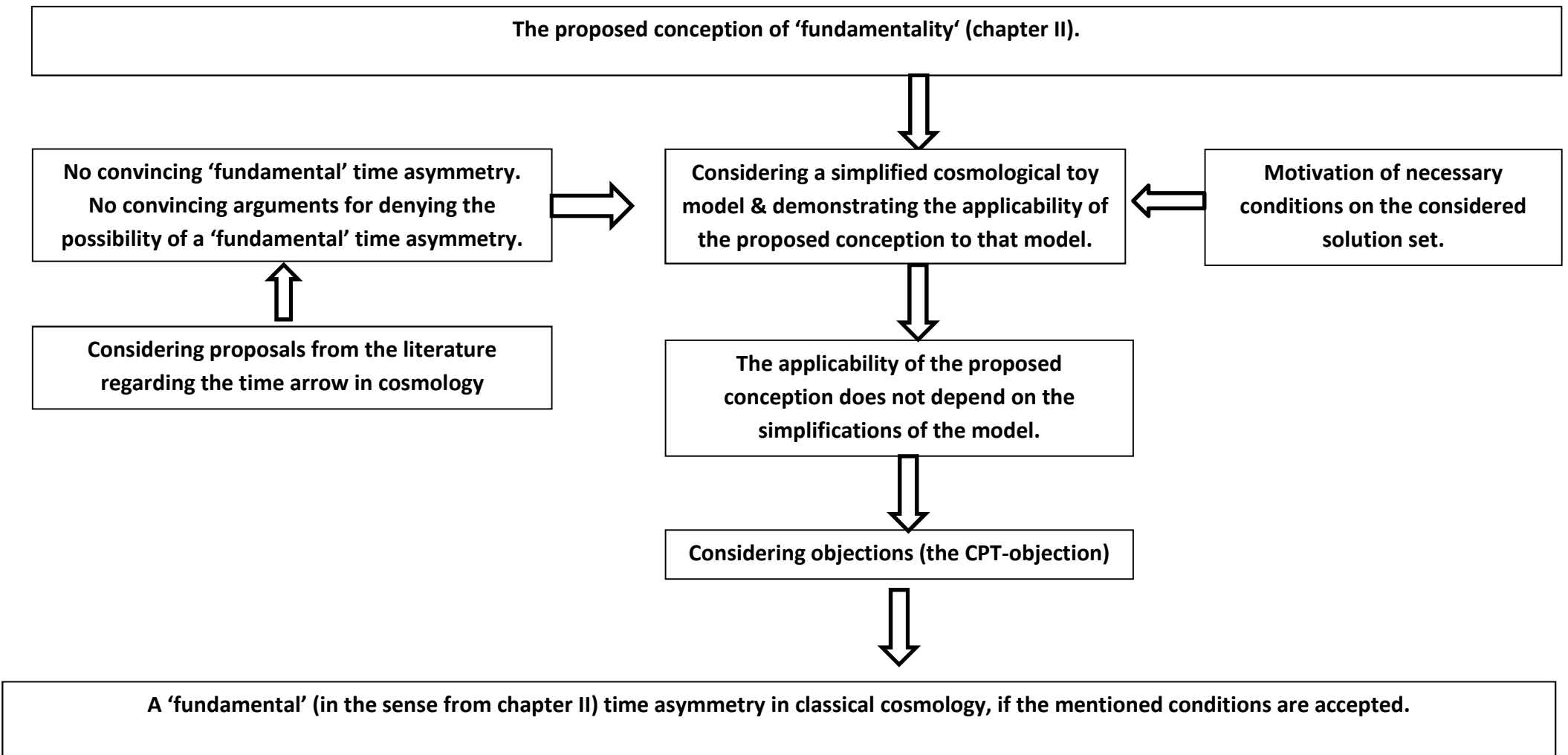
Time's Arrow in Cosmology

Cosmic time asymmetries   **Proper time asymmetries**

Motivation and Definition of ‚Fundamentality' in the Context of Time Asymmetries (Chapter II)

Investigating the Solution Set of Crucial Dynamic Equations in Cosmology (Chapter III and V)

Fundamental Time Asymmetry in the Solution Set of Crucial Dynamic Equations in Cosmology (Chapter III)

Time Asymmetric Behavior of the Expectation Value of the Particle Number Operator in Hyperbolic Curved Spacetimes (Chapter V)

The Entropic Time Arrow; Understood as a Consequence of the Fundamental Time Asymmetry in Cosmology (Chapter V)

Time Asymmetric Behavior of the Relativistic Energy Flux in Spacetimes similar to ours (Chapter IV)

The Arrow of Radiation; Understood as a Consequence of Time Asymmetric Behavior of the Relativistic Energy Flux (Chapter IV)

The Traditional Arrow of Time in Quantum Mechanics; Understood as a Consequence of the Time Asymmetric Behavior of the Relativistic Energy Flux (Chapter VI)

Motivating the Rigged Hilbert Space approach to non-Relativistic Quantum Mechanics (Chapter VI)

Time Asymmetric Decoherence Processes; Understood as a Consequence of the Time Asymmetric Behavior of the Relativistic Energy Flux (Chapter VI)

30

Further, I must mention that this chapter as well as chapter V are the only chapters (see diagram) in which the proposed conception of fundamentality leads directly to a fundamental time asymmetry in a field of physics. Thus, in addition to the introduction and the diagram above, I will also present, as in chapter II, a short and simplified version of the organisation of this chapter. As in chapter II, some of the connections shown in the following figure may not be clear now, but they provide a useful overview of the chapter's organisation. At the end of this chapter, I hope the reader will be able to trace the crucial arguments of this chapter more easily with the use of this diagram.

The proposed conception of 'fundamentality' (chapter II).

No convincing 'fundamental' time asymmetry. No convincing arguments for denying the possibility of a 'fundamental' time asymmetry.

Considering a simplified cosmological toy model & demonstrating the applicability of the proposed conception to that model.

Motivation of necessary conditions on the considered solution set.

Considering proposals from the literature regarding the time arrow in cosmology

The applicability of the proposed conception does not depend on the simplifications of the model.

Considering objections (the CPT-objection)

A 'fundamental' (in the sense from chapter II) time asymmetry in classical cosmology, if the mentioned conditions are accepted.

III.1. Different Views

III.1.1. The Time Symmetric View

In this subsection, I briefly examine the suggestions of Price (1996) for favouring a time symmetric view. Because the arguments do not appear plausible or compelling to me, I think the search for a 'fundamental' arrow of time is not doomed to fail, at least not for the reasons given by Price.

Price begins his cosmological analysis with the observation that the early state of the universe is special in one interesting way: the universe near the Big Bang is smooth [Price (1996)].[5] A matter distribution consisting of a number of black holes would be much more likely than a smooth distribution if classical gravity is the dominant force, as assumed in classical cosmology. Thus, according to classical thermodynamics, Price argues that this fact shows that the 'early' universe has very low entropy.

Now, Price (1996, p. 78) goes on to argue that entropic behaviours (as well as all other properties) cannot provide a fundamental time arrow in classical cosmology. His arguments are based on the fact that all statistical considerations, which in fact yield the second law of thermodynamics, are also valid in the reverse time order. Thus, the time asymmetry of entropic behaviour is based on the initial low-entropy conditions of the Big Bang [see also

---

[5] However, on this point, a crucial question arises that is important for understanding Price's suggestions: It is not clear what the word 'early' [which Price uses many times; see, for example, Price (1996), pp. 79 and 80) means in this context. Price does not seem to refer to a proper time period with respect to the constituents of the universe when he uses the phrase 'early state of the universe'. Instead, it seems that he refers to a) cosmic time or b) the geometrical fact that these states of the universe are 'near' [according to a metric such as the Friedmann–Robertson–Walker (FRW) metric] the Big Bang. To make this interpretation of Price plausible, we need a definition of the distance between a spacetime point p (or the singularity) and a three-dimensional spacetime region R (a state of the three-dimensional universe). However, this distance can be defined in many possible ways, for example, taking the nearest (according to a metric) spacetime point p' of a region R and calculating the distance between p and p' according to the metric. Of course, we can also construct many definitions of distance for this task. Such definitions would be more or less plausible, but they all define what it means to say that 'a spacetime region is *near* a spacetime point or the Big Bang'. It seems fair to assume that Price refers to the most acceptable meaning of 'early'. That is, it seems most unproblematic to define his use of the term 'early' as 'near the Big Bang according to a metric and a plausible definition of distance between a spacetime point p and a spacetime region R'.

Albert (2000)]. Therefore, Price (1996, pp. 81–99) argues that according to entropic behaviour (or other statistical reasoning), a closed universe with low-entropy boundary conditions in the 'future' is time symmetric, and the low-entropy boundary in the 'future' is as likely as the one in the 'past', which seems to be given in our actual world. Additionally [see Price (1996), pp. 95–96], Price argues that we are not concerned with whether our particular spacetime is closed or open, but with whether a closed spacetime with symmetric boundary conditions is possible given the laws of classical cosmology.

> *'This point [the possibility of open spacetime geometries] is an interesting one, but it should not be overrated. For one thing, if we are interested in whether the Gold universe [a type of time symmetric closed spacetime] is a coherent possibility, the issue as to whether the actual universe recollapses is rather peripheral. […] Of course, if we could show that a recollapsing universe is impossible, given the laws of physics as we know them, the situation would be rather different.'* [Price (1996), p. 95]

Thus, Price seems to argue that if the existence of such a universe is possible given the laws of classical cosmology, there is no reason to assume that time behaves asymmetrically in a fundamental sense. Even if our particular spacetime has boundaries that yield an entropic time asymmetry, this would not affect the question of a fundamental time direction because the direction, in such a spacetime, would be given by (perhaps accidental) boundary conditions.

My critique can be outlined very briefly. Price argues mainly that in a possible closed spacetime, there would be no physical parameter that distinguishes a Big Bang from a Big Crunch (and thus could be used to define a time's arrow). This assumes that the scale factor (Price calls it the radius of the universe), which behaves symmetrically in a closed spacetime, is the only fundamental property to distinguish between 'initial' and 'final' singularities. Thus, time would be symmetric in the cosmological description. However, his crucial assumption seems to be that the scale factor (or the radius) is the only basic property that could be used to distinguish between the two singularities in a closed spacetime, given the standard theories of classical cosmology [see Price (1996), pp. 86–111]. Thus, I argue that this assumption is implausible and that other physical properties (not statistical considerations) than the scale factor must be considered for defining a fundamental time direction.

Thus, I conclude that if it is possible to define physical properties that are as basic as the scale factor but independent of them (e.g. matter fields, see section 2), Price's arguments are no longer plausible; thus, the time directedness in classical cosmology could be fundamental.

Nevertheless, this brief investigation was necessary to show that the search for a fundamental time arrow is not doomed to fail on the grounds of Price's arguments, if there are fundamental properties of the universe other than the scale factor. This, in fact, seems to be the case according to modern cosmological models. I consider this point more precisely when I argue for a new understanding of the origin of the cosmological time arrow.

However, before I show that the cosmological time asymmetry is understandable in the fundamental sense defined in chapter II, I briefly examine some approaches to the arrow of time in classical cosmology in order to show that the suggestions of which I am aware cannot define a convincing *fundamental* time asymmetry in classical cosmology.

III.1.2. Entropy-Based Approaches

In this subsection, I show that the various entropy-based approaches of which I am aware cannot define a fundamental arrow of time in classical cosmology; that is, the time direction in these approaches is not derived from the basic properties of the cosmological models.

In this section, I focus on approaches that define the arrow of time in cosmology via the time asymmetric behaviour of entropy. More precisely, the future direction in such approaches is given by the time direction in which entropy increases according to the second law of thermodynamics. Of course, statistical mechanics shows that most types of entropy would also (theoretically) increase in the time mirrored direction. Thus, it is necessary to set some boundary conditions (or provide other explanations) for the past, specifically, that the universe has low entropy in the past [see, for example, Albert (2000)].

However, approaches that focus on the temporal behaviour of entropy can be subsumed under two different types, as analysed by Price (2002):

> 1. Causal–general approaches: This class of approaches seeks to explain the low-entropy state near the Big Bang by fundamental physical laws.

However, no current physical theory can explain the specialness of the early universe *only* by invoking dynamic laws.[6] It seems that with our current knowledge of physics, we cannot formulate approaches that deduce an arrow of time *only* from dynamic laws [see also Wald (2006)]. Thus, causal–general approaches seem unsuccessful in deducing a fundamental

---

[6] Inflation and models of the multiverse are considered later. Also, there are unable to provide a time asymmetry *only* by invoking dynamic laws.

time direction in classical cosmology as long as boundary conditions are not understood as fundamental properties. I also believe that many philosophers and physicists are aware of and agree with this view; hence, this point is not discussed in more detail.

> 2. Acausal–particular approaches: This class of approaches describes the specialness (or the low entropy) of states near the Big Bang by the existence of boundary conditions.

Because boundary conditions should not be understood as fundamental properties of physics, neither of these two approaches explains the occurrence of a fundamental (in some reasonable sense) arrow of time in classical cosmology.

However, note that some authors have argued for another possibility [for example, Ćirković and Miloševic-Zdjelar (2004)]. They argued that some types of multiverse theories could provide another possibility for defining the direction of time, because some inflation theories [see, for example, Linde (1990)] could explain the fact that our universe has a very smooth matter distribution near the Big Bang. In particular, Ćirković and Miloševic-Zdjelar (2004) argue that, in addition to the entropy-based approaches mentioned, there is another possibility:

> *'[…] we propose a third approach to the explanation of the thermodynamic asymmetry, which could be labelled the Acausal-Anthropic approach. […] It is essentially a Dicke-like approach, applied to the nature (entropy-wise) of the cosmological initial conditions.'* [Ćirković and Miloševic-Zdjelar (2004), pp. 5, 6]

In the class of multiverse theories, we have more than one universe, where each universe can be called a cosmic domain. One of these domains is our universe. Also, each of these domains could have different initial conditions. Thus, if the occurrence of smooth states near the Big Bang has a probability of, for example $1 : 10^{10^{123}}$ [Penrose (1979)], the prediction of the existence of cosmic domains that include such smooth early states seems very plausible, as long as the number of cosmic domains is assumed to be much larger than the reciprocal of the probability of their occurrence. This prediction is as plausible as the prediction of getting a 6 at least once if a fair six-sided dice is rolled *n* times, where *n* is much larger than six.

Authors who support this view occasionally call it an anthropic approach [Ćirković and Miloševic-Zdjelar (2004)] for the following reasons. At first glance, it may appear very surprising in such a theory that the observable universe belongs to this minority of domains, which are only as likely as $1 : 10^{10^{123}}$ (for example). If other domains are more likely, why do we not observe one of them in our cosmic environment? At this point, it becomes possible to consider the weak anthropic principle. The answer to the question would be that the

more likely domains cannot be observed, because no human being can survive in such a universe, in which the existence of stars or even atoms is unlikely.

Thus, the anthropic principle is not used to explain a cosmological fact. The fact is explained by physical theories, independently of any anthropic considerations. The anthropic principle is used only to clarify that it is not surprising that our particular universe belongs to a small minority because otherwise it would not have been *our* particular universe. However, what is interesting is that we find a physical theory that explains the occurrence of certain initial conditions in a particular cosmic domain. Thus, according to such types of multiverse theories, an acausal–particular approach could be understood as fundamental in the sense that the basic properties of the laws of the universe, in such theories, show that some particular boundary conditions occur (in some domains) and provide an arrow of time.

Thus, in addition to causal–general and acausal–particular approaches, which were also presented by Price (2002), we could consider this acausal–anthropic approach in our study. I give this name because it is used in Ćirković and Miloševic-Zdjelar (2004). Henceforth, I will also refer to the approach as the entropic–anthropic approach. However, upon closer examination, this time asymmetry cannot be understood as being based on the fundamental laws of the considered models, for the following reason.

The multiverse theories predict a very large number of cosmic domains. In addition, the time parameter, which is fundamental in this context, is a quantum parameter independent of particular cosmic times in some cosmic domains. The laws that give rise to the fundamental processes of creating different cosmic domains (which could have different cosmic times) are processes in physics (described e.g. in string theory, M-theory or ordinary QFT) that are understandable as time symmetric according to the fundamental time parameter. The fundamental laws and mechanisms of those theories also allow many cosmic domains, which could be time symmetric in terms of their cosmic times. In such domains, therefore, the behaviour of entropy is symmetric, or the value of entropy is constant apart from fluctuations.

Thus, according to the entropic–anthropic approach, we find an explanation for the occurrence of time asymmetric behaviour in our particular cosmic domain, but this asymmetry occurs by accident. The entropic–anthropic approach explains only that it is not *surprising* that we find ourselves in a cosmic environment such as the observable universe. Nevertheless, the temporal direction is not based on basic properties of the theory that is treated as fundamental (here, e.g. QFT) and time symmetric domains are still possible in general.

Thus, we have seen that all the approaches described in this section fail to describe the existence of a time arrow that could reasonably be called fundamental (for some reasonable understanding of fundamentality in this context). This holds for:

> a) all the causal–general accounts,

> b) the acausal–particular accounts and

> c) the acausal–anthropic considerations.

Note however, that approach c) is rejected only as an explanation of the origin of a cosmological time arrow that is *fundamental*. The motivation for the acausal–anthropic account of Ćirković and Miloševic-Zdjelar (2004) was instead as follows:

'Having already obtained from modern cosmology a useful notion of the global system that includes all the different (perhaps infinitely many) cosmic domains (such as our particular spacetime region), we could as well employ it to account for the prima facie extremely improbable 'choice' of initial conditions. In other words, we imagine that everything that exists represents a 'Grand Stage' for the unfolding of thermodynamical histories of chunks of matter. Moreover, the entire system, seen as the entire multiverse consisting of all the cosmic domains, immediately solves the problem of the extraordinarily improbable endpoints of those chunks we observe in our vicinity. Entropy in the entire system is high almost everywhere. Our particular cosmic domain represents a natural fluctuation (according to the set of possible initial conditions); however, the anthropic selection effect answers the question of why we find ourselves on an upward slope of such a fluctuation. Hence, what we must explain is not that such fluctuations exist, nor that the local initial condition has an extremely low probability, but the fact we happen to live in such an atypical region of the entire system, which is almost always at equilibrium. This can be explained by determining why the observed entropy gradient is required for our existence as intelligent observers.'

Given that motivation and the claims mentioned above, I agree entirely with the position, but the point here is simply that the consideration of this type of cosmological models, including a large or infinitely large number of cosmic domains, is not sufficient to show that the cosmic time asymmetry is given for fundamental reasons. According to such models there are many other cosmic domains, which could have also cosmic times, but there could behave symmetrically (probably the entropy value could fluctuate about a maximum value). Hence the entropic time asymmetry is not given for fundamental reasons but by accident (a lucky one, because otherwise we would not exist in this cosmic domain). Thus, the particular asymmetry of cosmic time in our cosmic domain is not a fundamental asymmetry (it is not based on fundamental properties of the theories or models used to describe the multiverse).

However, I shall discuss another prominent approach to defining an arrow of time in cosmology. This approach is independent of the behaviour of entropy.

III.1.3. Hyperbolic Curved Spacetimes

Many modern cosmological models of the evolution of our particular universe have a property that we have not yet discussed in this chapter. Observations and theoretical work support the idea that the three-dimensional universe exhibits accelerated expansion in cosmic time [e.g. Riess et al. (1998) on supernovae observations; Barlett and Blanchard (1996) on the cosmic virial theorem; Fan, Bahcall and Cen (1997) on mass indicators in galaxy clusters; Bertschinger (1998) on large-scale velocity maps and Kochanek (1995) or Coles and Ellis (1994)]. This could be described in general relativity as a large positive value of the cosmological constant. Moreover, this seems to indicate that it is plausible to assume that the universe is an open one, not only because the matter and energy density are too low to overcome expansion but also because the expansion is accelerated by a force described by the cosmological constant.

Thus, given the empirical data from observational astrophysics, the origin of the cosmological time arrow can be identified simply by the fact that the universe has, for physical reasons, an accelerated expansion and thus a time asymmetric (regarding cosmic time) open spacetime. For example Ćirković and Miloševic-Zdjelar (2004), in criticizing Price (2002):

> '[…] the massive evidence for a large positive cosmological constant […] obliterates prospects for any recollapsing universe, and a fortiori the prospects for a very special case of recollapsing universe, such as Gold's. […] Ignoring this development, as well as the entire tradition of observational cosmology in at least 30-odd years of history of attempts to measure the cosmological density fraction Ω, certainty deserves a dictum of Earman […].' [Ćirković and Miloševic-Zdjelar (2004), p. 11]

However, the question arises of whether this large value of the cosmological constant occurs accidentally or owing to some law of physics. If the accelerated expansion of the universe were the result of a fundamental law, an arrow in cosmic time could be defined as the direction pointing to the open end of spacetime, and this arrow could be based on the properties of the fundamental law. However, I must argue that this is not convincing given our current knowledge of cosmology.

Note first that the value and origin of the cosmological constant remain a crucial issue in fundamental physics and cosmology. Nevertheless, I briefly examine one possible account because Ćirković and Miloševic-Zdjelar (2004) attempted to argue that a particular account of the origin of the cosmological constant yields a special understanding of the arrow of time in cosmology. This account is that the cosmological constant is a result of vacuum polarisation. The effect of vacuum polarisation is surely very fundamental; moreover, it arises from the fundamental structure of QFT.

Nevertheless, I shall argue that this assumption does not allow us to conclude that the open geometry of spacetime yields a fundamental asymmetry of cosmic time based on the properties of a fundamental law. This is because the force of gravity, which opposes the cosmological constant, depends on the matter and energy density of the universe. Thus, the critical value that the cosmological constant must exceed to yield an accelerated universe in such a semi-classical model depends on the matter and energy density. This density does not seem to be determined by QFT. Moreover, most types of cosmological theory suggest that this density could vary among cosmic domains. If such variation is possible, it follows that the effect of the accelerated expansion of the universe could be used to define an arrow of time only in some particular cosmic domains. However, I think this arrow should not be understood as fundamental, because we notice the same situation in standard classical cosmology, where a spacetime could be open or closed, depending on the mass and energy density. Thus, an account based on the positive value of our particular cosmological constant yields the same problems as the accounts mentioned above: A closed universe is *possible* even if the value of the cosmological constant is assumed to be positive for fundamental reasons. As long as we prefer to call an arrow of time fundamental only if the time direction is based on fundamental properties of the theory used to create the model, Ćirković and Miloševic-Zdjelar (2004) cannot provide new insight on this question by attributing the large value of the cosmological constant to vacuum polarisation. As mentioned, this is because a closed spacetime geometry is not ruled out (in principle) by a positive cosmological constant. Thus, the fact that our particular universe or cosmic domain seems to exhibit an accelerated expansion cannot be used to conclude that we can define a cosmological time asymmetry based on *fundamental* laws of physics.[7]

---

[7] Accordingly, it may be mentioned that the current overwhelming evidence that the universe will not recollapse represents empirical data that should certainly be considered in any discussion of the asymmetry of time. In addition, there are also theoretical reasons that the closed universe is no longer a likely option for *our particular* spacetime. The list of references containing the empirical evidence for $\Omega_{matter} < 1$ and $\Omega_{\Lambda} > 0$ certainly represents a convincing set of cosmological data. Thus, the account of the arrow of time in cosmology based on the hyperbolic curvature of spacetime originates in cosmology as far as possible for any type of ontological question. This all may be true, and my goal is not to question the validity of the hyperbolic curved spacetime view. However, as shown, this view alone is not sufficient for

Consequently, the attempts to define an arrow of time in cosmology that we have discussed so far do not yield a fundamental understanding of some sort of cosmological time asymmetry. In the following section, I present my suggestion for understanding the origin of the arrow of time in classical cosmology.

III.2 On a Fundamental Time's Arrow

III.2.1 A New Approach

In this section, I present my suggestion regarding the problem of *fundamental* time directions in the context of classical cosmology. I show, by adoption Castagnino, Lara and Lombardi (2003a) and Castagnino and Lombardi (2009), that an arrow of time can be deduced from the structure of the solution set of crucial dynamical equations. Moreover, I shall demonstrate under which conditions such an arrow could be understood as fundamental (according to chapter II).

This section is organized as follows:

1. I present two conditions on the considered solutions set that are necessary to support my argument. These conditions are motivated by methodological and physical considerations, as described later.

2. The set of solutions of the crucial equations that satisfy these conditions (point 1) contains time symmetric and time asymmetric solutions. I show that under some physically motivated conditions, the space of the time symmetric solutions has dimension $m < n$, where n is the dimension of the total space of solutions that satisfy the same conditions (following Castagnino, Lara and Lombardi (2003a) and Castagnino and Lombardi (2009)). Thus, in almost all[8] solutions, we could say that time asymmetry is a generic property of almost all spacetimes under the assumed conditions. Thus, condition i) of definition I is satisfied.

---

providing a *fundamental* time direction in cosmology. Even if the empirical evidence combined with the theoretical work gives us reason to assume that our particular cosmic domain has a time arrow in cosmic time, the empirical evidence is not sufficient for concluding that this time arrow is a fundamental property of nature or the physical models we use to describe nature. In fact, as shown, it turns out to be not fundamental even if it occurs in our particular universe.

[8] In the sense given by i) in definition I, chapter II

3. Furthermore, I show that the space of solutions that satisfies the assumed conditions also satisfies condition ii) of definition I from chapter II.

4. I reject a possible objection to my conclusion.

III.2.2. Crucial Conditions

The first of two conditions that I shall introduce in this analysis is that when investigating the solution set of the crucial dynamical equation, I consider only those solutions for which cosmic time can be defined.

This assumption, of course, seems prima facie extremely critical. Solutions of the Einstein equations that are not time orientable exist; hence, cosmic time cannot be defined in such spacetimes. At first glance, it appears that an approach that assumes the definability of cosmic time could not lead to a fundamental time asymmetry because this assumption is not motivated by the underling theory.

I will show why I believe that the assumption of the definability of cosmic time is acceptable. In classical cosmology, the only time coordinates that appear at a fundamental level of description are the proper times of different elementary physical systems. Consider the simple example of two parallel world lines. Their proper times could be synchronized, but their directions (i.e. the distinction between past and future semi-light cones) are generally independent of each other. If the directions of proper times on different world lines can, in principle, be connected, the time orientability of spacetime must be assumed. Thus, if we try to deny this assumption in order to achieve greater generality, we cannot discuss the direction of time in general but only the direction of time according to one world line. This is because of the existence of possible spacetimes that represent possible solutions of the Einstein equations that are not time orientable.

Moreover, note that in addition to the assumption of time orientability, we must assume the definability of cosmic time (which implies time orientability). This is because otherwise we cannot refer to a time parameter that allows the conception of time asymmetries for more than *one world line* in general.

To make this point clear, my suggested application of definition I to classical cosmology cannot define a direction of time that fulfils condition i) of definition I without this additional assumption. However, if the notion of time directions for more than *one world line* is self-

consistent (which means that the condition is fulfilled), cosmology provides a 'fundamental' time asymmetry.

In summary, I claim to show that my suggestions produce the following result:

> *By assuming the possibility of producing even a hypothetical sense of a time directions valid for* more than one world line*, we find that classical cosmology reveals that a fundamental arrow of time exists, given the restrictions described above.*

The second additional condition is that, in the considered set of spacetimes, the dynamics of spacetime as a whole is only describable by using more than one independent dynamic variable. This seems plausible on the basis of physics. The point is that a classical toy model that includes only dynamic variations in the scale factor cannot describe the dynamics of the energy and matter content of spacetime, which is usually described by the dynamic behaviour of an additional matter field. Thus, in the simplest case of a scalar matter field, we need at least two independent dynamic variables, not just the scale factor, in order to describe the dynamics of spacetime. As mentioned in section III.1.1, if we assume that the scale factor is the only dynamic variable at a fundamental level [see, for example Price (1996)], it follows that spacetime is generically time symmetric on a fundamental level. Thus, all solutions of the dynamic equation for such a spacetime are time symmetric. This situation changes in a more physical model that includes more dynamic variables in addition to the scale factor. This also shows that the critique of the account in Price (1996) in section III.1 is based on physical considerations. These consideration can be sketched in terms of the need to describe the dynamics of the energy and matter content of spacetime, which is impossible in a model that has only the scale factor [or the radius, in the terminology of Price (1996)] as a fundamental quantity (e.g. the Gold-model, which is considered by Price (1996)).

However, in the next subsection, I show in detail how we can define a fundamental arrow of time in classical cosmology by using the conditions described above.

III.2.3. Symmetric Spacetimes

To define a fundamental time asymmetry, I have to analyse the types of spacetimes that allow the definition of cosmic time and are time *symmetric* with respect to cosmic time. I shall show, by combining the outcomes from Castagnino, Lara and Lombardi (2003a) and Castagnino and Lombardi (2009) with crucial singularity theorems, that such spacetimes belong to a subset of solutions having a lower dimension than the entire set of solutions

(according to a dynamic equation that describes spacetime as a whole). This fulfils part i) of definition I in chapter II. Moreover, I also show that condition ii) of definition I is satisfied by the solution space. In order to do so, I will first concentrate to a very simple toy example and I will argue that, in that toy model, condition i) from definition I is fulfilled. Additionally, I will argue that the applicability of condition i) from definition I does not depend on the simplifications made in that toy model hence, that condition I of definition I is a generic property of classical cosmology if the mentioned conditions regarding the definability of cosmic time and the number of dynamic variables are fulfilled.

Most open spacetimes are time asymmetric according to cosmic time. This is because we can define the time arrow of cosmology in an open spacetime according to the asymmetric behaviour of the scale factor. Thus, we will not consider open spacetimes in this section, although they seem to yield the correct description for our particular universe. In the context of classical cosmology, such spacetimes are time asymmetric with respect to cosmic time, and we seek the origin of time symmetric spacetimes in classical cosmology (in order to show that there belong to a subspace of measure zero).[9] Consequently, there is no need to consider open spacetimes in order to examine the mathematical origin of time symmetry with respect to cosmic time.[10]

Therefore, I must concentrate on closed spacetimes. According to singularity theorems [Hawking and Penrose (1970); Hawking and Ellis (1973)], such spacetimes, at least in classical cosmology, have just one maximum of the scale factor. Hence, this class of spacetimes could be time symmetric. Thus, we concentrate on such spacetimes to demonstrate the mathematical origin of cosmic time symmetry. This starting point makes it possible to follow parts of the analysis from Castagnino, Lara and Lombardi (2003a p. 374-375) and Castagnino and Lombardi (2009):

---

[9] Note that there is one exception in classical cosmology: a static universe that is open but also time symmetric. However, we will not consider the static solution of the Einstein equations because it requires fine tuning of the cosmological constant and the energy and matter content (and distribution) of the universe. Thus, according to classical cosmology, this solution is a special type that belongs to a subspace of solutions to the Einstein equation that surely has a lower dimension than the entire solution space.

[10] Spacetimes that have an open but time symmetric and non-static geometry are open to the past and future. I do not consider them because they require a change in the value of the cosmological constant, and in the context of classical cosmology, the cosmological constant is constant in cosmic time. This may change in string or loop cosmology, but that is beyond the scope of this chapter. In classical cosmology, a contracting spacetime always has a Big Crunch [see Hawking and Ellis (1973) and Hawking and Penrose (1970)]. Thus, in classical cosmology, a spacetime cannot be open in two directions of cosmic time if the spacetime is not static.

Consider, for simplicity, the very simple toy example where the dynamics of spacetime is described by the scale factor $a(t)$ and a scalar matter field $\phi(t)$, which depend on cosmic time t. In Hamiltonian mechanics, the dynamic equations (and thus the Hamiltonian) depend on the dynamic variables and their first derivatives with respect to t. Thus, in our example, we have four arguments in the Hamiltonian, $a(t), \dfrac{da}{dt}, \phi(t), \dfrac{d\phi}{dt}$. Analytical mechanics always allows us to describe one of these variables as a function of the others, and the choice of variable that depends on the others is just a matter of description. Thus, for simplicity, I choose $a(t) = f\left(\dfrac{da}{dt}, \phi(t), \dfrac{d\phi}{dt}\right)$, where $\dfrac{da}{dt}, \phi(t), \dfrac{d\phi}{dt}$ are now independent variables.[11]

If we try to construct a time symmetric spacetime, all the variables together must behave in a time symmetric way. According to the singularity theorems of classical cosmology, we know that $a(t)$ has just one maximum. Next, we can choose the mathematical origin of cosmic time. For simplicity, I choose this origin so that $a(0)$ is exactly the maximum value of the scale factor. Thus, $a(t)$, as a function of cosmic time, is symmetric in relation to the a axis at the point $t = 0$. From this, it is obvious that $\dfrac{da}{dt}$ is symmetric in relation to the point $\left(t = 0; \dfrac{da}{dt} = 0\right)$. However, for such a spacetime to be time symmetric, the behaviour of $\phi(t)$ and $\dfrac{d\phi}{dt}$, together with that of $\dfrac{da}{dt}$, must also be symmetric. Thus, in this example, only two possible behaviours of $\phi(t)$ and $\dfrac{d\phi}{dt}$ at the cosmic time point $t = 0$ yield an entire spacetime to be time symmetric. Those possibilities are given by the triplets $\left\{\left.\dfrac{da}{dt}\right|_{t=0} = 0, \phi(t=0), \left.\dfrac{d\phi}{dt}\right|_{t=0} = 0\right\}$, which is a symmetric solution of $\phi(t)$ with respect to the ϕ axis at $t = 0$, and $\left\{\left.\dfrac{da}{dt}\right|_{t=0} = 0, \phi(t=0) = 0, \left.\dfrac{d\phi}{dt}\right|_{t=0}\right\}$, which is a symmetric solution of ϕ with respect to the point $\left(t = 0; \phi(t=0) = 0\right)$.[12]

---

[11] The Hamilton equation in this case is given by $H\left(\dfrac{da}{dt}, \phi, \dfrac{d\phi}{dt}\right) = 0$.

[12] The symmetric solution of ϕ with respect to the point $\left(t = 0; \phi(t=0) = 0\right)$ is anti-symmetric. But, the sign of ϕ is not necessarily physical meaningful. Hence, the

Therefore, all symmetric solutions can be constructed using these triplets. Thus, we can construct a subspace of time symmetrical solutions: $span\left\{ \begin{pmatrix} \left.\dfrac{da}{dt}\right|_{t=0}=0 \\ \phi \\ \left.\dfrac{d\phi}{dt}\right|_{t=0}=0 \end{pmatrix}, \begin{pmatrix} \left.\dfrac{da}{dt}\right|_{t=0}=0 \\ \phi(t=0)=0 \\ \left.\dfrac{d\phi}{dt}\right| \end{pmatrix} \right\}$. The complete space of solutions of the dynamic equation is given by $span\left\{ \begin{pmatrix} \dfrac{da}{dt} \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \phi \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ \dfrac{d\phi}{dt} \end{pmatrix} \right\}$.

Thus, the time symmetric behaviour of a spacetime appears only in a subspace of solutions having a lower dimension than the entire solution space, even if we only consider closed spacetimes. This also makes the criticisms of Price (1996) explicit. The result shows that even almost all closed spacetimes could be time asymmetric. The meaning of this fact is illustrated in the geometrical drawing in Fig. II.1 in chapter II.

Therefore, according to the given toy model, we see that time asymmetry is a generic property. However, this also holds if we add more dynamic variables, because the calculation in those cases would be similar, and the entire space of solutions always has a higher dimension than the subspace of time symmetric solutions [except if we have only one dynamic variable, e.g. the scale factor, which is the case in the analysis of Price (1996)]. Also, the particular form of the Hamiltonian can vary and additional dynamic variable's, describing different cosmological properties or also quantum fluctuations (e.g. of the metrical tensor), can be added to the Hamiltonian. All this would not change the result from the simplified toy model, which is that almost all solutions of the dynamic equation are intrinsically time asymmetric. Thus, as long as the mentioned conditions are fulfilled, time asymmetry is a generic property of the set of possible spacetimes.

But, we have to examine now whether condition ii) of definition I is also satisfied by the solution set.

---

combinations $\left\{ \left.\dfrac{da}{dt}\right|_{t=0}=0, \phi(t=0)=0, \left.\dfrac{d\phi}{dt}\right|_{t=0} \right\}$ has to be considered in order to capture all symmetric possibilities.

III.2.4. Solution Set

The solution space for closed spacetimes that allow the definition of cosmic time and that have physical dynamics[13] is, for mathematical reasons, built of time-mirrored pairs of functions $f(t)$ and $f(-t)$, as is the case for many other time-dependent dynamic equations. Thus, every time asymmetric solution $f(t)$ has a pair function $f(-t)$, which is also a solution to the dynamic equation. They are intrinsically asymmetric, but the directions of the asymmetries seem to be mirrored. We are certainly not worried about a physical superposition of both spacetimes, but we could argue that the approach I have suggested so far will not explain why $f(t)$ occurs instead of $f(-t)$ or vice versa.

This concern motivates me to choose a cosmological example to demonstrate the applicability of definition I to physical theories. Otherwise, I would have no argument for favouring one time direction, and I must refer to boundary conditions [as in Albert (2000)], undiscovered laws of nature [as suggested in Price (1996)] or accidental facts or anthropic reasoning's. However, in this case, I can present an argument to defend my example and the suggested understanding of fundamentality.

I shall provide an argument for the proposal from Castagnino, Lara and Lombardi (2003a) and Castagnino and Lombardi (2009), which is that the pairs $f(t)$ and $f(-t)$ are physically identical; thus, they do not refer to physically different worlds. My argument combines a Leibniz argument in three steps with the argumentation that Leibniz principle is applicable to the given situation. The steps are the following:

> i) The solution $f(t)$ does not include intrinsic properties that are not included in the same way in $f(-t)$. This is because they are only mirrored geometrical objects (spacetimes).

> ii) Both are global solutions that describe spacetime as a whole. This implies that there is no time parameter (or other physical parameter) outside of the geometrical objects $f(t)$ or $f(-t)$. Thus, they are not related to an outstanding circumstance.

Thus:

> iii) Points i) and ii) together show that two time-mirrored spacetimes $f(t)$ and $f(-t)$ do not differ in their intrinsic properties, *and* they do not differ in any external relationship. Thus, they describe the same physical world.

---

[13] This phrase means: 'are described by more than one independent dynamic variable'.

Here, it is not the case that we just cannot *distinguish* between the two solutions. Taking general relativity (or empirically equivalent spacetime theories) as the fundamental theory in classical cosmology, my argument shows that there cannot be any property that differs between the two solutions. The minus sign in front of *t* refers only to the notion of an assumed, but not physical, absolute Newtonian background time direction. Thus, if both *f(t)* and *f(−t)* are identical in all external and internal physical properties, and the only difference is the time alignment of the solutions in a purely mathematical and non-physical absolute Newtonian coordinate system, we can conclude that both solutions describe the same physical world according to the physical theory that is treated as fundamental (general relativity or empirical equivalent spacetime theories). Note also that this argument arises not because of some accidental reason for the fact that there is no background time in theory, but because there *cannot* be an absolute background time or space, which is one of the fundamental lessons of general relativity. The fact that there are no relational structures is not merely accidental in general relativity; it is a most crucial lesson that I believe should be treated seriously. Thus, Leibniz's argument seems to be appropriate here, and as mentioned earlier, the conclusion is that the solutions are identical and are only labelled differently. Consequently, the solution space is not built from *physical* time-mirrored pairs.

However, there is an objection to the argument that I have presented see Earman (1974). I discuss this objection in detail in the next subsection.

III.2.5. The CPT Objection

Earman presented a general objection against some types of Leibniz arguments of the form I used above. He showed that according to CPT-symmetry,[14] it follows from the conclusion that *f(t)* and *f(−t)* describe the same physical world that the CP signs are also not physically meaningful. This seemed incorrect to Earman, and he concluded that the argument presented above cannot be accurate. However, there is a loophole in Earman's argument: again, we have no external background. Here, I shall show in greater detail how we could use this loophole to establish a similar form of Leibniz argument regarding the C and P transformations.

---

[14] C is the charge transformation, P is the parity transformation and T is the time transformation. These transformations reverse the signs of the charge (C), parity (P) and time (T). Physics seems to be invariant under combined transformations of C, P and T. This is the meaning of CPT-symmetry.

But firstly, I must mention a debate on whether to interpret parity as an intrinsic or a relational property, which has a long tradition in philosophy [e.g. Leibniz (1686), (1714) or Kant (1768), (1770)] and also in the modern philosophy of science [see e.g. Pooley (2002) or Frederick (1991) for a modern introduction to this field]. I claim that this interesting discussion will not affect my arguments, regardless of whether a relational or an intrinsic interpretation of parity is favoured. We will see this soon.

However, I begin with the C transformation. The sign of a charge is clearly a conventional label. It is also clear that a positive and a negative charge in an everyday environment differ physically from each other. In classical electrodynamics, the direction of the field lines is reversed if we change the sign of the associated electric charge. However, without a background, the direction of the field lines has no physical sense. Again, similar to the situation with time-mirrored pair functions, the following holds:

> i) The intrinsic properties of two spacetimes including electromagnetic fields (including the sources) that differ from each other only in the direction of the field lines (and thus in the sign of the electric charges) are identical, because they are only mirrored.

> ii) If there are no physical surroundings, there are also no relational structures that differ between two C-mirrored pair spacetimes.

Thus

> iii) Both field configurations (formally distinguished by the signs of the sources) are physically identical if we do not consider the external environment.

Outside of spacetime, no electromagnetic fields exist (note that the solutions of the dynamic equation discussed here describe spacetimes as a whole); thus, we cannot define extrinsic relationships. Again, this is not accidental but a crucial lesson from general relativity. Thus, we find the same situation for the direction of time as for the direction of electric field lines (or magnetic field lines, which are equivalent in special relativity). Consequently, we find the same situation for C as for T transformations. Moreover, I believe that Leibniz's argument is as appropriate as in the former case, and for exactly the same reasons.

Now I will show that the same holds for the P transformation. The parity transformation that is used to formulate the CPT-theorem [for example, Gross (2004)] is a special type of Lorentz transformation given by $P(t) = t$ and $P(x) = -x$, where t is a time coordinate, and x represents spatial coordinates. Thus, the P transformation is the mirroring of space. Again, consider a solution of a global dynamic equation that describes the properties of spacetime as a whole, $f(t,c,p)$ (where t, c and p now refer to a time coordinate, the sign of C and the parity, respectively). The P transformation of this global solution is given by $P(f(t,c,p)) = f(t,c, -p)$.

Here, the minus sign in front of p indicates that all space coordinates are mirrored. However, they are mirrored according to the notion of a former absolute (and thus not physical) coordinate system. In the context of general relativity, we clearly do not have absolute directions of space outside of spacetime. Hence, again, we find the following:

> i) All intrinsic properties of two spacetimes, $f(t,c,p)$ and $f(t,c,-p)$, are identical and are just mirrored in an assumed mathematical background space.

> ii) No external relationships exist for the spacetimes $f(t,c,p)$ or $f(t,c,-p)$.

Thus

> iii) Both solutions, $f(t,c,p)$ and $f(t,c,-p)$, are physically identical.

Note that this indicates nothing about the possibility that spacetime itself (as one particular solution of the Einstein equations) could include an asymmetry of space (or time or electric field lines). In fact, the aim of this chapter is to show that such a time asymmetry appears in almost all spacetimes. Consequently, one spacetime can have the property that p and $-p$ (for example, left- and right handedness) differ from each other in a physical sense. However, the p-mirrored spacetime of such a space asymmetric spacetime has the same asymmetry, and both spacetimes are identical, because there *cannot* be an absolute space independent of spacetime that defines external relationships. Hence, this argument could not be made to show that, for example, a relational interpretation of parity is more plausible than an intrinsic one, because in a particular spacetime, the parity transformation could have a physical meaning. Nevertheless, we are again in the same situation as outlined for the T or C transformation, and we can use the outlined Leibniz's argument to conclude that C, P and thus the CP sign (in our case) have no physical meaning.

Note that my entire argument works only for such global solutions of global dynamic equations that describe spacetime as a whole (this is why I refer to cosmological equations). In the context of particle physics, where the CPT-theorem was developed, we usually discuss the transformation of particle properties or systems constructed from particles (or field excitations). In this case, the P value is important, as seen by the effect of parity violation in weak interactions. However, in the case discussed above, Earman's counterargument can be rejected only for a global dynamic equation and the associated solution space of *global* solutions. This also shows why the adequacy of the CPT-theorem seems unable to change the whole issue of time directions in physics. This is simply because the theorem is adequate regarding the description of fundamental interactions between particles (or field excitations).

More precise: If we would use the CPT-theorem in order to define a time asymmetry, the consideration, I think, should be similar to the following.

1. If we consider a physical system and its associated time evolution, find two formal distinguished evolutions (in two time directions) distinguished by the sign of t.

2. Both evolutions can be distinguished, using the CPT-theorem, by the CP sign.

3. This means that there is a physical difference between both evolutions, because the system associated with the positive time evolution would interact differently with a physical environment than the system associated with the negative time evolution (simple because of the CPT-theorem). Hence, we have a direction of time.

But, obviously, this account depends on the interaction with a physical environment. Point three is simply wrong if the considered system is the whole system (the universe). A way of making this more vivid, I think, is to consider point three and additionally assume that the system associated with the positive time evolution (CP-sign:=(+,r)) interacts with a physical environment, which evolves in the positive time direction. The system associated with the negative time evolution (CP-sign:=(-,l)) interacts also with a physical environment, which evolves in the negative time direction (all CP-signs switched) (this is equivalent to the assumption that the whole system is T, respectively CP transformed). Now, the assumed physical differences in the behaviour of both systems vanishes (both system interacting identical with their associated physical environments), which shows that point three seems not adequate if the considered system is global. That, indeed, seems crucial if the time direction should be a *fundamental* time direction. Thus, the CPT-theorem seems not directly connected to the question of *global* and *fundamental* time directions.

