Preface to the Third and Preliminary Fourth Edition

– Prelim. 4th edtn. (Feb 01): www.time-direction.de

The third (1999) edition of the Direction of Time offered far more revisions and additions than the second one in 1992. During the seven years in between, several fields of research related to the arrow of time had shown remarkable progress. For example, decoherence proved to be the most ubiquitous manifestation of the quantum arrow, while articles on various interpretations of quantum theory (many of them with inbuilt time-asymmetric dynamical aspects) can and do now regularly appear in reputed physics journals. Therefore, most parts of Chap. 4 were completely rewritten and some new sections added, while the second part of Chap. 3 was affected by these changes in order to prepare for the discussion of measurements and *dynamical maps* within the framework of classical ensemble theory.

However, *all* parts of the book have been revised, and some of them completely rewritten, whilst essentially maintaining the book's overall structure. Some of its new aspects may be listed here:

The Introduction now attempts to distinguish rigorously between those time *asymmetries* which still preserve dynamical determinism, and the various 'irreversiblities' (arrows of time proper) which are the subject of this book.

In Chap. 2, the concept of forks of causality is contrasted to that of forks of indeterminism (to be used in Chaps. 3 and 4), while the treatment of the radiation reaction of a moving charge (Sect. 2.3) had to be updated.

Sects. 3.2–3.4 have been given a new structure, while a discussion of semigroups and their physical meaning has been added to Sect. 3.4.

In Chap. 4, only Sects. 4.1 and 4.5 (the former Sect. 4.3 on exponential decay) are not entirely new. In particular, there is now an extended separate Sect. 4.3 on decoherence. Sects. 4.4 (on quantum dynamical maps) and 4.6 (on the time arrow in various *interpretations* of quantum theory) is now added.

In Chap. 5, the thermodynamics of acceleration is now presented separately (Sect. 5.2), while Sect. 5.3 on the expansion of the universe contains a discussion of the consistency of cosmic two-time boundary conditions. The dynamical interpretation of general relativity with its concept of *intrinsic time* is discussed in Sect. 5.4.

Chap. 6 now covers all aspects of quantum cosmology and thus includes, as Sect. 6.1, the material of the former Sect. 5.2.2 on phase transitions of the vacuum with their consequences on entropy capacity. In Sect. 6.2 on quantum gravity, emphasis is on *timelessness*, which is enforced by quantization of a reparametrization invariant theory. There is a new Sect. 6.2.2 on the emergence of classical time along the lines of the Tomonaga-Schwinger equation, while Sect. 6.2.3 describes some speculations on the impact of quantum cosmology on the concept of black holes and their thermodynamical properties. A numerical toy model has been appended after the Epilog in order to illustrate some typical arguments of stastistical mechanics.

I also hope that most disadvantages which had resulted from the fact that I previously had (very unfortunately) translated many parts of the first edition from the German lecture notes that preceded it (Zeh 1984), have now been overcome. Two new books on the arrow of time (Price 1996 and Schulman 1997) have recently appeared. They are both well written, and they discuss many important aspects of 'irreversible' physics in a consistent and illuminating manner — often nicely complementing each other as well as this book. However, I differ from their views in two respects: I regard gravity (not least its quantized form) as basic for the arrow of time, as I try to explain in Chaps. 5 and 6, and I do not think that the problem of quantum measurements can be solved by means of an appropriate final condition in a satisfactory way (see Footnote 4 of Chap. 4).

I wish to thank Julian Barbour, Erich Joos, Claus Kiefer, Joachim Kupsch, York Ramachers, Huw Price, Fritz Rohrlich, Paul Sheldon and Max Tegmark for their comments on early versions of various parts of the manuscript.

Heidelberg, April 1999 *H.-D. Zeh*

The fourth edition contains revisions throughout the whole book. There are many new formulations and arguments, several new comments and references, and three minor error corrections (on page 22, 112 and 146 of the third edition). It is now planned to be published in Spring 2001. Therefore, this preliminary internet version will not be further updated.

For this edition I am grateful to David Atkinson (for a very helpful discussion of radiation damping — Sect. 2.3), to Larry Schulman (for comments on the problem of simultaneous arrows of time — Sect. 3.1.2), and to Paul Sheldon (for a discussion of the compatibility of closed time-like curves with quantum theory — Chap. 1). The most efficient help, this time, came from John Free, who carefully edited the whole book (not only for matters of English language).

Heidelberg, February 2001 *H.-D. Zeh*

Contents

Introduction

Prelim. 4th edtn. (Nov 00): www.time-direction.de

The asymmetry of nature under a 'reversal of time' (that is, a reversal of motion and change) appears only too obvious, as it deeply affects our own form of existence. If physics is to justify the hypothesis that its laws control everything that happens in nature, it should be able to explain (or consistently describe) this fundamental asymmetry which defines what may be called a *direction in time* or even — as will have to be discussed — a direction *of* time. Surprisingly, the very laws of nature are in pronounced contrast to this fundamental asymmetry: they are essentially symmetric under time reversal. It is this discrepancy that defines the enigma of the direction of time, while there is no lack of asymmetric formalisms or pictures that go beyond the empirical dynamical laws.

It has indeed proven appropriate to divide the formal dynamical description of nature into *laws* and *initial conditions*. Wigner (1972), in his Nobel Prize lecture, called it Newton's greatest discovery, since it demonstrates that the laws by themselves are far from determining nature. The formulation of these two pieces of the dynamical description requires that appropriate *kinematical* concepts (formal *states* or *configurations* z, say), which allow the unique mapping (or 'representation') of all possible states of physical systems, have already been defined on empirical grounds.

For example, consider the mechanics of N mass points. Each state z is then equivalent to N points in three-dimensional space, which may be represented in turn by their $3N$ coordinates with respect to a certain frame of reference. States of physical *fields* are instead described by certain *functions* on three-dimensional space. If the laws of nature, in particular in their relativistic form, contain kinematical elements (that is, constraints for kinematical concepts that would otherwise be too general), such as $\mathrm{div}\boldsymbol{B} = 0$ in electrodynamics, one should distinguish them from the dynamical laws proper. This is only in *formal* contrast to relativistic spacetime symmetry (see Sect. 5.4).

The laws of nature, thus refined to their purely dynamical sense, describe the time dependence of physical states, $z(t)$, in a general form — usually by means of differential equations. They are called *deterministic* if they uniquely determine the state at time t from that (and possibly its time derivative) at any earlier or later time, that is, from an appropriate initial or final condition. This symmetric causal structure of dynamical determinism is stronger than the traditional concept of *causality*, which requires that every event in nature must possess a specific *cause* (in its past), while not necessarily an *effect* (in its future). The *Principle of Sufficient Reason* can be understood in this asymmetric causal sense that would depend on an absolute direction of time.

However, only since Newton do we interpret uniform motion as 'eventless', while acceleration requires a *force* as the modern form of *causa movens* (sometimes assumed to act in a retarded, but hardly ever in an advanced manner). From the ancient point of view, terrestrial bodies were regarded as eventless or 'natural' when at rest, celestial ones when moving in circular orbits (including epicycles), or when at rest on the celestial ('crystal') spheres. These motions thus did not require any dynamical causes according to this picture, similar to uniform motion today. None of the traditional causes (neither physical nor other ones) ever questioned the fundamental asymmetry in (or of) time, as there were no conflicting symmetric dynamical laws yet.

Newton's concept of a force as determining acceleration (the *second* time derivative of the 'state') forms the basis of the formal Hamiltonian concept of states in *phase space* (with corresponding dynamical equations of *first* order in time). First order time derivatives of states in configuration space, required to define momenta, can be freely chosen as part of the initial conditions. In its Hamiltonian form, this part of the kinematics may appear as dynamics, since the definition of canonical momentum depends in general on a dynamical concept (the Lagrangean).

Physicists after Newton could easily recognize friction as a possible source of the apparent asymmetry of conventional causality. While different motions starting from the same unstable position of rest require different initial perturbations, friction (if understood as a fundamental force) could deterministically bring different motions to the same rest. States at which the symmetry of determinism may thus come to an end (perhaps asymptotically) are called *attractors* in some theories.

The term 'causality' is unfortunately used with quite different meanings. In physics it is often synonymous with determinism, or it refers to the relativistic speed limit for the propagation of causal influences (hence of information). In philosophy it may refer to the existence of laws of nature in general. In (phenomenological) mathematical physics, dynamical determinism is often understood to apply in the 'forward' direction of time only (thus allowing attractors — see Sect. 3.4). Time reversal-symmetric determinism was discovered only with the laws of mechanics, when friction could either be neglected, or was recognized as being based on thermodynamics. An asymmetric concept of 'intuitive causality' that is compatible with (though different from) symmetric determinism will be defined and discussed in the introduction to Chap. 2.

The time reversal symmetry of *determinism as a concept* does not require symmetric dynamical laws. For example, the Lorentz force $e\boldsymbol{v}\times\boldsymbol{B}$, acting on a charged particle, and resulting from a *given* external magnetic field, changes sign under time reversal (defined by a replacement of t with $-t$, hence as a reversal of *motion*[1]), as it is proportional to the velocity $\boldsymbol{v}$. Nonetheless, determinism applies in both directions of time.

This time reversal asymmetry of the equation of motion would be cancelled by a simultaneous space reflection, which would reverse the magnetic field. Similar 'compensated asymmetries' may be found in many other situations, with more or less physical symmetry operations (see Sachs 1987). As an example, the formal asymmetry of the Schrödinger equation under time reversal is cancelled by complex conjugation of the wave function on *configuration* space. This can be described by an anti-unitary operation T that leaves the configuration basis unchanged, $Tc|q\rangle = c^*|q\rangle$, for complex numbers c. For technical reasons, T may be chosen to contain other self-inverse operations, such as multiplication with the Dirac matrix β. As a further trivial application, consider the time reversal of *states* in classical phase space, $\{q,p\} \to \{q,-p\}$. This transformation restores symmetry under a *formal* time reversal $p(t), q(t) \to p(-t), q(-t)$. In quantum theory it corresponds to the transformation $T|p\rangle = |-p\rangle$ that *results* from the complex conjugation of the wave function e^{ipq} which defines the state $|p\rangle = (2\pi)^{-1/2}\int dq e^{ipq}|q\rangle$.

For trajectories $z(t)$, one usually includes the transformation $t \to -t$ into the action of T rather than applying it only to the state z: $Tz(t) := z_T(-t)$, where $z_T := Tz$ is the 'time-reversed state'. In the Schrödinger picture of quantum theory this is automatically taken care of by the anti-unitarity of T when commuted with the time translation e^{iHt} by means of a time reversal invariant Hamiltonian H. In this sense, 'T invariance' means time reversal invariance. When discussing time reversal, one usually assumes invariance under *translations* in time, in order not to specify an arbitrary origin for the time reversal transformation $t \to -t$.

The time reversal asymmetry characterizing weak forces responsible for K-meson decay is balanced by an asymmetry under CP transformation, where C and P are charge conjugation and spatial reflection, respectively. The latter do *not* reflect a time reversal elsewhere (such as the reversal of a magnetic field that is caused by external currents). Only if the compensating symmetry transformation represents an observable, such as CP, and is not the consequence of a time reversal elsewhere, does one speak of a *violation* of time reversal invariance.

The possibility of compensating for a dynamical time reversal asymmetry by another asymmetry (observable or not) reflects the prevailing *symmetry of determinism.* This is in fundamental contrast to genuine 'irreversibilities', which form the subject of this book. No time reversal asymmetry of deterministic laws would be able to explain such irreversibilities.

---

[1] Any distinction between reversal of time and reversal of motion (or change, in general) is meaningful only with respect to some concept of absolute or external time (see Chap. 1). An asymmetry of the fundamental dynamical laws would *define* (or presume) an absolute direction of time — just as Newton's equations define absolute time up to linear transformations (including a reversal of its sign, which is thus *not* absolutely defined in the absence of asymmetric fundamental forces).

All known *fundamental* laws of nature are symmetric under time reversal after compensation by an appropriate symmetry transformation, $\hat{T}$, say, since these laws are deterministic. For example, $\hat{T} = CPT$ in particle physics, while $\hat{T}\{\boldsymbol{E}(\boldsymbol{r}), \boldsymbol{B}(\boldsymbol{r})\} = \{\boldsymbol{E}(\boldsymbol{r}), -\boldsymbol{B}(\boldsymbol{r})\}$ in classical electrodynamics. This means that for any trajectory $z(t)$ that is a solution of the dynamical laws there is a time-reversed solution $z_{\hat{\mathrm{T}}}(-t)$, where $z_{\hat{\mathrm{T}}}$ is the 'time-reversed state' of z, obtained by applying the compensating symmetry transformation.

'Initial' conditions (contrasted to the dynamical laws) are understood as conditions which fix the integration constants, that is, which select particular solutions of the equations of motion. They could just as well be regarded as final conditions, even though this would not reflect the usual *operational* (hence asymmetric) application of the theory. These initial conditions are to select the solutions which are 'actually' found in nature. In modern versions of quantum field theory, even the boundary between laws of nature and initial conditions blurs. Certain parameters which are usually regarded as part of the laws (such as those characterizing the mentioned CP violation) may have arisen by *spontaneous symmetry-breaking* (an indeterministic irreversible process of disputed nature in quantum theory — see Sects. 4.6 and 6.1).

An individual (contingent) trajectory $z(t)$ is generically *not* symmetric under time reversal, that is, not *identical* with $z_{\hat{\mathrm{T}}}(-t)$. If $z(t)$ is sufficiently complex, the time-reversed process is not even likely to occur anywhere else in nature within reasonable approximation. However, most phenomena observed in nature violate time reversal symmetry in a less trivial way if considered as whole *classes of phenomena*. The members of some class may be found abundant, while the time-reversed class is not present at all. Such symmetry violations will be referred to as 'fact-like' — in contrast to the mentioned CP symmetry violations, which are called 'law-like'. In contrast to what is often claimed in textbooks, this asymmetric appearance of nature cannot be explained by statistical arguments. If the laws are invariant under time reversal when compensated by another symmetry transformation, there must be precisely as many solutions in the time-reversed class as in the original one (see Chap. 3).

General classes of phenomena which characterize a direction in time have since Eddington been called *arrows of time*. The most important ones are:

1. Radiation: In most situations, fields interacting with local sources are appropriately described by retarded (outgoing or defocusing) solutions. For example, a spherical wave is observed *after* a point-like source event, propagating away from it. This leads to a damping of the source motion (see Item 5). For example, one may easily observe 'spontaneous' emission (in the absence of incoming radiation), while absorption without any outgoing radiation is hardly ever found. Even an ideal absorber leads to *retarded* consequences in the corresponding field (shadows) — see Chap. 2.

2. Thermodynamics: The Second Law $dS/dt \geq 0$ is often regarded as a law of nature. In microscopic description it has instead to be interpreted as *fact-like* (see Chap. 3). This arrow of time is certainly the most important one. Because of its applicability to human memory and other physiological processes it may be responsible for the *impression* that time itself has a direction (related to the *apparent flow of time* — see Chap. 1).

3. Evolution: Dynamical 'self-organization' of matter, as observed in biological and social evolution, for example, may appear to contradict the Second Law. However, it is in agreement with it if the entropy of the environment is properly taken into account (Sect. 3.4).

4. Quantum Mechanical Measurement: The probability interpretation of quantum mechanics is usually understood as a fundamental indeterminism of the future. Its interpretation and compatibility with the deterministic Schrödinger equation constitutes a long-standing open problem of modern physics. Quantum 'events' are often dynamically described by a collapse of the wave function, in particular during the process of measurement. In the absence of a collapse, quantum mechanical interaction leads to growing entanglement (quantum nonlocality) — see Chap. 4.

5. Exponential Decay: Unstable states (in particular quantum mechanical 'particle resonances') usually fade away exponentially with increasing time (see Sect. 4.5), while exponential growth is only observed in self-organizing situations (cf. Item 3 above).

6. Gravity seems to 'force' all matter to contract with increasing time according its attractivity. However, this is another prejudice about the *causal* action of forces. Gravity leads to the acceleration of contraction (or deceleration of expansion) in *both* directions of time, since acceleration is a second time derivative. The observed contraction of complex gravitating systems (such as stars) against their internal pressure is in fact controlled by thermodynamical and radiation phenomena. Such gravitating objects are characterized by a negative heat capacity, and classically even by the ability to contract without limit in accordance with the Second Law (see Chap. 5). In general relativity this leads to the occurrence of asymmetric *future horizons* through which objects can only *dis*appear. The discussion of quantum fields in the presence of such *black holes* during recent decades has led to the further conclusion that horizons must possess fundamental thermodynamical properties (temperature and entropy). This is remarkable, since horizons characterize spacetime, hence time itself. On the other hand, *expansion* against gravity is realized by the universe as a whole. Since it represents a unique process, this cosmic expansion does not define a *class* of phenomena. For this and other reasons it is often conjectured to be the 'master arrow' from which all other arrows may be derived (see Sects. 5.3 and 6.2.1).

In spite of their fact-like nature, these arrows of time, in particular the thermodynamical one, have been regarded by some of the most eminent physicists as even more fundamental than the dynamical laws. For example, Eddington (1928) wrote:

"The law that entropy always increases holds, I think, the supreme position among the laws of nature. If someone points out to you that your pet theory of the universe is in disagreement with Maxwell's equations – then so much the worse for Maxwell's equations. ... but if your theory is found to be against the second law of thermodynamics, I can give you no hope; there is nothing for it but to collapse in deepest humiliation."

And Einstein (1949) remarked:

"It" (thermodynamics) "is the only physical theory of universal content concerning which I am convinced that, within the framework of the applicability of its basic concepts, it will never be overthrown."

Regardless of whether these remarks will always remain valid, their meaning should be understood correctly. They were hardly meant to express doubts in the *derivability* of this 'fundamental' thermodynamical arrow of time from *presumed* dynamical laws in the manner of Boltzmann (see Chap. 3). Rather, they seem to express their authors' conviction in the invariance of the derived results under modifications and generalizations of these laws. However, the derivation will be shown to require important assumptions about the initial state of the universe. *If* the Second Law is fact-like in this sense, its violation or reversal must at least be compatible with the dynamical laws.

The arrows of time listed above describe an asymmetry in the *history* of the physical world under a formal reversal of time. This history can be considered as a whole, like a complete movie film sitting on the desk, or an ordered stack of static picture frames ('states'), without any selection of a present (one specific 'actual' frame) or an external distinction between beginning and end. This is sometimes called the 'block universe view' (cf. Price 1996), and contrasted to that of an *evolving* universe (based on the concept of a 'flow of time', picture by picture, as seen by an *external* movie viewer as a definer of 'absolute' time for the movie).

It appears doubtful that these different *view points* should have different power of explaining an asymmetry in the *content* of the movie, even though they are regarded as basically different by many philosophers, and also by some physicists (Prigogine 1980, von Weizsäcker 1982) — see also Chap. 1. The second point of view is related to the popular position that the past be 'fixed', while the future is 'open' and does 'not yet exist'. The asymmetry of history is then regarded as the 'outcome' (or the consequence) of this time-directed 'process of coming-into-being'. (The abundance of quotation marks indicates how our language is loaded with prejudice about the flow of time.) The fact that there are documents, such as fossils, only about the past, and that we cannot remember the future, [2] appears as evidence for this 'structure of time' (as it is called), which is also referred to as the 'historical nature' (*Geschichtlichkeit*) of the world.

[2] "It's a bad memory that only works backwards" says the White Queen to Alice.

However, the asymmetry of the movie film can be discussed regardless of its presentation or production (if any). The *relation* between the memory that is contained in an individual frame and the history itself is no more than a property of this movie that is sitting on the desk. If such asymmetric relations exist(!) throughout the whole movie, an internal observer could *know* its content only at the (intrinsically defined) 'end'. Nonetheless he could conceive of a 'potential' whole movie even in ('during') the asymmetric movie, in particular if he discovers dynamical laws. He could neither argue against the *existence* of the remaining part of the movie, nor prove it, as *his* time is part of the content of the movie. It has to be read from clocks showing up on the picture frames — but not from the watch of an external movie viewer or any frame numbers (see also Chap. 1 and Sect. 5.4). The concept of existence is here evidently used with various meanings, while the debate essentially becomes one about words. Within our world movie, concepts like *fixed* and *open* (or *actual* versus *potential*) can be meaningful only either as statements about *practical abilities* of predicting and retrodicting, or about dynamical *models* used to describe it.

The objection that the historical nature of the world be a prerequisite (in the Kantian sense) for the fact that it can be experienced at all does not exclude the possibility (or necessity) of describing it in terms of those laws and concepts that have been extracted from this experience. They may then be hypothetically extrapolated to form a 'world model', whereby the historical nature may turn out *not* to apply to other spacetime regions (see Sect. 5.3).

In classical physics, the Second Law (or its microscopic roots) is usually regarded as the physical basis of the historical nature of the world. Its statistical interpretation would then mean that this 'structure of time' (that is, its apparent direction) is merely the consequence of contingent facts which characterize our specific world. For example, one has to explain thermodynamically why there are observations, but no 'un-observations' in which initially present information (memory about the future) would disappear by means of a controlled interaction between the observing and observed systems. Such conceivable processes have to be distinguished from the familiar ones ('forgetting'), which represent a loss of information characterized by an increase of entropy in the memory device (see Sect. 3.3.)

The concept of information would thus arise as a consequence of thermodynamics (and not the other way, as is sometimes claimed). The inconsistency of presuming extra-physical concepts of information or 'operation' has often been discussed by means of *Maxwell's demon*. In particular, the 'free will' of the experimenter should not be misused to explain the specific (low entropy) initial conditions that he prepares in his laboratory. If the experimenter were not required to obey the thermodynamical laws himself (as exemplified by the demon), his actions could readily *create* the thermodynamical arrow of time observed in his experiments. Nonetheless, there has always been speculation that the intervention by conscious beings must require new fundamental laws — a possibility that cannot, of course, be absolutely excluded.

The indeterminism observed in *quantum* phenomena has often been interpreted as evidence for such a fundamental asymmetry characterizing observations (by humans?). This may be documented by many statements, such as Heisenberg's famous remark in the spirit of idealism that "a particle trajectory is created by our act of observing it," [3] or von Weizsäcker's "what has been observed exists with certainty." [4] One can similarly understand Bohr's statement: "Only an observed quantum phenomenon is a phenomenon." Bohr insisted that a quantum measurement cannot be analyzed as a dynamical physical process ("there is no quantum world"). A similar view can be found in Pauli's letter to Born (Einstein and Born 1969): "The appearance of a definite position x_0 during an observation ... is then regarded as a creation existing outside the laws of nature." [5] Born often expressed his satisfaction with quantum mechanics, as he felt that his probability interpretation saved free will from the determinism of classical laws.

The extra-physical time arrow is clearly used in all *operational* formulations of quantum theory (such as probabilistic relations connecting preparations and measurements — thus restricted to laboratory physics). A careful distinction between temporal and logical aspects of actual and 'counterfactual' measurements can be found in Mermin (1998).

Most of these formulations rely on a given (absolute) direction of time. This should then be reflected by the dynamical description of quantum measurements and 'measurement-like processes' — even in the block universe picture. The impact of such concepts (provided they are justified) on the formal physical description should therefore be precisely located.

Much of the *philosophical* debate about time is concerned with linguistic problems, some of them simply arising from the pre-occupied usage of the tenses, particularly for the verb 'to be' (see Smart 1967, or Price 1996). However, Aristotle's famous (pseudo-)problem regarding the meaning of today's statement that there *be* a sea battle tomorrow does not only survive in *Sein und Zeit*, but even in quantum theory in the form of an occasional confusion of logic with dynamics ('logic of time' — see Butterfield and Isham 1999).

The prime intention of this book is to discuss the relations between different arrows of time, and to search for a common *master arrow*. Toward this end, open problems will have to be pointed out. In the traditional fields they have often been pragmatically put aside, while they may become essential in conjunction with more recent theoretical developments, which seem to have fundamental cosmological implications (see Chaps. 5 and 6).

---

[3] "*Die Bahn entsteht erst dadurch, daß wir sie beobachten.*"

[4] "*Was beobachtet worden ist, existiert gewiß.*"

[5] "*Das Erscheinen eines bestimmten Ortes x_0 bei der Beobachtung ... wird dann als außerhalb der Naturgesetze stehende Schöpfung aufgefaßt.*"

1. The Physical Concept of Time

– Prelim. 4th edtn. (Nov 00): www.time-direction.de

The concept of time has been discussed since the earliest records of philosophy, when science had not yet become a separate subject. It is rooted in the subjective experience of the 'passing' present or moment of awareness, which appears to 'flow' through time and thereby dynamically to separate the past from the future. This has led to the formal representation of time by the real numbers, and of the present by a point that 'moves' in the direction characterized by their sign.

The *mechanistic concept of time* avoids any subjective foundation, as it is defined by the motion of objects (in particular celestial bodies). It is often attributed to Aristotle, although he seems to have regarded this presumably more ancient concept as insufficient. [1] A concept of time *defined* (not merely measured) by motion may appear to be a circular conception, since motion is defined as change with (that is, dependence on) time, thus rendering the metaphor of the *flow of time* a tautology (see e. g. Williams 1951). However, this concept remains a convenient tool for *comparing* simultaneous motions, provided an appropriate concept of simultaneity of different events (such as

[1] "Time is neither identical with movement nor capable of being separated from it" (Physics, Book IV). This may sound like an argument for some absoluteness of time. However, the traditional philosophical debate about time is usually linked to (and often confused with) the psychological and epistemological problem of the *awareness* of time 'in the soul' (and hence related to the problem of consciousness). This is understandable for ancient philosophers, who could neither have anticipated the role of physico-chemical processes (i. e., molecular *motions*) in the brain as 'controlling the mind', nor were in possession of reasonable clocks to give time a precise operational meaning for fast phenomena. According to Flasch (1993), Albertus Magnus (ca. 1200-1280) was the first philosopher who supported a rigorously 'physical' concept of time, since he insisted that time exists in nature, while the soul merely perceives it: "Ergo esse temporis non dependet ab anima, sed temporis perceptio."

Another confusing issue of time in early philosophy, represented by some of Zeno's paradoxes, was the mathematical problem of the real numbers, required to characterize the continuum. Before the discovery of calculus, mathematical concepts ('instruments of the soul') were often thought to be restricted to the natural (or rational) numbers, while reality would correspond to the conceptually inaccessible continuum. Therefore, periodic motion was essential for 'counting' time in order to grasp it, not only to provide a measure. *Uniform* circular motion then appears natural.

Since Newton, and even more so since Einstein, the concept of 'time in nature' has almost exclusively been elaborated by physicists. The adjective 'physical' in the title of this chapter is thus not meant as a restriction.

points on different orbits in space) is available. In pre-relativistic physics, the simultaneity of events is operationally defined by their simultaneous observation (which presumes a *local* concept of simultaneity), while it is formally characterized by the separate concepts of space and time. The possibility of comparing different motions provides a sufficient conception for all *meaningful* temporal statements about nature, while clocks are based on motion (or change), too. All 'properties of time' must then be abstractions from relative motion and its empirical laws.

Physicists contributing to the concept of time have in general been careful to avoid any hidden regress to the powerful prejudice of *absolute time*. Their modern conception of time may indeed be described as the *complete elimination* of absolute time, and hence also of absolute motion. This elimination may be likened to the derivation of 'timeless orbits' from time-dependent ones, such as the function $r(\phi)$, obtained from the time-dependent coordinates $r(t)$ and $\phi(t)$ for motion in a plane by eliminating t. In this way, all motions (in their general sense of all changes) $q_i(t)$ are in principle replaced with 'timeless' trajectories $q_i(q_0)$ in a global configuration space, where q_0 may be chosen as the position of the hand of an appropriate 'clock'.

These time-less trajectories may yet be described by means of a *physically meaningless* parameter λ in the form $q_i(\lambda)$, where equal values of λ characterize the simultaneity of different q_i's. This parametric form was explicitly used in Jacobi's formulation of mechanics (Sect. 5.4), since astronomers without precise clocks (such as atomic ones) had to define time operationally as global *ephemeris time* in terms of the celestial orbits which result from their calculations. Since Jacobi's principle was applied to Newton's theory, its success nonetheless confirmed Newton's hypothetical absolute time, defined as a parameter that simplifies the equations of motion (Poincaré 1902). It is a non-trivial empirical aspect of nonrelativistic physics that such a preferred time parameter is defined uniquely up to linear transformations. This property of Newton's absolute time then allows one to compare *different* time intervals.

Even the topology (ordering) of time may be regarded as the consequence of this *choice of an appropriate time parameter* that, in particular, allows motion to appear continuous. According to this radical view, the 'timeless history' of the whole universe is equivalent to an unordered 'heap of states' (or a stack of shuffled movie frames) that *can* be uniquely ordered and given a measure of distance only by the relations between their *intrinsic* structures (Barbour 1986, 1994a). This view will lead to entirely novel aspects in quantum gravity (see Sect. 6.2). If certain states from the stack (called 'time capsules' by Barbour) contain intrinsically consistent correlations representing 'memories', they may give rise to the impression of a flow of time to observers intrinsic to the system (such as the universe), since these observers would 'remember' properties of those global states which they interpret as forming 'their present past'.

The concept of absolute motion thus shares the fate of the flow of time. 'Time reversal' can only be meaningfully defined as *relative* reversal of motion (for example, relative to such physiological processes which physically control the subjective awareness of time and memory). He who regards this mechanistic concept of time as insufficient should be able to explain what a reversal of *all* motion could *mean*. Ancient versions of a concept of time based on motion may instead have been understood as a 'causal control' of all motion on earth by the 'external' motion of (and on) the celestial sphere — an idea of which astrology is still a relic.

According to *Mach's principle* (see Barbour and Pfister 1995), the concept of absolute time is not only *kinematically* redundant — it should not even play any dynamical role as a preferred (heuristic) parameter, such as it does in Newton's theory.[2] Similarly 'relativistic' ideas (still in the sense of an absolute global concept of simultaneity) were already formulated by Leibniz, Huygens, and later by Berkeley. They may even have prevented Leibniz from co-discovering Newton's mechanics. Leibniz's conception led him to a definition of time by the motion of the universe as a whole. An exactly periodic universe would then describe the recurrence of the *same* time. This concept is far more rigorous than its ancient predecessor in not ascribing a special role to the motion of the celestial bodies (for example, as *causing* motion on earth). Even Newton found the motivation for his functional concept of absolute physical time in the empirical laws of motion which he was discovering, and which were later modified by Einstein, among other things on the basis of Mach's principle.

The concept of mechanistic time (based on motion according to the laws of nature) neither specifies a direction nor any point that might represent the present. By taking into account thermodynamic phenomena (including friction, which, in contrast to Newton's understanding, is not a *fundamental* force), one may complement it with a phenomenological direction, thus arriving at the concept of a *thermodynamico-mechanistic time*. The empirical basis of this concept is the observation that the thermodynamical arrow of time always and everywhere points in the same direction. Explaining this fact (or possibly its range of validity) must be part of the *physics of time asymmetry*. As will be explained, this problem is a physical one, whereas physics does not even offer any conceptual means for deriving the concept of a *present* that would distinguish the past from the future (see also the Epilog).

The concept of a present seems to have as little to do with the concept of time itself as color has to do with light (or with the nature of objects reflecting it). Both the present and color are merely aspects of how we perceive

---

[2] Mach himself was not very clear about whether he intended to postulate what is now often called his principle, or rather to prove such a principle meaningless (see Norton 1995). A related confusion between the trivial invariance of a theory under a mere rewriting of the laws in terms of new coordinates ('Kretzschmann invariance') and the nontrivial *invariance of the laws* under coordinate transformations has even survived in general relativity (Norton 1989).

time and light, respectively. Indeed, just as most information that is contained in the frequency spectrum of observed light has been lost in the eye or the visual cortex long before it can cause any brain activities associated with consciousness, all information about events which are separated in time by as little as two or three seconds seems to be integrated into certain neuronal 'states of being conscious' (see Pöppel, Schill, and von Steinbüchel 1990). The 'moments of awareness' may even be discrete rather than reflecting the time continuum in terms of which the corresponding physical brain activities are successfully described. Continuous time then appears epistemologically as a heuristic abstraction — just as *all* concepts describing 'reality'. Similarly, the topology of colors (forming the color circle), or the identification of very different frequency mixtures of light as the same color, may be readily understood in terms of neuronal structures and activities. However, neither their subjective appearance (such as 'blue') nor the subjective appearance of the present can be derived from physical and physiological concepts. Nonetheless, the subjectively experienced *direction* of the apparent 'passage' of time may be understood as a consequence of the *thermodynamical* (objective) arrow that controls neurobiological processes, and thus allows memories *of the past and only of the past* to occur in those states-of-being-conscious.

In Einstein's special theory of relativity, the mechanistic or thermodynamico-mechanistic concept of time is applied locally, that is, along timelike world lines. These *proper times*, although being anholonomous (that is, path-dependent — as exemplified in the twin paradox), still possess the hypothetical absoluteness of Newton's time, since they are assumed to be defined (or to 'exist') even in the absence of anything that moves and thus may represent a clock. The absolute claim of proper time to control all conceivable motion is then formulated in the *principle of relativity.* However, any simultaneity of *moments of time* (events) at different locations in space is wholly arbitrary, as it would merely represent the choice of coordinates in spacetime. Objective geometrical and physical properties can instead be defined 'absolutely', that is, independent of any choice of coordinates. An example is the *abstract* spacetime metric (to be distinguished from its basis-dependent representation by a matrix $g_{\mu\nu}$) with its derived geometric concepts such as proper times and the light cone structure. Hence, one may still define a spacetime future and past *relative* to every spacetime point P (see Fig. 1.1), and unambiguously compare this orientation of light cones at different space time points by means of their path-independent parallel transport. This allows one to distinguish globally between forward and backward directions of time, and thus again to define a global thermodynamico-mechanistic concept of time. [3]

These concepts are still applicable to *general relativity* if one excludes non-orientable manifolds, which would permit the continuous transport of

[3] While superluminal objects ('tachyons') may be compatible with the relativistic light cone structure, they would pose severe problems to thermodynamics or the formulation of a physically reasonable boundary value problem (cf. Sect. 2.1).

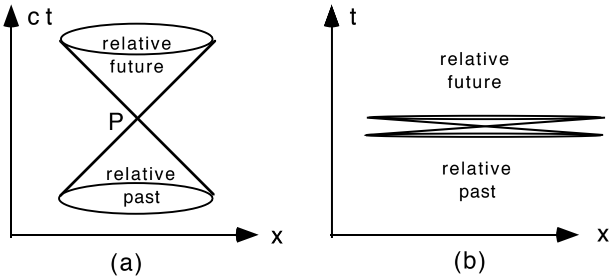


Fig. 1.1. (a) Local spacetime structure according to the theory of relativity. Spacetime future and past are defined *relative* to every event P, and independent of any choice of a frame of reference. (b) In conventional units (large numerical value of the speed of light) the light cone opens widely, so its exterior seems to degenerate into a space-like hypersurface of 'absolute' simultaneity. What we *observe* as an apparently objective present is in reality the backward light cone with respect to the subjective *Here-and-Now* P. Since only nonrelativistic speeds are relevant in our macroscopic neighbourhood, this apparent simultaneity *seems* also to coincide with the spacetime border to events which we (now) can in principle affect by our 'free will' (or things we can 'kick')

forward light cones into backward ones. One may also have to exclude solutions of the Einstein equations with closed time-like curves (world lines which return into their own past). Although compatible with general relativity, such spacetimes would not allow one to distinguish the past of an event on this curve from its future. What cosmic boundary conditions to the metric are required to exclude such spacetime geometries 'fact-like'?

If local states of matter are unambiguously defined at each spacetime point, a closed time-like curve must lead back to the *same* physical state (including all memories and clocks). This would be inconsistent with a persisting thermodynamical arrow and/or 'free will' along closed world lines, and thus eliminates the infamous paradox of murdering one's own grandfather. A 'spacetime traveler' would either have to stay on a loop in a periodic manner, or to meet his older self already at his *first* arrival at their meeting point — both in conflict with an evolutionary universe. Spacetime 'travel' is a misconception, related to that of the flow of time. It is, therefore, not surprising that spacetime geometries with closed time-like curves seem to be *dynamically unstable* (and thus could never *arise*) in the presence of thermodynamically normal matter (Penrose 1969, Friedman et al. 1990, Hawking 1992, Maeda, Ishibashi, and Narita 1998). Nonetheless, such scenarios are apparently quite popular even among certain well known relativists.

If closed time-like curves can in fact be neglected, then our spacetime can still be time-ordered by means of a monotonic foliation. While there have been speculations about 'time warps' in *quantum gravity* (see Morris, Thorne and Yurtsever 1988, Frolov and Novikov 1990), their consistent description would have to take into account the rigorous revision of the concept of time that is required in this theory (see Sect. 6.2). Any *quasi-classical* spacetime would here already require the time arrow of decoherence (see Sects. 4.3 and 6.2.2). In quantum theory, the dynamically evolving states are generically strongly entangled, that is, nonlocal (Sect. 4.2). Therefore, they cannot evolve locally (along time-like *curves* in spacetime).

In contrast to flat (Minkowski) spacetime, in general relativity world lines may begin and end on spacetime singularities at finite values of their proper times. This prevents the general applicability of Zermelo's recurrence objection that was raised against Boltzmann's statistical mechanics (see Chap. 3).

The most important *novel* aspect of general relativity for the concept of time is the *dynamical* role played by spacetime geometry. It puts the geometry of space-like hypersurfaces in the position of 'physical objects' that evolve dynamically and interact with matter (see Sect. 5.4). In this way, spatial geometry becomes a *physical* clock, and the spirit of Leibniz and Mach may finally be fulfilled by completely eliminating any relic of absolute time. Traditionally, proper times (defined by the abstract metric) are regarded as a *prerequisite* for the formulation of dynamical laws rather than as properties of a dynamical object. In general relativity, however, the spacetime metric is not the *exclusive* definer of time as a controller of motion (although geometry still dominates over matter because of the large value of the Planck mass — see Sect. 6.2.2). This situation is reminiscent of Leibniz's elimination of the special role played by the celestial bodies, when defining time by *all* relative motion in the universe. The dynamical role of geometry also permits (and requires) the quantization of time (Sect. 6.2).

This *physicalization of time* (that may formally appear as its *elimination*) in accordance with Mach's principle then allows one to speak of a direction *of* time instead of a direction *in* time, provided an appropriate asymmetry in the spacetime history of our universe does exist. However, as a consequence of the quantization of gravity, even the concept of a history of the universe as a parametrizable succession of global states has to be abandoned. The conventional concept of time can only be upheld as a quasi-classical approximation.

General literature: Reichenbach 1956, Mittelstaedt 1976, Whitrow 1980, Denbigh 1981, Barbour 1989

2. The Time Arrow of Radiation

– Prelim. 4th edtn. (Feb 01): www.time-direction.de

After a stone has been dropped into a pond, one observes concentrically diverging ('defocussing') waves. Similarly, *after* an electric current has been switched on, one finds a retarded electromagnetic field that is moving *away* from its source. Since the fundamental laws of nature, which describe these phenomena, are invariant under time-reversal, they are equally compatible with the reverse phenomena, in which concentrically focussing waves (and whatever was caused by the stone — such as heat) would 'conspire' in order to eject the stone out of the water. Deviations from the time reversal symmetry of the laws would modify this argument only in detail, as one merely had to alter the reverse phenomena correspondingly (cf. the Introduction). Such reverse phenomena have, however, never been observed in nature. The absence of focusing processes in high-dimensional configuration space may similarly describe the time arrow of thermodynamics (Chap. 3) or, when applied to wave functions, even that of quantum theory (see Sect. 4.6).

Electromagnetic radiation will here be considered as an example for wave phenomena in general. It may be described in terms of the four-potential A^μ, which in the Lorentz gauge obeys the wave equation

$$-\partial^\nu \partial_\nu A^\mu(\boldsymbol{r},t) = 4\pi j^\mu(\boldsymbol{r},t) \quad \text{with} \quad \partial^\nu \partial_\nu = -\partial_t^2 + \Delta \quad , \tag{2.1}$$

with $c = 1$, where the notations $\partial_\mu := \partial/\partial x^\mu$ and $\partial^\mu := g^{\mu\nu}\partial_\nu$ are used together with Einstein's convention of summing over identical upper and lower indices. When an appropriate boundary condition is imposed, one may write A^μ as a functional of the sources j^μ. For two well known boundary conditions one obtains the *retarded* and the *advanced* potential,

$$A^\mu_{\text{ret}}(\boldsymbol{r},t) = \int \frac{j^\mu(\boldsymbol{r},t-|\boldsymbol{r}-\boldsymbol{r}'|)}{|\boldsymbol{r}-\boldsymbol{r}'|} d^3\boldsymbol{r}' \quad , \tag{2.2a}$$

$$A^\mu_{\text{adv}}(\boldsymbol{r},t) = \int \frac{j^\mu(\boldsymbol{r},t+|\boldsymbol{r}-\boldsymbol{r}'|)}{|\boldsymbol{r}-\boldsymbol{r}'|} d^3\boldsymbol{r}' \quad . \tag{2.2b}$$

These two functionals of $j^\mu(\boldsymbol{r},t)$ are related to one another by a time reversal transformation (see (2.5) below). Their linear combinations are also solutions of the wave equation (2.1).

At this point, many textbooks argue somewhat mysteriously that 'for reasons of causality', or 'for physical reasons', only the retarded fields, derived from the potential (2.2a) according to $F^{\mu\nu}_{\text{ret}} := \partial^\mu A^\nu_{\text{ret}} - \partial^\nu A^\mu_{\text{ret}}$, may occur. This is evidently an independent condition, based on causal experience. It is *added* to the deterministic laws such as (2.1), which emerged historically from the traditional concept of causality. This example allows us to formulate in a preliminary way what seems to be meant by this *intuitive notion of causality:* correlated effects (that is, *non-local* regularities such as coherent waves) must always possess a *local* common cause (in their past).[1] However, this asymmetric notion of causality is a major *explanandum* of the physics concerned with the direction of time. As pointed out in the Introduction, it cannot be derived from known dynamical laws.

The popular argument that advanced fields are not found in nature because of their *improbable* initial correlations is known from statistical mechanics, but absolutely insufficient (see Chap. 3). The observed retarded phenomena are precisely as improbable among *all possible* ones, since they contain equally improbable *final* correlations. Their 'causal' explanation from an initial conditions would just beg the question.

Some authors have claimed that retarded waves represent emission, while advanced ones have to be used to describe absorption. However, this interpretation ignores the vital fact that absorbers give rise to *retarded* shadows (that is, to retarded waves that interfer destructively with incoming ones). In spite of the retardation, energy may thus flow from the electromagnetic field into an antenna. When incoming fields are present (as is the generic case), retardation does not necessarily mean emission of energy (see Sect. 2.1).

At the beginning of the last century, Ritz proposed a radical solution of the problem by postulating the exclusive existence of retarded waves *as a law.* Such time-directed *action at a distance* is equivalent to fixing the boundary conditions for the electromagnetic field in a universal manner. The field would then not possess any degrees of freedom.

This proposal, probably supported by many physicists at that time, led to a famous controversy with Einstein, who favored the point of view that retardation of radiation can be explained by thermodynamical arguments. Einstein, too, argued by means of an action-at-a-distance theory (see Sect. 2.4), which had played an important role historically because of its analogy with

---

[1] In the case of a *finite* number of effects (correlated future 'events') resulting from one local cause, this situation is often described as a 'fork' in spacetime (cf. Horwich 1987, Sect. 4.8). This *fork of causality* should not be confused with the *fork of indeterminism* (in configuration space and time), which points to different (in general global) *potential states* rather than to different events (see Footnote 7 of Chap. 3 and Fig. 3.8). The fork of causality ('intuitive causality') also characterizes measurements and the documentation of their results, that is, the formation and distribution of information. It is related to Reichenbach's (1956) concept of *branch systems*, and to Price's (1996) *principle of independence of incoming influences* (PI^3). Insofar as it describes the cloning and spreading of information, it represents an *overdetermination* of the past by the future (Lewis 1986), observed as a *consistency of documents* about the *macroscopic* past. These correlations let the past appear 'fixed' (unaffected by *local* events), while complete documents about *microscopic* history would be in conflict with thermodynamics and quantum theory.

Newton's non-retarded gravitation. At the end of their dispute, both authors published a short letter in order to formulate their different opinions. After an introductory sentence, according to which retarded and advanced fields are equivalent "in some situations", the letter reads as what appears also to be a verbal compromise (Einstein and Ritz 1909 — my translation):[2]

"While Einstein believes that one may restrict oneself to this case without essentially restricting the generality of the consideration, Ritz regards this restriction as not allowed in principle. If one accepts the latter point of view, experience requires one to regard the representation by means of the retarded potentials as the only possible one, provided one is inclined to assume that the fact of the irreversibility of radiation processes has to be present in the laws of nature. Ritz considers the restriction to the form of the retarded potentials as one of the roots of the Second Law, while Einstein believes that the irreversibility is exclusively based on reasons of probability."

Ritz thus conjectured that the thermodynamical arrow of time might be explained by the retardation of electrodynamics because of the latter's universal importance for all matter. However, if one does not want to modify mechanics in a similar way, the retardation of mechanical waves (such as sound) would have to be explained quite differently (for example, by using thermodynamical arguments again).

In Maxwell's local field theory the problem does not appear to be so rigorously defined, since in every bounded region of spacetime there may be 'free fields' which do not possess any sources *in this region.* Therefore, one can consistently understand Ritz's hypothesis only cosmologically: *all* fields must possess advanced sources ('causes' in their past) somewhere in the universe. While the above-discussed examples demonstrate that the time arrow of radiation does not just reflect the way how boundary conditions are posed, in cosmological context this becomes even more evident from the time-reversed question: "Do all fields in the universe possess a *retarded source* (a sink in time-directed terms), that is, will they all vanish somewhere in the future?" This assumption would be physically equivalent to the *absorber theory of radiation*, a T-symmetric action-at-a-distance theory to be discussed in Sect. 2.4). The observed asymmetries would then require an asymmetry in the distribution of sources in the distant past and future.

[2] The original text reads: "*Während Einstein glaubt, daß man sich auf diesen Fall beschränken könne, ohne die Allgemeinheit der Betrachtung wesentlich zu beschränken, betrachtet Ritz diese Beschränkung als eine prinzipiell nicht erlaubte. Stellt man sich auf diesen Standpunkt, so nötigt die Erfahrung dazu, die Darstellung mit Hilfe der retardierten Potentiale als die einzig mögliche zu betrachten, falls man der Ansicht zuneigt, daß die Tatsache der Nichtumkehrbarkeit der Strahlungsvorgänge bereits in den Grundgesetzen ihren Ausdruck zu finden habe. Ritz betrachtet die Einschränkung auf die Form der retardierten Potentiale als eine der Wurzeln des Zweiten Hauptsatzes, während Einstein glaubt, daß die Nichtumkehrbarkeit ausschließlich auf Wahrscheinlichkeitsgründen beruhe.*"

2.1 Retarded and Advanced Form of the Boundary Value Problem

In order to distinguish the indicated pseudo-problem that concerns only the definition of 'free' fields from the physically meaningful question, one has to investigate the general boundary value problem for hyperbolic differential equations (such as the wave equation). This can be done by means of Green's function, defined as a solution of the specific inhomogeneous wave equation

$$-\partial^\nu \partial_\nu G(\boldsymbol{r}, t; \boldsymbol{r}', t') = 4\pi\delta^3(\boldsymbol{r} - \boldsymbol{r}')\delta(t - t') \quad , \tag{2.3}$$

with an appropriate boundary condition in space and time. Some of the concepts and methods to be devoloped below will be applicable in a similar form in Sect. 3.2 to the investigation of solutions of the Liouville equation (Hamilton's equations for ensembles). Using (2.3), a solution of the inhomogeneous wave equation (2.1) may then be written in the form of pure *source terms*

$$A^\mu(\boldsymbol{r}, t) = \int G(\boldsymbol{r}, t; \boldsymbol{r}', t') j^\mu(\boldsymbol{r}', t')\, d^3r'\, dt' \quad . \tag{2.4}$$

The boundary conditions for $G(\boldsymbol{r}, t; \boldsymbol{r}', t')$ determine here those for $A^\mu(\boldsymbol{r}, t)$. Special (retarded or advanced) solutions are obtained from Green's functions G_{ret} and G_{adv}, which are defined as

$$G_{\substack{\mathrm{adv} \\ \mathrm{ret}}}(\boldsymbol{r}, t; \boldsymbol{r}', t') := \frac{\delta(t - t' \pm |\boldsymbol{r} - \boldsymbol{r}'|)}{|\boldsymbol{r} - \boldsymbol{r}'|} \quad . \tag{2.5}$$

The potentials A^μ_{ret} and A^μ_{adv} resulting from (2.4) are thus functionals of sources on the past or future light-cone of their arguments, respectively.

In contrast, Kirchhoff's formulation of the boundary value problem allows one to express *any* solution $A^\mu(\boldsymbol{r}, t)$ of the wave equation by means of Green's function $G(\boldsymbol{r}, t; \boldsymbol{r}', t')$ with *arbitrary* boundary conditions. This can be done by using the three-dimensional Green's theorem (valid for all sufficiently regular functions G and A^μ)

$$\begin{aligned} &\int_V [G(\boldsymbol{r}, t; \boldsymbol{r}', t')\Delta' A^\mu(\boldsymbol{r}', t') - A^\mu(\boldsymbol{r}', t')\Delta' G(\boldsymbol{r}, t; \boldsymbol{r}', t')]\, d^3r' \\ &\quad = \int_{\partial V} [G(\boldsymbol{r}, t; \boldsymbol{r}', t')\nabla' A^\mu(\boldsymbol{r}', t') - A^\mu(\boldsymbol{r}', t')\nabla' G(\boldsymbol{r}, t; \boldsymbol{r}', t')] \cdot d\boldsymbol{S}' \ , \end{aligned} \tag{2.6}$$

where $\Delta = \nabla^2$, while ∂V is the boundary of the spatial volume V, and t' is a fixed time. Multiplying (2.3) with $A^\mu(\boldsymbol{r}', t')$, integrating over the δ-functions on the right hand side (RHS), twice integrating by parts with respect to t' on the left hand side (LHS), and using (2.1) and (2.6), one obtains

$$
\begin{aligned}
A^\mu(\boldsymbol{r},t) &= \int_{t_1}^{t_2}\int_V G(\boldsymbol{r},t;\boldsymbol{r}',t')j^\mu(\boldsymbol{r}',t')\,d^3r'\,dt' \\
&\quad - \frac{1}{4\pi}\int_V [G(\boldsymbol{r},t;\boldsymbol{r}',t')\partial_{t'}A^\mu(\boldsymbol{r}',t') - A^\mu(\boldsymbol{r}',t')\partial_{t'}G(\boldsymbol{r},t;\boldsymbol{r}',t')]\,d^3r'\,\Big|_{t_1}^{t_2} \\
&\quad + \frac{1}{4\pi}\int_{t_1}^{t_2}\int_{\partial V}[G(\boldsymbol{r},t;\boldsymbol{r}',t')\nabla' A^\mu(\boldsymbol{r}',t') - A^\mu(\boldsymbol{r}',t')\nabla' G(\boldsymbol{r},t;\boldsymbol{r}',t')]\cdot d\boldsymbol{S}'\,dt' \\
&\equiv \text{‘source term’} + \text{‘boundary terms’} \quad . \qquad (2.7)
\end{aligned}
$$

For every spacetime point $P(\boldsymbol{r},t)$ within the boundary, both (past and future) light cones will in general contribute to the terms occurring in (2.7), as indicated in Fig. 2.1.

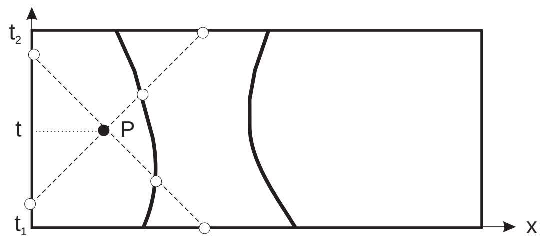


Fig. 2.1. Kirchhoff's boundary value problem, including initial, final and spatial boundaries. Sources (thick world lines) and boundaries contribute to the electromagnetic potential A^μ at the spacetime point P from events (open circles) on the light cones (dashed lines) according to any freely *chosen* Green's function

The T-symmetry of this *representation* of the potential as a sum of a source term and boundary terms in the past *and* future can be broken by the choice of one or the other asymmetric Green's functions (2.5). The spacetime boundary required for determining the potential at time t then assumes one of the two forms indicated in Fig. 2.2. Hence, the *same* potential can be written according to one or the other RHS of

$$
\begin{aligned}
A^\mu = \text{source term} + \text{boundary terms} &= A^\mu_{\text{ret}} + A^\mu_{\text{in}} \\
&= A^\mu_{\text{adv}} + A^\mu_{\text{out}} \quad . \qquad (2.8)
\end{aligned}
$$

For example, A^μ_{in} is here that solution of the *homogeneous* equations which coincides with A^μ for $t = t_1$. A^μ_{ret} and A^μ_{adv} vanish by definition for $t = t_1$ or $t = t_2$, respectively. *Any* situation can therefore be described as an *initial*

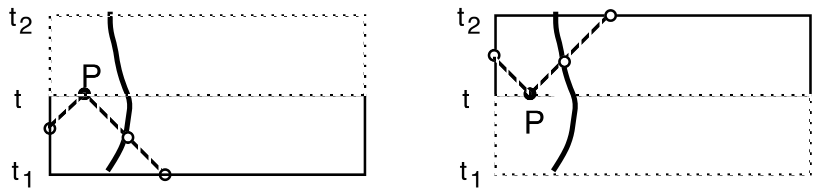


Fig. 2.2. Two representations of the *same* electromagnetic potential at time t by means of retarded *or* advanced Green's functions. They require data on boundaries (indicated by the solid lines) corresponding to an initial or a final-value problem, respectively

or *final value problem.* This symmetry represents T-symmetric determinism, while 'causality' is often used as an *ad hoc* argument for choosing G_{ret}.

However, *two* temporal boundary conditions would in general lead to consistency problems even if individually incomplete (see also Sect. 5.3). Fields formally resulting from different past and future sources (that is, retarded and advanced fields) *do not add* — except in mixed representations of the *same* field. In field theory, no (part of the) field 'belongs to' a certain source (in contrast to action-at-a-distance theories). Rather, sources in a bounded spacetime volume determine only the *difference* between the ('real') outgoing and incoming fields (as an anti-symmetric relation between them — similar to the action of $S-1$ in the interaction picture of the S-matrix). In causal *language*, where A^{μ}_{in} is assumed to be 'given', the source always 'creates' precisely its retarded field.

From a *physical* point of view, spatial boundary conditions always represent an *interaction* with the spatial environment. For infinite spatial volume, $V \to \infty^3$, (when the light cone cannot reach ∂V within finite time $t - t_1$), or in a closed universe, one loses this boundary term in (2.7), and thus obtains the *pure* initial value problem ($t > t_1$),

$$
\begin{aligned}
A^{\mu} = A^{\mu}_{\mathrm{ret}} + A^{\mu}_{\mathrm{in}} \equiv \int\limits_{t_1}^{t} \int\limits_{R^3} G_{\mathrm{ret}}(\boldsymbol{r}, t; \boldsymbol{r}', t') j^{\mu}(\boldsymbol{r}', t') \, d^3r' \, dt' \quad + \\
\frac{1}{4\pi} \int\limits_{R^3} \left[G_{\mathrm{ret}}(\boldsymbol{r}, t; \boldsymbol{r}', t_1) \partial_{t_1} A^{\mu}(\boldsymbol{r}', t_1) - A^{\mu}(\boldsymbol{r}', t_1) \partial_{t_1} G_{\mathrm{ret}}(\boldsymbol{r}, t; \boldsymbol{r}', t_1) \right] d^3r' \,,
\end{aligned}
\tag{2.9a}
$$

and correspondingly the pure final value problem ($t < t_2$),

$$
\begin{aligned}
A^{\mu} = A^{\mu}_{\mathrm{adv}} + A^{\mu}_{\mathrm{out}} \equiv \int\limits_{t}^{t_2} \int\limits_{R^3} G_{\mathrm{adv}}(\boldsymbol{r}, t; \boldsymbol{r}', t') j^{\mu}(\boldsymbol{r}', t') \, d^3r' \, dt' \quad - \\
\frac{1}{4\pi} \int\limits_{R^3} \left[G_{\mathrm{adv}}(\boldsymbol{r}, t; \boldsymbol{r}', t_2) \partial_{t_2} A^{\mu}(\boldsymbol{r}', t_2) - A^{\mu}(\boldsymbol{r}', t_2) \partial_{t_2} G_{\mathrm{adv}}(\boldsymbol{r}, t; \boldsymbol{r}', t_2) \right] d^3r' \,.
\end{aligned}
\tag{2.9b}
$$

The different signs at t_1 and t_2 are due to the fact that the gradient in the direction of the outward-pointing normal vector has now been written as a derivative with respect to t_1 (inward) or t_2 (outward).

One finds the retarded potential $A^\mu = A^\mu_{\mathrm{ret}}$ precisely if $A^\mu_{\mathrm{in}} = 0$. (Only its 'Coulomb part', required by Gauß's law, and often regarded as kinematics, must always be contained in A^μ_{in} as a consequence of charge conservation.) In scattering theory, a condition fixing the incoming wave (usually as a plane wave) is called a *Sommerfeld radiation condition*. Aside from its gauge contribution it describes the *actual* physical situation. Therefore, the physical problem is not which of the two forms, (2.9a) or (2.9b), is *correct* (both are), but:

1. Why does the Sommerfeld radiation condition $A^\mu_{\mathrm{in}} = 0$ (in contrast to $A^\mu_{\mathrm{out}} = 0$) approximately apply in most situations?
2. Why are initial conditions more *useful* than final conditions?

The second question is related to what in the Introduction has been called the *historical nature* of the world. Answers to these questions will be discussed in Sect. 2.2.

The form (2.7) of the four-dimensional boundary value problem, characteristic of determinism in field theory, applies to partial differential equations of the hyperbolic type (that is, with a Lorentzian signature $-+++$). Elliptic type equations would instead lead to the Dirichlet or von Neumann problem, which require values of the field *or* its normal derivative on a complete *closed* boundary (which in spacetime would include past *and* future). Only hyperbolic equations lead generally to 'propagating' solutions that allow one to impose free initial conditions. They are thus responsible for the concept of a *dynamical state* of the field. (See also Sect. 6.2 for an application to 'static' quantum gravity).

As is well known, the wave equation with its hyperbolic signature can be derived from Newton's equations as the continuum limit for a spatial lattice of mass points, kept at their positions by means of harmonic forces. For a linear chain, $md^2q_i/dt^2 = -k[(q_i - q_{i-1} - a) - (q_{i+1} - q_i - a)]$ with $k > 0$, this is the limit $a \to 0$ for fixed ak and m/a. The crucial restriction to 'attractive' forces ($k > 0$) may appear surprising, since Newton's equations are *always* deterministic, and allow one to pose initial conditions regardless of the type or sign of the forces. Vibrating systems are, however, characterized by a *stable* equilibrium (here described by the lattice constant a). An elliptic differential equation (with signature $++++$) would result in the same limit $a \to 0$ from the dynamics of variables q_i with repulsive forces ($k < 0$) for deviations from their *unstable* equilibrium. Their singular repulsion away from this equilibrium would cause the particle distances $q_i - q_{i-1}$ to explode instantaneously. The unstable trivial solution $q_i - q_{i-1} = a$ is in this case the only eigensolution of the Dirichlet problem with eigenvalue 0 (derived from the condition of a bounded final state). Mathematically, the dynamically diverging solutions simply do not 'exist' any more in the continuum limit.

For second order wave equations, the hyperbolic signature forms the basis for all (exact or approximate) conservation laws, which give rise to the continuity of 'objects' in time (including the 'identity' of observers if described as dynamical structures on a continuum). For example, the free wave equation has solutions with a *fixed form* $f(z \pm ct)$, while the Klein-Gordon equation with a positive variable $m^2 = V(\boldsymbol{r}, t)$ has unitary solutions $i\partial\phi(\boldsymbol{r}, t)/\partial t = \pm\sqrt{-\Delta + V}\phi(\boldsymbol{r}, t)$. These dynamical consequences of the metric, which lead to 'wave tubes' rather than wave packets, seem to be crucial for what appears as the inevitable 'progression of time' (in contrast to our freedom to move in space). However, the *direction* of this apparent flow of time requires *additional* conditions.

In this Section, the boundary value problem has been discussed for fields in the presence of *given* sources. In general, these sources would have to be dynamically determined in turn from these fields by means of the Lorentz force. For a charge continuum, the resulting coupled system of equations (magneto-hydrodynamics) is, of couse, still T-symmetric, while all considerations regarding the retardation of the electromagnetic fields in this and the following Section remain valid. New problems will arise with the idealization of point charges or rigid charged 'objects' (see Sect. 2.3).

2.2 Thermodynamical and Cosmological Properties of Absorbers

Wheeler and Feynman (1945, 1949) took up the Einstein-Ritz controversy about the relation of the two time arrows of radiation and thermodynamics in two important papers. Their analysis essentially confirms Einstein's point of view, provided his 'reasons of probability' are replaced with 'thermodynamical reasons'. Statistical reasons allone will prove insufficient for deriving a thermodynamical arrow (see Chap. 3.) The major part of Wheeler and Feynman's work is again based on a T-symmetric action-at-a-distance theory, which is particularly suited for representing the arguments in their historical context. From the point of view of local field theory (that for good reasons is preferred today), this description may appear somewhat disturbing or even misleading. Their *absorber theory of radiation* will therefore be postponed to Sect. 2.4, while the relevant physical properties of absorbers will now be discussed in the language of conventional field theory.

A relation between the two time arrows can be derived from presumed thermodynamical aspects of absorbers (in the 'ideal' case with infinite heat capacity at absolute zero). They can be described by the following definition:

A spacetime region is an '(ideal) absorber' if any radiation in it is (immediately) thermalized at the absorber temperature T $(= 0)$.

The process of *thermalization* used in this definition is based on the arrow of time defined by the Second Law. For electromagnetic waves it can be phenomenologically described in terms of the Maxwell equations by means of a complex refractive index, whereby the sign of the imaginary part characterizes the thermodynamical time direction. The above definition means that no radiation propagates within ideal absorbers, and in particular that no radiation may *leave* the absorbing region (along forward light-cones). This property can be applied to the boundary value problem (see Fig. 2.3):

Absorbers forming (parts of) a spacetime boundary in (2.7) do *not* contribute by means of the *retarded* Green's function (except for their thermal radiation).

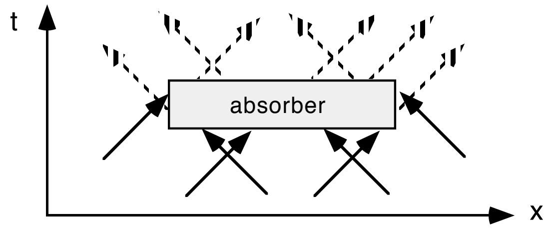


Fig. 2.3. An ideal absorber as a boundary does not contribute by means of G_{ret}. It would formally contribute (destructively) as a 'source' *inside* the spacetime region of interest. (Arrows represent here the formal time direction of retardation.)

This boundary condition simplifies the *initial* value problem. If the spacelike part ∂V of the boundary required for the retarded form of the boundary value problem depicted in Fig. 2.2 consists entirely of ideally absorbing walls (as is usually the case in a laboratory in the frequency range well above that of relevant heat radiation), the condition $A^{\mu}_{\mathrm{in}} = 0$ is effective shortly after the initial time t_1 that is used to define the 'incoming' field. In this case, precisely the retarded fields of sources inside the laboratory will be present. On the other hand, absorbers on the boundary do *not* restrict the contribution of G_{adv}; in the nontrivial case one has $A^{\mu}_{\mathrm{out}} \neq 0$. In this laboratory situation, the radiation arrow may thus easily be derived from the thermodynamical arrow that characterizes absorbers.

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 $A^{\mu}_{\mathrm{in}} \approx 0$, although the present universe is transparent to visible light. Can the darkness of the night sky then be understood in a realistic cosmological model? For the traditional (static and homogeneous) model of the universe this was impossible, a situation called *Olbers' paradox* after one of the first astronomers who recognized this problem. The total brightness B of the sky would then be given by

$$B = 4\pi \int_0^\infty \varrho L_{\mathrm{a}}(r) r^2 \, dr \quad , \tag{2.10}$$

where ϱ is the number density of sources (essentially the fixed stars), while $L_\mathrm{a}(r) = \bar{L}/r^2$ is their mean apparent luminosity. In the static and homogeneous situation ($\bar{L}, \varrho =$ constant) this integral diverges linearly, and the night sky should be infinitely bright. The screening of distant sources by closer ones would reduce this result to a finite value, corresponding to a sky as bright as the mean surface of a star. It would not help to take into account dark matter in the universe, since this would have to be in thermal equilibrium with the radiation in this case. An eternal homogeneous universe should therefore be found in the situation of a 'heat death' (see Chap. 3).

Olbers' paradox could be resolved on empirical grounds after Hubble had discovered the cosmic redshift of galaxies, proportional to their distance. They indicated an expanding universe, and therefore its origin in a *big bang* at a finite time of order 10^{10} years in the past. According to Wien's displacement law $T_\mathrm{a} \propto \lambda^{-1}$, the increase of all wave lengths λ with the expansion parameter $a(t)$ leads to the reduction of the apparent temperature T_a of the sources. Because of Stefan and Boltzmann's law, $L \propto T^4$, the apparent brightness of the stars, L_a, would then decrease with distance faster than r^{-2}. In a homogeneous (though expanding) universe, the brightness of the sky is given by

$$B \propto \int_0^{\tau_\mathrm{max}} \varrho(t_0 - \tau)\bar{L}(t_0 - \tau)\left[\frac{a(t_0 - \tau)}{a(t_0)}\right]^4 d\tau \quad , \tag{2.11}$$

where t_0 is the present, while τ_max is the age of the transparent universe. If the total number of sources had not changed, their density would have varied with $\varrho(t_0 - \tau) \propto [a(t_0 - \tau)]^{-3}$. If their mean absolute luminosity, $\bar{L}$, had also remained constant, the integrand would be proportional to $a(t_0 - \tau)/a(t_0)$. The integral (2.11) converges in most realistic cosmologies, in particular for a finite cosmic age τ_max, while times as far back as $10^{23}y$, that is, 10^{13} times the Hubble age, would have to dominate in the integral in order to produce a night sky as bright as the surface of stars (Harrison 1977).

Starlight is therefore completely negligible for the background radiation of the night sky. While this argument resolves Olbers' paradox, it is neither sufficient to explain the cosmological condition $A^\mu_\mathrm{in} \approx 0$, nor is it realistic, since the nature of the sources must have changed drastically during the early history of the universe. During its 'radiation era' our universe was very hot. Matter was ionized and almost homogeneous (cf. Wu, Lahav, and Reeves 1999), representing a non-ideal absorber with temperature of several thousand degrees (see Fig. 2.4). Because of the expansion, its thermal radiation has now cooled down to its observed value of 2.7 K, compatible with the darkness of the night sky.

The cosmic expansion, which is vital for this quantitative argument, is thus also essential for the drastic thermodynamical non-equilibrium on the cosmic scale, that is, the contrast between cold interstellar space and the hot stars (which are gaining their energy from gravitational contraction — see Chap. 5). For this reason, the expansion of the universe has often been

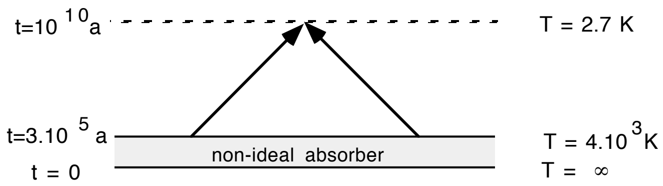


Fig. 2.4. The cosmological initial value problem for the electromagnetic radiation. The thermal contribution of the non-ideal absorber represented by the hot, ionized matter during the radiation era has now cooled down to the measured background radiation of 2.7 K (which can be neglected for most purposes)

proposed as the *master arrow* of time. However, it would be quite insufficient to use such *causal* arguments for this purpose. For indeed, in order to reverse thermodynamical processes in a *contracting* universe, incoming radiation would have to form precisely the advanced fields of its future sources, and thus interfere with their retarded fields in a destructive manner. The scenario of fields and phenomenological absorbers in an expanding universe is far too simple to describe a master arrow. This discussion will be resumed in Sects. 5.3 and 6.2.

The thermodynamical arrow entered this derivation of the radiation arrow through the assumption of negligible (in the ideal case vanishing) thermal radiation on the *future* 'light shadows' of absorbers, in particular the absorber represented by the cosmic radiation era. Hogarth (1962) suggested instead that the time arrow of radiation requires a *change* in the opacity of intergalactic matter (cosmic absorbers) during the evolution of the universe. However, this assumption is neither sufficient nor necessary, since absorbers specify a direction in time even when they do *not* essentially change. Their thermal equilibrium is not equivalent to the absence of any *microscopic* asymmetry (see the Appendix for an example). Hogarth's proposal was apparently motivated by the T-symmetric (and therefore thermodynamically inappropriate) definition of absorbers used by Wheeler and Feynman (Sect. 2.4).

These conclusions regarding the retardation of electromagnetic radiation apply in principle to all wave phenomena. The only exception may be gravity, since even the radiation era may have been transparent to gravitons.[3] Ritz's conjecture of a *law-like* nature of the retardation of electrodynamic fields will therefore be reconsidered and applied to gravity in Chap. 5.

[3] The retardation of gravitational waves has been indirectly confirmed by double pulsars (see Taylor 1994).

2.3 Radiation Damping

The emission of electromagnetic radiation by a charged particle that is accelerated by an external force requires that the particle reacts by *losing* the irradiated energy. Similar to the force of friction, this *radiation reaction* in the *effective equation of motion* must change sign under time reversal. As will be explained, this can be understood as a consequence of the retardation of the field when acting on its own source, even though the retardation seems to disappear at the position of a point source. However, the self-interaction of point-like charges leads to singularities (infinite mass renormalization) which need care when being separated from that part of the interaction which is responsible for radiation damping. While these problems can be avoided if any self-interaction is eliminated by means of the action-at-a-distance theory (described in Sect. 2.4), others then arise in their place.

Consider the trajectory of a charged particle, represented by means of its Lorentzian coordinates z^μ as functions of proper time τ. The corresponding four-velocity and four-acceleration are $v^\mu := dz^\mu/d\tau$ and $a^\mu := d^2z^\mu/d\tau^2$, respectively. From $v^\mu v_\mu = -1$ one obtains by differentiation $v^\mu a_\mu = 0$ and $v^\mu \dot{a}_\mu = -a^\mu a_\mu$. In a rest frame, defined by $v^k = 0$ (with $k = 1, 2, 3$), one has $a^0 = 0$. The four-current density of this point charge is given by

$$j^\mu(x^\nu) = \mathrm{e} \int v^\mu(\tau)\delta^4[x^\nu - z^\nu(\tau)]\, d\tau \quad . \tag{2.12}$$

Its retarded field $F^{\mu\nu}_{\mathrm{ret}} = 2\partial^{[\mu} A^{\nu]}_{\mathrm{ret}} := \partial^\mu A^\nu_{\mathrm{ret}} - \partial^\nu A^\mu_{\mathrm{ret}}$ is known as the Liénard-Wiechert field. The retarded or advanced fields can be written in an invariant manner (see, for example, Rohrlich 1965) as

$$\begin{aligned} F^{\mu\nu}_{\mathrm{ret/adv}}(x^\sigma) &= \pm \frac{2\mathrm{e}}{\varrho} \frac{d}{d\tau} \frac{v^{[\mu} R^{\nu]}}{\varrho} \\ &= \frac{2\mathrm{e}}{\varrho^2} v^{[\mu} u^{\nu]} + \frac{2\mathrm{e}}{\varrho} \left\{ a^{[\mu} v^{\nu]} - u^{[\mu} v^{\nu]} a_u \pm u^{[\mu} a^{\nu]} \right\} , \end{aligned} \tag{2.13}$$

with v^μ and a^μ taken at times τ_{ret} or τ_{adv}, respectively.

In this expression,

$$R^\mu := x^\mu - z^\mu \left(\tau_{\mathrm{ret/adv}} \right) =: (u^\mu \pm v^\mu)\varrho \quad , \tag{2.14}$$

with $u^\mu v_\mu = 0$ and $u^\mu u_\mu = 1$, is the light-like vector pointing from the retarded or advanced spacetime position z^μ of the source to the point x^μ where the field is considered. Obviously, ϱ is the distance in space and in time between these points in the rest frame of the source, while $a_u := a^\mu u_\mu$ is the component of the acceleration in the direction of the spatial distance vector u^μ. Retardation or advancement are enforced by the condition of R^μ being light-like, that is, $R^\mu R_\mu = 0$.

On the RHS of (2.13), second line, the field consists of two parts, the first one proportional to $1/\varrho^2$, the second one to $1/\varrho$. They are called the generalized Coulomb field ('near-field') and the radiation field ('far-field'), respectively. Since the stress energy tensor

$$T^{\mu\nu} = \frac{1}{4\pi}\left(F^{\mu\alpha}F_{\alpha}{}^{\nu} + \frac{1}{4}g^{\mu\nu}F^{\alpha\beta}F_{\alpha\beta}\right) \tag{2.15}$$

is quadratic in the fields, it consists of *three* parts characterized by their powers of ϱ. For example, one has

$$\begin{aligned} T^{\mu\nu}_{\text{ret}} := T^{\mu\nu}(F^{\mu\nu}_{\text{ret}}) = {} & \frac{\mathrm{e}^2}{4\pi\varrho^4}\left(u^\mu u^\nu - v^\mu v^\nu - \frac{1}{2}g^{\mu\nu}\right) \\ & + \frac{\mathrm{e}^2}{2\pi\varrho^3}\left\{a_u\frac{R^\mu R^\nu}{\varrho^2} - [v^{(\mu}\,a_u + a^{(\mu}]\frac{R^{\nu)}}{\varrho}\right\} \\ & + \frac{\mathrm{e}^2}{4\pi\varrho^2}(a_u^2 - a^\lambda a_\lambda)\frac{R^\mu R^\nu}{\varrho^2} \quad , \end{aligned} \tag{2.16}$$

where braces around pairs of indices define symmetrization: $v^{(\mu}\,R^{\nu)} := (v^\mu R^\nu + v^\nu R^\mu)/2$. Here, $T^{\mu\nu}$ is the ν-component of the current of the μ-component of four-momentum. In particular, T^{0k} is the Poynting-vector in the chosen Lorentz system, and $T^{\mu\nu}d^3\sigma_\nu$ is the flux of four-momentum through an element $d^3\sigma_\nu$ of a hypersurface. If $d^3\sigma_\nu$ is space-like (a volume element), this 'flux' describes its energy-momentum ('momenergy') content, otherwise it is the flux through a surface element during an element of time.

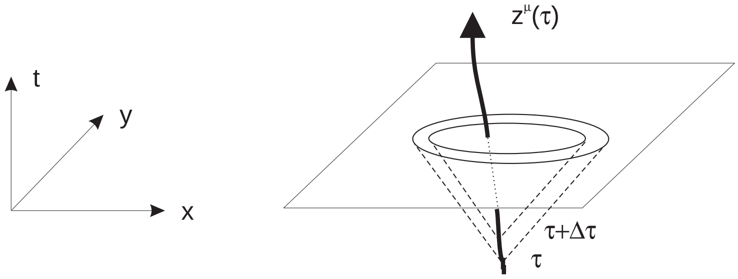


Fig. 2.5. Spacetime support of the retarded field of a world line element $\Delta\tau$ of a point charge is located between two light cones (co-axial only in the rest frame of the source). Flux of field momentum crosses light cones in the near-field region of the charge

The retarded field caused by an element of the world line of the point charge between τ and $\tau + \Delta\tau$ has support between forward light cones of the two end points, that is, on a thin four-dimensional conic shell (see Fig. 2.5). The intersection of the cones with a space-like hyperplane forms a spherical shell (concentric only in the rest frame at time τ, and in the figure depicted two-dimensionally as a narrow ring). The integral of the stress-energy tensor over this spherical spatial shell,

$$\Delta P^\mu = \int T^{\mu\nu}\, d^3\sigma_\nu \quad , \tag{2.17}$$

is the four-momentum of the field on this hyperplane 'caused' by the world line element Δz^μ. In general, this momentum is not conserved along light cones, since (2.16) contains a momentum flux orthogonal to the cones, due to the dragging of the near-field by the charge. (Therefore, Teitelboim 1970 suggested a different splitting of the energy-momentum tensor. It leads to a time-asymmetric electron dressing — valid only for 'given' F_{in}.) However, this flux component vanishes in the far-zone, where $T^{\mu\nu}$ is proportional to $R^\mu R^\nu$. In this region the integral (2.17) describes the four-momentum radiated *away* from the trajectory of the charge during the interval $\Delta\tau$,

$$\Delta P^\mu \underset{\varrho\to\infty}{\to} \Delta P^\mu_{\mathrm{rad}} = \frac{2}{3}\mathrm{e}^2 a^\lambda a_\lambda v^\mu \Delta\tau =: \Re v^\mu \Delta\tau \quad . \tag{2.18}$$

The quantity $\Re = \frac{2}{3}e^2 a^\lambda a_\lambda$ is called the *invariant rate of radiation.* In the co-moving rest frame ($v^k = 0$) one recovers the non-relativistic Larmor formula,

$$\Delta P^0_{\mathrm{rad}} = v_\mu \Delta P^\mu_{\mathrm{rad}} = \frac{2}{3}e^2 a^\lambda a_\lambda \Delta t = \frac{2}{3}e^2 \boldsymbol{a}^2 \Delta t \quad . \tag{2.19}$$

One confirms that the energy transfer into radiation in a positive interval of time cannot be negative. This is a consequence of the presumed retardation. It means that an accelerated charged particle must *lose* energy regardless of the direction of the driving external force.

Larmor's formula led to a certain confusion when it was applied to a charged particle in a gravitational field by means of the principle of equivalence. Because of its dependence on acceleration, (2.18) is restricted to inertial frames. In general relativity, inertial frames are freely falling. According to the principle of equivalence, a freely falling charge should then *not* radiate, while a charge 'at rest' in a gravitational field (under the influence of non-gravitational forces) should do so. This problem was not understood until Mould (1964) demonstrated that the response of a detector to radiation, too, depends on its acceleration (see also Fugmann and Kretzschmar 1991).

In general relativity, the principle of equivalence is only locally meaningful. However, a *homogeneous* gravitational field (as it would result from a massive plane) is described by flat spacetime. It is *globally* equivalent to a *rigid field of uniform accelerations* a^μ on Minkowski spacetime. This field corresponds to a set of 'parallel' hyperbolic trajectories with constant (but different) accelerations $a^\mu a_\mu$. These trajectories define rigid accelerated frames, since they preserve their distance in their common comoving rest frames. Together with their proper times they define the curved *Rindler coordinates* in spacetime (see (5.15) and Fig. 5.5 in Sect. 5.2).

The equivalence principle can therefore be applied globally to a *homogeneous* gravitational field. It requires that an inertial (freely falling) detector is *not* affected by an inertial charge, while a detector 'at rest' is excited. The latter would remain idle in the presence of a charge being 'equivalently at rest' (that is, at a fixed or even uniformly changing distance). A detector-independent definition of *total radiation* also turns out to depend on acceleration (as it should for consistency) because of the occurrence of spacetime horizons for truly uniform acceleration (see Boulware 1980 and Sect. 5.2).

The emission of energy according to (2.19) thus requires a deceleration of the point charge in order to conserve total energy. It should be possible to derive this consequence directly from the fundamental dynamical equations. They are given by the Lorentz force,

$$\mathcal{F}^{\mu}_{\mathrm{self}}(\tau) = eF^{\mu\nu}_{\mathrm{ret}}[z^{\sigma}(\tau)]v_{\nu}(\tau) \quad , \tag{2.20}$$

of the particle's self-field (its field at the position of the point charge itself). However, this expression leads into problems caused by the fact that the electromagnetic force acts only on the point charge, where the self-field is singular (its Coulomb part even with $1/\varrho^2$), while an essential part of the accelerated mass is contained in the co-moving Coulomb field. Dirac (1938) was able to show that the T-symmetric part $\bar{F}^{\mu\nu}$ of the retarded field,

$$F^{\mu\nu}_{\mathrm{ret}} = \tfrac{1}{2}(F^{\mu\nu}_{\mathrm{ret}} + F^{\mu\nu}_{\mathrm{adv}}) + \tfrac{1}{2}(F^{\mu\nu}_{\mathrm{ret}} - F^{\mu\nu}_{\mathrm{adv}}) \equiv: \bar{F}^{\mu\nu} + F^{\mu\nu}_{\mathrm{rad}} \quad , \tag{2.21}$$

is alone responsible for the infinite mass renormalization, while its T-antisymmetric part, $F^{\mu\nu}_{\mathrm{rad}}$, remains regular and does indeed describe the radiation reaction when treated properly.

In order to prove the second part of this statement, one has to expand all quantities in (2.13) up to the third order in terms of the retardation $\Delta\tau_{\mathrm{ret}} = \tau_{\mathrm{ret}} - \tau$, for example

$$\begin{aligned} v^{\nu}(\tau_{\mathrm{ret}}) = v^{\nu}(\tau + \Delta\tau_{\mathrm{ret}}) &= v^{\nu}(\tau) + \Delta\tau_{\mathrm{ret}}a^{\mu}(\tau) \\ &\quad + \tfrac{1}{2}\Delta\tau^2_{\mathrm{ret}}\dot{a}^{\mu}(\tau) + \tfrac{1}{6}\Delta\tau^3_{\mathrm{ret}}\ddot{a}^{\mu}(\tau) + \cdots \quad . \end{aligned} \tag{2.22}$$

All terms which are singular at the position of the point charge itself cancel from the antisymmetric difference of retarded and advanced fields, and one obtains (cf. Rohrlich 1965, p. 142)

$$F^{\mu\nu}_{\mathrm{rad}} = -\frac{4e}{3}\dot{a}^{[\mu}\,v^{\nu]} \quad . \tag{2.23}$$

The resulting T-antisymmetric Lorentz self-force (the *Abraham four-vector*)

$$\mathcal{F}^{\mu}_{\mathrm{rad}} := eF^{\mu\nu}_{\mathrm{rad}}v_{\nu} = \frac{2e^2}{3}(\dot{a}^{\mu} + v^{\mu}\dot{a}^{\nu}v_{\nu}) = \frac{2e^2}{3}(\dot{a}^{\mu} - v^{\mu}a^{\nu}a_{\nu}) \tag{2.24}$$

(using $a^\nu v_\nu = 0$ in the second step) should then describe the *radiation reaction* of a point charge. It leads to a nonlinear equation of motion (the *Lorentz-Dirac equation*). However, while the second term is in accordance with (2.18), the $\dot{a}^\mu$-term is ill-defined (see below). Together with the mass renormalization term due to $\bar{F}^{\mu\nu}$ it must describe the required four-momentum transfer from the point charge to its singular near-field.

In a rest frame (with $v^k = 0$ and $a^0 = 0$) one obtains

$$\mathcal{F}^0_{\mathrm{rad}} = -\frac{2e^2}{3}\boldsymbol{a}^2 \quad \text{and} \quad \mathcal{F}^k_{\mathrm{rad}} = \frac{2e^2}{3}\frac{da^k}{dt} \quad . \tag{2.25}$$

Therefore, the radiation reaction represents non-relativistically a force proportional to the change of acceleration, da^k/dt, while its fourth component describes the energy *loss* in accordance with the non-negative invariant rate of radiation (2.18). The latter was defined by the energy flux through a distant sphere on the future light-cone. This is often used to 'derive' the radiation reaction (2.25). However, global conservation laws may be used only if all their contributions are taken into account. For example, there is no similar conservation of *three-momentum* of the bare point charge and its far-field. The reason is the aforementioned momentum flux of the near field orthogonal to the future light-cone of the moving charge, which is required to keep the Coulomb field co-moving. For this reason, the uniformly accelerated charge may radiate with $\Re \neq 0$ even though the 'radiation reaction' $\mathcal{F}^\mu_{\mathrm{rad}}$ (including its ill-defined term) vanishes in this case, as can be shown separately for its two relevant components, $\mathcal{F}^\mu_{\mathrm{rad}} a_\mu$ and $\mathcal{F}^\mu_{\mathrm{rad}} v_\mu$.

If the boundary condition $F^{\mu\nu}_{\mathrm{in}} = 0$ does not hold, the complete electromagnetic force on a point charge is given by

$$ma^\mu = \mathcal{F}^\mu = \mathcal{F}^\mu_{\mathrm{in}} + \mathcal{F}^\mu_{\mathrm{rad}} = \mathcal{F}^\mu_{\mathrm{out}} - \mathcal{F}^\mu_{\mathrm{rad}} \tag{2.26}$$

(cf. Sect. 2.1 and (2.21)). The renormalization terms, caused by the T-symmetric part of the self-field, have now been brought to the LHS in the form $\Delta m a^\mu$. Equation (2.26) still exhibits the T-symmetry, but the latter may be broken fact-like by the given *initial* condition $F^{\mu\nu}_{\mathrm{in}}$ (in contrast to the uncontrollable outgoing radiation contained in $F^{\mu\nu}_{\mathrm{out}}$). The Lorentz-Dirac equation, based on (2.24) and the *first* RHS of (2.26), may then be written in the form

$$m(a^\mu - \tau_0 \dot{a}^\mu) = K^\mu(\tau) := \mathcal{F}^\mu_{\mathrm{in}} - \Re v^\mu \quad , \tag{2.27}$$

where $\tau_0 = 2e^2/3mc^2$ is the time required for light to travel a distance of the order of the 'classical electron radius' e^2/mc.

Both terms of (2.27) that result from the radiation reaction (2.24) change sign under time reversal (or the replacement of retarded with advanced fields). While the second one represents the friction-type radiation damping $-\Re v^k$, required for the conservation of energy, the first one (called the *Schott term*) is proportional to the *third* time-derivative of the position in an inertial frame. A solution to the equation of motion (2.27) would thus require *three* initial vectors as integration constants (the initial acceleration in addition to the usual initial position and velocity). Evidently, information has been lost through differentiation by the expansion (2.22). Even for $\mathcal{F}^\mu_{\text{in}} = 0$, the Lorentz-Dirac equation (2.27) admits *runaway* solutions, non-relativistically in the form of an exponentially increasing *self-acceleration*, $a^k(t) = a^k(0)\exp(t/\tau_0)$.

Because of this information loss, the Lorentz-Dirac equation can only describe a necessary condition for the motion of the point charge. In the free case, unphysical runaway solutions could simply be eliminated by fixing the artificial integration constant by the condition $a^k(0) = 0$. Unfortunately, this would still lead to runaway as soon as an external force is turned on. Since the formal solution to (2.27) with respect to a^μ is

$$ma^\mu(\tau) = \mathrm{e}^{\tau/\tau_0}\left[ma^\mu(0) - \frac{1}{\tau_0}\int\limits_0^\tau \mathrm{e}^{-\tau'/\tau_0} K^\mu(\tau')d\tau'\right] \quad , \tag{2.28}$$

Dirac proposed to fix the initial acceleration in terms of the *future force* according to $ma^\mu(0) = (1/\tau_0)\int_0^\infty \mathrm{e}^{-\tau'/\tau_0}K^\mu(\tau')d\tau'$. The substitution $\tau' \to \tau' + \tau$ then leads to Dirac's equation of motion,

$$ma^\mu(\tau) = \int\limits_0^\infty K^\mu(\tau + \tau')\frac{\mathrm{e}^{-\tau'/\tau_0}}{\tau_0}\,d\tau' \quad . \tag{2.29}$$

It represents a Newtonian (second order) equation of motion with an 'acausal force' K^μ. This force contains the required nonlinear deceleration $-\Re v^k$ (the radiation reaction proper), but acts according to a distribution in time that ends with a sharp maximum at the 'correct' time $\tau' = 0$ after growing exponentially with a time constant τ_0 in advance of the 'correct' force. How could this 'acausality' be derived from the Lorentz force by means of retarded fields?

Moniz and Sharp (1977) demonstrated that the pathological behavior of this 'classical electron' is the consequence of a mass renormalization that exceeds the physical electron mass (so that the bare mass becomes negative). For example, if the point charge is replaced with a rigid charged sphere of radius r_0 in its rest frame, one obtains, by using the now everywhere regular retarded field, an equation of motion that was first proposed by Caldirola (1956), and later derived by Yaghjian (1992) as an approximation. It reads

$$m_0 a^\mu(\tau) = \mathcal{F}^\mu_{\text{in}}(\tau) + \frac{2e^2}{3r_0}\frac{v^\mu(\tau - 2r_0) + v^\mu(\tau)v^\nu(\tau)v_\nu(\tau - 2r_0)}{2r_0} \quad , \tag{2.30}$$

where m_0 is the bare mass. The retardation $2r_0$ in the arguments is the travel time of the electromagnetic field within this model electron. It would change sign for advanced fields (consistent only in conjunction with given $\mathcal{F}^\mu_{\text{out}}$). Taylor expansion of (2.30) with respect to $2r_0$ (equivalent to (2.22)), and using $v^\mu v_\mu = -1$ and its time derivatives, leads in first order to a finite mass renormalization (4/3 of the electrostatic mass), and in second order back to the Lorentz-Dirac equation. The expectation that this limit of a point charge ($r_0 \to +0$) might lead to a T-symmetric equation of motion (Rohrlich 1998) is thus unjustified (Zeh 1999a). While (2.30) can be regarded as a non-Markovian master equation (see Sect. 3.2), the Lorentz-Dirac equation corresponds to its Markovian limit, valid for slowly varying fields. In this sense, the radiation reaction has to be calculated from the given history in order to *determine* the acceleration (rather than its derivative) towards the future (right derivative).

The force acting on the rigid 'electron' according to (2.30) is the asymmetric difference between a decelerating and an accelerating friction type force. For positive bare and physical masses it does not lead to runaway, although it may possess complicated non-analytic solutions, in particular for forces varying on a time scale shorter than the light travel time within the charged sphere. Dirac's preacceleration of the center of mass, now explained by the extended charge distribution, is avoided in (2.30) by the presumed rigidity of the sphere, which requires additional forces of constraint.

A discussion of the history of electron theory may be found in Rohrlich (1997). It indicates that the concept of a non-inertial point charge is inconsistent with classical electrodynamics, while external forces acting on a dynamically stable charge *distribution* would disturb its shape and structure. A quantum mechanical ground state may be protected against deformation by its large excitation energy. Before praising advantages of QED compared to the classical theory, however, one should pay attention to the question which of the dangerous terms have simply been omitted in this theory. Because of quantum entanglement, composite quantum objects (such as particle plus field) cannot be described in three-dimensional space (see Sect. 4.2).

General literature: Rohrlich 1965, 1997, Levine, Moniz and Sharp 1977, Boulware 1980

2.4 The Absorber Theory of Radiation

Ritz's retarded action-at-a-distance theory, mentioned at the beginning of this chapter, eliminates all electromagnetic degrees of freedom by postulating the *cosmological* initial condition $A^\mu_{\mathrm{in}} = 0$. Since electromagnetic forces would then act only on the forward light-cones of their sources, this theory cannot be compatible with Newton's third law, which requires the equivalence of action and reaction. The reaction to a retarded action must be advanced.[4]

[4] In *field theory*, sources and fields interact *locally* in spacetime. For this reason the self-force (2.24) could not be derived from the flux of field momentum in the far-zone.

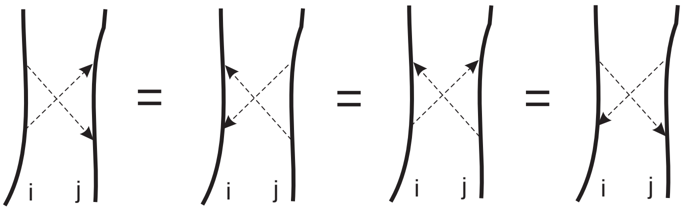


Fig. 2.6. Different interpretations of the *same* interaction term of the Hamiltonian for a pair of particles

In order to warrant energy-momentum conservation, an action-at-a-distance theory has to be formulated in a T-symmetric way, as done by Fokker (1929) by means of his action,

$$\begin{aligned} I = \int (T - V)\, dt = \sum_i m_i \int d\tau_i \\ - \frac{1}{2} \sum_{i \neq j} e_i e_j \iint v_i^\mu v_{j\mu} \delta[(z_i^\nu - z_j^\nu)(z_{i\nu} - z_{j\nu})]\, d\tau_i\, d\tau_j \quad . \end{aligned} \tag{2.31}$$

Here, indices i and j are particle numbers. (A sum over $i \neq j$ defines a double sum excluding equal indices, while a sum over $i(\neq j)$ indicates a sum over i only, excluding a given value j.) In (2.31), the particle positions z_i^μ and velocities v_i^μ have to be taken at the proper time τ_i of the corresponding particle, for example $z_i^\mu = z_i^\mu(\tau_i)$.

Expanding the δ-function in the potential energy according to

$$\delta(\Delta z^\nu \Delta z_\nu) = \delta(\Delta z_0^2 - \Delta \boldsymbol{z}^2) = \frac{1}{2|\Delta \boldsymbol{z}|}[\delta(\Delta z_0 - |\Delta \boldsymbol{z}|) + \delta(\Delta z_0 + |\Delta \boldsymbol{z}|)] \tag{2.32}$$

(with $\Delta z^\nu = z_i^\nu - z_j^\nu$) preserves its symmetric form. By integrating either over τ_i or over τ_j, one obtains, respectively, the first or second of the following expressions (first two graphs of Fig. 2.6),

$$\begin{aligned} \frac{e_i}{2} \int [A^\mu_{\mathrm{ret},j}(z_i^\sigma) + A^\mu_{\mathrm{adv},j}(z_i^\sigma)] v_{i\mu}\, d\tau_i \\ \equiv \frac{e_j}{2} \int [A^\mu_{\mathrm{adv},i}(z_j^\sigma) + A^\mu_{\mathrm{ret},i}(z_j^\sigma)] v_{j\mu}\, d\tau_j \quad . \end{aligned} \tag{2.33}$$

$A^\mu_{\mathrm{ret},j}$ and $A^\mu_{\mathrm{adv},j}$ are the retarded and advanced potential of the j-th particle according to (2.2a) and (2.2b). However, if the integral is always carried out with respect to the particle on the backward light-cone of the other one, one obtains, in spite of the preserved T-symmetry of the theory, only contributions in terms of retarded potentials (third graph),

$$\frac{e_i}{2} \int A^\mu_{\mathrm{ret},j}(z_i^\sigma) v_{i\mu}\, d\tau_i + \frac{e_j}{2} \int A^\mu_{\mathrm{ret},i}(z_j^\sigma) v_{j\mu}\, d\tau_j \quad , \tag{2.34}$$

and similarly, but time reversed, for the advanced potentials (fourth graph). Einstein seems to mean this equivalence of different *forms* of the interaction (in his Letter with Ritz quoted in the introduction to this chapter).

However, the Euler-Lagrange equations resulting from (2.31) always lead to T-symmetric forces which are in accordance with Newton's third law:

$$ma_i^\mu = \frac{e_i}{2} \sum_{j(\neq i)} [F^{\mu\nu}_{\mathrm{ret},j}(z_i^\sigma) + F^{\mu\nu}_{\mathrm{adv},j}(z_i^\sigma)] v_{i,\nu} \quad . \tag{2.35}$$

According to (2.8), this would require the cosmological boundary condition $F^{\mu\nu}_{\mathrm{in}} + F^{\mu\nu}_{\mathrm{out}} = 0$ in Maxwell's theory. These derived equations of motion differ from the empirically required ones,

$$ma_i^\mu = e_i \sum_{j(\neq i)} F^{\mu\nu}_{\mathrm{ret},j}(z_i^\sigma) v_{i,\nu} + \frac{e_i}{2} [F^{\mu\nu}_{\mathrm{ret},i}(z_i^\sigma) - F^{\mu\nu}_{\mathrm{adv},i}(z_i^\sigma)] v_{i,\nu} \quad , \tag{2.36}$$

not only by the replacement of half the retarded with half the advanced forces, but also by the missing radiation reaction $\mathcal{F}^\mu_{\mathrm{rad},i}$ (the asymmetric self-force). While the problem of a mass renormalization has disappeared, (2.35) seems to be in drastic conflict with reality. Moreover, it contains a complicated dynamical meshing of the future with the past that does not in any obvious way permit the formulation of an initial-value problem.

The two equations of motion, (2.35) and (2.36), differ precisely by a force that would result from the sum of the T-asymmetric fields of *all* particles, $F^{\mu\nu}_{\mathrm{rad,total}} = \sum_j \frac{1}{2}(F^{\mu\nu}_{\mathrm{ret},j} - F^{\mu\nu}_{\mathrm{adv},j})$. Since the retarded and advanced fields appearing in this expression possess identical sources, their difference solves the *homogeneous* Maxwell equations, and thus represents a *free* field. Therefore, it (the difference) may consistently be assumed to vanish for all times. As there are no retarded fields at the beginning of the universe, this would require $\sum_j F^{\mu\nu}_{\mathrm{adv},j}(\text{big bang}) = 0$ *as a global constraint to all future sources.*

If the condition $F^{\mu\nu}_{\mathrm{rad,total}} = 0$ were valid, the advanced effects of all charged matter in the universe would precisely double the retarded forces in (2.35), cancel the advanced ones, and imitate a self-interaction that is responsible for radiation damping. This is an example for the equivalence of apparently quite different dynamical representations of *deterministic* theories, such as causal or teleological, local or global ones.

Instead of referring to a cosmic initial condition, Wheeler and Feynman (1945) tried to derive the vanishing sum of asymmetric fields from the assumption that the total charged matter in the universe behaves as an 'absorber' in a sense that is very different from Sect. 2.2. They assumed that the T-symmetric field $\bar{F}$ resulting from *all* particles in the universe, and according to (2.35) determining the force on an additional 'test particle', vanishes *for statistical reasons* (destructive interference) in a presumed empty space surrounding all matter of this 'island universe'. This assumption, in the form

$$\sum_j \bar{F}_j^{\mu\nu} := \sum_j \frac{1}{2} [F^{\mu\nu}_{\mathrm{ret},j} + F^{\mu\nu}_{\mathrm{adv},j}] \;\Rightarrow\; 0 \quad , \tag{2.37}$$

constitutes their cosmic *absorber condition.* Since the retarded and advanced fields vanish by definition in the asymptotic past or future, respectively, Wheeler and Feynman concluded by means of the homogeneous Maxwell equations that the total asymmetric field must vanish everywhere.

The consistency of this procedure is far from being obvious. A similar problem would arise for an expanding and recollapsing universe that were sandwiched between two *thermodynamically opposite* radiation eras (absorbers with opposite thermodynamical arrows of time) — see Sect. 5.3. As explained in Sect. 2.1, the compatibility of double-ended (two-time) boundary conditions is highly nontrivial, and similar to an eigenvalue problem. The consistency problem is particularly severe for a classical universe that remains optically transparent (Davies and Twamley 1993).

In contrast to the physical absorbers of Sect. 2.2, the *absorber condition* is T-symmetric. This has led to many misunderstandings. For example, rather than adding the vanishing antisymmetric term to (2.35), one might as well subtract it in order to obtain the time-reversed representation

$$ma_i^\mu = e_i \sum_{j(\neq i)} F^{\mu\nu}_{\mathrm{adv},j}(z_i^\sigma)v_{i,\nu} - \frac{e_i}{2}[F^{\mu\nu}_{\mathrm{ret},i}(z_i^\sigma) - F^{\mu\nu}_{\mathrm{adv},i}(z_i^\sigma)]v_{i,\nu} \quad . \tag{2.38}$$

Though being as correct as (2.36) under the absorber condition, (2.38) describes *advanced* actions and a radiation reaction that leads to reversed damping (exponential growth).

Therefore, Wheeler and Feynman's absorber condition cannot explain the *observed* radiation arrow. Neither (2.36) nor (2.38) would describe the asymmetric empirical situation. Evidently, in general only a limited number of 'obvious sources' contribute noticeably to the *retarded* sum (2.36). Otherwise, retardation would have never been recognized. This means that the *retarded* contribution of *all 'other'* sources (those which form the 'universal absorber') must interfere destructively,

$$\sum_{i\,\in\,\mathrm{absorbers}} F^{\mu\nu}_{\mathrm{ret},i} \approx 0 \quad \text{'inside' universal absorber} \tag{2.39}$$

(see Fig. 2.7). This is possible (except for the remaining thermal radiation) if the absorber particles approach thermal equilibrium by means of collisions *after* having been accelerated by the retarded field. Therefore, one can *not* expect

$$\sum_{i\,\in\,\mathrm{absorbers}} F^{\mu\nu}_{\mathrm{adv},i} \approx 0 \quad \text{'inside' universal absorber} \tag{2.39$'$}$$

to hold in a symmetric way.

In order to justify the applicability of (2.36) (in contrast to that of (2.38)), one still needs the *asymmetric* condition that has been derived in Sect. 2.2 from the thermodynamical arrow of time under certain cosmological assumptions. Any motion of absorber particles is dissipated as heat *after* it has been induced. While in field theory the field may be regarded

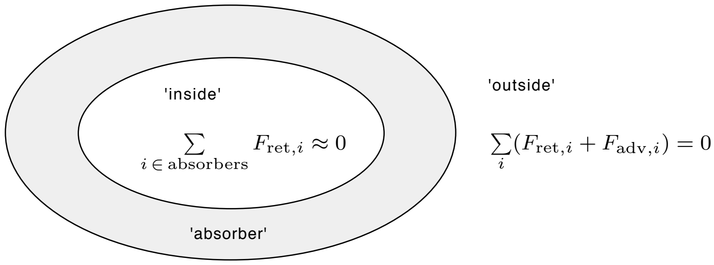


Fig. 2.7. T-symmetric ('outside') and T-asymmetric ('inside') absorber conditions of a model universe with an action-at-a-distance electrodynamics

as 'matter' with its own temperature, action-at-a-distance theory ascribes thermodynamical properties only to the charges. In the former description, the relation between electromagnetic and thermodynamical arrows is just an example of the *universality* of the thermodynamical arrow (see Sect. 3.1.2). As will be argued in Chap. 5, the observed arrows may have emerged from an initial homogeneous thermal equilibrium (the radiation era) by means of inhomogeneous gravitational contraction.

With these remarks, I also hope to put to rest objections raised by Popper (1956) against the thermodynamical foundation of the radiation arrow — see also Price (1996). The only 'unusual' aspect of electromagnetic fields (when regarded as matter) is their weak coupling, which may greatly delay their thermalization in the absence of absorbers. It is this property that allows light and radio waves to serve as an excellent information medium.

Therefore, the T-symmetric 'absorber condition' (2.37) would explain the equivalence of various forms of electrodynamics, but *not* the time arrow of radiation. In action-at-a-distance theory there is no free radiation, while the radiation reaction is the effect of advanced forces 'caused' by future absorbers. If the universe remained transparent for all times in some direction and at some frequency, an appropriately beamed emitter of electromagnetic radiation should not draw any power according to the absorber theory. As it always seems to do so (Partridge 1973), the absorber theory may be indeed be ruled out in a forever expanding universe. If similarly applied to *gravitational* fields, it would be in conflict with the observed energy loss of double pulsars.

General literature: Wheeler and Feynman 1945, 1949, Hoyle and Narlikar 1995

3. The Thermodynamical Arrow of Time

– Prelim. 4th edtn. (Feb 01): www.time-direction.de

The thermodynamical arrow of time is characterized by the increase of entropy according to the Second Law, postulated by Clausius in 1865 on the basis of Carnot's cycle. It can be written in the form

$$\frac{dS}{dt} = \left\{\frac{dS}{dt}\right\}_{\text{ext}} + \left\{\frac{dS}{dt}\right\}_{\text{int}} \quad ,$$

$$\text{with} \quad dS_{\text{ext}} = \frac{dQ}{T} \quad \text{and} \quad \left\{\frac{dS}{dt}\right\}_{\text{int}} \geq 0 \quad , \tag{3.1}$$

where S is the phenomenological entropy of a bounded system, while dQ is the total inward flow of heat through its boundary (see also the local form (3.1') on page 57). The heat flux is usually not written as a derivative dQ/dt, since its integral would not represent a 'function of state' $Q(t)$, while it does, of course, define the time-integrated flux in the actual process. The first term of (3.1) vanishes by definition for 'thermodynamically closed' systems. Since the whole universe is defined as an absolutely closed system (even if infinite), its total entropy, or the mean entropy of co-expanding volume elements, should according to this law evolve towards its maximum — the so-called *Wärmetod* (heat death). The phenomenological thermodynamical concepts used in (3.1) or (3.1'), in partucular the temperature, apply only in situations of partial (local) equilibrium.

Statistical physics is expected to provide a foundation of phenomenological thermodynamics — including the Second Law. While in principle *all* physical concepts are phenomenological, this term is here used to emphasize the existence of a (conceivably complete) 'microscopic' description that awaits the application of statistical methods.

Statistical considerations are indeed essential for the understanding of the mechanisms of 'irreversible' processes (those with $\{dS/dt\}_{\text{int}} > 0$). However, statistics as an aspect of counting has nothing a priori to do with dynamics. Therefore, it cannot by itself single out a direction in time. The statistical justification of 'irreversible' processes requires *additional* assumptions, which often remain unnoticed, since they appear 'natural' to our prejudiced causal way of thinking. They cannot be derived by means of statistical arguments from the symmetric dynamical laws, and therefore have to be analyzed in order to reveal the true origin of the thermodynamical arrow.

The attempt to explain this fundamental asymmetry on the basis of the 'historical nature' of the world, that is, from the assumption that the past be 'fixed' (and therefore neither requires nor allows statistical retrodiction) would clearly represent a circular argument. This view must itself be rooted in the time asymmetry of the physical world. The existence of knowledge or information about the past corresponds to a time-asymmetric relation between documents and their objects, equivalent to the asymmetric 'causal' relation between retarded electromagnetic fields and 'their' sources (Sect. 2.1). For example, light contains information about objects in the more or less recent past, as we can *see* when we open our eyes. Thus, documents represent an asymmetry in the physical world, and do not simply reflect the way the boundary conditions are posed (as initial or final ones). On the other hand, a certain (objectivizable) observer-relatedness of statistical concepts seems to be required for these concepts to be justified.

In a statistical description, 'irreversible' processes are of the form

$$\text{improbable state} \underset{t}{\rightarrow} \text{probable state,}$$

where the probability ratio is usually a huge number. These probabilities are defined by the size of certain *sets* of elementary states (called 'representative ensembles' by Tolman 1938) which contain the *real* state of the considered system (a point in its configuration space) as a member. This measure of probability changes while a state moves along its trajectory through different sets. If the representative ensembles are *operationally* defined, for example by means of a preparation procedure, they are often called macroscopic or thermodynamical *states.* Their physical justification is a major objective of the theory.

Since the above *processes* possess microscopic realizations, they are more probable than those of the kind

$$\text{improbable state} \underset{t}{\rightarrow} \text{improbable state.}$$

Their overwhelming occurrence in nature can thus be statistically derived from the presumption of an improbable *initial* state. In the operational approach, such an assumption is simply taken for granted. In a cosmological context, it has occasionally been called the *Kaltgeburt* (cold birth) of the universe, although a low temperature need not be its essential aspect (see Sect. 5.3). However, such an assumption appears quite unreasonable precisely for statistical reasons, since (1) there are just as many processes of the type

$$\text{probable state} \underset{t}{\rightarrow} \text{improbable state,}$$

and (2) far more of the kind

$$\text{probable state} \underset{t}{\rightarrow} \text{probable state.}$$

The latter describe equilibrium. Hence, for statistical reasons we should expect to *find* the world in the situation of a heat death.

The first of these two objections is Loschmidt's *Umkehreinwand* (*reversibility objection*). It is based on the property that each trajectory has precisely one time-reversed counterpart.[1] If, for example, $z(t) \equiv \{q_i(t), p_i(t)\}_{i=1\ldots 3N}$ describes a trajectory in $6N$-dimensional phase space (*Γ-space*) according to the Hamiltonian equations, then the time-reversed trajectory, $z_{\mathrm{T}}(-t) \equiv \{q_i(-t), -p_i(-t)\}$, is also a solution of the equations of motion (cf. the Introduction). If the entropy S of a state z can be defined as a function of state, $S = F(z)$, with $F(z) = F(z_{\mathrm{T}})$, then Loschmidt's objection means that for every solution with $dS/dt > 0$ there is precisely a corresponding one with $dS/dt < 0$. In statistical theories, $F(z)$ is defined as a monotonic function (conveniently the logarithm) of the measure of the mentioned set of states to which z belongs. The property $F(z) = F(z_{\mathrm{T}})$ is then a consequence of the fact that the transformation $z \rightarrow z_{\mathrm{T}}$ represents a symmetry, while the stronger objection (2) of above means that there are far more solutions with $dS/dt \approx 0$.

In order to describe the thermodynamical arrow of time statistically, one either has to derive the *Kaltgeburt* in some form from a new and fundamental assumption, or simply to postulate it. The Second Law is by no means *incompatible* with deterministic or T-symmetric dynamical laws; it is just extremely improbable, and therefore in conflict with unbiased statistical reasoning. The widespread 'double standard' of readily accepting improbable initial conditions while rejecting similar final ones has been duly criticized by Price (1996).

Another historically relevant objection (Zermelo's *Wiederkehreinwand* or *recurrence objection*) is mathematically correct, but does not apply to a sufficiently young universe (as ours seems to be). It can be based on a theorem by Poincaré which states that every bounded mechanical system will return as close as one wishes to its initial state within a sufficiently large time. The entropy of closed systems would therefore have to return to its former values, provided only the function $F(z)$ is continuous. This is a special case of the *quasi-ergodic theorem* which asserts that every system will come arbitrarily close to any point on the hypersurface of fixed energy (and possibly of fixed other analytical constants of the motion) within finite time.

While these theorems are mathematically correct, the *recurrence objection* fails to apply, since the age of our universe is much smaller than the Poincaré times of a gas consisting of as few as ten or twenty particles. Their recurrence to the vicinity of their initial states (or their coming close to any other similarly specific state) can therefore be practically excluded. (The formal objection that Zermelo's argument does not apply to *ensembles* — which

[1] Often T (or CPT) symmetry of the dynamics is used in this argument. This has misled to the by no means justified expectation that the difficulties in deriving the Second Law may be overcome by dropping this symmetry. However, as already pointed out in the Introduction, the crucial point in Loschmidt's argument is *determinism*, which is often reflected by the possibility of compensating for T asymmetry by another symmetry violation (see also Sect. 3.4).

are used to describe macroscopic systems — does in turn not apply to reality, since an ensemble does not represent the *real* physical state.) Many 'foundations' of irreversible thermodynamics are based on a formal idealization that leads to *infinite* Poincaré recurrence times (for example by using the 'thermodynamical limit' of infinite particle number). They are quite irrelevant in our universe of finite age, and they would not invalidate the reversibility objection (or the equilibrium expectation). Rather, they illustrate that some kind of *Kaltgeburt* is required in order to derive the thermodynamical arrow.

The theory of thermodynamically irreversible processes thus has to address two main problems:

1. The investigation of realistic *mechanisms* which describe the dynamical evolution away from certain (presumed) improbable initial states. This is usually achieved in the form of *master equations*, which mimic a *law-like* T-asymmetry — analogous to Ritz's retarded action-at-a-distance in electrodynamics. Their ensemble dynamics is equivalent to a stochastic process for individual states (applicable in the 'forward' direction of time). These asymmetric dynamical equations may then even describe the emergence of *order* (Sect. 3.4).
2. The precise characterization of the required improbable initial states. This leads again to the quest for an appropriate *cosmic* initial condition, similar to $A^{\mu}_{\mathrm{in}} = 0$ for the radiation arrow (cf. Sects. 2.2 and 5.3).

3.1 The Derivation of Classical Master Equations

Statistical physics is concerned with systems consisting of a huge number of microscopic constituents. These are known to obey quantum mechanics. However, this theory has always been haunted by interpretational problems regarding the stochastic nature of 'quantum events'. This probabilistic nature of quantum theory is in general understood as representing a *fundamental* (albeit not precisely defined) irreversible dynamical element, that would have to be taken into account in a microscopic description of thermodynamical systems. In contrast, classical mechanics is deterministic and well defined. Therefore, classical statistical mechanics will be discussed in this chapter for conceptual consistency and later comparison with quantum statistical mechanics — even though it is based on an incorrect microscopic theory. Most thermodynamical properties of a gas, for example, can in fact be modelled by a system of interacting classical mass points (see (4.20)). While the present section follows historical routes, a more general and systematic formalism will be presented in Sect. 3.2.

3.1.1 μ-Space Dynamics and Boltzmann's H-Theorem

The *state* of a mechanical system of N classical (distinguishable) particles can either be represented by one point in its $6N$-dimensional phase space ('Γ-space'), or by N ordered points in six-dimensional 'μ-space' (the single-particle phase space). If the particles are *not* distinguished from one another, these N points form a discrete *distribution* in μ-space. It is equivalent to the ensemble of $N!$ points in Γ-space, which result from all particle permutations. Because of the large number of particles forming macroscopic systems (of the order 10^{23}), Boltzmann (1866, 1896) used continuous (smoothed) distributions (or *phase space densities*) $\varrho_\mu(\boldsymbol{p}, \boldsymbol{q})$. This harmless looking difference will turn out to have important consequences. Two types of argument have been proposed for its justification:

1. The formal *thermodynamical limit* $N \to \infty$ represents an idealization for which the mathematically disturbing Poincaré recurrence times become infinite. Mathematical proofs may then appear to be rigorous, where in fact they are approximations. While often convenient, this procedure may conceal physically important aspects, in particular when the above limit is interchanged with others.
2. Slightly 'uncertain' positions and momenta, corresponding to small volume elements in Γ-space, $\Delta V_\Gamma = (\Delta V_\mu)^N$ (that is, infinite *ensembles* of states), may also lead to smooth distributions, since $N!$ volume elements easily overlap even for a diluted gas according to $N!\Delta V_\Gamma \approx (N\Delta V_\mu)^N$. Although uncertainties slightly larger than distances between the particles are sufficient for the smoothing, they have drastic dynamical consequences in the generic situation of interacting particles. However, these uncertainties *cannot* be based on the quantum mechanical uncertainty relations with their corresponding phase space cells of size h^{3N}, since equivalent problems reappear in quantum theory if phase space points are consistently replaced with wave functions (see Sect. 4.1.1).