So, we conclude that the solution space of a dynamic equation that describes spacetime as a whole does not consist of physically different pairs $f(t)$ and $f(-t)$. Therefore, we find a fundamental time asymmetry in the considered solution space. This, so it seems, is a generic property of classical cosmology as long as the mentioned conditions on cosmic time and the number of dynamic variables are fulfilled.

So, we see that the definition of fundamental time directions in chapter II is applicable to classical cosmology and that the time-reversal invariance of the Einstein equation and of the crucial Hamiltonians is irrelevant for a fundamental time asymmetry based on the structure of the equation's solution space. This perhaps surprising fact can be explained in a bit more detail. In order to do so, I present a very short digression (Digression III.B at the end of this chapter). This is not crucial for the entire investigation, but, I think, it could be fruitful to examine more closely how the main philosophical thesis from chapter II solves the crucial problem regarding cosmic time asymmetries.

However, I think it is fruitful to stress some general advantages of the suggested non-entropic understanding of the cosmological time asymmetry in order to make clear that there are independent motivations, in addition to the definition of fundamentality from

chapter II, that motivate me to use the proposed non-entropic definition of the cosmological time arrow. However, I also think that my suggestion reflects the tradition of geometrical attempts to define the cosmological time arrow.[15] Although I use a new concept of fundamentality (chapter II) and present an argument to show that this concept of fundamentality is applicable in classical cosmology, the models and general ideas I use are often discussed in the literature; see, for example: Castagnino, Giacomini and Lara (2000), Castagnino, Giacomini and Lara (2001), Hawking (1994) on the *Physical Origins of Time Asymmetry,* Castagnino (1989), Castagnino and Mazzitelli (1990), Castagnino and Lombardo (1993), Lichnerowicz (1955), Visser (1996) on *Lorentzian Wormholes,* Barceló and Visser (2002) on *Twilight for the energy conditions?*, Landau and Lifshitz (1970), Castagnino, Gadella, Gaioli and Laura (1999), Castagnino and Lombardi (2009), Aquilano, Castagnino and Eiroa (1999) or Vilenkin (1988).

So, I will begin by stress some advantages of the proposed non-entropic account of the cosmological time arrow that are independent of the applicability of my definition of fundamentality from chapter II.

III.2.6. Advantages of the Proposed Non-Entropic Arrow of Time in Classical Cosmology

The first advantage is that the time asymmetries (as explications from the fundamental asymmetry) are *global* cosmic time asymmetries. In an entropy-based approach, it is possible that, apart from fluctuations, the maximum value of entropy is reached; hence, the time arrow would vanish if it is defined by the second law of thermodynamics.

Moreover, according to classical thermodynamics, crucial problems arise for a global understanding of an entropic time arrow in cosmology, because the motivation for the entropy-based approach is based on the second law of thermodynamics. Consider a closed

---

[15] But, I shall show later that in contrast to the global and geometrical proposal from e.g. Castagnino, Lara and Lombardi (2003b) or Castagnino and Lombardi (2009), my proposal (i) is independent from an ontic interpretation of spacetime geometry, and (ii) does not aim to show that the consideration of time asymmetric cosmology, as above, is sufficient to define many time arrows regarding proper times, because the connection between cosmic time asymmetries and asymmetries regarding proper times is rather unclear. This connection, so I shall argue, has to be relieved before the cosmological considerations can be used to define time directions of proper times in an fundamental sense (or as by-products of fundamental asymmetries).

system that has an entropy $S(t_0) < S_{max}$, where S_{max} is the maximal entropy of this system, and t_0 is a time point in cosmic time. Now, for every time point $t_1 > t_0$, thermodynamics predicts that it is very likely that the entropy of the closed system $S(t_1) > S(t_0)$. We could then try to define the future direction of cosmic time as the direction from t_0 to t_1. Note, however, that if we consider a cosmic time point $t_2 < t_0$, classical thermodynamics again predicts that $S(t_2) > S(t_0)$ is very likely. Thus, the time direction t_0 to t_2 meets the same standards as the time direction t_0 to t_1, and we have no argument in the framework of classical thermodynamics to use reach any asymmetric conclusion regarding cosmic time. Hence, we need boundary conditions such as the past hypothesis suggested by Albert (2000) in order to create a cosmic time asymmetry.

If we would like to use the thermodynamic time direction, what really fixes the direction of time is the boundary condition [such as the past hypothesis of Albert (2000)] or other explanations (i.e. in a multiverse) for a special 'initial' state of our cosmic domain. However, boundary conditions can be reasonably and *fundamentally* explained only for local systems because only in this case can we refer to the environment to motivate the occurrence of some special boundary condition. Moreover, as argued above, the multiverse picture seems unable to add new lights on the issue if the considered time asymmetry should be understood as a *fundamental* one. Thus, such an asymmetry, if it is to be fundamental, is a local asymmetry of time valid for a system within an environment, which explains the boundary conditions. Consequently, the proposed non-entropic approach seems to have an additional advantage over the entropy-based approach: A fundamental arrow of time can be a *global* one.

Another advantage that we will see more clearly in chapter V is that the proposed non-entropic approach allows us to deduce other time asymmetries in our particular universe, for example, some particular entropic asymmetries or time asymmetric behaviours of the particle number operator. We will see in chapter V that, in the proposed non-entropic approach, the entropic time asymmetry can be reconstructed and understood as a by-product of the more fundamental time asymmetry.

To keep the structure of the chapter as simple as possible, I do not consider the discussion of a non-entropic versus an entropic attempt to define the cosmological arrow of time in more detail. Moreover this question is not a truly essential question in this investigation, because if it were possible to apply the entropic approach to the cosmological time arrow successfully in order to provide a fundamental time arrow, I would be convinced that both strategies achieve the similar conclusions. However, I still would think (as outlined above) that the proposed non-entropic account would have crucial advantages over the entropy-based understanding of the cosmological time arrow.

However, as mentioned, I try to distinguish between the main claim of this chapter, which is that the definition of fundamentality from chapter II can be successfully applied in classical cosmology, and the debate between the non-entropic and entropy-based approaches in general. Thus, after this chapter I present a short digression (Digression III.A) that tries to show why I think that the non-entropic approach to the arrow of time in cosmology has, in addition to the advantages described, also a *conceptual priority*.

However, given that my proposal is based on the structure of a crucial solution set, whereby every solution corresponds to a spacetime, my proposal could be seen as located in the geometrical tradition. Nevertheless, there is a crucial difference between traditional geometrical approaches and my own proposal. In the geometrical tradition it seems prima face crucial that spacetime geometry is interpreted as a physical or ontic structure. Castagnino, Lara and Lombardi (2003a) put it in the following way:

> *Traditional discussions about the arrow of time in general involve the concept of entropy. In the cosmological context, the direction past-to-future is usually related to the direction of the gradient of the entropy function of the universe. But the definition of the entropy of the universe is a very controversial matter. Moreover, thermodynamics is a phenomenological theory. Geometrical properties of spacetime provide a more fundamental and less controversial way of defining an arrow of time for the universe as a whole. (Castagnino, Lara and Lombardi 2003a p.1)*
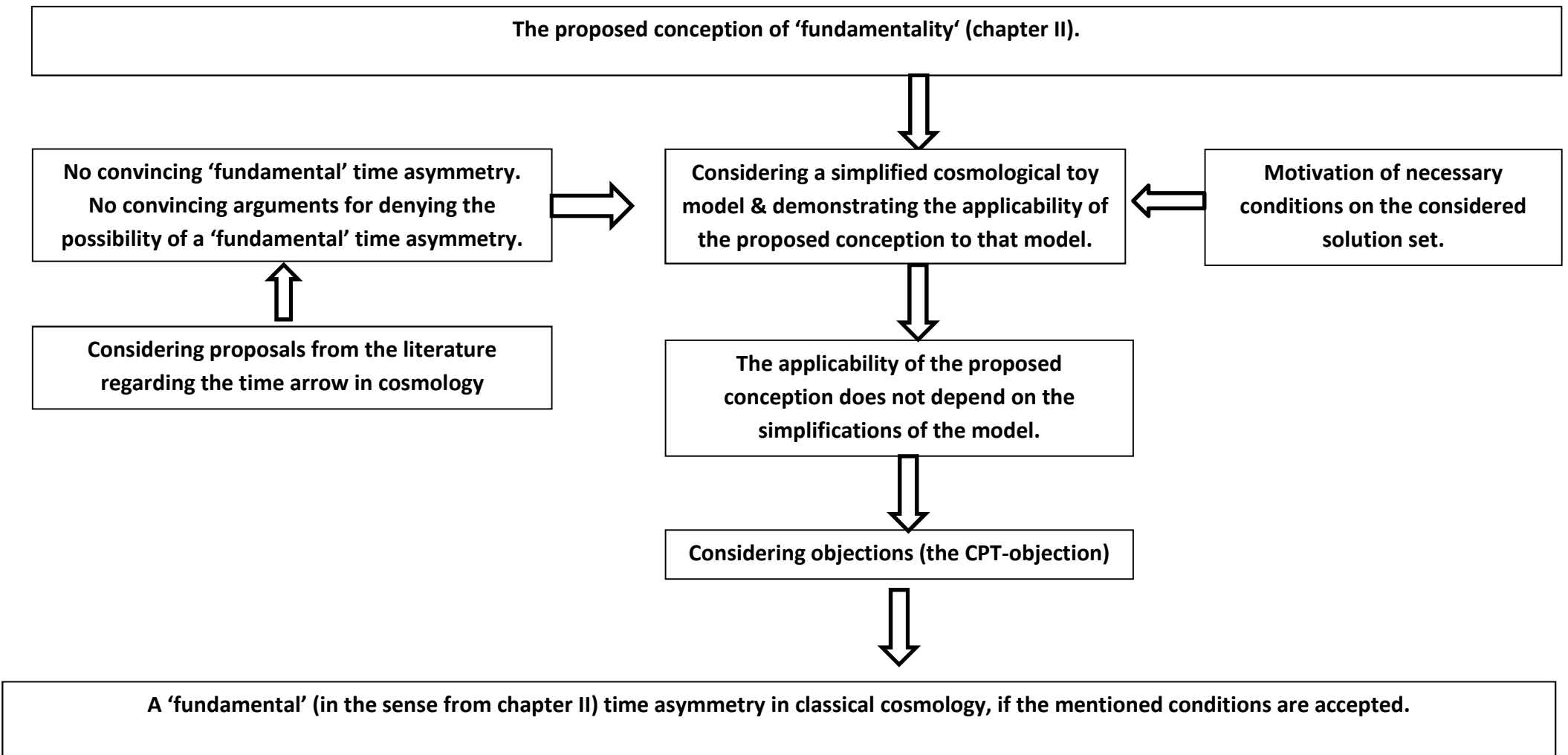
In contrast to that view, my proposal is independent from the ontic status of spacetime geometry. The time asymmetry is explicated in the behaviour of physical entities (e.g. the matter field). Moreover the considered solution set is invariant if we switch from a particular theory (like general relativity) to an empirically equivalent theory with different geometry and different dynamical laws; simply because the theories are empirically (in the solutions and regarding the physical entities) equivalent. This makes the proposal independent from an ontic interpretation of spacetime geometry itself.

This, I think, is an additional advantage of the proposed view also with respect to traditional geometrical accounts to the cosmological time arrow.

III.3. Summary and conclusion

This chapter has demonstrated that the interpretation of fundamentality described by definition I in chapter II can be successfully applied in physical theories, although the

definition is applicable to classical cosmology only if two additional conditions are fulfilled. The resulting situation does not yield a fundamental arrow of time in general, but it shows that if we assume that a time direction could be *possible* for more than one world line (which means that a crucial condition is fulfilled; see section III.2), then classical cosmology provides a fundamental time asymmetry. This asymmetry yields explications in almost all spacetimes, which are physical time asymmetries. However, before the construction of the time asymmetry in this context, I demonstrated that other classes of approaches to the problem of constructing cosmic time asymmetries in classical cosmology (or denying the existence of a fundamental time asymmetry) are misleading, at least when attempting to define a *fundamental* time direction. The structure presented at the beginning of this chapter is also shown in the following diagram, and the critical reader may use it to determine the validity of the arguments and conclusions presented in the analysis above.

The proposed conception of 'fundamentality' (chapter II).

No convincing 'fundamental' time asymmetry. No convincing arguments for denying the possibility of a 'fundamental' time asymmetry.

Considering a simplified cosmological toy model & demonstrating the applicability of the proposed conception to that model.

Motivation of necessary conditions on the considered solution set.

Considering proposals from the literature regarding the time arrow in cosmology

The applicability of the proposed conception does not depend on the simplifications of the model.

Considering objections (the CPT-objection)

A 'fundamental' (in the sense from chapter II) time asymmetry in classical cosmology, if the mentioned conditions are accepted.

Following parts of the physical analysis of Castagnino, Lara and Lombardi (2003a) as well as by the philosophical reflection Castagnino and Lombardi (2009), and a proposed connection to the singularity theorems, I showed that the solution space of the dynamic equation that describes the dynamics of spacetime contains almost exclusively time asymmetric functions. This asymmetry is necessarily embedded in the dynamics of almost all considered spacetimes. Thus, because of the structure of the solution space, I conclude that such a situation satisfies condition i) of definition I from chapter II. Moreover, I showed that condition ii) of definition I is also fulfilled, and hence that the resulting time asymmetry could be understood as fundamental.

I also defended this interpretation against a possible objection. In classical cosmology, I proposed that this CPT objection can be rejected because of the global nature of the solutions of the considered dynamical equation.

To summarize, with the suggested understanding of fundamentality (chapter II), it is possible to construct a fundamental time asymmetry in classical cosmology with respect to cosmic time. This asymmetry occurs in almost all considered spacetimes and it occurs necessarily if the definability of cosmic time is assumed. Also, this condition is motivated by the fact that it is a necessary condition if we want to construct even a self-consistent notion of time asymmetries that is valid for more than one world line.

The next chapter shall consider the understanding of the arrow of radiation (as a local and proper time asymmetry) in some characterisations and in spacetimes similar to ours. We will see this more precisely in during the next chapter.

Digression III.A: Conceptual Priority of a Non-Entropic Arrow of Time in Classical Cosmology

As noted above, this small digression is not relevant to the main topic of this chapter. It is, however, interesting to consider the general discussion between advocates of a non-entropic attempt to define time asymmetries in cosmology and advocates of an entropy-based approach. Thus, readers who are not particularly interested in the discussion of these two strategies may skip the digression.

The entropy-based approaches are discussed at the beginning of this chapter. In this digression, I try to show that the proposed non-entropic approach to construct cosmic time asymmetries in cosmology has conceptual priority over these entropy-based approaches, independent of the question of fundamentality and independent of successful applications in the field. Therefore, for the purposes of this digression, it can be assumed that both approaches could be successful. Here I argue that, in addition to the questions of fundamentality and applicability, the proposed non-entropic approach has conceptual priority over the entropy-based approach. Thus, in order to show that, I will demonstrate two parts independently of each other, as follows.

> i) The proposed non-entropic approach requires some additional assumptions to define a cosmic time asymmetry. If the proposed non-entropic approach should have conceptual priority over the entropy-based approaches, it should not have additional assumptions beyond those of the entropy-based approaches.

> ii) The properties that are used to define the time asymmetry (in the proposed non-entropic approach) should be grounded in more fundamental considerations (where general relativity or empirically equivalent spacetime theories are understood as the fundamental theories in classical cosmology) than the theory of thermodynamics.

Now, for the entropy-based approaches, cosmic time is clearly a very important time coordinate. In classical cosmology, a fundamental level of description includes only the proper times of elementary physical systems. In the entropy-based approach (as well as in the proposed non-entropic approach), another time parameter, cosmic time, is important. However, cosmic time is, at first glance, just a mapping from spacetime states to the real numbers. The sequence of different states of expansion[16] of the three-

---

[16] Note that the sequence of expansion is not necessarily used at this point but is traditional; other physical properties could also be used to define a cosmic time.

dimensional universe, if the universe is nearly homogeneous and isotropic,[17] could be observed by all different observers in spacetime in the same order. The mapping t (cosmic time) is a mapping from the continuous set of expansion states of the three-dimensional universe to the real numbers, and for all observers in spacetime, the sequence in the image of t has the same order. Moreover, **grad**t must be timelike everywhere. Then, this mapping is called cosmic time.

Thus, if we accept that the entropy of the three-dimensional universe could be defined, it is interesting only for the problem of temporal direction according to cosmic time. This is because, in the entropy-based approaches, the variations in entropy are visible to all observers in the same order according to their proper times only if the variations could be described according to cosmic time. Thus, the entropy-based approach could be *valid* only if spacetime has such properties that cosmic time can be defined. Otherwise, not all observers will see the *same* time direction on the basis of the behaviour of the assumed entropy according to their proper times.[18] Thus, the prominent entropy-based approaches make an important assumption regarding considered spacetimes, namely, that cosmic time could be defined.

This assumption is one of two that I used in the proposed non-entropic approach. The second condition was that the dynamics of the entire spacetime is described not only by the scale factor but also, for the matter and energy content, by additional dynamic variables. However, this assumption is strictly motivated for physical reasons, specifically, by the motivation to construct models that can describe the dynamics of the energy and matter content of spacetime. Thus, they are also necessary in the entropy-based approaches, because otherwise the four-dimensional distribution of entropy cannot be time asymmetric when the scale factor is time symmetric. [This is essentially the argument of Price (1996).] Thus, the number of crucial assumptions in the proposed non-entropic approach and the hypothetical entropy-based approach with a minimal set of additional conditions is identical. Hence, the first criterion for concluding that the proposed non-entropic approach has conceptual priority over the entropy-based approach is fulfilled.

Moreover, according to general relativity and empirically equivalent spacetime theories, the used physical entities in the proposed non-entropic account are at least

---

[17] These are not additional assumptions because they are needed only if we use the expansion states of the universe to define cosmic time, which is not necessary. Only the used (for the definition of cosmic time) cosmic field has to be homogeneous and isotropic.

[18] This is important only if the entropy-based approach is used to define a *global* asymmetry of time. Otherwise, different observers could see different local time asymmetries, but this is not a problem if the asymmetry need not be a global one. However, in this section we compare the entropy-based and the proposed non-entropic approaches with respect to defining a *global* asymmetry of time; hence, cosmic time must be assumed for the entropy-based approach.

more basic than thermodynamic considerations (in fact there are the cosmic variables in the Hamiltonian). Thus, a time asymmetry that is defined by the used entities seems more directly linked to the fundamental theory, as is the case of the entropy-based approach. Thus, point b) is also satisfied.

Therefore, I think these are good reasons to believe that the proposed non-entropic approach has *conceptual priority* over the entropy-based approach. Moreover, this seems to be the case independent from a particular interpretation of general relativity.

Digression III.B: On Time Reversal Invariant Equations

In chapter III, we saw that fundamental time asymmetries could be found in classical cosmology. But, it was crucial that the dynamic equation that describes a physical law is global. In this digression, I shall show how the philosophical concept developed in chapter II has solve the crucial problem of fundamental cosmic time directions in chapter III, without considering all the essential investigations in classical cosmology.

As I argued in chapter II, a physical law could be seen as described by a differential equation, but the equation has a set of solutions. A process that we observe in nature is described by the solutions of a differential equation, and hence not *directly* by the equation itself. The solution set of the equation yields the set of physically possible solutions (processes or spacetimes) that could be observed or predicted as processes in nature without breaking physical laws. As mentioned in chapter II, a crucial time asymmetry could arise directly in a differential equation if and only if this equation is not time-reversal invariant. Given the laws of nature as we know them, it seems that we have no differential equation that is not time-reversal invariant *and* describes a *fundamental* law of physics.[19]

However, a time-dependent differential equation has a set of solutions that can contain

> i) time symmetric solutions and

> ii) time asymmetric solutions (in which case, the *t*-mirrored function of a solution is also a solution to the differential equation).

So, a time-reversal variant differential equation generally can have such *t*-mirrored pairs in the set of solutions as well. Thus, the issue that I had to address in responding to the CPT objections and in order to demonstrate that condition ii) is fulfilled does not arise because the laws of nature are time-reversal invariant. Moreover, one motivation for the view in chapter II would *still be given* in a world that includes time-reversal variant laws. This motivation is that, in the proposed view, we would refer to physical *processes* (solutions of differential equations), which are *observable* in nature, and not to abstract physical laws. Thus, my example of classical cosmology would also be similar if the Einstein equations were not time-reversal invariant because I would deal with the same problems as in the former case, and one

---

[19] Perhaps such equations describe quantum laws, but this view depends on a special interpretation of quantum dynamics, e.g. treating the complex conjugate of the Schrödinger equation as a physically different equation from the Schrödinger equation itself. Thus, without a special interpretation of quantum physics, we have only TRILs in fundamental physics, according to our best physical theories. The CPT-theorem, as argued, does not change this situation.

philosophical motivation for applying the defined fundamentality from chapter II to classical cosmology would not change. In fact, this motivation is only slightly connected to the form of the Einstein equations. Thus, I think, we could say that the view in chapter II is attractive not only because it avoids some problems that arise in the 'law-fundamental' view. Instead, I think, it is *intuitive* to try to define the 'fundamentality' of time asymmetries according to physical *processes* instead of laws because processes are what we can *observe* in nature as well as (so it seems) the asymmetries of time.

Note also that it seems remarkable that, for a global dynamical equation, as we have seen, the *t*-mirrored solutions do not refer to *physically* different worlds. Consequently, we can apply the view from chapter II. Thus, it seems that condition ii) from chapter II is applicable only for global equations; otherwise, the Leibniz argument above could not be formulated. Thus, the crucial problem for every non-global dynamic equation disappears in the *global* case, where a solution describes a complete spacetime. In every other context, e.g. in the context of Maxwell's equations, we could not argue that the fully advanced and fully retarded solutions refer to the *same* physical world, because relations to a physical background are possible. Hence, we can see (even without considering cosmological models or theories) why the application of the definition in chapter II to classical cosmology was successful and why we find a fundamental time asymmetry in this field.

Chapter IV

The Arrow of Radiation

This chapter demonstrates that the arrow of radiation, in some characterisations (Frisch (2000)), is understandable as a by-product of a time asymmetric energy flux in our particular spacetime [see Castagnino, Lara and Lombardi (2003) and Castagnino and Lombardi (2009)]. Additionally, as in chapter III, I discuss the prominent accounts regarding the arrow of radiation: Rohrlich (2005) and similarly Jackson (1999), Frisch (2000), Price (1996), Price (2006) and Zeh (1999). I argue that these accounts cannot provide a satisfactory understanding of the origin of the arrow of radiation and that the understanding proposed here may be more attractive.

The chapter is organised as follows:

> Section IV.1 demonstrates that classical electrodynamics is understandable as a time symmetric theory, in contrast to Albert (2000). This section also discusses and rejects a prominent approach (see e.g. Jackson (1999) or Rohrlich (2005)) that aims to understand the origin of the arrow as a consequent of the asymmetry of causation.

> Section IV.2 focuses on a proposal from Frisch (2000) and rejects it in the original form. Nevertheless, I shall argue later on, that my own proposal can be seen as a continuation of Frisch's original proposal.

> In section IV.3 I discuss two proposals from Price that aim to understand the radiation arrow in terms of macroscopic thermodynamics (Price (1996)) and the low entropy conditions of our universe (Price (2006)).

> Section IV.4 focuses on the approach from Zeh (1999), which tries to understand the arrow as a consequent of astrophysical boundary conditions. I shall argue that this view is not convincing.

> Finally, in Section IV.5 I present my own proposal in which I demonstrate that the radiation arrow can be seen as an occurring consequence of a time asymmetric energy flux (see also Castagnino, Lara and Lombardi (2003) and Castagnino and Lombardi (2009)), provided some conditions on the energy-momentum tensor are fulfilled. However, I argue that the relation to the cosmic time asymmetries, investigated in chapter III, remain unclear. In contrast to Castagnino and Lombardi (2009), I argue that the alignment of the arrow of radiation (as well as the time asymmetry of the relativistic energy flux) with respect to the explications from

fundamental cosmic time asymmetries, cannot be understood based on the proposed understanding.
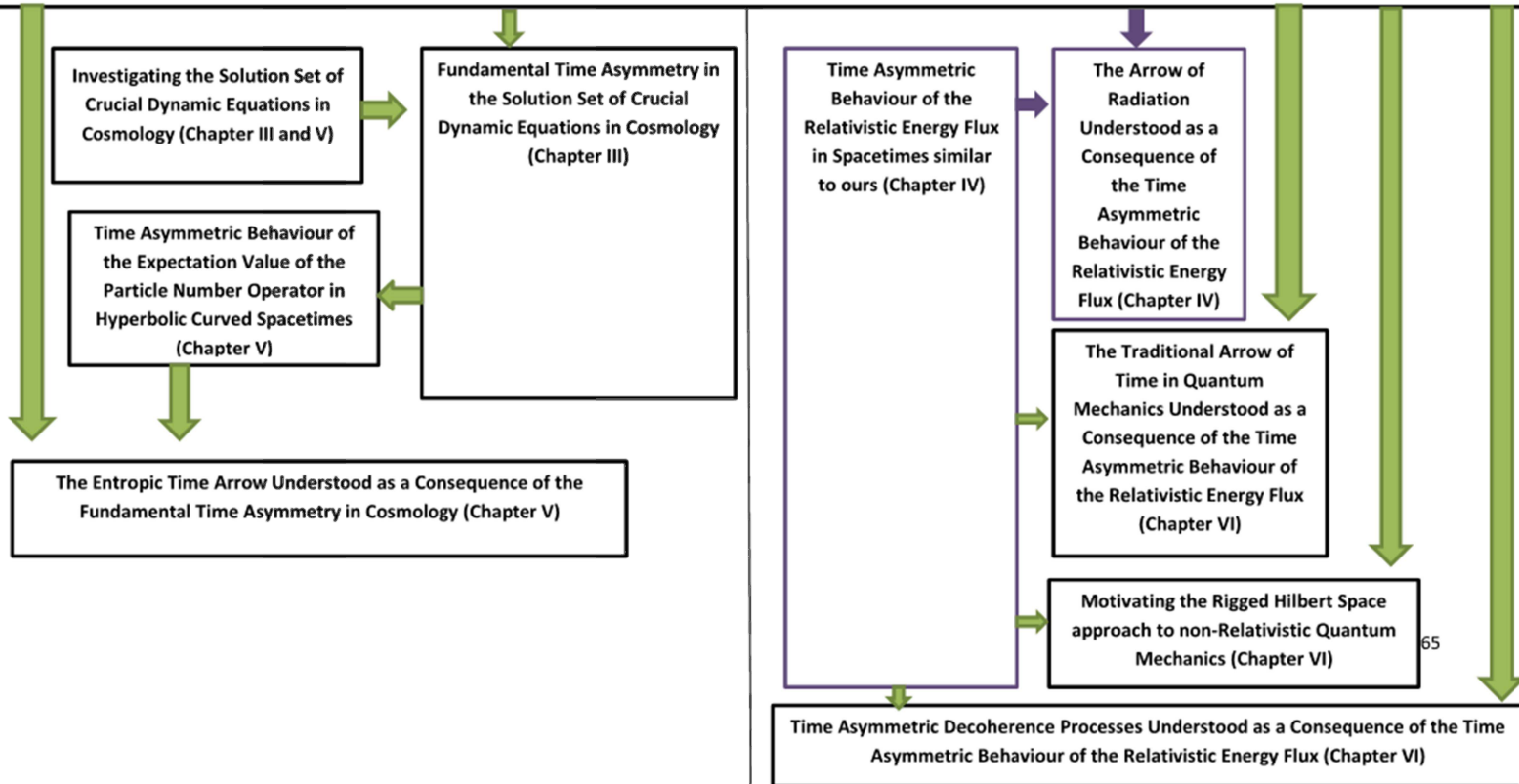
The digressions at the end of the chapter focus on interesting aspects related to the discussed accounts, but these digressions are not necessary for presenting my main critiques of the discussed approaches or for presenting my proposal regarding the origin of the arrow of radiation. Nevertheless, I think, they are particularly useful for readers interested in particular aspects of the arrow of radiation.
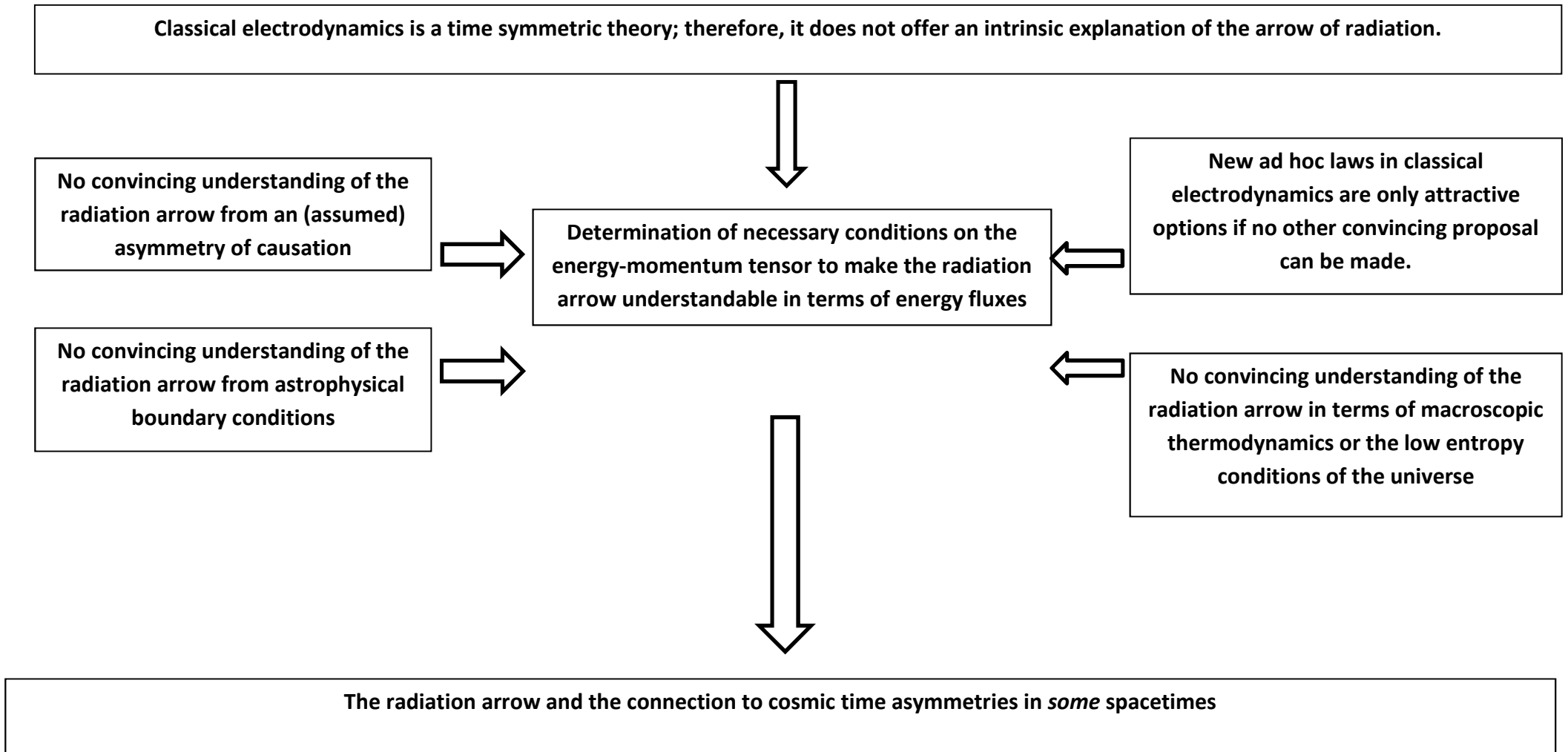
Figure 4.1 illustrates those parts of the entire analysis that are considered in this chapter. Figure 4.2 presents an overview of the structure of this chapter.

The Arrow of Radiation

Cosmic time asymmetries

Proper time asymmetries

Motivation and Definition of 'Fundamentality' in the Context of Time Asymmetries (Chapter II)

Investigating the Solution Set of Crucial Dynamic Equations in Cosmology (Chapter III and V)

Fundamental Time Asymmetry in the Solution Set of Crucial Dynamic Equations in Cosmology (Chapter III)

Time Asymmetric Behaviour of the Expectation Value of the Particle Number Operator in Hyperbolic Curved Spacetimes (Chapter V)

The Entropic Time Arrow Understood as a Consequence of the Fundamental Time Asymmetry in Cosmology (Chapter V)

Time Asymmetric Behaviour of the Relativistic Energy Flux in Spacetimes similar to ours (Chapter IV)

The Arrow of Radiation Understood as a Consequence of the Time Asymmetric Behaviour of the Relativistic Energy Flux (Chapter IV)

The Traditional Arrow of Time in Quantum Mechanics Understood as a Consequence of the Time Asymmetric Behaviour of the Relativistic Energy Flux (Chapter VI)

Motivating the Rigged Hilbert Space approach to non-Relativistic Quantum Mechanics (Chapter VI)

65

Time Asymmetric Decoherence Processes Understood as a Consequence of the Time Asymmetric Behaviour of the Relativistic Energy Flux (Chapter VI)

Classical electrodynamics is a time symmetric theory; therefore, it does not offer an intrinsic explanation of the arrow of radiation.

No convincing understanding of the radiation arrow from an (assumed) asymmetry of causation

New ad hoc laws in classical electrodynamics are only attractive options if no other convincing proposal can be made.

Determination of necessary conditions on the energy-momentum tensor to make the radiation arrow understandable in terms of energy fluxes

No convincing understanding of the radiation arrow from astrophysical boundary conditions

No convincing understanding of the radiation arrow in terms of macroscopic thermodynamics or the low entropy conditions of the universe

The radiation arrow and the connection to cosmic time asymmetries in *some* spacetimes

IV.1. Time-Reversal Invariance and the Arrow of Radiation

The first approach to the arrow of radiation that I consider seems to be the most prominent in physics. It appears in many textbooks on classical electrodynamics, such as Jackson (1999), and also in research papers: Rohrlich (2005) or, for related sometimes implicit views, Jauch and Rohrlich (1976) on *Theory of Photons and Electrons*, Rohrlich (1990) on *Classical Charged Particles,* Dirac (1938), Teitelboim (1992), Spohn (2000), Rohrlich (2002) or Landau and Lifshitz (1951) on *The Classical Theory of Fields.* In these approaches, a special notion of causality is taken for granted (sometimes implicitly). We will see that this class of approaches, which tries to explain the arrow of radiation using a causality principle, seems implausible if we cannot show that causality is given in only one time direction. Price (1996) showed explicitly that the notion of backwards causality is not prima facie implausible; thus, it is not reasonable to deduce a physical time arrow from a time directed notion of causality if we cannot show that causality should have a time direction independent of an physical arrow of time and *additionally* that physical processes are led by this conception of causality. Thus, I shall begin this analysis with the most popular view (as far as I know) of the arrow of radiation in physics [see Jackson (1999) or Rohrlich (2005)].

But firstly, we must note that Maxwell's theory is *really* time symmetric in all crucial properties. Thus, in classical electrodynamics, the time asymmetry of radiation is not grounded on the established laws of classical electrodynamics. I will demonstrate that in more detail in the following subsection because occasionally authors argue that the theory of classical electrodynamics includes some crucial time asymmetric aspects (see for example Albert (2000)). However, before I come to that demonstration, I sketch the popular textbook view of this problem in more detail.

One known solution of Maxwell's equations is the fully advanced solution, but this radiation, according to the textbook view, seems to contradict a time directed notion of causality. In this approach, the time directed notion of causality is, I believe [see Rohrlich (2005)] understood as a law-like fact of nature (or physics), and according to this fact, so Rohrlich argues, fully advanced radiation is forbidden. All other known solutions of Maxwell's equations, which could include free fields, superposition of advanced fields with free fields and also retarded fields, are consistent not only with Maxwell's equations but also with the assumed time direction of causality. Thus, all these solutions could be seen as the set of solutions that could be used to describe empirical observations of electromagnetic radiation. Hence, there is an arrow of time in classical radiation, given by the time direction of causality. For example Rohrlich:

'Finally, I must return to the issue of radiation dissipation. The typical emission of radiation by moving charges is dissipative in the sense that the energy and momentum of radiation absorbed from other sources is negligibly small compared to the radiation emitted. Only two alternatives could weaken that dissipation: reabsorption of the emitted radiation, or arrival and absorption of advanced radiation. The latter is excluded because it would have to come from sources in the future going in the negative time direction and arriving at the particle on a future light cone. This violates causality. The reabsorption of the charge's own radiation is possible by suitable reflection but is always smaller than the emitted radiation. The limiting case in which the emitted radiation is fully reflected back to the moving charges and is fully reabsorbed is the case of hyperbolic motion. A perfectly reflecting cylinder concentric to a uniformly accelerated particle beam reflects the emitted radiation back onto the particles resulting in complete reabsorption. With that exception, radiation emission is dissipative. It follows that there exists an arrow of time of electromagnetic radiation.' [Rohrlich (2005), p. 3]

In the next subsection, I argue in more detail that this view is not a convincing option for understanding the origin of the radiation arrow. Additionally, the next subsection also claims to demonstrate that the *theory* of classical electrodynamics is time symmetric in all properties. This, I think, must be shown rigorously if we are to demonstrate that proposals, such as from Albert (2000), which claim that classical electrodynamics is intrinsically time asymmetric and hence could provide a radiation arrow by its own, are not fruitful. I hope this demonstration will explicate the puzzle of the radiation arrow in a more vivid form.

IV.1.1. The Arrow of Radiation, Time-Directed Causation and the Theoretical Symmetry of Classical Electrodynamics

In this subsection, I first demonstrate why classical electrodynamics is a time symmetric theory in all crucial properties. To do so, I show that this conclusion follows from the time-reversal invariance of Maxwell's equations and the time symmetry of the equation of motion for classical sources (see also Spohn (2000)). Thus, the arrow of radiation is not understandable in terms of classical electrodynamics itself. Hence, we have to seek other explanations. At the end of this section, I show that the standard account in physics textbooks, which aims to provide such an explanation, seems to be a misleading description of the origin of the arrow of radiation.

So firstly, I demonstrate that classical electrodynamics is a time symmetric theory in all relevant properties. Here it is necessary to consider the structure of classical electrodynamics in more detail. For this task, the formal operator that mirrors the time coordinate will be labelled T and works on a doublet (t, x) (where t is the time coordinate and x is the space coordinate) as follows:

$$T : (t, x) \rightarrow (-t, x) .$$ (4.1)

Special relativity, which is treated as the fundamental theory in classical electrodynamics, defines the four-velocity of a point charge as $v^{\mu}(\tau) = \dfrac{d}{d\tau} z^{\mu}(\tau)$. Here, $z(\tau)$ is the world line of a classical point charge. The proper time τ of the charge is given by $d\tau = \dfrac{dz^0}{\gamma}$ with $\gamma = \dfrac{1}{\sqrt{1-v^2}}$, where v is the three-velocity of the point charge; thus, the four-velocity is given by $v^{\mu} = \left(\gamma(\tau), \gamma(\tau) v(\tau) \right)$. Now, v^{μ} and v can be transformed under the transformation T . For the three-velocity, we get $v(\tau) \xrightarrow{T} -v(\tau')$, where τ' is the transformed proper time of the charge. For γ, we get $\gamma(\tau) \xrightarrow{T} \dfrac{d}{d\tau'} z^{\prime 0}(\tau') = \dfrac{d}{d\tau} z^0(\tau) = \gamma(\tau) \geq 1$. Hence, under the transformation T , the four-velocity clearly transforms to $v^{\prime\mu}(\tau') = \left(\gamma(\tau'), -\gamma(\tau') v(\tau') \right)$. Thus, if we use the metric $(-1, 1, 1, 1)$, we can conclude that for each spacetime point x^{μ} in Minkowski space M and for each four-velocity $v^{\mu}(\tau)$, the following transformations are crucial: $M \xrightarrow{T} M'$, $x^{\mu} \xrightarrow{T} x_{\mu}$ and $v^{\mu}(\tau) \xrightarrow{T} -v_{\mu}(\tau')$. Thus, for each electric charge, the mirroring of proper time τ leads to an inversion of the direction of the charge motion. Also, in special relativity, the motion of charges leads to a four-current $j^{\mu}(x, t)$, which is given by

$$j^{\mu}(x, t) = q \int d\tau \delta_4 (x - z) v^{\mu}(\tau) .$$ (4.2)

For simplicity, Gaussian units are used here. Therefore, the speed of light in vacuum is set to 1, and the function $\delta_4(x - z)$ is the four-dimensional delta function. The four-current transforms like the four-velocity [see (4.2)] under the transformation T . Thus, not surprisingly, the direction of the electric flux is also mirrored by the transformation T .

Now, I will use the electromagnetic four-potential $A^\mu(\phi, \vec{A})$ to describe the time-mirrored fields $\vec{E}, \vec{B}$, which could be seen as basic entities in the theory. Moreover, I use the identity $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ and the Lorenz gauge for all following calculations.[20] In the Lorenz gauge, the four-potential satisfies the equation $\partial_\mu A^\mu = 0$. Therefore, Maxwell's equation can be written as $\Box A^\mu = -j^\mu$ (where $\Box$ is the d' Alembert operator). Additionally, the transformation of the four-potential A^μ is given by the behaviour of the four-current; thus, we get:

$$A^\mu(x) \rightarrow -A_\mu(x').$$
(4.3)

Therefore, Maxwell's equation ($\Box A^\mu = -j^\mu$) is invariant under the transformation T. Thus, one fundamental law of classical electrodynamics seems to be time-reversal invariant. However, to show that the entire theory is time symmetric, it is necessary to consider other physically relevant quantities of classical electrodynamics.

Consider a moving point charge q that produces an electromagnetic field and is assumed to be located at the spacetime point Q, which is time dependent and given by the world line $z(\tau)$. In addition, consider another spacetime point P located at x. The vector from Q to P is given by $x - z(\tau)$. The magnitude of the four-vector is given by $\rho(x, \tau) = U^\mu(\tau)\big(x - z(\tau)\big)_\mu$, where $U^\mu(\tau)$ denotes the unit vectors from Minkowski space.

Additionally, the retarded potential A^μ_{ret} at the spacetime point x^μ is given by

$$A^\mu_{ret}(x, \tau) = q \frac{v^\mu(\tau)}{\rho(x, \tau)}.$$
(4.4)

We have already seen that under the transformation T, the future semi-light cone of a charge on a spacetime point is transformed to the past semi-light cone. This is because the direction of the four-velocity is mirrored under the transformation T. Hence, T is considered to represent the mirroring of proper times and thus the direction of each movement. However, under T, a retarded potential is still retarded with respect to the transformed proper time τ' [see (4.4)]. This is because, although the direction of the four-

---

[20] The argument is not gauge dependent. Gauge fixing at this point only simplifies the calculation. This is possible because, in classical electrodynamics, the fields $\vec{E}, \vec{B}$ (not the potential) are treated as physical entities.

velocity is mirrored under T, the distance $\rho(x,\tau)$ is invariant. The four-potential transforms like the four-velocity [see (4.4)]. Therefore, according to the new proper time τ', which describes the proper time of the moving charge under T, we find that a former potential that is retarded with respect to τ remains retarded under T with respect to τ'. This shows that the time direction of emitted electromagnetic waves is not changed by T because the time direction of the four-velocity is also mirrored.

Nevertheless, to confirm that classical electrodynamics is a time symmetric theory, a brief mathematical consideration of the behaviour of the quantities in classical electrodynamics must be presented. The behaviour of $F_{ret}^{\mu\nu}$ under T is anti-symmetric (because $F_{ret}^{\mu\nu}$ is given by the four-potential, as sketched above):

$$F_{ret}^{\mu\nu}(x) \xrightarrow{\ T\ } -F_{\mu\nu}^{ret}(x') \ . \tag{4.5}$$

However, the equation of motion for a point charge is not given by the relativistic formulation of the Newtonian equation of motion. In the Newtonian formulation, we would obtain the equation $m\dfrac{dv^{\mu}}{d\tau} = G^{\mu}$, where G^{μ} describes an external force. However, in the traditional formulation of the equation of motion for a point charge [see, for example, Dirac (1938)], we obtain

$$m_{bar}\frac{dv^{\mu}}{d\tau} = G^{\mu} + F_{ret}^{\mu}, \tag{4.6}$$

where $F_{ret}^{\mu} = qF_{ret}^{\mu\alpha}v_{\alpha}$ is used. $F_{ret}^{\mu\nu}$ can be separated into a trivial linear combination $F_{ret}^{\mu\nu} = \dfrac{1}{2}\left(F_{ret}^{\mu\nu} + F_{adv}^{\mu\nu}\right) + \dfrac{1}{2}\left(F_{ret}^{\mu\nu} - F_{adv}^{\mu\nu}\right) = F_{+}^{\mu\nu} + F_{-}^{\mu\nu}$. The term $F_{+}^{\mu\nu}$ represents the corrections to the internal field, and $F_{-}^{\mu\nu}$ represents the self-interaction field. Thus, the self-interaction force Γ^{μ} is given by $F_{-}^{\mu\nu}$, where $F_{+}^{\mu\nu}$ represents a force that is analogous to a Newtonian force:

$$qF_{+}^{\mu\alpha}v_{\alpha} = -m_{elm}\frac{dv^{\mu}}{d\tau} \tag{4.7}$$

and

$$qF_-^{\mu\alpha}v_\alpha = \Gamma^\mu \,. \tag{4.8}$$

Additionally, for a charge with mass m, the mass is given by $m = m_{bar} + m_{elm}$. This allows equations (4.7) and (4.8) to be combined to:

$$m\frac{dv^\mu}{d\tau} = G^\mu + \Gamma^\mu \,. \tag{4.9}$$

Now it appears to be useful to describe the self-interaction force Γ^μ in terms of the dependence of its first and second derivatives on the four-velocity with respect to the proper time of the charge. Dirac did this (1938), and the expression he obtained for the self-interaction force Γ^μ is *not* time-reversal invariant. In the philosophical discussion of the arrow of radiation, a time variant equation (4.9) would allow the definition of a fundamental arrow of radiation. However, such an approach is misleading. To show this fact more precisely, it is necessary to consider some aspects of QED *or* the classical results from Spohn (2000), who also showed that the classical description of the self-interaction force Γ^μ differs from Dirac's description [see Dirac (1938)]. From Spohn (2000), we get

$$\Gamma^\mu(\tau) = m\tau_0(\eta^{\mu\nu} + v^\mu v^\nu)\frac{dG_\nu}{d\tau} \,, \tag{4.10}$$

where $\eta^{\mu\nu}$ is the metrical tensor, and $\tau_0 = \frac{2}{3}\frac{q^2}{m}$. Note that (4.10) is necessary in classical electrodynamics to ensure that the part of the energy-momentum emission, which is not covered by the external Newtonian force, is captured. Thus, (4.10) is motivated by the empirical adequacy of classical electrodynamics. This can be shown in a formal context by considering the relation

$$\tau_0 v^\mu v^\alpha \frac{dG_\alpha}{d\tau} = m\tau_0 v^\mu v^\alpha \frac{d^2v_\alpha}{d\tau^2} = -m\tau_0 v^\mu \frac{dv^\alpha}{d\tau}\frac{dv_\alpha}{d\tau} \,. \tag{4.11}$$

This expression describes the classical rates of energy and momentum emitted by an excited charge (a relativistic form of the Larmor formula). Now, with the corrected self-interaction force and the result that, under time reversal, retarded fields remain retarded and advanced ones remain advanced, the equation of motion (4.9) is time-reversal invariant (see also Rohrlich (2005)).

However, to demonstrate the time symmetry of (4.9) (and hence of the whole theory) it is still necessary to analyse the external forces G^μ. Therefore, consider an external electromagnetic force given by $G^\mu = qF_{ext}^{\mu\alpha}v_\alpha$. Under T, $G^\mu(\tau)$ transforms to $G_\mu(\tau')$. Again, τ' is the transformed proper time of a point charge. We can now use the equation of motion (4.9), together with equations (4.7) and (4.10), to determine the behaviour of the equation of motion under T:

$$m\frac{d^2x^\mu}{d\tau^2} = G^\mu + \Gamma^\mu \xrightarrow{\ T\ } m\frac{d^2}{d\tau'^2}x'^\mu(\tau') = G'^\mu(\tau') + \Gamma'^\mu(\tau'),$$
(4.12)

and of course

$$m\frac{d^2}{d\tau'^2}x_\mu(\tau') = G_\mu(\tau') + \Gamma_\mu(\tau').$$
(4.13)

Now, I think, we see the time-reversal invariant as clear as possible. Thus, the theory of classical electrodynamics *is* time-reversal invariant in all crucial quantities, and the arrow of radiation, *therefore*, cannot be deduced from classical electrodynamics itself.[21]

However, on this point it seems necessary to consider different characterisations of the arrow of radiation in order to make clear what exactly has to be understood. Up to now, we have seen that classical electrodynamics includes no intrinsic time asymmetry, so the precise characterisation of the arrow was not needed. But, the discussion of proposals to solve the puzzle clearly requires a precise characterisation of the radiation arrow. Traditionally, the arrow of radiation is characterised by the fact that fully advanced radiation, as an empirical fact, does not (or only very rarely) occur in the observable universe.

However, what is crucial, I think, is that not that all types of fully advanced radiation need to be ruled out by a suitable characterisation. Instead, I shall argue that only the non-occurrence of a special type of fully advanced radiation is needed to obtain a suitable characterisation of the arrow of radiation. To demonstrate that, consider the following types of fully advanced radiation:

---

[21] Note that an analysis analogous to the cosmological case from chapter III (for the nonlinear Maxwell–Lorentz equations, which would require numerical calculations) would not be fruitful. This can be seen easily at one point: In the case of classical electrodynamics we do not deal with global solutions. Hence, condition ii) of definition I (chapter II) cannot be fulfilled (see also chapter III).

a) Source-free fields coming from '$\pm$ infinity'. These can be combined in a way that makes the fully advanced description applicable to the phenomenon.

b) Fully retarded emitters can be arranged in a special geometry such that the combined field becomes describable, in a special region, as fully advanced.

c) An accelerated charge is supposed to radiate but the associated radiation field could be fully advanced *or* fully retarded. In nature only (or almost) the fully retarded solution to Maxwell's equation seems to occur.

I shall argue that only the non-occurrence of fully advanced radiation from type c) provides a suitable characterisation of the arrow of radiation.

Consider fully advanced radiation of type a). Experimentally, we detect quasi source-free radiation from the microwave background, but no radiation has been observed coming from the cosmic future. This could be interpreted as a time asymmetry or not (see for example Price (2006)). Nonetheless, given the cosmological models and the observationally well-established assumption that our particular cosmic domain (or the universe) accelerates its expansion, the non-occurrence of microwave radiation from the cosmic future seems not too surprising. Even if the lack of this radiation were interpreted as an observable time asymmetry, it seems to be a time asymmetry that is grounded on cosmological boundary conditions of our particular observable universe (or cosmic domain). This time asymmetry will therefore not be understood as the arrow of radiation in this investigation, because the non-occurrence of other kinds of fully advanced radiation, as I shall argue, provide a different and more plausible characterisation of a time asymmetry of electromagnetic phenomena. This time asymmetry (the non-occurrence of fully advanced radiation from type a), instead, can be understood as a consequence of the accelerated expansion of our particular cosmic domain or universe (which seems not *fundamentally* given, see chapter III). Moreover, in electromagnetic shielded regions, the asymmetry (with respect to fully advanced radiation from type a) is non-existent. Nevertheless, the arrow of radiation seems also crucial for such regions.

Now consider fully advanced radiation of type b). Radiation of type b) is not a special phenomenon of electromagnetic waves: all types of classical waves show this type of behaviour. A special geometry of fully retarded emitters can provide a wave field that converges coherently (in a special region) and can be described as a fully advanced wave field (in that region). But the *total* wave field of the retarded emitters is not appropriately described as fully advanced. This possibility of special emitter geometries is not connected to a suitable characterisation of the arrow of radiation. The special emitting geometries are built of fully retarded emitters. Thus, the fundamental emitters are fully retarded and are merely arranged so as to produce a radiation field that can be described (in a special region)

as fully advanced. Hence, the arrangement of the emitters is crucial and in nature the absence of such arrangements is well understood (even if not in *fundamental* terms) by thermodynamic considerations (see for example Popper (1956) or Price (2006) or the digressions at the end of this chapter).

But there seems to be another possible kind of fully advanced radiation, which is not observed in nature. The absence of this radiation, type c, seems to provide a basic time asymmetry in classical electrodynamics (not obviously provided from thermodynamic considerations or boundaries) and hence should be used to characterise the term 'arrow of radiation'. This, I think, avoids the confusion that could occur if fully advanced radiation from type a or b is taken into account. Thus, in this investigation, the term 'arrow of radiation' will be understood as the fact that radiating accelerated charges seem not (or only rarely) to be associated with fully advanced radiation, even in electromagnetically shielded regions.

Now, coming back to the mentioned textbook account in physics [Jackson (1999); Rohrlich (2005)], the absence of fully advanced radiation in nature is explained by the following special notion of time-directed causation. Rohrlich (2005) puts it this way:

> *'The latter [fully advanced radiation] is excluded because it would have to come from sources in the future going in the negative time direction and arriving at the particle on a future light cone. This violates causality.'* [Rohrlich (2005), p. 3]

Thus, Rohrlich makes an unjustified assumption, which is: There is a time arrow of causation and this arrow guides physical phenomena. However, as shown by Price (1996), the notion of time-directed causality is *not* necessary; Therefore, this popular approach appears to be implausible if not circular.[22] Moreover, electrodynamics, as a theory, gives no reason to assume the applicability of a time-directed causation to electromagnetic phenomena in such a way that fully advanced radiation, an electromagnetic process, is forbidden by this time-directed causation. Thus, even if a time-directed notion of causality were plausible in some context, it is not clear why electromagnetic processes should be guided by this understanding of causality. Consequently, this standard account of the origin of the arrow of radiation seems not fruitful at all.

Moreover, some other prominent approaches to the radiation arrow depending on the assumed concept of causation as well [Frisch (2002) together with Frisch (2004), Frisch (2005), (2006), (2007) and (2009)]. Even if the approach to the causation arrow taken by Frisch is quite different from the one discussed above, the dependence of the radiation arrow on the conception of causation stays the same. Frisch argues, given his approach to the conception of physical theories, that causality is always a primitiv concept in physical

---

[22] The account would become circular if the time direction of causality itself were explained by the arrow of radiation.

theories. Nevertheless, I shall argue that the presupposing of a time asymmetric "causality arrow" is not necessary to understand the origin of the radiation arrow. Therefore, I will demonstrate that my own proposal does not need such presuppositions and seems, therefor, more attractive than those approaches.

Nevertheless, I shall first refer to other formulated approaches in order to show that these prominent accounts, which do not refer to causality, also cannot describe the origin of the time asymmetry in classical electrodynamics.

IV.2. On the Retardation Condition

I begin by discussing the approach taken by Frisch (2000) to determine the origin of the arrow of radiation. Frisch argues that in nature, fully advanced radiation cannot be observed because it is contrary to an existing law in classical electrodynamics. This law, suggested by Frisch (2000), is called the retardation condition and indicates that the electromagnetic radiation associated with moving charges cannot be fully advanced.

First, it is obvious that Frisch's approach can describe a possible origin of the arrow of radiation in a fundamental way. In principle, however, it is possible to formulate a law for all observations in nature. The question, therefore, concerns the plausibility of assuming an additional law in classical electrodynamics, namely the retardation condition. Frisch (2000) argued that all observations in the area of applications of classical electrodynamics fulfil the retardation condition. He further argued that the same argument is used to show the plausibility of Maxwell's equations in classical electrodynamics. Therefore, the empirical data and observations would support Maxwell's equations as well as the retardation condition.

However, I consider this argument as misleading. I agree with Frisch (2000) that, while the retardation condition does appear to be ad hoc, this is not sufficient reason to reject a proposed law. Moreover, the fact that physical problems are easily 'solved' by proposing new ad hoc laws is not sufficient to reject Frisch's account.

But, according to physics, classical electrodynamics can be seen as an approximation of QED. QED, as far as we know, is the most fundamental theory of pure electromagnetism[23] and the fundamental laws of electrodynamics are those of QED. The classical limit of QED yields

---

[23] Electromagnetism alone, electro-weak theories are not considered here, but they would not change the argument.

approximations consistent with the classical descriptions. Therefore, an analysis of the laws of QED in the classical limit supports Maxwell's equations, but the retardation condition is not obtained. Thus, as classical laws, Maxwell's equations are supported by QED, but the retardation condition is not.

With respect to Frisch's approach (2000), it may be thought that the retardation condition should be assumed not only in classical electrodynamics but also in QED. In this case, the retardation condition would appear in the classical limit of QED. However, this proposition appears to be less plausible than the assumption of the retardation condition in classical electrodynamics. It seems to me that QED lacks a sufficient empirical basis for favouring the retardation condition [see, for example, Gross (2004)].

Moreover and disconnected from the assumption of the retardation condition in QED, a plausible explanation of the arrow of radiation that is formulated *without* proposing new laws in QED or classical electrodynamics would be, I think, more attractive than that of Frisch. I will try to present such an account later in this chapter. Note, however, that the suggestions of Frisch (2000) are not prima face implausible, even considering the thoughts I present here. In fact, my suggestion can also be seen as a possible explanation of Frisch's retardation condition. I will argue that the non-occurrence of fully advanced radiation (of type c)), i.e. the retardation condition, can be understood as a result of the energy flux in some time asymmetric spacetimes. Given the following quote, my suggestions, I think, can be seen as a proposed improvement of the account of Frisch (2000):

> *'According to the textbook[24] account, electromagnetic fields have to meet an additional general constraint beside those imposed by the Maxwell equations. The account says that not all solutions to the Maxwell equations but only those that satisfy the retardation condition, according to which electromagnetic fields associated with a charge Q propagate along the future light cone of Q, can represent physically possible situations. In physics textbooks the retardation condition is sometimes presented as a causal constraint [...]. So one might understand by 'the textbook account' an account that justifies or explains the retardation condition by appealing to a principle of causality and a temporal asymmetry that is supposed to be implied by that principle. This is not, however, the kind of account I want to advocate here. Rather, the account I wish to advocate simply stipulates that, in addition to the Maxwell equation, electromagnetic fields associated with electric charges satisfy the retardation condition without offering any explanation as to why this condition should hold. [...] Of course, anyone who thinks that the puzzle of the asymmetry of radiation presents a genuine puzzle will not be satisfied with an 'account' that offers*

---

[24] Physics textbooks, such as Jackson (1999)

little more than the statement that the asymmetry does in fact hold.' [Frisch (2000), p. 25]

On the basis of that statement, I claim to provide an explanation for the retardation condition.

But firstly, though, I shall make note of an argument that is presented frequently in discussions of the arrow of radiation, mostly, I think, in the physics literature. This argument stipulates that any view that suggests that the arrow of radiation can change its direction (in different spacetime regions) is ruled out for empirical reasons [see, for example, Zeh (1999) or Gell-Mann and Hartle (1994)]. This argument is grounded on the fact that the night sky appears to be mostly black. Price (1996), in contrast, argues for the possibility that the arrow of radiation, in a closed spacetime, could change its direction and so produces a time symmetric entire spacetime. Now, advocates of the 'Night Sky argument' try to show that, if there were a spacetime region in the cosmic future of our epoch in which the arrow of radiation has changed its direction, we should see light coming from that region. Thus, they try to conclude that the arrow of radiation, at least in our particular spacetime, cannot change its direction. Readers who believe that argument would likely assume that the time symmetric view from Price, which in fact predicts that the arrow of radiation would change direction in a closed spacetime, is ruled out by empirical data, at least for our particular spacetime.

I think most of the modern literature acknowledges that this argument is invalid. Price himself (1996) and also (2006) shows in detail that his time symmetric view is unaffected by this argument. Thus, if some readers of the next sections feel that I ignore this traditional argument so as to show that Price's suggestions seem misleading, I refer to Price (1996) or (2006) to show that this traditional argument is irrelevant to the debate.[25] Nevertheless, in Digression IV.A, I also give a brief discussion of the main strategy used to defend a time symmetric view in which the arrow of radiation could change its direction. I treat this argument only briefly because it is irrelevant to the main discussion in this chapter; the modern literature provides more than sufficient strategies to defend views that include a changing electromagnetic time direction against this argument [particularly Price himself (1996) and (2006)]. Nevertheless, for the sake of completeness, a brief discussion of this strategy can be found at the end of this chapter.

---

[25] Not only because the empirical data are obviously not sufficient to rule out a change of the arrows direction in *some possible* closed spacetime.

IV.3. On the Time Symmetric View from Price (1996) and (2006)

IV.3.1. On Price (1996)

Price (1996) characterizes the arrow of radiation (in contrast to my characterisation) as a purely macrophysics phenomena, provided by classical thermodynamics. Precisely, according to Price (1996), the radiation arrow is characterised by a statement which stipulates that: In macrophysics, all emitted electromagnetic waves are coherent, and all absorbed electromagnetic waves are incoherent (statement A hereafter).

However, in addition to A, Price (1996) also describes the time asymmetry in classical electrodynamics by the statement

> *'Maxwell's theory clearly permits both kinds of solutions [i.e. retarded and advanced solutions] but nature appears to choose only one. In nature it seems that radiation is always retarded rather than advanced. Why should that be so?'* [Price (1996), p. 50]

This quotation is not equivalent to A. The following slight modification makes it consistent with the suggested characterisation of the time asymmetry in classical electrodynamics:

> *'Maxwell's theory clearly permits both kinds of solutions [i.e. retarded and advanced solutions] but nature appears to choose only one. In nature it seems that radiation is always retarded rather than [fully] advanced. Why should that be so?'*[My modification of Price (1996), p. 50]

However, the modified quotation and the original version are both not equivalent to A. Moreover, Price does not refer to this characterisation in the discourse following the quoted text. Instead, Price (1996) again refers to A as a characterisation of the time asymmetry in classical electrodynamics.

My aim here is to demonstrate the implausibility of Price's view. The ability to explain A in macrophysics is not sufficient to confirm that this is also the origin of the arrow of radiation (for example in microphysics). In fact, Price (1996) does not describe the arrow of radiation in microphysics. Instead, he suggests that classical microphysics is time symmetric and attempts to show that by using a new interpretation of the absorber theory of Wheeler and Feynman (1945). Price (1991a), (1991b), (1994) and (1996) argues that his interpretation shows that classical microphysics is time symmetric. This would imply that the origin of the arrow of radiation can be found only in macrophysics., which would be extremely important. However, Frisch (2000) has already shown that

i) The 'interpretation' put forward by Price is not a reinterpretation of Wheeler and Feynman (1945), but a new physical absorber theory.

ii) The absorber theory in Price (1996) contradicts either the empirical data or Maxwell's equations.

iii) Thus, Price's absorber theory does not show that classical microphysics is time symmetric.

Therefore, I will not present the interesting arguments from Frisch (2000) in much detail here (there are presented in detail in Digression IV.B together with additional arguments against the suggestions from Price (1996)) because the arguments above are sufficient to support the following conclusion: Price cannot show that his suggestion for the origin of A in macrophysics contributes to the understanding of the origin of the arrow of radiation (for example in microphysics). Therefore, Price's suggestion, which focuses on macrophysics, is not a plausible option for describing the origin of the arrow of radiation in his suitable characterisation (section IV.1). Nevertheless, for the sake of completeness, I discuss the arguments of Frisch (2000) against Price's suggested reinterpretation of the absorber theory and I will add some thoughts in Digression IV.B in order to make Frisch's critique even more precise. Moreover, I will show in more detail why the consideration of classical thermodynamics appears not helpful in the context of the radiation arrow. Nevertheless, for the purposes of the main chapter it is sufficient to conclude that Price (1996) focuses only on macrophysics and his statement A. Hence, his attempts cannot describe the origin of the radiation arrow characterised by the absence of fully advanced radiation from type c.

I will now discuss the slightly different proposal from Price (2006), which differs from his considerations in (1996).

IV.3.2. On the Arrow of Radiation as a By-Product of Classical Thermodynamics; Price (2006)

The claim of Price (2006), I think, can be outlined as follows. For wave phenomena in classical wave media, for example water waves, the low-entropy past of our universe can be reasonably assumed to play an important role. In addition, Popper (1956) argued in a similar way (although his view is not identical to Price's). Both Price (2006) and Popper (1956) argued that the arrow of time in classical thermodynamics, which requires the assumption of a low-entropy past [see also Albert (2000)], is responsible for the time asymmetry of wave

phenomena in classical wave media. The connection between wave mechanics and thermodynamics appears because the wave equations are statistical approximations of the motions of each constituent of the wave media. In a classical wave medium, a fully advanced wave would be contrary to the arrow of time in classical thermodynamics because it would require the correlated motion of many constituents at the moment the wave 'starts'.

According to Price (2006), an analogy can be established for the case of electromagnetic radiation. He puts it as follows:

> 'Then the lesson of the argument is that the observed asymmetry of radiation depends on an asymmetry in the environment in which wave media are embedded: the asymmetry is that the environment supplies large ''kicks'' in one time-sense but not in the other. In our ordinary time-sense, it adds large amounts of energy to the media (''all in one go''—in a coherent way, in other words) much more frequently than it subtracts or removes large amounts of energy, in a similar coherent fashion.' [Price (2006), pp. 21, 22]

Note, however, that in the relativistic formulation of classical electrodynamics, there is no wave medium, no ether, and therefore no constituents of the wave medium. There is no way of adding statistical mechanics to a single classical electromagnetic wave because the wave is not built by the correlated motion of the constituents of a wave medium. In Price (2006), a proposal is formulated; this proposal stipulates to add the statistical considerations of the wave medium to QED. Hence, the statistical behaviour of the entropy of a wave medium is said to be analogous to that of the entropy of a quantum field in QED.

Here I present two points to demonstrate that the entropy behaviour of electromagnetic radiation is not analogous to that of a wave medium. Firstly, as mentioned, according to classical electrodynamics, Price's arguments are invalid because, in classical relativistic electrodynamics, electromagnetic waves occur not only in a wave medium. So secondly, QED itself is a quantum field theory. Now, QFT can be seen as analogous to classical field theories in some aspects. However, in classical and quantum field theories, the fundamental entities are the fields. To determine their entropy behaviour in the classical case, we refer to statistical descriptions of the motions of the constituents of the wave medium. But, the fundamental entities in QED are quantum fields. Field excitations of a photonic field can be seen as photons, but electromagnetic waves in the classical QED limit are *not* described as the correlated motions of constituents in a photonic wave medium. In the QED case, we cannot simply refer to a deeper physical structure, which is composed of constituents of the quantum field.

Therefore, the analogy between wave phenomena in various media and electromagnetic radiation, as Price (2006) suggests, is misleading.

Of course, this may raise the following question: If the time asymmetric behaviour of wave phenomena in classical wave media is not analogous to the time asymmetric behaviour of electromagnetic fields, why is the time asymmetry of wave phenomena in classical wave media, which is surely important, not considered in its own chapter in this investigation? The answer is that, in the case of electromagnetic fields, as argued so far, no formulated account seem successful and plausible to me. Consequently, this seems to be an outstanding question in the philosophy of physics. In contrast, given the time asymmetric behaviour of wave phenomena in classical wave media, we have well-established accounts that explain this asymmetry using the arrow of time in thermodynamics and the second law. This arrow of time is considered in chapter V. Moreover, the question of how the time asymmetry of wave phenomena in classical wave media can be understood by using the second law of thermodynamics is investigated by Popper (1956) and also by Price (2006) in similar but still distinctive ways. Thus, I will not reconsider the entire investigation here. For the sake of completeness, however, and because the time asymmetry of wave phenomena in classical wave media is surely an important aspect of our everyday experience, I give in Digression IV.C a short overview of the connection between the arrow of time in thermodynamics, which is considered in detail in chapter V, and the time asymmetric behaviour of wave phenomena in classical wave media. There the analysis from Price (2006) is discussed in more detail (but restricted to time asymmetries in classical wave media).

However, in the context of the main chapter, the proposal from Price (2006) seems misleading, if it is understood as an attempt to describe the origin of the arrow or radiation.

IV.4. Boundary Conditions for the Arrow of Radiation

In this section, I focus on the approach of Zeh (1989) and (1999) in which he attempts to deduce the arrow of radiation from astrophysical boundary conditions which are related to the Sommerfeld condition.[26] To show that Zeh's approach does not provide a satisfactory understanding of the origin of the radiation arrow, I first describe the account in more detail.

It is well known that mathematical boundary conditions can be formulated to produce a situation where the *total* radiation field in a finite spacetime region cannot be fully advanced

---

[26] The Sommerfeld condition is: '*There are no incoming fields in a considered spacetime region*'. Formally, this leads to the situation that the general description of a wave field (including retarded, advanced, and free fields: $F = F_{ret} + F_{in} = F_{adv} + F_{out}$) becomes possible as fully retarded (because $F_{in} = 0$).

but fully retarded. This condition is called the Sommerfeld condition.[27] However, Sommerfeld has shown only the type of mathematical condition that can create a situation where no *total* wave field can be described as fully advanced. Zeh tries to show that such a condition is also physically plausible because of astrophysical considerations. He formulates a boundary condition that affects a region of spacetime in the cosmic past[28] of our observable spacetime region. The condition he formulates is that: 'this spacetime region is a perfect absorber'. This, in Zeh's approach, is the astrophysical assumption that should realize the Sommerfeld condition in nature. In such a case, Zeh argues that a spacetime region that is bounded by a perfect absorber in the past cannot include fully advanced *total* radiation fields.

Firstly, I mention an argument formulated by Frisch (2000) that shows that Zeh's approach depends on the arrow of time in thermodynamics. Secondly, I present other arguments that demonstrate that Zeh's suggestion is unattractive for describing the origin of the arrow of radiation in general.

Frisch (2000) argues that, in some ways, Zeh's approach seems to be circular. This can be shown as follows. Consider a moving charge in the spacetime region beyond (in the cosmic past) the perfect absorber. Now, consider that the electromagnetic field associated with the moving charge is fully advanced. In this case, the field in the region of the perfect absorber is not equal to zero. Therefore, the Sommerfeld condition is not fulfilled, and it is possible that fully advanced total radiation fields may exist in our particular spacetime region. Hence, Zeh must assume that the electromagnetic fields in the spacetime region beyond (in the cosmic past) the perfect absorber cannot be fully advanced in order to conclude that the Sommerfeld condition is fulfilled. Hence, Zeh can conclude that there are no fully advanced total radiation fields in our spacetime region only if he assumes that no fully advanced total radiation fields occurred in the past.

However, there is an objection to this argument. Zeh could argue that the assumption of a fully advanced field in the spacetime region in the past is implausible because if such a field exists, it would imply that the motions of all absorber particles, in the assumed absorber, are correlated to each other. Such a correlation can be described by classical thermodynamics. In this context, classical thermodynamics predicts that such a correlation is very unlikely for a very huge astrophysical absorber. Thus, the arrow of radiation is again understood (according to Zeh's suggestion) as a by-product of the thermodynamic time arrow.

---

[27] It may also be called the 'hard' Sommerfeld condition if a convincing critique by Price (2006) regarding the usual formulation of the Sommerfeld condition is considered. However, I show that regardless of the form of the Sommerfeld condition used in Zeh's account, his suggestions cannot plausibly describe the origin of the arrow of radiation.

[28] 'Cosmic past' should mean 'according to cosmic time' in the 'past' of our present cosmological epoch.

However, there is another more important argument that counters this approach. Zeh's approach does not provide any argument that demonstrates that a source in our particular spacetime region (or in electromagnetic shielded regions) could not be associated with fully advanced radiation (type c). This is because, even if the Sommerfeld condition is fulfilled in our particular spacetime region, we can describe a smaller sub-region that includes only one accelerated classical source. This source, according to Maxwell's equations, may in principle be associated with fully advanced radiation. There is no known correlation between the existence of the huge astrophysical absorber and the possibility of a single accelerated source being associated with fully advanced radiation (if the radiation field associated with the source is not the *total* field of the entire bounded spacetime region). Zeh's approach would be convincing if the total electromagnetic wave field in the observable spacetime region were the only wave field that could be described. However, we can also describe smaller sub-regions, and fully advanced radiation would be allowed in these sub-regions according to Maxwell's equations. Therefore, Zeh's approach does not convincingly explain the origin of the fact that accelerated sources in nature seem not to be associated with fully advanced radiation, which seems to be the most suitable characterisation of the arrow of radiation. Hence, Zeh's account seems unattractive for describing the origin of the radiation arrow in the proposed characterisation.

Besides, notice that Price (2006) has presented a discussion of the Sommerfeld condition in general and its importance for time asymmetries in wave phenomena. Nonetheless, the argument outlined above is sufficient for rejecting Zeh's approach regarding the origin of the arrow of radiation. However, for the sake of completeness, in Digression IV.D I present a more detailed discussion regarding the Sommerfeld condition and its importance for time asymmetries of wave phenomena in general.

IV.5. On The Arrow of Radiation and Time Asymmetric Spacetimes

I will argue by following parts of Castagnino, Lara and Lombardi (2003) and Castagnino and Lombardi (2009) that the arrow of radiation is understandable as a by-product of a time asymmetric energy flux. In addition, I argue that the radiation arrow and its alignment with respect to the cosmological explications from the fundamental time asymmetry (chapter III) are only understandable if some kind of connection between the alignments of proper and cosmic time asymmetries is assumed.

Regarding connections between cosmic and proper time asymmetries it is well known [see Earman (1974), Castagnino, Lara and Lombardi (2003) or Castagnino and Lombardi (2009)] that we can use a non-vanishing, continuous timelike vector field on a time-orientable spacetime to distinguish between the semi-light cones.

> *'Assuming that spacetime is temporally orientable, continuous timelike transport takes precedence over any method (based on entropy or the like) of fixing time direction; that is, if the time senses fixed by a given method in two regions of spacetime (on whatever interpretation of regions you like) disagree when compared by a means of transport that is continuous and keeps timelike vectors timelike, then if one sense is right, the other is wrong.'* [Earman (1974), p. 22]

With a non-vanishing, continuous timelike vector field in an asymmetric spacetime, we can understand the difference between the past and future semi-light cones at one arbitrary spacetime point as physical, because the global properties are different (physically) in each direction of cosmic time. Therefore, they differ in each direction in which the vector field can point. However, the conventional labels 'cosmic past' and 'cosmic future' can be combined with the 'arrowhead' of the four-vector as well as with the mirrored direction. Thus, what remains is a physical difference between the directions of the timelike vector field and its mirrored direction. Nevertheless, if a proper and local time asymmetry is defined according to this physical difference, the alignment of this asymmetry with respect to the cosmic asymmetry is arbitrary. Thus, not all properties of the local asymmetry can be understood by considering even fundamental cosmic time asymmetries in spacetime. Note also that this is one important point that seems to be overlooked in the analysis from Castagnino, Lara and Lombardi (2003) and Castagnino and Lombardi (2009)

Nevertheless, the considered vector field is finite at each point in spacetime. Thus, the difference between the labels future and past semi-light cone, given by the direction of this timelike vector field, remains the same at each point in spacetime if one conventional labelling is chosen at one point. However, at first glance, this difference appears to be more technical than physical because the vector field is, so far, just a mathematical construction. However, the possibility of defining such a vector field illustrates that the differences between the past and future semi-light cones could have local physical and time asymmetric consequences.

But if we want to come closer to the origin of the arrow of radiation, we must analyse the physical consequences of this difference between the semi-light cones in more detail. To do so, we must first identify physical candidates for the role of the continuous, non-vanishing timelike vector field. As we will see below, it is useful to consider the energy-momentum tensor:

$$T_{\mu\nu} = \frac{1}{8\pi}\left(R_{\mu\nu}(g) - \frac{1}{2} g_{\mu\nu} R(g) - \Lambda g_{\mu\nu} \right).$$

(4.14)[29]

The components of $T_{\mu\nu}$ as they stand in (4.14) do not generally act as a continuous, non-vanishing timelike vector field, but we can add two conditions to $T_{\mu\nu}$, namely:

a) $T_{\mu\nu}$ is a type I energy-momentum tensor,[30] and

b) This tensor satisfies the dominant energy condition $T^{00} \geq \left| T^{\mu\nu} \right|$ for any orthonormal basis.

In this case, with condition a), we can write (4.14) as

$$T_{\mu\nu} = s_0 V_\mu^0 V_\nu^0 + \sum_{i=1}^{3} s_i V_\mu^i V_\nu^i \, ,$$

(4.15)

where $\left\{ V_\mu^0, V_\mu^i \right\}$ is an orthonormal tetrad, and, as in the standard notation, V_μ^0 is timelike and V_μ^i is spacelike with $i \in \left\{ 1, 2, 3 \right\}$.

Now, I will use condition b), which shows that $s_0 \geq 0$ and that s_i is given by s_0 or $-s_0$. Therefore, where s_0 is not zero:

$V_\mu^0(x)$ (where x represents the spacetime coordinates) is a continuous, non-vanishing timelike vector field. Moreover, $T^{0\mu}$ (as usual) can be interpreted as the physical energy flux described by a continuous, non-vanishing timelike vector field.[31]

---

[29] Here $R_{\mu\nu}$ is the Ricci tensor, R is the Ricci curvature, Λ is the cosmological constant and $g_{\mu\nu}$ is the metrical tensor.

[30] This means that the tensor can be described in normal orthogonal coordinates. See Hawking and Ellis (1973) and also equation (4.15) for the meaning of 'type I' or 'normal' in this context.

[31] This interpretation appears to be canonical in the context of general relativity or empirical equivalent spacetime theories, but there are exceptions that show that this understanding of $T^{0\mu}$ is not valid in general. But the exceptions appear only when quantum cosmology or QFT is considered. Critical points are, for example, the Casimir effect, squeezed vacuum or

Also, for our particular spacetime, conditions a) and b) seem fulfilled (at least in the context of classical physics), which is sufficient for applying the strategy of Earman (1974) without any further consideration. Types of considerations (e.g. Castagnino, Lara and Lombardi (2003)) regarding other timelike vector fields, provided from pseudo tensors, seem not necessary and moreover less plausible if the asymmetry of the field should be interpreted as an local energy flux.

But note that we cannot make any assumptions about the type of the energy-momentum tensor in general because we do not know anything about the phenomenology of other possible universes, and I will not consider speculations about quantum gravity at this point. As a result, all my results have a 'reduced' fundamentality. The conclusions are nevertheless interesting because they help us to understand the origin of the energetic time asymmetry by determining the conditions on which it depends. Nevertheless, I now show how this time asymmetry, without further consideration of other vector fields (in contrast to e.g. Castagnino, Lara and Lombardi (2003)), is connected to the arrow of radiation.

In spacetimes where $T^{\mu\nu}$ is given by a type I tensor that satisfies the dominant energy condition, we can interpret $T^{0\mu}$ as the local energy flux. In fact, our particular universe seems to satisfy this condition, at least for classical processes. Hence, $T^{0\mu}$ can be understood as the local energy flux, at least in spacetimes similar to ours. Moreover, the physical time asymmetry from chapter III is an intrinsic property of $T^{0\mu}$. This yields a local energy flux that is always directed from the proper past to the proper future. Thus, fully advanced radiation is forbidden by the time directed energy flow $T^{0\mu}$ because fully advanced radiation would imply an energy flux from the proper future to the proper past. But note that this is not sufficient to conclude that the radiation arrow is understandable as a by-product of the fundamental time asymmetry. The connection between the cosmic and the proper time asymmetry, as far as analysed above, shows only that the proper time asymmetry, leading to the radiation asymmetry, can be imbedded in the cosmic asymmetry. But the alignment of the proper asymmetry, with respect to the fundamental cosmic one, is not yet fixed by the understood analysis. Hence, this analysis can demonstrate a possible origin of the arrow of radiation but the alignment of this arrow (the pointing of the proper future in the cosmic future direction) cannot be explained by the proposed analysis so far. This needs to be stressed here. I will not try to answer this question in this investigation; nevertheless, I think that the proposed investigation is quite useful for understanding the origin of the radiation arrow and for helping us see what is needed to understand this arrow as a consequence of the fundamental cosmic time asymmetry. Similar accounts, which also

---

Hawking evaporation. [See, for example, Visser (1996), (2002).]. Thus, the analysis in this chapter is restricted to classical physics.

motivated my investigation [for example Castganino and Lombardi (2009) and Castagnino, Lara and Lombardi (2003)] have failed to notice this crucial lack:

> *"...the local energy flow defines the future light semicone: the energy emitted at x must be contained in or must belong to $C_+(x)$. Therefore, in any case the twin corresponding to this kind of energy flow is the member of the pair that must be retained as physically meaningful. For instance, in electromagnetism only retarded solutions fulfill this condition since they describe waves which cannot propagate outside of the future semicone "* [Castganino and Lombardi (2009) p. 23, 24]

Hence, even if I am very sympathetic to their investigations, apart from other crucial disagreements, I must reject the claim that the general strategy is able to understand the alignment of the radiation arrow (with respect to the cosmic asymmetry) as a consequence of the cosmic asymmetry itself.

Thus, we can understand parts of the properties of arrow of radiation in terms of the explanation for its occurrence, as shown above. But, the alignment of the arrow, with respect to the cosmic asymmetry, is not explained by the cosmic time asymmetry. Thus, the radiation arrow is not yet understandable as a by-product of the physical explications of the fundamental time asymmetry in cosmology (see chapter III). But the analysis showed what is needed to create such an understanding, namely the connection between the alignments of the cosmological explications of the fundamental time asymmetry in cosmology and the proper time asymmetry of the relativistic energy flows in spacetimes similar to ours.

IV.6. Summary

This chapter has shown the following:

1) The formulated prominent approaches to the problem of the arrow of radiation, as discussed, cannot describe its origin in a plausible and satisfactory manner.

2) The arrow of radiation in classical electrodynamics can be understood as a result of the relativistic energy flux in spacetimes similar to ours but not directly as a by-product of the physical explications of the fundamental time asymmetry in cosmology.

Regarding the first point:

i) To reach claim 1), I considered the standard approach from physics, as advocated, for example, by Jackson (1999) or Rohrlich (2005). I showed that this standard approach cannot provide a plausible description of the origin of the arrow of radiation. The reason is that the standard approach needs to refer to a time-directed notion of causality to exclude fully advanced radiation in nature.

ii) I investigated the approach taken by Frisch (2000) and showed that parts of his approach appear to be inadequate. I agree that the origin of the arrow of radiation should be understood in terms of the retardation condition; however, in contrast to Frisch (2000), I think this condition should be explained by deeper considerations. As it happens, the time asymmetry in the local energy flux can be understood as an explanation of the retardation condition in classical electrodynamics.

iii) I discussed two accounts that were formulated by Price (1991a), (1991b), (1994), (1996) and Price (2006). I showed that neither of these approaches to describing the origin of the arrow of radiation seems to be plausible. My arguments (regarding Price (1996)) are based on the fact that Price focused on macrophysics and argued that the origin of the arrow of radiation is on this level. However, Frisch (2000) has already shown that Price's arguments for favouring a time symmetric understanding of classical microphysics are misleading because of Price's physically implausible absorber theory. Thus, Price cannot explain the arrow of radiation in microphysics *or* the absence of this arrow in classical microphysics. I also argued that the suggestion in Price (2006) seems implausible given that in relativistic electrodynamics (and also in QED), electromagnetic fields are not described analogous to the correlated motions of the constituents of a classical wave medium.

iv) Finally, I discussed the approach of Zeh (1989) and (1999). I showed that Zeh's account cannot establish that single accelerated sources in our spacetime region (or in electromagnetic shielded regions) are not associated with fully advanced radiation. Thus, it fails to describe a possible origin of the arrow of radiation in nature because it focuses on the total wave field in a huge spacetime region.

To make the second claim, I showed that the arrow of radiation in classical electrodynamics can be understood as a consequence of the time asymmetric energy flux in spacetimes similar to ours but not yet as a by-product of the fundamental (according to definition I in chapter II) time asymmetry in cosmology (see chapter III). Nevertheless, the arrow necessarily occurs in a spacetime that satisfies conditions a and b after (4.14). Both conditions seem fulfilled in our particular universe, at least in classical physics. Thus, I believe that this clearly advances the philosophical understanding of proper time directions in classical electrodynamics even with respect to the similar proposals from Castagnino, Lara and Lombardi (2003) and Castagnino and Lombardi (2009).

The following diagram shows again the coarse structure of this chapter in order to provide a brief overview of the didactical structure.

Classical electrodynamics is a time symmetric theory; therefore, it does not offer an intrinsic explanation of the arrow of radiation.

No convincing understanding of the radiation arrow from an (assumed) asymmetry of causation

No convincing understanding of the radiation arrow from astrophysical boundary conditions

Determination of necessary conditions on the energy-momentum tensor to make the radiation arrow understandable in terms of the energy flux chapter III

New ad hoc laws in classical electrodynamics are only attractive options if no other convincing proposal can be made.

No convincing understanding of the radiation arrow in terms of macroscopic thermodynamics or the low entropy conditions of the universe

The radiation arrow and the connection to cosmic time asymmetries in *some* spacetimes

So far, however, the suggested understanding of the time directions of proper times (the arrow of radiation) and cosmic times (the fundamental time arrow in cosmology and the explications in almost all spacetimes, see chapter III) seem to depend on considerations in classical physics. These considerations were discussed in chapter III for the cosmological and fundamental time direction, and in this chapter, taking into account only general relativity (or empirically equivalent spacetime theories) and classical electrodynamics. Thus, one claim of the next chapter is to consider quantum field theoretical models in cosmology and to show possible explications of the fundamental time asymmetry (chapter III) in those models.

Furthermore, I will show that one prominent time arrow, the arrow of time in thermodynamics, can be understood on the fundamental level of quantum thermodynamics as a necessarily occurring by-product of the cosmic time asymmetry. I will show that this time arrow in quantum thermodynamics exhibits some interesting properties that are not associated with this time direction in classical and traditional considerations. One of these properties is that this time arrow is based on a deeper time asymmetry of the particle number operator.

Moreover, in the traditional accounts of the thermodynamic time arrow, the entropy also increases with decreasing time (theoretically), but in the understanding suggested in the next chapter, the entropy (for some crucial definitions of entropy) decreases with decreasing cosmic time. This allows the formulation of a much more physically effective second law of quantum thermodynamics. Moreover, the crucial time asymmetry, as it is presented in the next chapter, will be objective and physically effective regardless of whether an epistemic or an ontic interpretation of entropy is favoured. We will see that more precisely in the next chapter.

Digression IV.A: Implications from the Night Sky?

The argument that there are direct empirical data to support the view that the arrow of radiation cannot switch its direction, even if our particular spacetime were a recollapsing one, is sometimes sound in the discussion [see, for example, Gell-Mann and Hartle (1994) or Zeh (1989), (1999)]. For example, Zeh argues that data from astronomy seem to support the view that typical situations in the laboratory (here, the existence of an approximately perfect absorber) are also realized in our particular universe:

> 'Do similar arguments also apply to situations outside absorbing boundaries, in particular in astronomy? The night sky does in fact appear black, representing a condition [$F_{in} \approx 0$], although the present universe is transparent to visible light. Can the darkness of the night sky then be understood in a realistic cosmological model?' [Zeh (1999), p. 25]

Various authors [for example, Gell-Mann and Hartle (1994)] have argued further that the darkness of the night sky is an empirical hint that no fully advanced radiation exists in the cosmic future of our particular spacetime region.

> 'Consider the radiation emitted from a particular star in the present epoch. If the universe is transparent, it is likely to reach the final epoch without being absorbed or scattered. There it may either be absorbed in the stars or proceed past them toward the final singularity. If a significant fraction of the radiation proceeds past, then by time-symmetry we should expect a corresponding amount of radiation to have been emitted from the big bang. Observation of the brightness of the night sky could therefore constrain the possibility of a final boundary condition time-symmetrically related to the initial one.' [Gell-Mann and Hartle (1994), pp. 326–327]

This argument is discussed in detail by Price (1996) and Price (2006) because he needs to argue that this argument is invalid if the Gold model, as he suggested, is a plausible cosmological model. I think Price's discussion and conclusion are very convincing regarding this point. Therefore, I sketch this argument only briefly to show that we cannot simply refer to direct empirical data to argue that there is no fully advanced radiation in a hypothetical recollapsing phase of our particular universe.

Price's example shows that the darkness of the night sky has no implications for the arrow of radiation. The argument is very simple and, I think, convincing; it can be sketched as follows. Consider a galaxy in the cosmic future.[32] Now, assume that this galaxy emits fully advanced radiation. The radiation converges coherently on the galaxy; if it were measured in the cosmic present, there is no way it can converge coherently on the galaxy in the future. Thus, there is no way that fully advanced radiation from an astrophysical object in the cosmic future could be measured in the cosmic present. Price (1996) and Price (2006) illustrates that argument as shown in Figure IV.1.

Reverse galaxy

$-x$ $+x$

O
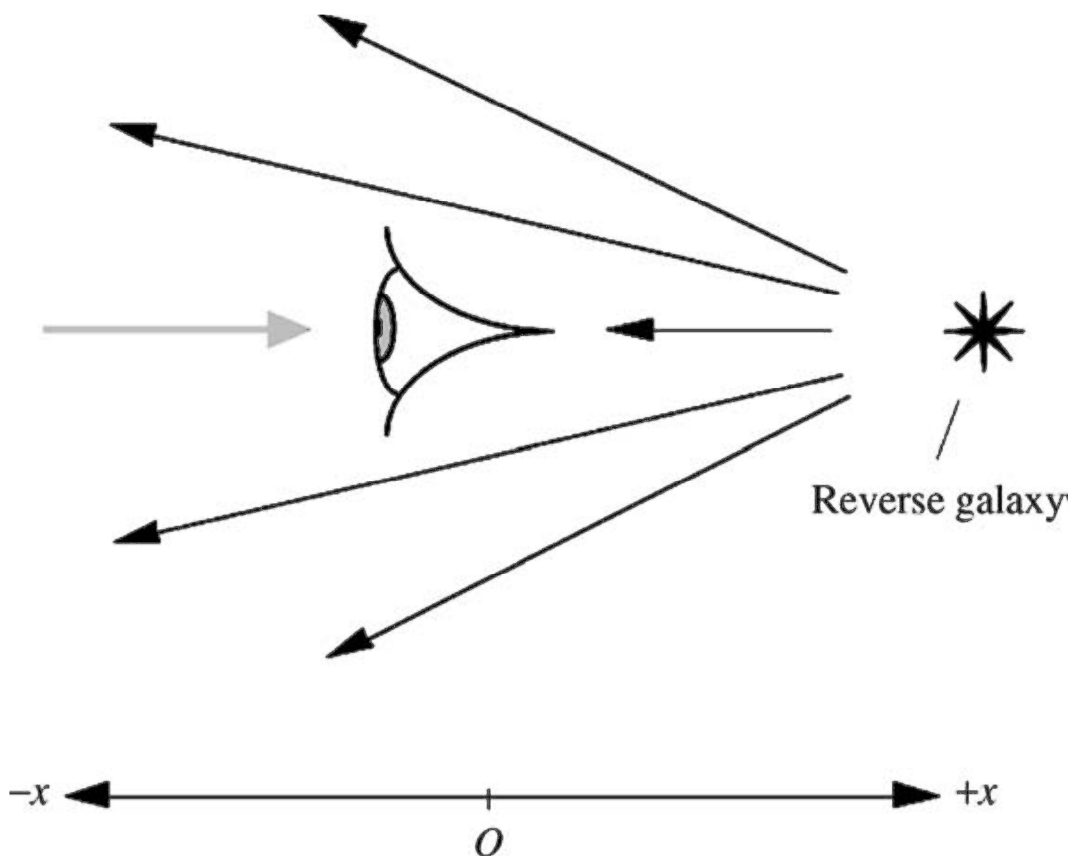
Fig. IV.1 Figure obtained from Price (2006). The human eye symbolizes an observer in the cosmic present O, and $-x$ and $+x$ are the directions toward the cosmic past and future, respectively. A galaxy

---

[32] 'Cosmic future' is understood according to the fundamental time asymmetry from chapter III. In the original discussion (Price (1996) and Price (2006)), of course, it would simply be a galaxy in the phase opposite to that of a human observer in a closed spacetime (for Price, contracting or expanding are only conventions).

that emits fully advanced radiation can do so only if this radiation is not absorbed in the cosmic present. If this were the case, the radiation would not converge coherently on the galaxy.