The time dependence of an individual point $\{p_i(t), q_i(t)\}$ in Γ-space (with $i = 1 \ldots 3N$), described by Hamilton's equations, is equivalent to the simultaneous time dependence of all N points in μ-space. Therefore, the time dependence of an ensemble in Γ-space (represented in general by a distribution ϱ_Γ) determines that of the corresponding smooth density ϱ_μ. In contrast to the dynamics in Γ-space (Sect. 3.1.2), however, this dynamics is not 'autonomous': the time derivative of a non-singular density ϱ_μ is *not* determined by ϱ_μ. The reason is that ϱ_Γ cannot be recovered from ϱ_μ in order to determine the latter's time derivative from that of the former. The mapping of Γ-space distributions on μ-space distributions cannot be uniquely inverted, as it destroys information about correlations between the particles (see also Fig. 3.1 below and the subsequent discussion). The smooth μ-space distribution may characterize a 'macroscopic state' in the sense mentioned in the introduction to this chapter. Therefore, the envisioned chain of computation

$$\varrho_\mu \to \varrho_\Gamma \stackrel{H}{\to} \frac{d\varrho_\Gamma}{dt} \to \frac{\partial \varrho_\mu}{\partial t} \quad , \tag{3.2}$$

which would be required to derive an autonomous dynamics for ϱ_μ, is broken at its first link. Boltzmann's attempt to bridge this gap by statistical arguments will turn out to be the source of the time direction asymmetry in his statistical mechanics, and similarly in other descriptions of irreversible processes. His procedure specifies a direction in time in a phenomenologically justified way, while it was apparently meant to represent an approximation rather than a modification of the T-symmetric Hamiltonian dynamics. The question is, under what circumstances it may be valid.

Boltzmann postulated a dynamical law of the form

$$\frac{\partial \varrho_\mu}{\partial t} = \left\{\frac{\partial \varrho_\mu}{\partial t}\right\}_{\text{free+ext}} + \left\{\frac{\partial \varrho_\mu}{\partial t}\right\}_{\text{collision}} \quad . \tag{3.3}$$

Its first term is defined to describe motion under external forces only. It can be written as a continuity equation in μ-space,

$$\begin{aligned} \left\{\frac{\partial \varrho_\mu}{\partial t}\right\}_{\text{free+ext}} &= -\mathrm{div}_\mu\, j_\mu := -\nabla_q \cdot (\dot{\boldsymbol{q}}\varrho_\mu) - \nabla_p \cdot (\dot{\boldsymbol{p}}\varrho_\mu) \\ &= -\nabla_q \cdot \left(\frac{\boldsymbol{p}}{m}\varrho_\mu\right) - \nabla_p \cdot (\boldsymbol{F}_{\text{ext}}\varrho_\mu) \quad , \end{aligned} \tag{3.4}$$

where j_μ is the 6-dimensional current density in μ-space. In the absence of particle interactions this equation describes the dynamics of the 'phase space fluid' exactly. It represents the *local* conservation of probability in μ-space according to the deterministic Hamiltonian equations, which hold separately for each particle in this case. Each point in μ-space (each single-particle state) moves continuously on its orbit, ruled by the external forces $\boldsymbol{F}_{\text{ext}}$, and thereby retaining its individual probability that is determined by the initial condition for ϱ_μ.

For the second (non-trivial) term Boltzmann proposed his *Stoßzahlansatz* (collision equation), which will here be formulated under the following simplifying assumptions:

1) $\boldsymbol{F}_{\text{ext}} = 0$ 'no external forces'
2) $\varrho_\mu(\boldsymbol{p}, \boldsymbol{q}, t) = \varrho_\mu(\boldsymbol{p}, t)$ 'homogeneous distribution'

The second condition is dynamically consistent for translation-invariant interactions. From these assumptions one obtains $\{\partial\varrho_\mu/\partial t\}_{\text{free+ext}} = 0$. The *Stoßzahlansatz* is then written in the plausible form

$$\frac{\partial \varrho_\mu}{\partial t} = \left\{\frac{\partial \varrho_\mu}{\partial t}\right\}_{\text{collision}} = \text{gains} - \text{losses} \quad , \tag{3.5}$$

that is, as a *balance equation.* Its two terms on the RHS are explicitly defined by means of transition rates $w(\boldsymbol{p}_1\boldsymbol{p}_2;\boldsymbol{p}'_1\boldsymbol{p}'_2)$ for particle pairs from $\boldsymbol{p}'_1\boldsymbol{p}'_2$ to $\boldsymbol{p}_1\boldsymbol{p}_2$. They are usually (in a low density approximation) assumed to be determined by the two-particle scattering cross sections, and have to satisfy conservation laws. Because of this description in terms of probabilities for a discontinuous change of momenta, the collisions cannot be represented by a *local* conservation of probability in μ-space, and therefore do not assume the form of a continuity equation (3.4).

The *Stoßzahlansatz* (3.5) reads explicitly

$$\frac{\partial\varrho_\mu(\boldsymbol{p}_1,t)}{\partial t} = \int [w(\boldsymbol{p}_1\boldsymbol{p}_2;\boldsymbol{p}'_1\boldsymbol{p}'_2)\varrho_\mu(\boldsymbol{p}'_1,t)\varrho_\mu(\boldsymbol{p}'_2,t) - w(\boldsymbol{p}'_1\boldsymbol{p}'_2;\boldsymbol{p}_1\boldsymbol{p}_2)\varrho_\mu(\boldsymbol{p}_1,t)\varrho_\mu(\boldsymbol{p}_2,t)]\, d^3p_2\, d^3p'_1\, d^3p'_2 \ . \tag{3.6}$$

It is the prototype of a *master equation* as an irreversible balance equation based on probabilistic transition rates. How can it be irreversible in spite of representing an approximation to the reversible Hamiltonian dynamics? This can clearly not be true in general, but only for special solutions which select an arrow of time. These solutions cannot even be probable.

For further simplification, invariance of the transition rates under *collision inversion*

$$w(\boldsymbol{p}_1\boldsymbol{p}_2;\boldsymbol{p}'_1\boldsymbol{p}'_2) = w(\boldsymbol{p}'_1\boldsymbol{p}'_2;\boldsymbol{p}_1\boldsymbol{p}_2) \tag{3.7}$$

will be assumed. It may be derived from invariance under space reflection *and* time reversal, although these two symmetries do not necessarily have to be separately valid. The *Stoßzahlansatz* then assumes the form

$$\frac{\partial\varrho_\mu(\boldsymbol{p}_1,t)}{\partial t} = \int [w(\boldsymbol{p}_1\boldsymbol{p}_2;\boldsymbol{p}'_1\boldsymbol{p}'_2)[\varrho_\mu(\boldsymbol{p}'_1,t)\varrho_\mu(\boldsymbol{p}'_2,t) - \varrho_\mu(\boldsymbol{p}_1,t)\varrho_\mu(\boldsymbol{p}_2,t)]\, d^3p_2\, d^3p'_1\, d^3p'_2 \quad . \tag{3.8}$$

In order to demonstrate the irreversibility described by the *Stoßzahlansatz*, it is useful to consider *Boltzmann's H-functional*

$$H[\varrho_\mu] := \int \varrho_\mu(\boldsymbol{p},\boldsymbol{q},t)\ln\varrho_\mu(\boldsymbol{p},\boldsymbol{q},t)\, d^3p\, d^3q = N\overline{\ln\varrho_\mu} \quad , \tag{3.9}$$

proportional to the mean logarithm of probability. The mean $\bar{f}$ of a function $f(\boldsymbol{p},\boldsymbol{q})$ is here defined as $\bar{f} := \int f(\boldsymbol{p},\boldsymbol{q})\varrho_\mu(\boldsymbol{p},\boldsymbol{q})\, d^3p\, d^3q/N$, in accordance with the normalization $\int \varrho_\mu(\boldsymbol{p},\boldsymbol{q})\, d^3p\, d^3q = N$. Because of this fixed normalization, the H-functional is large for narrow distributions, but small for wide ones. A discrete point (or δ-distribution), for example, would lead to $H[\varrho_\mu] = \infty$, while a distribution being constant on a volume of size V_μ, $\varrho_\mu = N/V_\mu$, gives $H[\varrho_\mu] = N(\ln N - \ln V_\mu)$. ($H$ is defined up to an additive constant, depending on the choice of a unit volume element in phase space.)

One may now derive Boltzmann's *H-theorem*,

$$\frac{dH[\varrho_\mu]}{dt} \leq 0 \quad , \tag{3.10}$$

by differentiating $H[\varrho_\mu]$ with respect to time, while using the collision equation in the form (3.8):

$$\begin{aligned}\frac{dH[\varrho_\mu]}{dt} &= V \int \frac{\partial \varrho_\mu(\boldsymbol{p}_1,t)}{\partial t}[\ln \varrho_\mu(\boldsymbol{p}_1,t)+1]\, d^3p_1 \\ &= V \int w(\boldsymbol{p}_1\boldsymbol{p}_2;\boldsymbol{p}_1'\boldsymbol{p}_2')[\varrho_\mu(\boldsymbol{p}_1',t)\varrho_\mu(\boldsymbol{p}_2',t) - \varrho_\mu(\boldsymbol{p}_1,t)\varrho_\mu(\boldsymbol{p}_2,t)] \\ &\quad \times [\ln \varrho_\mu(\boldsymbol{p}_1,t)+1]\, d^3p_1\, d^3p_2\, d^3p_1'\, d^3p_2' \quad .\end{aligned} \tag{3.11}$$

The last expression may be conveniently reformulated by using the symmetries under collision inversion (3.7), and under particle permutation, $w(\boldsymbol{p}_1\boldsymbol{p}_2;\boldsymbol{p}_1'\boldsymbol{p}_2') = w(\boldsymbol{p}_2\boldsymbol{p}_1;\boldsymbol{p}_2'\boldsymbol{p}_1')$. (Otherwise this combined symmetry would be required to hold for short *chains* of collisions, at least.) Rewriting the integral as a sum of the four different permutations of the integration variables, one obtains

$$\begin{aligned}\frac{dH[\varrho_\mu]}{dt} &= \frac{V}{4} \int w(\boldsymbol{p}_1\boldsymbol{p}_2;\boldsymbol{p}_1'\boldsymbol{p}_2')[\varrho_\mu(\boldsymbol{p}_1',t)\varrho_\mu(\boldsymbol{p}_2',t) - \varrho_\mu(\boldsymbol{p}_1,t)\varrho_\mu(\boldsymbol{p}_2,t)] \\ &\quad \times \{\ln[\varrho_\mu(\boldsymbol{p}_1,t)\varrho_\mu(\boldsymbol{p}_2,t)] \\ &\quad - \ln[\varrho_\mu(\boldsymbol{p}_1',t)\varrho_\mu(\boldsymbol{p}_2',t)]\} d^3p_1\, d^3p_2\, d^3p_1'\, d^3p_2' \leq 0 \quad .\end{aligned} \tag{3.12}$$

The integrand is now manifestly non-positive, since the logarithm is a monotonically increasing function of its argument. This concludes the proof of (3.10).

In order to recognize the connection of the H-functional with entropy, one may consider the *Maxwell distribution* ϱ_{M},

$$\varrho_{\mathrm{M}}(\boldsymbol{p}) := \frac{N}{V}\frac{\exp(-p^2/2mkT)}{\sqrt{(2\pi mkT)^3}} \quad . \tag{3.13}$$

Its H-functional $H[\varrho_{\mathrm{M}}]$ has two important properties:

1. It represents a *minimum for given energy*, $E = \int \varrho_\mu(\boldsymbol{p})[p^2/2m]\, d^3p \approx \sum_i \boldsymbol{p}_i^2/2m$. A proof will be given in a somewhat more general form in Sect. 3.1.2. (Unconstrained statistical reasoning would predict infinite energy, since the phase space volume grows non-relativistically with its $3N/2$-th power.) ϱ_{M} must therefore represent an equilibrium distribution under the *Stoßzahlansatz* if the latter is assumed to conserve energy.
2. One obtains explicitly

$$\begin{aligned}H[\varrho_{\mathrm{M}}] &= V \int \varrho_{\mathrm{M}}(\boldsymbol{p}) \ln \varrho_{\mathrm{M}}(\boldsymbol{p})\, d^3p \\ &= -N\left(\ln\frac{V}{N} + \frac{3}{2}\ln T + \text{constant}\right) \quad .\end{aligned} \tag{3.14}$$

If one compares this expression with the entropy of a mole of a monatomic ideal gas according to phenomenological thermodynamics,

$$S_{ideal}(V,T) = R\left(\ln V + \frac{3}{2}\ln T\right) + \text{constant} \quad , \tag{3.15}$$

one recognizes (up to the phenomenologically undefined, possibly N-dependent constant)

$$S_{ideal} = -kH[\varrho_{\mathrm{M}}] =: S_\mu[\varrho_{\mathrm{M}}] \quad , \tag{3.16}$$

where $k = R/N$.

If the entropy is thus interpreted as representing a measure of the width of the particle distribution in μ-space, the *Stoßzahlansatz* describes successfully the evolution towards a Maxwell distribution with its ensemble parameter T that is recognized as the *temperature* (which determines the mean energy in this ensemble).

This success seems to be the origin of the 'myth' of the statistical foundation of the thermodynamical arrow of time. However, statistical arguments can neither explain why the *Stoßzahlansatz* is a good approximation in one and only one direction of time, nor tell us whether S_μ is *always* an appropriate definition of entropy. It will indeed turn out to be insufficient when correlations between particles become essential, as is, for example, the case for real gases or solid bodies. Taking them into account requires more general concepts, which were first proposed by Gibbs. His approach will also allow us to formulate the exact ensemble dynamics in Γ-space, although it can still *not* explain the origin of the thermodynamical arrow of time.

3.1.2 Γ-Space Dynamics and Gibbs' Entropy

In the preceding section, Boltzmann's smooth phase space density ϱ_μ was identified as representing small uncertainties in positions and momenta: it is equivalent to an infinite number (a continuum) of states, described by a volume element ΔV_Γ (or any other Γ-space distribution) rather than a 'real' point. An objective state would have to be represented by a point (or a δ-distribution) in Γ-space, or by a sum over N δ-functions in μ-space (if particles are not distinguished). It would, therefore, lead to an infinite value of Boltzmann's H-functional.

However, the *finite* value of S_μ, derived from the smooth μ-space distribution, is *not* merely a measure of this *arbitrary* smoothing procedure. If points are replaced with small (but not too small) volume elements in phase space, this leads to a smooth distribution ϱ_μ whose width reflects that of the discrete distribution. Therefore, S_μ does characterize the real N-particle state. The formal 'renormalization of entropy', which is part of this smoothing procedure, adds an infinite positive contribution to the infinite negative entropy corresponding to a point — such that the remaining finite sum is physically meaningful. The 'representative ensemble' resulting from the smoothing defines a finite measure of probability in the sense of the introduction to this chapter for each point in Γ-space. It depends only negligibly on the precise smoothing conditions, provided the discrete μ-space distribution is already smooth in the mean.

The ensemble concept introduced by Gibbs (1902) is different from Boltzmann's from the outset. He considered probability densities $\varrho_\Gamma(p,q)$ (with $\int \varrho_\Gamma(p,q)\,dp\,dq = 1$ — from now on writing $p := p_1 \dots p_{3N}$, $q := q_1 \dots q_{3N}$ and $dp\,dq := d^{3N}p\,d^{3N}q$ for short) which are meant to describe incomplete *information* about microscopic degrees of freedom, for example characterizing a macroscopic (incomplete) preparation procedure. Boltzmann's H-functional is then replaced by Gibbs' formally analogous *extension in phase* η,

$$\eta[\varrho_\Gamma] := \overline{\ln \varrho_\Gamma} = \int \varrho_\Gamma(p,q) \ln \varrho_\Gamma(p,q)\,dp\,dq \quad , \tag{3.17}$$

that leads to an *ensemble entropy* $S_\Gamma := -k\eta[\varrho_\Gamma]$. For a probability density being constant on a phase space volume element of size ΔV_Γ (while vanishing otherwise) one has $\eta[\varrho_\Gamma] = -\ln \Delta V_\Gamma$. The entropy $S_\Gamma = k \ln \Delta V_\Gamma$ is a logarithmic measure of the size of this volume element that is uniformly occupied by the ensemble, but does not characterize a real many-particle state (a point), as Boltzmann's entropy is supposed to do.

For a smooth distribution of statistically independent particles, $\varrho_\Gamma = \prod_{i=1}^N [\varrho_\mu(\boldsymbol{p}_i, \boldsymbol{q}_i)/N]$, one obtains nonetheless

$$\begin{aligned}\eta[\varrho_\Gamma] &= \sum_{i=1}^N \int [\varrho_\mu(\boldsymbol{p}_i,\boldsymbol{q}_i)/N] \ln[\varrho_\mu(\boldsymbol{p}_i,\boldsymbol{q}_i)/N]\, d^3p_i\, d^3q_i \\ &= \int \varrho_\mu(\boldsymbol{p},\boldsymbol{q})[\ln \varrho_\mu(\boldsymbol{p},\boldsymbol{q}) - \ln N]\, d^3p\, d^3q = H[\varrho_\mu] - N \ln N \quad .\end{aligned} \tag{3.18}$$

In this important special case one thus recovers Boltzmann's statistical entropy S_μ (with all its advantages) — except for the term $kN \ln N \approx k \ln N!$, that can be interpreted as the *mixing entropy* of the gas with itself. Although merely an additive constant in systems with fixed particle number, it would lead to observable contributions at variance with experiments in situations where the particle number may vary dynamically. Large particle numbers would then acquire far too large statistical weights. This physically inappropriate mixing entropy (that led to confusion in Gibbs' approach although it had already been known as a problem to Maxwell) disappears when Gibbs' ensemble concept is applied to quantum states defined in the occupation number representation.[2] After borrowing this result from quantum theory,

[2] The popular argument that this term has to be dropped because of the indistinguishability of particles is certainly wrong, since kinematically different (even though op-

one can therefore justify Boltzmann's entropy in terms of ensemble entropy for statistically uncorrelated particles.

Furthermore, S_Γ is maximized under the constraint of fixed mean energy, $\bar{E} = \int H(p,q)\varrho_\Gamma(p,q)\,dp\,dq$, by the *canonical* (or Gibbs') distribution $\varrho_{\mathrm{can}} := Z^{-1}\exp(-H(p,q)/kT)$. The latter can be derived from a variational procedure with the additional constraint of fixed normalization of probability, $\int \varrho_\Gamma(p,q) = 1$, that is, from

$$\begin{aligned}\delta\{\eta[\varrho_\Gamma] + \alpha \int \varrho_\Gamma(p,q)\,dp\,dq + \beta \int H(p,q)\varrho_\Gamma(p,q)\,dp\,dq\} \\ = \int \{\ln \varrho_\Gamma(p,q) + (\alpha+1) + \beta H(p,q)\}\delta\varrho_\Gamma(p,q)\,dp\,dq = 0\ ,\end{aligned} \tag{3.19}$$

with Lagrange parameters α and β. Its solution is

$$\varrho_{\mathrm{can}} = \exp\{-[\beta H(p,q) - \alpha - 1]\} =: Z^{-1}\exp\{-\beta H(p,q)\} \quad , \tag{3.20}$$

and one recognizes $\beta = 1/kT$ and the *partition function (sum over states)* $Z := \int \mathrm{e}^{-\beta H(p,q)}\,dp\,dq = \mathrm{e}^{-\alpha-1}$. By using the *ansatz* $\varrho = \mathrm{e}^{\chi+\Delta\chi}$ with $\mathrm{e}^{\chi} := \varrho_{\mathrm{can}}$, an arbitrary (not necessarily small) variation $\Delta\chi(p,q)$, the above constraints, and the general inequality $\Delta\chi\mathrm{e}^{\Delta\chi} \geq \Delta\chi$, one may even show that the canonical distribution represents an *absolute* maximum of this entropy. In statistical thermodynamics (and in contrast to phenomenological thermodynamics) entropy is thus a more fundamental concept than temperature, which applies only to special (canonical or equivalent) probability distributions, while a formal entropy is defined for *all* ensembles.

One can similarly show that S_Γ is maximized by the micro-canonical ensemble $\varrho_{\mathrm{micro}} \equiv \delta(E - H(p,q))$ if constrained by the condition of *fixed energy*, $H(p,q) = E$. Although essentially equivalent for most applications, the canonical and the microcanonical distribution characterize two different situations: systems with and without energy exchange with a heat bath.

For non-interacting particles, $H = \sum_i[\boldsymbol{p}_i^2/2m + V(\boldsymbol{q}_i)]$, one obtains a factorizing canonical distribution $\varrho_\Gamma(p,q) = \prod_i[\varrho_\mu(\boldsymbol{p}_i,\boldsymbol{q}_i)/N]$ (as considered in (3.18)) with a μ-space distribution given by $\varrho_\mu(\boldsymbol{p},\boldsymbol{q}) \propto N\exp\{-[\boldsymbol{p}^2/2m + V(\boldsymbol{q})]/kT\}$. This is a Maxwell distribution multiplied by the barometric formula. However, the essential advantage of the canonical Γ-space distribution

erationally indistinguishable) states would have to be counted individually for statistical purposes. Different classical states related by particle permutation would dynamically retain their individual probabilities. The use of μ-space distributions, such as in Boltzmann's statistical mechanics, is similarly unjustified from a classical point of view, unless the probabilities were multiplied by the *empirically excluded* weight factors $N!$ again. Indistinguishability and identity (assumed for 'quantum particles') are different concepts. The absence of mixing entropy from systems with variable particle number can only be understood in terms of quantum mechanical concepts (see (4.20)). Even the difference $N\ln N - \ln N! \approx N - \ln N$, usually neglected in this argument, can be *physically* understood, as it counts those particle number fluctuations which occur in open systems described by a *grand* canonical distribution with a chemical potential (see page 69).

(3.20) over Boltzmann's is its ability to describe equilibrium correlations between particles. This has been demonstrated in particular by the cluster expansion of Ursell and Mayer (see Mayer and Mayer 1940), in more recent terminology called an expansion by N-point functions, and technically a predecessor of Feynman graphs. The distribution (3.20) must, however, not include macroscopic degrees of freedom (such as the shape of a solid body — see also Sect. 3.3.1). In the case of a rotationally symmetric Hamiltonian, for example, the solid body in thermodynamical equilibrium would otherwise have to be physically characterized by a symmetric distribution of all its orientations in space rather than by a definite orientation. Similarly, its center of mass would always have to be expected close to the minimum of an external potential (see also Fröhlich 1973). However, these macroscopic variables are *dynamically robust* rather than being ergodic. When calculating an appropriate representative ensenble according to (3.19), one has to impose the additional constraint of 'given' values for them (see also Sect. 3.3.1).

Gibbs' extension in phase η thus appears superior to Boltzmann's H-functional (3.9). Unfortunately, the corresponding *ensemble entropy* S_Γ has two (related) defects, which render it entirely unacceptable for representing physical entropy: (1) in blunt contrast to the Second Law it remains constant under exact (Hamiltonian) dynamics, and (2) it is obviously not a local or additive concept (that would define an entropy *density* in space).

The exact ensemble dynamics may be formulated in analogy to (3.4) by using the $6N$-dimensional continuity equation,

$$\frac{\partial \varrho_\Gamma}{\partial t} + \mathrm{div}_\Gamma(\varrho_\Gamma \boldsymbol{v}_\Gamma) = 0 \quad . \tag{3.21}$$

It describes the conservation of probability for each volume element as it moves through Γ-space as a bundle of trajectories. The $6N$-dimensional velocity $\boldsymbol{v}_\Gamma$ may be substituted using the Hamiltonian equations,

$$\begin{aligned} \boldsymbol{v}_\Gamma &\equiv (\dot{p}_1, \ldots, \dot{p}_{3N}, \dot{q}_1, \ldots, \dot{q}_{3N}) \\ &= \left(-\frac{\partial H}{\partial q_1}, \ldots, -\frac{\partial H}{\partial q_{3N}}, \frac{\partial H}{\partial p_1}, \ldots, \frac{\partial H}{\partial p_{3N}}\right) \quad . \end{aligned} \tag{3.22}$$

After rewriting (3.21) by employing the identity $\mathrm{div}_\Gamma(\varrho_\Gamma \boldsymbol{v}_\Gamma) = \varrho_\Gamma \mathrm{div}_\Gamma \boldsymbol{v}_\Gamma + \boldsymbol{v}_\Gamma \cdot \mathbf{grad}_\Gamma \varrho_\Gamma$, one may use the *Liouville theorem*,

$$\begin{aligned} \mathrm{div}_\Gamma \boldsymbol{v}_\Gamma = &-\frac{\partial^2 H}{\partial p_1 \partial q_1} - \cdots - \frac{\partial^2 H}{\partial p_{3N} \partial q_{3N}} \\ &+ \frac{\partial^2 H}{\partial q_1 \partial p_1} + \cdots + \frac{\partial^2 H}{\partial q_{3N} \partial p_{3N}} \equiv 0 \quad , \end{aligned} \tag{3.23}$$

which characterizes an 'incompressible flow' in Γ-space. In this way one obtains the *Liouville equation*,

$$\begin{aligned}\frac{\partial \varrho_\Gamma}{\partial t} &= -\boldsymbol{v}_\Gamma \cdot \mathbf{grad}_\Gamma \varrho_\Gamma \\ &= \sum_{n=1}^{3N} \left(\frac{\partial H}{\partial q_n} \frac{\partial \varrho_\Gamma}{\partial p_n} - \frac{\partial H}{\partial p_n} \frac{\partial \varrho_\Gamma}{\partial q_n} \right) = \{H, \varrho_\Gamma\} \quad , \end{aligned} \tag{3.24}$$

where $\{a, b\}$ defines the Poisson bracket. This equation represents the exact Hamiltonian dynamics for ensembles $\varrho_\Gamma(p, q, t)$ under the assumption of individually conserved probabilities.

From this picture of an incompressible flow one may expect that the ensemble entropy S_Γ (the measure of 'extension in phase') remains constant. This can indeed be formally confirmed by differentiating (3.17) with respect to time, inserting (3.24), and repeatedly integrating by parts:

$$\begin{aligned}\frac{dS_\Gamma}{dt} &= \int (\ln \varrho_\Gamma + 1) \dot{\varrho}_\Gamma \, dp \, dq \\ &= \int (\ln \varrho_\Gamma + 1) \sum_{n=1}^{3N} \left(\frac{\partial H}{\partial q_n} \frac{\partial \varrho_\Gamma}{\partial p_n} - \frac{\partial H}{\partial p_n} \frac{\partial \varrho_\Gamma}{\partial q_n} \right) dp \, dq \\ &= - \int \sum_{n=1}^{3N} \left(\frac{\partial H}{\partial q_n} \frac{\partial \ln \varrho_\Gamma}{\partial p_n} - \frac{\partial H}{\partial p_n} \frac{\partial \ln \varrho_\Gamma}{\partial q_n} \right) \varrho_\Gamma \, dp \, dq \\ &= - \int \sum_{n=1}^{3N} \left(\frac{\partial H}{\partial q_n} \frac{\partial \varrho_\Gamma}{\partial p_n} - \frac{\partial H}{\partial p_n} \frac{\partial \varrho_\Gamma}{\partial q_n} \right) dp \, dq = 0 \quad . \end{aligned} \tag{3.25}$$

A more instructive proof may be obtained by multiplying the Liouville equation (3.24) with i in order to cast the dynamics into a form that is analogous to the Schrödinger equation,

$$\mathrm{i} \frac{\partial \varrho_\Gamma}{\partial t} = \mathrm{i}\{H, \varrho_\Gamma\} =: \hat{L} \varrho_\Gamma \quad . \tag{3.26}$$

The operator $\hat{L}$ (acting on probability densities) is called the *Liouville operator*. In accordance with this analogy one may use the formal solution $\varrho_\Gamma(t) = \exp(-\mathrm{i}\hat{L}t)\varrho_\Gamma(0)$, valid if $\partial \hat{L}/\partial t = 0$ (cf. Prigogine 1962). The Liouville operator is Hermitean with respect to the inner product $\langle \varrho_\Gamma, \varrho_\Gamma{}' \rangle := \int \varrho_\Gamma^* \varrho_\Gamma' \, dp \, dq$ (that is, $\langle \varrho_\Gamma, \hat{L} \varrho_\Gamma' \rangle = \langle \hat{L} \varrho_\Gamma, \varrho_\Gamma' \rangle$), as can again be shown by partial integration. This means that the Liouville equation conserves these inner products. For example, for $\varrho_\Gamma' = \ln \varrho_\Gamma$ one has

$$\frac{d}{dt} \langle \varrho_\Gamma, \ln \varrho_\Gamma \rangle = \frac{d}{dt} \overline{\ln \varrho_\Gamma} = 0 \quad , \tag{3.27}$$

since the Liouville operator, when applied to a function $f(\varrho_\Gamma)$, satisfies the same Leibniz chain rule $\hat{L} f(\varrho_\Gamma) = (df/d\varrho_\Gamma) \hat{L} \varrho_\Gamma$ as the time derivative.

The norm corresponding to this inner product, $\|\varrho_\Gamma\|^2 = \langle\varrho_\Gamma, \varrho_\Gamma\rangle = \int \varrho_\Gamma^2 \, dp \, dq = \overline{\varrho_\Gamma}$, is then also dynamically invariant. It represents a linear measure of extension in phase (a *linear ensemble entropy*[3]), and thus has to be distinguished from the probability norm $\int \varrho_\Gamma \, dp \, dq = \overline{1} = 1$. The conservation of these measures under a Liouville equation confirms in turn that the Γ-space volume is an appropriate measure for non-countable sets of states (Ehrenfest and Ehrenfest 1911): the thus defined 'number' of states must not change under an appropriately defined determinism. A more fundamental justification of this measure can be derived from the conservation of probabilities of discrete *quantum* states (see Sect. 4.1).

The conservation of ensenble entropy that exact dynamics imposes is unacceptable for physical entropy. Thus Gibbs introduced a more subtle concept of entropy; one that he derived from his famous *ink drop analogy:* A drip of ink into a glass of water is assumed to behave as an incompressible fluid when the water is stirred. Although its volume has then to remain constant, the whole glass of water will appear homogeneously colored after a short while. Only a microscopic examination would reveal that the ink had simply rearranged itself in many thin tubes, which still occupied a volume of its initial size.

Therefore, Gibbs defined his new entropy S_{Gibbs} by means of a *coarse-grained* distribution ϱ^{cg}, obtained by averaging over non-overlapping small (but fixed) $6N$-dimensional volume elements ΔV_m $(m = 1, 2, \ldots)$ which cover the whole Γ-space:

$$\varrho^{\text{cg}}(p,q) = \frac{1}{\Delta V_m} \int\limits_{\Delta V_m} \varrho(p',q') \, dp' \, dq' =: \frac{\Delta p_m}{\Delta V_m} \quad \text{for} \quad p,q \in \Delta V_m \quad . \tag{3.28}$$

The resulting ensemble entropy is then given by

$$S_{\text{Gibbs}} := -k\eta[\varrho^{\text{cg}}] = -k \sum_m \Delta p_m \ln \frac{\Delta p_m}{\Delta V_m} \quad . \tag{3.29}$$

As already mentioned in connection with the smoothing of Boltzmann's μ-space distributions, the justification of this procedure by the uncertainty relations, and thus choosing the size of these phase space cells as h^{3N} (or $N!h^{3N}$), may be tempting, but would clearly be inconsistent with classical mechanics. The consistent quantum mechanical treatment (Chap. 4) leads again to the conservation of ensemble entropy (now for ensembles of wave functions). 'Quantum cells' of size h^{3N} can be justified only as convenient *units* of phase space volume in order to obtain the same normalization of entropy as in the classical limit of quantum statistical mechanics, where ensemble entropy vanishes for pure states (those that correspond to *one* cell - see (4.20)).

[3] See Wehrl (1978) for further measures, which are, however, not always monotonically related to one another. The conventional logarithmic measure is usually preferred because of the resulting additivity of the entropies of statistically independent subsystems.

The increase of Gibbs' entropy can now be understood according to the ink drop analogy. While the volume of the compact ink droplet is only slightly increased by moderate coarse-graining, that of a dense web of thin tubes (obtained by stirring) is considerably enlarged. Even though the coarse-graining itself is quite artificial, its efficiency depends on the shape of the volume to which it is applied. This is reminiscent of Boltzmann's μ-space densities, which characterize properties of the *discrete* particle distributions. Now, since there are evidently far more droplet shapes with a large surface than compact ones, the former have to be regarded as more probable. For *statistical* reasons one should hardly ever find a compact droplet (which is known to be true for three-dimensional droplets in the absence of surface tension).

However, there remains an essential difference between a droplet of ink in water and a dynamical volume element in phase space. While extension and shape of a droplet are real physical properties, the individual state of a classical mechanical system is always represented by a *point* in phase space. Course graining of the ink distribution may be likened to Boltzmann's smoothing procedure (that preserves properties of the *real* particle distribution), while Gibbs' entropy for a real state p, q, $S_{\mathrm{Gibbs}} = f(p,q) := k \ln \Delta V_{m_0}$ (resulting if $p, q \in \Delta V_{m_0}$), does *not* characterize this state.

Gibbs' procedure is therefore usually applied to phase space densities representing (incomplete) *information* from the outset. His entropy measures the enlargeability by coarse-graining of some status of knowledge. Its increase, $dS_{\mathrm{Gibbs}}/dt \geq 0$, under a deterministic (information-conserving) dynamical law describes transformation of coarse-grained (macroscopic) information, assumed to be present *initially*, into fine-grained information, which is thus regarded as 'irrelevant' (Sect. 3.2). However, this information concept conflicts with the idea of entropy as a objective physical quantity that is independent of any information held by an observer. This fundamental problem will be further discussed in Sect. 3.3 and later chapters.

Similar to smoothing in μ-space, the operation of coarse-graining cannot be uniquely inverted, since it destroys information. The intended chain of calculation,

$$\varrho^{\mathrm{cg}} \;\rightarrow\; \varrho \;\;\stackrel{\hat{L}}{\rightarrow}\;\; \frac{\partial \varrho}{\partial t} \;\rightarrow\; \frac{\partial \varrho^{\mathrm{cg}}}{\partial t} \quad , \tag{3.30}$$

in analogy to (3.2), is again broken at its first link. A new autonomous dynamics has therefore been proposed for ϱ^{cg}, in analogy to the *Stoßzahlansatz*, by complementing the Hamiltonian dynamics with a *dynamical* coarse-graining, applied in small but finite time steps Δt:

$$\left\{ \frac{\partial \varrho^{\mathrm{cg}}}{\partial t} \right\}_{\mathrm{master}} := \frac{[\mathrm{e}^{-\mathrm{i}\hat{L}\Delta t} \varrho^{\mathrm{cg}}]^{\mathrm{cg}} - \varrho^{\mathrm{cg}}}{\Delta t} \quad . \tag{3.31}$$

In this form it may be also regarded as a variant of the *Unifying Principle* thas was proposed by Lewis (1967). Instead of dynamically applying Gibbs' coarse-graining, Lewis himself suggested maximizing the entropy under the constraint of certain fixed 'macroscopic' quantities (see also Jaynes' theory in Sect. 3.3.1).

Equation (3.31) is reasonable if the corresponding probability increments $\Delta\Delta p_m$ (cf. 3.28) are proportional to Δt for small but finite time intervals Δt, thus describing transition *rates* between the cells ΔV_m. This important condition will be discussed in a more general form in Sect. 3.2, and when deriving the Pauli equation (4.17). Master equations such as (3.31) ensure a monotonic entropy increase. Their approximate validity requires that the microscopic (fine-grained) information remains *dynamically* irrelevant for the evolution of the coarse-grained one. Except in the case of equilibrium, this can at most be true in one direction of time.

Boltzmann's dynamics, described by his *Stoßzahlansatz*, can be similarly interpreted, as it neglects all particle correlations *after* they have formed in collisions (similar to fine-grained information). It, too, is based on the assumption that the interval Δt is finite and large in terms of collision times. The effect of a collision on the phase space distribution may be illustrated in two-dimensional momentum space (Fig. 3.1): a collision between two particles with small momentum uncertainties Δp_1 and Δp_2 leads deterministically to a *correlating* (deformed) volume element of the same size ΔV_Γ. (In a complete description, momenta would also be correlated with particle positions.) Subsequent neglect of the arising correlations will then enlarge this volume element ($\Delta V'_\Gamma > \Delta V_\Gamma$). However, neglecting such statistical correlations has no effect on the real phase space *point*.

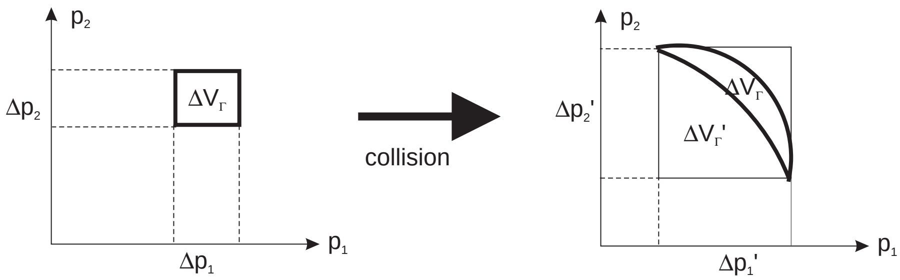


Fig. 3.1. Transformation of information about particle momenta into information about correlations between them as the basis of the H-theorem.

The question under what precise mathematical conditions certain systems are indeed 'mixing' in the sense of the plausible ink drop analogy (in a stronger version referred to as *K-systems* after Kolmogorov) is investigated in ergodic theory (see Arnol'd and Avez 1968, or Mackey 1989). Non-ergodic systems are usually pathological in forming sets of measure zero, or in being unstable against unavoidable perturbations. Otherwise they define robust (usually macroscopic) properties. Such situations have been claimed to exist generically (Yoccoz 1992), although their general physical meaning and relevance has never been established.

The strongest mixing is required for the finest conceivable coarse-graining. This is defined by its nontrivial limit $\Delta V_\Gamma \to 0$, which can also be formalized as a *weak convergence* for measures on phase space. It may be used to obtain infinite Poincaré recurrence times for isolated systems. However, this is neither required in a universe of finite age, nor would it be realistic, since quantum theory limits the entropy capacity available in the form of unlimited fine-graining of phase space. For this quantum mechanical reason there can be no 'overdetermination' of the *microscopic* past in spite of the validity of microscopic causality (cf. Footnote 1 of Chap. 2 and the end of Sect. 5.3). However, it is important to notice that all *concepts* of mixing are T-symmetric. In order to explain the time asymmetry of the Second Law ('irreversibility'), they have to be used dynamically in a specific direction of time.

Dynamical coarse-graining (as in (3.31)) may also be represented by an 'uncertain' Hamiltonian. An *ensemble of Hamiltonians* describes an indeterminism similar to stochastic dynamical models (used for calculating 'forward' in time). Even very small such uncertainties may be sufficient to completely destroy fine-grained information within a short time interval. For example, Borel (1924) estimated the effect of a changed gravitational force caused on earth by the displacement of a mass of the order of grams by a few centimeters at the distance of Sirius. He could demonstrate that it would completely alter the microscopic state of gas molecules in a vessel under normal conditions within seconds. Although the distortions of individual molecular trajectories are extremely small, they are amplified in each subsequent collision by a factor of the order of l/R, the ratio of the mean free path over the molecular radius. This extreme sensitivity to the environment describes in effect a *local microscopic indeterminism.*[4] In many situations, the microscopic distortions may even co-determine macroscopic effects (thus causing an effective *macroscopic indeterminism*), as discussed at length in the *theory of chaos* ('butterfly effect').

The essence of Borel's argument for our purpose is that microscopic states of macroscopic systems, aside from the whole universe, may never be treated as dynamically isolated — even when they are thermodynamically closed in the sense of $dS_{ext} = 0$. The *dynamical* coarse-graining that is part of the master equation (3.31) may indeed be ascribed to perturbations by the environment — provided they can be treated stochastically in the forward direction of time. This important dynamical assumption is yet another form of the *intuitive causality* that has been discussed at the beginning of Chap. 2 as a

[4] While the effect of Borel's gravitational distortions is drastically reduced in *quantum* gravity (or for any other relevant long range quantum interaction), other environmental effects (such as 'decoherence') become important unless there are stabilizing energy gaps in the considered systems (see Sects. 4.3.3 and 5.3).

manifestation of the arrow of time. The *representative ensembles* that are used in statistical thermodynamics may be understood as those which arise (and are maintained) by this stochastic nature of the unavoidable perturbations, while 'robust' properties are regarded as macroscopic.

While the *intrinsic dynamics* of a macroscopic physical system may thus transform coarse-grained information into fine-grained, interaction with the environment transforms the resulting fine-grained information very efficiently into practically useless correlations with distant systems. The sensitivity of microscopic states of macroscopic systems to their environment strongly indicates that simultaneously existing *opposite arrows of time* in different regions of the universe would be inconsistent with one another. This universality of the arrow of time seems to be its most important property. It has been regarded as an example of global *symmetry breaking*, although this classification would not yet exclude the far more probable symmetric (equilibrium) universe.

Schulman (1999) has recently challenged this conclusion by presenting an explicit counter example. However, his model is based on unrealistic and possibly inconsistent assumptions. It presumes (partial) two-time boundary conditions which are not dynamically independent (similar to the Wheeler-Feynman type boundary conditions — cf. Sects. 2.1 and 2.4). Such conditions may furthermore be insufficient to define arrows of time unless they had both causally evolved from their respective (opposite) 'pasts' (see also Casati, Chirikov and Zhirov 2000). This would dramatically enhance the consistency problem. Most importantly, Schulman's specific model is not based on Hamiltonian dynamics, but rather on 'cat maps', which do *not* reflect Borel's extreme sensitivity to very weak interactions.

In order to reverse the thermodynamical arrow of time in a bounded system, it would, therefore, not suffice to "go ahead and reverse all momenta" in it, as ironically suggested by Boltzmann as an answer to Loschmidt. In an interacting Laplacean universe, the Poincaré cycles of a subsystem could only be those of the entire universe, since a (quasi-)periodicity of the subsystem state would require the same periodicity of its effective Hamiltonian (which in turn depends on its environment). Time reversal with thermodynamical aspects has nonetheless been achieved for quantum mechanical spin waves, which are strongly protected against external perturbations by their weak interaction (similar to electromagnetic waves), and which furthermore allow a sudden sign reversal of their Hamiltonian in order to simulate time reversal (Rhim, Pines and Waugh 1971).

Spin wave experiments also demonstrate that an exactly closed system in thermodynamical equilibrium may still contain an arrow of time in the form of 'hidden correlations'. If the system has just reached equilibrium with respect to a certain kind of (generalized) coarse-graining, it appears T-symmetric, although its fine-grained information would still determine the direction and distance in time to its closest low-entropy state (see the Appendix for a numerical example). Systems affected by Borel's argument can *not* be reversed by local manipulations.

3.2 Zwanzig's General Formalism of Master Equations

Boltzmann's *Stoßzahlansatz* (3.6) for μ-space distributions and the master equation (3.31) for coarse-grained Γ-space distributions can thus be understood in a similar way. There are indeed many other master equations following this general strategy, which are to suit different purposes. Zwanzig (1960) succeeded in formalizing them in a general and instructive manner, that also reveals their analogy with retarded electrodynamics as another manifestation of the arrow of time (see (3.39) – (3.46) below).

The basic concept of Zwanzig's formalism are idempotent mappings $\hat{P}$, defined on probability distributions $\varrho(p,q)$,

$$\varrho \;\rightarrow\; \varrho_{\mathrm{rel}} := \hat{P}\varrho \quad \text{with} \quad \hat{P}^2 = \hat{P} \quad \text{and} \quad \varrho_{\mathrm{irrel}} := (1-\hat{P})\varrho \quad . \tag{3.32}$$

Their interpretation and motivation will be explained by means of a number of examples below. If these mappings *reduce* the information content of ϱ to what is then called its 'relevant' part ϱ_{rel}, they may be regarded as a *generalized coarse-graining*. In order to interpret ϱ_{rel} as a probability density again, one has to require its non-negativity and, for convenience,

$$\int \varrho_{\mathrm{rel}}\, dp\, dq = \int \varrho\, dp\, dq = 1 \quad , \tag{3.33a}$$

that is,

$$\int \varrho_{\mathrm{irrel}}\, dp\, dq = \int (1-\hat{P})\varrho\, dp\, dq = 0 \quad . \tag{3.33b}$$

Reduction of information means

$$S_\Gamma[\hat{P}\varrho] \geq S_\Gamma[\varrho] \tag{3.34}$$

(or similarly for any other measure of ensemble entropy).

Lewis' master equation (3.31) may then be written in the more general form

$$\left\{\frac{\partial \varrho_{\mathrm{rel}}}{\partial t}\right\}_{\mathrm{master}} := \frac{\hat{P}\mathrm{e}^{-\mathrm{i}\hat{L}\Delta t}\varrho_{\mathrm{rel}} - \varrho_{\mathrm{rel}}}{\Delta t} \quad . \tag{3.35}$$

It describes a monotonic increase of the corresponding entropy $S[\varrho_{\mathrm{rel}}]$. However, in contrast to Zwanzig's approach, to be discussed below, phenomenological master equations such as (3.35) are often understood as describing a *fundamental* indeterminism that would replace the Hamiltonian dynamics.

In most applications, Zwanzig's idempotent operations $\hat{P}$ are linear and Hermitean with respect to the inner product for probability distributions defined above (3.27). In this case they are projection operators, which project out some 'relevant component' of the original distribution. If such a *projection* obeys (3.33a) for every ϱ, it must leave the equipartition invariant, $\hat{P}1 = 1$, as can be seen from the inner product of this equation with an arbitrary distribution ϱ.

However, Zwanzig's dynamical formalism may also be meaningful for non-Hermitian or even non-linear idempotent mappings $\hat{P}$ (cf. Lewis 1967, Willis and Picard 1974). These mappings are then not projections any more: they may even *create new information.* A trivial example for the creation of information is the nonlinear mapping of all probability distributions onto a fixed one, $\hat{P}\varrho := \varrho_0$ for all ϱ, regardless of whether they contain ϱ_0 as a component. The physical meaning of such generalizations will be discussed in Sects. 3.4 and 4.4. In the following we will concentrate on information-reducing mappings.

Zwanzig's 'projection' concept is deliberately kept general in order to permit a wealth of applications. Examples introduced so far are coarse-graining, $\hat{P}_{\mathrm{cg}}\varrho := \varrho^{\mathrm{cg}}$, as defined in (3.28)), and the neglect of correlations between particles described in μ-space:

$$\hat{P}_\mu \varrho(p,q) := \prod_{i=1}^{N} \frac{\varrho_\mu(\boldsymbol{p}_i, \boldsymbol{q}_i)}{N} \quad , \quad \text{with}$$
$$\varrho_\mu(\boldsymbol{p}, \boldsymbol{q}) := \sum_{i=1}^{N} \int \varrho(p,q)\delta^3(\boldsymbol{p}-\boldsymbol{p}_i)\delta^3(\boldsymbol{q}-\boldsymbol{q}_i)\, dp\, dq \quad . \tag{3.36}$$

(As before, boldface letters represent three-dimensional vectors, while p, q is a point in Γ-space.) This particular application of the general concept defines a non-linear though information-reducing 'Zwanzig projection'. Most arguments, valid for linear operators $\hat{P}$, can here still be applied to the linearly resulting objects $\varrho_\mu(\boldsymbol{p}, \boldsymbol{q})$ (which do not live in Γ-space) rather than their products $\hat{P}_\mu \varrho(p,q)$ (which do). In quantum theory, this procedure is related to the Hartree or mean field approximation. Boltzmann's complete 'relevance concept' can be written as $\hat{P}_{\mathrm{Boltzmann}} = \hat{P}_\mu \hat{P}_{\mathrm{cg}}$, which leads to the convenient *smooth* μ-space distribution even when applied to points (or a sum of δ-functions) in Γ-space. An obvious generalization of $\hat{P}_{\mathrm{Boltzmann}}$ consists of a projection onto two-particle correlation functions. In this way a complete hierarchy of relevance concepts in terms of *n-point functions* (equivalent to a cluster expansion) can be defined.

A particularly important concept of relevance that is often not even noticed is *locality* (see e.g. Penrose and Percival 1962). It is required for the additivity of entropy that has been presumed in the fundamental phenomenological equation (3.1), and is required for the concept of an entropy *density*, $S = \int s(\boldsymbol{r})\, d^3r$. The corresponding Zwanzig projection of locality can be symbolically written as

$$\hat{P}_{\mathrm{local}}\varrho := \prod_k \varrho_{\Delta V_k} \quad . \tag{3.37}$$

It neglects all statistical correlations beyond an appropriate small distance in three-dimensional space. The probability distributions $\varrho_{\Delta V_k}$ are here defined by integrating out all degrees of freedom outside each respective volume element ΔV_k. The volume elements have to be assumed large enough to contain a sufficient number of particles in order to preserve dynamically relevant short range correlations (as required in real gases). For volume elements ΔV_k with physically open boundaries, their probability distributions $\varrho_{\Delta V_k}$ in (3.37) have to include variable particle number.

Locality is also presumed when one writes (3.1) in its differential (local) form as a 'continuity inequality' for the entropy density s,

$$\frac{\partial s}{\partial t} + \mathrm{div}\boldsymbol{j}_s \geq 0 \quad , \tag{3.1$'$}$$

with an entropy current $\boldsymbol{j}_s$. Further assumptions and approximations lead to phenomenological entropy-producing terms on the RHS (such as $\kappa(\mathrm{grad}T)^2/T^2$ in the case of heat conduction), and thus to a continuity *equation* with phenomenological *sources* for the entropy defined by this relevance concept (cf. Landau and Lifschitz 1959, Glansdorff and Prigogine 1971).

The general applicability of (3.1$'$) demonstrates that the concept of *physical entropy* is always based on the relevance of locality. Therefore, production of physical entropy can be generally understood as the transformation of local information into nonlocal correlations (similar to Fig. 3.1). This is in accordance with the conservation of ensemble entropy and intuitive causality. The Second Law thus depends crucially on the dynamical irrelevance of microscopic correlations *for the future* (as assumed in the *Stoßzahlansatz*, for example). Since this 'microscopic causality' cannot be observed as easily as the retardation of macroscopic radiation, its validity under all circumstances has been questioned (Price 1996). However, it (or its generalization to quantum correlations — see Sect. 4.2) is not only confirmed by the success of the *Stoßzahlansatz* and similar master equations, but also by microscopic scattering experiments, or the widely-observed phenomenon of exponential decay (Sect. 4.5). Irreproducible conspiratorial behavior cannot be excluded in principle, though.

The Zwanzig projection of locality is again ineffective on δ-functions in Γ-space — just as $\hat{P}_\mu$, but unlike $\hat{P}_{\mathrm{Boltzmann}}$. Therefore, it is not able to define a non-singular entropy as a function of *state*, that is, in the form $S_\Gamma[\hat{P}\delta^{6N}]$. Real (microscopic) states of classical systems are always local, as they define states of their local subsystems. This will be fundamentally different in quantum mechanics (Chap.4).

Boltzmann's relevance concept $\hat{P}_\mu$ evidently neglects *all* correlations between particles. The coarse graining over Δp and Δq, required to obtain smooth distributions in μ-space, may remain small for individual particles in a dense gas, even though the corresponding *distance* in configuration space, $(3N)^{1/2}\overline{\Delta q}$, written in terms of the mean coordinate difference $\overline{\Delta q}$, is quite large. 'Collective' coordinate differences may similarly become large in spite of small differences of particle coordinates. They may be *relevant* for certain macroscopic phenomena, such as lattice vibrations. In quantum theory they may even give rise to new 'particle' concepts (phonons).

As mentioned shortly after Fig. 3.1, coarse-graining as a relevance concept may also enter in a hidden form (corresponding to its nontrivial limit $\Delta V_\Gamma \to 0$ taken before any formal limit $t \to \infty$) by considering only *non-singular measures* on phase space (thus excluding δ-functions). The dynamical effect of this formal idealization may be mathematically signalled by a 'unitary inequivalence' between the Liouville equation and the resulting master equation (see Misra 1978 or Mackey 1989).

Effective *equations of motion* for individual particles can also be understood as master equations if environmental degrees of freedom, such as heat or emitted radiation (described in Sect. 2.3), are 'eliminated' in an appropriate way. Further examples of Zwanzig projections will be defined throughout the book, in particular in Chap. 4 for quantum mechanical applications, where the relevance of locality leads to the important concept of decoherence. Different schools and methods of irreversible thermodynamics may even be categorized by the concepts of relevance which they prefer, and which they typically regard as 'natural' or 'fundamental' (cf. Grad 1961). However, the mere *conceptual* justification of a relevance concept ('paying attention') does not yet warrant its *dynamical* autonomy in the form of a master equation. (See the Appendix for an explicit example.) For example, locality is relevant because of the locality of interactions.

The dynamical locality of interactions is therefore essential for the very concept of *systems*, and in this way even for that of a *local observer* as the ultimate referee for what is relevant. Different local systems may thus possess different thermodynamical equilibrium parameters (such as space-dependent temperature). Exceptional (weakly interacting) systems may even possess different temperatures (of different things) at the same place. An example is the solar corona (an electron gas surrounding the sun), which is much hotter than the baryonic solar surface. In the extreme case, dynamically decoupled systems would represent 'separate worlds'.

Rather than *postulating* a phenomenological master equation (3.35) in analogy to Boltzmann, Zwanzig reformulated the exact Hamiltonian dynamics of ϱ_{rel} regardless of any specific choice of $\hat{P}$. It cannot generally be autonomous (of the form $\partial\varrho_{\mathrm{rel}}/\partial t = f(\varrho_{\mathrm{rel}})$), [5] but has to be written as

$$\frac{\partial \varrho_{\mathrm{rel}}}{\partial t} = f(\varrho_{\mathrm{rel}}, \varrho_{\mathrm{irrel}}) \quad , \tag{3.38}$$

[5] In mathematical physics, the term 'autonomous dynamics' is often understood as the absence of any explicit time dependence in the dynamics (regardless of its origin). Moreover, 'states' are there usually operationally defined (by a preparation procedure, for

in order to eliminate ϱ_{irrel} by means of certain assumptions. The procedure is analogous to the elimination of the electromagnetic degrees of freedom by means of the condition $A^{\mu}_{\mathrm{in}} = 0$ when deriving a retarded action-at-a-distance theory (Sect. 2.2). In both cases, empirically justified boundary conditions which specify a time direction are assumed to hold for the degrees of freedom that are to be eliminated.

Toward this end the Liouville equation $\mathrm{i}\partial\varrho/\partial t = \hat{L}\varrho$ is decomposed into its relevant and irrelevant parts by multiplying it by $\hat{P}$ or $1-\hat{P}$, respectively,

$$\mathrm{i}\frac{\partial\varrho_{\mathrm{rel}}}{\partial t} = \hat{P}\hat{L}\varrho_{\mathrm{rel}} + \hat{P}\hat{L}\varrho_{\mathrm{irrel}} \tag{3.39a}$$

$$\mathrm{i}\frac{\partial\varrho_{\mathrm{irrel}}}{\partial t} = (1-\hat{P})\hat{L}\varrho_{\mathrm{rel}} + (1-\hat{P})\hat{L}\varrho_{\mathrm{irrel}} \quad . \tag{3.39b}$$

This corresponds to representing of the Liouville operator by a matrix of operators

$$\hat{L} = \begin{pmatrix} \hat{P}\hat{L}\hat{P} & \hat{P}\hat{L}(1-\hat{P}) \\ (1-\hat{P})\hat{L}\hat{P} & (1-\hat{P})\hat{L}(1-\hat{P}) \end{pmatrix} \quad . \tag{3.40}$$

Equation (3.39b) for ϱ_{irrel}, with $(1-\hat{P})\hat{L}\varrho_{\mathrm{rel}}$ regarded as an inhomogeneity, may then be formally solved by the method of the variation of constants (interaction representation). This leads to

$$\begin{aligned}\varrho_{\mathrm{irrel}}(t) = {} & \mathrm{e}^{-\mathrm{i}(1-\hat{P})\hat{L}(t-t_0)}\varrho_{\mathrm{irrel}}(t_0) \\ & - \mathrm{i}\int_0^{t-t_0} \mathrm{e}^{-\mathrm{i}(1-\hat{P})\hat{L}\tau}(1-\hat{P})\hat{L}\varrho_{\mathrm{rel}}(t-\tau)\,d\tau \quad ,\end{aligned} \tag{3.41}$$

as may be confirmed by differentiation.