Thus, the observed darkness of the night sky does not support the view that there is no fully advanced radiation in our cosmic future.

Digression IV.B: On the Reinterpretation of the Absorber Theory

In this digression, I show in detail that the reinterpretation of the absorber theory by Price (1996) is not really a reinterpretation. I argue, partly on the basis of Frisch (2000), that his suggestion is more than that; in fact, I shall demonstrate that it is an independent physical theory. Moreover, I shall show that this theory seems implausible because it contradicts either Maxwell's equations or the empirical data. To show this, however, we must take a closer look at the Wheeler–Feynman theory (1945).

The central assumption of this theory is that all electric charges in our universe are surrounded by a perfect absorber. This assumption is crucial [see Wheeler and Feynman (1945)]. Moreover, it is a very questionable assumption, as seen by modern astrophysical data. This is the most serious criticism of the theory in general, I think. However, if we wish to argue that Price's suggestions are implausible in the context of classical electrodynamics, it appears to be helpful to follow Price by first accepting the Wheeler–Feynman assumption that every moving charge in our particular universe is surrounded by a perfect absorber.

Because we accept this assumption, it follows for all field components of any electromagnetic field inside the boundaries that, in a spacetime region outside of the absorber, the following is true:

$$\sum_n F_{n;ret}(x) + \sum_m F_{m;adv}(x) = 0; x > x_{abs}.$$

(4.B.1)

The retarded field components $F_{n,ret}$ can be seen as wave components coherently radiating away from the source, and the advanced field components $F_{m,adv}$ can be seen as coherently arriving components. However, fully destructive interference between the field components is not possible without a special assumption about the geometry of the system. Thus, (4.B.1) could not be realized by interference effects alone, and we find from (4.B.1) that both field components are equal to zero independently. Therefore, the following is also true:

$$\sum_n F_{n;ret}(x) - \sum_m F_{m;adv}(x) = 0; x > x_{abs}.$$

(4.B.2)

Dirac (1938) proved that, if the difference between all advanced and all retarded field components is zero in one spacetime region, this difference must be zero in every spacetime region. Thus, from (4.B.2), it follows

$$\sum_n F_{n;ret} - \sum_m F_{m;adv} = 0 \ . \tag{4.B.3}$$

Now, consider a charge q that is surrounded by a perfect absorber. According to the Wheeler–Feynman theory, we write the electromagnetic field of this charge as one-half of the sum of all retarded *and* advanced field components for all absorber particles. Wheeler and Feynman (1945) showed that this field is identical to the sum over all the retarded field components for all absorber particles $\sum_{k \neq q} F_{ret}^k$ plus one-half of the difference between the retarded (F_{ret}^q) and advanced (F_{adv}^q) field components for the charge q. The sum $\sum_{k \neq q} F_{ret}^k$ gives the retarded field component of the entire system except that of the charge q itself. Note also that, half of the difference between F_{ret}^q and F_{adv}^q is exactly the self-interaction term from the Dirac theory (1938). Thus, we obtain the same empirical predictions (with respect to the charge) from the Wheeler–Feynman theory (1945) as from the Dirac theory (1938).

However, there is one important difference between the formulations. In the Wheeler–Feynman formulation, the electromagnetic field associated with the moving charge is half retarded *and* half advanced, so no time asymmetry is imbedded in the electromagnetic field. Note, however, that all arguments for constructing a time symmetric electromagnetic field associated with the moving charge in the Wheeler–Feynman theory can likewise be made for constructing a time symmetric electromagnetic field associated with the *absorber* particles, including the self-interaction of all of them. Such a field is not observable in nature, and the absorber theory alone gives no reason that the electromagnetic field of these *absorber* particles should behave differently from that of a moving charge. Thus, we again have a time asymmetry in the behaviour of the absorber particles. Thus, the Wheeler–Feynman theory simply shifts the question of explaining the occurrence of the arrow of radiation from the moving charge to the absorber particles. Wheeler and Feynman (1945) tried to explain this fact by referring to classical thermodynamics. Specifically, they tried to argue that the electromagnetic field associated with absorber particles (q') could not have the form of the sum of all the advanced field components $\sum_{k \notin q'} F_{adv}^k$ associated with the absorber particles minus the self-interaction term. The thermodynamic argument goes as follows. The absorber, as a macroscopic entity, has a finite temperature. Thus,

all the absorber particles are in chaotic motion as long as there is no interaction with the electromagnetic field of a source. Because of the chaotic motion of the absorber particles, the sum over the advanced fields from all the absorber particles $\sum_{k \notin q} F_{adv}^k$ is zero on average. Therefore, the reason that, according to Wheeler and Feynman (1945), there is no time symmetric field associated with all the absorber particles is given not by classical electrodynamics, but by classical thermodynamics, namely, by the finite temperature of the macroscopic absorber. However, as Price (1996) and also Frisch (2000) argued, this reasoning seems misleading. This can be seen in the fact that, *after* the interaction of the absorber particles with the half-advanced, half-retarded field from the source, the absorber particles are, of course, not in chaotic motion, because their movements are correlated via the electromagnetic field from the source. If this is true there is no reason for the assumption that the advanced components of the field associated with the source will not correlate the movements of the absorber particles *in advance*, simply in the same way as the retarded field components do. Thus, the sum over all the advanced field components of all the absorber particles $\sum_{k \neq q} F_{adv}^k$ would *not* be zero, because they are *not* in chaotic motion and this is independent of the temperature of the absorber. Thus, there would be no reason, in classical electrodynamics or in classical thermodynamics, that the advanced field components of the absorber and of the charge would not be observable in nature. Thus, the Wheeler–Feynman theory (1945) fails to explain the origin of the arrow of radiation. So far, I agree with Price's analysis of the absorber theory [see Price (1996)].

Now, however, I shall argue that the reinterpretation of the Wheeler–Feynman theory (1945) by Price (1996) is even more implausible for providing an improved understanding of the radiation arrow. Price argued that the theory would show that the fully retarded field of a charge is identical with the sum over all the advanced field components from all the absorber particles $\sum_{k \neq q} F_{adv}^k$. If this were true, it would be obvious that electromagnetic fields in nature can always be seen as, in fact, totally time symmetric. Hence, there would be no empirical evidence for the existence of an arrow of radiation in nature. Therefore, I shall argue now that Price's reinterpretation of the Wheeler–Feynman theory (1945) contradicts some laws of classical electrodynamics and thus is not plausible.

As noted above, the first step in the 'reinterpretation' by Price (1996) is to assume that the fully retarded field of a charge is identical to the sum over all the advanced field components of all the absorber particles $\sum_{k \neq q} F_{adv}^k$. Price puts it this way:

> *'the real lesson of the Wheeler-Feynman argument is that the same radiation field may by described equivalently either as a coherent wave front diverging from [a charge q], or as the sum of coherent wave fronts converging on the absorber particles.'* [Price (1996) p. 71]

Thus, the 'reinterpretation' has a basic equation:

$$F_{ret}^{q} = \sum_{k \neq q} F_{adv}^{k} \ . \tag{4.B.4}$$

I argue that (4.B.4), as a general statement, is physically implausible. Firstly, I shall ask how the equation of motion would look if (4.B.4) is assumed. There are two possibilities for constructing the equation of motion:

> i) We could use the Dirac theory (1938) and the self-interaction term or the theory from Spohn (2000). Here, we would assume that the equation of motion is constructed by these theoretical considerations in the context of classical electrodynamics.

> ii) We could use the Wheeler–Feynman theory (1945) and assume that the equation of motion could be constructed in terms of the interaction between the time symmetric fields from the source and the absorber particles.

In case ii), as we will see, formal difficulties arise if we try to construct a charge's equation of motion without contradicting (4.B.4). This is evident from the following considerations.

Let us assume two electromagnetic sources q and q'. Assume that q' has a finite distance in space from q. In particular let q' be one of the absorber particles surrounding the charge q. Now, according to (4.B.4) it follows that

$$F_{ret}^{q} = F_{adv}^{q'} + \sum_{k \notin \{q,q'\}} F_{adv}^{k} \ . \tag{4.B.5}$$

Thus, from (4.B.5), it follows that the field at the source q' is given by F_{ret}^{q} and the sum in (4.B.5), so the field at q' is independent of the self-interacting term, which is given by one-half of the difference between $F_{ret}^{q'}$ and $F_{adv}^{q'}$. However, as a general prediction of (4.B.5), this contradicts the empirical data in classical electrodynamics. In the standard formulation of classical electrodynamics, whether relativistic or not, (4.B.5) is fulfilled only if q' is a particle that could not be excited. So, there is no

reason, of course, that two excitable charges could not be at a finite distance from each other; in fact, it seems that we observe such relationships between two charges in nature. Thus, (4.B.5) contradicts the empirical observations. Therefore, in the theory suggested by Price (1996), the equation of motion of a source could not be constructed if we try to use the Wheeler–Feynman theory (1945) together with Price's suggested basic equation (4.B.4). Consequently, Price must assume i) as the only plausible possibility for constructing a source's equation of motion, so his theory must use the Dirac theory (1938) *and* the additional equation (4.B.4).[33] However, this account is also not fruitful for constructing the equation of motion. From (4.B.4), it follows that the sum of F_{ret}^{q} and F_{adv}^{q} is equal to the sum of all the advanced field components of all the particles in the system $\sum_{k} F_{adv}^{k}$. Therefore, without special assumptions regarding the system, though, the sum of F_{ret}^{q} and F_{adv}^{q} will not go to zero for an observer far (in any finite distance) from the system. Thus, $\sum_{k} F_{adv}^{k}$ will also not go to zero for such an observer. However, this fact contradicts (4.B.1) and therefore the Wheeler–Feynman absorber theory (1945), of which (4.B.1) is the crucial assumption. This is because (4.B.1) says that a charge is surrounded by a perfect absorber. Thus, if Price will not suggest (4.B.5) because it contradicts the empirical observations, his theory becomes inconsistent with (4.B.1) and thus with the Wheeler–Feynman theory, which is the theory he claimed to reinterpret. Now, this seems very puzzling, because (4.B.1) is the most important assumption of the absorber theory, and (4.B.4) is the most important assumption of Price's theory.

Altogether, Price's suggestion contradicts either the empirical observations or the theory that he tried to reinterpret. Thus, it is obvious that this suggestion is actually a new physical theory of classical electrodynamics, independent of the Wheeler–Feynman theory and (4.B.1). Of course, the fact that (4.B.4) is not a reinterpretation but a new theory that contradicts the Wheeler–Feynman theory (4.B.1) does not mean that Price's suggestions is implausible. But, I shall demonstrate that (4.B.4), and thus the entire theory, is not a physically plausible theory for classical electrodynamics at all.

First, if we wish to add some thoughts about the physical validity of Price's theory, we should consider Maxwell's equation:

$$dF = 4\pi j \, . \tag{4.B.6}$$

---

[33] Alternatively, he could use the theory of Spohn (2000) instead of Dirac (1938), but the arguments will be the same in both cases. Thus, for simplicity and familiarity, I construct the argument only for Dirac (1938). To translate the argument, see section IV.1 or Spohn (2000).

(4.B.6) means that the divergence of the electromagnetic field F is given by the four-current j. Additionally, consider a source q surrounded by a perfect absorber. It is trivial that the retarded field from this charge has a source in every neighbourhood of the charge; this source is the charge q itself. However, the advanced field of the absorber particles has no source in a sufficiently small neighbourhood of the charge. Because there is a source in every neighbourhood of the charge, dF_{ret}^q does not equal zero. In contrast, the sum over all the divergences of all the advanced field components of all the absorber particles $\sum_{k \neq q} dF_{adv}^k$ is zero, because the advanced fields of the absorber particles do not have a source in every neighbourhood of the charge. Thus, if (4.B.4) is true (and it is the central assumption of the Price's theory), this means that at least one of the two electromagnetic fields, F_{ret}^q or $\sum_{k \neq q} F_{adv}^k$, contradicts (4.B.6), and (4.B.6) is the famous Maxwell equation, which is usually understood as the most fundamental equation in classical electrodynamics. If Price predicts that $\sum_{k \neq q} F_{adv}^k$ contradicts (4.B.6), this would mean that it does not contradict empirical observations because we have no observations of fully advanced field components of absorber particles. But firstly, Maxwell's equations seem to be one of the most fundamental laws in classical physics. Therefore, it seems very problematic to assume a new theory of classical electrodynamics that contradicts it. Moreover and more importantly, Price's theory does not explain why only the advanced fields contradict Maxwell's equation and not the retarded fields. Note, however, this would be a time asymmetry in classical electrodynamics, which seems not to be explained by the suggested theory (4.B.4). Even if we accept (4.B.4) and deny the validity of Maxwell's equations for fully advances fields, we would again have an unexplained time asymmetry in classical electrodynamics. Price can eliminate this asymmetry only if he denies that *only* $\sum_{k \neq q} F_{adv}^k$ contradicts Maxwell's equations. In this case, he must assume that the retarded field of the charge also contradicts Maxwell's equations. So, in this case, Maxwell's equations would also be false for the retarded fields, which seem to be inconsistent with the empirical observation again.

Hence, to save his suggested theory (4.B.4), Price must predict either that Maxwell's equations are wrong for both the retarded and advanced fields, or that there is also a time asymmetry in classical microphysics that is not explained by (4.B.4).

Thus, I think it is well motivated to reject Price's suggestion (1996). However, Frisch (2000) seems to argue that there could be another way of understanding Price's theory, for which (4.B.4) is not necessary. It seems to me that Price (1996) does not want to suggest this understanding of his theory because (4.B.4) seems to be the formalism

that follows from the quotation above. Nevertheless, we shall consider this interpretation.

Frisch (2000) suggested that (4.B.4) could be applied only in a special spacetime region. This, of course, has the potential to eliminate the contradiction between (4.B.4) and Maxwell's equations. We can see this by considering the following. The Wheeler–Feynman absorber theory (1945) predicts that

$$\sum_{k \neq q} F_{adv}^{k} = F_{ret}^{q} - F_{adv}^{q} \Leftarrow \sum_{k \neq q} F_{ret}^{k} = 0 \, . \tag{4.B.7}$$

Thus, it follows that (4.B.4) is also true in every spacetime region where F_{adv}^{q} is identical to zero and additionally $\sum_{k \neq q} F_{ret}^{k}$ is also zero. Thus, if we understand Price's theory in this way, it does not contradict Maxwell's equations. However, if this restriction on (4.B.4), which is visible in (4.B.7), is assumed, the following quotation from Price would make no sense.

> *'the real lesson of the Wheeler-Feynman argument is that the same radiation field may by described equivalently either as a coherent wave front diverging from [a charge q], or as the sum of coherent wave fronts converging on the absorber particles.'* [Price (1996), p. 71]

A radiation field generally cannot be described equivalently as a coherent wave front diverging from a charge or as the sum of coherent wave fronts converging on the absorber particles. As (4.B.7) shows, this would be true only if some conditions are fulfilled, so it is not true for all electromagnetic fields. Moreover, if we assume Frisch's (2000) interpretation of Price's suggestions (1991a), (1991b), (1994) and (1996), this theory cannot provide new information on the arrow of radiation because it would predict only the possible existence of spacetime regions in which the arrow of radiation is hidden and could not be observed.

Thus, I think it is plausible to reject the Price theory (1991a), (1991b), (1994) and (1996) in all forms, including (4.B.4) and the interpretation from Frisch (2000).

Digression IV.C: Thermodynamics and the Time Asymmetry in Wave Mechanics

The account discussed here pertains to the connection between the thermodynamic time arrow and the time asymmetry of wave phenomena in classical wave media. It is given in Price (2006) and similarly in Popper (1956). To outline this account and to provide independent arguments for its validity, I shall follow Price (2006) in examining the process of damping.

Consider a spatial volume, let us say a cylinder. Assume additionally that this volume is filled with an ideal gas that is in thermodynamic equilibrium. The walls of the cylinder are constructed so that a particle, having a much higher kinetic energy than the gas particles in equilibrium, can pass through them and enter the cylinder. If such an external particle arrives in the cylinder, collisions will occur; on average, the external particle will lose kinetic energy, and the other gas particles, on average, will increase their kinetic energy. Thus, on average, after a time period (the equilibration time), the average kinetic energy of the gas particles has increased. This process is illustrated in Figure IV.2, which is obtained from Price (2006).
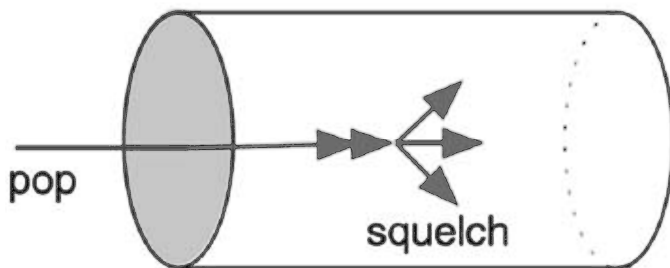

Fig. IV.2. Illustration of a damper. The 'pop' is an incoming particle with sufficient kinetic energy to pass the wall, and the 'squelch' in the cylinder illustrates the collision process.

According to the laws of classical mechanics, the time inverted process of the squelch in the figure is possible. In this process, all particles involved in a collision collide in such a way that one of them increases in kinetic energy. In this process, the particle with the increased kinetic energy, after the collision, has sufficient kinetic energy to leave the cylinder. This process would be an antidamping process. According to our

experience, in nature we find a very huge number of damping processes. In contrast, antidamping seems to be very rare in nature.

Now, we can begin to investigate how this numerical asymmetry can be responsible for the asymmetry in wave phenomena and how it is related to classical thermodynamics. To demonstrate this, we can follow Price (2006) and set up another thought experiment. Price (2006) combines a damper and an antidamper in the system illustrated in Figure IV.3.
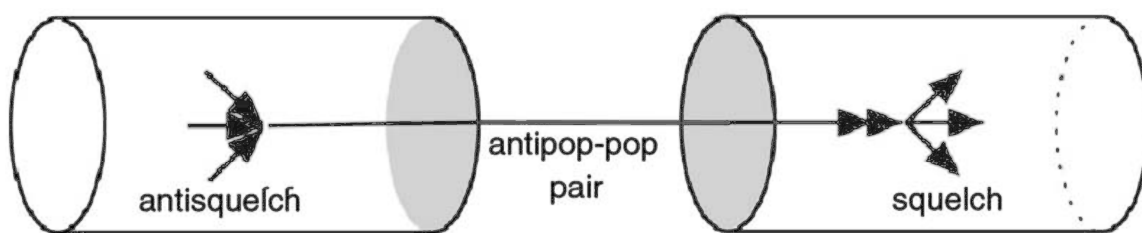

Fig. IV.3 Antidamping appears on the left; the antisquelch produces a particle with sufficient kinetic energy to leave the antidamping cylinder. The outgoing particle hits the damper and produces a squelch. If we assume that the two cylinders are isolated from other particle showers, we see only antipop–pop pairs and no pop events alone.

In a system like the one illustrated in Figure IV.3, we find damping processes only when an antidamping process has taken place. Thus, in the toy system in Figure IV.3, there is no numerical asymmetry that favours damping processes. Consequently, we find that there is a qualitative difference between the toy model in Figure IV.3 and our particular universe: In nature, we observe a numerical asymmetry between damping and antidamping. This is one of the facts that force us to assume that in the cosmic past of our particular universe, we find a state of very low entropy; otherwise, damping and antidamping processes would occur more or less equally in nature. The connection between the low-entropy cosmic past of our particular universe and the numerical asymmetry in damping and antidamping is explained as follows. If the cosmic past of the universe exhibits low entropy, there are many constituents with higher kinetic energy than the average (if not, the entropy would be higher). Thus, we have many particles that can produce a damping process, which increases the entropy of the system. However, the antidamping process is as unlikely as before; thus, the low-entropy past of the system is a necessary condition for the numerical asymmetry between dampers and antidampers.

Now, however, we must investigate how this time asymmetry is connected to the asymmetry of wave phenomena in wave media.

Absorber, emitter, sources and sinks in classical thermic fields

The sources and sinks of a thermic field can be seen as the dampers and antidampers of the last subsection. Of course, they are not identical with the sources and sinks of an electromagnetic field. An antidamper is a thermic source because it is a source of free energy for the system;[34] a damper is a thermic sink because it decreases the free energy of the system.[35] Moreover, an antidamper works as a particle emitter, and a damper acts as an absorber. Now, according to our particular universe, the low-entropy past produces an asymmetry in the number of dampers and antidampers, which leads to a numerical asymmetry in the existence of emitters and absorbers. This is all we need to see that a numerical difference exists between wave sources and sinks in a wave medium. Thus, the damping process, e.g. a particle which 'causes' a water wave, is common in nature, but the antidamping process, e.g. a water wave which coherently converges and 'causes' a particle to jump out of a lake (or from the water surface), is not (or only rarely) observed in nature.

Of course, this account for the time asymmetry in macroscopic classical wave media is essentially provided by the thermodynamic time arrow and the assumption of the low-entropy cosmic past. Chapter V covers the thermodynamic time arrow in much greater detail. But, for the purposes of this digression, it is sufficient to sketch the connection between the thermodynamic time asymmetry (respectively the low entropy state in the 'early' universe) and the time asymmetry in classical wave observations, which was done above.

---

[34] Here, we ignore the inner energy of a damper or antidamper.

[35] This is consistent with energy conservation, because the inner energy of the damper is increased by absorption, but this is not counted as kinetic energy here because it is hidden in the average energy of the gas particles in the damping cylinder.

Digression IV.D: The Sommerfeld Condition

The aim of this digression is to clarify whether the Sommerfeld condition and some particular wave phenomena in classical wave media (specifically, the fact that we cannot observe fully advanced *total* wave fields in nature) are connected. Also, this digression aims to clarify the nature of this connection, if there is any, and to outline the actual status of the debate regarding the Sommerfeld condition in general.

As I argued in the main chapter, the Sommerfeld condition is relevant only for *total* wave fields in a given spacetime region. Consequently, e.g. Zeh's account cannot explain the occurrence of the arrow of radiation for *local* wave fields in e.g. physically shielded regions or everyday observations. However, as noted in the chapter, the discussion of the Sommerfeld condition and its connection to the arrow of radiation has additional aspects, which are discussed here.

Firstly, I consider Price (2006) regarding the connection between the Sommerfeld condition and time asymmetric wave phenomena. Price (2006) argues that the Sommerfeld condition is neither necessary nor sufficient for producing time asymmetric wave phenomena. My claim shall be to demonstrate that his view and his arguments are not convincing in this regard.

I begin by examining how Price's argument works for wave phenomena in classical wave media. For simplicity, we follow Price by considering water waves on an infinite water surface. Now, we focus on a finite region of the water surface. The finite region has boundaries, and on these boundaries we can construct boundary conditions such as the Sommerfeld condition.

Now, the total amplitude of the water wave $\Phi(\vec{r}, t)$ can be described, in general, as the sum of all retarded fields F_{ret} and all incoming fields F_{in} or as the sum of all advanced fields F_{adv} and all outgoing fields F_{out}. Thus, we find:

$$\Phi(\vec{r}, t) = F_{ret} + F_{in} = F_{adv} + F_{out} . \tag{4.D.1}$$

In the example of the water region, F_{ret} is the sum of all fully retarded waves having sources in the region. These waves originate in the past semi-light cone of the point $(\vec{r}, t)$. F_{in} is the sum over all fields having no sources in the region and entering the

region. F_{adv}, then, is the sum over all fully advanced waves coherently converging to a 'sink' in the region. Here, however, the waves originate in the future semi-light cone of the point $(\vec{r}, t)$. Finally, F_{out} describes the sum over all source-free fields exiting the region. Thus, the difference between the fully retarded and fully advanced fields is the time direction of the water wave. Also, the amplitude (4.D.1) is a wave field because it is the sum of two wave fields. Thus, the wave field in every finite region on the water surface is described, in general, by the following equation:

$$F = F_{ret} + F_{in} = F_{adv} + F_{out} . \tag{4.D.2}$$

We see that the wave field (4.D.2), like (4.D.1) of course, is symmetric. The total wave field can be described by retarded and advanced fields. The Sommerfeld condition is important here because it makes (4.D.2) a time asymmetric equation. The Sommerfeld condition is: *There are no incoming fields in the described region*. Thus, it follows immediately from (4.D.2) that every wave field can be described as a fully retarded field but not, in general, as a fully advanced one. The same holds for electromagnetic waves; thus, the Sommerfeld condition describes the boundary condition that can provide time asymmetric wave dynamics as well as the fact that we cannot observe fully advanced *total* wave fields (see e.g. Zeh (1999)). To express this formally, from the Sommerfeld condition together with (4.D.2), it follows that

$$F = F_{ret} . \tag{4.D.3}$$

This is also true in the absorber theory of Wheeler and Feynman (1945). I think it is important to notice this because otherwise the connection between the Sommerfeld condition and the arrow of radiation would depend on the status of the absorber theory. In this theory the total wave field in a spacetime region that includes a charge q is

$$F_{tot}^q = F^q + R^q = F_{ret}^q . \tag{4.D.4}$$

The fields F_{ret}^q in (4.D.4) as well as the field F_{ret} in (4.D.3) are fully retarded fields, which is completely consistent with empirical observations. Thus, at first glance, it seems that wave fields in nature behave in such a way that the Sommerfeld condition would be fulfilled. Now the question is whether this impression is true, or, more precisely, whether the Sommerfeld condition is equivalent to the non-occurrence of fully advanced waves in nature. One obvious reason why this is not trivially true is

that, if the Sommerfeld condition is fulfilled, it does not follow that there are no fully advanced total fields in nature.

The hard Sommerfeld condition

It is important for our discussion that the Sommerfeld condition in the original form, that *there are no incoming waves*, is not sufficient to conclude that we cannot observe fully advanced total waves fields in general. This is simple to see in (4.D.2), from which it follows immediately that if the Sommerfeld condition is fulfilled and if, additionally, there are no outgoing waves in a region, the field can be described as fully retarded or fully advanced. Thus, even if the Sommerfeld condition is fulfilled, the wave phenomena are not necessarily asymmetric. Now, as Price (2006) also mentioned, there is a quite simple and obvious way to modify the Sommerfeld condition so that this problem will not arise. This modified 'hard' Sommerfeld condition says, in addition to the original Sommerfeld condition, that: *there are no incoming waves in a region and there must be outgoing waves from a region*. Thus, from (4.D.2) we see immediately that an asymmetry *always* follows from this hard Sommerfeld condition. Not only can a total wave field be described as fully retarded; it also *cannot* be described as fully advanced. Thus, the hard Sommerfeld condition seems to be the mathematical condition that produces a time asymmetric situation for the total classical wave field.

However, Price (2006) tried to argue that even the hard Sommerfeld condition is not equivalent to an asymmetry of *total* wave fields. If this were true, Price would have made an additional argument against the account stipulated by Zeh (1999), independent of the fact that Zeh's account explains only the origin of an electromagnetic time asymmetry in the *total* wave field of a bounded spacetime region. Thus, even if we remember that Zeh's account does not successfully explain the arrow of radiation for local sub-regions of the observable spacetime region, it seems appropriate to discuss the additional considerations from Price (2006).

Price's arguments are based on a set of thought experiments. The first one is the following: Consider, again, a water surface, and let us say that the wave field on this surface is not zero. Now, we create another wave on the surface by excitations. In this case, Price pointed out that the hard Sommerfeld condition is not necessarily fulfilled. Additionally, he argued that in such a situation we will observe a time asymmetry in the wave phenomena in this experiment. He stated that the new wave behaves asymmetrically, which means that it diverges coherently from the centre of the excitation. This demonstrates that time asymmetric wave phenomena appear even if the hard Sommerfeld condition is not fulfilled.

But remember, the asymmetry of the Sommerfeld condition is an asymmetry of the *total* wave field. The local field, associated only with the excitation, may still be asymmetric (in fact this is my argument against Zeh (1999) and I totally agree with Price in that point) but the *total* wave field would be symmetric; symmetric in the meaning that it can be described according to *both* possibilities in (4.D.2). This arises simply for mathematical reasons from (4.D.2) regarding the *total* wave field of the surface.

Again, if Price (2006) argued that the new additional wave behaves asymmetrically, independent of the original background wave field, *I agree*, but this has nothing to do with the Sommerfeld condition. In this sense, Price (2006) gives the same argument as me against the view that the Sommerfeld condition yields a plausible explanation for the radiation arrow. Nevertheless, the *total* wave field, in the outlined experiment, is a symmetric one (in the meaning from (4.D.2)).

Moreover and more importantly, Price (2006) tried to show that the hard Sommerfeld condition is also insufficient for producing a time asymmetry in wave phenomena. He argues again on the basis of a thought experiment. Again, consider a water surface. Additionally, assume an absorber (having some special properties) surrounding a finite region on the surface. The absorber is assumed to absorb all incoming waves but no outgoing ones. Thus, the hard Sommerfeld condition is seen as fulfilled in this region. Now, Price (2006) argued that in such a case the asymmetry of wave phenomena would be *inverted* (see Price (2006)). This seems puzzling to me. In contrast, it seems plausible to me to assume that in such an experiment, the wave will propagate along the water surface and could be described according to (4.D.2). If we imagine a source in the region, the total wave field is the sum of a source-free outgoing wave and a coherent wave that converges towards the source. However, as (4.D.2) shows, the total wave field is also equivalently described by a fully retarded wave. Thus, I do not see in which sense there would be an *inverted* asymmetry. This seems puzzling because the wave would not be a fully advanced wave, and the entropy behaviour of the wave medium does not disobey the second law of thermodynamics.[36] Thus, I think this experiment carries no useful information regarding wave phenomena in classical wave media or the connection between those phenomena and the entropy behaviour in time.

But, it should be mentioned that, to show that the hard Sommerfeld condition is not sufficient to yield asymmetric wave phenomena, Price presented another thought experiment. Nevertheless, I find this second experiment no more convincing than the first. In my rephrasing: Again, consider a water surface, but now, there is no wave field on the surface. Now, a shower of dust particles occurs in a finite region on this surface. When a dust particle hits the water surface, it behaves as a source for a wave field. Consider also that on the entire surface, the dust particles provide the only wave

---

[36] Even because the assumed surface is not a closed system

sources. Thus, there are no incoming waves in the finite region on the surface. However, because we have sources in the region, we have outgoing waves, so the hard Sommerfeld condition is fulfilled. Now, Price (2006) argued that there will be *no* asymmetry in this case. If I understand him correctly, he argued that the amplitude of the waves is too low for observation, but this is only the case because human eyes are not sensitive enough. This does not means that there is no asymmetry, and in the context of this experiment we can assume much better eyes than we have or other means of observing the waves on the surface. The point here is not the asymmetry of the waves, but the sensitivity of observation. I think it is not convincing to assume that there would be no asymmetry only because unaided human eyes are not sensitive enough to observe a wave phenomenon. Thus, this second experiment also does not seem to support the conclusion that the hard Sommerfeld condition is not sufficient to produce a time asymmetry of wave phenomena for *total* wave fields. Instead, I think that, given (4.D.2), it simply seems mathematically true that there cannot be such a thought experiment, because it would be inconsistent with the basic mathematical description of waves. Thus, I think I can fully agree with Price (2006) that there are other crucial time asymmetries (local asymmetries, not restricted to the *total* wave field), which are obviously not connected to the Sommerfeld condition. Nevertheless, I think it is crucial to acknowledge that the hard Sommerfeld condition produces, for simple mathematical reasons, always an asymmetry of the *total* wave field.

Nevertheless, given the aim of this digression, which is to outline the discussion regarding the importance of the Sommerfeld condition, it seems also appropriate to discuss two different proposed interpretations of the Sommerfeld condition (see Price (2006)).

Counterfactual interpretation of the Sommerfeld condition

Price (2006) discussed a counterfactual interpretation of the Sommerfeld condition and he argued that this interpretation provides a time asymmetry of its own, and thus this interpretation seems implausible if the asymmetry should come from the phenomena and not from the interpretation of the boundary conditions. This situation seems not too surprising because counterfactuals are well known to be asymmetric constructions; thus, if we describe a phenomenon or condition in time using a counterfactual, a time asymmetry will surely appear. Thus, the relevant question is whether it is plausible to use counterfactuals to interpret the Sommerfeld condition. It seems convincing to me, as Price (2006) argued, that the asymmetry of a phenomenon must be shown independently of a counterfactual description or interpretation. So, we can conclude that a counterfactual interpretation could be plausible only if the phenomenon itself is

asymmetric. Thus, I agree with Price (2006) that we should not introduce a time asymmetry from a counterfactual interpretation of the Sommerfeld condition. Even if the main chapter has argued that there is, indeed, a time asymmetry in classical electrodynamics, as shown, this asymmetry is understandable without introducing the Sommerfeld condition. Hence, in any case, an interpretation of the Sommerfeld condition, which does not presuppose an asymmetry would be more appropriate than the counterfactual interpretation.

Hence, in contrast to the counterfactual interpretation of the Sommerfeld condition, Price (2006) suggested another interpretation.

Comparative interpretation of the Sommerfeld condition

The comparative interpretation of the Sommerfeld condition differs from the counterfactual one primarily in its actual formulation. Price (2006) suggested an interpretation in which three types of wave configurations are compared. Configuration C is characterized such that it describes the wave configuration empirically observed in nature. Configuration C_{ret} is defined as the wave configuration in a spacetime region that has no incoming waves. The third configuration, C_{adv}, is defined as a wave configuration in a spacetime region that has no outgoing waves. We can generally assume that outgoing waves can appear in a spacetime region; thus, in configuration C_{ret}, the total wave can be described as fully retarded but not as fully advanced. In configuration C_{adv}, the situation is time mirrored: The total wave can generally be described as fully advanced but not as fully retarded. Now, the comparative interpretation of the Sommerfeld condition is that C_{ret} can be observed nearly as often as C, if the condition is fulfilled. In contrast, C_{adv} is never (or very rarely) observed in nature.

I think that Price (2006) suggested a very convincing interpretation of the (hard) Sommerfeld condition in this context. Thanks to his formulation, this interpretation introduces no time asymmetry in the *interpretation* of wave phenomena. Nevertheless, as shown before in this digression, Price (2006) thinks, because of his thought experiments, that the hard Sommerfeld condition does not necessarily yield to a time asymmetry in wave phenomena. Consequently, he proposed another formulation of the Sommerfeld condition.

Although we had already seen that Price's conclusion regarding the hard Sommerfeld condition seems misleading, it appears suitable, for this digression, to discuss Price's approach regarding the reformulation of the Sommerfeld condition.

The weak Sommerfeld condition

Price (2006) suggested, instead of the hard Sommerfeld condition, a weak formulation in which it is assumed only that the amplitude of the incoming waves is negligibly small. However, Price does not give a value in comparison to which the amplitudes should be small. The most natural choice would be to require that the amplitudes of the incoming waves are small compared to other wave fields in the system. However, Price's thought experiment with the dust particles gives the impression that Price could also have been thinking of a comparison with amplitudes that human beings can observe with their unaided eyes. The last possibility seems misleading, as I argued above. Modern science has many tools that make it possible to observe phenomena that are invisible to the unaided human eye. Thus, I interpret the weak Sommerfeld condition as being fulfilled if the incoming wave amplitudes are small in comparison to the average wave amplitudes in the system.

In this case, the weak Sommerfeld condition clearly does not provide an exact symmetry. This can be seen in (4.D.2). If the incoming wave amplitude is small compared to the other amplitudes, it follows that the total wave field can be described as consisting mostly of a fully retarded wave field with a small correction from the incoming small wave field. In a description including fully advanced wave fields, the outgoing wave field makes, in general, more than only *small* corrections to the fully advanced field. Thus, we see a quantitative time asymmetry if the weak Sommerfeld condition is fulfilled.

But, of course, this is what we expect if we change an absolute qualitative asymmetry, as given by the hard Sommerfeld condition, into an approximate quantitative asymmetry. Thus, I don't think this result provides some interesting insights regarding the connection between the Sommerfeld condition and the time asymmetry of wave phenomena. Instead, I think, it seems reasonable, as argued in this digression, to conclude that

> i) The hard Sommerfeld condition is sufficient to provide a time asymmetry in the *total* wave field of a system.

ii) Nevertheless, the Sommerfeld condition, in any formulation, is unable to describe the origin of the radiation arrow in its local characterisation (no occurring of fully advanced radiation from type c)).

iii) The proposed weak Sommerfeld condition clearly provides a quantitative time asymmetry for the *total* wave field, but this time asymmetry is expected given that the weak formulation of the Sommerfeld condition is only an approximation of the hard formulation.

The aim of this digression is now fulfilled and all the different discussed aspects regarding the understanding of the Sommerfeld condition are captured. Nevertheless, remember that the purpose of this digression was only to outline modern discussions regarding the Sommerfeld condition. The main part of the chapter has already shown that the puzzle of the radiation arrow can be reformulated and at least partly (modulo the alignment of the arrow regarding the explications from the fundamental cosmic time asymmetry) solved without invoking the Sommerfeld condition. Moreover, it was also shown that the consideration of the Sommerfeld condition, as Zeh (1999) tries it, is unable to describe the origin of the radiation arrow in its most convincing characterisation.

Chapter V
Time Asymmetries in Quantum Cosmology, Entropy and the Second Law of (Quantum) Thermodynamics

This chapter develops and applies a new proposal regarding the origin of the arrow of time in quantum thermodynamics. For this purpose, I will first, in section 2 and after some introductory thoughts in section 1, consider the account of Allahverdyan and Gurzadyan (2002) as representative (at least in parts) for the standard approaches to the thermodynamic time asymmetry. I shall argue that their approach does not convincingly establish an understanding of the time arrow in thermodynamics that can be seen as part of nature, or at least part of the models of nature in physics. I consider this approach in order to present some crucial issues in the discussion of time asymmetries in statistical physics, which are, I think, occasionally misunderstood.

More precise, I shall show that Allahverdyan and Gurzadyan used epistemically motivated assumptions and artificial definitions, together with statistical approximation methods, to draw physical conclusions about time directions. I focused especially on Allahverdyan and Gurzadyan (2002) because those authors are very explicit and transparent regarding the used assumptions and the associated motivations. Also, at least in the considered part of their work, I think that the used assumptions and motivations from Allahverdyan and Gurzadyan (2002) are often shared, at least implicitly or unwittingly, in the standard accounts to the thermodynamic time asymmetry.

Further, I shall isolate the critical points of their account in order to highlight the contrast to my proposal. In this proposed account (section 3), motivated and based on the physical analysis of Castagnino and Laciana (2002), I demonstrate that time directions in quantum thermodynamics that are based on entropy behaviour can be constructed without using epistemically motivated assumptions or artificial definitions.
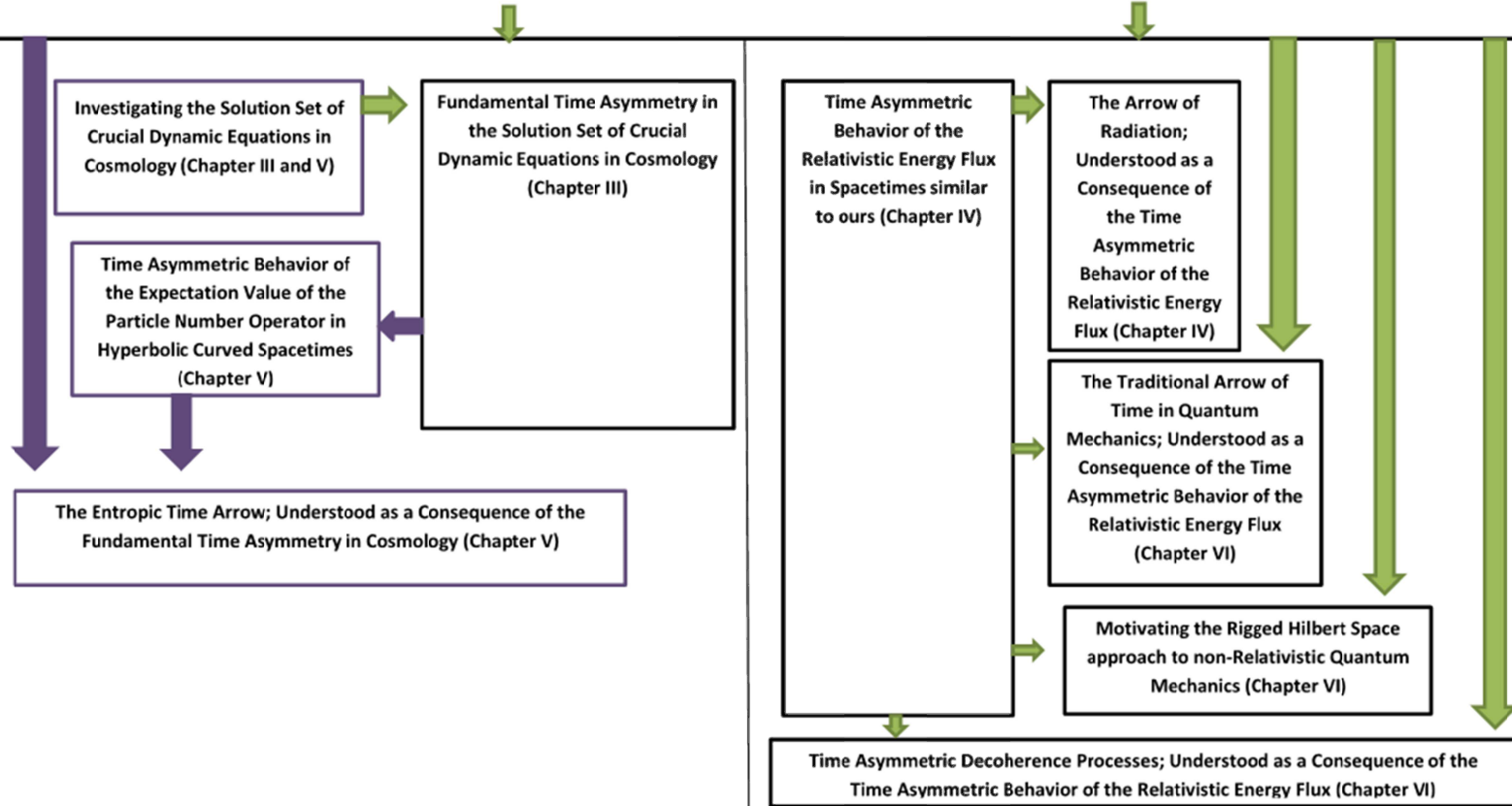
Moreover, I shall demonstrate that some crucial details of the proposed account providing some unexpected philosophical advantages in the understanding of the origin of (quantum) thermodynamic time asymmetry. Those advantages will be demonstrated during the investigation and there will be summarised again in section 4, together with the main conclusion of this chapter.

The first following diagram illustrates, as usual, which parts of the entire analysis are considered in this chapter. The second following diagram illustrates the structure of this particular chapter and I shall return to this diagram at the end of the chapter again.

Time Asymmetries in Quantum Cosmology, Entropy and the Second Law of (Quantum) Thermodynamics

Cosmic time asymmetries

Proper time asymmetries

Motivation and Definition of 'Fundamentality' in the Context of Time Asymmetries (Chapter II)

Investigating the Solution Set of Crucial Dynamic Equations in Cosmology (Chapter III and V)

Fundamental Time Asymmetry in the Solution Set of Crucial Dynamic Equations in Cosmology (Chapter III)

Time Asymmetric Behavior of the Expectation Value of the Particle Number Operator in Hyperbolic Curved Spacetimes (Chapter V)

The Entropic Time Arrow; Understood as a Consequence of the Fundamental Time Asymmetry in Cosmology (Chapter V)

Time Asymmetric Behavior of the Relativistic Energy Flux in Spacetimes similar to ours (Chapter IV)

The Arrow of Radiation; Understood as a Consequence of the Time Asymmetric Behavior of the Relativistic Energy Flux (Chapter IV)

The Traditional Arrow of Time in Quantum Mechanics; Understood as a Consequence of the Time Asymmetric Behavior of the Relativistic Energy Flux (Chapter VI)

Motivating the Rigged Hilbert Space approach to non-Relativistic Quantum Mechanics (Chapter VI)

Time Asymmetric Decoherence Processes; Understood as a Consequence of the Time Asymmetric Behavior of the Relativistic Energy Flux (Chapter VI)

Consideration of typical approaches to the entropic time arrow in statistical physics; I shall consider, as representative to similar standard approaches, the account from Allahverdyan and Gurzadyan (2002).

⇩

Approaches like the one taken by Allahverdyan and Gurzadyan (2002) are unable to provide an understanding of the thermodynamic time arrow, which is independent from epistemically motivated assumptions / definitions.