If $t > t_0$, (3.41) is analogous to the *retarded form* (2.9a) of the boundary value problem in electrodynamics. In this case, $\tau \geq 0$, and $\varrho_{\mathrm{rel}}(t-\tau)$ may be interpreted as an advanced source for the 'retarded' $\varrho_{\mathrm{irrel}}(t)$. Substituting this formal solution (3.41) into (3.39a) leads to three terms on the RHS, *viz.*

$$\begin{aligned}\mathrm{i}\frac{\partial\varrho_{\mathrm{rel}}(t)}{\partial t} &= I + II + III \\ &\equiv \hat{P}\hat{L}\varrho_{\mathrm{rel}}(t) + \hat{P}\hat{L}\mathrm{e}^{-\mathrm{i}(1-\hat{P})\hat{L}-(t-t_0)\varrho_{\mathrm{irrel}}(t_0)} - \mathrm{i}\int_0^{t-t_0} \hat{G}(\tau)\varrho_{\mathrm{rel}}(t-\tau)\,d\tau \;.\end{aligned} \tag{3.42}$$

The integral kernel of the last term,

$$\hat{G}(\tau) := \hat{P}\hat{L}\mathrm{e}^{-\mathrm{i}(1-\hat{P})\hat{L}\tau}(1-\hat{P})\hat{L}\hat{P} \quad , \tag{3.43}$$

---

example). In general they represent *ensembles* of microscopic states. This may easily lead to confusion, in particular in a 'statistical interpretation' of quantum theory (Chap. 4).

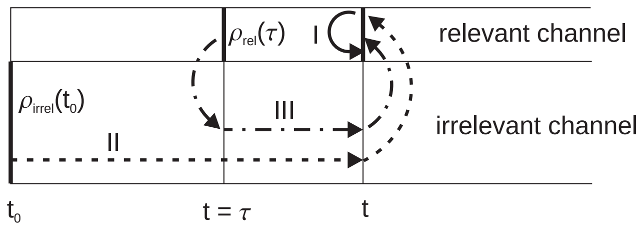


Fig. 3.2. Retarded form of the exact dynamics for the relevant information according to Zwanzig's *pre-master equation*. In addition to the instantaneous 'self-interaction' I, there is the contribution II arising from the 'incoming' irrelevant information, and the retarded term III in analogy to electromagnetic action at a distance, resulting from 'advanced sources' in the whole time interval between t_0 and t (cf. the left part of Fig. 2.2.)

corresponds to the retarded Green's function of Sect. 2.1 (if $t > t_0$).

Equation (3.42) is exact and does not as yet describe time asymmetric physics. Since it forms the first step in this derivation of master equations, it is known as a *pre-master equation*. The meaning of its three terms is illustrated in Fig. 3.2. The first one is the 'trivial part', which describes the direct interaction of ϱ_{rel} with itself. In Boltzmann's μ-space dynamics (3.3), it would correspond to $\{\partial\varrho_\mu/\partial t\}_{\mathrm{free+ext}}$. It vanishes if $\hat{P}\hat{L}\hat{P} = 0$ (as is often the case). [6]

The second term is usually omitted by presuming the absence of irrelevant initial information: $\varrho_{\mathrm{irrel}}(t_0) = 0$. If *relevant* information happens to be present initially, it can be transformed into irrelevant information. (Because of the asymmetry between $\hat{P}$ and $1 - \hat{P}$, irrelevant information would have to be measured by $-S[\varrho] + S[\varrho_{\mathrm{rel}}]$ rather than $-S[\varrho_{\mathrm{irrel}}]$.)

The vital third term is *non-Markovian* (non-local in time), as it depends on the whole time interval between t_0 and t. Its retarded form (for $t > t_0$) is compatible with the intuitive concept of causality. This term becomes approximately *Markovian* (while in general remaining time asymmetric) if $\varrho_{\mathrm{rel}}(t-\tau)$ varies slowly over a relatively short 'relaxation time' τ_0 during which $\hat{G}(\tau)$ becomes negligible by dissipating its information. In (3.42), $\hat{G}(\tau)$ may then in lowest order be assumed to be proportional to a δ-function in τ. This assumption is also contained in Boltzmann's *Stoßzahlansatz*. In analogy to action-at-a-distance theories, this third term of the pre-master equation assumes the

[6] Since the (indirectly acting) non-trivial terms contribute only in second or higher order of time, the time derivative defined by the master equation (3.35) would vanish in the limit $\Delta t \to 0$. This corresponds to what in quantum theory is known as the *quantum Zeno paradox* (Misra and Sudarshan 1977), also called *watched pot behavior* or the *watchdog effect*. It describes an immediate loss of 'information' from the irrelevant channel (or its dynamically relevant parts — see later in the discussion), such that it has no chance to react on its relevant counterpart. Fast information loss may be induced by an efficient coupling to the environment, for example. Since this efficiency depends on the level density (Joos 1984), the Zeno effect is mainly important in quantum mechanical systems.

form of an effective direct interaction (instantaneous in this nonrelativistic case) between the relevant degrees of freedom. (In electrodynamics, charge positions are the 'relevant' variables.) In statistical physics, this 'interaction' describes the dynamics of *ensembles*.

Markovian behavior may may now be derived from assumptions which simultaneously explain the general *dyanamical* applicability of the initial condition $\varrho_{\mathrm{irrel}} \approx 0$ at *all* times (provided it holds in the very distant past — again in analogy to $A^{\mu}_{\mathrm{in}} \approx 0$ in electrodynamics). Consider the action of the operator $(1-\hat{P})\hat{L}\hat{P}$ that appears on the RHS of the kernel (3.43). Because of the structure of a typical Liouville operator, it transforms information from ϱ_{rel} only into *specific* parts of ϱ_{irrel}. In the theory of nuclear reactions such parts are called *doorway states* (Feshbach 1962). For example, if the Hamiltonian contains no more than two-particle interactions, $\hat{L}\hat{P}_{\mu}$ creates at most two-particle correlations. Only the subsequent application of the propagator $\exp[-\mathrm{i}(1-\hat{P})\hat{L}\tau]$ is then able to produce states 'deeper' in the irrelevant channel (many-particle correlations in this case) — see Fig. 3.3. However, recurrence from the depths of the irrelevant channel is related to Poincaré recurrence times, and may be neglected in the normal case. If the *relaxation time*, now defined as the time required for the transfer of information from the doorway 'states' into deeper parts of the irrelevant channel, is of the order τ_0, say, one may assume $\hat{G}(\tau) \approx 0$ for $\tau \gg \tau_0$, as required for the Markovian δ-function approximation $\hat{G}(\tau) \approx \hat{G}_0\delta(\tau)$.

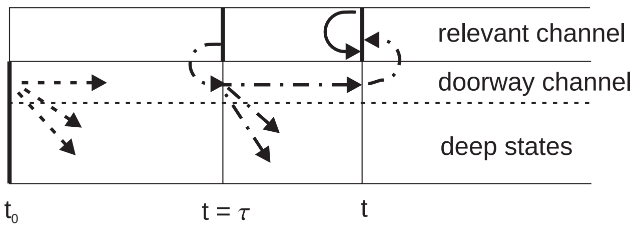


Fig. 3.3. The large information capacity of the irrelevant channel and the specific structure of the interaction together facilitate the disappearance of information into the depth of the irrelevant channel if an appropriate initial condition holds

Essential for the validity of this picture is the large information capacity of the irrelevant channel (similar to that of the electromagnetic field in Chap. 2, but exceeding it by far). For example, correlations between particles require far more information than the single particle distribution ϱ_{μ}. A fundamental *cosmological* assumption,

$$\varrho_{\mathrm{irrel}}(t_0) = 0 \quad , \tag{3.44}$$

at a time t_0 in the *finite* past (similar to the cosmological $A^{\mu}_{in} = 0$ at the big bang) is therefore quite powerful — even though it is a *probable* condition. Any irrelevant information formed later from the initial 'information'

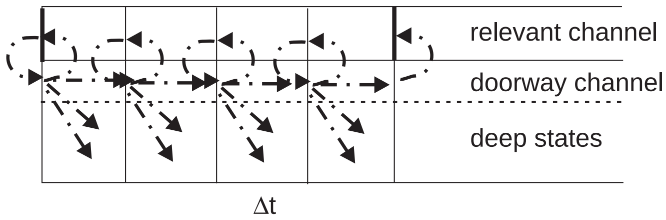


Fig. 3.4. The master equation represents 'alternating dynamics', usually describing a monotonic loss of relevant information

contained in $\varrho_{\mathrm{rel}}(t_0)$ (that is, from any specification of the initial state) may be expected to remain dynamically negligible in (3.42) for a very long time. It would be essential, however, for calculating backwards in time.

The assumption $\varrho_{\mathrm{irrel}} \approx 0$ has thus to be understood in a *dynamical* sense: any newly formed contribution to ϱ_{irrel} must remain irrelevant in the 'forward' direction of time. The dynamics for ϱ_{rel} may then appear autonomous (while it cannot be exact). For example, all correlations between subsystems seem to require advanced local *causes*, but no similar (retarded) *effects*. Otherwise they would be interpreted as a *conspiracy*, the deterministic version of *causae finales*.

Under these *assumptions* one obtains as a first step from (3.42) the non-Markovian dynamics

$$\frac{\partial \varrho_{\mathrm{rel}}(t)}{\partial t} = -\int_0^{t-t_0} \hat{G}(\tau)\varrho_{\mathrm{rel}}(t-\tau)\,d\tau \quad . \tag{3.45}$$

The upper boundary of the integral can here be replaced by a constant T that is large compared to τ_0, but small compared to any (theoretical) recurrence time for $\hat{G}(\tau)$. If $\varrho_{\mathrm{rel}}(t)$ may now be assumed constant over time intervals of the order of the relaxation time τ_0, corresponding to an already prevailing partial (e. g. local) equilibrium, one obtains the nontrivial Markovian limit ($\tau_0 \to +0$),

$$\frac{\partial \varrho_{\mathrm{rel}}(t)}{\partial t} \approx -\hat{G}_{\mathrm{ret}}\varrho_{\mathrm{rel}}(t) \quad , \tag{3.46a}$$

with

$$\hat{G}_{\mathrm{ret}} := \int_0^T \hat{G}(\tau)\,d\tau \quad . \tag{3.46b}$$

A similar nontrivial limit of vanishing retardation led to the Lorentz-Dirac equation with its asymmetric radiation reaction in Sect. 2.3. The integral (3.46b) can be formally evaluated after inserting (3.43), although it is more conveniently computed after applying this operator to a specific $\varrho(t)$. (See the explicit evaluation for discrete quantum mechanical states in Sect. 4.1.2.)

The autonomous master equation (3.46) describes again an *alternating dynamics* of the type (3.35) (see Fig. 3.4). Irrelevant information may be disregarded within short but finite time intervals Δt (now representing the relaxation time τ_0). If $\hat{P}$ only destroys information, the master equation describes never-decreasing entropy,

$$\frac{dS_\Gamma[\varrho_{\text{rel}}]}{dt} \geq 0 \quad . \tag{3.47}$$

This corresponds to a *positive* operator $\hat{G}_{\text{ret}}$ (as can most easily be shown for the *linear* measure of entropy).

For example, a phenomenological probability-conserving Markovian master equation for a system with 'macroscopic states' described by a (set of) 'relevant' variable(s) α, that is, $\varrho_{\text{rel}}(t) \equiv \varrho(\alpha, t)$, (see also Sects. 3.3 and 3.4) can be written as

$$\frac{\partial \varrho(\alpha, t)}{\partial t} = \int [w(\alpha, \alpha')\varrho(\alpha', t) - w(\alpha', \alpha)\varrho(\alpha, t)]\, d\alpha' \quad . \tag{3.48}$$

The transition rates $w(\alpha, \alpha')$ define $\hat{G}_{\text{ret}}$ in the form of the integral kernel $\hat{G}_{\text{ret}}(\alpha, \alpha') = -w(\alpha, \alpha') + \delta(\alpha, \alpha') \int w(\alpha, \alpha'')d\alpha''$. If they satisfy a generalized time inversion symmetry,

$$\frac{w(\alpha, \alpha')}{\sigma(\alpha)} = \frac{w(\alpha', \alpha)}{\sigma(\alpha')} \quad , \tag{3.49}$$

where $\sigma(\alpha)$ may represent the density of microscopic states with respect to the variable α, that is, $\sigma := dn/d\alpha$, an H-theorem can again be derived in analogy to (3.12) for the *generalized H-functional*

$$H_{gen}[\varrho(\alpha)] := \int \varrho(\alpha) \ln \frac{\varrho(\alpha)}{\sigma(\alpha)}\, d\alpha = \overline{\ln p} \quad . \tag{3.50}$$

The second RHS is correct, since probabilities p for the individual microscopic states (when assumed to depend only on α) are given by $p(\alpha) = \varrho(\alpha)/\sigma(\alpha)$. The entropy $-kH_{gen}$ is also known as the *relative entropy* of $\varrho(\alpha)$ with respect to the measure $\sigma(\alpha)$. It is often introduced phenomenologically as part of the macroscopic description.

For $w(\alpha', \alpha) = f(\alpha)\delta'(\alpha - \alpha')$, one obtains the deterministic 'drift' limit of the master equation (3.48) — corresponding to the first term of (3.42). It defines the first order of the *Kramers-Moyal expansion*, equivalent to an expansion of $\varrho(\alpha', t)$ in terms of powers of $\alpha' - \alpha$. The second order, $w(\alpha', \alpha) = f(\alpha)\delta'(\alpha - \alpha') + g(\alpha)\delta''(\alpha - \alpha')$, leads to the *Fokker-Planck equation* as the lowest irreversible approximation (see de Groot and Mazur 1962, Röpke 1987). It is formally analogous to the Lorentz-Dirac equation as the lowest non-trivial order in the Taylor expansion of the Caldirola equation (2.30). A master equation is generally equivalent to a (stochastic) *Langevin* *equation* applied to the *individual* trajectories $\alpha(t)$ that would form an ensemble represented by $\varrho(\alpha, t)$.

In contrast to the Liouville equation (3.26), the master equation (3.46) or (3.35) cannot be unitary with respect to the inner product for probability distributions defined previous to (3.27). While total probability may be conserved, that of the individual trajectories is *not* in this case (see also Sect. 3.4). Information-reducing master equations describe an indeterministic evolution, which in general determines only an ever-increasing ensemble of different *potential* successors for each macroscopic state (such as a point in α-space).[7] As discussed above, this is compatible with microscopic determinism if — as is perhaps statistically reasonable under the assumption (3.44) — that information which is transformed from relevant to irrelevant in the course of time no longer has any relevant effects for future times of interest. The validity of this assumption depends on the dynamics and on the specific initial conditions.

Time-reversed ('anti-causal') effects would have to be derived from a *final* condition $\varrho_{\mathrm{irrel}}(t_f) = 0$ in the far future by applying the corresponding approximations to (3.42). However, it is an *empirical fact* that such a condition, analogous to $A^{\mu}_{\mathrm{out}} = 0$ in electrodynamics, does not describe our observed universe. An exact boundary condition $\varrho_{\mathrm{irrel}}(t_0) = 0$ at some *accessible time* t_0 would for similar statistical reasons lead to a non-*de*creasing entropy for $t > t_0$, but to non-*in*creasing entropy for $t < t_0$, hence to an entropy minimum at $t = t_0$. This emphasizes once again the T-symmetry of statistical considerations.

While the (statistically probable) assumption (3.44) led to the master equation (3.46), it would by itself *not* characterize an arrow of time. Without an improbable assumption about $\varrho_{\mathrm{rel}}(t_i)$, the approximate validity of the equality sign in (3.47) would be overwhelmingly probable. The condition (3.44) could then be dynamically appropriate in both directions of time. *Improbable*, though required as the *vital* assumption for an arrow of time, is only an initial condition $S[\varrho_{\mathrm{rel}}] \ll S_{\mathrm{max}}$. Retarded action-at-a-distance electrody-

[7] The frequently used idea of a 'fork', defined in configuration space, to characterize a dynamical indeterminism may be misleading, since it seems to imply unique predecessors. This is wrong, as can be recognized, for example, in an equilibrium situation. In the case of a stochastic dynamical law that is defined on a closed set of states, a state must, in general, also be *reached from* different predecessors. This would define an inverse fork. Inverse forks may also describe a pure forward-determinism (a 'semigroup', that contains *attractors*). All these structures are meant to characterize the *dynamical law.* They are neither properties of the (f)actual history (which is assumed to evolve along a *definite* trajectory regardless of the dynamical law), nor of an evolving ensemble that represents incomplete knowledge.

However, only unique predecessors may give rise to recordable histories (where history can be reconstructed from the records). The *historical nature* of the world is thus based on the existence of unique past histories for *those* macroscopic quantities which represent documents or can be reliably documented. (See also Footnote 1 of Chap. 2, Fig. 3.8, and the discussion which concludes this Chapter.) On the other hand, a macroscopic history that was completely determined from its macroscopic past would be in conflict with the notion of (an apparent) free will.

namics would be trivial, too, (and equivalent to its advanced counterpart) if *all* sources were absorbers in the thermodynamical sense. The low entropy initial condition is responsible for the formation of that irrelevant information which would be required for correctly calculating $\varrho_{\mathrm{rel}}(t)$ backwards in time.

The main conclusions derived in this and the previous section can thus be listed as follows:

1. The *ensemble entropy* S_Γ does not represent physical entropy, since (a) it would diverge for a real physical state (a point in phase space), (b) it is otherwise not additive for composite systems (in particular, it is not an integral over an entropy density), and (c) it remains constant under deterministic dynamics (in contrast to the Second Law). For indeterministic dynamics it would increase from a given value in *both* directions of time (except in equilibrium). This demonstrates drastically indeed that ensemble entropy, representing incomplete information, is not an objective physical quantity (see Kac 1959).
2. Coarse-grained (or 'relevant') entropy, when defined as a function of the deterministically evolving microscopic state, would in general fluctuate, most of the time remaining close to its maximum value. However, it may increase during a very long time (by far exceeding the present age of the universe) if the universe had begun in an appropriate state of extremely low entropy (see Sect. 5.3). While a Zwanzig projection can be *chosen for convenience* in order to derive a master equation (if dynamically consistent), the initial condition must be *specified* as a real condition characterizing this universe.
3. Only a relevance concept of locality, such as (3.37), is able to describe the additivity of phenomenological entropy.
4. Any coarse-grained entropy could be forced *never* to decrease by an appropriate modification of the corresponding *ensemble* dynamics, such as in (3.35). This may either represent new physics or an approximation to the situation described in item 2.

General literature: Jancel 1963, Balian 1991

3.3 Thermodynamics and Information

3.3.1 Thermodynamics Based on Information

As explained in the previous section, Gibbs' probability densities or ensembles ϱ_Γ represent incomplete ('macroscopic') *information* about the microscopic state (point in phase space) that in classical mechanics is assumed to represent the physical state. Similarly, Zwanzig's projection operators (generalized coarse-grainings) $\hat{P}$ are motivated by incomplete observability (or preparation) of macroscopic systems. Parameters characterizing ensembles, in particular the entropy, would then be fundamentally observer-related (that is, meaningless for the *real* physical states). In Gibbs' approach the ensembles seem to describe *actual* knowledge, while for Zwanzig (or Boltzmann) they may be based on a shared ('objectivized') limitation of knowledge common to a certain class of potential observers, such as those who can only utilize information contained in the single-particle distribution ϱ_μ.[8] For this reason, the (generalized) course-graining $\hat{P}$ is kept fixed, and not regarded as co-moving according to the Liouville equation. However, the concepts of information or knowledge would be extra-physical ones if the carrier of this information were not described as a physical system, too, (see Sect. 3.3.2).

Jaynes (1957) generalized Gibbs' statistical methods by rigorously applying Shannon's (1948) formal information concept to physical systems. Shannon's measure of information for a probability distribution $\{p_i\}$ over a set of elements distinguished by an index i,

$$I := \sum_i p_i \ln p_i \leq 0 \quad , \tag{3.51a}$$

is defined in analogy to Boltzmann's H, and therefore also called *negentropy*. However, since it is meant to measure lacking information, it corresponds more closely to Gibbs' η. It is often normalized *relative* to its value for minimum information, $p_i = p_i^{(0)}$ (where $p_i^{(0)} = 1/N$ if $i = 1 \dots N$, unless different statistical weights of the 'elements' i arise from a more fundamental level of description):

$$I_{rel} = I(p_i|p_i^{(0)}) := \sum_i p_i \ln(p_i/p_i^{(0)}) = \ln N + \sum_i p_i \ln p_i \geq 0 \quad . \tag{3.51b}$$

This renormalized measure of information may remain finite for diverging I in the limit $N \to \infty$. With an appropriate modification it can even be applied

[8] The term 'objectivized' presumes the basically subjective (observer-related) status of what is to be objectivized. In contrast, the term 'objective' is in physics often used synonymously with 'real', and then means the assumed or conceivable existence regardless of its (f)actual observation.

to a continuum. While working with renormalized quantities may simply be *convenient* from a mathematical point of view, renormalization is physically *meaningful* even if it did involve infinities.

Jaynes based his approach on the assumption that in general only a small subset of all state functions $\alpha(p,q)$ of a macroscopic system can be *known*. Therefore, he introduced new 'representative ensembles', $\varrho_\alpha(p,q)$ or $\varrho_{\bar\alpha}(p,q)$, defined to maximize the entropy functional $S_\Gamma[\varrho]$ under the constraints of fixed values α or fixed mean values $\bar\alpha := \int \alpha(p,q)\varrho(p,q)\,dp\,dq$, respectively. They represent minimum information under this constraint. The entropy thus becomes a function of α or $\bar\alpha$ according to $S(\bar\alpha) := S_\Gamma[\varrho_{\bar\alpha}]$, for example. This generalization of Gibbs' ideas has turned out useful for many applications.

As mentioned in Sect. 3.1.2, this information-theoretical foundation contrasts strongly with the understanding of entropy as an observer-independent physical quantitiy that can be objectively measured. On the other hand, the dependence on a certain basis of information may be quite appropriate. For example, the numerical value of $S_\Gamma[\varrho]$ depends in a reasable way on whether or not ϱ contains information about actual density fluctuations, or about the isotopic composition of a gas. The probability $p_{\text{fluct}}(\alpha)$ for the occurrence of some quantity α in thermodynamical equilibrium was successfully calculated by Einstein in his theory of Brownian motion from the expression

$$p_{\text{fluct}}(\alpha) = \frac{\exp\{S(\alpha)/k\}}{\exp\{S[\varrho_{\text{can}}]/k\}} \quad , \tag{3.52}$$

thus exploiting the interpretation of entropy as a measure of probability. The probability for other quantities to be found immediately after the occurrence of this fluctuation would then have to be calculated from the 'conditioned' ensemble ϱ_α rather than from ϱ_{can}.

Similarly, a star cluster (that is, a big enough collection of macroscopic objects) possesses meaningful temperature and entropy $S \neq 0$ from the *point of view* that the motion of the individual stars is regarded as 'microscopic'. The same statistical considerations as used for molecules then show that their velocity distribution must be Maxwell's. Imagine, in contrast, the viewpoint of an (only classically conceivable) external *super-observer* of the individual molecules in a gas! Entropy is here indeed defined as depending on the basis of consideration. Its objectivity in thermodynamics can only be understood as reflecting a common perspective shared by us human observers. This perspective must represent our situation as physical systems: we can hardly observe molecular dynamics without disturbing it, while the reaction of stellar dynamics to our observational intervention is negligible.

In order to regard ensembles as consistently representing *actual* information, one has to allow for 'external dynamics' affecting the probabilities, such as $\varrho_\Gamma(p,q,t)$. For example, in addition to the observer-independent Liouville equation that describes the motion of each member of the ensemble according to the Hamiltonian equations, one would have to take into account any change of information, such as its increase by a new observation or its reduction during a loss of memory. The corresponding *physical* processes now occur in an external observer or storage and measurement device, but not in the considered system itself (although possibly induced by interaction with them — see the following subsection). A loss of memory, for example, might be associated with dissipation in the physical information carrier. However, in contrast to Jaynes' interpretation, the *physical concept of entropy* is based on the assumption that certain ('easily accessible' or 'controllable') quantities, such as mean density, temperature and other *macroscopic* variables, are *always given* (but nothing else). This physical entropy does then not depend on the information actually at hand. For example, it does not depend on whether or how accurately the temperature has been measured; it is simply *defined as a function* of temperature.

Let $\alpha(p,q)$ represent such (sets of) 'relevant' quantities (cf. (3.48)) which are assumed to be given (possibly within some uncertainty $\Delta\alpha$). The Hamiltonian $H(p,q)$ is in general one of them. The subsets of states p,q within intervals $\alpha_0 - \Delta\alpha < \alpha(p,q) < \alpha_0$, with widths $\Delta\alpha$ defined, for example, by Jaynes' representative ensembles for given mean values, define subvolumes of Γ-space. (Any finer resolution would regard fluctuations as being relevant, unless α is a constant of the motion for the considered system.) For a single parameter α, these volume elements can be written as $\Delta V_\alpha := (dV/d\alpha)\Delta\alpha$, with $V(\alpha_0) := \int_{\alpha(p,q)<\alpha_0} dp\,dq$. In high-dimensional N-particle phase space, the size of the interval $\Delta\alpha$ is quite irrelevant, since contributions to the volume integral for a compact region $\alpha(p,q) < \alpha_0$ are strongly peaked just below the surface defined by α_0. In reasonable units, the term $\ln\Delta\alpha$ can then be neglected under the logarithm, $\ln\Delta V_\alpha$, that defines entropy.

One may now define a new useful Zwanzig projection $\hat{P}_{\mathrm{macro}}$ by averaging over these subsets:

$$\begin{aligned}\hat{P}_{\mathrm{macro}}\varrho(p,q) &:= \frac{\Delta p_\alpha}{\Delta V_\alpha} \\ &:= \frac{1}{\Delta V_\alpha}\int\limits_{\Delta V_\alpha} \varrho(p',q')\,dp'\,dq' \quad \text{for} \quad p,q \in \Delta V_\alpha \quad . \end{aligned} \tag{3.53}$$

If discrete values α_i are defined for convenience according to $\alpha_i + \Delta\alpha = \alpha_{i+1}$, the entropy $S_\Gamma[\hat{P}_{\mathrm{macro}}\varrho]$ splits into two sums,

$$\begin{aligned}S_\Gamma[\hat{P}_{\mathrm{macro}}\varrho] &= -k\int \hat{P}_{\mathrm{macro}}\varrho \ln(\hat{P}_{\mathrm{macro}}\varrho)\,dp\,dq \\ &= -k\sum_i \Delta V_{\alpha_i}\frac{\Delta p_{\alpha_i}}{\Delta V_{\alpha_i}}\ln\frac{\Delta p_{\alpha_i}}{\Delta V_{\alpha_i}} \\ &= -k\sum_i \Delta p_{\alpha_i}\ln\Delta p_{\alpha_i} + \sum_i \Delta p_{\alpha_i} k\ln\Delta V_{\alpha_i} \quad . \end{aligned} \tag{3.54}$$

(Notice the similarity to the concept of relative information (3.50) or (3.51b) — see also Schlögl 1966.) The first term in the last line describes the entropy corresponding to the lacking *macroscopic information* about α. The second term is the mean 'physical entropy' with respect to this macroscopic ensemble. The physical entropy $S(\alpha) := k \ln \Delta V_\alpha \approx S_\Gamma[\varrho_\alpha]$ thus measures the size of Jaynes' representative ensembles ϱ_α, or, in Planck's language, the *number of complexions*, that is, the number of microscopic states which may represent the macroscopic 'state' (here defined by the interval $\Delta\alpha$). In the special case $\alpha(p,q) := H(p,q)$ one obtains the entropy of the canonical ensemble as a function of the mean (or maximum) energy. If $\Delta\alpha = \Delta E$ is chosen infinitesimal, one obtains the entropy of the micro-canonical ensemble, relevant for energetically closed systems. [9]

Although the first term on the RHS of (3.54) is usually much smaller than the second one, it is essential for a complete and consistent discussion of information processing and measurement (see Sect. 3.3.2). An example of the partition of the ensemble entropy into physical entropy and entropy of lacking information, one that is not plagued by renormalization problems, is provided by the particle number in a grand canonical ensemble, $Z^{-1}\exp[-(H-\mu N)/kT]$. This particle number is assumed to be 'given' (although in general not known) once the vessel that was in equilibrium with a particle reservoir characterized by the chemical potential μ has been closed. The system is thereafter represented by a *canonical* ensemble with fixed N, while the relative contribution of that part of the original ensemble entropy which has now become entropy of lacking information about the exact particle number N is of the order $\ln N/N$ (Casper and Freier 1973). This contribution to the entropy is often neglected by using the 'approximation' $N! \approx N^N$. The argument demonstrates, however, that this different choice of ensembles is dynamically motivated (by their robustness), and that the difference between the number of permutations, $N!$, of a fixed number N of particles and the factor N^N arising from the grand canonical ensemble with *mean* particle number N (see (4.30)) has an interpretation (cf. Footnote 2).

While the concept of *physical* entropy, defined above, does not depend on actual information any more, the choice of 'macroscopic' subsets, characterized by functions of state $\alpha(p,q)$, is still motivated by their 'controllability', and therefore may appear artificial from a strictly objective point of view. In general, those variables α are regarded as macroscopic which are represented by subspaces of phase space that are densely populated by a trajectory (in the sense of quasi-ergodicity) within a short time (in relevant units). The quantitative aspect of quasi-ergodicity within these *robust* subspaces relies on a measure of Γ-space distance that — as has been pointed out before — is not invariant under canonical transformations. The macroscopic variables α themselves (or their mean values) are instead assumed to vary slowly

[9] The infinite renormalization which is required for the definition of an entropy *density* or a corresponding measure as a function of α is due to the fact that the classical entropy has no lower bound (as it would arise from a fundamental minimum size of phase space cells). This allows the measure of information about a continuous quantity to grow beyond all limits (cf. 3.51b).

and controllably — even under the influence of 'normal' perturbations, or during observation. Robust quantities, represented by such subspaces, define 'approximate constants of the motion' (or adiabatically changing collective variables).

Since this definition of robustness is based on qualitative aspects, the related concepts cannot usually be defined with mathematical rigor in realistic cases. For example, the positions and shapes of droplets forming in a condensation process, or even more so those of the walls of a vessel, are evidently robust properties, although they do not represent exact constants of the motion. A certain freedom in the choice of macroscopic quantities remains, such as whether or not fluctuations are to be taken into account. One may even speculate about the possibility of replacing the apparently universal concept of thermodynamical entropy with a hierarchy of similar concepts (depending on various degrees of robustness, defined by their relaxation time scales, for example).

An actual microscopic trajectory $q(t)$ determines all macroscopic ones that are defined as functions of state, $\alpha(t) := \alpha(p(t), q(t))$. As discussed in Sect. 3.1.2, the macroscopic dynamics is in general not deterministic by itself (not autonomous), since different trajectories starting from the same $\alpha(t_0)$ may lead to different $\alpha(t_1)$ — depending on the microscopic state $p(t_0), q(t_0)$ which happens to represent $\alpha(t_0)$ in reality. This *macroscopic indeterminism* is relevant, in particular, for fluctuations or phase transitions.

In contrast, the determinism of a dynamical *model* (such as Laplacean mechanics) is defined by the mere *existence* of a unique mapping of appropriate initial (or final) states onto complete trajectories. This concept of determinism is independent of the availability of an (analytic or algorithmic) *procedure* for explicitly constructing these trajectories in terms of conventional coordinates ('integrability'). It is therefore also independent of any *practical* limitation in computability, which forms the basis of Kolmogorov's (1954) entropy and is often used for the definition of *chaos* (see Schuster 1984, or Hao-Bai-Lin 1987). In classical mechanics, this deterministic dynamical mapping is generically *defined* by Newton's equations (cf. Bricmont 1996 for his lucid criticism of the popular misuse of the concept of chaos in this connection).

Trajectories could be described in principle in a trivial way in terms of the constants of motion. The latter (to be used as new coordinates or 'co-evolving grids' — see App. B of Zurek 1989) are mathematically singular (and therefore often denied to exist) only in their *relation* to conventional (local) coordinates, but not in any absolute sense. It was indeed one of the great lessons from the theory of relativity that physics and spacetime geometry ('reality') are independent of the choice of coordinates, while the ancient Greeks were not even able to overcome Zeno's paradox of Achilles and the tortoise by a transformation to more appropriate 'coordinates' of description. We should similarly be able to overcome all mathematical limitations in the construction of canonical transformations, and again rely on the assumption of a coordinate-free 'reality' (at least in classical mechanics).

These mathematical difficulties may nonetheless reflect the complex and non-trivial *physical* relation between the universe and its 'observing parts'. Observers are evidently *not* in any simple way related to the constants of the motion — the reason why we feel 'time flow'.[10] Some authors have related the problems of a universe that contains its observers (*physical* self-reference) to Gödel's undecidability theorems, which apply to logical systems that allow *formal* self-reference (see Wheeler 1979, Smullyan 1981). However, one cannot conclude that the assumption of an observer-independent reality is excluded just because of the latter's limited observability — as has even been suggested in order to explain quantum 'uncertainty' (Popper 1950, Born 1955, Brillouin 1962, Cassirer 1977, Prigogine 1980). There is a fundamental difference between the *impossibility of ever knowing* the precise classical state of the universe and the *incompatibility of its existence with certain empirical facts*. While the former is often derived precisely by using classical concepts, the latter is a consequence of the crucial experiments.

3.3.2 Information Based on Thermodynamics

Macroscopic indeterminsim, represented by Einstein's probabilities for fluctuations (3.52), may in principle give rise to a decrease of *physical* entropy $S(\alpha)$ in accordance with microscopic determinism. This requires a transformation of irrelevant into relevant information, which is illustrated by the first step of Fig. 3.5 for the case of a measurement of a (classical) microscopic quantity. Problems may arise if the physical realization of information by documents, computers or brains is here not properly taken into account.

As is well known since the days of *Maxwell's demon*, any change of information must be described as a physical process together with its thermodynamical consequences. He assumed this demon to operate a microscopic sliding door between two compartments in such a way that only fast molecules may enter the first compartment, but slow molecules the second one. His action would then lead to a temperature and pressure difference, and thus admit the construction of a perpetuum mobile of the second kind.

Evidently, the demon has here to invest *his* information about the momenta and positions of the molecules. Smoluchowski (1912) objected that the demon, as a physical system, would himself have to obey the Second Law: his 'operational' description has to be replaced with a dynamical one. In phenomenological terms, the lowering of the entropy of the gas, corresponding to the arising temperature difference, must at least be compensated for by an increase in the demon's entropy. If the demon is assumed to be a finite

[10] "Time goes, you say? Ah no! Alas, time stays, we go." (Austin Dobson – discovered in Gardner 1967)

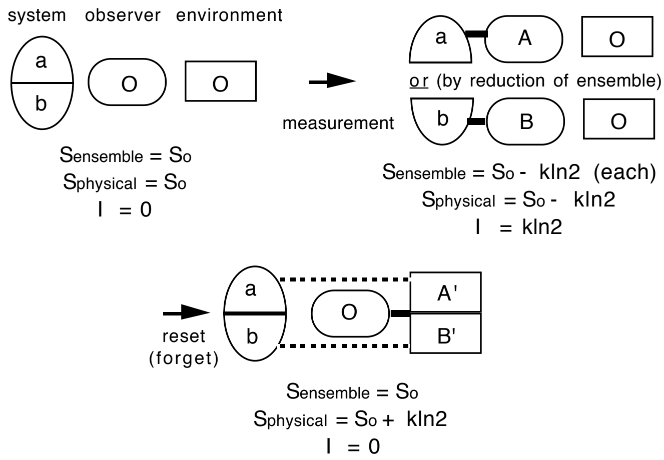


Fig. 3.5. Entropy relative to the status of information during a classical measurement. In the first step in the figure the state of the observer changes depending on that of the system. The second step represents the subsequent reset of the 'observer' or device (Bennett 1973), required in order to repeat the measurement. Areas represent sets of microscopic states of the subsystems (while those of uncorrelated combined systems would be represented by their direct products). The lower case letters a and b characterize the property to be measured; 0, A and B the corresponding 'memory states' of the observer, while A' and B' are their effects in the thermal environment, required for a deterministic reset. The 'physical entropy' (*defined* to add for subsystems) measures the phase space of all microscopic degrees of freedom, including the property to be measured. Because of this presumed additivity, the physical entropy neglects statistical correlations (dashed lines, which indicate *sums* of products of sets) as being 'irrelevant' in the future — hence $S_{\text{physical}} \geq S_{\text{ensemble}}$. I is the amount of information held by the observer. S_0 is at least $k \ln 2$ in this simple case of two equally probable values a and b. (From Chap. 2 of Giulini et al. 1996)

and thermodynamically closed system, his increasing Brownian motion would ultimately prevent him from acting properly (by making his 'hands tremble').

The increase in the demon's entropy may also be interpreted as a decrease in his information about the gas molecules. Therefore, from this *gedanken* experiment with a physical demon, Szilard (1929) extracted a fundamental information-theoretical aspect. In accordance with Maxwell's demon he assumed that an 'intelligent being' must be in possession of the required information about the molecules, and able to utilize it. Szilard concluded that, in order to lower the entropy of the gas by an amount $-\Delta S > 0$, the demon needs the minimum information

$$\Delta I = -\frac{\Delta S}{k} \quad , \tag{3.55}$$

compatible with Boltzmann-Gibbs' interpretation of entropy, or Einstein's expression (3.52).

Szilard's main argument was based on a model 'gas' consisting of a single molecule in a vessel of volume V. Thermodynamical aspects are introduced by means of collisions of the molecule with the walls, leading to equilibration of the molecule's motion with a surrounding heat bath (see Fig. 3.6). A piston is then vertically inserted without using energy in order to separate two partial volumes V_1 and V_2. This partition of the volume is *robust* in the sense of Sect. 3.3.1. According to (3.54) this procedure therefore transforms part of the entropy of the 'gas' into entropy of *lacking information*. If the experimenter knows (only) in which partial volume i the molecule resides, corresponding to a Shannon measure $\Delta I_i = \ln[(V_1 + V_2)/V_i]$, with $i = 1$ or 2, he is able to retrieve the mechanical energy

$$\Delta A_i = \int_{V_i}^{V_1+V_2} p\,dV = \int_{V_i}^{V_1+V_2} \frac{kT}{V}\,dV = -kT \ln \frac{V_i}{V_1 + V_2} \tag{3.56}$$

by allowing the piston to move away from the region where the molecule is. The molecule's mean kinetic energy may thereby be kept constant by the reversible transfer of heat from the external reservoir with temperature T. This process lowers the entropy of the reservoir by an amount

$$\Delta S_i = -\frac{\Delta A_i}{T} = k \ln \frac{V_i}{V_1 + V_2} = -k\Delta I_i \quad , \tag{3.57}$$

in accordance with (3.55).

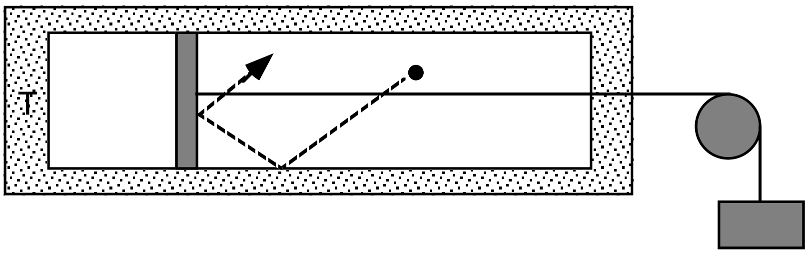


Fig. 3.6. Szilard's *Gedanken* engine completely transforms thermal energy into mechanical energy by using information

According to Smoluchkowski, one could avoid referring to knowledge or information by using a 'mechanical rectifier' (such as a ratchet) that causes the piston to move in the appropriate direction. However, this rectifier would ultimately have to perform thermal motion large enough to make it useless, corresponding to the demon's trembling hands (see also Feynman, Leighton and Sands 1963, Vol. I, p. 46-1). One may conclude that utilizing knowledge for making decisions (for example in the brain) is equivalent to the operation of a 'rectifier'. It is here essential that the rectifier cannot be reset to its original state without getting rid of entropy — usually in the form of heat (Bennett 1987). For this reason the mechanism cannot work reversibly in a closed system.

Brillouin (1962), when elaborating on ideas originally presented by Gabor in lectures given in 1952 (see Gabor 1964), emphasized that Szilard's 'intelligent being' may have to acquire its information. Since this process must also be compatible with the Second Law, Brillouin postulated his *negentropy principle*

$$\Delta S' - k\Delta I \geq 0 \quad , \tag{3.58}$$

which states that any *gain in information* ΔI has to be accompanied by some process of dissipation leading to the production of thermodynamical entropy $\Delta S'$. Gabor and Brillouin exemplified this by the transfer of entropy to the information medium (usually light). Thereby they referred to its quantum aspect (photons), since classical light would represent an infinite entropy capacity. Because of the minimum information required according to Szilard, the construction of a perpetuum mobile of the second kind would be excluded. However, the relation (3.58) is also confirmed by the above example of a directly coupled mechanical rectifier, without any explicit reference to an information medium.

All non-phenomenological arguments are here based on two assumptions: (1) Global determinism, which requires that an ensemble of N different states (or N ensembles of equal measure) must have N different successors, which have to be counted by the total ensemble entropy. Different states may evolve into the same final state only by an appropriate interaction with the environment that transfers this difference to the latter (for example, in the form of heat). (2) Intuitive causality, which asserts that uncontrollable 'perturbations' by the environment can only *enlarge* the ensemble. It gives rise to the *in*equality sign in equations such as (3.58). If thermodynamical concepts apply, the transfer of ensemble entropy ΔS must be accompanied by a transfer of energy according to $\Delta Q = T\Delta S$. This conclusion has led to the interpretation of entropy as a measure of *degradation of energy.*

This equivalence of information and negative entropy also suggests that pure information *processing* (in a computer) can in principle be performed reversibly. However, arithmetical operations are often *logically irreversible* in the sense that two factors cannot be recovered from their product. (In a mechanical computer this operation may require friction.) In the theory of computing this led to the conjecture that a minimum entropy $k \ln 2$ has to be produced for each bit of information in each elementary calculational step (Landauer 1961). It was refuted by Bennett (1973 — see also Bennett and Landauer 1985). In their analysis, the logically lost information ('garbage bits') — even if randomized — is still regarded as 'relevant' in the thermodynamical sense (that is, not yet as heat). For this reason, the entropy creation is deferred to the reset or clearing of the memory which is required in order to allow the computer to work reversibly (see the second step of Fig. 3.5).

All these conclusions support the view that information has to be physically realized (and therefore to be compatible with the laws of thermodynamics), rather than representing an extra-physical concept that has to be *presumed* in the statistical foundation of thermodynamics. On the other hand, *mathematical* theorems do *not* represent information (as, for example, assumed by Landauer 1996), since logic deals exclusively with tautologies (or 'analytical judgements') from an epistemological point of view — though in mathematics with highly complex and nontrivial ones.

General literature: Denbigh and Denbigh (1985), Bennett (1987), Leff and Rex (1990)

3.4 Semigroups and the Emergence of Order

In physical systems, 'order' is represented by states of low entropy. For example, the rectifier used in the previous section for representing Maxwell's demon must display ordered (regular) behavior. The emergence of order from disorder in nature, also called *self-organization of matter*, may appear to contradict the Second Law with its general trend towards disorder or *chaos*. This has often been misinterpreted as a 'discrepancy between Clausius and Darwin'. However, the fundamental phenomenological equation (3.1) permits entropy to decrease *locally*. Its first term allows physical entropy (again *defined* as an extensive quantity by means of $\hat{P}_{local}$) to be transferred to the environment. If this environment is not in complete thermal equilibrium, and is characterized by at least two different temperatures, T_1 and T_2, a local decrease of entropy $dS_{\text{ext}} = dQ_1/T_1 + dQ_2/T_2 < 0$ does not even require any net transfer of heat, compatible with $dQ_1 + dQ_2 = 0$. (The sign of dS is here always meant to refer to positive time increments dt.) This local decrease of entropy is thus *not* in conflict with its global increase according to the Second Law (see also Sect. 5.3).

In statistical terms, the number of states in a *dynamically representative ensemble* (cf. Sect. 3.1.2) may decrease in accordance with determinism *and* intuitive causality, provided the ensemble characterizing the state of the environment increases accordingly — precisely as during the 'reset' of a memory device, indicated in Fig. 3.5. In Laplacean description, the outcome of evolution would be determined by the microscopic initial state of the whole universe.

An important special case is the situation of stationary non-equilibrium (*Fließgleichgewicht*), characterized by $dS = dS_{\text{int}} + dS_{\text{ext}} = 0$ *and* positive entropy production, $dS_{\text{int}} > 0$, (Bertalanffi 1953). It may support ordered states as *dissipative structures*. The standard example, known as *Bénard's instability*, describes convective heat transfer through a thin horizontal layer of a liquid in the form of ordered convection cells, which optimize the process of thermal equilibration between two reservoirs of different temperatures. In a finite universe, this stationary situation can only represent a transient local phenomenon. The emergence of structure is often connected with the breaking of symmetries (in particular the symmetry under translations), similar to a phase transition. In deterministic description, an initial microscopic asymmetry is thereby amplified to the macroscopic scale.

For similar reasons, Boltzmann suggested that biological processes here on earth are facilitated by the temperature difference between the sun (with its 6000 K surface temperature) and the dark universe (at 2.7 K, as we know today). At the earth, the solar radiation has an energy density much lower than that of a black body with the same spectrum (temperature). Since photon number is not conserved (not even robust), a canonical distribution $\exp(-H/kT)$ in the occupation number representation does not only determine the spectral distribution as a function of temperature, but also the intensity (photon density). A gas with conserved particle number would instead allow one to choose the mean density independently — either by fixing the particle number by closing the vessel, or by fixing the chemical potential (in a grand canonical ensemble) by connecting the vessel to a particle reservoir. A photon from the sun can be very efficiently transformed into many soft photons, which together possess much higher physical entropy.

While order may appear to be an objective property (in contrast to information), an absolute concept of order that is not simply defined by means of phenomenological entropy is as elusive as an objective concept of information or relevance (see Denbigh 1981, p.147, or Ford 1989). Low entropy, on the other hand, may characterize complex as well as simple states (which do not have any potential for order). For reasons similar to those discussed towards the end of Sect. 3.3.1, the definition of order in terms of 'computability' would again depend on the choice of 'relevant' variables. For example, the obvious order observed in a crystal lattice is not invariant under canonical transformations in their most general sense. How, then, may the order of an organism be conceptually and *absolutely* distinguished from the 'chaotic' correlations arising from molecular collisions in a gas? Robustness may represent an essential attribute of ordered complexity. One could alternatively define correlations operationally as representing order if they have the potential of causing local macroscopic *effects* (such as interference in X-ray scattering from an atomic lattice). This would readily describe the sudden 'decay of order' in a dying organism, but it would *presume* an absolute direction in time — related to that of intuitive causality.

Many self-organizing systems include chemical reactions. They are then phenomenologically described by irreversible rate equations which determine the change of concentrations $X, Y, \ldots$. These concentrations are 'macroscopic' in the sense of Sect. 3.2 or 3.3.1 if they are robust. In statistical terms, rate equations form a generalized *Stoßzahlansatz* that includes rearrangement collisions between different kinds of molecules, which are assumed to be in partial equilibrium (at the same temperature T). These rate equations are therefore special master equations (as derived in Sect. 3.2) for these 'relevant' degrees of freedom $X, Y, \ldots$.

Rate equations determine trajectories in the configuration space of concentrations (here representing the macroscopic variables α of Sect. 3.1.1).[11] For closed systems these trajectories may eventually approach that point in their configuration space which describes equilibrium. Symmetric determinism must come to an end at such *attractors* (see Fig. 3.7a), although this may require infinite time. A *mechanical* example of an attractor in the presence of friction is provided by the phase space point $v := dx/dt = 0$ and $V(x) = V_{\mathrm{min}}$. The corresponding equation of motion, such as $dv/dt = -av - \mathrm{grad}V$, neglects any stochastic response from the energy-absorbing degrees of freedom (which would in principle be required for non-vanishing temperature by the *fluctuation-dissipation theorem*). Similar to the Lorentz-Dirac equation of Sect. 2.3, this equation is deterministic (for finite times) even though it is asymmetric under time reversal.

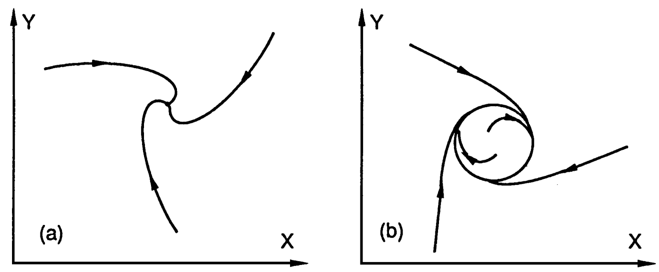


Fig. 3.7a,b. Standard representation of an attractor (a) and a limit cycle (b) as examples of phenomenologically irreversible dynamics in the configuration space of macroscopic variables $\alpha \equiv X, Y$

Points in the space of macroscopic variables X, Y or x, v (macroscopic 'states' α, in general) describe the physical states incompletely. They represent large subspaces of the complete Γ-space (that may here include the environment). Volume elements of equal size in the space of macroscopic variables may therefore possess very different ensemble measure. Parameter space volume is not dynamically conserved in general — neither exactly nor under a rate equation. For example, the immediate vicinity of an equilibrium 'state' X_0, Y_0 (such as $v = 0$, $V(x) = V_{min}$ in the mechanical example) covers most of the whole Γ-space of the complete system (or some subspace representing conserved quantities) in the sense of the fundamental measure. A parameter space should therefore be clearly distinguished from the fundamental space of *elementary states.*

[11] As the rate equations are of first order in time, this configuration space is often called a phase space.

In the specific mechanical example of friction, the *modified* macroscopic phase space measure $dx\,dv/v$ nonetheless happens to be dynamically invariant. Time reversal can here be compensated for by the transformation $v \to 1/v$ in order to restore a formal T-symmetry (similar to the 'symmetric violations' of time-reversal invariance mentioned in the Introduction). The remaining symmetry is reflected in a conserved *generalized H-functional* (cf. (3.50) and (3.51b)),

$$H_{gen} := \int \varrho(v,x) \ln[|v|\, \varrho(v,x)]\, dv\, dx \quad . \tag{3.59}$$

This defines a *reference density* $\varrho_0 = |v|^{-1}$ that describes a formal equilibrium measure on this macroscopic phase space. In situations of quasi-stationary non-equilibrium, macroscopic trajectories described by effective irreversible equations of motion may approach certain closed curves which do *not* correspond to maximum entropy. They may represent dissipative structures (see Fig. 3.7b), and are called *limit cycles*.

Open systems are also described by means of phenomenological *semigroups*, defined as *dynamical maps* acting on ensembles. These maps are essentially equivalent to time-integrated master equations, and thus again applicable only in the 'forward' direction of time (in contrast to the reversible group of time translations, valid for dynamically closed systems). Occasionally, such ensembles are regarded as the *fundamental* kinematical objects of the theory, which need not consist of any *elements* that would describe reality completely. 'Determinism' is then understood as a forward determinism for these formal ensembles. *Maps* are called irreversible if they form genuine semigroups, that is, if they cannot be uniquely inverted *as maps on ensembles*.

This irreversibility of maps may not only represent a dynamical indeterminism of elementary states: it also represents resets and attractors (decreasing entropy).[12] In order to describe a reset, the master equation (3.46a) has to be based on a *non-Hermitean* Zwanzig projector that 'creates' relevant information in the formalism. In a globally deterministic context, its microscopic realization must contain some way of getting rid of information (as discussed in Sect. 3.3.2). As can be seen from the second step of Fig. 3.5, the reset transforms local information into nonlocal correlations (also depicted in Fig. 3.1), where it is particularly irrelevant (uncontrollable). This transformation is interpreted as a *production of physical entropy*, while the ensemble entropy is conserved. The presumed absence of 'conspiratorial' correlations

[12] Therefore, some authors have proposed a redefinition of entropy that would *always* allow subsystem entropy to grow under a semigroup, even when local entropy is reduced in accordance with (3.1) — see Mackey (1989). For example, the *relative entropy of a system* (cf. 3.51b) with respect to a canonical distribution with temperature of an external heat bath would increase even when the temperature of the system is *lowered* by thermal contact. Such a redefinition is certainly physically misleading, even though it may be useful for certain purposes.

(which could dynamically reduce physical entropy) describes the irreversibility of the semigroup in microscopic terms.

As semigroups are defined to act on ensembles (which represent incomplete information), their inversion does in general *not* represent a reversal of the microscopic dynamics ('time reversal'). Similarly, their forward determinism is not equivalent to microscopic determinism. A map may not be invertible *as a map* even though the underlying dynamical transformation of microscopic states can be reversed (not necessarily uniquely — just as the individual forward evolution is here not necessarily presumed to be uniquely determined).

Individual indeterminism and attractors are illustrated on a finite set of states in Fig. 3.8. An *asymmetric* dynamical indeterminism (b) is represented by a diverging fork (cf. Footnote 7), while an attractor is characterized by a converging (or 'inverse') fork (c). The indeterminism characterizing an incompletely defined Hamiltonian is dynamically symmetric (a). On a continuum of states, a measure would first have to be *chosen*, usually by means of its invariance under an assumed *fundamental* deterministic dynamics of the closed system. (This may represent a problem if determinism is to be given up fundamentally.) Semigroups may be conveniently studied on various discrete state spaces (where measures of states are trivial), for example by means of the model of 'deterministic cellular automata' (cf. Kauffman 1991). Their merging trajectories (representing attractors) are equivalent to the shrinking of phase space volume elements in the continuum model with friction that led to the generalized H-functional (3.59).

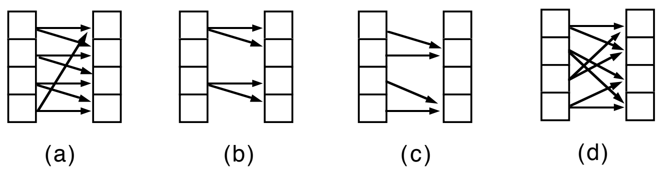


Fig. 3.8. Dynamical transformation of states on a discrete and finite 'phase space' consisting of four states only: (a) T-symmetric indeterminism (representing an incompletely determined Hamiltonian); (b) asymmetric indeterminism, representing a law-like increase of ensemble entropy (cannot be defined everywhere on finite sets); (c) attractor (reset); (d) discrete caricature of a Frobenius-Perron map (see text). The symmetric indeterminism (a) would appear asymmetric (similar to (b)) when *applied to* a low-entropy initial ensemble (for example, an individual state) in a given direction of time. It would then describe the usual increase of ensemble entropy by uncontrollable 'perturbations'. The distinction between (b) and (c) requires an absolute direction of time

Forward-deterministic dynamical maps are often defined by means of nonlinear transformations of individual states (points in phase space). A popular (though not very physical) one-dimensional toy model of a semigroup is the *Bernoulli shift*, defined by the mapping $\alpha \rightarrow 2\alpha \; mod \, 1$ on the interval $(0, 1]$. (Its α-measure is invariant under translations in α.) The dynamical increase of this 'phase space volume' $d\alpha$ by multiplication by 2 in this map can be uniquely inverted on the continuum although it represents a physical indeterminism. The second term, $mod1$, characterizes the semigroup property as in Fig. 3.8c. Both dynamical parts are here combined in order to form a *Frobenius-Perron map*, defined everywhere on the interval in spite of representing a genuine semigroup (symbolically indicated on the discrete set by Fig. 3.8d). Multiplication by 2 has been added to the semigroup operation $mod1$ in order to compensate for the decrease of ensemble entropy. The forward indeterminism, obvious in the discrete case, is often overlooked on the continuum, where the topology-conserving stretching of 'phase space' α may appear deterministic without a measure. However, the 'topological time asymmetry' represented by the Frobenius-Perron map may be phenomenologically relevant, as it is able to describe the formation of macroscopic diversity. Realistic attractors must be of mixed type in order to comply with the fluctuation-dissipation theorem.

Many similar dynamical maps are discussed in the literature. They are (at most) of phenomenological value, and have little explanatory power from a fundamental statistical point of view. Their investigators often seem to regard the underlying individual microscopic dynamics as irrelevant. The ensembles being mapped dynamically are then treated as *real states* of physical objects. This must, of course, lead to confusion from a fundamental point of view. 'Statistical theories' based on dynamical maps are occasionally even used in a 'minimal' interpretation of quantum mechanics (see Sect. 4.4). The misuse of purely formal ensembles as describing *physical* states is thereby reversed by identifying wave functions (that is, elementary quantum mechanical *states*) with ensembles. However, the conclusion that quantum phenomena cannot be explained in any such 'simple way' was already drawn by Bohr before the advent of matrix and wave mechanics (when his theory with Kramers and Slater had failed).

The description of thermodynamical systems far from equilibrium (cf. Glansdorff and Prigogine 1971) remains usually phenomenological. This unsatisfactory situation may partly be due to the fact that spontaneous symmetry breaking, related to the emergence of new structures, seems to be based on the *fundamental quantum mechanical indeterminism* (see Chap. 4 and Sect. 6.1). The onset of structure is phenomenologically described by means of unstable fluctuations of certain quantities α, whose probabilities are calculated by means of Einstein's formula (3.52). An instability arises for them when the second derivative $\partial^2 S/\partial\alpha^2$ at a stationary point of $S(\alpha)$ become negative in value during the adiabatic change of an external parameter. In this way, new *robust* quantities in the sense of Sect. 3.3.1 (cf. also Sect. 4.3.2) emerge, and physical entropy is transformed into entropy of *lacking information*, distinguished according to (3.54).