⇩

Hence, a proposal which offers an understanding, which is not based on epistemically motivated assumptions / definitions, seems much more satisfying than such standard approaches.

⇩

An alternative understanding of the entropic time asymmetry, independent of epistemically motivated assumptions / definitions but based on an explication from the fundamental time asymmetry (chapter III) in quantum cosmological models.

⇩

The quantum cosmological model and the intrinsic (cosmic) time asymmetry as an explication from the fundamental time asymmetry from chapter III

⇩

Time asymmetric behaviour of the expectation value of the particle number operator in the considered model

⇩

Entropic time arrows arising from the behaviour of the particle number operator.

⇩

Some unfamiliar properties of the entropic time asymmetries

V.1. Introductory Thoughts

Before considering the work of Allahverdyan and Gurzadyan (2002) in any detail, I shall present a short introduction to it and the related critique. The goal of the following section will be to:

a) show that Allahverdyan and Gurzadyan do not succeed in constructing a plausible understanding of a thermodynamic time arrow *in nature* (or in parts of the physical models that are assumed to describe nature correctly).

b) analyse the critical steps in their arguments in order to avoid implicit usage of them in my own suggestion. This seems important to me because it seems to me that many critical steps are quite often used in the standard view of thermodynamic time asymmetries in statistical physics.

The work of Allahverdyan and Gurzadyan (2002) is discussed here in great detail because the authors make their assumptions very transparent, so criticisms that focus on these assumptions can be made quite explicit without too much effort. Notice, however, that similar assumptions are common in some of the literature regarding the thermodynamic time arrow. For a broader discussion (implicit as well as explicit) see e.g. Allahverdyan and Gurzadyan (2002), Gold (1962), Penrose (1979) on *'General Relativity: An Einstein Centenary Survey',* Penrose (1989) in *'The Emperor's New Mind',* Penrose (1994), Hawking (1985), Page (1985), Petrosky and Prigogine (1997), Halliwell, Perez-Mercader and Zurek (1996) on *'Physical Origins of Time Asymmetry',* Zeh (1999) on *'The Physical Basis of the Direction of Time',* and Kiefer and Zeh (1995), or also Allahverdyan and Gurzadyan (2000) on *'The Chaotic Universe.*

In the second part of this chapter (section 3), I develop my suggestions regarding the understanding of thermodynamic time asymmetries. I attempt to show that the arrow of time in quantum thermodynamics can arise from an explication from the fundamental time asymmetry (chapter III) in quantum cosmological models. This understanding, as I will argue, provides some crucial advantages:

i) The proposed account should be independent of epistemically motivated assumptions or artificial definitions [in contrast to e.g. Allahverdyan and Gurzadyan (2002)]. I shall discuss those assumptions and definitions in greater detail during the first part of this chapter.

ii) The suggested account will be neutral with respect to an ontic or epistemic interpretation of entropy. Nevertheless, the crucial time asymmetry will be physically effective.

iii) The proposed understanding of the thermodynamic time asymmetry suggests an alternative formulation of the second law of thermodynamics.

Furthermore, as mentioned in chapter IV, I shall show that the fundamental time asymmetry in the solution space of the crucial dynamic equations in cosmology leads also to time asymmetric explications in quantum cosmological models (as well as in the classical case, which was demonstrated in chapter III). This is an additional goal of this chapter and is mostly motivated by the need to show that the fundamental time asymmetry (from chapter III) can provide time asymmetric explications in quantum physics, which is understood as more fundamental than classical physics.

Because both claims, regarding the thermodynamic- and the quantum cosmological time asymmetry, are strongly connected to each other, I present them together in this chapter. Thus, this chapter not only aims to provide an understanding of a cosmological time arrow in quantum cosmology (similar to chapter III), but also to show that this time asymmetry can yield to a new understanding of the thermodynamic time arrow in quantum thermodynamics.

So, I shall begin in the next section, as mentioned, with considering the work of Allahverdyan and Gurzadyan (2002). The main goal of the next section will be my critics regarding typical approach to the thermodynamic time asymmetry. As mentioned earlier, Allahverdyan and Gurzadyan (2002) are quite explicit regarding there assumptions and motivations and hence, I shall use parts of their work in order to explicate my general critics regarding traditional approach to the thermodynamic time asymmetry.

V.2. On the Traditional Understanding of the Thermodynamic Asymmetry

V.2.1. Allahverdyan and Gurzadyan; The Abstract Setup

To describe a typical construction of time asymmetries in thermodynamics, I begin with a formal description of a physical system that interacts with a heat bath. This system is also

used by Allahverdyan and Gurzadyan (2002) (but also in many textbooks) to motivate their account. The Hamiltonian of the entire considered system is given by

$$H = H_S + H_B + H_W ,$$
(5.1)

where H_S is the Hamiltonian of the system itself, H_B is the Hamiltonian of the heat bath and H_W is the Hamiltonian that describes the interaction between the system and the heat bath. Although the distinction between the subsystems ('physical system' and heat bath) seems artificial, there are no arguments that such a distinction should not be made in order to provide a simple or common description of a system. Moreover, this distinction alone is not an approximation or an assumption but only a special type of description. Thus, even if the distinction between the two subsystems is motivated by epistemic considerations, it is not critical for drawing physical conclusions. However, we shall see below how this distinction works together with other assumptions in such a way that it becomes crucial for a time asymmetric description.

In order to describe this shift from a purely descriptive tool to a crucial part of a time asymmetric claim regarding the involved physics, it seems necessary to consider the formal description of a quantum system. The state of the entire considered system is described by a density matrix $D(t)$. Additionally, the well-known von Neumann equation gives the time evolution:

$$i\partial_t D(t) = \left[H, D(t) \right] .$$
(5.2)

The density matrix is given by

$$D(t) = e^{-it\frac{H}{\hbar}} D(0) e^{it\frac{H}{\hbar}} .$$
(5.3)

One crucial and common condition of the density matrix is that at the initial time point $t = 0$, it is expressed as the direct product of the density matrix of the physical system, $D_S(0)$, and that of the heat bath, $D_B(0)$. This is the formal description of the assumption that at the initial time point, where the physical system and the heat bath 'begin' to interact, the subsystems are independent of each other. At first glance, this assumption seems reasonable because neither subsystem has time to interact with the other; thus, the subsystems are independent of each other. However, this assumption is also time asymmetric; after the initial time point, the systems are no longer independent, but before

$t = 0$ they are. This is the initial assumption that introduces the first time asymmetry into the entire description of the thermodynamic system.

There are at least two crucial objections to this assumption. The first is that 'backwards causation', which is not ruled out at this level of description [see, for example, Price (1996)], is forbidden by this assumption. The second and more essential point is that the distinction between the physical system and the heat bath seems artificial and based on not physical but epistemic or anthropocentric reasons. These reasons could be that one part of the entire system is of special interest for scientific investigation and thus is descriptively distinguished from the environment and named the 'physical system', where the environment is named the 'heat bath'. Thus, the assumption of initial independence turns an epistemic distinction (that between the subsystems) into a crucial time asymmetry.

Nevertheless, an advocate of this assumption could point out that the assumption seems natural in many laboratory experiments where two prepared subsystems, at one time point, begin to interact. However, in such a setting, time asymmetry is imposed by the experimentalists, who prepare the two subsystems independently of each other and then bring them together. Thus, in this case, this description is well motivated; however, to shed light on the origin of time asymmetries that are *independent* of human beings and applicable in *global* systems (such as the universe), the physical description should not be motivated by epistemic reasons or laboratory procedures. Although such assumptions are typical in theoretical descriptions of thermodynamic systems and well-motivated for laboratories or local descriptions, they are not founded on any fundamental (or statistical) law or global consideration.

Moreover, the assumption of the initial independence of two subsystems defined artificially and for epistemic reasons (or on the basis of a laboratory procedure) is only the initial setup in the statistical description. In some accounts of the time arrow in thermodynamics and especially in quantum thermodynamics, this assumption is used to draw ontic or at least physical conclusions about time directions. Allahverdyan and Gurzadyan (2002) put it as follows:

> *'The crucial assumption on the initial state [...] means that at the initial time t = 0 the system and the bath were completely independent. The state of the system at arbitrary positive time t is described by the corresponding partial density matrix [...]. The important point of the system–bath approach*—as well as any statistical physics approach which derives the thermodynamical arrow—*is its dependence on incomplete observability: although the system and the bath constitute a closed system, one is* interested *in the state of the system only.'* [Allahverdyan and Gurzadyan (2002), p. 5]

I shall demonstrate how a typical consideration works after this first time asymmetric assumption is accepted. This will allows us to examine all the points that are critical for drawing conclusions about a physical time direction in typical thermodynamic descriptions.

If we accept the assumption [hereafter, the *initial independence assumption* (IIA)], the state of the system without the heat bath at a special time point t can be described as

$$D_S(t) = Trace_B D(t) \, . \tag{5.4}$$

Using (5.4), we obtain another crucial issue for physical considerations in thermodynamics. Common approximation methods in quantum thermodynamics describe the state of a system according to (5.4), but the physical system also depends on the state of the heat bath, even if this dependence is small compared to that covered in (5.4). The dependence on the state of the heat bath is traced out by (5.4), so a state defined by (5.4) is, strictly speaking, an entire set of states with different sub states in the heat bath. Thus, (5.4) describes an entire set of states as one system state. This is typical for statistical physics, and we will see how this leads to a more crucial time asymmetry in the statistical description when (5.4) and the IIA are applied together.

In this case, according to (5.4), the evolution operator in time, T, becomes non-unitary:

$$D_S(t) = T(t,0)D(0) = \sum_{\alpha,\beta} A_{\alpha\beta} D(0) A_{\alpha\beta}^{\dagger} \, . \tag{5.5}$$

$A_{\alpha\beta}$ represents operators defined in the Hilbert space of the system without the heat bath; Nevertheless, they are obtained by spectral analysis of the heat bath's density matrix before the interaction 'begins'. More precise:

$$A_{\alpha\beta} = \sqrt{\lambda_\beta} \left\langle \alpha \left| e^{-it\frac{H}{\hbar}} \right| \beta \right\rangle \, . \tag{5.6}$$

Here, (5.5) and (5.6), together with the von Neumann equation (5.2), provides:

$$\sum_{\alpha,\beta} A_{\alpha\beta}^{\dagger} A_{\alpha\beta} = 1 \, . \qquad\qquad (5.7)$$

Thus, the crucial formal requirements on the formalism are fulfilled by the time evolution operator (5.5). So, we have all the equations crucial for the approximate dynamics of the system in interaction with the heat bath.

Nevertheless, equation (5.5) shows that the evolution operator is not unitary, so in general there is no inverse operator, and the *formalism* used here simply *cannot* describe a time mirrored evolution process in general. But, this time asymmetry, which looks physically important on the first glance, is simply obtained from the IIA together with the statistical approximation method. Hence, there seems no convincing argument for assuming that this asymmetry describes an ontic or global physical property of the entire system.

Another very important aspect of this formalism is that the operators $A_{\alpha\beta}$, given by (5.6), does not depend on the state of the observed system 'before' the interaction begins. Thus, the dynamics of the observed system in this formalism is independent of its initial conditions. This fact is a direct consequence of the IIA in the statistical approximation method; it is time asymmetric but, nevertheless, provided by epistemic motivated assumptions and approximation technics.

Moreover, (5.5) indicates another notable property of the formalism. The evolution of $T(t_f, t_i)$ with $t_f > t_i$ is not given by $T(t_f - t_i)$, which is the case for unitary evolution operators. This is another consequence of the IIA together with (5.4), which leads to time asymmetries in the thermodynamic description of an observed system. Hence, because of the dependence on (5.4), this asymmetry is also provided by the IIA together with statistical approximation methods.

All together: The non-unitary evolution operator, which provides a formal time asymmetry in the description method, is only a consequence of the von Neumann equation if the epistemic motivated IIA and approximation (5.4) is applied. Hence, the time asymmetry in the dynamical description is also provided by the epistemic consideration (the distinction between the subsystems), the IIA and the statistical approximation (5.4).

These time asymmetries are thus asymmetries in formal description methods that are motivated by artificial definitions. So, it appears to be interesting to quantify the difference between different descriptions methods. A possible statistical quantification of the formal difference between two formal descriptions is given by the quantity of relative entropy. This value will become crucial, in particular for the approach from Allahverdyan and Gurzadyan.

However, before I shall consider the value of relative entropy, it seems crucial to consider some intrinsic properties of this quantity.

In order to present some properties of relative entropy, assume that, before the interaction begins, the density matrix of the entire considered system is given by the direct product of a matrix $R_S(0)$ and the density matrix $D_B(0)$, where $R_S(0)$ is not the density matrix of the observed system $D_S(0)$. The relative entropy between the different dynamic evolutions:

a), where the entire density matrix is given by $D_S(0)$ and $D_B(0)$, and

b), where $D_S(0)$ is replaced by $R_S(0)$,

is given by:

$$S(D_S(0)\|R_S(0)) = Trace(D_S(0)\ln D_S(0) - D_S(0)\ln R_S(0)) .$$ (5.8)

Equation (5.8) shows that the relative entropy $S(D_S(0)\|R_S(0))$ is never negative and is zero only when $R_S(0)$ is equal to $D_S(0)$. It appears that an epistemic interpretation of the relative entropy is given by a measure of the information that is needed to distinguish between the two evolution processes. Also, there is a theorem [Allahverdyan and Saakian (1998); Lindblad (1975); Schlögl (1980)] which shows that with increasing time, the relative entropy never increases:

$$S(D_S(0)\|R_S(0)) \geq S(D_S(t)\|R_S(t)) \equiv S(T(t,0)D_S(0)\|T(t,0)R_S(0)) .$$ (5.9)

Inequality (5.9) is an equality only when the evolution operator is a unitary operator [which allows a time symmetric dynamics according to (5.1)] or in the case mentioned above, i.e. when $R_S(0)$ is equal to $D_S(0)$.

However, in the following I shall demonstrate how Allahverdyan and Gurzadyan (2002) use this formal time asymmetry of relative entropy in their particular account. My goal here is again to isolate the critical steps that turn artificial assumptions into physical conclusions regarding time asymmetries. This will be important for avoiding these steps in the second part, where I suggest a physical understanding of thermodynamic time direction.

V.2.2. An Epistemic Time Arrow

Equation (5.9) represents a statistical quantity used to compare two formal evolutions with different initial conditions and/or different time evolutions. Also, it can be asymmetric in time. However, this time asymmetry is on the formal level at which the formal difference between two different systems with different initial conditions is described, and not on the level of the considered system itself. Thus, (5.9) would not define a time arrow in thermodynamics because it does not refer to the description of a considered system directly; instead, it is a mathematical construct for comparing two different physical models. However, Allahverdyan and Gurzadyan (2002) mentioned different possibilities for constructing a time arrow from (5.9). I investigate these possibilities in order to show that they seem not plausible to me.

One possibility, investigated by Allahverdyan and Gurzadyan (2002), is that the heat bath can be a special system in such a way that, for sufficiently large times, the dependence of the evolution operator (5.5) on the state of the heat bath would be negligible. This means that, after a time period, the dynamics of the system described by the evolution operator (5.5) would be independent of the initial conditions of the *entire* system. Such a dynamics is called a 'dynamics without memory' (DWM) [Allahverdyan and Gurzadyan (2002)].

Additionally, there is another suggested possibility for constructing a time arrow according to (5.9) by using the statistical approximation method. This possibility arises from the density matrix itself when the density matrix of the observed system becomes a static matrix after a finite period of time:

$$T(t)D_{S;Stat} = D_{S;Stat} \, , \tag{5.10}$$

with

$$D_S(t) \xrightarrow{\;t \to t_{finite}\;} D_{S;Stat} \; . \tag{5.11}$$

To further investigate the possibilities, we will follow Allahverdyan and Gurzadyan (2002) by assuming that the evolution operator (5.5) becomes independent of the initial conditions of the system. This assumption yields an interesting possibility, which is crucial for Allahverdyan and Gurzadyan (2002). This is that we can use (5.9) for all times $\theta > 0$ to calculate the relative entropy between the DWM case and the other time asymmetric possibility, described by (5.10) and (5.11). In this case, we obtain

$$S(D_S(t)\|D_{S;Stat}) \geq S(T(\theta)D_S(t)\|T(\theta)D_{S;Stat}) \equiv S(D_S(t+\theta)\|D_{S;Stat}) \; . \tag{5.12}$$

Equation (5.12) indicates that the relative entropy $S(D_S(t)\|D_{S;Stat})$ decreases monotonically with increasing time. Thus, the difference between the two formal approximation methods decreases with increasing time. Crucially, Allahverdyan and Gurzadyan (2002) assumed that if the relative entropy goes to zero, the difference in the physical dynamics also vanishes. This, again, cannot be justified by physical considerations because the relative entropy itself is just a mathematical approximation method, and it seems obvious that an ontic interpretation of it is implausible because it is a relative value between two hypothetical evolution processes. However, according to Allahverdyan and Gurzadyan (2002), the relative entropy becomes crucial for the physical systems themselves. Their reasoning is as follows:

The exact properties of $S(D_S(t)\|D_{S;Stat})$ depend on the properties of the static density matrix $D_{S;Stat}$. Therefore, (5.12) is used to predict the entropy behaviour of the *physical system* [Allahverdyan and Gurzadyan (2002)]. The described evolutions (the DWM case as well as the 'static density matrix case') are assumed to capture the physical dynamics approximately for sufficiently large time scales. This is the point at which an epistemic property, the relative entropy, is assumed to provide physical consequences. Allahverdyan and Gurzadyan seem to overlook this questionable aspect of their account. They do not even mention the gap between the epistemic considerations and the physical conclusions:

> *'one can apply [(5.9)] for any θ as [(5.12)] and deduce that the function $\left[S\left(D_S(t)\|D_{S;Stat}\right) \right]$ is monotonically decreasing with time, since $\theta > 0$ was arbitrary.'*
> [Allahverdyan and Gurzadyan (2002), p. 7]

But the relative entropy is now used to draw physical conclusions about the time behaviour of thermodynamic systems in general.

For example and in order to stress this important aspect: Allahverdyan and Gurzadyan (2002) arguing that:

> when $D_{S;Stat}$ describes a microcanonical distribution, it follows that $D_{S;Stat} \propto 1$. With this assumption about the system, (5.12) shows that the entropy of the described system becomes, in the limit of large times, approximately the von Neumann entropy, which is given by the negative trace of the product of $D_S(t)$ and $\ln D_S(t)$. In this case, we find that the entropy of the system increases monotonically with time. This is the second law of phenomenological thermodynamics in this special quantum thermodynamic case.

However, this law holds only because we used epistemic motivated assumptions (such as the IIA and the assumption that both time asymmetric evolutions captures the physical dynamics of the system for sufficiently large times) and took epistemic considerations (the relative entropy and the distinction between the two subsystems) as evidence of physical behaviour.

Not surprisingly, the second law will not hold only if we restrict the static density matrix to a microcanonical ensemble. For example, when $D_{S;Stat}$ describes a canonical distribution, $D_{S;Stat} \propto e^{-\frac{H_S}{T}}$. (Here T is the temperature of the observed system and not the evolution operator of course.) In this case, we can describe the free energy F of the observed system as

$$F = U(t) - TS_N(t),$$ (5.13)

with

$$U(t) = Trace\big(D_S(t)H_S\big).$$ (5.14)

Thus, the free energy of the observed system decreases monotonically with increasing time, where $U(t)$ is the average energy of the system.

To summarize, (5.13), (5.14) as well as the consequence of considering a microcanonical ensemble define only *formal* time asymmetries in some statistical approximation methods. However, the entire account is based on the assumption that the approximate formal time evolution covers the physical dynamics of the described system for sufficiently large times.

Moreover, this assumption is motivated only by the fact that the relative entropy, a purely epistemic mathematical construct, goes to zero for finite times.

I believe that this is the main failure of Allahverdyan and Gurzadyan (2002), but it is very useful to see in detail how a concrete arrow of time can be explicated using the initial conditions (the IIA) and the dynamical conditions (DWM). This shall be useful in order to avoid similar reasoning's in my own proposal. Hence, I will demonstrate how a concert explication of a thermodynamic time asymmetry works in the framework of Allahverdyan and Gurzadyan (2002).

Thus, I will present a brief example that shows precisely how Allahverdyan and Gurzadyan (2002) try to apply their concept of a quantum thermodynamic time arrow to physical systems.

V.2.3. The DWM Explication

Consider, again, an observed system in interaction with a heat bath. Let us assume additionally that the interaction between the heat bath and the observed system is weak. Here, 'weak' means that the interaction is small compared to the dynamical effects of the observed system alone. Additionally, I follow Allahverdyan and Gurzadyan (2002) in assuming that the correlation function has a very short relaxation time, where 'short' means shorter than the time scales for the dynamic processes of the observed system alone. This means that correlations that are associated with the constituents of the heat bath give rise (approximately) only to processes that vanish after a time period that is very short compared to the dynamic processes of the observed system alone. Thus, the effects of correlations with the constituents of the heat bath are small compared to the effects of the dynamics of the observed system alone. Notice that this assumption focuses on the properties of the heat bath, and all of these assumptions are independent of the specific form of the interaction between the observed system and the heat bath. Nevertheless, they are typical approximations for a statistical description method.

According to the quantum dynamical description (5.1), it seems reasonable to construct the Liouville operators for the system. This simplifies the description greatly, and I think it is also helpful because of the familiarity of the description. The Liouville operators, according to (5.1), are given by

$$L_k(t) = \frac{1}{i\hbar}\left[H_k, ..\right] \text{ with } k \in \{S, B\}.$$

(5.15)

In analogy to (5.15), we can also construct the Liouville operator for the interaction $L_W(t)$ between the observed system and the heat bath. However, in this case I use the description involving the Heisenberg operator $H_W(t)$. We will see that this description makes some later arguments more obvious. The Heisenberg operator for the interaction is given by

$$H_W(t) = e^{\frac{it}{\hbar}(H_S + H_B)} H_W e^{-\frac{it}{\hbar}(H_S + H_B)}.$$

(5.16)

Now, the dynamics of the entire system is described by the density matrix. According to (5.15), the density matrix is given by

$$D(t) = e^{t(L_S + L_B)} \hat{T} e^{\int_0^t d\theta L_W(\theta)} D_B(0) \otimes D_S(0).$$

(5.17)

Additionally, according to (5.17) and (5.4) we can calculate the expectation value for an arbitrary operator χ and thus also for every observable of the observed system (again, a typical approximation in statistical physics):

$$\langle \chi \rangle = Trace_B(\chi D_B(0)).$$

(5.18)

Therefore, according to (5.17) it is possible to calculate the important expectation value $\hat{T} e^{\int_0^t d\theta L_W(\theta)}$ from (5.18). To create the most obvious analogy to classical thermodynamics, it is possible to write the expectation value $\hat{T} e^{\int_0^t d\theta L_W(\theta)}$ in the form:

$$\left\langle \hat{T} e^{\int_0^t d\theta L_W(\theta)} \right\rangle = \hat{T} e^{\int_0^t d\theta F(\theta)}.$$

(5.19)

F , of course, is an operator that could be developed in perturbation theory from (5.19). This means that we could write F as the sum over all the F_k operators. Here, F_k are the operators for every order of the perturbation theoretical description of (5.19):

$$\left\langle e^{-\frac{i}{\hbar}\int_0^t d\theta L_W(\theta)} \right\rangle = 1 + \sum_{k=1} \int_0^t d\theta_1 \int_0^{\theta_1} d\theta_2 \int_0^{\theta_2} d\theta_3 ... \int_0^{\theta_{k-1}} d\theta_k \left\langle L_W(\theta_1)L_W(\theta_2)L_W(\theta_3)...L_W(\theta_k) \right\rangle .$$
(5.20)

The first two terms in (5.20) are given by

$$F_1(t) = \left\langle L_W(t) \right\rangle$$
(5.21)

and

$$F_2(t) = \int_0^t d\theta \left(\left\langle L_W(t)L_W(\theta) \right\rangle - \left\langle L_W(t) \right\rangle \left\langle L_W(\theta) \right\rangle \right).$$
(5.22)

With the assumptions described at the beginning of this subsection and (5.20), we see that the higher orders of the perturbation theoretical description, as expected, are small compared to (5.21) and (5.22). Thus, it is possible to approximate (5.20) as the sum of (5.21) and (5.22). This yields an approximation for the dynamics of the system.

To simplify the description, we can assume that $F_1(t)$ in (5.21) is zero[37]. Allahverdyan and Gurzadyan (2002) showed that, with (5.22) and a vanishing $F_1(t)$ in (5.21), the derivation of the density matrix of the observed system is given by

$$\dot{D}_S(t) = \frac{1}{i\hbar} \left[H_S(t), D_S(t) \right] + e^{tL_S} F_2(t) e^{-tL_S} D_S(t) .$$
(5.23)

---

[37] In fact this assumption has already been made at the beginning. There it is assumed that the correlation function has a very 'short' relaxation time. We will see the connection to a vanishing $F_1(t)$ more precisely below.

Because (5.23) is a non-Markovian differential equation in time, it shows a time asymmetry. But, at least so far, the change in the value of the thermodynamic potential could be non-monotonic. This shows that the approximate dynamics of this toy system has not lost its memory. To create a crucial DWM toy model, Allahverdyan and Gurzadyan (2002) add an assumption regarding the Heisenberg operator for the interaction between the heat bath and the observed system:

$$H_W = S \otimes B,$$
(5.24)

where S and B are defined only in the Hilbert spaces of the observed system respectively the heat bath. According to (5.24), we can change the form of (5.23) and write the time derivative of the density matrix of the observed system depending on S. According to this assumption, the Heisenberg operators for the heat bath $B(t)$ and the observed system $S(t)$ can be defined using S and B:

$$S(t) = e^{it\frac{H_S}{\hbar}} S e^{-it\frac{H_S}{\hbar}}$$
(5.25)

for the observed system and

$$B(t) = e^{it\frac{H_B}{\hbar}} B e^{-it\frac{H_B}{\hbar}}$$
(5.26)

for the heat bath. Now, however, the differential equation (5.23), written in terms of (5.25) and (5.26), depends also on the correlation function $K(t,\theta)$. As we assumed above, the relaxation time τ for the constituents of the heat bath is very short. Here 'short' should mean that the constituents of the heat bath always interact with the system approximately as if no correlation with the heat bath itself occurs. Hence, we obtain for the correlation function:

$$K(t,\theta) \approx \langle B(t)\rangle \langle B(\theta)\rangle = 0; |t-\theta| \gg \tau.$$
(5.27)

Of course, this is true only as long as the assumption that $F_1(t)$ is zero is true. Hence, we see that this mathematical requirement for the perturbation was motivated by the assumptions regarding the heat bath, and even if it is an approximation, this does not seem to preclude

the drawing of physical conclusions. Thus, the assumption that $F_1(t)$ vanishes is equivalent to the physical assumption that the correlations from the heat bath have an approximately short lifetime. The formal assumption is only the formal description of the physical assumption mentioned at the beginning of this subsection. However, according to this assumption, Allahverdyan and Gurzadyan (2002) showed that the density matrix of the observed system is given approximately by

$$D_S(t) = e^{tL_{eff}} D_S(0) \,.$$
(5.28)

Here, L_{eff} is the effective Liouville operator [see also Allahverdyan and Gurzadyan (2002)]. Thus, after a time period larger than τ, (5.28) does not depend on the initial conditions of the entire system. Thus, we create a DWM case by using (5.28). The dynamics of the toy model described by (5.28) includes an arrow of time as an explication of (5.12). Nevertheless, the crucial critics regarding the more abstract proposal (5.12) are also valid regarding (5.28). The time arrow is provided by a special initial condition (the IIA) and the dynamic conditions / approximation (the DWM).

Therefore, Allahverdyan and Gurzadyan (2002) suggest an application to physics of an approximation method that can make predictions on laboratory experiments but cannot draw global physical or ontic conclusions about the time arrow in nature or in fundamental models of nature. The reason is that even if we accept the argument that leads to (5.12), the applicability to quantum thermodynamic models would depend on approximations that fulfil the DWM assumption. As in the more general case (5.12), Allahverdyan and Gurzadyan seem to overlook the crucial issue of applying an epistemic motivated assumption together with approximation methods to physical and ontic considerations:

'the thermodynamical arrow of time has been established as follows from equation [(5.12)]. The decoupling property [(5.27)] is seen to be connected with the dynamics of the free bath, see [(5.26)], and hence needs a concrete physical mechanism for its validity. The standard mechanism for this is to take a very large bath, consisting of many nearly independent pieces. Another possible mechanism is the intrinsic chaoticity of the bath, which leads to decoupling of correlators.' [Allahverdyan and Gurzadyan (2002), p. 9]

It seems fair to mention at this point that Allahverdyan and Gurzadyan (2002) also showed that the mixing process of geodesics in hyperbolic curved spacetimes can be approximately

described by their formal account.[38] Nevertheless, they have not justified the drawing of physical or ontological consequences from their epistemic considerations or epistemic motivated assumptions such as the IIA and DWM. Thus, I think it is not necessary to consider the cosmological part of their investigation here.

Moreover, to be fair regarding the suggestions of Allahverdyan and Gurzadyan (2002), I should mention that they claim only to suggest that, in a hyperbolic curved spacetime, the description method that is discussed and criticized here can be applied and then, in such spacetimes, the origin of the thermodynamic time arrow would be understandable. In fact, my own suggestion shares that motivation. However, the contrast between the two accounts is that the description method that Allahverdyan and Gurzadyan applied to hyperbolically curved spacetimes was not based on fundamental models of the physical theories but on epistemic motivated approximation methods and assumptions. This chapter criticizes this point and this point only. In contrast, the description method that I will try to apply to hyperbolically curved spacetimes is motivated by quantum cosmological models alone, and the time asymmetry will be given in terms of the exaptation value of the particle number operator.

Additionally, I must mention that both the assumptions and the statistical approximation methods are, of course, useful and powerful tools for physics. My aim was never to criticise the usefulness of this tools. It is, of course, reasonable to use them to predict the behaviour of a system. However, the mathematical structure that is produced by the assumptions and the approximations is not valid for drawing physical or ontic conclusions regarding the involved dynamics in general; especially if, as it seems to be the case, the goal is an explanation for the occurring of an ontic thermodynamic arrow of time, which is based on physical properties.

---

[38] More precisely, Allahverdyan and Gurzadyan (2002) showed that the negative curvature in a Friedmann–Robertson–Walker universe and the effect of geodesic mixing can fulfill the approximate condition necessary for the emergence of the thermodynamic arrow of time in theirs account. Moreover, this mechanism can describe a situation in which the cosmic microwave background contains a major fraction of the entropy of the universe. If this were the origin of the thermodynamic arrow, the thermodynamics in flat and positively curved universes need not be strongly time asymmetric, and we observe this situation because we happen to live in a universe having negative curvature. Thus, as I shall explicate in the next section, I agree with Allahverdyan and Gurzadyan (2002) regarding the *connection* between the hyperbolic curvature of spacetime and the dynamics and boundaries that are required to provide an arrow of time in thermodynamics. However, I do not think that their investigation can describe this connection because their description method simply cannot explain why epistemic motivated approximation methods and initial assumptions can be used to draw ontic conclusions regarding thermodynamic and cosmic time directions, even if the approximations are 'good' for some cosmological models, there are still purely descriptive approximations.

In contrast to the criticised account, I shall propose an alternative understanding of the entropic time asymmetry in the next section.

V.3. On an Physical Time Arrow in Thermodynamics

This section aims to demonstrate that, for cosmological reasons, the Landau entropy as well as a possible definition of phenomenological entropy obeys the second law of thermodynamics in hyperbolically curved spacetimes (or empirically equivalent spacetimes). Those spacetimes can reasonably be considered as candidates for a physical description of our particular universe. Additionally, as mentioned, I shall argue that a quantum cosmological description of such a spacetime includes a time asymmetry as an explication of the fundamental time asymmetry from chapter III. In addition, this section shows that, according to quantum cosmology, the second law of thermodynamics is not valid in all spacetime regions, but it describes the time behaviour of the Landau entropy and the phenomenological entropy correctly if the scale factor a is described by $a \propto t^{\alpha}$ for $\alpha < 2$, which seems the case in most phases of our cosmic evolution. Moreover, by developing my thermodynamic approach, I will try to show that neither the IIA nor the DWM is needed to create a thermodynamic arrow of time on the basis of physical considerations. This, I think, is demonstrated by Castagninio and Laciana (2002) and this section aims to clarify the philosophical understanding of the thermodynamic time asymmetry sketched by them. Also, in the proposed account the time asymmetric behaviour of entropy will be provided by the physical dynamics of the system, at least in particular parts of the considered spacetimes.

The next two subsections will describe the quantum cosmological model that I shall use and which seems to be a valid candidate for an appropriate semi-classical description of our particular universe.

V.3.1. On the Arrow of Time in Quantum Cosmology

It is important to describe the type of thermodynamic arrow that I will suggest. I will argue that an arrow of time in quantum thermodynamics is caused by a cosmological time asymmetry in hyperbolical curved spacetimes. This cosmological time asymmetry, as I will demonstrate, is understandable as an explication of the fundamental time asymmetry in the solution set of Einsteins' equation (described in chapter III). However, it can provide the arrow of time in thermodynamics only in those spacetime regions that are not too 'close' to the Big Bang (according to cosmic time).

The section is structured as follows:

> Firstly, I discuss the quantum cosmological model. I will argue that in the used model, we find an explication of the fundamental cosmological time asymmetry (chapter III). Also, I will restrict the considered spacetimes to hyperbolic curved spacetimes with a positive cosmological constant (or empirical equivalent spacetime descriptions) and finite expectation value of the energy–momentum tensor. Both considerations are motivated by physical considerations, and both assumptions seem to be fulfilled in our particular spacetime.

> Secondly, for the restricted set of spacetimes, I will show that the expectation value of the particle number increases with cosmic time. This occurs only for $a \propto t^{\alpha}$ with $\alpha < 2$. Thus, the time asymmetry does not appear in very early stages of the universe. Nevertheless, it will be a necessarily occurring time asymmetry in the set of considered quantum cosmological models.

> Thirdly, on the basis of this time asymmetry of the expectation value of the particle number, I will show that the Landau entropy as well as a particular phenomenological entropy behaves asymmetrically in cosmic time. Moreover, the second law, according to those entropies, can be formulated as explicitly asymmetric in time. This means that the entropy value would not only increase with increasing time but also decrease with decreasing time (apart from fluctuations and within the boundaries of maximum and minimum values).

Hence, I shall begin with describing the quantum cosmological model that I shall use in this section.

V.3.2. On the Quantum Cosmological Model

In this subsection, I present a simplified version of the quantum cosmological description of spacetimes. The simplifications are only mathematical, and they do not introduce time asymmetries. At the end of the chapter, I will return to the crucial simplifications. There I will argue that the arguments presented in the simplified model are also valid in sophisticated models and that my conclusions do not depend on the most simplifications in the model. To check the plausibility of the simplifications, see, for example, Castagnino, Giacomini and Lara (2000) and (2001), for a more detailed introduction in the model see Cstagnino and Laciana (2002). The simplified model that I will use assumes that spacetime allows the definition of cosmic time, which is crucial also for applying the considerations from chapter III. The justification of this assumption, of course, is therefore identical to the one given in chapter III.

Hence, it seems possible to use a Friedman Robertson metric in this model:

$$ds^2 = dt^2 - a^2(t)\left(dx^2 + dy^2 + dz^2\right).$$

(5.29)

As argued in chapter III, the fundamentality of the basic time asymmetry in the solutions space of the crucial dynamic equations is still well established especially under those assumptions.

Also, the basic mathematical reasoning from chapter III can be applied to quantum cosmology as well, as long as the expectation values of the fields are assumed to be finite, which seems to be a physically motivated assumption. Quantum effects such as fluctuations in the fields can be described as additional parameters in the Hamiltonian of the entire spacetime. Thus, one of the claims of this chapter can be established reasonably quickly by referring to the reasoning in chapter III. Given the fact that the argumentation in chapter III can be applied to quantum cosmology as well (as long as gravity is not quantised), we find that, according to spacetimes which makes the definability of cosmic time possible, almost all[39] solutions of the associated dynamic equations include an intrinsic time asymmetry of cosmic time (see chapter III and the independence of the consideration from the particular form of the Hamiltonian).

---

[39] In the understanding from definition I in chapter II.

However, to make my claim regarding thermodynamic time directions, it appears necessary to add some crucial conditions on the spacetimes under consideration. This dialectical structure will be similar to that in chapter IV, where the arrow of radiation is considered in spacetimes similar to ours. The main difference, regarding the two structures, will be that the thermodynamic arrow, as an arrow regarding cosmic time, can be understood as a direct consequence of the explications of the fundamental time asymmetry from chapter III. In chapter IV, the additional conditions were assumptions regarding the type of energy–momentum tensor. In this quantum cosmological case, the crucial additional condition focuses on the matter and energy dynamics in spacetime. The condition is that it is possible to describe the energy / matter distribution and dynamics by a scalar matter field ϕ. This seems reasonable in our particular spacetime, even though other descriptions are not ruled out by astrophysical considerations (as far as I know). Thus, by assuming these conditions I restrict the set of spacetimes in which the results are valid to the set of spacetimes in which the matter and energy dynamics is described by a scalar field. Hence, if the claim of deducing the thermodynamic time directions from the cosmological time asymmetry (in the restricted set of considered spacetimes) were successful, the thermodynamic time direction would not itself be fundamental but would be a necessarily occurring by product of the cosmological time asymmetry.

Regarding the detailed consideration: The total action in the set of considered universes is given as:

$$S = S_\phi + S_g \, .$$

(5.30)

Here, the action of the gravitational field is characterized by the Einstein–Hilbert action. Additionally, I restrict the Einstein–Hilbert action to the case where the cosmological constant Λ is assumed to be positive or zero [see also Bergström and Goobar (1999) for the purely physical motivation for this restriction]. Thus, it seems that the set of considered spacetimes can again include our particular spacetime.

Here I shall begin the detailed analysis. The part of the total action given by the scalar matter field is expressed by

$$S_\phi = \frac{1}{2} \int dx^4 \sqrt{-g} \left(g^{\mu\nu} \phi_\mu \phi_\nu - \left(m^2 + \frac{R}{6} \right) \phi^2 \right).$$

(5.31)

The variation in the total action (5.30) yields the Einstein equations [Castagnino and Laciana (2002)]:

$$\frac{2\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} = -8\pi G p + 3\Lambda \quad \text{and} \quad \frac{\dot{a}^2}{a^2} = \frac{8\pi}{3} G \rho + \Lambda \ . \tag{5.32}$$

Note that the quantum cosmological description that I offer and that is often applied in cosmology is semi-classical. In fact, spacetime itself is treated as a classical spacetime, but all fields in spacetime, including the scalar matter field, are treated as quantum fields. In some ways, this description appears to be slightly unsatisfactory, but it seems the only possibility if the cosmological description shall not include speculations about quantum gravity. Therefore, I think it is acceptable to consider a semi-classical description here.

If we take the fields in spacetime as quantum fields, we must deal with the expectation values of the energy–momentum tensor $\langle T_\nu^\mu \rangle$. In an effective renormalized model, $\langle T_i^j \rangle$ is given by the product of the cosmic pressure p and the delta distribution δ_i^j. Moreover, $\langle T_0^0 \rangle$ is identical to the cosmic energy density ρ. If we assume that the expectation values of the energy–momentum tensor are finite or zero, which seems physically plausible, the two Einstein equations (5.32) become connected and are no longer independent. This was shown by Fulling and Parker (1974); they also showed that the Einstein equations (5.32) are related by the equation:

$$p = -\frac{1}{3a^2\dot{a}} \frac{d}{dt} (a^3 \rho) \ . \tag{5.33}$$

Thus, it is sufficient to solve only one of the Einstein equations (5.32). For simplicity, I consider the second Einstein equation in (5.32). I define the Hubble coefficient as usual by $H = \frac{\dot{a}}{a}$. In (5.32), the rhs of the equation must be positive because the lhs is given by the Hubble coefficient squared. For the same reason, the square root of the rhs of the second Einstein equation in (5.32) must be a real valued function of cosmic time.

The rhs of the second Einstein equation in (5.32) will be called γ^2. Thus, the equation becomes:

$$H^2 = \gamma^2.$$ (5.34)

Equation (5.34) provides two solutions for the Hubble coefficient: $\pm\gamma$. As mentioned above, γ is either positive or zero. Hence, if we do not consider a static spacetime, γ must be a positive real function. As in the classical case, the static solution belongs to a subset of cosmological solution that has a lower dimension than the entire considered solution set of the Einstein equations, so it is not considered here. Thus, we obtain two time asymmetric (because of the accelerated expanding in realistic quantum cosmological models) solutions of the Einstein equation, one according to γ and the other according to $-\gamma$.

Traditionally, the $-\gamma$ solution is understood as describing a hyperbolically curved collapsing universe because the Hubble coefficient is negative. In this understanding, the solution according to γ describes an expanding universe because of the positive value of the Hubble coefficient. However, both solutions describe an entire spacetime, and for fundamental reasons in general relativity (or empirically equivalent spacetime theories), there is no absolute time parameter or other physical value outside of spacetime that can be used to define external relations to spacetimes. Moreover, the solutions are time reversed geometrical objects of each other; thus, there are no intrinsic structures that differ from one solution to the other. Therefore, the solutions are identical in all intrinsic properties and in all external relations. This is a fundamental consequence of general relativity (or empirically equivalent spacetime theories), and according to this Leibniz reasoning applied in chapter III, the conclusion is that both solutions of the Einstein equation describe the same physical spacetime. Thus, the Einstein equation (5.34) describes a time asymmetry in the considered set of possible spacetimes. This time asymmetry is the explication of the fundamental time asymmetry from chapter III in this particular quantum cosmological model.

In the following subsections, my aim is to show that this time asymmetric solution of the Einstein equations provides a thermodynamic time arrow. To do so, I shall demonstrate that the expectation value of the particle number operator shows time asymmetric behaviour according to cosmic time. This, I think, is also interesting because it is a necessarily occurring time asymmetry in QFT descriptions of the considered set of quantum cosmological models.

V.3.3. On Time Asymmetric Behaviours of the Particle Number Operator

In this subsection, I investigate the time behaviour of the particle number operator with respect to cosmic time by following mostly the physical analysis of Castagnino and Laciana (2002). During this presentation of the analysis from Castagnino and Laciana I shall also develop the crucial philosophical advantages of this time asymmetry.

I will begin with the action of the scalar matter field in (5.31). The variation in the total action (5.30) allows us to deduce the field equation for the scalar matter field ϕ :[40]

$$\nabla^{\mu}\partial_{\mu}\phi + \left(m^2 + \frac{R}{6}\right)\phi = 0. \tag{5.35}$$

Parker (1969) showed that, according to (5.35), the field operator is given by

$$\phi(x) = \frac{1}{(La)^{\frac{3}{2}}}\sum_{k} A_k h_k(t)e^{ikx} + A_{-k}^{\dagger} h_k^*(t)e^{-ikx}. \tag{5.36}$$

Here, A_k is the annihilation operator and $A_{-k}^{\dagger}$ is the creation operator. As typical I denote the vacuum state as $|0\rangle$. In this simplified model, the vacuum state is assumed to be analogous to the classical vacuum in which space is empty, I shall return to this point much more detailed and I shall argue that the conclusion does not depend on this simplification.

A_k applied to the vacuum state $|0\rangle$ yields zero, whereas $A_{-k}^{\dagger}$ applied to $|0\rangle$ creates the one-particle state $|1_k\rangle$, where k denotes the quantum momentum number. The annihilation and creation operators also fulfil the usual commutation relations:

$$\left[A_k, A_{k'}\right] = \left[A_k^{\dagger}, A_{k'}^{\dagger}\right] = 0 \text{ and } \left[A_k, A_{k'}^{\dagger}\right] = \delta_{k,k'}. \tag{5.37}$$

Equations (5.35) and (5.36) provide a differential equation for $h_k(t)$, which is the crucial function if we want to analyse the time behaviour of the field operator (5.36):

---

[40] The used field equation does not consider self-interaction of the quantum field. This is one simplification of the model that I will use here but I will return to that issue at the end of this section. There I shall argue that the consideration of self-interactions of the matter field should not change the conclusion drawn in the simplified model.

$$\ddot{h}_k(t) + Q_k^2(t)h_k(t) = 0 \ . \tag{5.38}$$

In (5.38) $Q_k^2(t)$ is a short version of the sum of ϖ_k^2 , $-\frac{1}{4}H^2$ and $\frac{1}{2}\dot{H}$, and the frequency ϖ_k^2 is given by $\varpi_k^2 = \frac{k^2}{a^2} + m^2$. Castagnino, Harari and Núñez (1987) provided a general solution for (5.38):

$$h_k(t) = \frac{1}{\sqrt{2\varpi_k}} \alpha_k(t)\exp\left(-i\int_{t_0}^{t}\varpi_k(t')dt'\right) + \frac{1}{\sqrt{2\varpi_k}} \beta_k(t)\exp\left(i\int_{t_0}^{t}\varpi_k(t')dt'\right) . \tag{5.39}$$

For effective utilization of (5.39) in the context of a cosmological model, it is necessary to assume boundary conditions. Therefore, of cause, also Castagnino, Harari and Núñez (1987) used boundary conditions to apply their results. Thus, before we can simply use their results, we have to make sure that these boundary conditions neither include *independent* (from the cosmological time asymmetry) time asymmetries nor are motivated by epistemic considerations. If that were the case, the time asymmetry that I would like to demonstrate would depend on boundary conditions similar to the IIA or DMW assumptions described in the first part of this chapter.

The mathematical boundary conditions that Castagnino, Harari and Núñez (1987) assumed are directly associated with physical assumptions, which seem critical. Moreover, they had to use timelike boundary conditions; thus, they used initial conditions for (5.38). The crucial initial condition was:

$$\alpha_k(t_0) = 1 \text{ and } \beta_k(t_0) = 0 . \tag{5.40}$$

However, the interesting issue in this investigation is the justification of (5.40) for *physical* reasons. As we will see, one closer look at the mathematics is essential to confirm that the initial conditions (5.40) will not provide additional time asymmetries beyond the cosmological one. I shall argue that the assumptions are, in fact, not critical; they do not introduce *additional* time asymmetries beyond the cosmological one, nor is their motivation artificial or purely epistemic. Nevertheless, the simplest way to show this is by demonstrating the mathematical reasoning in a bit more detail. For this purpose, we can use (5.36) to obtain:

$$\phi(x) = \frac{1}{(La)^{\frac{3}{2}}} \sum_k A_k(t) h_{0k}(t) e^{ikx} + A^{\dagger}_{-k}(t) h^*_{0k}(t) e^{-ikx} . \tag{5.41}$$

Here, $h_{0k}(t)$ is given by

$$h_{0k}(t) = \frac{1}{\sqrt{2\varpi_k}} \exp\left(-i\int_{t_0}^{t} \varpi_k(t') dt'\right). \tag{5.42}$$

The operators $A_k(t)$ and $A^{\dagger}_{-k}(t)$ are determined by the products of $\alpha_k(t) A_k$, $\beta^*_k(t) A^{\dagger}_k$, $\alpha^*_k(t) A^{\dagger}_k$ and $\beta_k(t) A_{-k}$. However, for the operators $A_k(t)$ and $A^{\dagger}_{-k}(t)$, the usual commutation relation

$$\left[A_k(t), A_{k'}(t)\right] = \left[A^{\dagger}_k(t), A^{\dagger}_{k'}(t)\right] = 0 \text{ and } \left[A_k(t), A^{\dagger}_{k'}(t)\right] = \delta_{k,k'} \tag{5.43}$$

holds only if

$$|\alpha_k|^2 - |\beta_k|^2 = 1 \tag{5.44}$$

holds. Thus, only by using (5.44), we can make the necessary identification of the operator $A_k(t_0)$ with the annihilation operator.

Additionally and in order to simplify the arguments for the acceptability of the boundaries (5.40), I will consider an eigenvector of the particle number operator as the considered state of the scalar matter field ϕ at the initial time point t_0. This is, of course, a special choice, but it can be made without loss of generality for this argument, because the eigenvector of the particle number operator is unstable under time evolution in both time directions. Thus, it will not eliminate interesting philosophical questions. This can be understood as follows:

    i) We could assume a stable initial state for the scalar matter field, but this choice corresponds to a fine tuned situation that does not seem to describe our particular spacetime and, more importantly, corresponds to a low dimensional solution set of

the crucial Einstein equations in this model. Thus, I will not consider it for the same reason that I ignored the static solution of the Einstein equations.

ii) We could choose an unstable initial state for the scalar matter field that is not given as an eigenvector of the particle number operator. However, this choice will not advance the discussion. The state would be unstable under time evolution in both time directions, and the only thing that would change is that the mathematical description would become more complex.

Thus, the assumption of an eigenvector of the particle number operator at the initial time point is not critical. In fact, the only consequence of this choice is to simplify the mathematical reasoning.

Before returning to the initial assumption (5.40), I have to make a special choice regarding the initial state of the considered spacetime. But, this choice will be motivated only by the cosmological model. Given the Einstein Equations (5.32), we are, of course, in a Big bang scenario. Hence, the initial state of the universe is considered to be a quantum vacuum state. This, of course, is a time asymmetric choice, but it is a direct reasonable consequence of the cosmological time asymmetry in the model. Therefore, it can be seen as a consequence of the explication from the fundamental time asymmetry (in chapter III) in this particular quantum cosmological model. Hence, this choice is not only acceptable but motivated by cosmology and understandable as an explication from the abstract cosmological time asymmetry investigated in chapter III.

With this in mind I shall return to the initial assumption (5.40). However, it is useful, as I shall show now, to investigate the mentioned assumption of a *particular* vacuum state $|0\rangle$ as the initial state. Although another vacuum definition than the one used here (equivalent to the classical vacuum) could be assumed, we will see later that the time asymmetry of the particle number operator would also arise with other initial vacuum states. This will be shown more precisely at the end of this subsection. For the moment, I shall simply accept this choice and present the generalisation to other initial vacuum states after I present the simplest form of the argument. Now, to return to the question of the initial condition (5.40), it is useful to discuss some properties of the vacuum in QFT, which will demonstrate the direct connection between the initial assumption (5.40) and the vacuum choice mentioned above.

In QFT, the vacuum state $|0\rangle$ is observer dependent, like every other state in any description that respected special relativity. For simplicity, I use only the formal description of the vacuum state $|0\rangle$ for a comoving observer. As in classical special relativity, this will not

change the physical content of the description but will provide a formal simplification. Now, the chosen vacuum state becomes important for (5.40) if we consider the formal proof from Castagnino, Harari and Núñez (1987), which shows that this classical vacuum is equivalent to the diagonalization of the corresponding Hamiltonian. Thus, the description in a comoving framework with a classical vacuum is equivalent to the basis choice in which the Hamiltonian becomes diagonal. Moreover, Castagnino, Harari and Núñez (1987) showed additionally that the initial condition (5.40) is equivalent to the fixing of the basis for the Hamiltonian in the vacuum state $|0\rangle$ [see also Castagnino and Laciana (2002)].

This type of assumption, unlike the IIA or DWM, will not introduce any asymmetry beyond the already constituted cosmological time asymmetry, which forces the initial state to be a vacuum state.[41] Moreover, it is not an approximation method, although it does simplify the formal argument.

To see this more formally, the initial conditions (5.40) define the functions h_k and h_k^*. These functions are exactly those that define the basis of the Hamiltonian. Thus, the implications from the initial conditions (5.40) will not change the physical content of the formal description beyond the restriction that are already necessary for cosmological reasons (the assumption of an initial vacuum state), which is significantly different from the IIA and DWM. Hence, there are acceptable in this investigation.

I shall return now to the main topic of this subsection: the question of how the expectation value of the particle number operator, as a function of cosmic time, includes time asymmetries as a consequence of the time asymmetry of the considered spacetimes. To answer this question, consider the expectation value of the particle number operator in mode k at the cosmic time point t and in the finite volume $\left(La(t) \right)^3$:

$$n_k(t) = \left\langle 0 \left| A_k^\dagger(t) A_k(t) \right| 0 \right\rangle. \tag{5.45}$$

If we use (5.37) together with (5.39), (5.40) and (5.45), we obtain

$$n_k(t) = \left| \beta_k(t) \right|^2. \tag{5.46}$$

---

[41] The issue of the vacuum state is considered later. Here, the vacuum is simply assumed to be analog to the classical vacuum.

However, according to (5.36), Parker (1969) showed that the field operator for the scalar matter field satisfies the usual commutation relations only if the Wronski condition is satisfied. This is the following condition:

$$\dot{h}_k^* h_k - h_k^* \dot{h}_k = i .$$

(5.47)

Fortunately, with (5.39) and (5.40), we can directly prove that (5.47) is satisfied. Thus, the Wronski condition yields the field operators to the usual commutation relation. Moreover, Parker (1969) also showed that α_k and β_k can be parameterized in the following useful way:

$$\alpha_k = e^{-i\gamma_\alpha} \cosh\theta_k \text{ and } \beta_k = e^{i\gamma_\beta} \sinh\theta_k .$$

(5.48)

Using (5.48), we can deduce a set of differential equations for γ_α, γ_β and θ. Castagnino and Laciana (2002) showed that it is useful to introduce Γ into the equations, where Γ is given by $\Gamma = \gamma_\alpha + \gamma_\beta + 2\int_{t_0}^{t} \varpi \, dt'$.

Then, the set of crucial differential equations can be written in the following form [Castagnino and Laciana (2002)]:

$$\dot{\theta} = \frac{m^2 H}{2\varpi^2} \cos\Gamma ,$$

$$-\dot{\gamma}_\alpha \cosh\theta = \frac{m^2 H}{2\varpi^2} \sinh\theta \sin\Gamma ,$$

$$-\dot{\gamma}_\beta \sinh\theta = \frac{m^2 H}{2\varpi^2} \cosh\theta \sin\Gamma .$$

(5.49)

To simplify the differential equations, we can also introduce $\mu_k = \dfrac{m^2 H}{2\varpi_k^2}$.

Moreover, to eliminate θ_k, we can introduce χ_k with $\chi_k = \sinh\theta_k$. Then (5.49) directly provides a differential equation for χ :

$$\ddot{\chi} - \frac{\chi\dot{\chi}^2}{1+\chi^2} - \frac{\dot{\mu}}{\mu}\dot{\chi} + 2\varpi\sqrt{\mu^2(1+\chi^2)-\dot{\chi}^2} - \frac{1+2\chi^2}{\chi(1+\chi^2)}(\mu^2(1+\chi^2)-\dot{\chi}^2) = 0. \tag{5.50}$$

where $\sqrt{\mu^2(1+\chi^2)-\dot{\chi}^2}$ must be a real number. Thus, the term under the square root has to be positive or zero. First, I will consider the simple case in which $\sqrt{\mu^2(1+\chi^2)-\dot{\chi}^2}$ is identical to zero. In this case, (5.50) can be simplified using $\dot{\chi}_\pm$, which is given as the product of $\pm\mu$ and $\sqrt{1+\chi^2}$. Thus, we get

$$\ddot{\chi}_\pm - \chi_\pm\mu^2 \mp \dot{\mu}\sqrt{1+\chi_\pm^2} = 0. \tag{5.51}$$

We can obtain a particular solution of (5.50) by integrating $\pm\mu$ and $\sqrt{1+\chi^2}$. To simplify the integration, we can use $\kappa = \dfrac{k^2}{m^2}$ and $\sigma = a^2$. In this case, we get

$$\mu = \frac{\sigma H}{2(\sigma+\kappa)}. \tag{5.52}$$

It is also useful to substitute the derivation $\dfrac{d}{dt}$ for the derivation with respect to σ. If we do so, we see that $\dfrac{d}{dt}$ is given by the product of σH and the derivation $\dfrac{d}{d\sigma}$. Thus, the equation that must be integrated becomes

$$\frac{1}{\sqrt{1+\chi_\pm^2}}d\chi_\pm = \pm\frac{1}{4(\sigma+\kappa)}d\sigma. \tag{5.53}$$

Integrating (5.53) yields

$$\ln\left(\chi_\pm + \sqrt{1+\chi_\pm^2}\right) = \pm\frac{1}{4}\ln(\sigma+\kappa) + \ln C. \tag{5.54}$$

Here, C is the integration constant. The integration of (5.53) yields two particular solutions of (5.50). I shall demonstrate that it is of particular interest to investigate these solutions because, in this case, the particle number n_k is given by $n_k = \chi_k^2$, and we find

$$n_k = \frac{1}{4}\left(C_k \sqrt[4]{a^2 + \frac{k^2}{m^2}} - \frac{1}{C_k} \frac{1}{\sqrt[4]{a^2 + \frac{k^2}{m^2}}} \right)^2 . \tag{5.55}$$

According to (5.55), we can draw the following conclusions:

i) If no additional cosmological condition is used, n_k is an increasing function with the scale factor a iff the following inequality holds:

$$C_k \sqrt[4]{a^2 + \frac{k^2}{m^2}} \gg \frac{1}{C_k} \frac{1}{\sqrt[4]{a^2 + \frac{k^2}{m^2}}} . \tag{5.56}$$

ii) Moreover and more importantly, in the considered case (5.40) (as motivated by the cosmological model), n_k is an increasing function with cosmic time for $t > t_0$. This is easily obtained as follows: The initial state of the universes was considered to be a vacuum state (motivated by the cosmological models according to our particular spacetime). Thus, for $t = t_0$, for the particular vacuum choice, the universe is considered to be empty. This can be formulated as follows:

$$\forall k \in \mathbb{R} : n_k(t_0) = 0 \tag{5.57}$$

where $\mathbb{R}$ denotes real numbers. With this, we find

$$C_k^0 = \pm \frac{1}{\sqrt[4]{a^2(t_0) + \frac{k^2}{m^2}}} . \tag{5.58}$$

With this expression for C_k, we find, for (5.55), that the particle number n_k is an increasing function with the scale factor a for all cosmic times $t \geq t_0$. The generalisation to other initial vacuum states will be considered later.

However, (5.55) is only a *particular* solution for the particle number. Thus, I shall argue that every other solution $\tilde{n}_k$ provides a similar kind of time asymmetry. This can be obtained from the integration of the inequalities

$$\frac{d\tilde{\chi}_\pm}{d\sigma} \leq \pm \frac{\sqrt{1+\tilde{\chi}_\pm^2}}{4(\sigma+\kappa)} \quad \text{and}$$

$$\frac{1}{\sqrt{1+\tilde{\chi}_\pm^2}} d\tilde{\chi}_\pm \leq \pm \frac{1}{4(\sigma+\kappa)} d\sigma . \tag{5.59}$$

The equation set (5.59) is a direct result of (5.50) if $\sqrt{\mu^2(1+\chi^2)-\dot{\chi}^2}$ is a positive real number; (5.52) is also used here. The simplest way to create other solutions of (5.50) is to analyse χ in (5.50) itself. For this task, it is important to remember that the functional behaviour of $\chi^2(\sigma)$ is a direct analogy to the functional behaviour of the particle number $n(\sigma)$. To analyse the behaviour of χ in (5.50), we can substitute (5.50) by a differential equation in which all derivations depend on σ :

$$\frac{d^2\tilde{\chi}}{d\sigma^2} + \frac{m^2}{\varpi^2\sigma}\frac{d\tilde{\chi}}{d\sigma} - \frac{\tilde{\chi}}{1+\tilde{\chi}^2}\left(\frac{d\tilde{\chi}}{d\sigma}\right)^2 + \frac{\varpi}{2H^2\sigma^2}\sqrt{f} - \frac{1}{4H^2\sigma^2}\frac{1+2\tilde{\chi}^2}{\tilde{\chi}(1+\tilde{\chi})}f = 0 . \tag{5.60}$$

Here, f is a function that is required so that (5.50) and (5.60) are equivalent. Castagnino and Laciana (2002) showed that this function is unambiguously defined by the quantities μ^2 , $\tilde{\chi}^2$ and $\left(\frac{d\tilde{\chi}}{dt}\right)^2$. Moreover, Castagnino and Laciana (2002) show that:

$$f \geq 0 . \tag{5.61}$$

is always the case. The particular solution of (5.50), considered above, corresponds to the case where $f = 0$ in (5.61). In addition, Castagnino and Laciana (2002) showed that (5.61) is equivalent to the following inequality:

$$\left(\frac{d\tilde{\chi}}{d\sigma}\right)^2 \leq \frac{1+\tilde{\chi}^2}{16(\sigma+\kappa)^2}. \tag{5.62}$$

For this investigation, it is reasonable (according to the considered cosmological model) to consider the case $a^2 = \sigma \to \infty$, which clearly indicates that $\tilde{\chi}^2 \leq \chi^2$ holds. Hence, we obtain the following equation:

$$f \leq \frac{H^2}{4}\left(1+\chi^2\right) \approx \frac{1}{64}\frac{\dot{a}^2}{a}C_k^2. \tag{5.63}$$

Here the approximation is given by $n \approx \chi^2$. Moreover, by considering this, we obtain

$$f \leq lH^2 a, \tag{5.64}$$

where l is a positive constant. If we also consider the case $\dfrac{\dot{a}^2}{a} \xrightarrow{t\to\infty} 0$, it follows that $f \xrightarrow{t\to\infty} 0$ holds. Thus, if $\sigma \to \infty$ holds, all solutions converge to the particular solution investigated above. Moreover, because each solution of (5.50), $\tilde{n}_k$, is smaller than n_k but positive, the particular solution investigated above is the limit solution of (5.50). Thus, the limit solution of the general problem (5.50) is always described by $\tilde{n} \xrightarrow{\sigma\to\infty} n$. Moreover, $n \xrightarrow{\frac{d\chi}{d\sigma}\to 0} const.$ clearly holds for $n = \chi^2$. This does not contradict (5.60), because in this limit, $\dfrac{d^2\chi}{d\sigma^2}$ also becomes zero.

Hence, the set of other solutions, because there show a similar functional behaviour [see Castagnino and Laciana (2002)] and there converge, in the case $\dfrac{\dot{a}^2}{a} \xrightarrow{t\to\infty} 0$, to the investigated limit solution, shows a similar kind of time asymmetry for the particle number operator, even if the particular properties (for small times) of the asymmetry depending on the chosen solution of (5.60).

However, a short look to the involved mathematics shows that the time asymmetry is not given independently from the cosmological behaviour of the considered spacetimes. To simplify the mathematics, it is possible, without loss of generality here, to consider only the

case $\Lambda = 0$. Therefore, with the scale factor $a(t) = t^\alpha$, we obtain $\dfrac{\dot{a}^2}{a} = H^2 a = \dfrac{a^2}{t^{2-\alpha}} = t^{\alpha-2}$.

Thus, to make sure that $\dfrac{\dot{a}^2}{a} \xrightarrow{t \to \infty} 0$ is satisfied, it is necessary that $\alpha < 2$.

In a simple quantum cosmological model of the evolution of the universe, this condition is fulfilled only in some special phases of cosmic evolution. For example, in a universe that is in the radiation-dominated or matter-dominated phase, there we find $\alpha = \dfrac{1}{2}$ respectively $\alpha = \dfrac{2}{3}$. Thus, the conclusion of this somewhat technical discussion is that the particle number is an increasing function of cosmic time in the radiation- and matter-dominated phases in the considered particular hyperbolic curved spacetime model. Thus, we saw the existence of an arrow, with respect to cosmic time, that is given in all semi-classical spacetimes that have

a) A radiation- or matter-dominated phase in their cosmic evolution,

b) That are hyperbolic curved.

c) In which the total action is describable as the sum of the Einstein-Hilbert Action and the contribution from a *scalar* matter field and

d) In which the field operators have the considered form.

e) In which the initial vacuum state can be described in analogy to a classical vacuum.

However, it seems that there are some crucial issues in the conditions that yield to this conclusion. The conditions a) and b) seem uncritical, even if those conditions restrict the set of spacetimes in which the conclusion can be arrived, the restricted set of spacetimes, so it seems, can include our particular spacetime. One more crucial concern is that the model was simplified not only by assuming a *scalar* matter field but also by excluding self-interactions (condition d) and presupposing a particular initial vacuum state (condition e). The first point was already mentioned, and given modern cosmological models, it can be defended in such a way that the assumption of a scalar matter field seems to agree with the astrophysical observations, even if that does not rule out other forms of cosmic matter fields.[42] However, condition d), cannot be defended in a similar way.

---

[42] Investigating the generalization to cosmological models that include other forms of matter fields seem outside the scope of this investigation. Thus, this issue is left for further mathematical, physical and philosophical investigations.

The reason for excluding self-interactions was to make the technical analysis as simple as possible. However, to make the conclusion plausible, it is crucial to generalize the above discussion to the case of self-interacting matter fields. Fortunately, Audretsch and Spangehl (1987) showed that self-interaction only amplifies particle creation in a certain way (see Audretsch and Spangehl (1987)). Thus, the conclusion that the expectation value of the particle number increases with cosmic time in a radiation- or matter-dominated cosmic epoch would not be changed by introducing amplification due to self-interactions. The amplifications would behave similar to an increasing (with cosmic time) offset value. Thus, this issue is already eliminated by Audretsch and Spangehl (1987), at least for the set of reasonably considered models.

The last crucial assumption that was included in the analysis and *must* be justified is that we chose the initial vacuum state of the universe as equivalent to a classical vacuum where the universe is empty (condition e). In fact this particular choice for the initial vacuum state is in contradiction to the motivation for considering even any vacuum state as an initial state of the universe. Earlier I argued that, in the context of the considered quantum cosmological model, we have to assume a vacuum state as the initial state of the universe, because the cosmological model requires such an initial state in Big Bang scenarios. However, in quantum cosmological descriptions, using QFT, the vacuum cannot simply be in analogy to the classical vacuum. This, of course, is a fundamental lesson, empirically well established, from QFT. Thus, if we take QFT seriously, especially in the cosmological context, the investigated vacuum choice seems extremely questionable.

Very fortunately, regarding this issue, I can refer to Laciana (1998), who showed that other more realistic vacuum states can only amplify particle creation in comparison to the vacuum state that is analogous to the classical vacuum. In fact the involved structure is very similar to the one involved by amplifications due self-interactions [compare Audretsch and Spangehl (1987) with Laciana (1998)]. Considering more realistic vacuum states would not change the time asymmetry of the particle number operator, because the amplification of particle creation would lead to similar structures as the amplification due self-interactions. Thus, this assumption, condition e), is also not critical, and the conclusion that the particle number increases with cosmic time in all cosmic phases that are radiation or matter dominated will not change if we consider a vacuum state that includes quantum fluctuations and vacuum polarisation.

Therefore, those crucial simplifications in the considered model are not relevant for the reached general conclusion, even if the detailed physical properties of the time asymmetric behaviour of the particle number operator are depending strongly on the considered vacuum and the self-interaction.

However, there are, as we saw, exceptions in cosmic evolution, where the particle number operator will not provide the time asymmetry presented here. These exceptions include, for example, the inflation phase of cosmic evolution. This is not too surprising because this phase is also expected to invert the arrow of time in thermodynamics. Thus, if this chapter will be successful in showing that the arrow of time in quantum thermodynamics is understandable as a consequence of the time asymmetry of the particle number operator, we would find exactly what we expect; that is, the arrow of time in thermodynamics would occur only in cosmic phases of radiation or matter domination but not in the inflation phase. In the next subsection, I shall show that, for some reasonable definitions of entropy, the time arrow in quantum thermodynamics is a necessary consequence of the time asymmetry of the particle number operator.

V.3.4. On the Thermodynamic Time Arrow

In the previous subsection, I argued that the particle number in the volume $\left(La(t) \right)^3$ increases with cosmic time if $\alpha < 2$ is satisfied (according to the considered quantum cosmological models; see above). In this subsection, I shall consider different definitions of entropy in order to investigate there dependence on the expectation value of the particle number operator.

Regarding this task, it seems unclear which type of entropy definition should be considered as plausible for this investigation. Castagnino and Gunzig (1999) showed in detail that the definitions of entropy for a global system in non-equilibrium depend on the definition of the subsystems as well as the definition of the traces and projectors that are used. Thus, in arguing for one of these definitions, it would be hard to avoid artificial assumption or epistemically motivated definitions, which could lead to the same problems that are found in Allahverdyan and Gurzadyan (2002). Moreover, it seems far beyond the scope of this investigation to develop plausible criterions for 'allowed' entropy definitions (or even to develop a concept of 'possible' entropy definitions). Instead, I think it is reasonable here to concentrate only on some entropy definitions from the literature.

The claim of this subsection is that some plausible definitions of entropy that are used in quantum thermodynamics provide a time arrow as a consequence of the time asymmetry of the particle number operator. Additionally, we will see that this arrow is not grounded on artificial assumptions or epistemically motivated definitions. Together with an understanding of the origin of the thermodynamic time arrow (based on the time behaviour of the particle

number operator) this fact, I think, is the second main advantage of the suggested philosophical understanding of the nature of thermodynamic time directions.

For this task, I think, one plausible definition of entropy is that of Landau and Lifshitz (1970), which is given by

$$S = \sum_k (n_k + 1)\ln(n_k + 1) - n_k \ln n_k \; . \tag{5.65}$$

In this case, it is obvious that the entropy is an increasing function of the particle number n_k. The previous subsection showed that the expectation value of n_k is an increasing function of the scale factor a and thus also of σ. Thus, we obtain the following inequality:

$$\frac{dS}{d\sigma} \geq 0 \; . \tag{5.66}$$

In addition, σ (as well as a of course) is an increasing function of cosmic time. Thus, (5.65) is also an increasing function of cosmic time. Therefore, the entropy S obtained from (5.65) behaves *time asymmetrically*.

However, we have not yet considered the *physical relevance* of this entropy. We have a time asymmetric entropy in (5.65), but this entropy seems not physically relevant because each physical phenomenon that is affected by this entropy is influenced only by the particle number (see (5.65)). Thus, what is physically relevant is not entropy but the particle number, at least at the fundamental level of description. Nevertheless, the time asymmetry of the entropy (5.65) seems interesting because it describes a necessary thermodynamic time arrow in the radiation- and matter-dominated epochs of each considered spacetime. This seems to be relevant for many phenomenological accounts regarding time directions. The fact that definition (5.65) may be used to argue for a reductive understanding of entropy, because all properties of this entropy can be deduced from the more fundamental behaviour of the particle number, is independent of the arguments in this chapter. Moreover, this example is only *one* possible entropy definition from the literature. Thus, it does not consider the possibility that other time directed entropies that are phenomenologically defined can also define a thermodynamic time direction.

Therefor I will consider another definition in order to show that the time asymmetry from (5.65) is not a special property of this particular reductive definition. In contrast to entropy

(5.65), let us consider a purely phenomenological entropy from the literature, that of Glansdorff and Prigogine (1971) [see also Gunzig (1989)].

Glansdorff and Prigogine (1971) showed that the phenomenological description of a quantum system that includes variations in the particle number can be captured in the following expression:

$$d(\rho V) + PdV - \frac{h}{n}d(nV) = 0 \, , \tag{5.67}$$

where ρ denotes the density of the system, P denotes the pressure and V denotes the volume of the system. Here n denotes the particle density; thus, nV gives the number of particles in the described system. In addition, h represents the enthalpy of the system and is defined by $h = \rho + P$.

In this investigation, where the considered system is the entire spacetime, there is, of course, no heat flux that brings heat into the system. Thus, $dQ = 0$, where Q denotes the heat of the system. However, for this analysis, the three-dimensional volume of the system is variable and time dependent. Thus, in this phenomenological account, it seems adequate to describe the system as open because matter creation in the system acts as a source for the internal energy. Note, however, that the described spacetime is not actually seen as an open system, but according to phenomenological quantum thermodynamics, the spacetime description should be analogous to that of an open system because particle creation provides sources for the internal energy independent of any heat flux. The change in entropy in such a system, dS , is not zero as it would be in a closed system with $dQ = 0$. In our case, the behaviour of entropy can be understood as follows:

$$TdS = \frac{h}{n}d(nV) - \mu d(nV) = \frac{Ts}{n}d(nV) \, , \tag{5.68}$$

where T is the temperature of the system, and $\mu n = h - Ts$, where μ is the chemical potential, S is the entropy value and s is the entropy density. We saw in the previous subsection that for $\alpha < 2$, nV increases with cosmic time t. In addition, in this phenomenological description, it seems reasonable to assume that the temperature on the Kelvin scale is given by a positive real number. Thus, we find that the entropy from (5.68) also increases with cosmic time and decreases in the reversed cosmic time direction. Therefore, we also find an entropic time arrow for the phenomenological entropy definition

of Glansdorff and Prigogine (1971) that is independent of artificial or epistemically motivated definitions and *not* trivially reductive.

Nonetheless, like the time arrow of the entropy of Landau and Lifshitz (1970), the entropic time arrow for phenomenological entropy is a direct by product of the time arrow of particle creation in a hyperbolic curved spacetime. This, of course, says nothing about a possible ontic or epistemic interpretation of this phenomenological entropy. The proposed account is neutral with respect to an epistemic or ontic understanding of the phenomenological entropy from (5.68). I demonstrated that it is possible to find at least some definitions of entropy that show a time directed behaviour caused by the time asymmetry of particle creation.

Additionally, note that (5.68) as well as (5.65), in this framework, allows a formulation of the second law, which is explicitly time asymmetric. Traditionally, the second law is a statement about the entropy increasing future (if the considered state is not at maximum entropy), regardless of the entropy behaviour in the past direction. In the context of the proposed understanding, the second law can be formulated with an additional addendum. This addendum says that, apart from fluctuation and, if the considered state of the system is not at minimum entropy, the entropy value will decrees with decreasing cosmic time. I think, additional to the proposed understanding of the origin of the thermodynamic time arrow, the possibility of formulating such a second law is a clear advantage of the proposed account, at least in the context of *quantum* thermodynamics.

V.4. Summary of the Main Conclusions

In the first part of this chapter, I discussed parts of the account of Allahverdyan and Gurzadyan (2002) as representing, at least in parts, traditional approaches to the thermodynamic time asymmetry. The goal was to show that their account contains several crucial issues and cannot itself define a thermodynamic arrow of time on the basic physical models alone. Instead, their arrow of time arises as a consequence of

     a) artificial assumptions (the IIA) and

     b) statistical approximation methods (which leads for example to DWM cases).

I analysed this part of their investigation in detail in order to clarify the following:

i) It seem implausible, for physical reasons, to assume that such time asymmetries are associated with time asymmetries in nature or at least time asymmetries in the basic models of physics.

ii) My suggestions in the second part of this chapter do not use assumptions or boundary/initial conditions that play the role of the IIA or DWM of Allahverdyan and Gurzadyan (2002).[43]

I presented my proposals, based on the analysis of Castagnino and Laciana (2002), that the thermodynamic time arrow, at least for some plausible definitions of entropy, is understandable as a consequence of the time asymmetry of the expectation value of the particle number operator in a hyperbolically curved spacetime.[44]

To show that the entropic time asymmetry, which seems to play an important role in many phenomenological approaches, is understandable as a necessary consequence of the cosmological time asymmetry in quantum cosmology, I investigated two reasonable definitions of entropy in quantum thermodynamics. One, (5.65), was developed for quantum systems in non-equilibrium, and the other, (5.68), was used in a phenomenological description of quantum systems. The outcome was that both entropy definitions yielded time asymmetric behaviour that defines the second law of thermodynamics in a non-time-reversible way (even in theory). I showed that this entropic time asymmetry is a necessary consequence of the cosmological time asymmetry, if the dynamics of the matter and energy content of spacetime can be described by a scalar matter field and if the total action is given as the Einstein-Hilbert action plus the contribution from the matter field. This account yields some important properties of the thermodynamic time arrow:

i) The arrow is not provided by any pure epistemically motivated or artificial definitions.

ii) The arrow is understandable as a consequence of the cosmological time asymmetry.

iii) The arrow is physically effective, via the time asymmetry of the particle number operator, for an ontic and an epistemic interpretation of entropy.

---

[43] The crucial assumption of an initial vacuum state is motivated by the considered quantum cosmological models and not by pure epistemically motivated assumptions or artificial definitions.

[44] If the dynamics of the matter and energy content can be described by a scalar matter field and if the total action is given as the Einstein-Hilbert action plus the contribution from the matter field

iv) A second law of thermodynamics can be formulated, which is, according to the proposal, explicitly time asymmetric.

In addition, the formal time symmetry that arises from the two time reversed cosmological solutions of the Einstein equations (in the considered model) could be seen as broken by Leibniz argument. In the quantum cosmological case, the solutions $\pm\gamma$ can be seen as describing the same physical world, and the change in the sign of t refers only to the mathematical construct of an absolute Newtonian background time, which is not physical (see also chapter III). Hence, we saw that the fundamental time asymmetry from chapter III, which was first only explicated in the context of classical physics (chapter III and IV), also has explications in quantum cosmological models.

In summary (see also the following diagram):

i) I showed that, in contrast to the account of Allahverdyan and Gurzadyan (2002), it is possible to develop an understanding of thermodynamic time arrows independently of purely epistemically motivated assumptions or artificial definitions.

ii) We find a physical explanation for the occurrence of the thermodynamic time asymmetry that arises from more basic physical theories.

iii) The entropic time arrow, in the proposed understanding, is physical effective, via the particle number operator, even if the entropy value is considered as a purely epistemic value. Of course, in this case, the time asymmetry of e.g. the function (5.56) would probably better named a 'particle number' asymmetry than an entropic one. Nevertheless, the formal structure would be the same and the physical effectiveness could be described according to this structure, independent from the name of the time asymmetry.

Consideration of typical approaches to the entropic time arrow in statistical physics; I shall consider, as representative to similar standard approaches, the account from Allahverdyan and Gurzadyan (2002).

⇩

Approaches, like the one taken by Allahverdyan and Gurzadyan (2002), are unable to provide an understanding of the thermodynamic time arrow, which is independent from epistemically motivated assumptions / definitions.

⇩

Hence, a proposal which offers an understanding, which is not based on epistemically motivated assumptions / definitions, seems much more satisfying than such standard approaches.

⇩

An alternative understanding of the entropic time asymmetry, independent of epistemically motivated assumptions / definitions but based on an explication from the fundamental time asymmetry (chapter III) in quantum cosmological models.

⇩

The quantum cosmological model and the intrinsic (cosmic) time asymmetry as an explication from the fundamental time asymmetry from chapter III

⇩

Time asymmetric behaviour of the expectation value of the particle number operator in the considered model

⇩

Entropic time arrows arising from the behaviour of the particle number operator.

⇩

Some unfamiliar properties of the entropic time asymmetries

Hence, we saw how the fundamental time asymmetry in the solution space of the crucial cosmological equations leads to explications in quantum cosmological models, which can have crucial time asymmetric effects for the entropy behaviour.

Nevertheless, until jet, I have not considered one very prominent time arrow from the literature, the time arrow in standard quantum mechanics (see also chapter I). I will consider this arrow in the next chapter. I shall argue that three different quantum mechanical time arrows can be constituted in three levels of quantum mechanical description. Moreover, as announced in chapter I, I shall argue that all of them can be seen as consequences from the cosmological time asymmetry. Also, I will argue in the next chapter that one of this time arrows, in combination with the proposed understanding, provides strong and independent arguments for a minimal rigged Hilbert space formulation of ordinary quantum mechanics.

Chapter VI

Time Arrows in Ordinary Quantum Mechanics

This short chapter aims to demonstrate that the time asymmetry of the local energy flux (chapter IV) yields different time arrows in ordinary quantum mechanics,[45] as straight forward consequences.

More precise: The time asymmetric energy flux (see chapter IV) will be shown to provide three different time asymmetries in quantum mechanics. While considering quantum mechanics, it may be crucial to distinguish between three levels of the theory. The distinction is motivated by the necessity to avoid considerations of the measurement problem. The different levels can be denoted as follows:

> i) The first level of quantum mechanics considers only quantum mechanical processes without assuming even the possibility of any quantum measurement. Thus this will be the level of 'pure' quantum mechanical evolution.

> ii) The second level considers the measurement itself, but this consideration will be in the context of only one prominent attempt to describe parts of quantum measurements, the decoherence approach.

> iii) The third level is a description in which measurements are considered as black boxes. This level, of course, is particularly useful to describe and predict the outcomes of laboratory experiments.
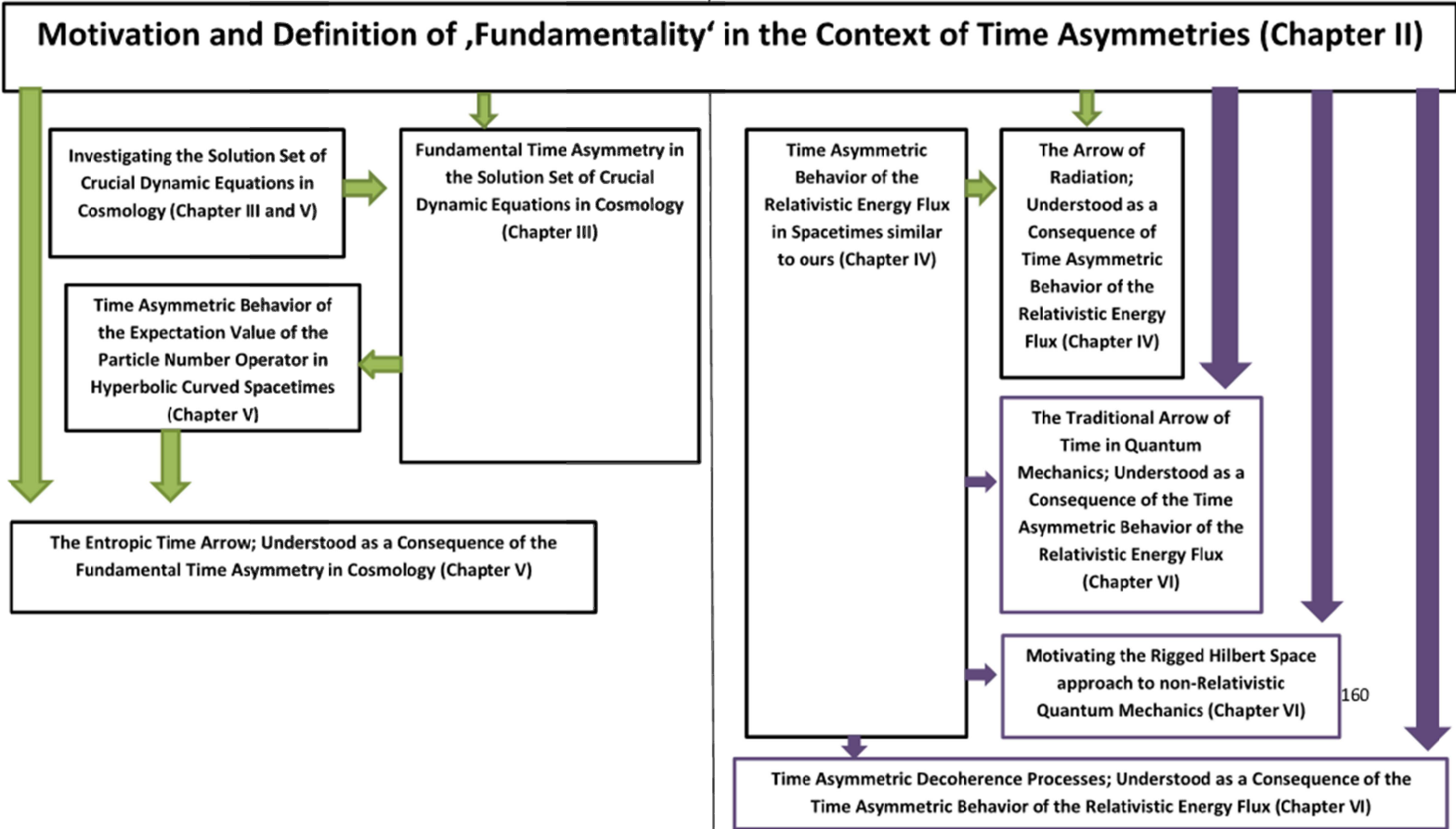
Motivated by the physical analysis of Castagnino, Lara and Lombardi (2003), I shall argue that the time asymmetry of the energy flux provides a time asymmetry at all three levels. Regarding this chapter's connection to the entire investigation, the following first diagram illustrates, as usual, which parts of the analysis are considered in this chapter and the second following diagram will sketch the internal structure of this chapter in a bit more detail.

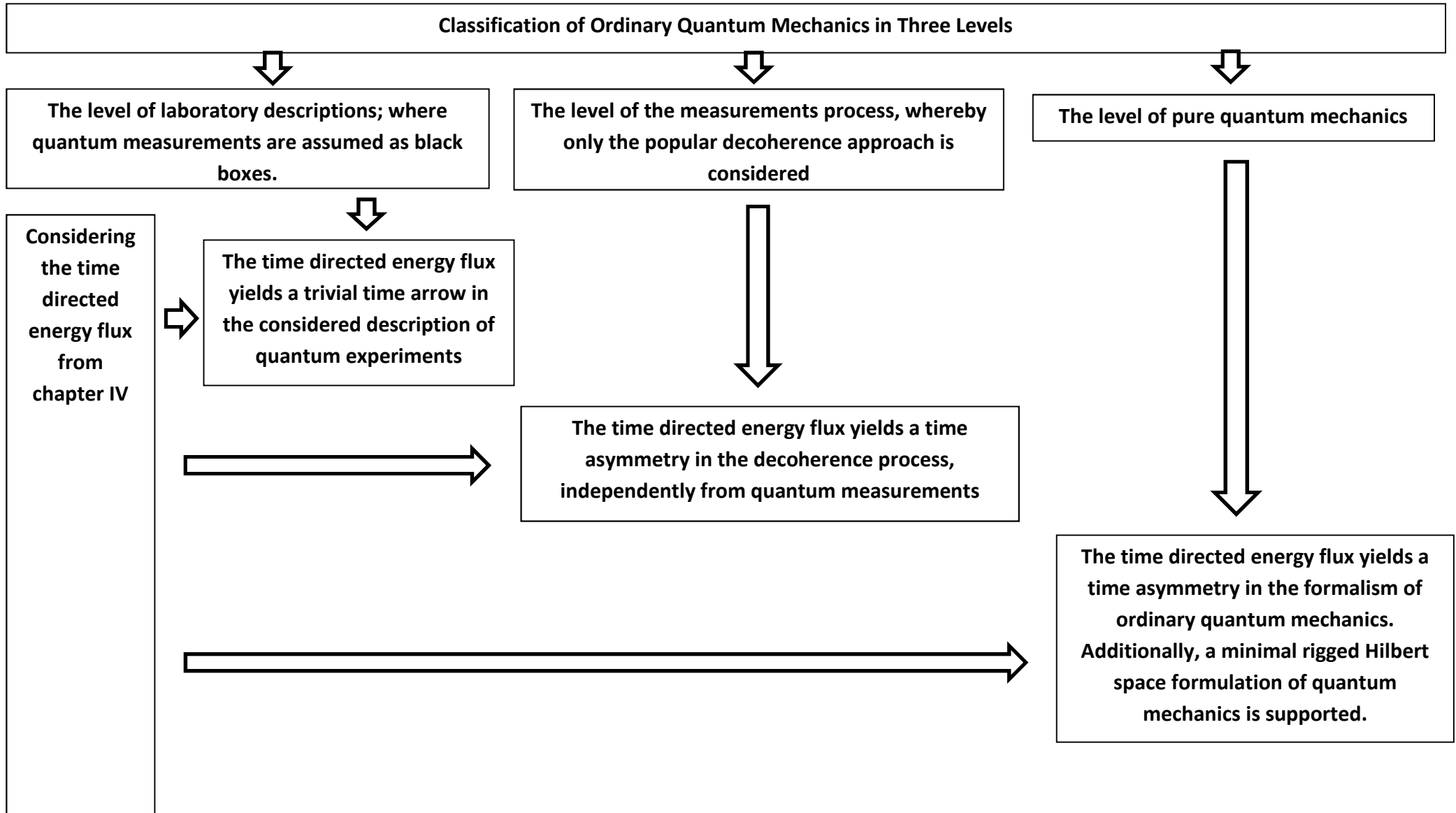[45] The scope of this chapter will not include field theoretical descriptions. This in necessary if we want to use the outcomes from chapter IV; because the conditions on the energy-momentum tensor, used in chapter IV, are not fulfilled generally in quantum field theoretical description. But, there seem fulfilled, in our particular spacetime as long as we only consider descriptions in ordinary quantum mechanics.

Time Arrows in Ordinary Quantum Mechanics

Cosmic time asymmetries

Proper time asymmetries

Motivation and Definition of ‚Fundamentality' in the Context of Time Asymmetries (Chapter II)

Investigating the Solution Set of Crucial Dynamic Equations in Cosmology (Chapter III and V)

Fundamental Time Asymmetry in the Solution Set of Crucial Dynamic Equations in Cosmology (Chapter III)

Time Asymmetric Behavior of the Expectation Value of the Particle Number Operator in Hyperbolic Curved Spacetimes (Chapter V)

The Entropic Time Arrow; Understood as a Consequence of the Fundamental Time Asymmetry in Cosmology (Chapter V)

Time Asymmetric Behavior of the Relativistic Energy Flux in Spacetimes similar to ours (Chapter IV)

The Arrow of Radiation; Understood as a Consequence of Time Asymmetric Behavior of the Relativistic Energy Flux (Chapter IV)

The Traditional Arrow of Time in Quantum Mechanics; Understood as a Consequence of the Time Asymmetric Behavior of the Relativistic Energy Flux (Chapter VI)

Motivating the Rigged Hilbert Space approach to non-Relativistic Quantum Mechanics (Chapter VI)

160

Time Asymmetric Decoherence Processes; Understood as a Consequence of the Time Asymmetric Behavior of the Relativistic Energy Flux (Chapter VI)

Classification of Ordinary Quantum Mechanics in Three Levels

The level of laboratory descriptions; where quantum measurements are assumed as black boxes.

The level of the measurements process, whereby only the popular decoherence approach is considered

The level of pure quantum mechanics

Considering the time directed energy flux from chapter IV

The time directed energy flux yields a trivial time arrow in the considered description of quantum experiments

The time directed energy flux yields a time asymmetry in the decoherence process, independently from quantum measurements

The time directed energy flux yields a time asymmetry in the formalism of ordinary quantum mechanics. Additionally, a minimal rigged Hilbert space formulation of quantum mechanics is supported.

161

VI.1. Brief Introductory Thoughts

When Eddington coined the phrase 'arrow of time' to describe the asymmetry of time directions according to physical phenomena in nature, the time direction in quantum physics was not considered because quantum physics was not well established in 1927. However, according to modern science, the fundamental level of description is given by models in fundamental physics, which are based primarily on quantum physics. In chapter V, we showed that quantum cosmology, as well as QFT and quantum thermodynamics, in the context of cosmological models, includes a time arrow as explications from the fundamental time asymmetry in cosmology (chapter III and V). However, as we have seen, this time arrow appears in the application of these fields to a cosmological description of large regions of considered spacetimes. Moreover, the fundamental time asymmetry (as well as the explications in quantum cosmology and quantum thermodynamics) is defined according to cosmic time; thus, this time asymmetry focuses explicitly on global descriptions of spacetimes.