General literature: Glansdorff and Prigogine 1971, Haken 1978, Cross and Hohenburg (1993)

I shall close this chapter with the discussion of an objection against the probability interpretation of entropy when applied to the whole universe and its evolution. It was first raised by Bronstein and Landau (1933), and later in a more explicit form by von Weizsäcker (1939).

Obviously, the present state of the universe does not only possess an entropy $S_{\alpha(\mathrm{now})}$ that is much smaller than its equilibrium value S_{equil}, but it also contains documents which consistently indicate that the entropy has been increasing during the past, $dS_{\alpha(t)}/dt > 0$. One may now compare the probability for the formation of such documents (including our private memories) from a mere chance fluctuation with that for their presumed evolutionary formation in a historical process, that seems to be supported by these documents. In the latter case, one has

$$S_{\alpha(\mathrm{yesterday})} < S_{\alpha(\mathrm{now})} \ll S_{\mathrm{equil}} \quad . \tag{3.60}$$

However, if Einstein's measure of probability in terms of entropy (3.52) were applicable to the universe, the formation of its present state in a chance fluctuation — as improbable as it may intuitively appear — would be far more probable than its evolution from a state with much lower entropy in the distant past. This objection undermines Boltzmann's explanation of the thermodynamical arrow of time as arising from a grand fluctuation that occurred in an eternal universe (see Sect. 5.3), since this fluctuation could be replaced with a more probable smaller one.

The argument just described requires that the left inequality of (3.60) remains valid if not only local ('physical') entroopy, but also those non-local correlations which represent the convincing *consistency of documents about history* are taken into account. Their existence in a historical universe is equivalent to Lewis' 'overdetermination of the past' (cf. Footnote 1 of Chap. 2).[13] Nonetheless, David Hume's fundamental insight that we can never *predict* anything with certainty (not even that the sun will rise again) applies to the past as well. The reliability of memories and documents is in principle as doubtful as that of predictions; only the *local present* (of which we are subjectively aware) cannot be questioned. Hence, even Kant's premise that we *are* making experience cannot be taken for granted. Not what has been observed, only our (perhaps deceiving) 'memory' is beyond doubt. St. Augustine already concluded in his *Confessiones* that the past and the future 'exist' only in the present — namely as memory and expectation 'in the soul'. This long-standing philosophical debate seems to be deeply affected by thermodynamical considerations.

St. Augustin's epistemologically rigorous concept of reality, that applies only to the present, is evidently too restrictive to form a 'world model', even though any generalization must remain hypothetical in principle (Poincaré 1902, Vaihinger 1911). The probabilistic objection against the evolutionary

[13] States containing consistent (though possibly deceiving) documents were called *time-capsules* by Barbour (1994a) — see Sect. 6.2.2.

picture of the world, even if formally correct, will thus hardly be accepted as demonstrating that causality is only an illusion, based on an accident. Einstein's probabilities (3.52) for the occurrence of non-equilibrium states α, based on the statistical interpretation of entropy, can indeed be justified only for those macroscopic properties α which occur repeatedly within relevant times ('quasi-ergodically') on a generic trajectory — that is, for properties which are not robust on relevant timescales.

Physical cosmology can fortunately be based on the more *economic* hypothesis of a universe of finite age. A homogenous (structureless) low entropy initial state appears as a more plausible assumption than a complex state with the same value of entropy. Probabilities for later states have then to be calculated as *probabilities for histories* that lead to them. For example, the folding of protein chains is calculated along trajectories of monotonically increasing entropy (according to a master equation). Final configurations not accessible through such histories would thus be excluded. (Quantum mechanically there is always a small tunneling probability.) Most probable are then those final states that are accessible through the most probable histories, but not those at an absolute entropy maximum. This picture would *explain* consistent documents.

Similar ideas have been applied to biological evolution by Lloyd and Pagels (1988). Whether the life-carrying situation of our world, including the existence of scientists observing it, is probable in this sense (that is, whether most of the microscopic states which are compatible with the initial conditions would lead to similar situations), or whether further 'anthropic' selection criteria are required, has never been estimated in a reliable and unbiased way. Only at a tremendously greater age of the universe than its present one could a state of maximum entropy be reached via improbable intermediate states or through quantum mechanical tunnelling (Dyson 1979). The early age of the universe would then remain the major unexplained improbable fact.

A 'plausible' low-entropy initial state of the universe will be considered in Sect. 5.3. To be realistic, a quantum mechanical description has to be used for this purpose. Quantum indeterminism, whatever its correct interpretation (see Sect. 4.6), may even allow the assumption of a *unique* 'initial' state of the universe (with a small entropy capacity) — see Chap. 6. However, the outcome of evolution (including ourselves) has to be presumed *as a possibility* in the configuration space defined by the fundamental kinematical concepts, regardless of all probability arguments.

4. The Quantum Mechanical Arrow of Time

– Prelim. 4th edtn. (Mar 01): www.time-direction.de

The dynamics of probability distributions on classical phase space, discussed under various aspects in Chap. 3, may be formally translated into quantum mechanics by means of the canonical quantization rules. Many authors of standard textbooks therefore maintain that the foundation of irreversibility in quantum mechanics is the same, in principle, as in classical physics. There could then only be quantitative differences arising from different spectral properties of the 'corresponding' Liouville operators. However, this approach to statistical quantum mechanics completely ignores the fundamentally different interpretation of concepts that *formally* correspond to one another (such as probability distributions and density operators — see Sect. 4.2). Therefore, it conceals important properties of quantum theory (compare the Introduction), which may be essential for irreversibility in general, *viz.*:

1. The quantum mechanical probability interpretation contains an *indeterminism* of controversial origin. Most physicists seem to regard it as representing an objective and law-like *dynamical* indeterminism (cf. Fig. 3.8), and some even as an extra-physical master arrow of time. Others have instead suggested that one may *explain* the unpredictability of quantum mechanical measurement results (or measurement-like events) in terms of conventional statistical arguments, that is, by means of thermodynamical fluctuations which occur during the required amplification process in the measurement device. If, however, this question is circumvented by interpreting the wave function as representing 'human knowledge as an intermediate level of reality' (Heisenberg 1956), Maxwell's demon, discussed in Sect. 3.3.2, may return through the quantum back door. Therefore, the foundation of irreversibility seems to be intimately related to the *interpretation* of quantum theory (see Sect. 4.6).
2. The quantum theory is *kinematically nonlocal*. The generic many-particle wave function $\psi(\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_N)$, which represents a 'pure' quantum state, describes *quantum correlations* between subsystems. They are *not* due to incomplete information (even though they may *lead to* such statistical correlations in measurements). Similarly, a state of quantum field theory is given by a wave functional of fields which are defined all over space. This 'entanglement' is a direct consequence of the superposition principle. In quantum theory, *the state of the whole generically does not define states of its parts.* This is in fundamental contrast to the completely determined

many-particle state of classical mechanics: a point in phase space remains a point when projected onto a subsystem. This kinematical indeterminacy of the parts describes a non-trivial 'wholeness' of nature, which cannot, as in classical physics, be interpreted as a mere *dynamical* interconnectedness (that may lead to *statistical* correlations in ensembles). Moreover, it has nothing to do with Heisenberg's uncertainty (or 'indeterminacy') relations, which signal the limited validity of *classical* concepts for describing physical states. The uncertainty relations apply even when individual quantum states *are* defined for all subsystems. Therefore, the Zwanzig projection of locality (3.37) has in general a nontrivial effect when applied to pure states; it defines non-vanishing local ('physical') entropy even for a completely determined ('real') state of the whole.

These basic differences between classical and quantum statistical physics will be discussed after their formal analogy has been set up in Sect. 4.1.

4.1 The Formal Analogy

4.1.1 Application of Quantization Rules

The *formal* transition from classical to quantum statistical mechanics, defined by 'quantization rules', replaces functions of state, $f(p,q)$, with 'corresponding' operators $F = f(P,Q)$, and Poisson brackets between them with commutators. In this way, the Liouville equation (3.26) becomes

$$\mathrm{i}\frac{\partial \varrho_\Gamma}{\partial t} = \mathrm{i}\{H, \varrho_\Gamma\} =: \hat{L}\varrho_\Gamma \quad \rightarrow \quad \mathrm{i}\frac{\partial \varrho}{\partial t} = [H, \varrho] =: \hat{L}\varrho \quad . \tag{4.1}$$

It is then called the *quantum Liouville* or *von Neumann equation*. The classical probability densities $\varrho_\Gamma(p,q)$ are thus replaced with *density operators* ϱ. The caret $\hat{}$ is here meant to distinguish the new operators, which act on the quantum mechanical Hilbert space operators (such as the density operators), from these Hilbert space operators themselves. In the formal analogy, the new 'superoperators' (as they are sometimes called) correspond to the operators that were defined in Sect. 3.1.2 as acting on probability densities. The Hilbert space operators form again a Hilbert space if an inner product $\langle\langle A, B\rangle\rangle := \mathrm{Trace}\{A^\dagger B\}$ is defined for them, in analogy to the inner product $\langle a, b\rangle = \int a^*(p,q)b(p,q)\,dp\,dq$ that applies to classical probability densities (cf. the text above (3.27)).

Furthermore, all mean values $\bar{f}$ for functions of state $f(p,q)$ with respect to the probability densities $\varrho_\Gamma(p,q)$ have to be replaced with *expectation values* $\langle F\rangle$,

$$\bar{f} := \int f(p,q)\varrho_\Gamma(p,q)\,dp\,dq \quad \rightarrow \quad \langle F\rangle := \mathrm{Trace}\{F\varrho\} \quad . \tag{4.2}$$

However, in spite of this formal analogy, a density operator can *not* be interpreted simply as representing an ensemble of states that would define ensembles of values for all functions of state (see below). Since (4.2) implies

$$\overline{\ln \varrho} \quad \rightarrow \quad \langle \ln \varrho \rangle = \mathrm{Trace}\{\varrho \ln \varrho\} \quad , \tag{4.3}$$

the quantum mechanical entropy functional corresponding to the ensemble entropy S_Γ becomes *von Neumann's entropy*,

$$S[\varrho] := -k\mathrm{Trace}\{\varrho \ln \varrho\} \quad . \tag{4.4}$$

The dynamics of the statistical operators, defined by the right equation (4.1), is unitary, with a formal solution

$$\varrho(t) = U(t)\varrho(0)U^\dagger(t) \quad , \tag{4.5}$$

and $U(t) = \exp(-\mathrm{i}Ht)$ for time-independent Hamiltonians. This warrants conservation of the von Neumann entropy under the von Neumann equation,

$$\begin{aligned}\mathrm{Trace}\{\varrho(t)\ln \varrho(t)\} &= \mathrm{Trace}\{U(t)\varrho(0)U^\dagger(t)U(t)\ln \varrho(0)U^\dagger(t)\} \\ &= \mathrm{Trace}\{\varrho(0)\ln \varrho(0)\} \quad .\end{aligned} \tag{4.6}$$

As *classical* determinism (the conservation of probabilities for individual trajectories) can similarly be described by the unitary time-dependence (3.26), the formal argument in (4.6) could also be used in classical ensemble mechanics.

The square of the Hilbert space norm of a density operator,

$$\|\varrho\|^2 := \langle\langle \varrho, \varrho \rangle\rangle = \mathrm{Trace}\{\varrho^2\} = \langle \varrho \rangle \quad , \tag{4.7}$$

defines again a *linear measure of (neg-)entropy* that is also conserved under the unitary dynamics (4.1). This entropy is often normalized according to $S_{lin} = \langle (1-\varrho) \rangle$ (such that $0 \leq S < 1$). In contrast, the *probability norm*, $\mathrm{Trace}\{\varrho\} = \langle 1 \rangle = 1$, characterizes the Banach space of *trace class operators* (those with a non-vanishing trace). This formal space is popular in open systems quantum mechanics (Sect. 4.4), since total probability must be conserved even under phenomenological stochastic equations of motion that describe an increase of ensemble entropy, S or S_{lin}.

In further formal analogy to classical ensemble mechanics, any coarse-grained (or relevant) information measured by $\mathrm{Trace}\{(\hat{P}\varrho)\ln(\hat{P}\varrho)\}$ is in general *not* conserved. Zwanzig projection operators $\hat{P}$ are here again idempotent operators on the Hilbert space of density operators, with additional properties $\mathrm{Trace}\{\hat{P}\varrho\} = 1$ and positive $\hat{P}\varrho$ for all ϱ, as in Sect. 3.2.

Statistical operators (that is, *density operators*) ϱ may be represented by various ensembles of wave functions ψ_α with probabilities p_α in the form $\varrho = \sum_\alpha |\psi_\alpha\rangle p_\alpha \langle \psi_\alpha|$ (see Sect. 4.2). In the time-dependent eigenbasis of $\varrho(t)$, where the states $\psi_\alpha(t)$ form an orthonormal set, $\|\varrho\|^2$ is given by the invariant sum $\sum_\alpha p_\alpha^2$. Its conservation (or that of $\langle \ln \varrho \rangle = \sum p_\alpha \ln p_\alpha$) thus represents the *individual* conservation of these diagonal elements in the moving basis in analogy to the conservation of a co-moving phase space volume in deterministic classical mechanics.

The matrix elements of the density operator with respect to a *random* basis $\{\phi_n\}$,

$$\varrho_{mn} = \sum_\alpha \langle \phi_m | \psi_\alpha \rangle p_\alpha \langle \psi_\alpha | \phi_n \rangle \quad , \tag{4.8}$$

are in general small for $m \neq n$ because of random phases in the sum over α. Pauli (1928) referred to this *random phase approximation* when he neglected off-diagonal matrix elements while deriving his master equation (4.17) below. However, they can in general be small only individually, while remaining essential as a whole. The neglect of off-diagonal elements *in a certain basis*,

$$\hat{P}_{\mathrm{diag}} \varrho_{mn} := \varrho_{mm} \delta_{mn} \quad , \tag{4.9}$$

defines instead the most important Zwanzig projection of quantum statistical mechanics, which regards these off-diagonal elements (or any interference between the states of this basis) as 'irrelevant'. It has nothing to do with the usual diagonalization of Hermitean operators in their eigenrepresentation. The inequality

$$\begin{aligned} \mathrm{Trace}\{\hat{P}_{\mathrm{diag}} \varrho \ln(\hat{P}_{\mathrm{diag}} \varrho)\} &= \sum_n \varrho_{nn} \ln \varrho_{nn} \\ &\leq \mathrm{Trace}\{\varrho \ln \varrho\} = \sum_{mn} \varrho_{mn} (\ln \varrho)_{nm} \end{aligned} \tag{4.10}$$

is called *Klein's lemma* — cf. (3.34). It is a consequence of the fact that $\hat{P}_{\mathrm{diag}}$ is a genuine projection operator (cf. Sect. 3.2).

An obvious generalization of (4.9) is

$$\hat{P}_{\mathrm{semidiag}} \varrho := \sum_n P_n \varrho P_n \quad , \tag{4.9$'$}$$

where $\{P_n\}$ (no caret!), with $P_m P_n = P_m \delta_{mn}$, is a complete set of projection operators on mutually orthogonal subspaces of the Hilbert space of quantum states. In quantum field theory, projections on 'unitarily inequivalent' separable subspaces of Hilbert space, sometimes even regarded as 'distinct Hilbert spaces', are often chosen for this purpose. However, these decompositions of non-separable Hilbert spaces are no less arbitrary than any other $\hat{P}_{\mathrm{semidiag}}$ (though often *useful* in the case of large numbers of effective degrees of freedom). If imposed axiomatically, the relevance concept (4.9$'$) represents a *superselection rule* (Wick, Wightman and Wigner 1952, Jauch 1968, Hepp 1972). This suggests that superselection rules may have a similar *dynamical* origin as the 'thermodynamically macroscopic' concepts of Chap. 3 — a possibility that will be further investigated and confirmed in Sect. 4.3.

4.1.2 Master Equations and Quantum Indeterminism

The Hamiltonian of a quantum mechanical system is often written in the form $H = H_0 + H_1$ in order to derive a master equation in terms of a perturbation expansion with respect to H_1. However, the main purpose of this split Hamiltonian is to define a relevance concept of type (4.9) or (4.9′) by means of the eigenbasis of H_0. It may then (but need not) *further* be used in a time-dependent perturbation expansion with respect to the off-diagonal elements of H in this representation (all contained in H_1).

The dynamics of $\hat{P}_{\text{diag}}\varrho$ is the dynamics of the diagonal elements of ϱ. According to (4.1) one has in any representation (now writing $\hat{P}_{diag} = \hat{P}$ for short)

$$\begin{aligned} \mathrm{i}\frac{d\varrho_{mm}}{dt} &= \sum_n (H_{mn}\varrho_{nm} - \varrho_{mn}H_{nm}) \\ &\equiv \sum_{n(\neq m)} (H_{mn}\varrho_{nm} - \varrho_{mn}H_{nm}) \mathrel{\widehat{=}} \hat{P}\hat{L}(1-\hat{P})\varrho \quad . \end{aligned} \tag{4.11}$$

Since the diagonal matrix elements of ϱ thus do not contribute to the RHS, the first term of Zwanzig's pre-master equation (3.42), representing $\hat{P}\hat{L}\hat{P}$, vanishes for this relevance concept. The terms remaining in (4.11) describe the coupling to the 'irrelevant' off-diagonal elements, and demonstrate that the diagonal elements can obey autonomous dynamics only in the trivial case (cf. Footnote 6 of Chap. 3, regarding the quantum Zeno effect). Because of the formal analogy, the rest of Zwanzig's method can then be applied, provided the required approximations are valid. The propagator $\exp[-\mathrm{i}(1-\hat{P})\hat{L}\tau]$ occurring in the operator $\hat{G}_{\text{ret}}$ of the Markovian approximation (3.46) defines here a closed (and therefore incomplete, though highly non-trivial) dynamics of the off-diagonal elements ϱ_{mn}.

Pauli's master equation is obtained from (3.46) and (3.43) by using a perturbation expansion in terms of the off-diagonal elements of the Hamiltonian for calculating $\hat{G}_{\text{ret}} = \int_0^T \hat{G}(\tau)\, d\tau$. These off-diagonal elements are thus assumed to be small, but not extremely small (the master equation would become trivial if they vanished exactly). This emphasizes again the *dynamical foundation of the concept of relevance.*

Consider the last three factors of the RHS of the integral kernel (3.43) when applied to ϱ,

$$(1-\hat{P})\hat{L}\hat{P}\varrho = (1-\hat{P})[H, \hat{P}\varrho] \mathrel{\widehat{=}} H_{mn}(\varrho_{nn} - \varrho_{mm}) \quad . \tag{4.12}$$

This expression depends only on the off-diagonal elements of H. The projection $1 - \hat{P}$ is ineffective, as $\hat{P}\hat{L}\hat{P} = 0$. Similarly, one has for the first three factors applied to any matrix X:

$$\hat{P}\hat{L}(1-\hat{P})X \mathrel{\widehat{=}} \sum_{k(\neq m)} (H_{mk}X_{km} - X_{mk}H_{km}) \quad . \tag{4.13}$$

Hence, $\hat{G}_{\mathrm{ret}}$ is at least of second order in the off-diagonal elements of H. When neglecting all higher orders according to Pauli, one may express the propagator $\exp[-\mathrm{i}(1-\hat{P})\hat{L}\tau]$ solely in terms of diagonal elements of H, $H_{mm} =: E_m^{(0)}$. This means

$$\mathrm{e}^{-\mathrm{i}(1-\hat{P})\hat{L}\tau} X \mathrel{\hat{=}} \mathrm{e}^{-\mathrm{i}(E_m^{(0)}-E_n^{(0)})\tau} X_{mn} \quad , \tag{4.14}$$

and one obtains

$$\begin{aligned}&\hat{P}\hat{L}(1-\hat{P})\mathrm{e}^{-\mathrm{i}(1-\hat{P})\hat{L}\tau}(1-\hat{P})\hat{L}\hat{P}\varrho \\ &\qquad \mathrel{\hat{=}} \sum_n |H_{mn}|^2 2\cos[(E_m^{(0)}-E_n^{(0)})\tau](\varrho_{mm}-\varrho_{nn}) \ .\end{aligned} \tag{4.15}$$

This result corresponds to a Born approximation in terms of the off-diagonal elements of the Hamiltonian. The time integral required to obtain $\hat{G}_{\mathrm{ret}}$ according to (3.46) leads to the resonance factor

$$\int_0^T \cos[(E_m^{(0)}-E_n^{(0)})\tau]\, d\tau = \frac{\sin[(E_m^{(0)}-E_n^{(0)})T]}{(E_m^{(0)}-E_n^{(0)})} \quad , \tag{4.16}$$

familiar from time-dependent perturbation theory. In the limit $T \to \infty$ it becomes a δ-function times π, and (3.46) can be written (Pauli 1928)

$$\begin{aligned}\frac{d\varrho_{mm}}{dt} &= 2\pi \sum_n |H_{mn}|^2 \delta(E_m^{(0)}-E_n^{(0)})(\varrho_{nn}-\varrho_{mm}) \\ &=: \sum_n A_{mn}(\varrho_{nn}-\varrho_{mm})\end{aligned} \tag{4.17}$$

This *Pauli equation* is similar to other master equations, such as (3.48), while the coefficients A_{mn}, defined on the RHS, are transition rates in analogy to Boltzmann's $w(\boldsymbol{p}_1\boldsymbol{p}_2, \boldsymbol{p}'_1\boldsymbol{p}'_2)$ of Sect. 3.1.1. If H_1 contains only two-particle interactions, the sum over n may again be written as a sum over particle pairs. According to the above definition, the coefficients A_{mn} conserve energy and satisfy the symmetry under collision inversion, $A_{mn} = A_{nm}$, (cf. (3.7)). Therefore, the Pauli equation conserves total probability, $\sum_n d\varrho_{nn}/dt = 0$.

One may use the explicit form of the Pauli equation to discuss, in this case, the validity of the approximations used when deriving the master equation (3.46) — see e. g. Jancel (1963). This validity depends on the spectrum of the Hamiltonian, but is limited by Poincaré recurrences *in practice* only in unrealistic models or exceptional situations. The discrete spectra of realistic macroscopic quantum systems do *not* lead to a behavior that is essentially different from that of a continuous spectrum, since the time scales of relevant quasi-periodicities would by far exceed the age of the universe. Quantum systems may even exhibit 'classical chaos' (Habib, Shizume and Zurek 1998). On the other hand, a continuous spectrum would by no means justify an arrow of time (as is often claimed). The negligibility of recurrences for all times of interest — whether they exist in principle or not — forms a T-symmetric property. The physical importance of the difference between discrete and continuous spectra seems to be grossly overestimated in mathematical statistical physics.

The δ-function in (4.17) is meaningful only under an integral over E, or as an approximation under a sum over m. Therefore, Pauli combined groups of states with macroscopically equivalent energies to form 'cells' (subspaces) with corresponding coarse-grained probabilities (see also van Kampen 1954). Joos (1984) was able to show that the off-diagonal elements ϱ_{mn} between such macroscopically different subspaces disappear by interaction with an environment that satisfies a Sommerfeld condition of lacking conspiratorial correlations ('decoherence' — see Sect. 4.3). This justifies Pauli's cells and the corresponding random phase 'approximation' dynamically.

When applied to a single initial state (with $\varrho_{00} = 1$), Pauli's equation (4.17) assumes the form of *Fermi's Golden Rule* in the Born approximation. Replacing the sum over initial states n with an energy integral and a sum over remaining quantum numbers β, that is, $\sum_n \cdots \to \sum_\beta \int \sigma_\beta(E) \ldots dE$ (with a partial density of states $\sigma_\beta(E)$), and similarly substituting $m \to E', \alpha$ for the final states, one obtains for the integrated probabilities, $\varrho_{\alpha\alpha} := \int \varrho_{E'\alpha,E'\alpha}\sigma_\alpha(E')dE'$,

$$\frac{d\varrho_{\alpha\alpha}}{dt} = 2\pi|H_{0\alpha}(E)|^2\sigma_\alpha(E) \quad \text{for} \quad \alpha \neq 0 \quad , \tag{4.18}$$

where $H_{\alpha 0}(E) := H_{\alpha E, 0E}$, while E is the energy of the initial state.

Although this Golden Rule can thus be derived as a formal approximation from the unitary dynamics (4.11), it is mainly used to describe stochastic decay and other 'quantum transitions'. (*Coherent* exponential decay according to the Schrödinger equation will be discussed in Sect. 4.5.) In contrast, Boltzmann's probabilistic transition rates $w(\boldsymbol{p}_1\boldsymbol{p}_2, \boldsymbol{p}'_1\boldsymbol{p}'_2)$ refer to *ensembles of individually-deterministic* collision trajectories (distinguished by their impact parameter). This difference reflects the fact that the formal concept of a density operator is already based on the probability interpretation. Nobody has ever been able to argue consistently how these fundamental quantum probabilities could be derived from an ensemble interpretation of the wave function. (Bohm's theory will be briefly discussed in Sect. 4.6.) In particular, the entropy (4.4) does not contain any contribution that could represent the missing information that would characterize this ensemble (as in Fig. 3.5 for classical measurements).

Pauli's equation does indeed resemble Born's original formulation of the probability interpretation by means of his *Born approximation* (Born 1926). Born used it initially to describe 'quantum jumps' between Schrödinger's stationary eigenstates of Hamiltonians H_0 that characterize isolated microscopic systems (such as atoms). [1] In quantum field theory, a similar splitting of

[1] Although Born may not have always been using his concepts consistently, in his third (here quoted) paper on the probability interpretation he evidently describes probabilities

the Hamiltonian is used to define the *interaction picture*. The special role attributed to the eigenstates of H_0 as representing the 'real' physical states, dynamically connected only through discrete jumps, was historically motivated by their correspondence to Bohr's discrete atomic electron orbits. Quantum jumps (or a 'collapse of the wave function') are, of course, incompatible with deterministic *trajectories in Hilbert space*, that is, with wave functions evolving according to a Schrödinger equation. The system Hamiltonians H_0 are thus assumed not to contain any interaction that is responsible for stochastic transitions. This early attempt to objectivize the probability interpretation (or the relevant observables) by a dynamical process is therefore based on an approximation. (Recall the trivial consequences for entropy according to the Pauli equation when exact energy eigenstates are used as a basis of relevance! Or could the Schrödinger equation nonetheless be exact? — See Sect. 4.6.)

The general structure of the Pauli equation is preserved even when the perturbation expansion in terms of the off-diagonal elements of H (in a certain basis) is not used. The new equation is known as Van Hove's 'exact' master equation (van Hove 1957). It represents Zwanzig's master equation for the projection (4.9) without any *further* approximation. In particular, if the chosen basis of relevance is the independent particle basis, the matrix elements H_{mn} appearing in the Pauli equation have merely to be replaced with the elements of the T-matrix, usually defined as $T := (S-1)/2\pi\mathrm{i}$, where S is the two-particle scattering matrix. This procedure presumes the negligibility of many-particle collisions, while now treating the two-particle collisions exactly (as in Boltzmann's *Stoßzahlansatz*). However, the adjective 'exact' for van Hove's equation is misleading even for diluted gases, as it refers only to the calculation of $\hat{G}_{\text{ret}}$, but not to the derivation of the master equation (3.46) itself. Similarly, Born's probability interpretation, when applied to *measurements*, depends on the choice of 'observables' even when the Born approximation is not used for calculating the corresponding probabilities.

In analogy to the H-theorem (3.10), one may again show that the entropy corresponding to the Zwanzig projection $\hat{P}_{\text{diag}}$ never decreases under the Pauli or Van Hove equation:

$$\frac{dS[\hat{P}_{\text{diag}}\varrho]}{dt} = -k\frac{d\left(\sum \varrho_{mm} \ln \varrho_{mm}\right)}{dt} \geq 0 \quad . \tag{4.19}$$

This entropy depends crucially on the chosen representation, that is, on the specific concept of relevance used in this master equation.

Because of the formal analogy, the classical canonical distribution, $\varrho_{\text{can}}(p,q) = Z^{-1}\exp[-H(p,q)/kT]$, is replaced with the canonical density

for jumps into new stationary *wave functions* — not probabilities for the occurrence of classical (such as particle) properties. In scattering or decay problems one has to identify the final stationary states with plane waves. Born then *associated* the latter with particle momenta according to de Broglie's relation. One year before the formulation of the uncertainty relations this was not recognized as being in conflict (in principle) with the measurement of the particle at the *position* of the detector.

operator, $\varrho_{\mathrm{can}} = Z^{-1}\exp(-H/kT)$. It can be derived precisely as in (3.19) by maximizing the entropy $S[\varrho]$ under the constraint of fixed mean energy. The so-called 'new statistics' (Bose or Fermi statistics) in terms of *particles* is then automatically obtained by applying this method to the quantum states of (free) *fields* (conveniently described in the *occupation number representation*), or equivalently to symmetric or antisymmetric particle wave functions. Only when expressed in terms of classical particle states does it appear as a new and *ad hoc* introduced method for counting them. The success of quantum statistics is indeed one of the strongest arguments against *particles* (in their original sense of point-like objects in space, distinguishable by their trajectories) as a fundamental kinematical concept.

This conclusion is further supported by the absence of Gibbs' self-mixing entropy (cf. page 69). The empirically correct measure on phase space, $d^{3N}p\,d^{3N}q/h^{3N}N!$, may be *derived*, for example, for the partition function Z of a grand canonical ensemble, $p_{E,N}(\mu,T) = \exp[-(E-\mu N)/kT]$. If this expression is evaluated by means of the familiar textbook approximation in the occupation number representation $|\{n_{\boldsymbol{k}}\}\rangle$ for 'single-particle' *wave functions* with wave numbers $\boldsymbol{k} = \boldsymbol{p}/\hbar$ on a large space volume V, one obtains for diluted gases (where $E = \sum_{\boldsymbol{k}} \varepsilon_{\boldsymbol{k}} N_{\boldsymbol{k}}$, with $\varepsilon_{\boldsymbol{k}} = p(\boldsymbol{k})^2/2m$ and $\varepsilon_{\boldsymbol{k}} - \mu \gg kT$)

$$
\begin{aligned}
Z(\mu,T) &= \sum_{\{n_{\boldsymbol{k}}\}} \exp\left[-\sum_{\boldsymbol{k}} \frac{(\varepsilon_{\boldsymbol{k}}-\mu)n_{\boldsymbol{k}}}{kT}\right] \\
&\approx \sum_N \frac{V^N}{h^{3N}N!}\exp\left(\frac{N\mu}{kT}\right)\int \exp\left(-\sum_{i=1}^N \frac{\boldsymbol{p}_i^2}{2mkT}\right) d^{3N}p \\
&\approx \sum_N \left[\frac{V}{h^3 N}\exp\left(\frac{\mu}{kT}\right)\int \exp\left(-\frac{\boldsymbol{p}^2}{2mkT}\right) d^3p\right]^N
\end{aligned}
\tag{4.20}
$$

(with $N = \sum n_{\boldsymbol{k}}$). The factorials $N! \approx N^N$ in the denominator are here *required* in order to compensate for the sum over all permutations of the N momenta $\boldsymbol{p}_i$ that are contained in this N-fold integral, since they all represent the *same* occupation states of *wave modes*. The corresponding density matrix, and therefore the partition function, factorizes in terms of wave modes $\boldsymbol{k}$ ('single particle states') rather than in terms of particles. The factorials thus need no longer be introduced *ad hoc* as in a particle picture.

General literature: Jancel 1963

4.2 Ensembles *versus* Entanglement

Quantum wholeness is analyzable

In the previous section, the von Neumann equation was derived from the Liouville equation by applying formal quantization rules. The dynamics of the density matrix, obtained in this way, is unitary. Therefore, it conserves $S[\varrho]$, while the Pauli (or Van Hove) equation, albeit derived from it *as an approximation*, appears to be superior, as it is able to describe quantum indeterminism and the increase of ensemble entropy characterizing the unpredictable outcome of measurement-like events.

The Liouville equation itself was obtained in Sect. 3.1.2 by applying Hamilton's (that is, Newton's) equations to ensembles that represent incomplete knowledge of classical states. Since the quantization of the Hamiltonian dynamics of mechanical systems leads to the Schrödinger equation, one may as well first quantize and then consider ensembles of solutions ψ_α with corresponding probabilities p_α, now describing incomplete knowledge *about the wave function* (see Fig. 4.1). This procedure may offer deeper insight into the meaning of the density matrix than did its formal foundation.

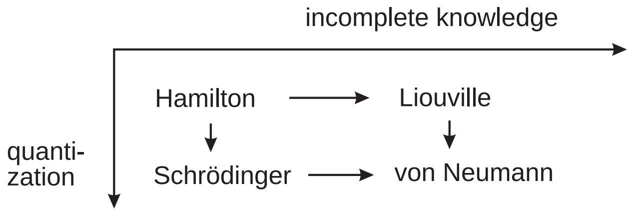


Fig. 4.1. Two routes from classical mechanics to the von Neumann equation

According to this ensemble interpretation, probabilities p_α, but not the density matrix $\varrho(q, q')$, correspond *conceptually* to the probability distributions $\varrho_\Gamma(p, q)$. The meaning of the density matrix can only be appreciated when considering ensemble expectation values of observables A, that is, *mean values of expectation values* with respect to different wave functions ψ_α:

$$\langle A \rangle := \sum_\alpha p_\alpha \langle \psi_\alpha | A | \psi_\alpha \rangle = \mathrm{Trace}\{A\varrho\} = \sum_n a_n \langle \phi_n | \varrho | \phi_n \rangle \quad , \tag{4.21}$$

with

$$\varrho := \sum_\alpha |\psi_\alpha\rangle p_\alpha \langle \psi_\alpha| \quad \text{and} \quad A := \sum_n |\phi_n\rangle a_n \langle \phi_n| \quad .$$

The symbol $\langle A \rangle$ denotes here a twofold mean; with respect to the ensemble of quantum states ψ_α with their probabilities p_α, *and* with respect to the quantum mechanical indeterminism of measurement results a_n with their probabilities $|\langle\phi_n|\psi_\alpha\rangle|^2$, valid for *given* quantum states ψ_α. In this way, the concept of a density matrix depends on the probability interpretation of the wave function — though *not* on any specific kinematical concepts (fundamental variables) representing the objects to which these probabilities apply,[2] or where and how they might dynamically arise (see Sect. 4.6).

An ensemble interpretation of the density matrix according to $\varrho = \sum_\alpha |\psi_\alpha\rangle p_\alpha \langle\psi_\alpha|$ does not require the members ψ_α of the ensemble of wave functions to be mutually orthogonal. They may in general even form an overcomplete set. Therefore, the ensemble cannot be uniquely recovered from the density matrix. Von Neumann's entropy (4.4) describes an ensemble entropy according to $S[\varrho] = -k\sum p_\alpha \ln p_\alpha$ only for the specific ensemble consisting of the orthonormal eigenstates of ϱ. Similar to classical statistical mechanics, its conservation reflects dynamical determinism — provided the Hilbert state norms are also conserved. This does not only require determinism, but also the *unitarity* of the Schrödinger equation (not just that of the von Neumann equation), since the formal density matrix does not distinguish between norm and probability of a wave function.

The mapping of ensembles of wave functions onto those which diagonalize their density matrix is an information-reducing idempotent operation, similar to a Zwanzig projection. Nonetheless, one may derive the autonomous von Neumann equation (4.1) from the ensemble interpretation and by using the further assumption that all wave functions defining the ensemble satisfy a Schrödinger equation $\mathrm{i}\partial\psi_\alpha/\partial t = H\psi_\alpha$ with one and the *same Hamiltonian* H. Although analogous arguments are used in classical statistical mechanics, presuming the Hamiltonian to be 'given' does not appear quite consistent while regarding *states* as incompletely known. The exact Hamiltonian would in general depend on the uncontrollably varying state of the environment.

Instead of describing an ensemble of wave functions, a density matrix may also represent the local (or 'reduced') perspective of *entangled* quantum systems. The generic quantum state of any two combined systems with variables x and y, say, may be written as

$$\psi(x,y) = \sum_{m,n} c_{mn}\phi_m(x)\Phi_n(y) \quad . \tag{4.22}$$

For spatially separate subsystems, this entanglement represents the fundamental concept of *quantum nonlocality* in its successful kinematical form. For example, it leads to Bell's theorem (in all its variants — cf. Bell 1964 and Greenberger, Horne, Shimony and Zeilinger 1990) and to so-called quantum teleportation. (The latter does not port anything, since what is to be 'ported' has to be carefully prepared as a component of the nonlocal initial state.)

[2] If these 'objects' of the probability interpretation were themselves represented by wave functions (such as eigenfunctions of 'observables'), the ensemble corresponding to all possible outcomes of all possible measurements would be quite different from the initial ensemble (which may be trivial in consisting of one state only). Nonetheless, probabilities for all these outcomes are appropriately *postulated* by the pragmatic rules used in (4.21).

All measurements performed at *one* subsystem (described by $\phi(x)$, say) can be characterized by expectation values for observables A_ϕ,

$$\langle A_\phi \rangle := \mathrm{Trace}\{A_\phi \varrho_{\mathrm{total}}\} = \mathrm{Trace}_\phi\{A_\phi \varrho_\phi\} \quad , \tag{4.23}$$

where ϱ_ϕ is defined as a partial trace, $\varrho_\phi := \mathrm{Trace}_\Phi\{\varrho_{\mathrm{total}}\}$, while ϱ_{total} may represent a wave function (or pure state), $\varrho_{\mathrm{total}} := |\psi\rangle\langle\psi|$. This *new* kind of density matrix is explicitly given by the expansion coefficients c_{mn} of the total state in (4.22),

$$(\varrho_\phi)_{mm'} := \langle \phi_m | \varrho_\phi | \phi_{m'} \rangle = \sum_n c_{mn} c^*_{m'n} \quad , \tag{4.24}$$

rather than by a statistical distribution p_α according to $\sum_\alpha c_{\alpha m} p_\alpha c^*_{\alpha m'}$.

Although the LHS of (4.24) cannot be distinguished by means of local *operations* from a density matrix describing an ensemble of states, it is thus evident that it does *not* represent one. This may be illustrated by the physically meaningful total angular momentum eigenstate of two distant particles. Therefore, the 'apparent ensemble' or 'improper mixture' (d'Espagnat 1966) must not be used to *explain* the probability interpretation (4.23) on that it was based. The fundamental difference between proper and improper mixtures cannot be overcome (though perhaps it can be obscured) by using the formal limit of an infinity of degrees of freedom (see Hepp 1972). Statistical operators, sometimes fundamentally postulated to describe open systems, are insufficient for a complete description, as they neglect nonlocal quantum correlations. Such formalisms remain blind to the measurement problem (see below and Sect. 4.6).

Any density matrix, such as ϱ_ϕ or ϱ_Φ, is Hermitean and can therefore be diagonalized in the form $\varrho_\phi = \sum_n |\tilde{\phi}_n\rangle p_n \langle\tilde{\phi}_n|$ that defines its eigenbasis $\{\tilde{\phi}_n\}$. This is a formal ensemble of *orthogonal* states, operationally meaningful only for the subsystem. By using this diagonal form and (4.24), one observes that all eigenvalues p_n are non-negative. Phenomenological *dynamical maps* (see Sect. 4.4) must therefore be chosen 'completely positive' by hand, that is, they have to conserve this property for *all* density matrices.

For an entangled state such as (4.22), the eigenbasis of both subsystem density matrices defines the *Schmidt canonical form*,

$$\psi(x,y) = \sum_k \sqrt{p_k}\,\tilde{\phi}_k(x)\tilde{\Phi}_k(y) \quad , \tag{4.25}$$

which, in contrast to (4.22), is a single sum (Schmidt 1907, Schrödinger 1935). All phase factors of the coefficients $\sqrt{p_k}$ in (4.25) have here been absorbed into the orthonormal states $\tilde{\phi}_k$ or $\tilde{\Phi}_k$. For given subsystems, this representation (and hence its time dependence — see Kübler and Zeh 1973 [3]) is defined by the state ψ of the total system (except for accidental degeneracy of the p_k's).

[3] In this and earlier papers, I translated Schrödinger's *Verschränkung* with 'quantum correlations', while the rate at which two factorizing systems start to become entangled (related to the decoherence rate) was referred to as a 'de-separation parameter'.

Complete neglect of the correlations between two subsystems can be described by a nonlinear (though information-reducing) Zwanzig projection,

$$\hat{P}_{\mathrm{sep}}\varrho := \varrho_\phi \otimes \varrho_\Phi \quad . \tag{4.26}$$

The stronger Zwanzig projection of locality, $\hat{P}_{\mathrm{local}}\varrho = \prod_k \varrho_{\Delta V_k}$, leading to a density matrix that factorizes, as in (3.37), with respect to small volume elements ΔV_k in space, would again be required in order to arrive at the approximate concept of an *entropy density* $s(\boldsymbol{r})$. Since indistinguishable particles can *not* be used to define subsystems, genuine entanglement does *not* include the formal correlations that describe symmetrization or antisymmetrization of the wave function. These pseudo-correlations are merely an artifact that remains from using classical particle concepts — cf. (4.20).

As a consequence of the nonlocality of quantum states, and in fundamental contrast to classical physics, the entropy $S[\hat{P}_{\mathrm{sep}}\varrho]$ (or $S[\hat{P}_{\mathrm{local}}\varrho]$) of a completely defined (pure) quantum state is in general nontrivial (it does not vanish). In this case the quantum entropy measures entanglement — not lacking information. As states of the subsystems do not exist, they cannot be merely *unknown*. The understanding of entanglement as "relative information" is an incomplete, often misleading analogy. The *objective* apparent ensembles, which are defined for all subsystems, may even define the *representative ensembles* often used (and sometimes questioned) in statistical thermodynamics (cf. Tolman 1938). However, an explanation is now required for the fact that (1) microscopic systems are usually found in pure states (such as eigenstates of their Hamiltonians H_0), and (2) that the macroscopic world is successfully described in terms of local (classical) concepts rather than in terms of their superpositions.

A local concept of relevance that, in contrast to $\hat{P}_{\mathrm{sep}}$, preserves all 'statistical' correlations (those based on incomplete information) while dropping quantum correlations (entanglement) may be defined by the Zwanzig projection of *classical correlations only*, $\hat{P}_{\mathrm{classical}}$. In the Schmidt-canonical representation it is defined as

$$\hat{P}_{\mathrm{classical}}(|\psi\rangle\langle\psi|) := \sum_k p_k |\tilde{\phi}_k\rangle\langle\tilde{\phi}_k| \otimes |\tilde{\Phi}_k\rangle\langle\tilde{\Phi}_k| \quad . \tag{4.27}$$

Quantum correlations would here require a double sum over k and k' (in another basis a sum over two *pairs* of indices). The RHS can be regarded as describing incomplete information about a presumed product state.

Quantum entanglement has also been discussed for 'mixed states' of the total ('bipartite') system (Werner 1989, Peres 1996). Its consequences are then suppressed either by *averaging* over the corresponding ensemble of completely defined states, or by *tracing out* the environment that gives rise to this reduced state. However, it would obviously be misleading to define such a situation as 'not entangled'.

A random *pure* state (4.22) does not lead to the 'statistical' result $\varrho_{\rm irrel} \approx 0$ (compare (3.44)) for the relevance concept (4.26), since only the improbable factorizing states would not contain any correlations. (The reader may wish to skip the rest of this somewhat technical paragraph.) For example, the linear entropy according to (4.7), given by $S_{\rm lin} = 1 - \sum_{kk'} |\varrho_{kk'}|^2$ if normalized to vanish for a pure state, assumes its maximum in a Hilbert space of dimension D, $S_{\rm lin}^{\rm max} = 1 - 1/D$, for the *mixture* $\varrho_{kk'} = \delta_{kk'}/D$. For *pure states* in the tensor product of two Hilbert spaces with dimensions M and N (hence with $D = MN$) one obtains for the *mean* linear subsystem entropy in either subsystem (Lubkin 1978)

$$\bar{S}_{\rm lin}^{sub} = 1 - \frac{M+N}{D+1} < 1 - \frac{1}{M} \quad \text{and} \quad < 1 - \frac{1}{N} \quad . \tag{4.28}$$

This vanishes only for $M = 1$ or $N = 1$. Since the *linear information* $I_{\rm lin} := 1 - S_{\rm lin}$ is multiplicative for products of density matrices (rather than being additive as is its logarithmic measure), its value for the total system is given by $I_{\rm lin}[\hat{P}_{\rm sep}\varrho_{pure}] = (I_{\rm lin}^{sub})^2$, as the entropies of the subsystems must be equal for a pure total state (see (4.25)). For $D \gg 1$ one has, according to (4.28), $\bar{I}_{\rm lin}[\hat{P}_{\rm sep}\varrho_{pure}] \approx \frac{(M+N)^2}{MN} I_{\rm lin}^{\rm min}$ ($\approx \frac{M}{N} I_{\rm lin}^{\rm min}$ for $M \gg N$), with $I_{\rm lin}^{\rm min} = 1 - S_{\rm lin}^{\rm max} = 1/MN$ (valid for mixtures). While $\varrho_{\rm irrel} = (1-\hat{P}_{\rm sep})\varrho$ vanishes in a random mixture $\sum_\alpha |\psi_\alpha\rangle p_\alpha \langle\psi_\alpha|$, the mean information for pure total states, $\sum_\alpha p_\alpha I_{\rm lin}[(1-\hat{P}_{\rm sep})|\psi_\alpha\rangle\langle\psi_\alpha|]$, does not. Since the irrelevant information is measured by $I[\varrho] - I[\hat{P}_{\rm sep}\varrho]$ (rather than by $I[(1-\hat{P}_{\rm sep})\varrho]$), the random pure state (with $S_{\rm lin}[\varrho] = S_{\rm lin}^{\rm min} = 0$) must possess large *local* entropy $S_{\rm lin}[\hat{P}\varrho]$, and therefore has to carry its information predominantly in $\varrho_{\rm irrel}$.

Similar relations hold for the logarithmic entropy (Page 1993, Foong and Kanno 1994). It is impossible to reach $S^{\rm max}$ for $M \gg N$ with a pure total state. Because of $I_\phi = I_\Phi$, the *local information* about the larger system cannot be entirely transformed into correlations. However, every (small) subsystem of the quantum universe that is completely described must essentially possess maximum entropy for a random global state.

In general, a reduced density matrix no longer obeys a von Neumann equation if the total wave function ψ evolves according to a Schrödinger equation. Its dynamics cannot be autonomous. Although it can be explicitly formulated (Kübler and Zeh 1973, Pearle 1979), its solution would in general require solving the Schrödinger equation for the total system. Indeed, ϱ_ϕ multiplied by the unit operator in Φ-subspace represents another (linear) Zwanzig projection,

$$\hat{P}_{\rm sub}\varrho_{\rm total} := \varrho_\phi \otimes 1_\Phi \quad . \tag{4.29}$$

Phenomenological master equations for ϱ_ϕ, introduced in analogy to Boltzmann's *Stoßzahlansatz*, are referred to as 'open systems' quantum dynamics (see Sect. 4.4). They are often derived by assuming an uncorrelated environment in the form of a heat bath (Favre and Martin 1968, Davies 1976). While this is analogous to Boltzmann's chaos assumption for *statistical* correlations, master equations must *explain* heat baths (that is, canonical ensembles described by a temperature parameter) rather than presuming their existence. (Cf. the concepts of *reference densities* and *relative entropy* in Sects. 3.3.1 and 4.4.)

The expectations values (4.21) and (4.23), which led to the concept of a density matrix, both rely on Born's probabilities for the outcome of measurements. Von Neumann (1932) proposed a model interaction that was meant to describe *ideal measurements* (or *measurements of the first kind*) dynamically. It is defined as a unitary transformation $\phi_n \Phi_0 \underset{t}{\rightarrow} \phi_n \Phi_n$, where ϕ_n is an eigenstate of an observable, $A = \sum_n |\phi_n\rangle a_n \langle\phi_n|$, while Φ_0 and Φ_n are the initial state of the apparatus and its final 'pointer positions', respectively. If the states ϕ_n are mutually orthogonal, this transition represents a fork of causality in classical *configuration* space (cf. Footnote 1 of Chap. 2). The observable A is thus *defined* by this interaction up to the scale factor a_n. If the microscopic system is initially in one of the eigenstates of A, it does not change during an ideal measurement, while the apparatus evolves into the corresponding pointer state Φ_n. (Non-ideal measurements could be similarly described by replacing ϕ_n on the RHS of the unitary transformation above with a different final state ϕ'_n.)

However, for a general initial state, $\sum_n c_n \phi_n$, one obtains for the same interaction, and for the same initial state of the apparatus,[4]

$$\left(\sum_n c_n \phi_n\right) \Phi_0 \quad \underset{t}{\rightarrow} \quad \sum_n c_n \phi_n \Phi_n =: \psi \quad . \tag{4.30}$$

The RHS is now an *entangled* state, while an *ensemble* of different measurement results (that is, of states $\phi_n \Phi_n$ with probabilities $|c_n|^2$), would require the fork of causality to be replaced with a fork of indeterminism. (The formal 'plus' characterizing the superposition would have to be replaced with an 'or'.) This discrepancy represents the *quantum measurement problem.* The subsystem density matrices resulting from these two types of fork are iden-

[4] The assumption that the initial quantum state Φ_0 of the apparatus does not depend on the state of the microscopic system before any interaction is another application of *intuitive causality* (see also Sect. 4.6). Schulman (1986, 1997) made an attempt to understand the measurement process, and even to *determine* its quasi-classical outcome, in terms of the Schrödinger equation by assuming that the apparatus 'knows' the measurement result in advance (equivalent to the existence of 'anti-causal correlations'). The speculative nature of this proposal is illustrated by the fact that Price (1996) made essentially the opposite suggestion, *viz.*, that the *wave* aspects of quantum mechanics be explained by means of advanced correlations between *classical* particles (see also Cramer 1986). None of these proposals explains how the 'special states' that they require could evolve (backwards in time) from a universal *final* condition (at the big crunch, for example) in a similar way as initial states are believed to form through evolution from a low entropy big bang. In *this* respect they represent wishful thinking. If, however, values of dynamical variables were allowed to be freely chosen or restricted at arbitrary times, this could be used to 'explain' *all* kinds of miracles. Our (seemingly-so) freedom to choose *initial* conditions only is an empirical manifestation of the arrow or time.

tical, since there is no way of distinguishing these different situations operationally by a *local* observation. As emphasized above, this argument does not *explain* the fork of indeterminism that is at the heart of the probability interpretation. If the pointer states Φ_n are also mutually orthogonal (as has to be assumed for a measurement), both sides of (4.30) are Schmidt-canonical.

This measurement problem is independent of the complexity of the measurement device (which may give rise to thermodynamically irreversible behavior), and of the presence of fluctuations or perturbations caused by the environment, since the states Φ may be defined to describe this complexity completely, and even to include the whole 'rest of the world'. The popular argument that the quantum mechanical indeterminism might, in analogy to the classical situation, be caused by thermal fluctuations during a measurement process (cf. Sect. 3.3 or Peierls 1985, for example) is incompatible with universal unitarity. It would require the existence of an initial ensemble of microscopic states which in principle had to determine the outcome. However, the ensemble entropy of the RHS of (4.30) does not characterize an ensemble that would allow the measurement to be interpreted as in Fig. 3.5.

If both systems in (4.30) are microscopic, the dynamics representing the fork of causality can even be reversed in practice in order to demonstrate that all components still exist. It may lead to observable consequences that depend on all of them, including their relative phases. This prohibits the interpretation of (4.30) as a dynamical fork of indeterminism (a fork between mere possibilities), even though von Neumann's fork of causality is defined (in terms of branching wave packets) in a *classical* configuration space. The transition from quantum to classical (Sect. 4.3) can be understood only if it explains why the fundamental arena for wave functions often *appears* as a space of 'classical' configurations.

The interaction (4.30) illustrates why subsystems by themselves cannot, in general, obey unitary dynamics. Similar arguments hold for interacting systems with merely classical correlations (where $\varrho = \hat{P}_{\text{classical}}\varrho$). In this case, the effective subsystem Hamiltonian H_ϕ, say, would depend on the state Φ_k of the other system through a partial expectation value,

$$H_\phi^{(k)}(t) = \langle \Phi_k(t)|H|\Phi_k(t)\rangle_{\text{Op}} \quad . \tag{4.31a}$$

This may be compared with the effective Hamiltonian of a genuinely classical subsystem, defined by

$$H_\phi(p_1 \dots q_N, t) = H\left(p_1 \dots q_n; p_{n+1}(t) \dots q_N(t)\right) \quad . \tag{4.31b}$$

Here, particle numbers $1 \dots n$ are meant to characterize the subsystem corresponding to the quantum states ϕ in (4.31a). Each element of the ensemble would then satisfy another Hamiltonian or Schrödinger equation. This is in contrast to the assumptions leading to the Liouville or von Neumann equation for the density matrix in (4.21). Neglecting these correlations dynamically by using $\hat{P}_{\text{sep}}$ in a master equation, such as (3.36), would amount to applying the whole resulting ensemble of sub-Hamiltonians (in the forward direction of time) to each individual element of the ensembles of states of the subsystems. However, only under the dynamically unstable classical assumption $\varrho = \hat{P}_{\text{classical}}\varrho$ would the quantum mechanical situation simply be equivalent to the classical one discussed in Sect. 3.1.2 (cf. Fig. 3.1).

It should be kept in mind, therefore, that the local concepts of relevance, $\hat{P}_{\text{sep}}, \hat{P}_{\text{local}}$ and $\hat{P}_{\text{classical}}$, appear 'natural' only to our classical prejudice. In the unusual situation of EPR/Bell type experiments, quantum correlations become relevant even for local observers. *Dynamical* locality, described by means of point interactions in the underlying classical field theory, merely warrants the dynamical consistency of these concepts of relevance, such as the approximate validity of autonomous master equations for $\hat{P}_{\text{local}}\varrho$.

General literature: d'Espagnat 1976, 1983

4.3 Decoherence

In science, new ideas are at first completely neglected,
later fiercely attacked, and finally regarded as well known
Konrad Lorenz

The concept of decoherence as a consequence of entanglement was originally born out of the idea of *spontaneous symmetry breaking in microscopic systems* (Zeh 1965, 1967) when applied to the universe as another closed quantum system (which contains its own observer, however). Using the terminology introduced above, decoherence may be technically defined as the justification of a *specific* $\hat{P}_{\text{semidiag}}$ for a given subsystem by presuming the relevance of the corresponding $\hat{P}_{\text{sub}}$ (cf. (4.29)). If this $\hat{P}_{\text{semidiag}}$ turns out to be dynamically valid under all normal circumstances, its eigenspaces characterize 'quasi-classical' properties. Thus, classical concepts *emerge* through unavoidable and practically irreversible interaction with the environment. They do not have to be presumed as an independent fundamental ingredient, required for an interpretation of the formalism.

In the theory of *decoherent histories* (Omnès 1992, Gell-Mann and Hartle 1993), decoherence is introduced as a *condition* for dynamically 'consistent' probabilities (Griffiths 1984). It is not necessarily based on entanglement (Omnès 1999), and insofar not equivalent to environmental decoherence.

Equation (4.30) was meant to describe the controllable interaction of a microscopic system ϕ with an appropriate measurement device (with 'pointer' states Φ_n). Its fact-like asymmetry (leading from factorizing to entangled states) could be reversed with sufficient effort if both subsystems were microscopic. In analogy to particle emission, this situation may be regarded as 'virtual decoherence'. For genuine quantum measurements, the pointer states

Φ_n must be macroscopic (cf. Gottfried 1966, for example). Assume now that the pointer states in (4.30) are replaced with states of an uncontrollable (such as thermal) environment of the ϕ-system, while the latter represents a macroscopic object, such as a pointer. Then (4.30) cannot be reversed, and decoherence may be regarded as 'real'. Measurements which lead to macroscopic pointer positions cannot be undone.

The uncontrollable (inter)action (4.30) is analogous to Boltzmann's *Stoßzahlansatz* in creating correlations which propagate away, while it describes the specific quantum aspect of *delocalizing* phase relations. Its time arrow may therefore be referred to as *quantum causality.* True measurements according to (4.30) must indeed be irreversible in order to avoid the possibility of *quantum erasure* (restoration of interference, or *recoherence*). 'Erasure' would require that every single scattered particle were recovered in order to undo the decoherence. Experiments demonstrating quantum erasure in certain microscopic systems (cf. Kwiat, Steinberg and Chiao 1992) do not necessarily recover the whole initial superposition — they may partially *rebuild* it by using an appropriate further collapse of the wave function.

The resulting local (reduced) density matrix ϱ_ϕ in the sense of (4.23) represents only an *apparent* (seemingly-so) ensemble of quasiclassical states. In contrast to 'phase averaging' based on incomplete information (as used by Pauli or van Kampen — cf. Sect. 4.1.2), decoherence is an objective (individual) dynamical process. The interaction with the environment, which leads to *ever-increasing entanglement*, is practically unavoidable for most systems in all realistic situations (Zeh 1970, 1971, 1973, Leggett 1980, Zurek 1981, 1982a, Joos and Zeh 1985). It is this universal and quantitative aspect of entanglement that seems to have been greatly overlooked for a long time, during which macroscopic objects were *unsuccessfully* described by a Schrödinger equation. Decoherence is extremely efficient, since it does not require an environment that *acts* on the system (as it would have to do in order to induce 'distortions' or Brownian motion). Neither the concepts of conserved momentum and energy nor that of an environmental heat bath are required. Decoherence may occur far from any (partial) equilibrium. In this respect it is comparable with Borel's unavoidable interaction at astronomical distances, mentioned in Sect. 3.1.2.