The time asymmetries considered so far that are valid for proper times in local descriptions are the time asymmetric behaviour of the energy flux and the arrow of radiation as a consequence of this time directed energy flux. None of those was considered in the context of quantum physics and, as mentioned in chapter IV, these asymmetries cannot yet be understood as consequences of the fundamental time asymmetry, because the alignment of the radiation arrow regarding the cosmic asymmetry is not well understood. We will find a similar situation regarding the local time asymmetries in ordinary quantum physics. So, in this chapter I focus on time asymmetries in local descriptions and in the theory of ordinary quantum mechanics.

Also, for many approaches, the time asymmetries in ordinary quantum mechanics are playing an very important role, see, for example Vilenkin (1988), Castagnino (1998), Narlikar (1995) in 'From Black Clouds to Black Holes', Adams and Laughlin (1997), Bergström and Goobar (1999) in 'Cosmology and Particle Astrophysics', Lax and Phillips (1979) in 'Scattering Theory', Zeh (1989) in 'The Physical Basis of the Direction of Time', Mackey (1989), Ehrenfest and Ehrenfest (1959) in 'The Conceptual Foundations of the Statistical Approach in Mechanics', Schulman (1997) in 'Time's Arrow and Quantum Measurements' or Ajiezer (1981).

In some sense, this chapter aims to clarify the question regarding the origin of different time asymmetries in ordinary quantum mechanics by using the time directed energy flux (chapter IV) and the formal description of quantum phenomena.

On one hand, the fundamental Schrödinger equation (fundamental in ordinary non-relativistic quantum mechanics)[46] is time-reversal invariant in the standard formulation and interpretation. On the other hand, quantum mechanics seems to provide an arrow of time on a very basic level. Thus, the main claim of this chapter is that the time asymmetry of the energy flux (see chapter IV) provides time arrows in quantum mechanics, independently of the measurement problem. Moreover, I will argue that, according to the decoherence account to the measurement problem, measurements itself could be seen as time asymmetric processes, where the origin of this time asymmetry is located in the energy flux.

VI.2. On the Motivation of the Arguments

I will present my arguments and suggestions about the time arrows in quantum mechanics without proposing speculations about quantum gravity, which may appear to be not really self-consistent. The time asymmetry in the energy flux (chapters IV) appears in classical or semi classical models, that is, without the quantisation of gravity. Given the considerations from chapter IV, it seems that general relativity (or empirically equivalent spacetime theories) is understood as the fundamental theory. Thus, it seems questionable to suggest that an analysis in which general relativity (or empirically equivalent spacetime theories) is treated as the fundamental theory can provide a self-consistent understanding of time arrows in quantum mechanics.

Consequently, I shall argue that for the parts of quantum mechanics that are relevant to the problem of time directions, it is possible to translate the time asymmetry in the relativistic energy flux in a quantum mechanical framework without losing self-consistency or the explanatory power of quantum mechanics. The reason is that philosophically, this analysis continues to treat the structure of general relativity as a basic structure. I will demonstrate that the time asymmetry discussed in chapter IV produces a time asymmetries in certain fields of quantum mechanics. Here, we can avoid the question of which theory should be treated as more fundamental because quantum mechanics, as well as ordinary QFT's, are not applied to the description of gravity. Thus, the critical question of quantum gravity does

---

[46] A generalization to the Klein-Gordon equation and the Dirac equation, in order to capture relativistic formulations, is, of course, possible. But it would not change the issue in question. Hence, I will focus on the Schrödinger equation, even if I occasionally will use the concept of proper times. The usage of proper times, in this chapter, is, of course, meant in the context of a relativistic formulation, but always excluding field theoretical descriptions.

not arise in this analysis as long as general relativity and quantum physics are treated as having different fields of application.

I shall first concentrate on level iii, the level of the description in laboratory experiments. The next subsection shall demonstrate that the fundamental time asymmetry in cosmology provides, via the time asymmetric energy flux (chapter IV), a time asymmetry in laboratory descriptions of quantum experiments if measurements are treated as black boxes. Thus, this time asymmetry will be independent of a hypothetical solution to the measurement problem and, as we will see, the asymmetry will arise in a very simple way as a consequence from a time asymmetric energy flux.

VI.3. The Arrow of Time in Laboratory Descriptions

To begin, let us consider a typical laboratory experiment such as the scattering experiment illustrated in the following Figure VI.1.
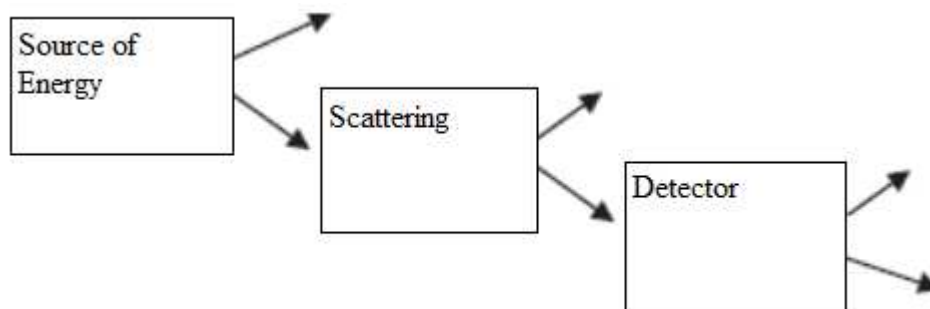
Fig. VI.1: Illustration based on Reichenbach–Davies diagrams; see also Castagnino, Lombardi and Lara (2003).

To prepare the system at the beginning of the experiment, we need a source of energy that enables the scattering process and the detection of the experimental outcomes. The boxes describe the different stages of the experiment. The arrows represent energy fluxes, and the diagram is oriented in time: The horizontal axis represents proper time. In addition, a flat spacetime is assumed to be adequate for the local environment of the experiment. This is important in order to avoid the effects of gravity on the scattering experiment itself. Although I used a Reichenbach–Davies-type diagram for its simplicity and ability to capture the relevant processes of the analysis, but note that this analysis is not based on Reichenbach's or Davies's or Castagnino's description of local time asymmetries. This is

because this analysis is not based on the definition of branches or subsystems *in general*, even if such definitions are presupposed in the laboratory description via the measurements (which makes the diagram VI.1 adequate).

I shall argue now that figure VI.1 contains all the information needed to construct a time asymmetry for typical laboratory descriptions in which quantum measurements are treaded as black boxes.

To describe such a scattering experiment (figure IV.1), we must consider the measurement apparatus. To make this analysis as simple as possible, all the laboratory apparatus is labelled 'measurement apparatus'; this includes the detector itself and the technical apparatus used to prepare the quantum state of the system. As shown in figure VI.1, we can assume that the apparatus used to prepare the quantum state for the experiment requires energy to do work. According to the discussion in chapter IV, the energy flux has an intrinsic time asymmetry. Thus, the energy needed to prepare the quantum state comes from the proper past of the quantum system. The detector reaches an excited state when a positive measurement is performed and then returns to a stable state. In the Reichenbach–Davies diagram in figure VI.1, the energy flux captures this process. Thus, the experiment, at least for such laboratory descriptions, can be understood as an energy cascade in the Reichenbach–Davies diagram. For the reason described in chapter IV, the reverse energy cascade (with reversed energy arrows in figure VI.1) is prohibited by the time asymmetry of the local energy flux.

Without the time asymmetry of the local energy flux, the description would be time symmetric because it seems that all relevant physical processes are guided by TRILs; however, according to chapter IV, the energy flux is time asymmetric, regardless of the alignment of this asymmetry in the cosmic background. Thus, the time-reversed process is forbidden by the time asymmetry of the local energy flux from chapter IV, which also explains why the energy arrows in the Reichenbach–Davies diagram are directed in the future direction of the system. Thus, the time asymmetry in laboratory descriptions can also be understood as a by-product of the time asymmetric energy flux, which becomes physically relevant (see chapter IV).

Moreover, the 'laboratory-arrow' occurs *necessarily* in spacetimes in which the crucial assumptions regarding the energy–momentum tensor are realized (see chapter IV). These assumptions are that the energy–momentum tensor is of type one and that the dominant energy condition is fulfilled. At least at the descriptive level of ordinary quantum mechanics and in our particular universe, those conditions seem fulfilled (but see also chapter IV for exceptions in quantum field theories). Also, even if this arrow seems a quite trivial consequence of the time directed energy flux, note that this time asymmetry is *independent* of a specific interpretation of quantum physics.

However, the proposed explanation of the arrow is valid only for special laboratory descriptions. Only on the basis of the above description of the experiment, where measurements are assumed to define the subsystems and are treaded as black boxes, we can avoid the measurement problem and use the proposed simple reasoning. So, the identified time asymmetry is on the '*classical*' level.

This can be seen more precisely: for a physical system in nature, according to usual quantum mechanics, it is unclear what counts as a 'measurement'. Thus, it would be unclear whether the suggested description can be applied to a quantum process in nature because we cannot say whether a 'measurement' (or an analogous physical process) is performed or not and the offered description, so far, was independent from the concrete quantum mechanical time evolution. Hence, time asymmetries in a quantum system in general cannot be understood in a similar way. In the considered description, this was only possible because we assumed a quantum measurement (as a black box), which gives us the subsystems which can be described as nearly classical. For this reason, the proposed understanding of time asymmetries is applicable only to laboratory descriptions where, by definition, it is clear whether a measurement is performed by a physical interaction and hence, can be treated as a back box.

The next section shall demonstrate how a quantum mechanical time asymmetry can be found in ordinary quantum mechanics. We will see that according to such ordinary models, the time asymmetry of the local energy flux from chapter IV can also provide a quantum mechanical time asymmetry even if we do not consider any measurement process. Hence, the next section considers level i of the quantum physical description, a pure quantum mechanical system.

VI.4. A Time Asymmetries in Ordinary Quantum Mechanics and the Rigged Hilbert Space Approach

To construct my arguments in this level it appears useful to consider the Hamiltonian of a quantum system with eigenvalues that describes the system's energy spectrum. The spectral analysis of every Hamiltonian and thus of every quantum system in ordinary quantum mechanics can be mapped to the complex plane. The poles in the lower half of the complex plane correspond to decaying unstable states of the system. Those in the upper half of the complex plane are symmetrically located with respect to the positions of the lower poles.

The symmetry here is provided by the formal time symmetry[47] of Schrodinger's equation. These upper poles correspond to unstable states of the system with unbounded growing energy. What is remarkable in this situation is that the poles are associated with energy fluxes; the growing states in the upper half of the complex plane provide an energy flux into the proper past of the system. In contrast, the decaying states in the lower half of the complex plane are associated with an energy flux into the proper future of the system (see Castagnino, Lara and Lombardi (2003)). Thus, according to the time asymmetry of the energy flux from chapter IV, we find that the growing states in the upper half of the complex plane are forbidden. This agrees perfectly with empirical observations because, traditionally, the growing states in the upper half of the complex plane are understood as unphysical due to their unusual energy behaviour, and such states are never observed in nature. Thus, apart from the time asymmetry itself, we find an explanation for the fact that the growing states in the upper half of the complex plane do not appear in our particular spacetime because they are ruled out by the time asymmetry of the energy flux.

Furthermore, I shall argue that the physical distinction between the poles in the upper and the lower halves of the complex plane supports a time asymmetric formulation of ordinary quantum mechanics. This seems desirable, because the traditional formulation, in theory, predicts energy-eigenstates of quantum systems which are never observed in nature. In contrast, a formulation in which all predicted energy-eigenstates are associated to physically possible states seems much more preferable to me.

To sketch such a formulation, we can divide the Hilbert space H into two subspaces, ϕ_+ and ϕ_-. The state vectors $|\varphi\rangle$ are now vectors in ϕ_+ or ϕ_- and are characterized by the following consideration:

The states $|\omega\rangle$ are the energy eigenstates of the system. The projections of the states $|\varphi\rangle$ on the energy eigenstates $|\omega\rangle$, are now functions of the Hardy class of the upper or lower half of the complex plane. Thus, the projections can be used to characterize the states $|\varphi\rangle$. [See Bishop (2004), Bohm, Gadella and Wickramasekara (1999) or Castagnino, Lara and Lombardi (2003).]

However, the Hilbert space itself is time-reversal invariant. Note that, so far, I have used the phrase 'time-reversal invariant' only to describe dynamic equations. For a vector space, 'time-reversal invariance' should mean that the following equation holds:

---

[47] In this case it is the symmetry between the Schrödinger equation and the complex conjugated Schrödinger equation.

$$KH = H \ . \tag{6.1}$$

Here, K is the anti-linear Wigner operator for time mirroring (see Gadella and Wickramasekara (1999) or Castagnino, Lara and Lombardi (2003)). Note, however, that the subspaces ϕ_+ and ϕ_- are not time-reversal invariant. For them the following equation holds:

$$K\phi_{\pm} = \phi_{\mp} \ . \tag{6.2}$$

According to (6.1), a crucial property of the traditional formulation of ordinary quantum mechanics in Hilbert space H is that it cannot be used to describe an intrinsically time-reversal *variant* phenomenon. I believe this is obvious from (6.1), but Castagnino and Gunzig (1999) present a more formal argument.

However, using the subspaces ϕ_+ and ϕ_-, it is possible to build a time asymmetric description of ordinary quantum mechanics. To change as few properties of the traditional formulation in Hilbert space as possible, we can use the substitution $H \rightarrow \phi_+$ or $H \rightarrow \phi_-$. The difference between these substitutions, which are standard in the description of rigged Hilbert space quantum mechanics, can be understood as a physical difference. In the traditional rigged Hilbert space description, the substitutions provide only a formal distinction between the situations; however, according to the analysis above and the time asymmetry of the energy flux from chapter IV, we find a physical difference between the two subspaces ϕ_+ and ϕ_-. Thus, this analysis strongly supports the rigged Hilbert space formulation of ordinary quantum mechanics, at least in spacetimes which fulfilling the conditions on the energy-momentum tensor (chapter IV). In the limits of the considered description, here in ordinary quantum mechanics, our particular spacetime seems to be a spacetime of this type.

Thus, according to this minimal formulation of rigged Hilbert Space quantum mechanics, we have identified a quantum theory that includes an arrow of time independently of any measurement but according to the time asymmetric energy flux from chapter IV.

To make the proposed formulation more precise, only the quantum states in the lower half of the complex plane are contained in the space ϕ_-; thus, only these states correspond to possible physical states. The other states are ruled out by the 'energy flux time asymmetry' and are no longer given in the new formulation of the rigged Hilbert space, if the substitution $H \rightarrow \phi_-$ is used. Only with this substitution do we find that the energy flux from a quantum state is contained *only* in the system's future semi-light cone.

To summarise: at the beginning of this chapter, we found that a time asymmetry appears in laboratory descriptions; thus, we avoid the measurement problem by using only measured physical states and treating the measurement itself as a black box.

In this section, we find that the asymmetric energy flux yields an arrow of time if we avoid the measurement problem by using only quantum mechanical descriptions in which no measurement is performed. In both cases, we find that the time asymmetric energy flux from chapter IV produces a time asymmetry.

However, current knowledge of physics also offers another level of ordinary quantum mechanics, the level of the measurement itself. So far, I have not examined this domain of quantum mechanics because providing a solution to the measurement problem seems totally outside the scope of this study. Nevertheless, I address this domain of ordinary quantum mechanics, focusing on only one account of describing quantum measurements, the decoherence approach, which seems very popular in modern physics. The next section shall show that, according to the decoherence approach, the asymmetry of the energy flux provides an intrinsically time asymmetry in the measurement process itself.

VI.5. A Time Arrow in Decoherence Effects

This section focuses on the time-directed aspects of the decoherence approach without discussing the various interesting points regarding the approach's relevance to the measurement problem. Therefore, I state only that in physics, the account is taken seriously although, I think, it has not yet been successful. I consider that its popularity renders it acceptable as a possible description of quantum measurements and hence as the considered approach in this section. I intend here to demonstrate, partly by applying Castagnino, Lara and Lombardi (2003), that the formalism of the decoherence approach together with the time asymmetry of the local energy flux can constitute a time asymmetry of quantum measurements. The analysis is based primarily on the quantum process of decoherence, in which quantum states with non-diagonal density operators become those with diagonal density operators.

Such a state, with diagonal density operators, corresponds to a classical ensemble. A time asymmetry in this approach, if the decoherence account is assumed to correctly describe quantum measurements, would itself be an intrinsic property of quantum measurements and hence, I think, would be show that the connection between quantum and classical

physics could be intrinsically time asymmetric. Note that this time asymmetry is independent from the specific way of defining physical subspaces (in contrast to a similar proposal from Castagnino, Lara and Lombardi (2003)). To argue for this proposal, in the following, I will sketch the decoherence approach for simple density operators and I shall demonstrate how the time directed energy flux (chapter IV) can yield time asymmetric consequences for the decoherence process itself.

VI.5.1. From Quantum Systems to Classical Ensembles; Decoherence

In this subsection, I focus on a simple quantum system with an energy spectrum bounded by $0 \leq \omega < \infty$. An arbitrary observable O should be part of the space O. The operator corresponding to the observable O can be described by the projectors of the eigenstates of the Hamiltonian $\{|\omega\rangle\}$. I designate ρ as the density operator. So, the density operator ρ of a quantum state is in the set $S \subset O$. For convex S sets, the density operator ρ can be described by the following equation:

$$\rho = \int \rho(\omega)(\omega] d\omega + \iint \rho(\omega,\omega')(\omega,\omega'] d\omega d\omega'.$$
(6.3)

In (6.3), $\{(\omega], (\omega,\omega']\}$ is the dual basis of $\{[\omega), [\omega,\omega')\}$, and (6.3) can be translated into the ordinary Dirac notation by using the equations $[\omega) = |\omega\rangle\langle\omega|$ and $[\omega,\omega') = |\omega\rangle\langle\omega'|$. Moreover, $\rho(\omega,\omega')$ represents regular functions, so we can use the Lebesgue theorem to solve (6.3). To see this more precise consider the expectation value of an operator O for a quantum system described by the density operator ρ. I denote this expectation value as $\langle O \rangle_\rho$.

$$\langle O \rangle_\rho = (\rho][O) = \int \rho(\omega)O(\omega)d\omega + \iint \rho(\omega,\omega')O(\omega,\omega')e^{-i(\omega-\omega')t} d\omega d\omega'.$$
(6.4)

Now equation (6.4) shows that we can use the Lebesgue theorem. According to this theorem, we can consider the limit case of (6.4) for infinitely large times t. For the expectation value given by (6.4), we find that the limit case for infinitely large times t is

given by $(\rho_*)[O]$. Here, O denotes an operator that corresponds to an arbitrary observable, and ρ_* denotes the diagonal density operator. Thus, for ρ_* we find

$$(\rho_*) = \int \rho_\omega [\omega] d\omega.$$ (6.5)

Also, corresponding to (6.5), we find the weak limit [see Castagnino, Lara and Lombardi (2003)]

$$W \lim_{t \to \pm\infty} (\rho) = (\rho_*).$$ (6.6)

(6.6) is the formal description of the physical decoherence in (6.4); (6.6) shows that the non-diagonal part of (6.4) (the double integral) vanishes for infinitely large times. Thus, a quantum system described by (6.3) for all observables (6.4) has a limit for $t \to \pm\infty$, which is describable as a classical ensemble.

This, I think, is the general motivation and the crucial point of the decoherence approach. For our analysis, (6.6) implies that a non-diagonal density operator ρ converges to the diagonal density operator ρ_* for $t \to \infty$ *and* $t \to -\infty$. This fact[48] is provided by the time-reversal invariance of the Schrödinger equation. So far, the decoherence approach seems time symmetric. However, according to the time asymmetry of the local energy flux, again, the time evolutions $t \to \infty$ and $t \to -\infty$ are physically different and thus, the evolution $t \to -\infty$ is forbidden by the time asymmetry of the local energy flux.

The time evolution $t \to -\infty$ is an evolution to the proper past of the system. According to this evolution, we also find an energy flux to the proper past. However, as shown in chapter IV, such an evolution is forbidden by the time asymmetry of the local energy flux. Thus, because of the time asymmetry of the energy flux, we find an time asymmetry pointing from the non-diagonal state of a quantum system (past) to the diagonal state (future) in the decoherence formalism. Of course, as outlined in chapter IV, the alignment of the labels 'proper past' and 'proper future' with respect to the labels 'cosmic past' and 'cosmic future' from the cosmological time asymmetry remains unclear, as it was the case for the radiation arrow in chapter IV.

---

[48] That is, the density operator converges in both time directions.

Also, all the properties we have used so far are physical properties and independent of any interpretation of quantum mechanics. However, as mentioned, the given time asymmetry is primarily interesting in the framework of the decoherence approach, where the described process is associated with a quantum 'measurement' or an analogous physical interaction. Thus, in the next subsection, I shall consider additional details of the decoherence approach to demonstrate that the time asymmetry considered above provides a time asymmetry in the measurement process itself, if the decoherence approach is considered to be successful in describing quantum measurements in general.

VI.5.2. The Measurement

According to the decoherence approach, the process described above is crucial to clarifying how a quantum state, in a time evolution $t \rightarrow \infty$, becomes describable as a classical ensemble. However, if the decoherence approach is applicable to the measurement problem, which I believe is why it is so popular, it is crucial to describe the evolution for *finite* times. Thus, the aim of the following analysis is to demonstrate that the time asymmetry obtained above is interesting not only for infinite times but also, at least in the decoherence approach, in the description of the measurement itself.

In the decoherence approach, the state with a diagonal density operator, which follows from the time evolution for the limit case of $t \rightarrow \infty$, is interpreted as a classical ensemble because it has the same formal description as any other classical ensemble in classical statistical physics. This results from the following reasons:

The diagonal limit state is interpreted as a set of classical states that has different parallel trajectories in phase space. These trajectories in phase space can be described as analogous to a classical ensemble, as demonstrated in detail by Castagnino and Laura (2000).

More precisely, a classical ensemble can be described by the Wigner function $\rho_\omega^W(q, p)$. The function $H^W(q, p)$ is the classical Hamilton function, and $A_i^W(q, p)$ represents the Wigner functions for the different commuting observables with the quantum numbers a_i. Those quantum numbers and the energy eigenvalues define the physical states $(\omega] = (\omega, a_1, a_2, \ldots a_N]$. Thus, the Wigner function $\rho_\omega^W(q, p)$ describes the classical trajectory given by the constant values of motion. These conserved quantities are given by

$$\omega = H^W(q, p) \text{ and } a_i = A_i^W(q, p); i \in \{1, 2, 3..., N\}. \tag{6.7}$$

Moreover, the Wigner function $\rho_*^W(q, p)$ also corresponds to $(\rho_*]$. Here, the connection is given by

$$\rho_*^W(q, p) = \int \rho_\omega \rho_\omega^W(q, p) d\omega. \tag{6.8}$$

According to classical statistical physics, this is the description of a classical ensemble. The state of the ensemble is $\rho_*^W(q, p)$, and it corresponds to the trajectories given by the conserved quantities $\omega = H^W(q, p)$ and $a_i = A_i^W(q, p); i \in \{1, 2, 3..., N\}$; the classical statistical probabilities are given by ρ_ω. This is why the quantum state with a diagonal density operator is interpreted as a classical ensemble in the decoherence approach. For all *finite* times, however, we find that each state corresponds to a small *region* in phase space, which contradicts the interpretation of the state as a classical ensemble. One crucial point for the application of the decoherence approach to the measurement problem is the following approximation.

If, even for finite times, an action S (e.g. a measurement) is performed on the system, where $S \gg \hbar$, then the small region in phase space becomes approximately a point, which yields the classical description of a classical state. Thus, the system is in a state where the classical limit $\hbar \to 0$ is approximately fulfilled. This crucial point allows us to describe the quantum system, approximately but even at finite times, as a classical ensemble that includes the time asymmetry mentioned above, if a sufficiently large action (e.g. a measurement) is performed.

However, to demonstrate the last necessary step, that is to demonstrate that the classical ensemble is reducible to a classical state, advocates of the decoherence approach focus on many aspects of it, such as chaotic statistical processes and approximations. I think, until jet, this last step is not demonstrated in the literature, which seems to yield the most serious critic on the decoherence approach in general.

Nevertheless, given that the decoherence approach is very popular, I find it important that the time asymmetry of the local energy flux provides a time asymmetry for decoherence processes itself, even for finite times. This time asymmetry, in the proposed understanding, is independent from the success of the decoherence approach regarding the last missing step.

Thus, assuming the success of the decoherence approach in describing quantum measurements, the time asymmetric energy flux would explain why a quantum measurement is asymmetric in time.

Assuming that the decoherence approach were unsuccessful in describing quantum measurements in general, the crucial effect of decoherence is still time asymmetric and physically important. The importance of the decoherence process is demonstrated by the fact that for large times, it prohibits certain quantum effects that we cannot discover in macrophysics: e.g. interference effects. Thus, I think that, even if physics should prove that the decoherence approach fails to describe quantum measurements in general, the decoherence effect would nevertheless be important for understanding parts of the connection between classical physics and quantum physics, and the investigation above demonstrates that this part of the connection can be seen as time asymmetric.

Thus, we find:

>a) that the description of a laboratory experiment, assuming laboratory descriptions where quantum measurements are handled as back boxes, is time asymmetric.

>b) that the pure quantum mechanical level, without assuming a measurement, is time asymmetric, if the energy flux from chapter IV is considered.

>c) that, according to one popular approach to quantum measurements in physics, the decoherence approach, the measurement process itself can also be time asymmetric.

All these asymmetries are derived from the asymmetry of the energy flux in spacetimes that fulfil the conditions mentioned in chapter IV, as our particular spacetime seems to do, at least as long as we focus on classical and ordinary quantum mechanical descriptions.
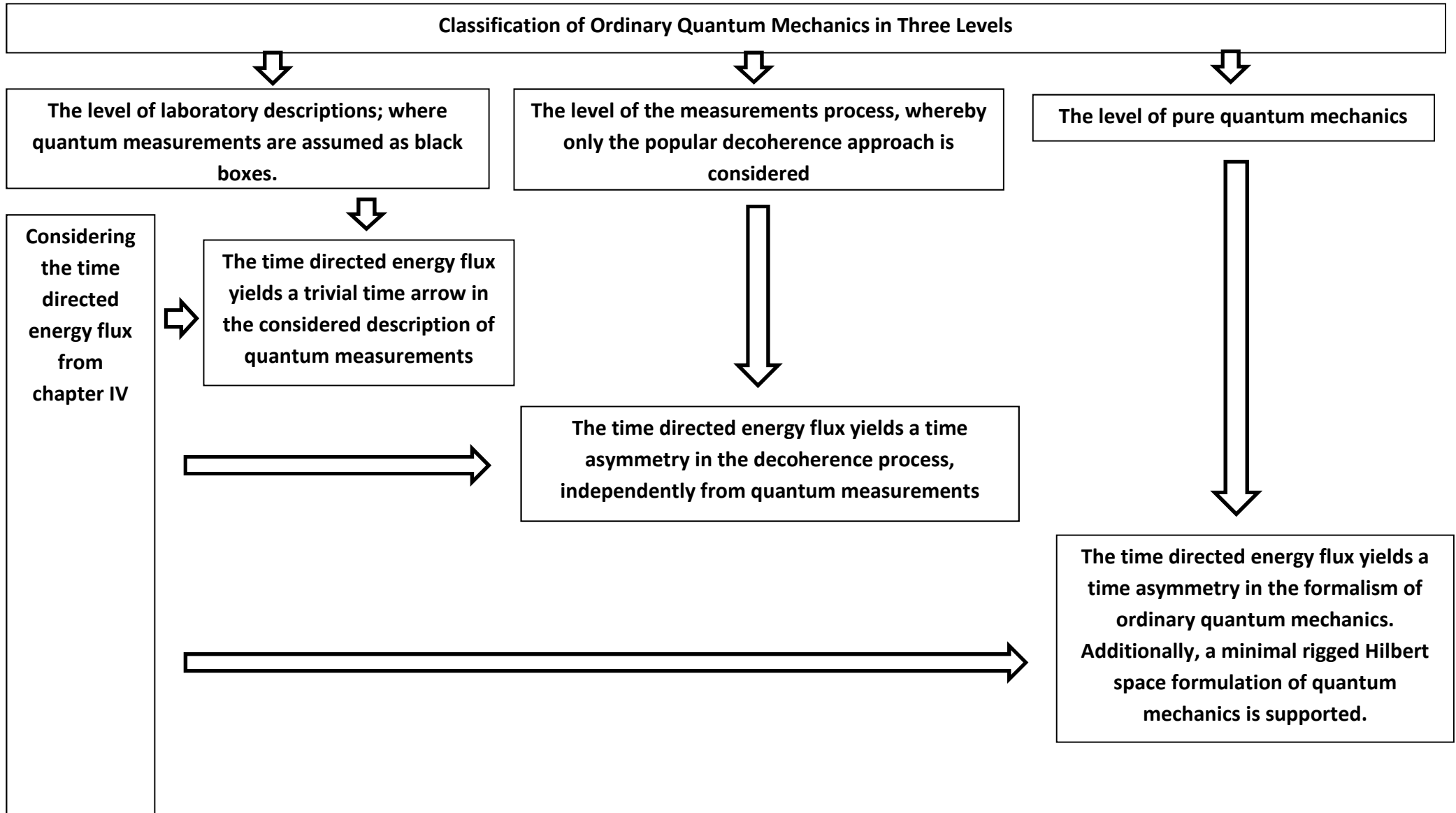
VI.6. Conclusions

This chapter showed that a time asymmetry of the energy flux in the considered spacetimes produces the following (see also the diagram on page 175):

>i) A reasonability time asymmetric description of laboratory experiments, if the measurement itself is treated as a black box

ii) A strong motivation for a time asymmetric formulation of ordinary quantum mechanics.

iii) A time asymmetry in the description of the quantum measurement itself, at least in the popular decoherence approach. Although we have sketched in the last subsection that the decoherence approach encounters a serious problem in solving the measurement problem, it appears very interesting that, according to this approach, a quantum measurement itself would be a time asymmetric process.

Classification of Ordinary Quantum Mechanics in Three Levels

The level of laboratory descriptions; where quantum measurements are assumed as black boxes.

The level of the measurements process, whereby only the popular decoherence approach is considered

The level of pure quantum mechanics

Considering the time directed energy flux from chapter IV

The time directed energy flux yields a trivial time arrow in the considered description of quantum measurements

The time directed energy flux yields a time asymmetry in the decoherence process, independently from quantum measurements

The time directed energy flux yields a time asymmetry in the formalism of ordinary quantum mechanics. Additionally, a minimal rigged Hilbert space formulation of quantum mechanics is supported.

176

Thus, we saw in this chapter that the time asymmetry of the energy flux (chapter IV) provides a time asymmetry on all levels of the quantum mechanical description: the level of laboratory descriptions in which the measurement is treated as a black box; the level of pure quantum mechanics; and the level of the measurement (or the level of the decoherence process) itself, at least according to the prominent decoherence approach.

As announced in chapter I, so far, I have considered the prominent time asymmetries in quantum mechanics (this chapter), quantum thermodynamics (chapter V), classical electromagnetism (chapter IV) and classical- and quantum cosmology (chapter III and V). The aim of the investigation was to demonstrate a new possible understanding of those time asymmetries by considering the conceptual proposal, regarding the understanding of fundamentality, in chapter II. In the case of the proper and local time asymmetries, which are the radiation arrow and all the quantum mechanical time asymmetries, the alignment of this asymmetries with respect to the explications from the fundamental cosmic time asymmetry, are unclear. Thus, the question of proper time arrows in not completely answered here. Nevertheless, it is show what is needed to answer the question, regarding all the considered prominent time asymmetries, if these asymmetries should be understood as consequences from the fundamental time asymmetry, which would show the alignment of those asymmetries in the cosmic background asymmetry. This, I think, is still an important advantage in order to understand the proper time asymmetry and directions in quantum physics and classical electrodynamics as fundamental and physically important do define the proper future and the proper past in non-conventional ways.

In the next and final chapter of this investigation I shall summarise the conclusions from the different chapters. Also, I will try to sketch the advantages of the proposed understanding of time asymmetries and arrows and briefly sketch some further research opportunities.

Chapter VII

Summary of the Main Conclusions

This final chapter shall summarize the main arguments and conclusions of the entire investigation in order to sketch the proposed understanding of time asymmetries. I shall begin, in section 1, by summarizing the results, motivations and arguments from previous chapters. In the last section, section 2, I will try to present a short overview of the main results of the investigation as well as possible connections to further research.

VII.1. Results, Motivations and Arguments

VIII.1.1. 'Fundamental' Time Asymmetries (Chapter II)

As outlined in chapter I, the main motivation of the entire investigation was the development of a new conception of fundamentality in the context of time asymmetries in physics. This issue is philosophically important because the difference between past and future in everyday experience seems to be one of the most fundamental and basic experiences in human life, but the traditional approaches to time directions in physics seem unsuccessful in grounding this difference on the fundamental structures of physical models and theories, which are used to describe nature. The main reason for this apparent failure is that the fundamental laws of physics, at least in the standard formulations and interpretations, are time-reversal invariant. Thus, in chapter II, I proposed a new notion of fundamentality in the context of time asymmetries that is not based on the time-reversal invariance of fundamental physical laws.

The central philosophical suggestion from chapter II can be seen as the following.

According to definition I from chapter II:

'Suppose L is a fundamental linear TRIL, and S(L) is the solution space with dim(S(L)) = n. I will call a time asymmetry 'fundamental' if and only if:

> *i) There is no more than a countable collection $S_i(L)$ of subspaces of dimensions $m_i < n$ and no more than an uncountable collection $S_i'(L)$ of subspaces of dimension*

$m_i' < n-1$, such that if $f(t) \in S(L)$ is time symmetric, then $f(t) \in S_i(L)$ or $f(t) \in S_i'(L)$ for some i, and if $f(t) \in S(L)$ is time asymmetric, then $f(t) \notin S_i(L)$ and $f(t) \notin S_i'(L)$ for all i.

ii) For time asymmetric solutions $f(t) \in S(L)$, the solution $f(-t) \in S(L)$ refers to the same physical world as $f(t)$ does.'

we find that the fundamentality of a time asymmetry can be based on the structure of the solution space of a crucial law-like dynamic equation. The structure of the solution space is, of course, as fundamental as the considered equation. Thus, according to definition I, it is possible to construct a fundamental time asymmetry based on a fundamental TRIL.

More precisely, If point i) from definition I is fulfilled, it follows that almost all the solutions of the fundamental law are asymmetric in time. If i) is fulfilled, all time symmetric solutions are contained in a subspace with lower dimension than the entire solution space. According to an ordinary measure, that means that the time symmetric solutions are in subspaces of measure zero.

If ii) from definition I is also fulfilled, it follows that the different asymmetric solutions, for example the pair $f(t) \in S(L)$ and $f(-t) \in S(L)$, are directed in the same physical way. If ii) is fulfilled, the sign of t refers only to a non-physical time parameter such as an absolute Newtonian background time. In the context of general relativity (or empirically equivalent spacetime theories), this time is not physical, and it is possible that the formally distinct functions can describe the same physical solution.

Thus, in that situation where i) and ii) are fulfilled, a time asymmetry appears for fundamental reasons owing to the structure of the solution space, even if the fundamental laws are time-reversal invariant. Thus, by following this suggestion, we avoid the crucial problems that traditionally arise in understanding time directions as fundamental properties of physical theories that are used to describe nature.

Of course, to show that this new concept provides any useful applications in the physical models that we use today, I concentrated on different fields of application for this philosophical idea. Thus, in chapter III I argued that the suggested understanding of 'fundamentality' is applicable in classical cosmology.

VII.1.2. Time Asymmetries in Cosmology (Chapter III)

In chapter III, I made some different claims. The first was to argue against a suggestion from Price (1996) and (2002). His view suggests that nature, if at least partially correctly described by modern physics, has no time asymmetry which is based on any fundamental physical consideration. Price (1996) and (2002) explicated his view also in the field of classical cosmology. Thus, I discussed his arguments at the beginning of chapter III in order to clarify that, in my view, these arguments are not compelling. In the context of classical cosmology, my main argument was based on the fact that Price (1996) considers only a description of spacetime that includes only one dynamic variable, the scale factor [which is called the radius of the universe in Price (1996)]. Only under this assumption can Price argue that the Big Bang and a hypothetical Big Crunch are not distinguishable on the basis of physical properties.

If this argument were right, of course, Price could argue that, on this basis, a cosmological time arrow cannot be defined using the physical properties of cosmological models. However, as I argue in chapter III, it seems implausible to use a description of spacetime where the only dynamic variable is the scale factor. According to cosmological models, the dynamics of the matter and energy content is described by an additional field, the matter field. Moreover, this additional field seems necessary if the cosmological model is supposed to include the dynamics of matter and energy, which cannot be described by the scale factor alone. Thus, for physical reasons, the cosmological model that Price assumes seems implausible, and his conclusions cannot be translated to a model with other dynamic variables in addition to the scale factor, because in this case the matter field would provide additional properties that could be used to distinguish between a Big Bang and a hypothetical Big Crunch. Therefore, the entire cosmological discussion of Price (1996) and also in parts of Price (2002) seems implausible.

Nevertheless, in chapter III, before I demonstrated an application of the new concept of 'fundamentality', I focused on various approaches to the time arrow in classical cosmology. This was useful for confirming that the application of the idea from chapter II provides an advantage in the philosophical understanding of time asymmetries in classical cosmology.

Thus, I first focused on accounts that try to define the cosmological time asymmetry via the time behaviour of entropy in the three-dimensional universe. I referred to Price (2002) and Ćirković and Miloševic-Zdjelar (2004) in order to distinguish between different types of accounts. I used the distinction suggested by Ćirković and Miloševic-Zdjelar (2004). According to their analysis, motivated, I think, by Price (2002), there are three different types of entropy-based accounts:

      i) causal–general accounts,

      ii) acausal–particular accounts and

      iii) acausal–anthropic accounts.

In chapter III, I showed that none of them can define a cosmological time asymmetry based on fundamental physical properties of the cosmological model.

Briefly:

The causal–general accounts seem simply unable to provide an explanation for the low-entropy past, so they cannot show that the entropy behaviour of our particular universe arises for non-accidental reasons.

The acausal–particular accounts, of course, refer to particular boundary conditions, so they yield not a fundamental physical understanding of the origin of the time asymmetry.

Moreover, the acausal–anthropic account of Ćirković and Miloševic-Zdjelar (2004) cannot provide a physical time direction in the cosmological model (see chapter III) because, based on multiverse theories in cosmology, it explains the occurrence of some particular boundary condition in our particular cosmic domain[49], but time symmetric boundary conditions are possible and in fact they are more likely than the time asymmetric ones. Consequently, Ćirković and Miloševic-Zdjelar (2004) also seem unable to define a time direction based on fundamental physical properties of the considered cosmology model, because time symmetric domains are more likely and hence, the origin of the observed asymmetry in our particular domain should be understood as accidental, even if it is not surprising in that account that our particular cosmic domain seems to have such boundaries.

Another claim of chapter III was that the well-known accounts of the cosmological time arrow, which are based on the accelerated expansion of the universe, also cannot provide a cosmological time asymmetry based on fundamental properties of the considered model. I investigated these types of accounts in contrast to the entropy-based accounts. However, they encounter the same problems as the traditional definition of the cosmological time arrow (according to a non-positive accelerated expansion of the universe). Because, in the considered model, a closed spacetime is still possible (see chapter III), the origin of the cosmological time asymmetry would not be fundamental but, again, accidental.

Thus, after showing that a successful application of the concept in chapter II to classical cosmology would provide a *more fundamental* understanding of cosmological time asymmetries, I started to develop my proposal in chapter III. I showed, by using the

---

[49] Probably, a better description is to say that there explain why it is not surprising that *our* cosmic domain has such unlikely boundaries.

cosmological outcomes from Castagnino, Lara and Lombardi (2003) and Castagnino and Lombardi (2009) that, given two crucial conditions, the crucial dynamical equation in classical cosmology provides a solution space that fulfils definition I from chapter II. The crucial conditions were:

> i) In all considered solutions it is possible to define cosmic time.

> ii) The cosmological model that is used to describe the dynamics of spacetime should include more than one dynamic variable.

A physical motivation for the second condition was easy to identify because without at least one additional scalar matter field, the dynamics of matter and energy in the universe could not be described adequately.

The first condition, in contrast, is motivated by methodological considerations. In fact, if condition i) were not fulfilled, we could consider only time directions referring to the proper time coordinate of only one world line. Even two parallel world lines with synchronised proper times could have independent time *directions*. Thus, if the consistent *possibility* of a time direction that is valid for more than one world line is required, we must refer to spacetimes where cosmic time is definable, at least in the context of classical cosmology.

The second part of chapter III shows that, if the crucial conditions are fulfilled, the solution space of the crucial dynamical equations (in this case the Hamilton equation) provides a situation in which definition I from chapter II is fulfilled.

Moreover, the sign of the formal time coordinate *t*, the sign of which could be used to distinguish between two time-mirrored solutions, is a mathematical construct, and the *sign* (not the coordinate itself) has no physical meaning in this context. So, I was able to argue that two formal, time-mirrored, asymmetric solutions describe the same physical spacetime. This is a consequence of a Leibniz argument: Because all internal and external properties of such spacetimes are identical, for fundamental reasons arising from general relativity (or empirical equivalent spacetime theories). Thus, in classical cosmology, given the two conditions, definition I from chapter II is applicable and defines a fundamental time asymmetry in cosmology.

As I argued in chapter V, this fundamental time asymmetry provides new perspectives on the prominent time arrow in statistical and thermic physics. But before, in chapter IV, I considered the field of classical electrodynamics and the arrow of radiation.

VII.1.3. The Arrow of Radiation (Chapter IV)

I demonstrate, firstly, the intrinsic time reversal invariants of classical electrodynamics. I began by demonstrating that classical electrodynamics should be seen as a time symmetric theory in all of its relevant properties. Not only are the fundamental equations, the Maxwell equations, time-reversal invariant, but also the equations of motion for moving charges in classical electrodynamics are time symmetric. This conclusion was demonstrated by (4.12):

$$m\frac{d^2 x^\mu}{d\tau^2} = G^\mu + \Gamma^\mu \xrightarrow{\ T\ } m\frac{d^2}{d\tau'^2} x'^\mu(\tau') = G'^\mu(\tau') + \Gamma'^\mu(\tau') \,.$$

(See chapter IV, section 1). An analogue analysis to the cosmological case (chapter III), was not considered in the case of classical electrodynamics. The reason for this omission was that

i) in the context of classical electrodynamics, two time mirrored solutions would surly describe different physical situation, simply because the solutions to Maxwell's equations are not representing a whole spacetime (as in the cosmological case). Thus, condition ii of definition I (chapter II) cannot be fulfilled in this case.

ii) an analogue consideration in the case of classical electrodynamics would require the numerical analysis of the solution set of the crucial equations, which is in any case, provided by the non-linearity of the Lorenz-Maxwell equations, not linear. Hence, the applicability of condition I (even in possible generalisations) would be questionable too.

Thus, given that situation, the first crucial question concerns a plausible definition of the empirical arrow of radiation. In this respect, my favoured characterisation of the arrow of radiation is identical to the one given by Frisch (2000) or Rohrlich (2005), and it is a result of the fact that fully advanced radiation, of a specific kind (see chapter IV) does not occur as associated radiation with an accelerated charge.

However, on the basis of this characterisation, I firstly described my criticism of the account by Rohrlich (2005). This account tries to explain the absence of fully advanced radiation by arguing that its occurrence would violate a time asymmetric causality principle. In chapter IV, I argued that this is not a valid explanation of the non-occurrence of fully advanced radiation because it is not clear in which physical way the 'time arrow of causality', if valid, should orientate electromagnetic phenomena in time. Additionally, even if such an arrow were given, it is unclear why electromagnetic processes should be guided by it and not vice

versa. Thus, I think the account of Rohrlich (2005) does not convincingly illuminate the origin of the arrow of radiation.

I also considered the account stipulated by Frisch (2000), which characterizes the arrow of radiation as I do in chapter IV. Further, Frisch (2000) simply argued that a new law in classical electrodynamics, the retardation condition, stipulates that electromagnetic fields associated with electric charges cannot be fully advanced. Frisch himself does not offer an explanation for this condition, which seems at least more ad hoc than necessary (see also chapter IV).

However, as in the investigation from chapter III, I also considered the time symmetric view of Price (1991a), (1991b), (1994), (1996) and (2006). I claimed, again, that his time symmetric view is not plausible for physical reasons, so an understanding of the arrow of radiation based on fundamental physics is not doomed to fail for the reasons given by Price (1991a), (1991b), (1994), (1996) and (2006).

Because his arguments in (1991a), (1991b), (1994) and (1996) seemed to be slightly different from that in Price (2006), I examined these investigations separately.

Specifically, Price (1991a), (1991b), (1994) and (1996) used a different characterisation of the arrow of radiation than I do [see chapter IV]. Price seemed to characterize the arrow of radiation according to the situation in which absorbed electromagnetic waves are not coherent whereas emitted ones are. Therefore, Price argued that the arrow of radiation is based on a numerical asymmetry between the occurrence of electromagnetic absorbers and emitters, which is only a macrophysical property. He argued that microphysical electrodynamics, especially if it is classical, is time symmetric, and he grounded his view on a new interpretation of the absorber theory from Wheeler and Feynman (1945) [see also Price (1991a), (1991b), (1994) and (1996)].