The process of decoherence acts on a much shorter time scale than thermal relaxation or dissipation (Joos and Zeh 1985, Zurek 1986). It means that *there are no approximately closed macroscopic systems* (save the whole universe). On the other hand, systematic decoherence requires the concept of a 'normal' environment that monitors *certain* properties (represented by subspaces). The latter then appear as 'classical facts', which seem to exist regardless of their observation, while their superpositions are never observed.

Some important applications of decoherence are described below.

General literature: Zurek 1991, Giulini et al. 1996

4.3.1 Trajectories

In an imagined two-slit interference experiment with 'bullets' (or perhaps more realistically with small dust grains or macro-molecules), not only their passage through the slits, but their whole path would be 'measured' by scattered molecules or photons under *all* realistic circumstances. For macroscopic objects one could simply confirm this by opening one's eyes. Therefore, no interference fringes could ever be observed — even if the resolution of the registration device were fine enough. (An interference experiment with *mesoscopic* molecules has recently been successfully performed by Arndt *et al.* 1999.) Classical objects resemble alpha particles in a cloud chamber (Mott 1929), since they can never be regarded as being isolated (as can be arranged for microscopic objects). Their unavoidable entanglement with their environment (Fig. 4.2) leads to a 'reduced' density matrix that can be conveniently represented by an ever-increasing ensemble of narrow wave packets following slightly stochastic trajectories. As a by-product, this result demonstrates that the concept of an S-matrix (with its presumed free asymptotic states) does not apply to macroscopic objects.

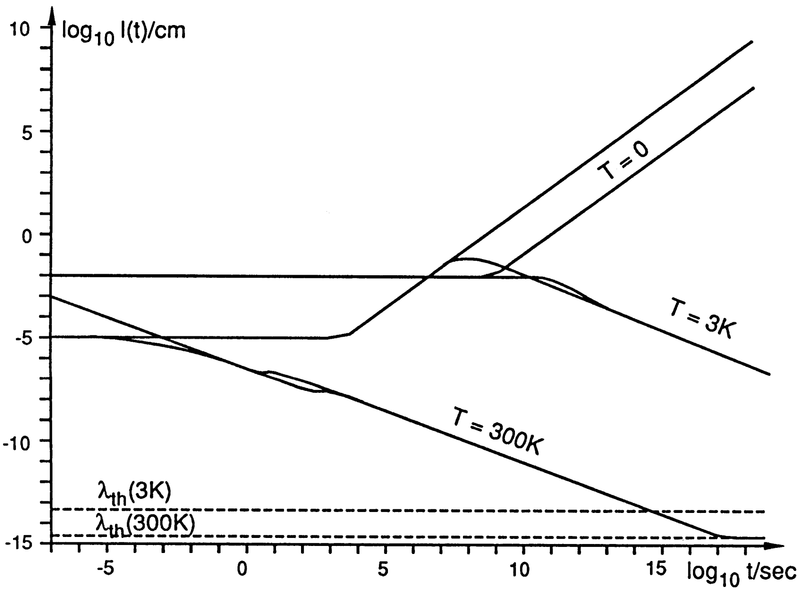


Fig. 4.2. Time dependence of the coherence length $l(t)$ for the center of mass of a small dust grain of 10^{-14} g with radius 10^{-5} cm under continuous measurement by thermal radiation. The six curves refer to two initially pure Gaussian wave packets, differing by their initial widths $l(0)$, and to three different temperatures T of the radiation. $T = 0$ describes the dispersion of the wave packet according to the Schrödinger equation, which holds otherwise as an approximation for a limited time only. Scattering of atoms and molecules is in general far more efficient than that of thermal photons — even in intergalactic space. Brownian motion becomes relevant only for $l(t) \to \lambda_{th}$. (From Joos and Zeh 1985)

For a continuous variable, decoherence competes with the dispersion of the wave packet that is reversibly described by the Schrödinger equation. Even the apparently small scattering rate of photons, atoms, or molecules in intergalactic space off small dust particles would suffice to destroy all coherence in a spreading center of mass wave packet (see Fig. 4.2). An otherwise free particle, for example, is then dynamically described by the master equation

$$i\frac{\partial\varrho(x,x',t)}{\partial t} = \frac{1}{2m}\left(\frac{\partial^2}{\partial x'^2} - \frac{\partial^2}{\partial x^2}\right)\varrho - i\lambda(x-x')^2\varrho \quad , \tag{4.32}$$

which can be *derived* from a universal Schrödinger equation by assuming the future irrelevance of all correlations with the environment (cf. Joos' Chap. 3 and App. 1 of Giulini et al. 1996). The coefficient λ is here determined by the rate of scattering, and its efficiency to orthogonalize states of the environment. One does not have to *postulate* a fundamental semigroup in order to describe open quantum systems (see Sect. 4.4). If the environment forms a heat bath, (4.32) describes the infinite-mass limit of *quantum Brownian motion* (cf. Caldeira and Leggett 1983, Zurek 1991, Hu, Paz and Zhang 1992, Omnès 1997). This demonstrates that even for entirely negligible recoil (which is responsible for noise and friction) there remains an important effect that is based on quantum nonlocality. Although Brownian fluctuations are required in a *thermal* environment, they affect the density matrix of macroscopic degrees of freedom much less than decoherence.

Classical properties (e. g. shapes and positions of droplets) thus *emerge* from the wave function (and are maintained) by a process that cannot practically be reversed. Particle aspects (such as tracks in a bubble chamber) are described by the reduced density matrix because of unavoidable interactions with the environment. The disappearance of interference between *partial waves* during a *welcher Weg* experiment (Scully, Englert and Walther 1991) does not require any wave-particle 'complementarity'. Similarly, there is no superluminal tunnelling (see Chiao 1998) in a consistent quantum description, since *all parts* of a wave packet propagate (sub-) luminally, while its group velocity does not represent propagation of any physical objects.

Master equations for open systems, such as (4.32), can also be derived by means of the *decoherence functional* (Feynman and Vernon 1963, Mensky 1979). Feynman's path integral is thereby used as a *tool* for calculating the propagation of a global density matrix, while the environment is again continuously traced out after getting entangled with the considered system. The intuitive concept of an *ensemble of paths* (representing different *possible* trajectories) is appropriate only *if* this treatment leads to their decoherence. A 'restriction' of the path integral (Mensky 2000) would introduce a phenomenological collapse of the wave function.

All quasi-classical phenomena, even those representing 'reversible' mechanics, are based on this *de facto* irreversible decoherence. It describes a perpetual production of objective physical entropy (increasing entanglement), which may be macroscopically negligible, but is large in terms of bits. If the quasi-classical trajectories are chaotic, this entropy production may be controlled by the classical Lyapunov exponent (Zurek and Paz 1994, Monteoliva and Paz 2000), even though this concept of entropy does not depend on any (initial) uncertainties that may grow *in the direction of calculation*, as in the classical theory of chaos (cf. Sect. 3.1.2).

General literature: Omnès 1997

4.3.2 Molecular Configurations as Robust States

Chiral molecules, such as the right- or left-handed sugars, represent another elementary property controlled by decoherence. A chiral state is described by a wave function, but in contrast to the otherwise analogous spin-3 state of an ammonia molecule, say, not by an energy eigenstate (see Zeh 1970, Primas 1983, Woolley 1986). The reason is that it is chirality (but not parity) that is continuously 'measured', for example by scattered air molecules. (For sugar molecules under normal conditions decoherence requires a time scale of the order $10^{-9}sec$ — see Joos and Zeh 1985.) Measurements of the parity of sugar molecules, or their preparation in energy eigenstates, are therefore practically excluded, since this would require an even stronger coupling to the corresponding apparatus.

As a dynamical consequence, each individual molecule in a bag of sugar must then retain its chirality, while a parity state — if it had come into existence in a mysterious or expensive way — would almost immediately 'collapse' into a definite chirality (with equal probabilities). Parity is thus not conserved for sugar molecules, while chirality would be confirmed 'without demolition' when measured twice. (A mixture of chiralities would be *operationally identical* to a mixture of parities only in the pathological case of *exactly* equal probabilities.)

This dynamical *robustness* of certain properties under the influence of the environment seems to characterize what we usually regard as 'real facts' (in the operational sense), such as spots on the photographic plate, or any other 'pointer states' of a measurement device. The concept of robustness is thereby compatible with a (regular) time dependence, as exemplified in the previous section for the center of mass motion of macroscopic objects. Since entropy production by interaction with the environment is least for a density matrix that is already diagonal in terms of robust states, this property has been called a 'predictability sieve', and proposed as a definition of classical states (Zurek, Habib and Paz 1993).

Robustness gives rise to quasi-classical trajectories in terms of wave packets, and it seems to be essential for the physical concept of memory (as in DNA, brains or computers) — with the exception of proposed quantum computers, which are extremely vulnerable to decoherence (Haroche and Raimond 1996, Ekert and Jozsa 1996 — see also Zurek 1998). Even 'states of being conscious' (cf. Chap. 1) that are able to communicate seem to be quasi-classical (Tegmark 2000). In contrast to these robust properties, which can be assumed (for all further consequences) to exist as 'facts' regardless of their measurement, *potentially* measurable quantities are called 'counterfactual'. Their superpositions must *not* be assumed to describe ensembles of definite (really existing though unknown) properties. These different situations can be well analyzed and understood in terms of decoherence.

Chemists know that atomic nuclei in large molecules have to be described classically (for example by rigid configurations, which may vibrate or rotate in a time-dependent manner), while the electrons are represented by stationary or adiabatically comoving wave functions. This asymmetric behavior is often attributed, by means of a Born-Oppenheimer approximation, to the large mass ratio. However, this argument is insufficient, since the same approximation can be applied to energy *eigenstates* of small molecules with their discrete energy bands. A similar case is now also studied in *quantum gravity*, where one claims that the 'classicality' of spacetime can be explained by merely employing a Born-Oppenheimer approximation with respect to the inverse Planck mass (see Sect. 6.2.2). It cannot be saved by means a WKB approximation for the massive variables, since this would *not* exclude broad wave functions (instead of the required narrow wave packets moving along trajectories). The WKB approximation can only explain the *propagation* of wave fronts according to geometrical optics, and therefore the stability of narrow wave packets *once they have formed*.

The formation of pseudo-classical wave packets for the atomic nuclei in large molecules or for the gravitational field can instead again be explained by decoherence (Joos and Zeh 1985, Unruh and Zurek 1989, Kiefer 1992 — see also Sect. 4.3.4). For example, the positions of nuclei are permanently monitored by scattering of (other) molecules. But why only the nuclei (or ions), and why not including those in very small molecules? The answer can only be quantitative, and it is based on a delicate balance between internal dynamics and interaction with the environment, whereby the density of states plays a crucial role (Joos 1984). Depending on the specific situation, one will either obtain an approximately unitary evolution, a master equation (with time asymmetry arising from quantum causality), or complete freezing of the motion (quantum Zeno effect). The situation becomes relatively simple only for a 'free' massive particle, which is described by (4.32).

General literature: Joos' Chap. 3 of Giulini et al. 1996

4.3.3 Charge Superselection

Gauß' law, $q = \frac{1}{4\pi}\int \boldsymbol{E}\cdot d\boldsymbol{S}$, tells us that every local charge is correlated with its Coulomb field on a sphere at any distance. A superposition of different charges,

$$\sum_q c_q \psi_q^{total} = \sum_q c_q \chi_q \Psi_q^{field} = \sum_q c_q \chi_q \Psi_q^{near} \Psi_q^{far} =: \sum_q c_q \chi_q^{dressed} \Psi_q^{far} \quad , \tag{4.33}$$

therefore represents an entangled state of the charge and its field. Here, χ_q describes the bare charge, while $\Psi_q^{field} = \Psi_q^{near}\Psi_q^{far}$ is the wave functional of its complete field, symbolically written as a tensor product of a near-field and a far-field (cf. Sect. 2.3). The dressed (physical) charge state would then be described by a density matrix of the form

$$\varrho_{local} = \sum_q |\chi_q^{dressed}\rangle |c_q|^2 \langle \chi_q^{dressed}| \tag{4.34}$$

if the states of the far field are mutually orthogonal (uniquely distinguishable) for different charge q. The charge is thus decohered by its own Coulomb field, and no charge superselection rule has to be *postulated* (see Giulini, Kiefer and Zeh 1995). The formal decoherence of the bare charge by its near-field remains unobservable, since experiments can only be performed with dressed charges.

While this result is satisfactory from a theoretical point of view, a more practical question is, at what distance and on what time scale the superposition of two different locations of a point charge (such as an electron during an interference experiment) are observationally destroyed by the corresponding state of the dipole field. A *classical* retarded Coulomb field would contain complete information about the precise path of the charge. However, since interference between different electron paths has been observed over distances of the order of millimeters at least (Nicklaus and Hasselbach 1993), the QED equivalent to Coulomb fields seems to contribute to decoherence only by its monopole component.

This conclusion (which emerged from discussions with E. Joos) can be readily understood in quantum mechanical terms, since photons with infinite wave length (representing static Coulomb fields) cannot distinguish different charge positions (even though the number of virtual photons may diverge). Static dipole (or higher) moments do not possess far-fields, defined to decrease with $1/r^2$. Therefore, of the Coulomb field, only the 'topological' Gauß constraint $\partial_\mu F^{\mu 0} = 4\pi j^0$ remains in QED. The *observed* (retarded) Coulomb field must then be completely described in terms of transverse photons, represented by the vector potential $\boldsymbol{A}$ (with $\mathrm{div}\boldsymbol{A} = 0$ in the Coulomb gauge), and restricted to states obeying the Gauß constraint. In this picture, only the spatial distribution of electric field lines — not their total flux — forms dynamical degrees of freedom that have to be quantized. In this sense, charge decoherence has been regarded as 'kinematical', although it might as well be assumed to be dynamically *caused* by the retarded field of the (conserved) charge in its past. Note, however, that a kinematical Coulomb constraint is in conflict with the concept of a physical Hilbert space that is spanned by direct products of local states.

Dipole and higher moments (which can define position *differences* of a point charge), can thus be measured by their environment either through emission (or scattering) of transversal ('real') photons, or by the irreversible polarization of nearby matter (Kübler and Zeh 1973, Zurek 1982a, Anglin and Zurek 1996). The decoherence of individual electrons within solid bodies is discussed in Imry (1997).

The emission of photons would require *acceleration*. For example, a *transient dipole* of charge e and maximum distance d, existing for a time interval t, involves accelerations a of at least the order d/t^2. According to Larmor's formula (cf. Sect. 2.3), the intensity of radiation is then at least $\frac{2}{3}e^2a^2$. In order to resolve the position difference, the emitted radiation has to consist of photons with energy greater than $\hbar c/d$ (that is, wave lengths smaller than d). The probability that information about the dipole is radiated away (by at least one photon) is then very small (of the order $\alpha Z^2(d/ct)^3$, where α is the fine structure constant and Z the charge number). In more realistic cases, such as interference experiments with electrons, stronger accelerations will be involved, but would still cause negligible decoherence in most situations. This limitation of the information capacity of an electromagnetic field by its quantum character must also give rise to an upper bound for the efficiency of interactions used in Borel's argument of Sect. 3.1.2.

The *gravitational field* of a point mass is analogous to the Coulomb field of a point charge. Superpositions of different mass should therefore be decohered by the quantum state of the monopole contribution of spatial curvature, and thus give rise to a mass superselection rule. However, superpositions of slightly differing masses evidently exist, since they form the time-dependent states of local systems, which would otherwise be excluded. This situation seems not to be sufficiently understood.

Nothing would remain of the Coulomb field in QED if the total charge of the quantum universe vanished exactly (cf. Giulini, Kiefer and Zeh 1995). The gravitational counterpart of this conclusion is the absence of time from a closed universe in quantized general relativity (see Sect. 6.2).

General literature: Giulini's Chap. 6 of Giulini et al. 1996

4.3.4 Classical Fields and Gravity

Not only are the quantum states of charged particles decohered by their fields, but also the quantum field states may in turn be decohered by the sources on which they (re)act. In this case, 'coherent states', that is, Schrödinger's time-dependent but dispersion-free Gaußian wave packets for the amplitudes of classical wave modes (eigenmodes of coupled oscillators), have been shown to be robust for similar reasons as chiral molecules or the wave packets describing the center of mass motion of quasi-classical particles (Kübler and Zeh 1973, Kiefer 1992, Zurek, Habib and Paz 1993, Habib et al. 1996). This explains why macroscopic states of neutral boson fields usually appear as *classical fields*, and why superpositions of macroscopically different 'mean fields' or different vacua (Sect. 6.1) have never been observed. In particular, quantum (field) theory must not and need not be reduced to a description of scattering processes, with probabilities interpretation for asymptotically isolated fragments.

These coherent (minimum uncertainty) harmonic oscillator states are defined for each wave mode k as the (overcomplete) eigenstates $|\alpha_k\rangle$ of the non-Hermitean photon annihilation operators a_k with complex eigenvalues α_k (that is, $a_k|\alpha_k\rangle = \alpha_k|\alpha_k\rangle$). They are Gaußian wave packets, centered at a time-dependent mean field $\alpha_k(t) = \alpha_k^0 e^{i\omega t}$, where $Re(\alpha_k)$ and $Im(\alpha_k)$ are analogous to the mean position and momentum of a mechanical oscillator wave packet. Since the Hamiltonian that describes the interaction with charged sources is usually linear in the field operators a_k or $a_k^\dagger$, these coherent states form a robust 'pointer basis' under normal conditions: they cause negligible entanglement with their environment (formed by their sources).

In contrast to these superpositions of many different photon numbers (that is, oscillator quantum numbers), one-photon states resulting from the decay of *one of several different* individual atoms (or even the n-photon states resulting from the decay of a *different number* n of atoms) are unable to interfere with one another, since they are correlated with mutually orthogonal final states (which have different atoms in their ground state). Two incoherent components of a one-photon state may then appear (using Dirac's language) as 'different' photons, although photons are not conceptually distinguishable from one another. In contrast, a coherent macroscopic ('collective') state *of the source* would react negligibly (judged in terms of the inner Hilbert space product) when *as a whole* emitting a photon. It would thus be able to produce the coherent ('classical') field states discussed above (see also Kiefer 1998). In his textbook, Dirac (1947) discussed also states of *two* (or more) photons correlated to the same state of the source, which would then have to contain two (or more) decayed atoms. They may be described by a symmetrized *product* of one-photon waves. Although these states form *one coherent component* of QED, their two (or more) *one-photon probabilities* (measured as light intensity) add again without interference (except for the exchange terms, which give rise to the Hanbury Brown-Twiss effect).

Superpositions of two *different* quasi-classical (coherent) states, $c_1|\alpha_1\rangle + c_2|\alpha_2\rangle$, have been produced and maintained for a short time as one-mode laser fields (k now omitted) in a cavity (Monroe et al. 1996). In order to prepare them, coherent field states are used as mesoscopic pointers in a 'measurement' that distinguishes two energy eigenstates $|\pm\rangle$ of a highly excited (Rydberg) atom by means of the von Neumann type interaction $|\pm\rangle|\alpha\rangle \to |\pm\rangle|\alpha e^{\pm i\theta}\rangle$. (The phase factor $e^{\pm i\theta}$ applies here to the eigenvalue α of the field mode, not to its Hilbert state $|\alpha\rangle$.) The common time dependence $e^{i\omega t}$ of these one-mode states cancels. Since the Rydberg states $|\pm\rangle$ can be rotated in their formal state space while being in 'Ramsey zones' (just as spin can be rotated in space in a magnetic field), they can be brought into a superposition, $|+\rangle\pm|-\rangle$, which upon measurement leads to

$$(|+\rangle \pm |-\rangle)\,|\alpha\rangle \to |+\rangle\,|\alpha e^{+i\theta}\rangle \;\pm\; |-\rangle\,|\alpha e^{-i\theta}\rangle \quad . \tag{4.35}$$

If θ is now chosen as $\pi/2$, and the atomic state 'rotated' again by $\pi/2$, the states $|\pm\rangle$ are correlated with field superpositions $|\alpha\rangle \pm |-\alpha\rangle$. They can be projected out by detecting that the atom is in the state $|+\rangle$ or $|-\rangle$. The field superpositions obtained in this way can then be probed and thus confirmed by means of a second Rydberg atom passing through the cavity. If the states $|\pm\alpha\rangle$ are regarded as classical, their superpositions may be likened to Schrödinger's much-discussed superposition of a cat being dead or alive. Unfortunately, a cat cannot be similarly transformed from alive to dead in a *reversible* process, since it would decohere too fast. The mesoscopic superposition must also decohere, although on a time scale that is slow enough to allow it to be monitored as a smooth process (Davidovich et al. 1996, Brune et al. 1996). There is evidently no discontinuous 'quantum jump' from the superposition into a definite quasi-classical state (see Sect. 4.3.5).

Similar arguments as those used above for electromagnetic fields apply to spacetime curvature in *quantum gravity* (Joos 1986, Kiefer 1999 — for applications to quantum cosmology see Chap. 6). One does not have to know its precise form (that may be part of an elusive unified quantum field theory) in order to conclude that the quantum states of matter and geometry (as far as this distinction remains valid) must be entangled and give rise to mutual decoherence. The classical appearance of spacetime geometry with its fixed light cone structure (that is *presumed* in conventional quantum field theory) is thus no reason *not* to quantize gravity. The beauty of Einstein's theory can hardly be ranked so much higher than that of Maxwell's as would be necessary in order to justify its exemption from quantization. An exactly classical gravitational field interacting with a quantum particle would even be incompatible with the uncertainty relations — as is known from the early Bohr-Einstein debate. The resulting density matrix (functional) for the gravitational tensor field as a 'quantum system' must therefore be expected to represent an apparent mixture of different quasi-classical curvature states (to which even the observer is correlated — see Sects. 4.6 and 6.2).

Moreover, the entropy and thermal radiation characterizing a black hole or an accelerated Unruh detector (Sects. 5.1 and 5.2) are consequences of the entanglement between relativistic vacua on two half-spaces separated by an event horizon. This entanglement entropy measures the same type of 'apparent' ensembles as the entropy produced according to the master equation (4.32) for a macroscopic mass point. An event horizon need not be different from any other quasi-classical property. Nonetheless, the disappearance of coherence behind an (even virtual) horizon has been regarded as a *fundamental* (law-like) violation of unitarity, and even as the final source of all irreversibility (see Sects. 4.4, 5.1 and 6.2). This does in no way appear justified (cf. Kiefer, Müller and Singh 1994).

General literature: Kiefer's Chap. 4 of Giulini et al 1996, Kiefer 1999

4.3.5 Quantum Jumps

Quantum particles are often observed as flashes on a scintillation screen, or heard as 'clicks' from a Geiger counter. These macroscopic phenomena are then interpreted as being caused by point-like objects passing through the observing instrument during a short time interval, while this is in turn understood as evidence for the discontinuous decay of an excited state (such as an atomic nucleus). A *rate* for stochastic 'decay events' is equivalent to a master equation. A *constant* rate describes exponential decay. Discrete quantum jumps between two energy eigenstates have also been directly monitored for single atoms in a cavity which are strongly coupled to an observing device (Nagourney, Sandberg and Dehmelt 1986, Sauter et al. 1986). Therefore, creation and annihilation operators are often understood as describing discrete events, even though they occur in a Hamiltonian.

The Schrödinger equation, on the other hand, describes a wave function (usually with angular distribution according to a spherical harmonic) smoothly leaking out of a decaying system (such as a 'particle' in a potential well). This contrast between observed discrete events and the Schrödinger equation is clearly the empirical root of the probability interpretation in terms of particles. Since probability is conserved, a wave function can exhibit exponential time dependence only in a limited region of space (usually an expanding sphere with radius determined by the time of formation of the decaying state and the speed of the decay fragments). The complete wave function represents a *superposition* (rather than an ensemble) of different decay times. Their interference and the dispersion of the outgoing wave lead to further deviations from an exponential law. Although they are too small to be observed in free decay, they have been confirmed as 'coherent state vector revival' for photons emitted into cavities with reflecting walls (Rempe, Walther and Klein 1987).

The appearance of 'particles' following tracks in a cloud chamber has been explained in Sect. 4.3.1 in terms of an apparent ensemble of narrow wave packets. Similar wave packets occur at counters or photographic plates, thus indicating 'quantum events'. Therefore, the same decoherence which describes localization in space also explains localization in time. Neither particles nor quantum jumps are required as *fundamental* concepts (Zeh 1993, Paz and Zurek 1999). Whenever decay fragments (or the decaying object before decay) interact appropriately with their environment, any interference between 'decayed' and 'not yet decayed' disappears on a very short (though finite) decoherence time scale, just as as in the case of a superposition of Schrödinger cats. This time scale is in general far shorter than the time resolution in genuine measurements. If it is even shorter than relaxation into a linear time dependence (defining a decay rate), decay may be strongly suppressed ('quantum Zeno effect') — see Joos (1984) for its non-phenomenological description.

If the decay status is permanently 'monitored' (in general uncontrollably), a set of identical decaying objects is more appropriately described by a decay rate than by a Schrödinger equation. After their formation this would define an exact exponential law, and it would not admit any interference between different decay times. Similarly, sufficiently distinct decay energies forming an initial superposition are usually absorbed into different final states of the environment. Microscopic systems with their discrete energy levels must therefore decohere into eigenstates of their own Hamiltonians. This explains why stationary states characterize the atomic world, and von Neumann spoke of an *Eingriff* (intervention) required for their change.

It seems that this situation of continuously monitored decay has led to the myth of quantum theory as a stochastic theory for fundamental *quantum events* (cf. Jadczyk 1995). Bohr (1928) remarked that "the essence" (of quantum theory) "may be expressed in the so-called quantum postulate, which attributes *to any atomic process* an essential discontinuity, or rather individuality ..." (my italics). This statement is in conflict with many macroscopic quantum phenomena that have now been observed. Heisenberg and Pauli similarly emphasized their preference for matrix mechanics because of its (evidently misleading) superiority in describing discontinuities. However, according to the Schrödinger equation (and recent experiments), the underlying entanglement processes are smooth. The short decoherence time scales mimic quantum *jumps* between energy eigenstates, while arising narrow wave packets are interpreted as particles or classical variables (even though classical properties have to be restricted in validity by the uncertainty relations in order to comply with the Fourier theorem).

While the description in terms of time-dependent entangled wave functions may now appear as a consistent picture, an important question remains open: how do the probabilities which were required to justify the concept of a density matrix in Sect. 4.2 have to be understood if they do *not* describe quantum jumps or the occurrence of classical properties in fundamental 'events'. This discussion will be resumed in Sect. 4.6.

4.4 Quantum Dynamical Maps

The phenomenological description of *open quantum systems* by means of semigroups leads to some novel aspects compared to their classical counterparts (Sect. 3.4). For example, quantum dynamical maps have been used to formalize von Neumann's 'first intervention' (the reduction of the wave function) as part of the dynamics (cf. Kraus 1971). This is possible, since semigroups can not only describe the transition of pure states into ensembles of measurement results, but also the 'selection' of an *individual* outcome. Otherwise they are equivalent to an entropy-enlarging Zwanzig-type master equation with respect to $\hat{P}_{sub}$ (or its equivalent in terms of path integrals — Feynman and Vernon 1963). Although the 'irrelevant' correlations with the environment, which would arise according to the exact global formalism, represent quantum entanglement, they are usually not distinguished from classical statistical correlations when it comes to applications. (This point of view is equivalent to a popular but insufficient 'naive' interpretation of decoherence, which pretends to derive genuine ensembles).

Quantum theory is sometimes even *defined* as describing open systems by means of a dynamical semigroup, that is as though it were a time-asymmetric local statistical theory. (Hence the term 'statistical operator' for the density operator.) However, this 'minimal statistical interpretation' is absolutely insufficient, as it entirely neglects the difference between genuine and apparent ensembles, and thus all consequences of entanglement beyond the considered systems (quantum nonlocality). The superposition principle has even been claimed to be derivable (cf. Ludwig 1990), although it is then simply re-introduced in some hidden form (for example by changing the laws of statistics).

Semigroups are certainly mathematically elegant and powerful. Therefore, they would form candidates for *new* theories if conventional (Hamiltonian) quantum theory should prove wrong as a universal theory. The question is whether mathematical elegance here warrants physical relevance or is merely convenient within a certain approximation. To quote Lindblad (1976): "It is difficult, however, to give physically plausible conditions ... which rigorously imply a semigroup law of motion for the subsystem. ... Applications ... have led some authors to introduce the semigroup law as the fundamental dynamical postulate for open (non-Hamiltonian) systems." Such a law would *fundamentally* introduce an arrow of time (see Sect. 4.6), but it would depend on the choice of systems (and in some cases contradict experiments that have already been performed).

The simplest quantum systems (such as spinors) are described by a two-dimensional Hilbert space. Their density matrix may be written by means of the Pauli matrices σ_i $(i = 1, 2, 3)$ in the form

$$\varrho = \frac{1}{2}(1 + \boldsymbol{\sigma} \cdot \boldsymbol{\pi}) \quad , \tag{4.36}$$

where the (mathematically) real *polarization vector* $\boldsymbol{\pi} = \mathrm{Trace}\{\boldsymbol{\sigma}\varrho\}$ — that is, the expectation value of *all* spin components — completely defines ϱ as a general Hermitean 2×2 matrix of trace 1. The latter is in turn equivalent to a (genuine or apparent) *ensemble* of orthogonal spinors. The length of $\boldsymbol{\pi}$ is a measure of purity, since $\mathrm{Trace}\{\varrho^2\} = (1 + \boldsymbol{\pi}^2)/2$, with $\boldsymbol{\pi}^2 \leq 1$. A pure state corresponds to a unit polarization vector, while an arbitrary density matrix (a general 'state' in the language of mathematical physics) is characterized by the mean value $\boldsymbol{\pi} = \sum_\alpha p_\alpha \boldsymbol{\pi}_\alpha$ of all unit vectors $\boldsymbol{\pi}_\alpha$ in an ensemble of spinors that may represent this density matrix.

A general trace-preserving linear operator $\hat{P}$ on ϱ must be defined on 1 and $\boldsymbol{\sigma}$ in order to be completely defined:

$$\hat{P}1 := 1 + \boldsymbol{\pi}_0 \cdot \boldsymbol{\sigma} \qquad \hat{P}\boldsymbol{\sigma} := \boldsymbol{A} \cdot \boldsymbol{\sigma} \quad , \tag{4.37}$$

with a real vector $\boldsymbol{\pi}_0$ and a linear vector transformation $\boldsymbol{A}$. $\hat{P}$ is idempotent (a Zwanzig 'projector') if $\boldsymbol{A}^2 = \boldsymbol{A}$ and $\boldsymbol{\pi}_0 \cdot \boldsymbol{A} = 0$ ($\boldsymbol{A} = 0$, for example). If $\boldsymbol{\pi}_0 \neq 0$, $\hat{P}$ creates new information — even from the unit matrix (cf. Sect. 3.2).

Dynamical combination of the projection $\hat{P}$ with a Hamiltonian evolution (describing a rotation of $\boldsymbol{\pi}$) in the form of a master equation leads to the *Bloch equation* for the vector $\boldsymbol{\pi}(t)$,

$$\frac{d\boldsymbol{\pi}}{dt} = \boldsymbol{\omega} \times (\boldsymbol{\pi} - \boldsymbol{\pi}_0) - \sum_i \gamma_i (\pi^i - \pi_0^i) \boldsymbol{e}_i \tag{4.38}$$

in a certain vector basis $\{\boldsymbol{e}_i\}$ (cf. Gorini, Kossakowski and Sudarshan 1976). Values of $\gamma_i < 0$ or $|\boldsymbol{\pi}_0| > 1$ would violate the positivity of the density matrix at some $t > 0$ (cf. Sect. 4.2), and thus have to be excluded.[5] The second term on the RHS describes anisotropic damping towards $\boldsymbol{\pi}_0$. This creation of information may describe very different situations, such as equilibration with an external heat bath of given temperature, or evolution towards a certain measurement result. $\boldsymbol{\pi}_0$ is often referred to as characterizing a *reference state*, while the *relative entropy* with respect to it (cf. (3.51b)) never decreases under the Bloch equation (even when the genuine local entropy does). However, hermiticity of $\hat{P}$ (corresponding to a genuine projection operator) would require $\boldsymbol{\pi}_0 = 0$ and $\boldsymbol{A} = \boldsymbol{A}^\dagger$, that is, a projection of vectors $\boldsymbol{\pi}$ in space.

If the two-dimensional Hilbert space describes something else than spin, such as isotopic spin or a $K, \bar{K}$ system, the polarization vector lives in an *abstract* three-dimensional space, with environmental conditions that cannot simply be 'rotated'. The abstract formalism can also be generalized to n-dimensional Hilbert spaces. For this purpose the Pauli matrices have to be

[5] As mentioned in Sect. 4.2, *all* subsystem density matrices remain positive under a global Hamiltonian dynamics, and even under a collapse of the global state vector. However, this property of 'complete positivity' has to be separately *postulated* for phenomenological quantum dynamical maps (cf. Kraus 1971), thus further illustrating that these maps do *not* describe a fundamental aspect of quantum theory.

replaced with the (n^2-1) Hermitean generators of $SU(n)$, while the real 'coherence vectors' (the generalizations of the polarization vector $\boldsymbol{\pi}$) now live in the vector space spanned by them. For example, SU(3) gives rise to the 'eight-fold way'. The most important difference is that there are now more than one (in fact, $n-1$) *commuting* Hermitean generators. They may contain a nontrivial subset that is decohered under all realistic environmental conditions, and thus may form the center of a phenomenological set of observables (the set of 'classical observables' — cf. Sect. 4.3). For example, maps of density matrices of dimension $n=4$ which decohere with respect to a certain basis would reproduce Figs. 3.8a, c and d.

In the infinite-dimensional Hilbert space of quantum mechanics, the *Wigner function*

$$\begin{aligned} W(p,q) &:= \frac{1}{\pi}\int e^{2ipx}\varrho(q+x,q-x)\,dx \\ &\equiv \frac{1}{2\pi}\int\int \delta\left(q-\frac{z+z'}{2}\right)e^{ip(z-z')}\varrho(z,z')\,dzdz' =: \mathrm{Trace}\{\Sigma_{p,q}\varrho\} \end{aligned} \tag{4.39a}$$

(written in analogy to $\boldsymbol{\pi} = \mathrm{Trace}\{\boldsymbol{\sigma}\varrho\}$) assumes the role of the coherence vector $\boldsymbol{\pi}$. Evidently,

$$\Sigma_{p,q}(z,z') := \frac{1}{2\pi}e^{ip(z-z')}\,\delta(q-\frac{z+z'}{2}) \tag{4.39b}$$

is a generalization of the Pauli matrices (with 'vector' index p,q). On a finite interval of length L, $\Sigma_{p,q}$ would require the additional term $-\frac{1}{2\pi L}\mathrm{e}^{ip(z-z')}$ in order to warrant tracelessness.

Therefore, the Wigner function is a continuous set of expectation values, which form the components (one for each point in phase space) of a generalized coherence vector. This 'vector' of expectation values again characterizes the density matrix ϱ completely, and regardless of its interpretation (Sect. 4.2). It does in general neither represent an individual quantum state nor a probability distribution on phase space. The latter follows from the negative values allowed for W, even though one may calculate all expectation values in the *form* of an ensemble mean, $< F >= \int f(p,q)W(p,q)dpdq$.

Lindblad (1976) was able to generalize the Bloch equation to infinite-dimensional Hilbert spaces. He wrote it (applicable to the density matrix) as

$$i\frac{\partial\varrho}{\partial t} = [H,\varrho] - \frac{i}{2}\sum_k\left(L_k^\dagger L_k\varrho + \varrho L_k^\dagger L_k - 2L_k\varrho L_k^\dagger\right) \quad , \tag{4.40}$$

with arbitrary generators L_k in Hilbert space. It describes *creation* (localization) of information, that is, a local decrease of the corresponding von Neumann entropy, if and only if some generators do not commute with their Hermitean conjugates $L_k^\dagger$. This can be shown by applying the non-Hamiltonian terms of (4.40) to the unit matrix $\varrho = 1$. Otherwise it describes information loss (a genuine Zwanzig projection). This can also be seen from the general representation of a Zwanzig projector in quantum mechanical Hilbert space, $\hat{P}\varrho = \sum_k V_k \varrho V_k^\dagger$, which is analogous to the square root of a positive operator in its eigenbasis for $V_k = V_k^\dagger$. If $L_k^\dagger = L_k$, the Lindblad terms can be written in the form of a double commutator, $L^2\varrho + \varrho L^2 - 2L\varrho L = [L, [L, \varrho]]$. For $L = \sqrt{2\lambda}x$ one recovers (4.32), that is, decoherence in the x-basis, which could be *derived* from unitary interaction with the environment (and shown to be practically unavoidable for macroscopic variables).

One may similarly describe other 'unread' measurements and their corresponding loss of phase relations. However, 'damping' towards a *pure* state (a semigroup proper), represented by the second term of (4.38) with a unit vector $\boldsymbol{\pi}_0$, allows one also to describe the evolution into a (freely chosen) definite measurement outcome (a 'collapse' — in contrast to local *or global* unitary evolution). This can then readily be dynamically combined with a stochastic formalism that represents an appropriate dice (or random number generator) that is defined to select pure states according to the Born-von Neumann probabilities (Bohm and Bub 1966, Pearle 1976, Gisin 1984, Belavkin 1988, Diósi 1988). If applied continuously, such as by means of the Itô process, this formalism describes measurements phenomenologically as smooth indeterministic processes.

Many explicit models have been proposed in the literature (see Stamatescu's Chap. 8 of Giulini et al. 1996). Some of them merely replace the apparent ensemble arising through decoherence for a bounded open system with a *genuine* one (Gisin and Percival 1992). The system under consideration is thereby assumed always to possess its own (in general unknown) state $\psi(t)$ that follows an indeterministic trajectory in its Hilbert space according to a quantum Langevin equation — in conflict with the exact global dynamics. Therefore, this *quantum state diffusion model* is essentially equivalent to what I have above called the 'naive interpretation' of decoherence. It may serve as an intuitive picture (or tool) for many practical purposes *if* (and insofar as) it selects the dynamically robust wave packets (see Diósi and Kiefer 2000). However, it would be seriously misleading if this formalism (based on the concept of a density matrix) gave the impression of *deriving* a real collapse just by taking into account the interaction with the environment.[6]

Many contributions in the literature remain ambiguous about their true intentions, or simply disregard the difference between genuine and apparent ensembles (proper and improper mixtures). In particular, the *quantum state diffusion model* is inappropriate for formulating a new (fundamental) dynamical law (such as an objective collapse). The pure state $\psi(t)$ that it postulates to exist for any open system would in general not define states for its subsystems (which could as well be chosen as *the* system, and would then lead

[6] Note that Schrödinger was worried about *correlated* superpositions of cats, $|0\rangle|dead\rangle + |1\rangle|alive\rangle$. He did not attempt to produce uncorrelated superpositions of cats as in the experiment discussed in Sect. 4.3.4. Therefore, he might not have been entirely satisfied with decoherence, although he would have had to take it into account.

to a different stochastic evolution). The picture of a trajectory of quantum states $\psi(t)$ for a macroscopic system that is not the whole universe is simply in drastic conflict with quantum nonlocality.

Other models have therefore been suggested to reproduce the observed statistical aspects of quantum theory by means of *fundamental* modifications of the Schrödinger equation. Since they cannot remove all entanglement, they cannot describe trajectories of wave functions for all systems. Measurements would then just form special applications of a more general stochastic quantum dynamics, describing an increase of ensemble entropy, while the subsequent reading of the outcome has to be described as in Fig. 3.5. These modifications of a *universal* Schrödinger equation were originally suggested in the form of stochastic 'hits' (jumps), assumed to act in addition to the unitary evolution, and to suppress coherence with growing distance (Ghirardi, Rimini and Weber 1986). They were chosen in such a way that they rarely affect individual particles, but are important for entangled aggregates of many particles in order to describe their classical aspects in an objective way. This proposal was later formulated as a continuous process, as indicated above (Pearle 1989, Ghirardi, Pearle and Rimini 1990). Since these models lead to novel predictions, they can be distinguished from a universally valid Schrödinger equation. In their original form they would either be completely camouflaged by environmental decoherence (Joos 1986, Tegmark 1993), or are ruled out by existing experiments (Pearle and Squires 1994). They cannot be excluded in general, provided one allows them to occur in the observational chain of interactions well *after* environmental decoherence has occurred (see Sect. 4.6). Their precise form would then be hard to guess in the absence of any empirical hints.

Several authors suggested searching for the root of a fundamental quantum indeterminism in gravity. Their main motivation is the apparently classical nature of spacetime curvature. However, as indicated in Sect. 4.3.4 (and further elaborated in Sect. 6.2), spacetime curvature must not be presumed to be classical. Collapse models along these lines have been proposed in a more or less explicit form (Penrose 1986, Károlyházy, Frenkel and Lukácz 1986). They either treat the quantum state of the gravitational field as an environment for matter in a specific quantum state diffusion model (as criticized above — see also Diósi 1987), or they use an arising event horizon as a 'natural' boundary to cut off entanglement dynamically (Hawking 1987, Ellis, Mohanty and Nanopoulos 1989). Even though this specific boundary between a 'system' and its environment may appear fundamental, it could still not be used to define an objective physical process. In particular, a horizon depends on the observer's trajectory (see Sect. 5.3). This proposal certainly does not lead to genuine ensembles of global states, unless explicitly so postulated in an invariant form as a *fundamental modification* of unitary quantum dynamics.

General literature: Alicki and Lendi 1987, Diósi and Lukács 1994, Kupsch's Chap. 7 of Giulini et al. 1996

4.5 Exponential Decay and 'Causality' in Scattering

There are only a few absolutely stable 'particles' (elementary quantum objects), while all others are observed as decaying on vastly different time scales. In quantum theory they can be formally represented by *complex energies* (with a negative imaginary part for the conventional sign of the time-dependent Schrödinger equation). Even though microscopic, they have to be regarded as *open* systems — an excited atom in vacuum, for example, as when coupled to a heat bath of zero temperature. Open space represents an irreversible absorber with infinite capacity according to this initial condition.

On the other hand, the S-matrix with its corresponding complex poles is usually regarded as describing *fundamental* microscopic dynamics (defined independently of any fact-like initial conditions). Exponential decay then seems to characterize a fundamental direction in time (cf. Prigogine 1980), similar to Ritz's retarded electrodynamics (Chap. 2). Since there are no energy eigenstates with complex eigenvalues ('Gamow vectors') in Hilbert space, this situation has even led to the introduction of the concept of 'rigged Hilbert spaces' as a fundamental generalization of the space of quantum mechanical states (Böhm 1978). As will be discussed, however, there is absolutely no reason for introducing new basic concepts. Decaying systems may be well described dynamically in conventional quantum mechanical terms, such as those used so far in this book, provided they are (for good reasons) assumed to decay only *approximately* according to an exponential law.

Exponential decay of a quantity A corresponds to a constant loss rate,

$$\frac{dA}{dt} = -\lambda A \quad , \tag{4.41}$$

for $\lambda > 0$, with a rate of change, dA/dt, thus completely determined by A itself. The asymmetry under time reversal described by this equation could be law-like, although it occurs usually as a consequence of a special initial condition, similar to that characterizing irreversible master equations. If, in particular, A is a conserved quantity, any back-flow that is in principle required must actually be negligible. This would represent a fact-like T-asymmetry that can be explained by means of a sufficiently large and *initially* empty reservoir (comparable to the irrelevant channel used in Sect. 3.2). If recurrence times are sufficiently large, the exponential law (4.41) remains an excellent approximation for a very long time.

This disappearance of a 'substance' A is an entirely classical picture. However, the time dependence (4.41) is best known from stochastic radioactive decay, where A represents the non-decay probability. It is then regarded as the standard example of *quantum indeterminism* — understood as fundamental and law-like. This interpretation of (4.41) would mean that decay *events* occur at unpredictable though definite instants in time (see below and Sect. 4.1.2).

The exponential law (4.41), if valid beyond the 'trivial range' $\lambda t \ll 1$, defines an elementary master equation (3.46a), with a Green's function $\hat{G}_{ret}$ given by the decay rate λ. Its foundation on time-symmetric fundamental dynamics (such as a universal Schrödinger equation) requires the usual assumptions, for example the negligibility of any back-flow into 'doorway states', which are directly coupled to A (cf. Fig. 3.4). Therefore, a conserved quantity has to disappear fast from such doorway states into 'deeper' (dynamically more distant) states, which must form a large reservoir.

A simple model is provided by the T-symmetric finite reaction chain

$$\frac{dA_n}{dt} = -(\lambda_n + \lambda_{n-1})A_n + \lambda_n A_{n+1} + \lambda_{n-1}A_{n-1} \tag{4.42}$$

with $n = 0 \dots N$, $\lambda_{-1} = \lambda_N = 0$, and the (statistically improbable) initial condition $A_{n\neq 0} \approx 0$. $n = 1$ corresponds here to the doorway channel of Sect. 3.2. For $\lambda_0 \ll \lambda_{n\neq 0}$ one obtains

$$\frac{dA_0}{dt} \approx -\lambda_0 A_0 \quad , \tag{4.43}$$

as long as $A_1 \ll A_0$. This requires only $\lambda t \ll N$, rather than $\lambda t \ll 1$, since all $A_{n\neq 0}$ will relax into partial equilibrium $A_{n\neq 0} \approx A_1$ on a short time scale (or just propagate away for $N \to \infty$).

Exponential decay can also be described by a deterministic wave equation on a continuum, where the small transition rate λ_0 is represented by a potential barrier. The Schrödinger equation does not describe the conserved quantity ('probability') itself, but rather its amplitude. An exact general time dependence according to a complex energy, $\psi(t) \propto \exp[-\mathrm{i}(E - \mathrm{i}\gamma)t]$, would not be compatible with unitarity, that is, with the hermiticity of the Hamiltonian. However, it may well represent an approximation that is valid in a bounded though possibly large spacetime region (Khalfin 1958, Petzold 1959, Peres 1980a) — similar to the reaction chain (4.42). Distant regions in space form a large and empty reservoir if a corresponding Sommerfeld condition holds.

In scattering theory, unstable states correspond to poles in the analytically continued S-matrix $S_{nn'}(k)$ at points $k = k_1 - \mathrm{i}k_2$ in the lower right half-plane ($k_1 > 0$ and $k_2 > 0$), where k is the wave number, $k^2 = 2mE$. In the restricted spacetime region where exponential behavior holds (after the incoming flux producing the decaying system has ceased), the wave function is dominated by the *Breit-Wigner* part (i. e., the pole contribution). This requires a (positive) *delay* during the scattering process, which can be described by means of the relevant partial wave $\psi_l(r,t)Y_{lm}(\theta,\phi)$. Its radial factor $\psi_l(r,t)$ may be expanded in terms of energy eigenstates, $\psi_l^{(k)}(r,t) := \phi_{k,l}(r)e^{-i\omega(k)t}$, in the form

$$\begin{aligned}\psi_l(r,t) &= \int_0^\infty f_l(k)\psi_l^{(k)}(r,t)\,dk \\ &\underset{r\to\infty}{\to} \int_0^\infty f_l(k)\frac{\mathrm{e}^{-\mathrm{i}kr} - (-)^l S_l(k)\mathrm{e}^{\mathrm{i}kr}}{r}\mathrm{e}^{-\mathrm{i}\omega(k)t}\,dk \quad , \qquad (4.44)\end{aligned}$$

where $S_l(k) = \mathrm{e}^{2\mathrm{i}\delta_l(k)}$ is the corresponding diagonal element of the S-matrix.

For sufficiently large values of t, the factor $\mathrm{e}^{-\mathrm{i}\omega(k)t}$ oscillates rapidly with k. This leads to destructive interference under the integral, except in regions of r and t where the phase $kr + \omega(k)t$ (for the incoming wave), or $kr - \omega(k)t + 2\delta_l(k)$ (for the outgoing one), is practically independent of k over the width of the wave packet (centered at k_0, say). For the outgoing wave, for example, this means

$$\frac{d}{dk}(kr - \omega(k)t + 2\delta_l(k))|_{k_0} \approx 0 \quad \Rightarrow \quad r \approx \frac{d\omega(k_0)}{dk_0}t - 2\frac{d\delta_l(k_0)}{dk_0} \quad . \qquad (4.45)$$

A noticeable delay compared to propagation with the group velocity $d\omega/dk$ requires a large value of $d\delta_l/dk$, such as in the vicinity of a pole. For sufficiently large times t but not too large distances r from the scattering center, and for initial momentum packets much wider than the resonance width, only the pole contribution remains. For it one may write

$$S_l(k) = \mathrm{e}^{2\mathrm{i}\delta_l(k)} \approx \frac{k - k_1 - \mathrm{i}k_2}{k - k_1 + \mathrm{i}k_2} \quad , \qquad (4.46)$$

and hence $k_0 = k_1$ for the surviving wave packet that represents the decaying state. In this spacetime region, the contribution of the pole to (4.46) is given by its residue, hence

$$\begin{aligned}\psi_l(r,t) \underset{t\to\infty}{\to} & -(-)^l f_l(k_1)\int_0^\infty \frac{k - k_1 - \mathrm{i}k_2}{k - k_1 + \mathrm{i}k_2}\frac{\mathrm{e}^{\mathrm{i}[kr - \omega(k)t]}}{r}\,dk \\ &\approx (-)^l 2\pi k_2 f_l(k_1)\frac{\mathrm{e}^{\mathrm{i}[k_1 r - \omega(k_1)t]}}{r}\exp\left[k_2\left(r - \frac{d\omega(k_1)}{dk_1}t\right)\right] \qquad (4.47)\end{aligned}$$

(assuming $k_2 \ll k_1$). In the last factor one recognizes the ‘imaginary part of the energy’, $\gamma = k_2\,d\omega(k_1)/dk_1$.

A positive delay (a ‘retardation’) of the scattered wave at the resonance requires

$$\frac{d\delta_l}{dk} \approx -\frac{d}{dk}\left(-\mathrm{arctg}\frac{k_2}{k - k_1}\right) = \frac{k_2}{(k - k_1)^2 + k_2^2} > 0 \quad . \qquad (4.48)$$

The pole must therefore reside in the lower half-plane. This condition is often referred to as *causality in scattering*, since the retardation specifies a direction in time related to intuitive causality. This position of the poles is also used for deriving *dispersion relations* in T- or TCP-symmetric quantum field theory. However, no time direction can result from the structure of the S-matrix, since this structure is a unique consequence of the time reversal-invariant Hamiltonian. The condition $k_2 > 0$ describes a fact-like direction in time that would be reversed for scattering states with a reversed boundary condition. One would thus have to fix the outgoing wave rather than the incoming one as a narrow packet in time or distance. Such exponentially *growing* states would require the 'decay products' to *come in* as coherent and exponentially increasing waves during a sufficient (usually very long) span of time. A similar asymmetry, caused by boundary conditions, occurred in the retarded radiation reaction of extended charges (2.30). Therefore, *the time arrow of exponential decay is determined by the boundary condition contained in* (4.44), *rather than by the position of poles in the S-matrix.* The latter represents the exact dynamics with all its symmetries.

The investigation of wave packets for massive particles in free space beyond the pure pole contribution (Petzold 1959, Winter 1961, Peres 1980a) shows that deviations from the exponential law become essential for small and for very large times. In the extreme limit (when the decaying wave function has become unobservably small), exponential decay is replaced with a power law. This conclusion is in conflict with the idea of local objects being stochastically emitted into infinite space. It must instead be interpreted as a coherent back-flow (dispersion) of the outgoing wave packet. While representing a very small effect in absolute terms, this back-flow is even further reduced (or at least decohered) whenever the decay products interact with surrounding matter. In the usual case of strong coupling to the environment (corresponding to measurement or absorption), the exponential law holds for all times, provided 'quantum causality' remains valid. On the other hand, decay by emission of weakly interacting photons inside the reflecting walls of a cavity leads to detectable deviations from exponential decay. This has been observed as 'coherent revival' of the decaying state (see Rempe, Walther and Klein 1987, Haroche and Kleppner 1989), and in other situations (Wilkinson et al. 1997). These deviations from the exponential law depend crucially on the density of available final states. For example, coherent decay into a single final state is well known for leading to harmonic oscillation.

The Breit-Wigner contribution (4.47) describes a non-normalizable wave packet. This result is an artifact of the pure pole approximation. The normalized state is correctly described by the wave packet $f_l(k)$ in (4.44), which has been replaced by a constant in (4.47). In an exact treatment, its square-integrable tails warrant normalizability even for large t by correcting the pure Breit-Wigner contribution.

In general, the scattering process would have to be described by a normalized time-dependent density matrix,

$$\varrho_{lm,l'm'}(r,r';t) \underset{r\to\infty}{\longrightarrow} \int_0^\infty\int_0^\infty \varrho_{lm,l'm'}(k,k')\frac{\mathrm{e}^{-\mathrm{i}kr}-(-)^l S_l(k)\mathrm{e}^{\mathrm{i}kr}}{r} \times \frac{\mathrm{e}^{\mathrm{i}k'r'}-(-)^{l'} S^*_{l'}(k')\mathrm{e}^{-\mathrm{i}k'r'}}{r'}\mathrm{e}^{-\mathrm{i}[\omega(k)-\omega(k')]t}\,dk\,dk' \quad , \tag{4.49}$$

with $\varrho_{lm,l'm'}(k,k')$ determined by a preparation procedure. After completion of the direct scattering process, and in the case of a resonance in the l_0-wave, this leads approximately to

$$\varrho_{lm,l'm'}(r,r';t) \underset{t\to\infty}{\longrightarrow} \delta_{ll_0}\delta_{l'l_0}\varrho_{l_0m,l_0m'}(k_1,k_1) \times \int_0^\infty\int_0^\infty \frac{k-k_1-\mathrm{i}k_2}{k-k_1+\mathrm{i}k_2}\frac{k'-k_1+\mathrm{i}k_2}{k'-k_1-\mathrm{i}k_2}\frac{\mathrm{e}^{\mathrm{i}(kr-k'r')}}{rr'}\mathrm{e}^{-\mathrm{i}[\omega(k)-\omega(k')]t}\,dk\,dk' \tag{4.50}$$

in the relevant space-time region. This approximation describes again a *pure* Breit-Wigner wave packet (or at most a mixture of magnetic quantum numbers in the case of rotational symmetry).

Hence, there are no exactly exponential states (Gamow vectors) which would require or justify the 'rigging' of the Hilbert space of quantum mechanics. Similarly, there are no exact energy (or particle number) eigenstates in reality, since their infinite exponential tails according to $\exp(-\sqrt{-2mE}r)$ can never form within finite time (in accordance with the time-energy uncertainty relation). Otherwise, for example if exact energy eigenstates did occur by means of instantaneous quantum jumps, they would lead to superluminal effects (as has even been found surprising — see Hegerfeldt 1994). Even nonrelativistically, stable or decaying states must form dynamically, that is, in accordance with a time-dependent Schrödinger equation. This holds similarly for controlled entanglement (as suggested for 'teleportation' experiments — cf. Sect. 4.2). An exactly bounded spatial support for a relativistic quantum state (such as the light cone describing 'relativistic causality') requires small (practically unobservable) uncertainties in energy and particle number, related to the Casimir effect or the Unruh radiation (see Sect. 5.2). Neglecting this consequence of quantum field theory (by arguing solely in terms of single-particle states, for example) leads to inconsistencies which illustrate the danger of remaining formal while not discriminating between fundamental and phenomenological concepts. (Cf. also Bohm, Doebner and Kielanowski 1998 for apparent time-asymmetries induced by presuming macroscopic *operations*.)

Radioactive decay is in practice investigated by means of collections of many identical decaying systems (atoms, say). If these atoms are distinguishable by their position, their total state may be described as a direct product.

According to the Schrödinger equation, each factor state will then evolve into a normalized superposition of its initial state (with an exponentially decreasing amplitude) and a direct product of its final state with a state for the emitted particle (including the latter's effect on the environment). The *precise* time-dependence of these components depends on what happens to wave function of the decay fragments (whether it is reflected or absorbed, or propagates towards infinity — as discussed above).

If the emitted particles (now assumed *not* to be reflected somewhere) are regarded as part of the environment (and thus traced out), the remaining N-atom density matrix describes a time-dependent *apparent ensemble* of product states with n atoms, say, in their final, and $N - n$ atoms in their initial state. Their probabilities in this ensemble, $p_n(t)$, reflects the exponential time-dependence of the Schrödinger amplitudes. (In general, components with the same number n but *differing* individual decayed atoms also decohere from one another — see Sect. 4.3.4). This apparent ensemble, represented by the probabilities $p_n(t)$, is dynamically well described by a master equation. Therefore, it is equivalent to an ensemble of solutions of a stochastic (Langevin type) equation that describes individual 'histories' $n(t)$. In terms of a universal Schrödinger equation, the number of decayed atoms n is a 'robust' property in the sense of Sect. 4.3.2 if decay is monitored by the environment. The various branches of the wave function arising by the fast but smooth action of decoherence then describe individual histories $n(t)$ which represent successions of almost discrete quantum jumps (as discussed in Sect. 4.3.5).