Motivated by Frisch (2000), I argued in chapter IV that such an interpretation of the Wheeler–Feynman absorber theory is not plausible for physical reasons. In the main chapter, I refer simply to the findings of Frisch (2000), which seem sufficient for rejecting Price view. Nevertheless, I broaden the discussion a bit in digression IV.B. Here it was demonstrated that his reinterpretation of the absorber theory is in fact a new physical theory, and this new theory contradicts Maxwell's equations or the empirical data from physical observations if the central equation of Price's observer theory is not understood as (4.B.7):

$$\sum_{k \neq q} F_{adv}^{k} = F_{ret}^{q} - F_{adv}^{q} \Leftarrow \sum_{k \neq q} F_{ret}^{k} = 0 \,.$$

However, even if Price suggestions are understood as described by (4.B.7) this seems puzzling, because in this case Price's absorber theory has no further relevance for the arrow

of radiation (see chapter IV, digression IV.B). Thus, based on the physical considerations discussed in chapter IV his characterisation of the arrow of radiation seems implausible and therefore also his proposed solution to the understanding of the radiation arrow.

Moreover, the investigation in chapter IV showed that the suggestions of Price (2006) also seem implausible. Price (2006) seemed to suggest that the arrow of radiation and the thermodynamic time asymmetry could have the same origin. To support this claim, he suggested a strong analogy between electromagnetic waves and waves in a classical wave media. Using thermodynamics, Price showed that some crucial time asymmetries for classical waves in a medium, specifically the difference in the numbers of coherent wave emitters and coherent wave absorbers, can be understood similar to time asymmetries in classical thermodynamics.

However, as I showed in chapter IV this analysis cannot be seen as relevant to our understanding of electromagnetic waves in relativistic theories, for the simple reason that, according to special relativity, electromagnetic waves have, in general, no wave medium. Price (2006) argues that QED, as a field theory, can provide the missing analogy between electrodynamics and wave phenomena in classical media. However, as argued in chapter IV, this suggestion is even more implausible because the quantum fields in QED are not the type of media in which waves propagate as in classical physics, where waves are constructed out of the motion of many constituents of the medium. Thus, together with the suggestions in Price (1994) and (1996), I also reject that in Price (2006).

Moreover, in addition to the suggestions in Price (1994), (1996) and (2006), I investigated the prominent account of Zeh (1989), (1999). Zeh's account of the arrow of radiation is based on the Sommerfeld condition. In fact, he argued that the Sommerfeld condition could be fulfilled in our particular spacetime region (provided by astrophysical boundaries), and thus the total radiation field in our particular spacetime region cannot be fully advanced.

I agree with Zeh on that point, but, as argued in chapter IV, the fact that the Sommerfeld condition is fulfilled is relevant only for the *total* wave field in our particular spacetime region. Some local electromagnetic wave fields, for example an emitting electron, could be associated with fully advanced radiation as well as with fully retarded radiation, given

classical electrodynamics *and* a fulfilled Sommerfeld condition. Thus, Zeh's account seems unable to explain the origin of the arrow of radiation for non-total wave fields.

My proposal was that two crucial conditions on the energy-momentum tensor alone (in contrast to Castagnino, Lara and Lombardi (2003)), which yielding an energetic time asymmetry, really yield the retardation condition of Frisch (2000).

These conditions are that the energy–momentum tensor is of type I and that the dominant energy condition is fulfilled (see chapter IV). The asymmetry of the energy flux defines the proper time directions but, in the proposed cosmological framework, the alignment of the crucial asymmetry in the more fundamental cosmic time asymmetry remains unclear.

Nevertheless, chapter IV demonstrated the crucial fact that, in the applied cosmological framework and with the fulfilled conditions, this energy flux is physically (in a non-conventional sense) time asymmetric with respect to proper times, even if the alignment of the asymmetry, with respect to the cosmic asymmetry, as well as the labelling is conventional. Thus, the proper time asymmetry of the radiation arrow cannot yet be understood as a consequence of the fundamental time asymmetry, but the consideration shows what is needed to put this understanding further: an understanding of the connection between the alignments of the proper time asymmetry with respect to the more fundamental cosmic time asymmetry.

However, until the end of chapter IV, the investigation was based mostly on classical physics, specifically, classical cosmology and classical electrodynamics. To generalize the ideas from chapters II and III, in chapter V I considered quantum cosmology and quantum thermodynamics.

VII.1.4 Time Asymmetries in Quantum Cosmology, Entropy and the Second Law of (Quantum) Thermodynamics (Chapter V)

Chapter V aims to make two main claims. The first was that the fundamental time asymmetry, investigated in chapter III, yields time asymmetric explications not only in classical cosmology but also in quantum cosmology, where quantum cosmology is understood as avoiding quantum gravity. Consequently, one of the main conclusions from

chapter V, which was established reasonably quick, was that in quantum cosmology, restricted by the conditions mentioned in chapter III, the fundamental time asymmetry from chapter III can be identified.

The second claim of chapter V was the development of a new understanding of what I consider the most prominent time arrow in physics, the arrow of time in thermodynamics. In chapter I, I outlined the view that the concept of the arrow of time in classical thermodynamics fails to provide a plausible concept for *fundamental* time directions based on entropy behaviour. Chapter V considered particular quantum cosmological models and demonstrated that the entropy behaviour for some prominent definitions of entropy is time asymmetric in the considered models.

To underline the advantage of the suggested understanding of a time arrow in quantum thermodynamics, I also considered the account of Allahverdyan and Gurzadyan (2002), also as representative, at least in parts, for standard accounts regarding thermodynamic time asymmetries. The aim of this section was to show how this account fails to provide a physical understanding of the origin of the time arrow in quantum thermodynamics. I showed (see chapter V) that Allahverdyan and Gurzadyan (2002) used artificial or epistemically motivated definitions and approximation methods to draw physical conclusions. I did so by investigating their IIA and DWM assumptions / description.

In contrast to Allahverdyan and Gurzadyan (2002), in the second part of chapter V I showed that some prominent definitions of entropy, in the considered cosmological models, produced an arrow of time in quantum thermodynamics, which is based on physical properties. Moreover, this arrow of time seems physically effective independently of an epistemic or ontic interpretation of the entropy value.

More precisely and regarding my proposals in chapter V:

Firstly, I sketch that, in quantum cosmology, a fundamental time asymmetry appears in a similar way to the fundamental time asymmetry in classical cosmology (chapter III). Additionally, I explicated the time asymmetric consequences of this abstract time asymmetry in a particular model. There, I showed that the structure of the Einstein equations, in the considered model, can be captured in (5.34):

$$H^2 = \gamma^2.$$

The lhs of (5.34) is the Hubble coefficient squared, and the rhs is a positive real function of cosmic time (or zero). According to (5.34), I argued that the two solutions for the Hubble coefficient, which are intrinsically time asymmetric in the considered model, describe the same physical world. This explicates the fundamental abstract time asymmetry in the considered model.

Next, the chapter turned to a perhaps very technical but necessary subject. I tried to demonstrate by considering many physical investigations that the expectation value of the particle number operator is time asymmetric in the considered model if additional assumptions are fulfilled. Prima faces, the additional assumptions were:

a) That there is a radiation- or matter-dominated phase in the cosmic evolution,

b) That the considered spacetimes are hyperbolic curved.

c) That the total action is describable as the sum of the Einstein-Hilbert Action and the contribution from a *scalar* matter field

d) That the field operators are independent of self-interactions.

e) That the initial vacuum state can be described in analogy to a classical vacuum.

The assumptions a) and b) seem reasonable for a cosmological model if it should be a candidate to describe also our particular spacetime.

Assumption c) seems, at least, a reasonable candidate for a semi classical description of our particular spacetime, even if other contributions to the total action or other forms of matter fields cannot be ruled out as correct descriptions of our particular spacetime.

Fortunately, the assumptions d) and e) has turned out to be not crucial for this investigation, as I argued in chapter V.

Thus, the time asymmetric behaviour of the particle number operator is not fundamental (in the sense from chapter II) but appears for physical reasons in spacetimes, described by the reasonable considered quantum cosmological model. It seems, therefore, that the expectation value of the particle number can be time asymmetric in our particular universe.

Motivated by this finding, the next step in building an arrow of time in quantum thermodynamics was to consider some prominent definitions of entropy. Because of the large set of possible definitions, the first problem was to decide which type of entropy to investigate in this context. To avoid certain crucial problems, I decided to refer only to certain possible definitions that are used in quantum thermodynamics. The first was formulated by Landau and Lifshitz (1970), and the second was formulated by Glansdorff and Prigogine (1971). In both cases, the entropies were shown to be increasing functions of the particle number.

Additionally, using the suggested understanding of the time arrow in quantum thermodynamics, I showed that several advantages over traditional understandings necessarily occur.

The first is that a second law of thermodynamics can be formulated in a way which ensures that the entropy value decreases with decreasing cosmic time. In traditional accounts, the entropy increases theoretically in both time directions, but according to the analysis in chapter V, the entropy increases in only one time direction, the future, and decreases in the other (apart from fluctuations, of course).

The second advantage was that, for both entropy definitions, the thermodynamic time arrow is understandable as an effective property of the physical model used to describe nature. The effectiveness of the time arrow is independent of an epistemic or ontic interpretation of the entropies. Even if the entropy were understood as a purely epistemic quantity, the time arrow would be physically effective because the arrow is grounded in the behaviour of the particle number. Thus, the arrow of time, or more precisely the physical effectiveness of the time asymmetry, is neutral with regard to an epistemic or ontic interpretation of entropy.

All together, the fundamental time asymmetry, investigated in chapter III, was easily translatable to quantum cosmological models in which physical explications produces, under some assumptions, thermodynamic time arrows for some entropy definitions.

VII.1.5 Time Arrows in Ordinary Quantum Mechanics (Chapter VI)

Before the investigation continues by considering the field of ordinary quantum mechanics in chapter VI, there was an important problem to avoid. The fundamental time asymmetry as well as the time asymmetry of the relativistic energy flux was found without considering the quantisation of gravity. Thus, in the entire investigation, general relativity (or empirically equivalent spacetime theories) is treated as the fundamental theory on which the time asymmetry is grounded and called 'fundamental'. However, if we consider quantum mechanics, which could also be seen as a fundamental theory, the question arises as to which theory should be treated as fundamental. Thus, in chapter VI, to avoid, among other, crucial problems with time definitions in quantum gravity, I considered both general relativity and quantum mechanics as fundamental but as having different fields of application. Specifically, I considered that quantum physics is not applied to gravity and focused only on ordinary non-relativistic quantum mechanics. On one hand, this seems slightly unsatisfactory, but, I think, on the other hand it is acceptable as long as fundamental physics lacks a well-established theory of quantum gravity.

Considering this restriction, chapter VI demonstrated that the time asymmetric energy flux (chapter IV) provides different time arrows in quantum mechanics. For this investigation, the field of quantum physics was divided into three different levels.

> One level, and in this chapter the most fundamental one, was the level of the quantum mechanical formalism itself, avoiding any quantum measurements or analogous physical processes.

> The other level of quantum physics that can avoid the measurement problem was the level of laboratory descriptions, in which quantum measurements are treated as black boxes. On this level, because laboratory descriptions are used, it was clear by definition whether a quantum interaction should be treated as a measurement or not.

> The third level was the measurement itself. Because it is completely outside of the scope of this investigation to solve the measurement problem, in chapter VI I considered only one approach, the very prominent decoherence approach, of the connection between quantum physics and classical physics.

Using these three levels of quantum mechanics, the investigation in chapter VI shows that a time asymmetry is imbedded in each level.

More precise:

The time asymmetry of the energy flux, considered in chapter IV, produces a time asymmetry in the description of laboratory experiments. This was demonstrated by the fact that laboratory experiments, in which measurements are described as black boxes and used to define subsystems, can be described as an energy cascade in cosmic time. The time-mirrored energy cascade is forbidden by the time asymmetry of the energy flux. Thus, quantum experiments in laboratory descriptions are time asymmetric, at least in the considered set of spacetimes, which, on this level of description, seems to include our own.

According to the pure formalism of ordinary quantum mechanics, chapter VI shows that even this level of quantum mechanics exhibits time asymmetry. This asymmetry arises from the fact that the eigenstates of a system can be mapped to the complex plane. Some correspond to decaying energy states, and their complex conjugated states correspond to growing energy states, where complex conjugation is identical to the mirroring of time. Here the symmetry of the eigenstates is due to the time-reversal invariance of the Schrödinger equation. However, it turns out, as consequents of the time asymmetric energy flux, that only one-half of the theoretical eigenstates are not forbidden. Moreover, this situation is also suggested in the rigged Hilbert space description of quantum mechanics. Thus, the

analysis also provides supporting arguments for the rigged Hilbert space formalism in quantum mechanics. This was argued according to (6.2):

$$K\phi_\pm = \phi_\mp \, .$$

Here, the two subspaces of the Hilbert space are exactly time-mirrored spaces of each other. Moreover, each subspace has one favoured time direction. If the Hilbert space were replaced by one of the subspaces in (6.2), the formalism becomes time asymmetric in exactly the way that is motivated by considering the time asymmetric energy flux from chapter IV. Moreover, the consideration of the time asymmetric energy flux solves two philosophical problems with the rigged Hilbert space approach.

Firstly, in the traditional approach, the replacement of the Hilbert space by only *one* of the subspaces in (6.2) is not explained but is only assumed in order to provide a time asymmetric formalism.

Secondly, it is also unclear which of the subspaces should be used to replace the traditional Hilbert space.

Both issues are clearly answered by the investigation in chapter VI.

The final argument in this investigation focused on the level of the connection between ordinary quantum mechanics and classical physics, at least on one description of parts of this connection, the prominent decoherence approach. The analysis in chapter VI showed that the formalism of the decoherence approach yields a time asymmetry (see also Castagnino, Lara and Lombardi (2003)). The crucial aspect of this point was that the time evolution of a quantum system, described by mixed density operators, for infinitely large times, comes to a limit description that is analogous to a classical ensemble.

In the traditional decoherence approach, time evolution, which produces the decoherence effects, is possible in both time directions with the same result. Here again, this formal symmetry is provided by the time-reversal invariance of the Schrödinger equation. However, the time evolution in the past direction seems forbidden by the time asymmetry of the energy flux from chapter IV. This shows at least that the physical process of decoherence in a complex quantum system is time asymmetric in the considered spacetime set, which is important in itself.

Thus, even if the decoherence approach cannot solve the measurement problem, the time asymmetry would be important because it shows that the decoherence process itself is time asymmetric.

Thus, the picture drawn according to the connection between the different time asymmetries in the different fields of quantum physics can be seen, I think, as a physically

plausible view of the difference between 'past' and 'future', at least in the considered spacetime set. Even if we have to note that, as well as in the radiation case, the alignment of the quantum physical time asymmetries, because there are consequences of the time asymmetry of the relativistic energy flux, in the explications of the fundamental time asymmetry in cosmology is still needed in order to understand the quantum physical time asymmetries as consequences from fundamental asymmetries.

VII.2. Summary and Motivations for Further Investigations

All together, the results of this investigation can be summarized very briefly:

i) A naturalistic definition of fundamentality in the context of time asymmetries that is not based on the time-reversal invariance of fundamental physical laws is possible and well-motivated (see chapter II).

ii a) This definition of fundamentality can be successfully applied to physical theories and models. This application provides a fundamental time asymmetry in the field of cosmology [see also Castagnino and Laciana (2002) and Castagnino, Lara and Lombardi (2003), Castganino and Lombardi (2009) as well as chapters III and V].

ii b) The crucial conditions for the applicability of definition I are methodologically and physically motivated. They are: The spacetimes under consideration should allow the definition of cosmic time, and the dynamics of the universe should be described by more dynamic variables than the scale factor alone (see chapters III and V).

iii a) An explanation of the retardation condition in Frisch (2000) and thus an explanation of the origin of the arrow of radiation can be found in the time asymmetry of the relativistic energy flux in spacetimes similar to our [see also chapter IV and Castganino and Lombardi (2009)].

iii b) The crucial conditions for producing the arrow of radiation are conditions on the particular form of the energy–momentum tensor that seem to be fulfilled in our particular universe, at least in classical electrodynamics.

iv a) The explication of the fundamental time asymmetry in certain quantum cosmological models provides a time asymmetry of the expectation value of the particle number operator [see also Castagnino and Laciana (2002) as well as chapter V].

iv b) The crucial conditions for the occurrence of time asymmetries in the behaviour of the particle number operator are

> a) That there is a radiation- or matter-dominated phase in the cosmic evolution,

> b) That the considered spacetimes are hyperbolic curved.

> c) That the total action is describable as the sum of the Einstein-Hilbert Action and the contribution from a *scalar* matter field

The assumption that this situation obtains in our particular spacetime seems at least plausible.

iv c) The time asymmetry of the particle number operator provides different time arrows in quantum thermodynamics. The difference between them is provided from the difference between the considered entropy definitions. It was argued that, in quantum thermodynamics, the time asymmetry of the particle number operator produces time arrows for the non-equilibrium entropy of Landau and Lifshitz (1970) and for the phenomenological entropy of Glansdorff and Prigogine (1971). The occurrence of those time arrows has the same requirements as the time asymmetry of the particle number operator.

iv d) The time arrows are physically effective properties of the quantum cosmological model independent of an epistemic or ontic interpretation of entropy; they arise from the time asymmetry of the particle number and are physically operative regardless of whether an epistemic or an ontic interpretation of entropy is favoured (see chapter V).

v a) The time asymmetric energy flux (chapter IV) provides several time asymmetries in ordinary quantum mechanics. [see also Castagnino, Lara and Lombardi (2003) as well as chapters IV and VI].

v b) The different time asymmetries in quantum mechanics were considered in

> a) laboratory descriptions, in which quantum measurements are treated as black boxes and are used to describe subsystems;

> b) in the formalism of quantum mechanics itself, where no measurement is assumed; and

> c) in the decoherence process of complex quantum systems (described by mixed density operators), which provides a time asymmetry in quantum measurements if the prominent decoherence approach of the measurement

problem is assumed to yield the correct description of quantum measurements (see chapter VI).

v c) The advantage of the time arrow in the quantum formalism over traditional accounts such as the ordinary rigged Hilbert space description is that the consideration of the time asymmetric energy flux in the ordinary Hilbert space formalism of quantum mechanics provides the formal structure of the well-known rigged Hilbert space description of quantum mechanics. Thus, the rigged Hilbert space formalism can be motivated by physical considerations of the energy flux in the considered set of spacetimes (see also Castganino, Lara and Lombardi (2003). But, the consideration of the time asymmetric energy flux provides only a physical distinction between the two possible time-mirrored replacements for the Hilbert space, if the alignment of the proper time asymmetry of the energy flux in the fundamental and physical time asymmetry in cosmology could be established in an nonconventional way. Thus, it is possible to make a physically motivated decision between the two possible formulations of the minimal rigged Hilbert space account only if this further question could be solved.

v d) The time asymmetry in the decoherence account is located in the decoherence process itself. The time asymmetry is important because the decoherence process itself is time asymmetric in the considered spacetime set. Thus, for long times, complex quantum systems would still exhibit an evolution in which the interference term vanishes. At least this part of the connection between quantum physics and classical physics, the decoherence itself, is, therefore, understandable as time asymmetric in the considered spacetime set.

Altogether, this investigation has shown that the most popular proper time arrows in physics could be understandable as explications of a more abstract time asymmetry in the solution set of cosmological equations or as a necessary by-product of such explications in the considered models, which can describe our own spacetime at least approximately, only if the alignment of the energy flux in the fundamental cosmic time asymmetry could be established in a nonconventional way. Thus, the difference between two cosmic time directions, according to cosmic time, can be understood on the basis of physics. However, the difference between two local and proper time directions can only understood in a similar manner if the alignment of the relativistic energy flux in the cosmic time asymmetries could be established nonconventional. Thus, the conjecture that time directions cannot be based on fundamental physics is wrong regarding cosmic time, especially regarding the thermodynamic time arrow in hyperbolical curved spacetimes (chapter V). However, regarding proper times directions, the property of the alignment of some prominent arrows is still yet an open question. Nevertheless, at least the *directedness* of crucial time

coordinates or parameters in some physical models can be seen as fundamental even if, of course, the labelling 'future' and 'past' remains conventional.

Additionally and as the final aspect of the entire investigation, I will try to sketch some fruitful possibilities for further research connected with the results of this investigation.

i) The crucial conditions on the energy–momentum tensor from chapter IV seem to be met in our particular spacetime, at least in classical descriptions. However, it seems also that those conditions are not met during e.g. Hawking evaporation, the Casmir effect or in squeezed vacuum [see Visser (1996) or Barcel'o and Visser (2002)]. To investigate whether the time asymmetry of the energy fluxes in chapter IV is understandable as more than just valid for physical models which exclude those effects and processes, it is necessary to investigate energy fluxes in QFT-models in general, which, of course, could led to the well-known problems with quantum gravity

ii) To provide the time asymmetry of the expectation value of the particle number operator, it was necessary to assume that the dynamics of the matter and energy content of the universe can be described by a *scalar* matter field. This seems to be a plausible assumption in our particular universe. Nevertheless, to confirm that this is really the case, or to analyse if the asymmetry occurs in models with different matter field, we need more sophisticated models for the dynamics of the matter (also dark matter) and energy content of our particular spacetime.

iii) An effective investigation of the possible and plausible definitions of entropy would be very helpful in order to investigate whether the time asymmetry of the particle number introduced a time asymmetric behaviour of entropy in an entire set (perhaps a set with some specific fundamental characteristics) of possible entropy definitions.

iv) The fundamental time asymmetry in cosmology, in both semi-classical and classical cosmology, requires crucial conditions on the set of considered spacetimes. One is that in this spacetimes, cosmic time can be defined. Chapter III shows that this requirement is motivated by methodological considerations because without this requirement it is impossible to define time directions which are physical important for more than one particular world line. Thus, if we want to discuss fundamental time directions that are valid for more than one particular elementary physical system, this requirement is necessary, at least in the context of general relativity or empirically equivalent spacetime theories. However, the Einstein equations, whose solution space is, of course, not fully accessed, allows non time-orientable spacetimes, so on the level of the fundamental equations and without any

requirements, we will not find a time direction for more than a local environment of a spacetime point according to one particular world line. Philosophically, I think it appears interesting to investigate the following ambivalent situation:

> a) if we set a *methodological* requirement for discussing the time direction (in any sense) for more than one world line, we find, according to physics, a fundamental time asymmetry that has crucial consequences in e.g. quantum thermodynamics. This time direction is not already imposed by the requirement; nevertheless,

> b) without this *methodological* requirement, physics cannot show any fundamental time asymmetry for more than a world line.

v) On the basis of the time asymmetric behaviour of some entropy definitions in quantum thermodynamics, it appears interesting to reconsider the discussion of an epistemic versus an ontic interpretation of entropy. The time arrows (see chapter V) are physically effective whether an epistemic or an ontic interpretation of entropy is favoured. Thus, it would be interesting to know whether the arguments in the discussion of the entropy interpretation are sensitive regarding to the results of this investigation. This could be the case in two different ways, I think.

> a) On one hand, because the physical effectiveness of the entropic time arrows (not obviously the content of entropy itself) in thermodynamics can be reduced to the behaviour of the particle number, an epistemic interpretation seems supported by the recognition that all time asymmetrical physical effects produced by the entropy behaviour represent, on a more fundamental level of description, the behaviour of the particle number operator.

> b) On the other hand, the results provide a physically effective time arrow associated with the entropy behaviour, which seems to be a good result for an ontic interpretation of entropy.

vi a) A fundamental time asymmetries, like those constructed in this investigation, seem to provide new arguments in the discussion of an understanding of causality. On the one hand, as Price (1996) showed, the asymmetry of causality must not be connected with physical time asymmetries. On the other hand, time asymmetries in physics could, perhaps, be used to define the asymmetry of causation as an objective property of nature or the physical model which are used to describe nature. The fundamental time asymmetry could be used to introduce an asymmetry between events in the 'past' and the 'future' (conventionally labelled of course, but physically distinguished). If they are related by a physical connection, this could be labelled as cause and effect, which could be used to define some type of causality.

vii) It appears interesting to me to investigate whether the subjective time experience of human beings in everyday life, as captured by the Newtonian limit of proper times, is directed because of the properties of the time asymmetric energy flux. Alternatively, even if fundamental time directions occur in nature, or the physical models that we use to describe nature, it would be useful to investigate whether the time-directed experience of everyday life arises from those physical structures or maybe from biological or other structures relevant for human beings but not necessarily from physics.

viii) Also it appears extremely important to me to clarify if the alignment of the proper and local time asymmetries with respect to the cosmic time asymmetry can be established in an nonconventional way. Only this step, of course, could finish the task of understanding local and proper time asymmetries as consequences from fundamental physics or at least some fundamental physical theories that we use to describe nature.

Literature

In 'Introduction and Motivation'

A. E. Allahverdyan and V. G. Gurzadyan (2002), *J. Phys. A: Math. Gen.* **35** 7243–7254

R. C. Bishop (2004), *International Journal of Theoretical Physics* **43**(7/8) 1675-1687

L. Boltzmann (1897), *Ann. Phys., Lpz* **60** 392.

M. Castagnino, F. Gaioli and E. Gunzig (1996), *Found. Cosmic Phys.* **16** 221.

M. Castagnino and E. Gunzig (1997), *Int. J. Theor. Phys.* **36** 2545.

M. Castagnino and R. Laura (1997), *Phys. Rev.* A **56** 108.

M. Castagnino (1998), *The global nature of time asymmetry and the Bohm–Reichenbach diagram Irreversibility and Causality.* In Proc. G. 21 Goslar 1996, eds. A Bohm, H Doebner and P Kielarnowski (Berlin: Springer).

M. Castagnino and E. Gunzig (1999), *Int. J. Theor. Phys.* **38** 47.

M. Castagnino, J. Gueron and A. Ordonez (2002), *J. Math. Phys.* **43** 705.

M. Castagnino, G. Catren and R. Ferraro (2002), *Class. Quantum Grav.* **19** 4729.

M. Castagnino and C. Laciana (2002), *Class. Quantum Grav.* **19** 2657–2670.

M. Castagnino, L. Lara and O. Lombardi (2003), *Class. Quantum Grav.* **20** 369–391.

M. Castagnino and O. Lombardi (2009), *Synthese* **169** (1):1-25.

M. M. Ćirković and V. Miloševic-Zdjelar (2004), *Found. Sci.* **9**(1) 1–24.

R. P. Feynman (1964), *The Feynman Lectures on Physics* vol 1 (New York: Addison-Wesley).

M. Frisch (2000), *Brit. J. Phil. Sci.* **51** 381–410.

J. D. Jackson (1999), *Klassische Elektrodynamik* (Berlin, New York: Walter de Gruyter).

L. D. Landau and E. M. Lifshitz (1970), *Statistical Physics* vol 5 (Oxford: Pergamon).

H. Price (1994), *Brit. J. Phil. Sci.* **45** 1023–1028.

H. Price (1996), *Time's Arrow and Archimedes' Point* (Oxford/New York: Oxford University Press).

H. Price (2002), *Burbury's Last Case: The Mystery of the Entropic Arrow.* In Time, Reality and Experience, ed. C Callender (Cambridge: Cambridge University Press).

H. Price (2006), *Stud. Hist. Philos. Mod. Phys.* **37** 498–527.

P. Glansdorff and I. Prigogine (1971), *Thermodynamic Theory of Structures, Stability and Fluctuations* (New York: Wiley).

F. Rohrlich (2005), *Phys. Rev.* E **72** 057601.

H. D. Zeh (1989), *The Physical Basis of the Direction of Time* (Berlin: Springer Verlag).

H. D. Zeh (1999), *The Physical Basis of the Direction of Time* (3rd edn) (Berlin: Springer Verlag).

In 'Fundamental Time Asymmetries'

A. E. Allahverdyan and V. G. Gurzadyan (2002), *J. Phys. A: Math. Gen.* **35** 7243–7254

R. C. Bishop (2004), *International Journal of Theoretical Physics* **43**(7/8) 1675-1687

A. Bohm, M. Gadella, and S. Wickramasekara (1999), Some little things about rigged Hilbert spaces and quantum mechanics and all that. In *Generalized Functions, Operator Theory, and Dynamical Systems,* ed. I. Antoniou and G. Lumer, 202–250 Chapman & Hall/CRC, Boca Raton.

M. Castagnino, L. Lara and O. Lombardi (2003), *Class. Quantum Grav.* **20** 369–391

M. Castagnino, M. Gadella and O. Lombardi (2005), Time's arrow and irreversibility in time-asymmetric quantum mechanics; *International Studies in the Philosophy of Science*, **19**: 223-243

M. Castagnino, M. Gadella and O. Lombardi (2006), Time-reversal, irreversibility and arrow of time in quantum mechanics; *Foundations of Physics*, **36**: 407-426

M. Castagnino and O. Lombardi (2009), *Synthese* **169** (1):1-25.

P. C. Davies (1994), *Stirring up trouble.* In Physical Origins of Time Asymmetry, eds. J J Halliwell, J Perez-Mercader and W H Zurek (Cambridge: Cambridge University Press)

J. Earman (1974), *Phil. Sci.* **41** 15

M. Frisch (2000), *Brit. J. Phil. Sci.* **51** 381 – 410

Literature

A. Grünbaum (1973), *Philosophical Problems of Space and Time* (2nd edn) (Dordrecht: Reidel)

S. Hawking and J. Ellis (1973), *The Large Scale Structure of Space-Time* (Cambridge: Cambridge University Press)

N. Kanekar, V. Sahni and Y. Shtanov (2001), *Phys. Rev.* D **63** 083520

G. Matthews (1979), *Phil. Sci.* **46** 82

R. Penrose (1979), *Singularities and time asymmetry*. In General Relativity: An Einstein Centenary Survey, eds. S Hawking and W Israel (Cambridge: Cambridge University Press)

H. Price (1996), *Time's Arrow and Archimedes' Point* (Oxford/New York: Oxford University Press)

H. Price (2006), *Stud. Hist. Philos. Mod. Phys.* **37** 498–527

H. Reichenbach (1956), *The Direction of Time* (Berkeley: University of California Press)

F. Rohrlich (2005), *Phys. Rev.* E **72** 057601

R. G. Sachs (1987), *The Physics of Time-Reversal* (Chicago: University of Chicago Press)

L. S. Schulman (1999), *Phys. Rev. Lett.* **83** 5419

In 'Time's Arrow in Cosmology'

D. Z. Albert (2000), *Time and Chance* (Cambridge, MA/London: Harvard University Press).

R. Aquilano, M. Castagnino and E. Eiroa (1999), *Phys. Rev.* D **59** 087301.

C. Barceló and M. Visser (2002), Twilight for the energy conditions? *Preprint* gr-qc/0205066.

J. G. Barlett and A. Blanchard (1996), *Astron. Astrophys.* **307** 1.

E. Bertschinger (1998), *Ann. Rev. Astron. Astrophys.* **36** 599.

M. Castagnino (1989), *Phys. Rev.* D **29** 2216.

M. Castagnino, M. Gadella, F. Gaioli and R. Laura (1999), *Int. J. Theor. Phys.* **38** 2823.

M. Castagnino, H. Giacomini and L. Lara (2000), *Phys. Rev.* D **61** 107302.

M. Castagnino, H. Giacomini and L. Lara (2001), *Phys. Rev.* D **63** 044003.

Literature

M. Castagnino and O. Lombardi (2004), *J. Phys.* A **37** 4445–4463.

M. Castagnino, L. Lara and O. Lombardi (2003a), *Class. Quantum Grav.* **20** 369–391.

M. Castagnino, L. Lara and O. Lombardi (2003b), *Found. Phys.* **22**(6).

M. Castagnino and F. Lombardo (1993), *Phys. Rev.* D **48** 1722.

M. Castagnino and F. D. Mazzitelli (1990), *Phys. Rev.* D **42** 482.

M. Castagnino and O. Lombardi (2009), *Synthese* **169** (1):1-25.

M. M. Ćirković and V. Miloševic-Zdjelar (2004), *Found. Sci.* **9**(1) 1–24.

P. Coles and G. Ellis (1994), *Nature* **370** 609.

J. Earman (1974), *Philos. Sci.* **41** 15–47.

J. Earman (2006), *Stud. Hist. Philos. Mod. Phys.* **37** 399–430.

X. Fan, N. A. Bahcall and R. Cen (1997), *Astrophys. J.* **490** L123.

R. E. Frederick (1991), Introduction to the argument of 1768, *in* J. Van Cleve & R. E. Frederick, (*eds*), '*The Philosophy of Right and Left'*, Kluwer, Dordrecht.

F. Gross (2004), *Relativistic Quantum Mechanics and Field Theory*. New York: Wiley.

S. W. Hawking (1994), *The No Boundary Condition and the Arrow of Time.* In Physical Origins of Time Asymmetry, eds. J J Halliwell, J Perez-Mercader and W H Zurek (Cambridge: Cambridge University Press) pp. 346-357.

S. W. Hawking and G. F. R. Ellis (1973), *The Large Scale Structure of Space-time* (Cambridge: Cambridge University Press).

S. W. Hawking and R. Penrose (1970), *Proc. R. Soc. London,* Ser. A **314** 529.

I. Kant (1768), 'Concerning the ultimate ground of the differentiation of directions in space', in *The Cambridge Edition of the Works of Immanuel Kant: Theoretical Philosophy, 1755–1770*, trans. and *ed.* D. Walford & R. Meerbote Cambridge University Press.

I. Kant (1770), 'Concerning the ultimate ground of the differentiation of directions in space', in *The Cambridge Edition of the Works of Immanuel Kant: Theoretical Philosophy, 1755–1770*, trans. and *ed.* D. Walford & R. Meerbote Cambridge University Press.

C. S. Kochanek (1995), *Astrophys. J.* **453** 545.

L. D. Landau and E. M. Lifshitz (1970), *Statistical Physics* vol 5 (Oxford: Pergamon).

G. W. Leibniz (1686), *Correspondence with Arnauld,* published by Jonathan Bennett (2010).

Literature

G. W. Leibniz (1715–1716), *Leibniz–Clark correspondence,* (reprint 1998), Manchester University Press.

A. Lichnerowicz (1955), *Théories Relativistes de la Gravitation et de l'Electromagnétisme* (Paris: Masson).

A. D. Linde (1990), *Inflation and Quantum Cosmology* (San Diego: Academic Press).

R. Penrose (1979), *Singularities and time asymmetry*. In General Relativity: An Einstein Centenary Survey, eds. S Hawking and W Israel (Cambridge: Cambridge University Press).

O. Pooley (2002), Handedness, parity violation, and the reality of space*, Philosophy of Science* (Preprint).

H. Price (1996), *Time's Arrow and Archimedes' Point* (Oxford/New York: Oxford University Press).

H. Price (2002), *Burbury's Last Case: The Mystery of the Entropic Arrow.* In Time, Reality and Experience, ed. C. Callender (Cambridge: Cambridge University Press).

A. G. Riess et al. (1998), *Astron. J.* **116** 1009 *.

A. Vilenkin (1988), *Phys. Rev.* D **37** 888.

M. Visser (1996), *Lorentzian Wormholes* (Berlin: Springer).

R. M. Wald (2006), *Stud. Hist. Philos. Mod. Phys.* **37** 394–398.

J. A. Wheeler and R. P. Feynman (1945), *Rev. Mod. Phys.* **17** 157–181.

In 'The Arrow of Radiation'

D. Z. Albert (2000), *Time and Chance* (Cambridge, MA/London: Harvard University Press).

M. Castagnino, L. Lara and O. Lombardi (2003), *Class. Quantum Grav.* **20** 369–391.

M. Castagnino and O. Lombardi (2009), *Synthese* **169** (1):1-25.

P. A. M. Dirac (1938), *Proc. R. Soc. London*, Ser. A **167**, 118.

J. Earman (1974), *Philos. Sci.* **41** 15–47.

M. Frisch (2000), *Brit. J. Phil. Sci.* **51** 381 – 410.

M. Frisch (2002), *Brit. J. Phil. Sci.* **53** 1-19.

Literature

M. Frisch (2004), *Philosophy of Science* **71:5** 696-706.

M. Frisch (2005), *Studies in the History and Philosophy of Modern Physics* **36, 4** 659-679.

M. Frisch (2006), *Studies in the History and Philosophy of Modern Physics* **37** 542-558.

M. Frisch (2007), "Causation, Counterfactuals and Entropy," in *Russell's Republic: The Place of Causation in the Constitution of Reality*, eds. Huw Price and Richard Corry, Oxford University Press, 351-395.

M. Frisch (2009), *Brit. J. Phil. Sci.* **60** 459-474.

M. Gell-Mann and J. Hartle (1994), *Time symmetry and asymmetry in quantum mechanics and quantum cosmology.* In Physical Origins of Time Asymmetry, eds. J J Halliwell, J Perez-Mercader and W H Zurek (Cambridge: Cambridge University Press) pp. 311–345.

F. Gross (2004), *Relativistic Quantum Mechanics and Field Theory* (New York: Wiley).

S. W. Hawking and G. F. R. Ellis (1973), *The Large Scale Structure of Space-time* (Cambridge: Cambridge University Press).

J. D. Jackson (1999), *Klassische Elektrodynamik* (Berlin/New York: Walter de Gruyter).

J. M. Jauch and F. Rohrlich (1976), *Theory of Photons and Electrons* (2nd edn) (New York: Springer-Verlag).

L. Landau and E. Lifshitz (1951), *The Classical Theory of Fields* (Cambridge: Addison-Wesley).

K. Popper (1956), *Nature* **177** 538.

H. Price (1991a), *Found. Phys.* **21** 8.

H. Price (1991b), *Brit. J. Phil. Sci.* **42** 111–144.

H. Price (1994), *Brit. J. Phil. Sci.* **45** 1023–1028.

H. Price (1996), *Time's Arrow and Archimedes' Point* (Oxford/New York: Oxford University Press).

H. Price (2006), Stud. Hist. Philos. Mod. Phys. **37** 498–527.

F. Rohrlich (1990), *Classical Charged Particles* (Redwood City: Addison-Wesley).

F. Rohrlich (2002), *Phys. Lett.* A **303**, 307.

F. Rohrlich (2005), *Phys. Rev.* E **72** 057601.

H. Spohn (2000), *Europhys. Lett.* **50** 287.

Literature

C. Teitelboim (1992), *Phys. Rev.* D **3** 297.

M. Visser (1996), *Lorentzian Wormholes* (Berlin: Springer).

M. Visser and C. Barcel´o (2002) *Twilight for the energy conditions? Preprint* gr-qc/0205066.

J. R. Wheeler and R. P. Feynman (1945), *Rev. Mod. Phys.* **17** 157.

H. D. Zeh (1989), *The Physical Basis of the Direction of Time* (Berlin: Springer Verlag).

H. D. Zeh (1999), *The Physical Basis of the Direction of Time* (3rd edn) (Berlin: Springer Verlag).

In 'Time Asymmetries in Quantum Cosmology, Entropy and the Second Law of (Quantum) Thermodynamics'

A. E. Allahverdyan and D. B. Saakian (1998), *Phys. Rev.* E **58** 1148

A. E. Allahverdyan and V. G. Gurzadyan (2000), *The Chaotic Universe.* eds. V G Gurzadyan and R Ruffini (Singapore: World Scientific)

A. E. Allahverdyan and V. G. Gurzadyan (2002), *J. Phys. A: Math. Gen.* **35** 7243–7254

J. Audretsch and P. Spangehl (1987), *Phys. Rev.* D **35** 2365

L. Bergström and A. Goobar (1999), *Cosmology and Particle Astrophysics* (New York: Wiley)

M. Castagnino, D. Harari and C. Núnez (1987), *J. Math. Phys.* **28** 184–192

M. Castagnino and E. Gunzig (1999), *Int. J. Theor. Phys.* **38** 47

M. Castagnino, H. Giacomini and L. Lara (2000), *Phys. Rev.* D **61** 107302

M. Castagnino, H. Giacomini and L. Lara (2001), *Phys. Rev.* **63** 044003

M. Castagnino and C. Laciana (2002), *Class. Quantum Grav.* **19** 2657–2670

S. A. Fulling and L. Parker (1974) *Ann. Phys. (NY)* **87** 176–204

P. Glansdorf and I. Prigogine, (1971), *Thermodynamic Theory of Structures, Stability and Fluctuations* (New York: Wiley)

T. Gold (1962), *Am. J. Phys.* **30** 403

E. Gunzig (1989), *Cosmological entropy generation*. In Proc. 3rd Int. Symp. Foundations of Quantum Mechanics *(Tokyo,* 1989*)* p. 349

J. J. Halliwell, J. Perez-Mercader and W. H. Zurek (1996) (eds), *Physical Origins of Time Asymmetry* (Cambridge: Cambridge University Press)

S. W. Hawking (1985), *Phys. Rev.* D **32** 2989

C. Kiefer and H. D. Zeh (1995), *Phys. Rev.* D **51** 4145

L. D. Landau and E. M. Lifshitz (1970), *Statistical Physics* vol 5 (Oxford: Pergamon)

C. Laciana (1998), *Ann. Phys. (NY)* **264** 95

G. Lindblad (1975), *Commun. Math. Phys.* **40** 147

D. Page (1985), *Phys. Rev.* D **32** 2496

L. Parker (1969), *Phys. Rev.* **183** 1057

R. Penrose (1979), *Singularities and time asymmetry*. In General Relativity: An Einstein Centenary Survey, eds. S Hawking and W Israel (Cambridge: Cambridge University Press)

R. Penrose (1989), *The Emperor's New Mind* (Oxford: Oxford University Press)

R. Penrose (1994), *J. Stat. Phys.* **77** 217

T. Petrosky and I. Prigogine (1997), *Adv. Chem. Phys.* **XCIX** 1

H. Price (1996), *Time's Arrow and Archimedes' Point* (Oxford/New York: Oxford University Press)

F. Schlögl (1980), *Phys. Rep.* **62** 268

H. D. Zeh (1992), *The Physical Basis of the Direction of Time* (Berlin: Springer Verlag)

In 'Time Arrows in Ordinary Quantum Mechanics'

F. Adams and G. Laughlin (1997), *Rev. Mod. Phys.* **69** 337

A. I. Ajiezer (1981), *Métodos de la Física Estadística* (Moscow: Mir)

L. Bergström and A. Goobar (1999), *Cosmology and Particle Astrophysics* (New York: Wiley)

R. C. Bishop (2004), Arrow of Time in Rigged Hilbert Space Quantum Mechanics, *Int. J. Theor. Phys.* **43**, 7/8

A. Bohm, M. Gadella and S. Wickramasekara, (1999). Some little things about rigged Hilbert spaces and quantum mechanics and all that. In *Generalized Functions, Operator Theory, and*

Dynamical Systems, ed. I. Antoniou and G. Lumer, 202–250 Chapman & Hall/CRC, Boca Raton.

M. Castagnino (1998), *Phys. Rev.* D **57** 750

M. Castagnino and E. Gunzig (1999), *Int. J. Theor. Phys.* **38** 47

M. Castagnino and R. Laura (2000), *Phys. Rev.* A **62** 022107

M. Castagnino, L. Lara and O. Lombardi (2003), *Class. Quantum Grav.* **20** 369–391

M. Castagnino and O. Lombardi (2009), *Synthese* **169** (1):1-25

P. Ehrenfest and T. Ehrenfest (1959), *The Conceptual Foundations of the Statistical Approach in Mechanics* (Ithaca: Cornell University Press) (original 1912)

P. D. Lax and R. S. Phillips (1979), *Scattering Theory* (New York: Academic Press)

M. C. Mackey (1989), *Rev. Mod. Phys.* **61** 981

J. V. Narlikar (1995), *From Black Clouds to Black Holes* (Singapore: World Scientific)

L. S. Schulman (1997), *Time's Arrow and Quantum Measurements* (Cambridge: Cambridge University Press)

A. Vilenkin (1988), *Phys. Rev.* D **37** 888

H. D. Zeh (1989), *The Physical Basis of the Direction of Time* (Berlin: Springer Verlag)

In 'Summary of the Main Conclusions'

A. E. Allahverdyan and V. G. Gurzadyan (2002), *J. Phys. A: Math. Gen.* **35** 7243–7254

C. Barcel´o and M. Visser (2002) *Twilight for the energy conditions? Preprint* gr-qc*/*0205066

M. Castagnino and C. Laciana (2002), *Class. Quantum Grav.* **19** 2657–2670

M. Castagnino, L. Lara and O. Lombardi (2003), *Class. Quantum Grav.* **20** 369–391

M. Castagnino and O. Lombardi (2009), *Synthese* **169** (1):1-25

M. M. Ćirković and V. Miloševic-Zdjelar (2004), *Found. Sci.* **9**(1) 1–24

M. Frisch (2000), *Brit. J. Phil. Sci.* **51** 381–410

Literature

P. Glansdorff and I. Prigogine (1971), *Thermodynamic Theory of Structures, Stability and Fluctuations* (New York: Wiley)

L. D. Landau and E. M. Lifshitz (1970), *Statistical Physics* vol 5 (Oxford: Pergamon)

H. Price (1991a), *Found. Phys.* **21**(8) 959–975

H. Price (1991b), *Brit. J. Phil. Sci.* **42** 111–144

H. Price (1994), *Brit. J. Phil. Sci.* **45** 1023–1028

H. Price (1996), *Time's Arrow and Archimedes' Point* (Oxford/New York: Oxford University Press)

H. Price (2002), *Burbury's Last Case: The Mystery of the Entropic Arrow.* In Time, Reality and Experience, ed. C. Callender (Cambridge: Cambridge University Press)

H. Price (2006), *Stud. Hist. Philos. Mod. Phys.* **37** 498–527

F. Rohrlich (2005), *Phys. Rev.* E **72** 057601

M. Visser (1996), *Lorentzian Wormholes* (Berlin: Springer)

J. R. Wheeler and R. P. Feynman (1945), *Rev. Mod. Phys.* **17** 157.

H. D. Zeh (1989), *The Physical Basis of the Direction of Time* (Berlin: Springer Verlag)

H. D. Zeh (1999), *The Physical Basis of the Direction of Time* (3rd edn) (Berlin: Springer Verlag)