4.6 The Time Arrow of Various Interpretations of Quantum Theory

> The truth could not be worth much
> if everybody was a little bit right

One often hears the remark that physicists who completely agree about all applications of quantum mechanics may not only differ entirely about its interpretation, but even on what the problem is and whether there really is one! It seems that the deepest discrepancies of this kind are still rooted in the difference between Heisenberg's and Schrödinger's approaches. While Heisenberg kept relying on fundamental classical concepts (suggesting only a limitation of their 'certainty'), Schrödinger described microscopic physical states by ('certain') wave functions on an (in general high-dimensional) classical configuration space, which would thereby replace three-dimensional space as the 'arena of reality'.[7] Whether wave functions or particular operators (representing 'observables') carry the dynamical time-dependence is merely a *consequence* of the thus chosen 'picture'.

Although both pictures are known to be equivalent in some pragmatic sense, most physicists seem to subscribe to one or the other (or perhaps a variant thereof) when it comes to interpretations ('probabilities for *what*?'). Typically, in the Schrödinger picture one regards the collapse of the wave function as a dynamical process, while in the Heisenberg picture it is viewed as an (extra-physical) increase of 'human knowledge'. However, any foundation of the concept of information or knowledge would require an *ensemble*. I hope that keeping this difference in mind for the rest of this section may help to avoid some misunderstandings that often lead to emotional debate. One should therefore give a consistent account of what is pragmatically *done* when the theory is so successfully applied. Which concepts are *fundamentally required*, rather than being approximately derived, or mere tradition and prejudice?

For example, physical entropy, which quantifies irreversibility by its increase, is in quantum statistical mechanics defined by means of von Neumann's functional of the density matrix (4.4). According to Sect. 4.2, it measures the size of (genuine or apparent) ensembles of mutually orthogonal (hence operationally distinguishable) *wave functions*. While only *genuine* ensembles represent incomplete information, the time-dependence of the density matrix determines that of local entropy in general. Conservation of global von Neumann entropy reflects the unitarity of the von Neumann equation (when applicable) — equivalent to the unitarity *and* determinism of the Schrödinger equation. No ensemble of *classical* nor any other variables representing the potential values of observables is 'counted' by von Neumann's entropy. Fig. 3.5, characterizing classical measurements, can therefore *not* be applied to quantum measurements. In terms of quantum states it has to be replaced with Fig. 4.3, which contains a collapse of the wave function. The transition from a superposition to an ensemble (depicted by the second step) affects the final value of von Neumann's 'ensemble' entropy (that would otherwise be reduced by the increase of information, as in Fig. 3.5).

A wave function *and* a set of classical configurations are used in Bohm's quantum theory (Bohm 1952, Bohm and Hiley 1993). This theory is often

[7] This contrast between the Heisenberg and the Schrödinger picture does *not* represent the 'dualism' between two competing classical concepts (particles and waves *in space*) that is often claimed to be part of *one* (the Copenhagen) interpretation. In classical theory, particle positions and wave amplitudes (fields) characterize *different physical objects*, both required to define states (or *configurations*) of general physical systems. A dualism (or 'complementarity'), used to characterize quantum objects, should simply be regarded as a conceptual inconsistency (while it is often interpreted as reflecting the absence of a microscopic physical 'reality'). Applying such a dualism dynamically (with *varying* concepts just as required) could readily be used to 'derive' the indeterministic master equations of Sect. 4.1.2. A critical historical account of the early development of these conceptual aspects of quantum theory can be found in Beller (1996, 1999).

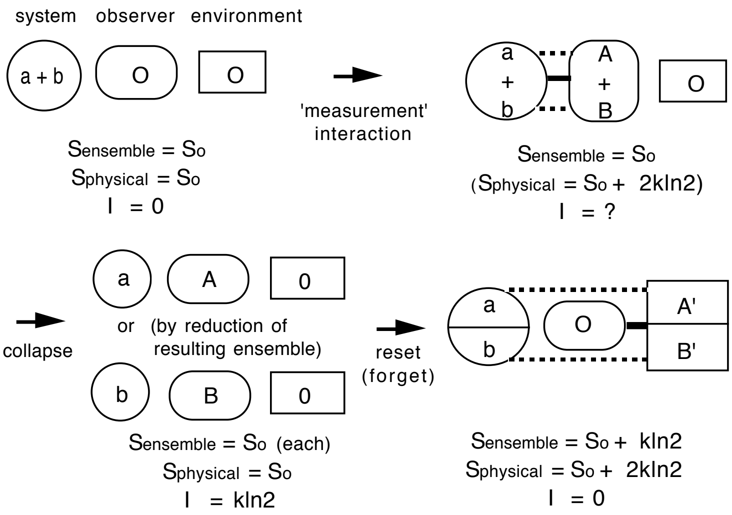


Fig. 4.3. Quantum measurement of a *superposition* $|a\rangle + |b\rangle$ by means of a *collapse* process, here assumed to be triggered by the macroscopic pointer position. The initial entropy S_0 is smaller by one bit than in Fig. 3.5 (and may in principle vanish), since there is no initial ensemble a/b for the property to be measured. Dashed lines before the collapse now represent quantum entanglement. (Compare the ensemble entropies with those of Fig. 3.5!) The collapse itself is often divided into two steps (see (4.51) below): first increasing the ensemble entropy by replacing the superposition with an ensemble, and then lowering it by reducing the ensemble (applying the 'or' — for macroscopic pointers only). The increase of ensemble entropy, observed in the final state of the Figure, is a consequence of the first step of the collapse. It brings the entropy up to its classical initial value of Fig. 3.5. (From Chap. 2 of Giulini et al. 1996.)

praised for exactly reproducing all predictions of conventional quantum theory in a deterministic way. However, this is not surprising, since it leaves Schrödinger's wave function entirely untouched, while the classical trajectories, which *determine* all observed quantities according to this model, remain unobservable ('surrealistic') in order to reproduce the empirically required quantum indeterminism by means of their *presumed* statistical distribution.

Although wave functions and points in configuration space are here both assumed to be real, [8] they are treated quite differently. While the former is usually regarded as 'given', the latter is always represented by an ensemble

[8] "No one can understand this theory until he is willing to think of ψ as a real objective field rather than just a 'probability amplitude' " (Bell 1981). This remark reflects Bell's conclusion drawn from the violation of his inequality (nonlocality). It applies also to the objective collapse theories which he later preferred in order to avoid the 'extravagant' or 'empty' wave function components that would exist according to Everett's as well as Bohm's theories. No classical concepts are fundamentally used in collapse theories.

(without thereby contributing to the entropy). Any increase of knowledge in a quantum measurement would here be taken into account by calculating the entropy (by means of the usual expression $S[\hat{P}|\psi\rangle\langle\psi|]$) from the *projection* of the total wave function onto the corresponding subspace (that is, from a *component* ψ) — even though there is no objective collapse in this theory. A justification of this asymmetric treatment of wave functions and classical variables seems hard to find (cf. Zeh 1999b). Moreover, this theory destroys the powerful symmetry under canonical transformations. Only positions represent their observed values, while velocities are unobservable, and the observed values of apparent momenta *arise* during measurements in an uncontrollable way. Bohm's trajectories in configuration space include the classical electromagnetic fields rather than photons (see Holland 1993). This has further painful consequences for other boson fields, or for 'intrinsic' degrees of freedom which do not define a classical configuration space.

Regardless of any interpretation, the density matrix of local macroscopic systems (such as measurement devices) is dynamically affected by decoherence; this is the most efficient and most ubiquitous irreversible process in nature. The process of decoherence is based on 'quantum causality', that is, on the future irrelevance of nonlocal entanglement — a kinematical aspect of *wave functions*. In this way, decoherence is able to *explain* classical properties and (apparent) stochastic quantum events in terms of smoothly but indeterministically evolving local wave packets.

While decoherence thus eliminates the basic motivation for the Heisenberg-Bohr interpretation (that *presumes* classical concepts, and their values as 'coming into being' during fundamental irreversible events — see Pauli's remark on the last page of the Introduction), this aspect of decoherence may not have been duly appreciated even by some authors who significantly contributed to it (Omnès 1988, 1992, Gell-Mann and Hartle 1990, 1993 — see also Omnès 1998). Guided by the Heisenberg picture they investigate consequences of decoherence on certain 'consistency conditions' (originally proposed by Griffiths 1984), which are assumed to regulate the applicability (or various kinds of 'truth') of varying classical concepts at successive (usually discrete) times within variable intervals of uncertainty. However, environmental decoherence allows one to *derive* quasi-classical concepts (all that is needed) in terms *wave packets* — close to what Schrödinger had originally in mind. Their apparent ensembles, formally described by the density matrix, obey master equations (such as (4.32)), and *in this way* consistently define probabilities for quasi-classical 'histories' (see Chap. 2 of Giulini et al. 1996).

In contrast to Bohm's or Everett's theory (see below), Griffith's consistent histories presume an *absolute* quantum arrow (see Hartle 1998) — just as collapse models do. Their selection by a formal 'consistency' requirement may indeed be inconsistent itself (Kent 1997), while *derived* probabilistic master equations do not hold exactly. These consistent histories have also been claimed to be equivalent to trajectories *of wave functions* that are assumed, according to the *quantum state diffusion model*, to exist for *all* systems. However, such an assumption is in conflict with quantum nonlocality (Diósi et al. 1995, Brun 2000 — cf. Sect. 4.4).

According to the theory of decoherence, physical objects may *appear* in terms of (more or less certain) values of 'observables', while they *act* (dynamically) as wave functions (for example by determining the scattering amplitudes of projectiles probing them). In this sense the wave function is 'actual', while quasi-classical variables *emerge* from it by means of decoherence. Any interpretation of quantum theory that does not assume the wave function to be 'real' has to *explain* its 'actuality'. In the Heisenberg picture one would have to explain coherence and nonlocality — not decoherence and localization. Most of the 'surprising' quantum phenomena that have been observed (and that have all been *predicted* from the superposition principle or the wave function) represent 'paradoxes' only when described by classical or other local concepts.

Wave functions are usually rejected as representing reality because they are not defined on three-dimensional space. However, precisely this nonlocality has been experimentally confirmed in many ways. This can only have surprised believers in particles or hidden variables. Although entanglement has been known since the advent of wave mechanics, it has until recently rarely been regarded as an aspect of reality. The wave function would then merely represent a dynamical tool that may even lose meaning after an 'individual atomic process' (or an experiment) ended. However, entanglement is crucial for many observable consequences, for example the very accurately known ground state energy of the Helium atom, total angular momentum of two combined systems, or the fractional quantum Hall effect. It has now been confirmed to exist, and to remain observable, over very large distances (Tittel et al. 1998).

Even Schrödinger (1935), in the title of an important paper about entanglement, referred quite insuffiently to 'probability relations' rather than nonlocal states. Quantum nonlocality is not only generic, but evidently *actual* in the literal sense used above: it has real *individual* (not only statistical) consequences. We should therefore face it as part of reality instead of retreating to Copenhagen irrationalism. In its uncontrollable ('noisy') and unavoidable form (Zeh 1970) it describes decoherence. Traditionally, these effects have been misinterpreted as 'distortion by collisions' (or momentum transfer by 'kicks' — see Knight 1998 for his critical remarks).

On the other hand, the concept of apparent ensembles of wave functions, or the density matrix in general, *presumes* a probability interpretation (as has been explained in Sect. 4.2), while the question 'probabilities for (or information about) *what*?' has not yet been answered at a fundamental level. The success of decoherence as well as many other quantum phenomena seem to favor the answer: wave packets in a high-dimensional 'configuration' space. For example, the consistent description of a spot on the photographic plate (regarded as a 'pointer') requires the concept of an altered *wave function* in order to take into account the physico-chemical basis of this spot. Similarly, a hydrogen atom has to be assumed to 'be' in a quantum state rather than in one of the classical configurations contained in a corresponding ensemble. The strongest support yet for this conclusion seems to come from an analysis of the decoherence of neuronal states in the brain (Tegmark 2000, see also Zeh 2000). These quasi-classical neuronal states would form the final link (or the ultimate pointer basis) in a chain of observational interactions, all describable in terms of interacting wave functions.

This situation was evidently the reason for postulating a collapse of the wave function as a real dynamical process (von Neumann's 'first intervention'). Despite frequent claims to the contrary, the collapse has never been part of the Copenhagen interpretation, even though quantum jumps are used as an argument against a wave function representing reality. The collapse may be formulated in two steps (cf. Fig. 4.3):

$$|\psi\rangle\langle\psi| = \underbrace{\sum_{mn} |\phi_m\rangle c_m c_n^* \langle\phi_n|}_{S=0} \;\rightarrow\; \underbrace{\sum_n |\phi_n\rangle |c_n|^2 \langle\phi_n|}_{S \geq 0} \;\rightarrow\; \underbrace{|\phi_{n_0}\rangle\langle\phi_{n_0}|}_{S=0\ .} \quad , \qquad \text{with}\ S=0 \rightarrow S\geq 0 \rightarrow S=0 \tag{4.51}$$

They represent (1) the replacement of a pure state with an ensemble (characterized by an increase of von Neumann's ensemble entropy), and (2) the selection of a specific state (thus lowering the ensemble entropy as depicted by Fig. 3.5). The first step can be described by a master equation that represents information loss, while the complete stochastic process (4.51) corresponds to a Langevin equation for wave functions (Sect. 4.4). Since the latter describes an individual physical process, it has to be used when calculating the changing *physical* entropy according to (3.54), while the master equation describes an ensemble that represents entanglement as well as lacking information.

If macroscopic properties α, say, are again regarded as 'always given' (as in Sect. 3.3.1), *physical entropy* can be characterized by a function $S(\alpha) = k \ln N_\alpha$, similar to the last term of (3.54), where N_α is now the dimension of the subspace representing a fixed value (or interval) of α. The time dependence of this entropy,

$$S(t) = S(\alpha(t)) \quad , \tag{4.52}$$

is then determined by the macroscopic dynamics $\alpha(t)$, which in general includes the collapse of the wave function, that is, its projection onto subspaces corresponding to definite macroscopic properties. This means that *the stochastic collapse is part of the macroscopic dynamics* $\alpha(t)$ — regardless of the interpretation of the wave function. It is dynamically objectivized by the process of decoherence, describing an *apparent* collapse. In contrast to classical Hamiltonian dynamics, which determines the time dependence of any macroscopic quantity, $\alpha(t) := \alpha(p(t), q(t))$ (cf. Sect. 3.3.1), the Schrödinger equation does *not* determine $\alpha(t)$. Therefore, the dynamics (4.30) cannot just describe a transformation of negentropy into information in accordance with the negentropy principle (3.58). In other words, not even *macroscopically* different states have to possess (macroscopically or microscopically) different predecessors ('sufficient reasons'). Such causal predecessors would have to be counted by the initial ensemble entropy as in Fig. 3.5.

Does the collapse, if *used* in this way as part of the dynamics of wave functions, now specify an arrow of time that could even be responsible for irreversible thermodynamics? The ensemble entropy (4.4) would increase only during the auxiliary first step of the collapse (describing incomplete knowledge of the outcome). For entangled systems, an individual ('real') collapse,

$$\psi = \sum_n c_n \phi_n \Phi_n \underset{t}{\rightarrow} \phi_{n_0} \Phi_{n_0} \quad , \tag{4.53}$$

occurring after an interaction of type (4.30), for example, does *not* alter the ensemble entropy. However, it specifies an arrow in so far as it transforms the entangled state into a factorizing one. Therefore, the additive ('physical') entropy *decreases* for this process (after it may have correspondingly *increased* during the interaction (4.30)), since

$$S[\hat{P}_{\text{sep}} |\psi\rangle\langle\psi|] \geq S[\hat{P}_{\text{sep}} |\phi_{n_0} \Phi_{n_0}\rangle\langle\phi_{n_0} \Phi_{n_0}|] = 0 \quad . \tag{4.54}$$

An objective collapse has never been observed as a *dynamical* process. It has nonetheless to be *taken into account* regardless of its interpretation before (or, at least, when) the observer becomes aware of the macroscopic pointer position. As he thereby becomes himself quantum correlated with the pointer state Φ_{n_0}, the corresponding 'state of being conscious' in his brain also becomes a pure state in this respect as a consequence (if not the cause) of this collapse. This 'observer state' (whatever it may be in detail) can thus be used for postulating a psycho-physical parallelism in accordance with von Neumann's intentions (see also London and Bauer 1939). At this point it is particularly helpful that neuronal states, which characterize *objective* brain activities, are themselves subject to decoherence (Tegmark 2000). There is no need for genuine classical (or any other novel) variables *anywhere in between* the observed microscopic system and the brain of the observer. On the other hand, spontaneous localization of the pointer position in an apparatus would not by itself lead to an entirely pure state of being conscious. While this description of observations by conscious beings remains vague in detail, it is all that is *in principle* required for a physical formulation in terms of wave functions. Although a 'real' collapse is usually assumed to occur as soon as the relevant phase relations have irreversibly become inaccessible in the environment, this assumption is merely convenient, as it comes close to a classical description.

Does the entropy-reducing collapse, wherever it may occur (or simply be *applied*), then have to be regarded as a quantum mechanical revival of Maxwell's demon, who could classically be exorcized in Sect. 3.3.2? Lubkin (1987) demonstrated that, just as in the classical case, this entropy decrease according to the collapse cannot be utilized in a cyclic process that would allow the construction of a perpetuum mobile of the second kind. However, it may have essential consequences during the unique process of cosmic evolution (see Sect. 6.1).

It is noteworthy that the conventional collapse contains an extension of *intuitive causality* to a region where by assumption it cannot be confirmed. In its usual form, the expression $|\langle\phi'_k|\phi_n(\Delta t)\rangle|^2$ (with $\phi_n(\Delta t) := \exp(-iH\Delta t)\phi_n$) defines the probability of finding the state ϕ'_k in an appropriate measurement at time $t_2 = t_1 + \Delta t$, provided the system was found in the state ϕ_n in a previous measurement (of the first kind) at time t_1. This interpretation assumes the wave function did collapse to the state ϕ_n during the first measurement, and then evolves unitarily according to the Hamiltonian H until it is measured again. An equivalent (though unconventional) time-reversed interpretation may be obtained from the identity of the above matrix element with $\langle\phi'_k(-\Delta t)|\phi_n\rangle$. This means that the wave function could as well be assumed to collapse during the first measurement in an 'acausal' manner from a 'conspiratorial' (advanced) state ϕ_n into a state $\phi'_k(-\Delta t)$ that will then evolve according to the Schrödinger equation into the state $\phi'_k = \phi'_k(0)$ just *before* the second measurement starts (see also Footnote 4 and Penrose 1979). These two versions of the collapse are indicated in Fig. 4.4. The second one is counterintuitive, since the observed system would have to 'know in advance' what kind of measurement will be performed, and when. Its exclusion is therefore again an application of *intuitive causality*, while the equivalence of the two descriptions is based on the T-symmetry of the formal quantum probabilities (Aharonov, Bergmann and Lebowitz 1964).

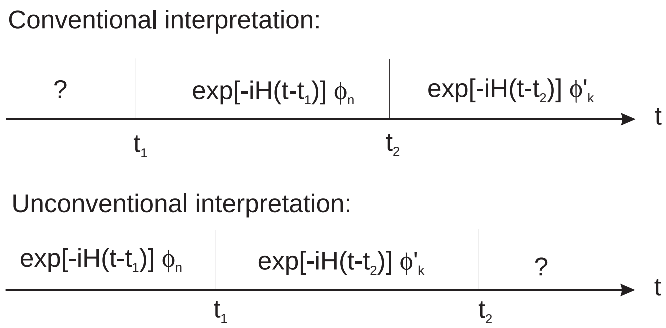


Fig. 4.4. Behavior of the wave function in the case of retarded (conventional) and advanced collapse. In contrast to classical waves, the choice of the usual ('retarded') interpretation of the collapse is a matter of pure convention

In contrast to the advanced electromagnetic fields (Sect. 2.4), which can be excluded empirically by means of small test charges, our preference for the causal version of the collapse is a pure matter of convention. However, 'macroscopic' registration devices Φ that are continuously monitored by the environment (in accordance with the time arrow of quantum causality) have to be assumed to possess objective pointer positions at all times, similar to a classical field but usually including a 'memory'. In this case, their states during two successive measurements have to be described, regardless of the chosen collapse version, by $\Phi_0(t)$ for $t < t_1$, by $\Phi_n(t)$ for $t_1 < t < t_2$, and by $\Phi_{nk}(t)$ for $t > t_2$ (see also Bläsi and Hardy 1995). In the *unconventional* collapse version, for example, the evolution of the complete system is depicted by Fig. 4.5.

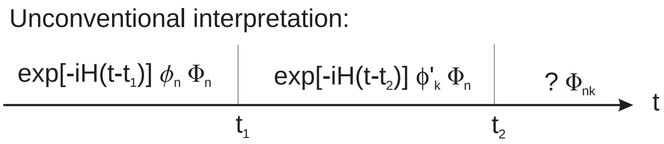


Fig. 4.5. Behavior of the total wave function under the unconventional assumption of advanced collapse and 'continuous measurement' of the pointer position Φ. Only the wave function of the microscopic system ϕ, but not that of the macroscopic pointer, depends on the convention

If the wave function were assumed to represent an *ensemble of some real and local variables* (including the classical ones), the possibility of their *postselection* by means of later measurements would lead to further 'surprising' consequences (cf. Aharonov et al. 1987, Peres 1994, Aharonov and Vaidman 1996, Kastner 1998). They add to other paradoxes that occur in this class of interpretations, such as the superluminal influence which would be required to explain quantum nonlocality. In the Schrödinger picture there simply *is no ensemble* for postselecting from. Rather, the wave function is assumed to collapse in an indeterministic and T-asymmetric way. In its conventional retarded form this cannot lead to any advanced consequences (not even unobservable ones) which would mimic postselection.

Choosing the retarded collapse in order to conform with intuitive causality even in the unobservable range of the quantum world model is an example of the *heuristic fictions* that are widely used in science (Poincaré 1902, Vaihinger 1911). In principle, the choice of the retarded electromagnetic fields (2.36) also served merely to simplify the description, since the uncontrollable 'distant' sources forming the cosmic absorber can then be neglected. However, it is a nontrivial *empirical fact* that (2.36), but not (2.38), leads to a simplification. In contrast, the freely evolving quantum state can by assumption *not* be confirmed in this sense.

These ambiguities in the dynamics of wave functions can be avoided if the asymmetric collapse is assumed not to affect the physical reality that is assumed to be described by the wave function. According to a proposal first seriously put forth by Everett, *all* components of an entangled superposition, as it may form in a measurement according to (4.30), are assumed to 'exist' as *one* universal superposition. There is then no law-like quantum arrow of time — but how can there be measurements with definite results?

It is appropriate to distinguish Everett's original *relative state interpretation* from the *many worlds interpretation* (DeWitt 1971, Deutsch 1997), which is essentially based on the Heisenberg picture, or alternatively on an *ensemble* of Feynman paths in configuration space (see Sokolovski 1998). In the many-worlds interpretation there is a time-dependent (dynamically growing) ensemble of many *classical* 'worlds'. This is *not* meant to describe lacking information about one 'real' world, but all worlds in the ensemble are assumed to exist simultaneously. A reduction of this ensemble can then only reflect the observer's indeterministically evolving identity (defined by his *position* in this 'multiverse'). This interpretation will be further discussed in the context of quantum gravity (towards the end of Sect. 6.2.2), where any time dependence disappears from the fundamental description.

Everett's interpretation is instead based on the Schrödinger picture: "This paper proposes to regard pure wave mechanics as a complete theory." (Everett 1957, see also Zeh 1970, 1973, 1979). In this description *there are no observables* on a fundamental level (only interactions in terms of a Schrödinger equation). Hence, there is no Heisenberg picture and no (certain or uncertain) classical world. It must therefore be conceptually compared with, but dynamically contrasted to, collapse models, which also describe pointer positions and states of observers in terms of wave packets.[9] The ultimate observer states (χ_n^{obs}, say) need not even represent quasi-classical states. Everett's conclusion from his assumption that the Schrödinger equation be *always* valid was that all components of (4.30) — but not any classical paths or histories — remain in existence in *one* superposition, $\sum \psi^{(n)} \chi_n^{obs}$. ("All components are actual.") The point is that they can be *experienced* only separately because of their different robust states χ_n^{obs}, which are here *postulated* on empirical grounds to represent subjective observers. (No definite observer state can be defined in the *global* Everett wave function.)

Since configuration space assumes the role of space as the arena of reality in all versions of the Schrödinger picture, the *fork of indeterminism* that seems to characterize quantum theory is here re-interpreted as a *fork of causality* (see Footnote 1 of Chap. 2). All observer states χ_n^{obs} (with different n) remember the same pre-measurement history. To all of them, the rest of the world is described by their respective 'relative states' $\psi^{(n)}$ (their co-factor states, which these observers may normalize for convenience). According to the global Schrödinger equation, these relative states represent that part of the quantum world that the pure observer states exclusively interact with once decoherence has dynamically decoupled the corresponding 'branches'. For example, when position of a quantum object that *is* in a broad

[9] Von Neumann used phenomenological observables in order to define his measurement interactions (4.30). Their algebra is fundamental only for constructing a specific Hilbert space and the quantum Hamiltonian by means of the quantization rules of Sect. 4.1.1.

wave function is measured by means of a macroscopic measurement device, its localized components form different quasi-classical 'worlds' (dynamically independent components with their own observer states). The formal 'plus' of a superposition is thereby objectively reduced to an effective 'and', while the 'or' has only subjective meaning (though objectivized with respect to correlated versions of different observers).

Only if both factors of all components of the Everett wave function were defined to be mutually orthogonal, would the sum of their products, $\sum \psi^{(n)} \chi_n^{obs}$, resemble the Schmidt representation (4.25) (Kübler and Zeh 1973, Zeh 1973, 1979, Albrecht 1992, 1993). Moreover, this representation would depend on the precise border line between observer and the rest of the world. Everett considered relative states with respect to general states of general subsystems, and he assumed the components to be dynamically autonomous after a measurement. However, only *continuous measurement* through the environment leads to decoherence, and defines robust pointer states or branches (cf. Tegmark 1998). Macroscopic systems are not dynamically isolated after completion of an interaction (in contrast to microscopic systems after completion of a scattering process).

Since 'the other' robust components, which exist according to the Schrödinger equation, are not observable to 'us', the assumption of their existence is operationally meaningless, and the Everett interpretation is thus indistinguishable from the collapse interpretation. However, the collapse is then equally meaningless. In the Everett interpretation this equivalence requires the additional assumption that components only *branch* into further ones with increasing time, but practically never (re)combine according to

$$\sum_n c_n \phi_n \Phi_n \quad \underset{t}{\rightarrow} \quad \left(\sum_n c_n \phi_n \right) \Phi_0 \quad . \tag{4.55}$$

Because of the T-symmetric Schrödinger dynamics, this requires that there are no suitable (conspiratorial) components $n \neq n_0$ available as an initial condition on the LHS in the real quantum world that is assumed to be described by the Everett wave function. The law-like arrow of the collapse thus becomes a fact-like arrow. This 'quantum causality' is analogous to Boltzmann's *Stoßzahlansatz*, which assumes that correlations are irrelevant *after* being formed in collisions. (This fact-like quantum arrow would also be required in Bohm's theory, as its 'empty' branches of the wave function may otherwise lead to recoherence, and guide the classical trajectory along *irreproducibly arising* new wave packets.)

One may postulate an appropriate initial condition by requiring that *all* existing nonlocal quantum correlations (such as those on the LHS of (4.55)) are retarded, that is, have been *caused* somewhere during the past history of the quantum universe. At the big bang ($t = 0$, say) one would then have something like

$$\psi(t \to +0) = \lim_{\Delta V_k \to 0} \prod_k \psi^0_{\Delta V_k} \quad , \tag{4.56}$$

with definite quantum states $\psi_{\Delta V_k}$ on small volume elements ΔV_k. The projector on this initial state is defined to be invariant under the Zwanzig projection $\hat{P}_{local}$. The assumption (4.56) appears 'natural' only to our causal prejudice (thus applying the *double standard* in Price's terminology). In contrast to its classical counterpart of initially absent (or future-irrelevant) correlations, this condition is not a *statistical* one (meaningful for an ensemble), but a condition to the real universal quantum state. It would explain both the absence of recoherence and the validity of thermodynamical master equations.

Since the 'Everett branching' is thus not based on an *objective* dynamical law (such as a collapse), it does not define a precise algorithm (as required by Kent and McElwaine 1997). The definition of Everett branches is instead related to von Neumann's problem of defining a psycho-physical parallelism in quantum mechanical terms (that is, the above-mentioned observer states ψ_{obs} — see Zeh 1970, 1979, 2000, Squires 1990, Lockwood 1996).

Even though Everett branches can be sufficiently defined, they do not yet possess any probabilities, since they are *all* assumed to exist once (with different albeit intrinsically meaningless norms). Everett's claim that the probability interpretation is a *consequence* of the formalism does not seem to be entirely justified. Therefore, Graham (1970) considered *series* of N equivalent measurements (N subsequent branchings — similar to the n decay events discussed at the end of Sect. 4.5). He was able to show that the total norm of all those resulting Everett branches which represent series of outcomes that differ significantly from the formal quantum mechanical probabilities vanishes in the limit $N \to \infty$ (see also Jammer 1974). While this result permits a compact formulation of the probability postulate (by assuming merely that we happen to live in an Everett branch of non-negligible norm), it *requires* that all branches contribute to this norm precisely according to their probabilities that are to be derived.[10] While stochastic collapse models postulate quantum probabilities as part of their dynamics, the Everett interpretation has *equivalently* to assume that 'we' are living in a branch that has been selected in accordance with the Hilbert space norm (see also Sect. 6.2.2).

General literature: Jammer 1974, Busch, Lahti and Mittelstaedt 1991, d'Espagnat 1995, Chap. 2 of Giulini et al. 1996

[10] This may be confirmed by the statistical distribution of frequencies of results obtained in N subsequent measurements which distinguish between states $|1\rangle$ and $|2\rangle$ in the same (independently prepared) initial superpositions $a|1\rangle + b|2\rangle$. Since the *number* of branches which contain n 'spin-ups', say, is given by the binomial coefficient $\binom{N}{n}$ regardless of a and b, the distribution of outcomes n, $p(n)$, forms a Poisson distribution centered at the required value only if each branch component contributes to the norm with weight $|a|^{2n}|b|^{2(N-n)}$. This is precisely its weight according to the probability postulate. (Example provided by Erich Joos)

5. The Time Arrow of Spacetime Geometry

– Prelim. 4th edtn. (Jan 01): www.time-direction.de

In the framework of general relativity, gravity is a consequence of spacetime curvature. Its dynamical laws are still symmetric under time reversal. However, if the actual global spacetime structure defines an arrow of time (as in an expanding universe), this must be reflected by the dynamics of all matter. While this had been well known, it came as a surprise during the early seventies that the most strongly gravitating systems possess properties that have to be described in thermodynamical terms, thus indicating an intimate connection between these apparently quite different fields of physics.

Gravitating systems are thermodynamically peculiar even in Newton's theory, since they possess negative heat capacities as a consequence of the universal attractivity of this force. In particular, forces depending homogeneously on the minus second power of distance, such as gravity and Coulomb forces, lead according to the virial theorem to the relation

$$\overline{E_{\mathrm{kin}}} = -\tfrac{1}{2}\overline{E_{\mathrm{pot}}} = -E \tag{5.1}$$

between the mean values of kinetic and potential energies, and therefore between them and the total energy. This virial theorem is usually valid in a statistical (ergodic) sense only. For bound systems it applies to mean values over a (quasi-) period of the motion, for quasi-bound systems it holds approximately for mean values over sufficiently large times. In quantum theory it applies to *expectation values* of proper (normalizable) energy eigenstates, as can be shown by using Fock's *ansatz* $\psi(\lambda\boldsymbol{r}_1, \ldots, \lambda\boldsymbol{r}_N)$ and the homogeneity of T and V in a variational procedure $\delta(\langle\psi|T + V|\psi\rangle/\langle\psi|\psi\rangle) = 0$ with respect to λ. The theorem must then also hold for mixtures which are diagonal in energy eigenstates, or (approximately) if non-diagonal terms can be neglected because of random phases. (For relativistic generalizations of the virial theorem see Gourgoulkon and Bonazzola 1994).

The anti-intuitive negative sign relating kinetic and total energy in (5.1) means, for example, that satellites are *accelerated* by weak friction in the earth's atmosphere, and that stars *heat up* by radiating energy away. (This second example is valid only as far as the quantum mechanical zero-point energy does not dominate $\overline{E_{\mathrm{kin}}} = \mathrm{Trace}\{\varrho T\}$ — as it does in white dwarf stars or solid bodies. Early astrophysicists believed that stars always cool down after exhausting their fuel.) It also means that the heat flow from hot to cold objects that are governed by gravity causes a thermal inhomogeneity to *grow*.

As an example, consider two monatomic ideal gases with entropies given according to (3.14) by

$$S_i = kN_i \left(\tfrac{3}{2}\ln T_i - \ln \varrho_i + C_i\right) \tag{5.2}$$

(with $i = 1, 2$, and constants C_i). Since the internal energy, $U = \overline{E_{\mathrm{kin}}}$, is here given by $U = (3/2)NkT$, the net change of entropy resulting from an exchange of energy, $\delta U_1 = -\delta U_2$, and of particles, $\delta N_1 = -\delta N_2$, becomes for fixed volumes V_i or for fixed densities $\varrho_i = N_i/V_i$

$$\delta S_{\mathrm{total}} = \delta S_1 + \delta S_2 = \left(\frac{1}{T_1} - \frac{1}{T_2}\right)\delta U_1 + k\left(\frac{3}{2}\ln\frac{T_1}{T_2} - \ln\frac{\varrho_1}{\varrho_2}\right)\delta N_1 \quad . \tag{5.3}$$

This describes entropy changes δS_1 and δS_2 with opposite sign, but cancelling only in thermodynamical equilibrium ($T_1 = T_2$ and $\varrho_1 = \varrho_2$). An entropy increase in accordance with the Second Law is in this normal situation achieved by a *reduction* of thermal and density inhomogeneities (aside from the transient *thermo-mechanical effect*, that is, a thermally induced pressure difference due to the temperature dependence of the second term).

However, the density of a star is not a free variable that can be kept fixed. Since a normal star, which will here for simplicity be assumed to be in thermal equilibrium, may in very good approximation be described as an ideal gas, its volume is related to the potential energy, and by means of the virial theorem then also to the temperature, according to

$$NT \propto U = \overline{E_{\mathrm{kin}}} \propto -\overline{E_{\mathrm{pot}}} \propto \frac{N^2}{R} \propto \frac{N^2}{V^{1/3}} \quad , \tag{5.4}$$

that is, $V \propto N^3/T^3$. The entropy (5.2) of a star is therefore

$$\begin{aligned} S_{\mathrm{star}} &= kN\left(\tfrac{3}{2}\ln T - \ln N + \ln V + C\right) \\ &= kN\left(-\tfrac{3}{2}\ln T + 2\ln N + C'\right) \quad . \end{aligned} \tag{5.5}$$

In the second line the signs of $\ln T$ and $\ln N$ are reversed. The total entropy change, $\delta S_{\mathrm{star}} + \delta S_{\mathrm{gas}}$, of a star embedded in an interstellar gas becomes (using the virial theorem $E_{\mathrm{star}} = -U_{\mathrm{star}}$ again)

$$\delta S_{\mathrm{total}} = \left(\frac{1}{T_{\mathrm{star}}} - \frac{1}{T_{\mathrm{gas}}}\right)\delta E_{\mathrm{star}} + k\ln\left(\frac{C'' N_{\mathrm{star}}^2 \varrho_{\mathrm{gas}}}{(T_{\mathrm{star}}T_{\mathrm{gas}})^{3/2}}\right)\delta N_{\mathrm{star}} \quad . \tag{5.3$'$}$$

While heat still flows from the hot star into the cold gas according to the Second Law, this leads now to further increase of the star's temperature, and to accretion of matter provided N_{star} is large enough. Thermal and density inhomogeneities thus grow in the generic astrophysical situation, although 'pathological' (non-ergodic) solutions do exist, such as gravitationally collapsing spherical matter shells or pressure-free dust spheres, where the virial theorem does not hold.

This demonstrates that the gravitational contraction of normal stars is dynamically controlled by thermodynamics, but *not* by gravity itself. If the thermodynamical arrow of time did change direction in a recontracting universe (as first suggested by Gold 1962 — see Sect. 5.3), stars and similar gravitating objects would have to re-expand during that epoch.

A homogeneous universe thus describes a state of very low entropy (improbable but 'simple' in the sense discused at the end of Chap. 3). This leads to the question whether its potential inhomogeneous contraction under gravitational forces represents an entropy capacity that is sufficient to explain the observed thermodynamical arrow of time. The *Kaltgeburt* could then be replaced with a *homogeneous birth*, while inhomogeneous contraction leads to the required thermodynamical non-equilibrium.

In order to estimate the improbability (negentropy) of a homogeneous universe, one has to know the maximum entropy that can be gained from gravitational contraction. Possible limits of the negative heat capacity are:

a) *Quantum degeneracy* (primarily of electrons). It is essential for the stability of solid gravitating bodies and of white dwarfs. By emitting heat these objects can cool down 'normally' rather than heating up.
b) *Repulsive short range forces.* They may be important in neutron stars, while they have similar consequences as a degenerate electron gas.
c) *Gravitation* itself. Even in Newton's theory, any radiation with bounded propagation velocity that is affected by gravity cannot escape from the surface of a sufficiently massive object. If this speed limit is as universal as gravity (as it is according to the theory of relativity), any further collapse of the mass distribution remains *irrelevant* to an external observer. The maximum radius from which light may escape for a given mass defines an *event horizon.* Matter disappearing behind the horizon — whatever will happen to it — cannot participate any more in the thermodynamics of the universe, while only its external gravitational field remains observable.

Such *non-relativistic black holes* were discussed as early as 1795 by Laplace, and even before him by J. Mitchel. In general relativity, black holes are described by specific spacetime structures. This theory leads to the further consequence that neither of the first two bounds to gravitational contraction may prevent an object of sufficiently large mass (that may always be formed by accretion of matter) from collapsing into a black hole. In case (b), repulsive forces give rise to a positive potential energy, which must eventually dominate as a source of gravity, while for (a) the increasing zero point pressure of the fermions forces them to combine into effective bosons, with all of them being able to occupy the same spatial state.

Therefore, black holes define an upper limit for the entropy production by gravitational contraction of matter from the point of view of an external observer. But what *is* the entropy of a black hole? This question cannot be answered by the investigation of relativistic stars, that is, of equilibrium systems, since the essential stages of the collapse proceed irreversibly. However, a unique and finite answer is obtained from a quantum aspect of black holes, *viz.* their Hawking radiation (Sect. 5.1).

Since the curvature of space represents the gravitational degrees of freedom in general relativity, it may carry entropy by itself. The dynamics of spatial curvature (regarded as the time-dependent state — see Sect. 5.4) is described by Einstein's equations

$$G_{\mu\nu} = 8\pi T_{\mu\nu} \quad , \tag{5.6}$$

(in units of $G = c = 1$), where $T_{\mu\nu}$ is the energy-momentum tensor of matter. These equations define an initial (or final) value problem again, since they are generically of hyperbolic type (cf. Sect. 2.1). The Einstein tensor $G_{\mu\nu}$ is a linear combination of the components of the Ricci tensor $R_{\mu\nu} := R^{\lambda}{}_{\mu\lambda\nu}$, that is, a trace of the Riemann curvature tensor. Forming this trace is analogous to forming the d'Alembertian in the wave equation (2.1) for the electromagnetic potential from the tensor of its second derivatives $\partial_\nu\partial_\lambda A^\mu$. Aside from the nonlinearities responsible for the self-interaction of gravity, the Riemann curvature tensor is similarly defined by the second derivatives of the metric $g_{\mu\nu}$, which thus assumes the role of the gravitational potential (analogous to A^μ in electrodynamics). In both cases, the trace of the tensor of derivatives is determined locally by the sources, while its trace-free parts represent the local field variables, which can be chosen freely as initial conditions.

Penrose (1969, 1981) used this freedom to conjecture that the trace-free part of the curvature tensor (the *Weyl tensor*) was zero when the universe began. This situation describes a 'vacuum state of gravity', that is, a state of minimum gravitational entropy, and a space as flat as is compatible with the sources. It is analogous to the cosmic initial condition $A^\mu_{\mathrm{in}} = 0$ for the electromagnetic field discussed in Sect. 2.2 (with Gauß' law as a similar constraint). Gravity would then represent an exactly retarded field, requiring 'causes' in the form of advanced sources. Since Penrose intends to explain the thermodynamical arrow from this initial condition (see Sect. 5.3), his conjecture revives Ritz's position in his controversy with Einstein (mentioned in Chap. 2) by applying it to gravity rather than electrodynamics.

In the big bang scenario, the beginning of the universe is characterized by a time-like curvature singularity (where time itself begins). Penrose used this fact to postulate his Weyl tensor hypothesis to past singularities in general, since this would allow only *one* (homogenous) past singularity (the big bang itself). However, in the absence of an *absolute* direction of time the past would then be distinguished from the future precisely and solely by this boundary condition. If the Weyl tensor condition could be derived from other assumptions (that did not *ad hoc* specify an asymmetry in time), it would have to be valid for future singularities as well.

5.1 Thermodynamics of Black Holes

In order to discuss the spacetime geometry of black holes, it is convenient to consider the static and spherically symmetric vacuum solution discovered by Schwarzschild, and originally expected to represent a point mass. In terms of spherical spatial coordinates this solution is described by the metric

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \left(1 - \frac{2M}{r}\right)^{-1} dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2) \quad . \tag{5.7}$$

Here, r measures the size of a sphere (but not the distance from $r = 0$). This metric form is singular at $r = 0$ and $r = 2M$, but the second singularity, at the *Schwarzschild radius* $r = 2M$, is merely the result of an inappropriate choice of these coordinates. The condition $r = 2M$ describes a surface with area $A = 4\pi(2M)^2$ (using Planck units $G = c = \hbar = k_{\mathrm{B}} = 1$ from now on). In its interior (that is, for $r < 2M$) one has $g_{tt} = 2M/r - 1 > 0$ and $g_{rr} = (1 - 2M/r)^{-1} < 0$. Therefore, r and t interchange their physical meaning as spatial and temporal coordinates. The internal solution is *not* static, while the genuine singularity at $r = 0$ represents a time-like boundary rather than the space point expected by Schwarzschild.

Physical (time-like or light-like) world lines, that is, curves with $ds^2 \leq 0$, hence with $(dr/dt)^2 \leq (1 - 2M/r)^2 \to 0$ for $r \to 2M$, can only approach the Schwarzschild radius parallel to the t-axis (see Fig. 5.1). Therefore, the interior region $r < 2M$ is physically accessible only via $t \to +\infty$ or $t \to -\infty$, albeit within finite proper time. These world lines can be extended regularly into the interior when t goes beyond $\pm\infty$. Their proper times continue into the finite future (for $t > +\infty$) or past (for $t < -\infty$) with the new time coordinate $r \to 0$. There are hence *two* internal regions (II and IV in the figure), with their own singularities at $r = 0$ (at a finite distance in proper times). These internal regions must in turn each have access to a new external region, also in their past or future, respectively, via different Schwarzschild surfaces at $r = 2M$, but with opposite signs of $t = \pm\infty$. There, proper times have to decrease with growing t. These two new external regions may then be identified with one another in the simplest possible topology (region III in the figure).

This complete Schwarzschild geometry may be regularly described by means of the *Kruskal-Szekeres coordinates* u and v, which eliminate the coordinate singularity at $r = 2M$. In the external region they are related to the Schwarzschild coordinates r and t by

$$u = \sqrt{\frac{r}{2M} - 1}\, \mathrm{e}^{r/4M} \cosh\left(\frac{t}{4M}\right) \tag{5.8a}$$

$$v = \sqrt{\frac{r}{2M} - 1}\, \mathrm{e}^{r/4M} \sinh\left(\frac{t}{4M}\right) \quad . \tag{5.8b}$$

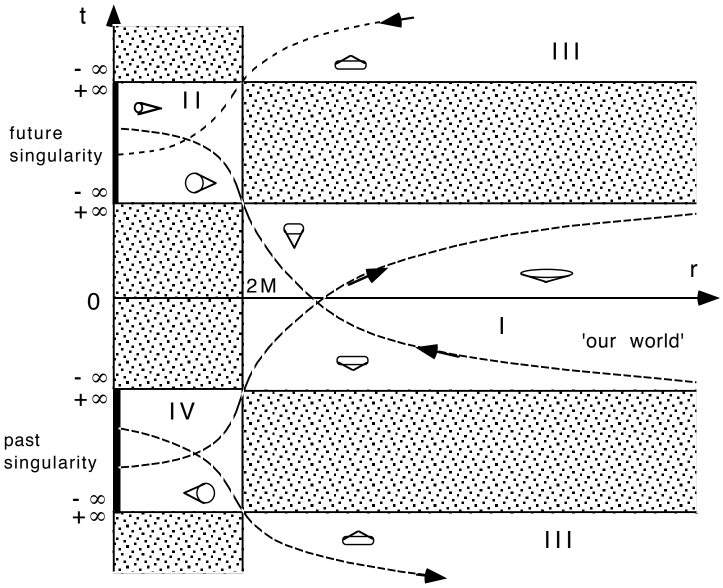


Fig. 5.1. Extension of the Schwarzschild solution from 'our world' beyond the *two* coordinate singularities at $r = 2M$, $t = \pm\infty$. Each point in the diagram represents a 2-sphere of size $4\pi r^2$. A consistent orientation of forward light cones (required from the continuation of physical orbits, such as those represented by dashed lines) is indicated in the different regions. There are also *two* genuine curvature singularities with coordinate values $r = 0$

The Schwarzschild metric in terms of these new coordinates reads

$$ds^2 = \frac{32M^2}{r} \mathrm{e}^{-r/2M}(-dv^2 + du^2) + r^2(d\theta^2 + \sin^2\theta\, d\phi^2) \quad , \tag{5.9}$$

with $r = r(u, v)$. It is evidently regular for $r \to 2M$ and $t \to \pm\infty$, where u and v may remain finite. The Kruskal coordinates are chosen in such a way that future light cones everywhere form an angle of 45° around the $+v$-direction (see Fig. 5.2). Sector I is again the external region outside the Schwarzschild radius ('our world'). One also recognizes the two distinct internal regions II and IV (connected only through the 'instantaneous sphere' that is represented by the origin, $u = v = 0$) with their two separate singularities $r = 0$. Both Schwarzschild surfaces are light-like, and thus represent one-way passages for physical orbits. Their interpretation as past and future horizons is now evident. Sector III represents the second asymptotically flat 'universe'. (It is *not* connected with the original one by a rotation in space, since u is not limited to positive values like a radial coordinate.)

This vacuum solution of the Einstein equations is clearly T-symmetric, that is, symmetric under reflection at the hyperplane $v = 0$ (or any other

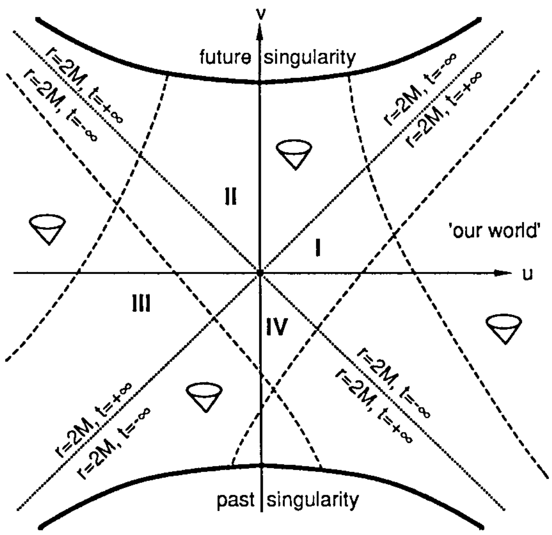


Fig. 5.2. Completed Schwarzschild solution represented in terms of Kruskal coordinates. Forward light cones appear now everywhere with a 45° opening angle around the $+v$-direction. Horizons are indicated by dense-dotted lines, possible orbits as dashed lines. Although the future horizon, say, moves in the outward direction with the speed of light from an inertial point of view, it does not increase in size. The center of the Kruskal diagram defines an 'instantaneous sphere' as a symmetry center, even though it does not specify a specific *external* time t_0

hyperplane $t =$ constant). Therefore, it does neither represent a black hole, nor would it be compatible with the Weyl tensor hypothesis. In the absence of gravitational sources, the Ricci tensor must vanish according to the Einstein equations (5.6), while a singular curvature can only be due to the rest of the Riemann tensor (the Weyl tensor).

A black hole is instead defined as an asymmetric spacetime structure that *arises* dynamically by the gravitational collapse of matter. For example, if the infalling geodetic sphere indicated by the dashed line passing through Sectors I and II of Fig. 5.2 represents the collapsing surface of a spherically symmetric star, the vacuum solution is valid only outside it. Neither a past horizon with its past singularity, nor a second asymptotically flat spacetime exist in this case. They are indeed excluded by regular initial conditions. The coordinates u and v can be extended into the interior only with different interpretation (see Fig. 5.3a, where $u = 0$ is chosen as the center of the collapsing star). This black hole is drastically asymmetric under time reversal, as it contains only a future horizon and a future singularity.

Because of the symmetry of the Einstein equations, a time-reversed black hole — not very appropriately called a *white hole* (Fig. 5.3b) — must also represent a solution. However, its existence in nature would be excluded by the Weyl tensor hypothesis. If it were the precise mirror image of a black

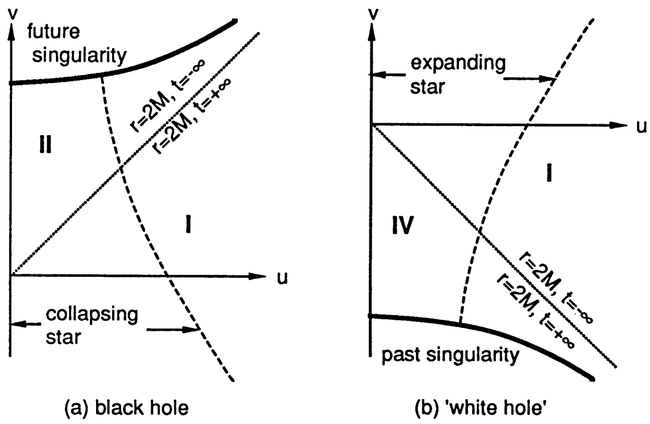


Fig. 5.3a,b. Geometry of a Schwarzschild *black hole* (a) which forms by the gravitational collapse of a spherically symmetric mass, and its time-reverse (b) — usually called a *white hole*

hole, the white hole could describe a star (perhaps with planets carrying life) emerging from a past horizon. This would be inconsistent with an arrow of time that is valid everywhere in the external region. If a white hole were allowed to exist, we could receive light from its singularity, although this light would be able to carry retarded *information* about the vicinity of the singularity only if our arrow of time extended into this region. This seems to be required for thermodynamical consistency, but may be in conflict with an initial singularity (see Sect. 5.3).

Similar to past singularities, also space-like singularities — so-called *naked singularities* — could be 'visible' to us. They, too, were assumed to be absent by Penrose. However, this *Cosmic Censorship* assumption cannot generally be imposed directly as an initial condition. Rather, it has to be understood as a conjecture about the nature of singularities which may *form dynamically* during a collapse from generic initial value data which comply with the Weyl tensor hypothesis. Although counterexamples (in which naked singularities form during a gravitational collapse from appropriate initial conditions) have been explicitly constructed, they seem either to form sets of measure zero (which could be selected by imposing exact symmetries that would be thermodynamically unstable in the presence of quantum matter fields), or to remain hidden behind black hole horizons (see Wald 1997, Brady, Moss and Myers 1998). In the first case, they may be compared with pathological mechanical systems that have occasionally been considered as counterexamples to ergodic behavior. This analogy may readily indicate a relationship between these aspects of general relativity and statistical thermodynamics.

The Schwarzschild-Kruskal metric may be generalized to become the *Kerr-Newman metric*, which describes axially symmetric black holes with angular momentum J and charge Q. This solution is of fundamental importance, since its external part characterizes the final stage of *any* gravitationally collapsing object. For $t \to +\infty$ (although very soon in excellent approximation during a collapse) every black hole may be completely described by the three parameters M, J and Q (except for Lorentz transformations and translations). This result is known as the *no-hair theorem*, as it means that black holes cannot maintain any external structure ('no hair'). It requires that the collapsing star radiates away all higher multipoles of matter and charge (in accordance with a Sommerfeld radiation condition), while conserved quantities connected with short-range forces, such as the lepton or baryon number, disappear from observability. A *white hole* would therefore require coherently incoming (advanced) radiation in order to 'grow hair'. For this reason, *white holes seem to be incompatible with the radiation arrow of our world.* (For their fate — if they had ever come into existence — see Eardley 1974, or Barrabès, Brady and Poisson 1993).

If the internal structure of a black hole is regarded as *irrelevant* for our future in the sense of statistical thermodynamics (in accordance with the no-hair theorem), the gravitational collapse *appears* to violate baryon and lepton number conservation. Even the entropy carried by collapsing matter would disappear from this point of view, in violation of the Second Law. A 'real' violation of any of these conservation (or non-conservation) laws would occur at the singularity that according to a *singularity theorem* must always arise behind a future horizon in the presence of 'normal' matter (see Hawking and Ellis 1973).

Spacetime singularities have particularly dramatic consequences in quantum theory because of the latter's kinematical nonlocality (cf. Sect. 4.2). Consider a global quantum state, propagating on space-like hypersurfaces ('simultaneities'), which define an arbitrary foliation of spacetime and thereby a time coordinate t. If these hypersurfaces somewhere met a singularity, not only the state on this singularity, but also its entanglement with the rest of the universe would cease to exist. In classical description, correlations can exist only as a consequence of incomplete information. Quantum mechanically, definite states would remain left for the non-singular rest of the universe only if the complete state approached the factorizing form $\psi = \psi_{\text{singularity}}\psi_{\text{elsewhere}}$ whenever a singularity forms. Unless all correlations with arising singularities vanished conspiratorially in this way (thus representing a strong final condition), the non-singular part of the universe would *for objective reasons* have to be described by a density matrix ϱ rather than a pure state. Therefore, several authors have argued that quantum gravity must violate unitarity and CPT invariance, and this idea has created a popular speculation ground for introducing an objective dynamical collapse of the wave function induced by gravity (Wald 1980, Penrose 1986, Károlyházy, Frenkel and Lukácz 1986, Diósi 1987, Ellis, Mohanty and Nanopoulos 1989, Percival 1997, Hawking and Ross 1997). Future horizons could then well explain the first step of (4.51).

This indeterminism of a global state vector would not only be inconsistent with canonical quantum gravity (Sect. 6.2), it may also be avoided in quantum field theory on a *classical* spacetime if the foliation defining a time coordinate is always chosen never to encounter a singularity. For example, the spherical Schwarzschild-Kruskal metric could be foliated according to Schwarzschild time t in the external region, and according to the new time coordinate $r < 2M$ (rather than the Kruskal time coordinate v) in its interior. This choice, which would always leave the whole black hole interior in *our* infinite future, has been advocated by 't Hooft (1990). A general singularity-free foliation is given by *York time*, which is defined by hypersurfaces of constant extrinsic spatial curvature scalar K (that describes the embedding into spacetime — see Qadir and Wheeler 1985). Such a foliation appears reasonable, in particular since consequences of an elusive unified field theory may become relevant close to curvature singularities.

This salvation of *global* unitarity is irrelevant to (local) observers who remain outside the horizon, since the reality accessible to them can be completely described by a density matrix ϱ_{ext} in the sense of a Zwanzig projection $\hat{P}_{\text{sub}}$ (cf. (4.26)) — regardless of the choice of a global foliation. The non-unitary dynamics of these density matrices has the same origin as it did in the quantum mechanical subsystems of Sect. 3.3: entanglement. The horizon appears only as a natural objectivization of the boundary, valid for all external observers. One may therefore appropriately describe the phenomenological properties of black holes (including their Hawking radiation — see below) regardless of their singularities. (The latter may even signal the need for a *new* theory). In a universe that is compatible with the Weyl tensor condition, this 'effective non-unitarity' of black holes mimics an indeterminism that would represent a quantum mechanical arrow of time, although this consequence does not *require* black holes and horizons (cf. Chap. 4 and Sect. 6.2.3).

From the point of view of an external observer, the information about matter collapsing under the influence of gravity becomes irreversibly irrelevant, except for the conserved observable quantities M, J and Q. However, the mass of a Kerr-Newman black hole is not completely lost (even if Hawking radiation is neglected). Its rotational and electromagnetic contributions can be recovered by means of a process discovered by Penrose (1969) — see Fig. 5.4. It requires boosting a rocket in the *ergosphere*, the region between the Kerr-Newman horizon, $r_+ := M + \sqrt{M^2 - Q^2 - (J/M)^2}$, and the *static limit*, $r_0(\theta) := M + \sqrt{M^2 - Q^2 - (J/M)^2 \cos^2\theta}$. In this ergosphere, the cyclic coordinate ϕ becomes time-like ($g_{\phi\phi} < 0$) as a consequence of extreme relativistic frame dragging. Because of the properties of this metric, ejecta from the booster which fall into the horizon may possess negative energy with respect to an asymptotic frame (even though this energy is locally positive). Similar arguments hold if the ejecta carry electric charge with a sign opposite to the black hole charge Q. In both cases the mass of the black hole may be reduced by reducing its charge or angular momentum.

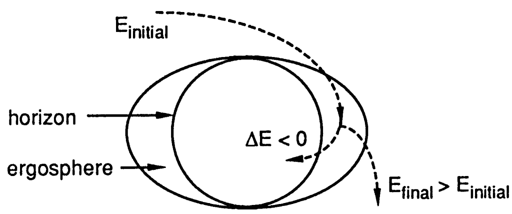


Fig. 5.4. Extraction of rotational energy from a black hole by means of the *Penrose mechanism*, using a booster in the *ergosphere* close to the horizon

The efficiency of this process of drawing energy from a black hole is limited — precisely as for a heat engine. According to a geometro-dynamical theorem (Hawking and Ellis 1973), the area A of a future horizon (or the sum of several such horizon areas) may never decrease. For all known processes which involve black holes, this can be written in analogy to thermodynamics as

$$dM = dM_{\rm irrev} + \varOmega\, dJ + \varPhi\, dQ \tag{5.10}$$

(Christodoulou 1970), where the 'irreversible mass change' $dM_{\rm irrev} \geq 0$ is given by the change of total area of future horizons, $dM_{\rm irrev} = \frac{\kappa}{8\pi}\, dA$. Here, κ is the *surface gravity*, which turns out to be constant on the horizon. $\varPhi$ is the electrostatic potential at the horizon, and $\varOmega$ the angular velocity defined by the dragging of inertial frames at the horizon. The last two terms in this equation describe work done reversibly at the black hole by adding angular momentum or charge, while the first one is analogous to $T\, dS$ because of the inequality $dA/dt \geq 0$. All quantities are defined not with respect to a local frame (where they may diverge), but rather with respect to an asymptotic rest frame of the black hole (where they remain regular because of their diverging red-shift). For the Schwarzschild metric, the surface gravity is $\kappa = 1/4M$. The quantities $\varPhi$ and $\varOmega$ are also constant on the horizon, in analogy to other thermodynamical equilibrium parameters, such as pressure and chemical potential, which appear in the expression for the work done on a thermodynamical system in the form $\mu dN - pdV$.

These analogies led to the proposal of the following *Laws of Black Hole Dynamics*, which form a complete analogy to the Laws of Thermodynamics (cf. Bekenstein 1973, Bardeen, Carter and Hawking 1973, Israel 1986):

0. The surface gravity of a black hole must approach an equilibrium value $\kappa(M, Q, J)$ everywhere on the horizon for $t \to \infty$.
1. The total energy of black holes and external matter, measured from asymptotically flat infinity, is constant in time.
2. The sum of the surface areas of all horizons, $A := \sum_i A(M_i, Q_i, J_i)$, never decreases:

$$\frac{dA}{dt} \geq 0 \quad . \tag{5.11}$$

3. It is impossible to reduce the surface gravity to zero by a finite number of physical operations.

The analogy between this version of the third law and its thermodynamical counterpart may have to be modified for other versions because of the negative heat capacity of a black hole. In particular, the surface area A does not disappear with vanishing surface gravity in a similar way as the entropy does with vanishing temperature.

Bekenstein conjectured that these analogies are not just formal, but indicate genuine thermodynamical properties of black holes. He proposed not only a complete *equivalence of thermodynamical and spacetime-geometrical laws and concepts*, but even their *unification*. In particular, in order to 'legalize' the transformation of thermodynamical entropy into black hole entropy A (when dropping hot matter into a black hole), he required that instead of the two separate Second Laws, $dS/dt \geq 0$ and $dA/dt \geq 0$, there be only one *Unified Second Law*

$$\frac{d(S+\alpha A)}{dt} \geq 0 \quad , \tag{5.12}$$

with an appropriate constant α. Its value remains undetermined from the analogy, since the term $\frac{\kappa}{8\pi}\,dA$, equivalent to $T\,dS$, may as well be written as $\frac{\kappa}{8\pi\alpha}\,d(\alpha A)$. The *black hole temperature* $T_{\mathrm{bh}} := \frac{\kappa}{8\pi\alpha}$ must classically be expected to vanish, since the black hole would otherwise have to emit heat radiation proportional to AT_{bh}^4 according to Stefan and Boltzmann's law. The constant α should therefore be infinite, and so should the *black hole entropy* $S_{\mathrm{bh}} := \alpha A$. This would require the absence of processes, otherwise allowed by the Unified Second Law, wherein black hole entropy is transformed into thermodynamical entropy (this would violate the area theorem).

Nonetheless, Bekenstein suggested a finite value for α (of the order of unity in Planck units). This was confirmed by means of quantum field theory by Hawking's (1975) prediction of *black hole radiation*. His calculation revealed that black holes must emit heat radiation according to the value $\alpha = 1/4$. This process may be described by means of virtual particles with negative energy tunnelling from a virtual ergosphere into the singularity (York 1983), while their correlated partners with positive energy may then propagate towards infinity. (Again, all energy values refer to an asymptotic frame of reference). The probabilities for these processes lead precisely to a black body radiation with temperature

$$T_{\mathrm{bh}} = \frac{\kappa}{2\pi} \quad , \tag{5.13}$$

and therefore to the black hole entropy [1]

[1] It is important to realize that this is a general result — independent of the precise nature of contributing fields. Therefore, it cannot be used to support any specific theory (such as M-theory). Physical theories can only be confirmed by comparison with *empirical* data.

$$S_{\mathrm{bh}} = \frac{A}{4} \quad . \tag{5.14}$$

The mean wave length of the emitted radiation is of the order $A^{1/2}$.

A black hole not coupled to any quantum fields ($\alpha = \infty$) would possess the temperature $T = 0$ and infinite entropy, corresponding to an ideal absorber in the sense of Sect. 2.2. This is indeed a general property of *classical black body radiation*, that is, of classical electromagnetic waves in thermal equilibrium (Gould 1987). According to (5.13), a black hole of solar mass would possess a temperature $T_{\mathrm{bh}} \approx 10^{-6}\,\mathrm{K}$. In the presence of the 2.7 K background radiation, it would absorb far more energy than it emits (even in the absence of dust or any other matter). Only a black hole of less than about 3×10^{-7} solar masses is sufficiently small, thus hot enough to *lose* mass at the present temperature of the universe (Hawking 1976). There also exists a solution that describes T-symmetric 'holes' embedded in an isotropic heat bath of its own temperature (Zurek and Page 1984). Instead of a future horizon and a future singularity it describes a *spatial* singularity at $r = 0$, representing negative mass. It seems to be unstable against quantum fluctuations of the metric in the region of the Schwarzschild radius, signalling the need for quantum gravity.

A small black hole may completely decay into thermal radiation. The resulting entropy has been estimated to be slightly larger than the black hole entropy if the radiation is emitted into empty space (Zurek 1982b). Since the future horizon and the singularity would thereby also disappear, one seems now confronted with a *genuine global indeterminism* — regardless of any choice of foliations. This problem is known as the *information loss paradox.*

Several possible resolutions of this paradox have been discussed in the literature. Conventional quantum theory would require that the entropy of the radiation is merely a result of its statistical description, which regards photon correlations as irrelevant. The photons in this black body radiation may be entangled to form a pure total state (as they would according to a Schrödinger equation that was formulated on a foliation with slices that never hit the singularity). A similar pure global state of radiation would arise according to the unitary description if a *macroscopic* body decayed from a high energy *pure* state into its ground state by emitting many photons (Page, 1980). This state of the radiation field can for all practical purposes (local or independent photon measurements) be assumed to be a thermal mixture. However, quantum correlations between photons of the incoming (advanced) radiation of a white hole would be dynamically relevant for it to 'grow hair'. A general correlation between the time arrows of horizons and radiation has been derived in the form of a 'consistency condition' for certain de Sitter type universes by Gott and Li (1997). Their model is remarkable in possessing *different* arrows of time in different spacetime regions (separated by an event horizon).

In spite of the absence of any explicit mechanism, this proposal of a unitary description of thermal radiation arising from black hole evaporation remains conventional (except for the possible violation of baryon and lepton number conservation). While not even the *Stoßzahlansatz* is in conflict with microscopic determinism, non-zero physical entropy has been shown to be compatible with a pure state in Sect. 4.2. However, the complete information representing the pure state can only be contained in the whole superposition of all Everett branches. It would be *absolutely* lost in a genuine collapse.

Therefore, the description of black holes by a probabilistic super-scattering matrix $ (Hawking 1976) could have similar roots as the apparent collapse of the wave function or the *Stoßzahlansatz*. An S- (or $-) matrix would in any case not represent a meaningful concept for describing black holes, which — like all macroscopic objects — never become 'asymptotically free' because of their never ending decoherence by their environment (Demers and Kiefer 1996). *Microscopic* 'holes', which could be described as quantum objects by means of an S-matrix, cannot possess the *classical* properties 'black' or 'white', that are analogous to the chirality of sugar molecules — cf. Sect. 4.3.2). They can only be expected to exist as T eigenstates in their symmetric or anti-symmetric superpositions of black and white.

According to other proposals, a *black hole remnant* that conserves all relevant properties (including lepton and baryon number and the complete entanglement with the environment) must always be left behind when a black hole is transformed into radiation. A third possibility will be discussed in Sect. 6.2.3.

General literature: Bekenstein 1980, Unruh and Wald 1982

5.2 Thermodynamics of Acceleration

While the time arrow of black holes is defined by their classical spacetime structure, Hawking radiation is a consequence of *quantum* fields on their spacetime. The existence of this radiation requires the presence of an event horizon, which in turn depends on the world line of an 'observer of reference' (or a family of observers, such as all inertial ones in a flat asymptotic spacetime). For example, the future horizon at $r = 2M$ would not exist for an inertial detector freely falling into the black hole, in contrast to one being held at a certain distance $r > 2M$. The latter would feel a gravitational force unless it were in appropriate rotational motion.

Homogeneous gravitational forces are 'equivalent' to uniformly accelerated frames of reference.[2] Must accelerated detectors then be expected to register heat radiation in an inertial vacuum — similar to the classical

[2] Homogeneous gravitational fields do *not* imply spacetime curvature. The popular error that they do has led to the folklore that accelerated systems have to be described

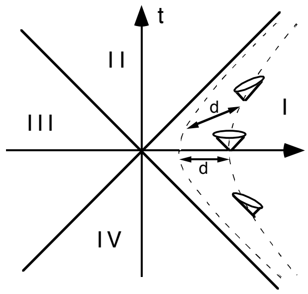


Fig. 5.5. Horizons appearing in Minkowski spacetime to uniformly accelerated observers (characterized by hyperbolic orbits $\varrho := (x^2 - t^2)/4 = \text{constant}$), with proper accelerations given by $a = (2\sqrt{\varrho})^{-1}$. The distance d between two parallel observers (who share co-moving rest frames and asymptotes) remains constant

electromagnetic radiation that they would measure in the Coulomb field of an inertial charge (Sect. 2.3) and equivalent to (5.13)? Horizons may indeed exist in flat spacetime for non-inertial observers. For example, there would be a past and a future horizon for a uniformly accelerated observer in flat Minkowski spacetime, given by the asymptotes of his hyperbolic relativistic orbit (see Fig. 5.5). This observer shares his horizons with a whole family of 'parallelly accelerated' ones (who require different accelerations in order to remain on parallel hyperbolae — similar to two observers at different fixed distances from a black hole). These observers also share comoving rest frames, and thus define accelerated rigid frames with fixed distances d in spite (or rather because) of their different acceleration.

In the x, t-plane this situation appears analogous to the Schwarzschild-Kruskal spacetime (Fig. 5.2), although it is singularity-free, since each point in the diagram now represents a flat R^2 rather than a 2-sphere. If the acceleration *began* at a certain finite time, no past horizon would exist (in analogy to a T-asymmetric black hole). The orbits of this family of observers can be used to define a new spatial coordinate $\varrho(x,t)$ that is constant along each orbit and may be conveniently scaled by $\varrho(x,0) = x^2/4$. Together with a new time coordinate $\phi(x,t)$ that is related to proper times τ along the orbits by $d\tau = \sqrt{\varrho}\, d\phi$, and the flat coordinates y and z, it defines the *Rindler coordinates* in flat spacetime. In region I of Fig. 5.5 they are related to the Minkowski coordinates by

$$x = 2\sqrt{\varrho}\cosh\frac{\phi}{2} \quad \text{and} \quad t = 2\sqrt{\varrho}\sinh\frac{\phi}{2} \quad . \tag{5.15}$$

by *general* relativity. However, all that is required is curved spacetime *coordinates* and the identification of inertial frames with freely falling (apparently accelerated) ones.

The proper accelerations a along $\varrho =$ constant are given by $a(\varrho) = (2\sqrt{\varrho})^{-1}$, while the resulting non-Minkowskian representation of the Lorentz metric,

$$ds^2 = -\varrho\, d\phi^2 + \varrho^{-1}\, d\varrho^2 + dy^2 + dz^2 \quad , \tag{5.16}$$

describes a coordinate singularity at $\varrho = 0$ that is analogous to $r - 2M = 0$ for the Schwarzschild solution. The Minkowski coordinates can therefore be compared with the Kruskal coordinates u and v of Fig. 5.2, while the Rindler coordinates are analogous to the Schwarzschild coordinates.

The Rindler coordinates are also useful for describing the uniformly accelerated point charge of Sect. 2.3 and its relation to a co-accelerated detector, (even though this situation does not possess the full cylinder symmetry of the coordinates). The radiation propagating along the forward light cone of an event on the accelerated world line of the charge must somewhere hit the latter's future horizon, and asymptotically disappear completely into region II from the point of view of an inertial observer. However, it would require *infinite* time for the radiation to reach the horizon if described in terms of co-accelerated (Lorentz-rotated) simultaneities $\phi = \mathit{constant}$, which must all intersect the horizon at the origin. These simultaneities also characterize co-accelerated detectors (those at a fixed distance $\Delta\varrho$ from the charge).

This may explain why, from the point of view of an inertial observer, but not for a co-accelerated one, the accelerated charge radiates (Boulware 1980). While Dirac's radiation reaction (2.24) vanishes for uniform acceleration, the distinction between near-fields and far-fields by their powers of distance according to (2.13), and therefore the definition of radiation, depends on the acceleration of the reference frame. Time reversal symmetry is expressed by the fact that in region I the total retarded field of the charge is identical with its advanced field (except on the horizons), while one has either only retarded outgoing fields in region II, or only advanced incoming fields in region IV (or any superposition of these cases). Therefore, even though global inertial frames are absolutely defined in special relativity, only *relative* acceleration between source and detector is relevant if uniform.

Unruh (1976) was indeed able to show — similar to Mould (1964) for classical radiation — that an accelerated 'particle' detector in the inertial vacuum of a quantum field must register isotropic thermal radiation with an *Unruh temperature*

$$T_{\mathrm{U}} := \frac{a}{2\pi} = \frac{a\hbar}{2\pi c k_{\mathrm{B}}} \quad . \tag{5.17}$$

This is precisely what had to be expected in analogy to (5.13) according to the principle of equivalence. However, the response of a detector cannot be a matter of perspective or definition (as was the distinction between radiation and near-field).

Unruh's result can also be derived by representing the inertial *Minkowski vacuum* $|0_{\mathrm{M}}\rangle$ in terms of 'Rindler particle states', which are defined as harmonic wave modes factorizing in the Rindler coordinates (with frequency

Ω with respect to the time coordinate ϕ).[3] If plane wave modes $e^{i(kx-\omega t)}$ (Minkowski modes) are expanded in terms of these Rindler modes, this expansion induces a *Bogoljubow transformation*, $a_k^+ \rightarrow b_{\Omega s}^+ := \sum_k (\alpha_{\Omega s,k} a_k^+ + \beta_{\Omega s,k} a_k)$, for the corresponding 'particle' creation operators. In this relation, the index $s =$ I or III specifies two Rindler modes (both with time dependence $e^{-i\Omega\phi}$) which vanish in the regions III or I of Fig. 5.5, respectively. On flat simultaneities through the origin they are thus complete on half-spaces with $x > 0$ or $x < 0$. These Bogoljubow transformations linearly combine creation and annihilation operators, since the non-linear coordinate transformations do not preserve the sign of the frequency (ω or Ω). These signs distinguish particle and anti-particle modes in the usual interpretation, and the two terms appearing in the Fourier representation of the field operator, $\Phi(\boldsymbol{r}, t) \propto \int \{\exp[i(kx + \omega t)] a_k + \exp[i(kx - \omega t)] a_k^+\} \, dk$, do not transform separately in this general case.

In terms of the Rindler particle representation, the Minkowski vacuum assumes the entangled form of a BCS ground state of superconductivity,

$$|0_\mathrm{M}\rangle = \prod_\Omega \left[\sqrt{1 - e^{-4\pi\Omega}} \sum_n e^{-2\pi\Omega n} \, |n\rangle_{\Omega,\mathrm{I}} \, |n\rangle_{\Omega,\mathrm{III}} \right] \tag{5.18}$$

(Bardeen, Cooper and Schrieffer 1957), where $|n\rangle_{\Omega,s} = (n!)^{-1/2} (b_{\Omega s}^+)^n |0_\mathrm{R}\rangle$ are the Rindler particle occupation number eigenstates. The *Rindler vacuum* $|0_\mathrm{R}\rangle$, defined by $b_{\Omega s}|0_\mathrm{R}\rangle = 0$ for all Ω and s, is therefore different from the Minkowski vacuum. In terms of Minkowski particles it still forms a *pure* state — not a thermal mixture. (The Hawking radiation remaining after the disappearance of a black hole may be expected to form a similar coherent superposition.) The global concepts of quantum particles and their vacua are thus not invariant under non-Lorentzian transformations. In contrast, the actual *quantum state* is invariant and physically defined ('real'), while its interpretation in terms of particle numbers depends on a 'reference basis'in Hilbert space that represents the kinematical status of the corresponding observational instruments. In particular, the Rindler basis (or any 'observable' defined in terms of it) characterizes measurements by means of correspondingly accelerated devices, while a specific 'vacuum' represents an actual state (defined or prepared by means of physical, usually cosmological, boundary conditions). This difference would be obscured in the Heisenberg picture.

Equation (5.18) is the *Schmidt canonical representation* (4.25) of nonlocal quantum correlations between the two sectors I and III (which together are spatially complete for hyperplanes intersecting the origin $x = t = 0$). It illustrates the kinematical nonlocality of a relativistic Minkowski vacuum (see

[3] According to the discussion in Sects. 4.3 and 4.6, the term 'particles' here means field modes with their occupation numbers (oscillator quantum numbers). The concept of *classical* particles should be avoided on a fundamental level of quantum theory. For example, Feynman diagrams merely form an intuitive scheme for calculating the propagation of *quantum states* (wave functionals) $\Psi(t)$.

also Gerlach 1988). The diagonal elements of the reduced density matrix that describes the corresponding 'mixed state' on sectors I or III in the sense of $\hat{P}_{\text{sep}}$ are proportional to $\exp(-4\pi n\Omega)$ for each frequency Ω. They represent a canonical distribution with dimensionless temperature $1/4\pi$ (compatible with the dimensionless time coordinate ϕ). Since proper times along the world lines $\varrho, y, z =$ constant are given by $d\tau = \sqrt{\varrho}\, d\phi = (2a)^{-1}\, d\phi$, the physical energies are $2an\Omega$. The (ϱ-dependent) temperature is therefore $T = a/2\pi$ — in accordance with (5.17). Disregarding quantum correlations with the other half-space thus leads to the *apparent ensemble* of states representing a heat bath. As the detector requires measurement times Δt of the order of $(a\Omega)^{-1}$ in order to resolve a frequency Ω, the acceleration has to remain approximately uniform during this interval of time in order to mimic the presence of an event horizon for this mode.

Although the result (5.17) might have been expected according to the principle of equivalence, the latter is in general only locally applicable. Non-inertial rigid systems are an exception valid for uniform acceleration in flat spacetime (cf. Sect. 2.3). Therefore, Unruh radiation cannot be *globally* compared with Hawking radiation. While the whole future light cone of an event on the world line of a uniformly accelerated object will asymptotically intersect the latter's horizon, only *part* of the future light cone of an event in the external region of a black hole will ever enter its internal region. Similarly, only part of the (past) celestial sphere of an observer may have come close to a black hole horizon, where Hawking radiation originates. (The horizon tends to cover the *whole* celestial sphere of an observer approaching a black hole as a consequence of spacetime curvature. He would have to speed towards the remaining 'hole in the sky' in order not to be swallowed.) Such geometric aspects also determine the generalized Sommerfeld radiation condition that characterizes a specific 'vacuum' (Unruh 1976). Only in the immediate neighbourhood of the horizon can the freely falling observer be equivalent to the inertial one in flat spacetime, and therefore experience a vacuum. While the Unruh radiation is isotropic and T-symmetric, the Hawking radiation specifies a direction in space as well as in time by its non-vanishing energy flux.

A real and observable Rindler vacuum of QED could be produced by a uniformly accelerated ideal mirror (Davies and Fulling 1977). A mirror, representing a plane boundary condition to the field, leads to the removal of an infinite number of field modes (those not matching the boundary condition) with their infinite zero-point energy. For an inertial mirror this leads to an infinite renormalization of energy, defining a 'dressed mirror'. The dressing would not be additive — though still meaningful — for *several* parallel mirrors set at a fixed distance in their rest frame, while the adiabatic variation of their distances would give rise to the finite and observable *Casimir effect* (a force between two closely-spaced mirrors). An *accelerated* mirror, acting as an accelerated boundary, produces a quantum state that would be experienced as a vacuum by a co-accelerated detector, but as a thermal bath by an inertial one. A uniformly accelerated mirror would completely determine this QED state on the concave side of its spacetime hyperbola in Fig. 5.5, while the convex side offers the freedom of additional boundary conditions in regions II or IV (similar to the classical field of a uniformly accelerated charge). According to the equivalence principle, the mirror would even have to 'drag' inertial frames if it were an ideal 'graviton mirror'. (A *virtual* accelerated graviton mirror can be mimicked by a massive plane, since the mass causes relative uniform acceleration between the innertial frames of the two half-spaces.)

The thermodynamical effects of acceleration or curvature are too small to be confirmed with presently available techniques. (For this reason, Bekenstein and Hawking could not yet receive the Nobel prize.) However, they are the consequence of combining two well established theories, and they are all required for the consistency of Bekenstein's generalized thermodynamics. This has been demonstrated by means of beautiful *gedanken* experiments with cyclic processes in the vicinity of black holes (Unruh and Wald 1982).

Although the boosted detector in flat spacetime is widely 'equivalent' to one at a fixed distance from a black hole, its excitation energy is provided by a quite different source. In the case of acceleration, the energy must be drawn from the booster, while for a black hole it would have to reduce the latter's mass, and ultimately lead to its disappearance. However, since the presence of Hawking radiation depends on the inertial state of the observer, the black hole's mass loss must then also be observer-dependent (Hawking's observer-dependent *back-reaction of the metric*). To quote Hawking (1976):

"*If spacetime is quantized, one has to abandon the idea of a metric which is independent of the observer. ... The reason is that to determine where one is in space-time one has to 'measure' the metric, and this act of measurement places one in one of the different branches of the wave function in the Everett-Wheeler interpretation.*"

This requires taking into account quantum gravity (Sect. 6.2), while similar arguments would apply to the fuel consumption of the booster according to quantum mechanics on classical (such as flat) spacetime.

General literature: Birrell and Davies 1983

5.3 Expansion of the Universe

Modern cosmology is based on the global spacetime structure, which is described by the same theory (Einstein's general relativity) that was used to describe black holes in Sect. 5.1. Since Hubble's discovery of 1923 we know that the universe is expanding. This is often regarded as a confirmation of general relativity (by an effect that Einstein missed predicting by introducing his cosmological constant, while it was present in Friedmann's solutions of 1922). However, a dynamical universe could as well have been discussed in terms of Newton's theory. In the spherically symmetric case, one would even have obtained the same expansion law (although this model would then have specified an inertial center). Apparently, applying the laws of mechanics and gravity to the whole universe met almost as much reservation (until the beginning of this century) as had their application to the celestial objects a few hundred years earlier. (Kepler and his contemporaries were surprised that planets 'fly like the birds'.) In particular, a *static* universe would require a new and fundamental repulsive force in order to compensate gravity — precisely what Einstein's cosmological constant was supposed to do before the discovery of the Hubble flow. In an open universe these consequences could at most be *obscured*, but not avoided. In Newton's theory, too, an expanding universe would require a big bang that provided the ordered initial kinetic energy.

In Einstein's theory, a homogeneous and isotropic universe is described by the Friedmann-Robertson-Walker (FRW) metric,

$$ds^2 = -dt^2 + a(t)^2 \left\{ d\chi^2 + \Sigma^2(\chi) \left[d\theta^2 + \sin^2\theta \, d\phi^2 \right] \right\} \quad , \tag{5.19}$$

with $\Sigma(\chi) = \sin\chi, \sinh\chi$, or χ, depending on the sign of the spatial curvature, $k = +1, -1$ or 0, respectively. The Friedmann time coordinate t in (5.19) represents the proper time of objects which are at rest in these coordinates ('comoving clocks'). This metric may remain valid at the big bang (for $a = 0$) in accordance with the Weyl tensor hypothesis. Its Weyl tensor does indeed *always* vanish, although this model can be generalized by means of a multipole expansion on the Friedmann sphere (see Halliwell and Hawking 1985, and Sect. 5.4).

The FRW metric depends only on the expansion parameter $a(t)$. Its dynamics, derived from the Einstein equations (5.6) with an added cosmological constant, then assumes the form of an 'energy integral' (with vanishing value of the energy) for $a(t)$, or for some function of it. For example, one obtains for its logarithm, $\alpha = \ln a$,

$$\frac{1}{2}\left(\frac{1}{a}\frac{da}{dt}\right)^2 = \frac{1}{2}\left(\frac{d\alpha}{dt}\right)^2 = -V(\alpha) \quad . \tag{5.20}$$

This logarithmic measure of spatial extension, α, which sends the big bang to minus infinity, will prove convenient in Sect. 6.2. The Friedmann potential $V(\alpha)$ is given by the energy density of matter, $\varrho(a)$, the cosmological constant Λ, and the spatial curvature k/a^2, as

$$V(\alpha) = -\frac{4\pi\varrho(e^{\alpha})}{3} - \frac{\Lambda}{3} + ke^{-2\alpha} \quad . \tag{5.21}$$

Essentially the same equation (without curvature term and cosmological constant, but with variable energy) is obtained from Newton's dynamics for the radius of a gravitating homogeneous sphere of matter (Bondi 1961).

The energy density ϱ may depend on a in various ways. In the matter-dominated epoch it is proportional to a^{-3}, while during the radiation era — at less than 10^{-3} times the present age of the universe — it changed with with a^{-4}, since all wave lengths expand with a. Much earlier (for extremely high matter density or curvature), novel and yet speculative theories may have been important. Several proposed quantum field theories imply that the vacuum state of matter fields went through phase transitions in this era (see Sect. 6.1). They may be characterized by an energy density that is independent of a (as during a condensation process). In this era, the matter term in the potential V would *simulate* a cosmological constant. Other contributions, due to a *dilaton field*, have been postulated in string theory as being relevant in the still earlier *Planck era* (of the order $a \lesssim 1$).

The Friedmann potential $V(\alpha)$ is indicated in Fig. 5.6. Just as in classical mechanics, allowed regions require positive kinetic energy $E - V = -V > 0$, while $V = 0$ defines turning points of the cosmic expansion. An upper turning point describes a recollapsing universe, while a lower turning point represents a 'bouncing' universe (without big bang or big crunch singularities).

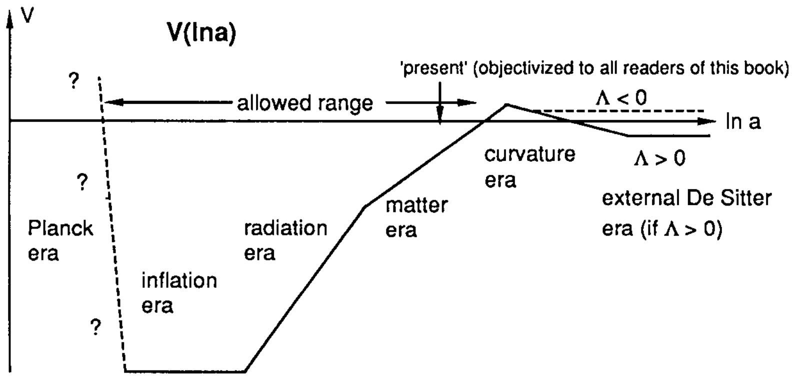


Fig. 5.6. Schematic behavior of the 'potential energy' for the dynamics of $\ln a$ (assuming positive spatial curvature). Turning points of the motion are defined by $V = 0$. Recent observations of distant supernovae indicate that the universe has already passed the maximum of the potential V (which would then have to lie below the axis). Therefore, V will always remain negative, while the dynamical effect of curvature remains negligible

Different eras, described by different equations of state $\varrho(a)$, may possess different analytical solutions $a(t)$. For a dominating (fundamental or simulated) cosmological constant, $a(t) = c\mathrm{e}^{\pm Ht}$, with a 'Hubble' constant $H = \dot{a}/a = \dot{\alpha}$. This situation is called a *de-Sitter era.* For a matter or radiation dominated universe, $a(t) = c' t^{2/3}$ or $a(t) = c'' t^{1/2}$, respectively, while for very large values of a (low matter densities), and in the absence of a fundamental cosmological constant, the curvature term would dominate. For positive curvature it would then require reversal of the expansion at some time. (See also the model (5.34) and Fig. 5.7 in the subsequent section!)

The finite age of an expanding universe with an initial singularity (a big bang) leads to the consequence that the backward light cones of two events might nowhere overlap. These events would then not possess any (partial) common cause. A sphere formed by the light front originating in a point-like event at the big bang (where $a(0) = 0$) is called a *causality horizon*. Its radius $s(t)$ at Friedmann time t is given by

$$s(t) = \int_0^t \frac{a(t)}{a(t')}\, dt' \quad . \tag{5.24}$$

For matter- or radiation-dominated universes this integral converges even though $a(t) \to 0$ for $t \to 0$, and for vanishing cosmological constant, s remains smaller than the semi-circumference of a closed Friedmann universe. Only (growing) *regions* of the universe may then be causally connected. Quantum entanglement may also be expected to require local *causes*, and therefore to be limited to distances within the causality horizon. This would explain the initial condition (4.56), while it leaves the homogeneity of the universe (that is, its simultaneous beginning) unexplained.

This latter *horizon problem* was the main motivation for postulating a phase transition of the vacuum that leads to an early de-Sitter phase (an exponentially expanding universe). The big bang singularity can then in principle be shifted arbitrarily far into the past, depending on the duration of this phase transition. However, in an extremely short time span (of the order of 10^{-33} sec) the universe, and with it all causality horizons, would *inflate* by a huge factor that was sufficient for the whole now observable 2.7 K background radiation to be causally connected (Linde 1979).

While the Friedmann model is an exact solution of the Einstein equations, and apparently an approximation to the large scale behavior of the real universe, it is not thermodynamically stable against density fluctuations (as discussed at the beginning of this chapter and in Sect. 5.1). This instability is expected to explain the formation of galaxies, galaxy clusters, and possibly larger structures in the present universe. Breaking translational symmetry may either require small inhomogeneities as a seed, or result from 'quantum fluctuations' that become 'real' by means of decoherence (see Calzetta and Hu 1995, Kiefer, Polarski and Starobinsky 1998, and Sect. 6.1). The second mechanism is also known to lead to a quantum limit for the retardation (hysteresis) of phase transitions. The onset of these structures is now believed being observed in the cosmic background radiation.

The arrow of time characterizing these irreversible processes is thus based again on an improbable cosmic initial condition: homogeneity. Boltzmann (1896) already discussed the Second Law in a cosmological context. Under the assumption of an infinite universe in space and time, he concluded that we, here and now, happen to live in the aftermath of a gigantic cosmic fluctuation.

Its maximum (that is, the state of lowest entropy) must have occurred in the distant past in order to explain the existence of fossils and other documents in terms of causal history and evolution. This *improbable* assumption (cf. the remarks at the end of Chap. 3) is an application of the *weak anthropic principle* (Carr and Rees 1979, Barrow and Tipler 1986): we could not exist at another place and time according to this mechanistic world model.

How improbable is the novel initial condition of homogeneity that Boltzmann did not even recognize as an essential assumption? We may calculate just how improbable from the entropy by means of Einstein's relation (3.52). The conventional physical entropy of a homogeneous state in local equilibrium is proportional to the number of particles, N. Gravitational entropy is negligible in the present state of our universe. However, the number of photons in the 2.7 K background radiation exceeds that of massive particles by a factor 10^8. (The mean entropy of a baryon is approximately the same as that of a photon under present conditions — both are of order k_B.) A standard model of a closed universe, with 10^{80} baryons (now often regarded merely as a causally connected 'bubble' in an infinite world), would therefore possess an entropy of order 10^{88} (the number of photons).

In Planck units, the area of a horizon of a neutral spherical black hole of mass M is given by $A = 4\pi(2M)^2$. Its entropy thus grows with the second power of its mass,

$$S_{\text{bh}} = 4\pi M^2 \quad . \tag{5.25}$$

Merging black holes will therefore produce a considerable amount of entropy. If the standard universe of 10^{80} baryons consisted of 10^{23} solar mass black holes (since $M_{\text{sun}} \approx 10^{57} m_{\text{baryon}}$), it would possess a total entropy of order 10^{100}, that is, 10^{12} times its present value. If most of the matter eventually formed a single black hole, this value would increase by another factor 10^{23}. The probability for the present, almost homogeneous universe of $2.7K$ is therefore a mere

$$p_{\text{hom}} \approx \frac{\exp(10^{88})}{\exp(10^{123})} = \exp(10^{88} - 10^{123}) \approx \exp(-10^{123}) \quad , \tag{5.26}$$

when compared with its most probable state in this logarithmic approximation (Penrose 1981). Gravitational contraction offers an enormous entropy capacity for structure and complexity to evolve out of homogeneity in accordance with the Second Law (even though the Second Law may not represent the complete story — cf. Figs. 3.5 and 4.3).

This improbable initial condition of homogeneity as the origin of the Second Law is different from Gold's (1962) proposal to derive the thermodynamical time arrow from expansion *in a causal way* (cf. Price 1996 and Schulman 1997 for critical discussions). While it is true that any non-adiabatic expansion of an inhomogeneous equilibrium system (such as droplets in a saturated gas) leads to a retarded non-equilibrium in our causal world, this would as well be true for non-adiabatic *contraction.* The growing space (and thus phase space, representing increasing entropy capacity) has often been proposed as the master arrow of time, although it is clearly insufficient to explain causality (such as non-equilibrium being *caused* by expansion or contraction). Non-adiabatic *compression* of a vessel would lead to retarded pressure waves emitted from the walls, but not to a reversal of the thermodynamical arrow. The quantitative arguments demonstrate that the phase space created by gravitational contraction (ultimately into black holes) is far more important than that due to the cosmic expansion.

There are other examples of *using* causality in thermodynamical arguments rather than *deriving* it in this cosmic scenario. Gal-Or (1974) discussed retarded equilibration caused by the slow nuclear reactions in stars. Even though nuclear fusion controls the time scale and energy production during early stages of stellar contraction, its entropy production is marginal compared to black hole formation. Another example of causal reasoning is provided by a positive cosmological constant that may counteract gravity and induce *accelerated* expansion. Can it then also reverse the gravitational trend towards inhomogeneity, or even causally explain the observed homogeneity?

This has indeed been suggested on the basis of the *cosmological no-hair conjecture* (Hawking and Moss 1982), which states that the universe must lose all structure and become homogeneous during inflation. However, this conjecture is questionable, since the long range forces described by a cosmological constant cannot force local gravitating systems, in particular black holes, to expand fast and enter a state of homogeneity. (Expanding white holes would require acausally incoming advanced radiation, as explained in Sect. 5.1.) The expansion of diluted parts of the universe would thus accelerate faster under the influence of a cosmological constant than that of its denser parts (unless this difference were compensated by a conspiratorial inhomogeneous initial expansion velocity), thus causing inhomogeneities to grow. 'Proofs' of this no-hair conjecture had therefore to exclude positive spatial curvature.

However, a cosmological constant simulated by a phase transition of the vacuum during the early stages of the universe could overcompensate the effect of gravity until strong inhomogeneities begin to form. Therefore, Davies (1983, 1984) suggested that the homogeneity of our universe was caused by its inflation (see also Page 1984, Goldwirth and Piran 1991). This would mean that the Weyl tensor 'cooled down' as a consequence of this driven spatial expansion — similar to the later occurring red-shifting of the electromagnetic radiation. While these *direct* implications of the expansion define reversible processes, equilibration (in the radiation era or during the phase transition) would be irreversible in the statistico-thermodynamical sense (based on microscopic causality).

This explanation of homogeneity would still have to exclude strong initial inhomogeneities (initial black holes, in particular). In order to represent a causal explanation, it would require the state that precedes inflation to be even less probable than the homogenous state after inflation. What would then happen backwards in time to those many *inhomogeneous* states which are 'now' conceivable?

These questions may be appropriately discussed in terms of a recollapsing universe. Its consistent analysis is meaningful regardless of what will ever happen to our own universe. Would the thermodynamical arrow have to reverse direction when the universe starts recollapsing towards the big crunch after having reached maximum extension? The answer would have to be 'yes' if the cosmic expansion represents the master arrow, but it is often 'no' on the basis of causal arguments continued into this region (as it proved sufficient to counter Gold's original argument). Several authors argued that the background radiation would first have to heat up through wave lenghth contraction (blue shifting), while the temperature gradient between interstellar space and the fixed stars would then be inverted only long after the universe has reached its maximum extension. However, this argument presupposes the overall validity of the 'retarded causality' in question, that is, the absence of conspiratorial correlations in the contraction phase. It would be justified if the relevant improbable initial condition, such as Penrose's Weyl tensor hypothesis, held at only one 'end' of this cosmic history. The absence or negligibility of irreproducible conspiratorial events in our present epoch seeems to indicate either that the universe is globally asymmetric in time, or that it is still 'improbably young' when related to its total duration.

Davies (1984) argued similarly that there can be no reversed inflation at the big crunch, since correlations required for an inverse phase transitions can be excluded as improbable. Instead of a homogeneous big crunch one would either obtain 'de-Sitter bubbles', which would reverse the cosmic contraction *locally* and lead to an inhomogeneous 'bounce', or inhomogenous singularities at variance with a reversed Weyl tensor condition, or both. This argument fails, however, if the required correlations are *caused in the backward direction of time* by a final condition that is thermodynamically symmetric in time to the initial one (cf. also Sect. 6.2.3). Similarly, if the big bang were replaced with a non-singular *homogeneous bounce* in some way, entropy must have decreased previous to the bounce. In all cases, an observer complying with the Second Law would *experience* an expanding universe.

On the other hand, a low entropy big crunch together with a low entropy big bang would constrain the possible histories of the universe far more than the low entropy condition at *one* end. The general boundary value problem (cf. Sect. 2.1) leaves only one complete (initial or final) condition free. Although a condition of low entropy does not define the initial state completely, statistically independent two-time boundary conditions would square the probability (5.26), that is,

$$p_{\mathrm{two-time}} = p_{\mathrm{hom}}^2 \approx [\exp(-10^{123})]^2 \approx \exp(-10^{123.301}) \quad . \tag{5.27}$$

The RHS may appear as a small correction to (5.26) in this exponential form, but an element of phase space as small as described by (5.27) would be much smaller than a Planck cell (that represents a pure quantum state). A two-time boundary condition of homogeneity may thus be inconsistent with 'ergodic' quantum cosmology.

The consistency of two-time boundary conditions has been investigated for simple deterministic (dynamically closed) systems (see Cocke 1967 for mechanical, and Schulman 1997 for wave mechanical systems). For example, an open Friedmann universe with vanishing cosmological constant would not even allow two boundaries with $a = 0$, but rather expand forever (in one time direction) in a drastically asymmetric way. Davies and Twamley (1993) discussed the situation of classical electromagnetic radiation in an expanding and recollapsing universe. According to their estimates, our universe will remain essentially transparent all the way between the two opposite radiation eras (in spite of the reversible frequency shift over many orders of magnitude in between). Following an argument by Gell-Mann and Hartle they then concluded that the light emitted causally by all stars before the 'turning of the tide' must propagate according to the Maxwell equations, and remain dominant until the reversed radiation era is reached. Therefore, our universe should possess a macroscopically asymmetric history even if it were bound to recontract.

Craig (1996) argued on this basis, but by *assuming* cosmic history to be essentially symmetric, that the brightness of our night sky at optical frequencies would have to contain a homogeneous component (that is, a non-Planckian high frequency tail in the isotropic background radiation) with intensity at least equalling the intensity of the light now observed from all stars and galaxies. This conclusion assumes tacitly that the retarded radiation emitted by all galaxies before the turning of the tide is *different* from the advanced radiation from (thermodynamically reversed) galaxies during the contraction phase. If this advanced radiation were identical with the retarded one (cf. Sect. 2.1), this would require the light from the expansion era to 'conspire' in order to focus on sources during the formal contraction era. Even though the star light would considerably spread in space during the life cycle of a long-lasting recollapsing universe, it would classically preserve all information about its origin, and thus not be compatible with local reversed sources in our distant future. Therefore, Craig assumed the star light from one side to continue into the radiation era on the temporally opposite side (so that the total intensity is doubled).

However, these conclusions are invalidated when the photon aspect of electromagnetic radiation is taken into account. The restricted information content of photons, emphasized already by Brillouin (1962), was essential also for establishing limits to Borel's argument of Sect. 3.1.2 (cf. Sect. 4.3.3). Each photon, even if emitted into intergalactic space as a spherical wave, disappears from the whole quasi-classical universe as soon as it is absorbed *somewhere*. There is an asymmetry between emission and absorption in the conventional quantum description (cf. Fig. 4.4). If the absorber is itself described quantum mechanically, the localization of the absorbed photon is a consequence of decoherence — reflecting the fact that wave functions live in configuration space rather than in space. In quantum cosmology (without taking into account quantum gravity yet), a reversal of this process would require many Everett components to conspire in order to recohere, while the corresponding final condition *for the total Everett wave function* would not be in any obvious conflict with the initial condition. If the Schrödinger dynamics were instead modified by a collapse of the wave function, this new dynamical law would also have to be reversed in time for Craig's considerations to remain valid.

This problem of consistent cosmic two-time boundary conditions will assume a conceptually novel form in the context of quantum gravity, since any concept of time then disappears from cosmology (Sect. 6.2).

5.4 Geometrodynamics and Intrinsic Time

In general relativity, the 'block universe picture' is traditionally preferred to a dynamical description, as its unified spacetime concept is then manifest. (A historically remarkable exception, aside from cosmology, is the Oppenheimer-Snyder model of gravitational collapse.) Only during the second half of the last century has the dynamical (Hamiltonian) approach to general relativity been appropriately explored, particularly by Arnowitt, Deser and Misner (1962). This has not always been welcomed, as it seems to destroy the beauty of relativistic spacetime symmetry by re-introducing a 3+1 (space and time) representation. However, only in this *form* can the dynamical content of general relativity be fully appreciated. A similarly asymmetric form in spite of relativistic invariance is known from the dynamical description of the electromagnetic field in terms of the vector potential **A** (cf. Chap. 21 of Misner, Thorne and Wheeler 1973).

This dynamical reformulation of the theory requires the separation of unphysical gauge degrees of freedom (which in general relativity simply represent the choice of coordinates), and the skilful handling of boundary terms. The result of this technically demanding procedure turns out to have a simple interpretation. It describes the *dynamics of spatial three-geometry*, ${}^{(3)}G(t)$, that is, a propagation of the intrinsic curvature of space-like hypersurfaces with respect to a time coordinate t that labels a foliation of the dynamically arising spacetime. This foliation has to be *chosen* during the construction of the solution. The extrinsic curvature, describing the embedding of the three-geometries into spacetime, is represented by the corresponding canonical momenta. The configuration space of the three-geometries ${}^{(3)}G$ has been dubbed *superspace* by Wheeler. Trajectories in this superspace define four-dimensional spacetime geometries ${}^{(4)}G$.

This 3+1 description may appear ugly not only as it hides Einstein's beautiful spacetime concept, but also since the required foliation by means of space-like hypersurfaces (on which ${}^{(3)}G$ is defined) is arbitrary. Hence, many trajectories ${}^{(3)}G(t)$ represent the same spacetime ${}^{(4)}G$, which is the physically meaningful object. Only in special situations (such as the FRW metric (5.19)) may there be a 'preferred choice' of coordinates, which reflects their exceptional symmetry. The time coordinate t of a given foliation is just one of the four arbitrary (physically meaningless) spacetime coordinates. As a parameter labelling trajectories it could as well be eliminated and replaced with one of the dynamical variables (cf. Chap. 1), such as the size of an expanding universe. The resulting four-geometry then defines all spacetime distances — including *all* proper times of real or imagined local objects ('clocks'). Classically, spacetime may always be assumed to be filled with a 'dust of test clocks' of negligible mass (cf. Brown and Kuchař 1995). However, they are not required to *define* proper times (except operationally); in general relativity, time as a metric property is itself a dynamical variable (see below). According to the principle of relativity, proper time then assumes the role of Newton's absolute time as a controller of motion for all material clocks.

Einstein's equations (5.6) are second order differential equations, similar to the wave equation (2.1). They may therefore be expected to determine the metric $g_{\mu\nu}(x, y, z, t)$ by means of two boundary conditions for $g_{\mu\nu}$ — at t_0 and t_1, say. (For $t_1 \to t_0$ this would correspond to $g_{\mu\nu}$ and its 'velocity' at t_0. This pair may then also determine the extrinsic curvature.) Since the time coordinate is physically meaningless, its value on the boundaries is irrelevant; two metric functions on three-space, $g^{(0)}_{\mu\nu}(x, y, z)$ and $g^{(1)}_{\mu\nu}(x, y, z)$, are sufficient. Not even their order is geometrically essential, since there is no *absolute* direction of light cones. Similarly, the t-derivative of $g_{\mu\nu}$, resulting in the limit $t_1 \to t_0$, is required only up to a numerical factor (that specifies a meaningless initial 'speed of three-geometry' in superspace).

If one similarly eliminates all *spatial* coordinates from the metric $g_{\mu\nu}$, it describes precisely the coordinate-independent three-geometry ${}^{(3)}G$. In the 'normal' situation, the coordinate-independent content of the Einstein equations should then determine the complete four-dimensional spacetime geometry in between, and in general also beyond, two spatial geometries, ${}^{(3)}G^{(0)}$ and ${}^{(3)}G^{(1)}$. Although the general existence and uniqueness of a solution to this boundary value problem remains an open mathematical problem (Bartnik and Fodor 1993, Giulini 1998), this does not seem to be physically relevant.

The procedure is made transparent by writing the metric with respect to a chosen foliation as

$$\begin{pmatrix} g_{00} & g_{0l} \\ g_{k0} & g_{kl} \end{pmatrix} = \begin{pmatrix} N^i N_i - N^2 & N_l \\ N_k & g_{kl} \end{pmatrix} \quad . \tag{5.28}$$

The submatrix $g_{kl}(x, y, z, t)$ (with $k, l = 1, 2, 3$) on a hypersurface $t =$ constant is now its spatial metric there, while the *lapse function* $N(x, y, z, t)$ and the three *shift functions* $N_i(x, y, z, t)$ define arbitrary increments of time and space coordinates, respectively, for a normal transition to a neighboring space-like hypersurface. These four *gauge functions* have to be *chosen* for convenience when solving an initial value problem.

The six functions forming the remaining symmetric matrix $g_{kl}(x, y, z, t)$ still contain three gauge functions representing the spatial coordinates. Their initial choice is specified by the initial matrix $g^{(0)}_{kl}(x, y, z)$, while the free shift functions determine their change with time. The three remaining, geometrically meaningful functions may be physically understood as representing the two (spin) components of gravitational waves and the 'many-fingered' (local) physical time of relativity, that is, a function that defines the increase of proper times along all time-like world lines connecting points on two neighboring space-like hypersurfaces. These three degrees of freedom need not always be practically separable; all three are gauge-free dynamical variables. In contrast, the lapse function $N(x, y, z, t)$, here in conjunction with the shift functions, determines merely how a specific time coordinate is related to this many-fingered physical time.

Therefore, the three-geometry ${}^{(3)}G$, representing the dynamical *state* of general relativity, is itself the 'carrier of information on physical time' (Baierlein, Sharp and Wheeler 1962): it *contains* physical time rather than *depending* on it. By means of the Einstein equations, ${}^{(3)}G$ determines a *continuum of physical clocks*, that is, all time-like distances from an 'initial' ${}^{(3)}G_0$ (provided a solution of the corresponding boundary value problem does exist). Given yesterday's geometry, today's geometry could not be tomorrow's — an absolutely non-trivial statement, since ${}^{(3)}G_0$ by itself is not a complete initial condition that would determine the solution of (5.6) up to a gauge. A mechanical clock can meaningfully go 'wrong', while for a rotating star one would have to know the initial orientation *and* the initial rotation velocity in order to read time from motion. However, a speed of three-geometry (in contrast to its *direction* in superspace) would be as tautological as a 'speed of time'. Time cannot be time-dependent in any nontrivial way.

In this sense, Mach's principle (not only with respect to time)[4] is anchored in general relativity: time must be realized by physical (dynamical) objects. Dynamical laws (as required by Mach) that do *not* specify (or depend on) an absolute time are characterized by their *reparametrization invariance*, that is, invariance under monotonic transformations, $t \to t' = f(t)$. In general relativity, the time parameter t labels trajectories in superspace by the values of a time coordinate. No specific choice may 'simplify' the laws according to Poincaré (cf. Chap. 1), and no distinction between active and passive reparametrizations remains meaningful (see Norton 1989).

Newton's equations are, of course, *not* invariant under such a reparametrization. His time t is not merely an arbitrary parameter, but represents a dynamically preferred (absolute) time. Its reparametrization would be characterized by 'Kretzschmann invariance', that is, the trivial invari-

[4] See Barbour and Pfister (1995) for various interpretations of Mach's principle.

ance of the theory under a rewriting of the dynamical laws (for example, by adding a pseudo-force). Nonetheless, Newton's equations can be brought into a reparametrization-invariant *form* by parametrizing the time variable t itself, $t(\lambda)$, and treating it as an additional dynamical variable with respect to λ. If $L(q, \dot{q})$ is the original Lagrangean, this leads to the new variational principle

$$\delta \int \tilde{L}\left(q, \frac{dq}{d\lambda}, \frac{dt}{d\lambda}\right) d\lambda := \delta \int L\left(q, \frac{dq}{d\lambda}\frac{d\lambda}{dt}\right) \frac{dt}{d\lambda} d\lambda = 0 \quad , \tag{5.29}$$

where $t(\lambda)$ has to be varied as a dynamical variable, too. This procedure describes the true meaning of the 'Δ-variation' that often appears somewhat mysteriously in analytical mechanics. Evidently, (5.29) is invariant under the reparametrization $\lambda \to \lambda' = f(\lambda)$.

Eliminating the formal variable t from (5.29) leads to *Jacobi's principle* (see below), that was partially motivated by the pragmatic requirements of astronomers who did not have better clocks than the objects they were describing dynamically. Their best available clock time was *ephemeris time*, defined by stellar positions that were taken from *tables of ephemeris* which were produced by their colleagues. Since celestial motions are usually strongly 'perturbed' by other celestial objects, they do not offer any obvious possibility to define Newton's time operationally. Jacobi's principle allowed astronomers to solve the equations of motion without explicitly using Newton's time. The dynamical system (5.29) is now often used as a toy model for reparametrization invariant theories. Einstein's equations of general relativity, on the other hand, are invariant under reparametrization of their original time coordinate, $t \to t' = f(t)$, without any further parametrization $t(\lambda)$. There is no time beyond the many-fingered dynamical variable ${}^{(3)}G$ any more!

In (5.29), $dt/d\lambda =: N(\lambda)$ may be regarded as a Newtonian lapse function. For a time-independent Lagrangean L, t appears as a cyclic variable in this formalism. Its canonical momentum, $p_t := \partial\tilde{L}/\partial N = L - \sum p_i \dot{q}_i = -H$, which is conserved in this case, is remarkable only as its quantization leads to the time-dependent Schrödinger equation. However, the 'super-Hamiltonian' $\tilde{H}$ that describes the thus extended system is trivial:

$$\tilde{H} := \sum p_i \frac{dq_i}{d\lambda} + p_t \frac{dt}{d\lambda} - \tilde{L} = N\left(\sum p_i \frac{dq_i}{dt} - H - L\right) \equiv 0 \quad . \tag{5.30}$$

More dynamical content can be extracted from Dirac's procedure of treating $N(\lambda)$ rather than $t(\lambda)$ as a new variable. The corresponding momentum, $p_N := \partial\tilde{L}/\partial(dN/d\lambda) \equiv 0$, has to be regarded as a constraint, while the *new* super-Hamiltonian is

$$\tilde{H} := \sum p_i \frac{dq_i}{d\lambda} + p_N \frac{dN}{d\lambda} - \tilde{L} = NH \quad . \tag{5.30'}$$

Although $dN/d\lambda$ can here not be eliminated in the normal way by inverting the definition of canonical momentum, it drops out everywhere except in the derivative $\partial\tilde{H}/\partial p_N$, as it occurs only multiplied with the factor p_N which vanishes. The two new resulting Hamiltonian equations are (1) the identically fulfilled $dN/d\lambda = \partial\tilde{H}/\partial p_N = dN/d\lambda$, and (2) $dp_N/d\lambda = -\partial\tilde{H}/\partial N = -H$. Because of $p_N \equiv 0$ one obtains the (secondary) *Hamiltonian constraint* $H = 0$ (but not $\equiv 0$), characteristic for reparametrization invariant theories. In general relativity, there are also three *momentum constraints*, characterizing invariance under spatial coordinate transformations which are 'dynamically' related by the shift functions.

Because of (5.30), with $p_t = -H$, Hamilton's principle (5.29) can be written as $\delta \int (\sum p_i \dot{q}_i - H)dt = 0$. For fixed energy, $H = E$, the second term, H, can be omitted under the integral in the variation. For the usual quadratic form of the kinetic energy, $2T = \sum a_{ij}\dot{q}_i\dot{q}_j = \sum p_i\dot{q}_i = 2(E - V)$, the integrand under the variation can then be written in a form that is homogeneously linear in $dq_i/d\lambda$,

$$\delta \int \sqrt{2(E-V)\sum a_{ij}\frac{dq_i}{d\lambda}\frac{dq_j}{d\lambda}}d\lambda = 0 \quad . \tag{5.31}$$

This is Jacobi's principle (cf. Lanczos 1970), useful for given energy. It is manifestly invariant under reparametrization of λ, as it describes only timeless orbits. Even though the equations of motion could be simplified by using Newton's time, (5.31) does evidently not depend on the choice of λ.

Since the energy E depends on absolute velocities in Newton's theory, Jacobi's principle would be 'Machian' only if the fixed energy represented a universal constraint. Therefore, Barbour and Bertotti (1982) suggested an illuminating nonrelativistic model of Machian mechanics. Their action principle in analogy to (5.31),

$$\delta \int \sqrt{-VT}dt = 0 \quad , \tag{5.32}$$

is universally invariant under reparametrizations of t (just as general relativity). Nothing new can then be discovered from parametrizing t in order to vary $t(\lambda)$ as in (5.29). Barbour and Bertotti also eliminated absolute rotations from their configuration space. While this has other important consequences, it is irrelevant for the problem of time. In general relativity, this 'Leibniz group', consisting of time reparametrizations and spatial rotations, has to be generalized to the whole group of *diffeomorphisms* (general coordinate transformations). In order to eliminate any absolute meaning of a time *coordinate* on spacetime, the Hamiltonian constraint has to be a function on space.

This situation has been carefully analyzed by Barbour (1994a,b). He refers to the empty physical content of a parametrization of t in the form $t(\lambda)$, in contrast to (5.29) in Newton's theory, as *timelessness*. This terminology may be misleading — not only because there are various concepts of time (see also Rovelli 1995). I will argue in Sect. 6.2 that the *rigorous timelessness of the Wheeler-DeWitt equation* (*viz.*, the absence of any time *parameter*) is a specific *quantum* aspect, as it reflects the absence of classical trajectories, or the time-energy uncertainty relation in the case of a Hamiltonian constraint. In classical general relativity, even a constrained Hamiltonian would define a time parameter for a trajectory that describes cosmic history, but not any meaningful concept of a 'speed of cosmic history', which would have to include a 'speed of physical time'.

In a general sense, *global time* may thus be defined by the succession of global physical states (a trajectory through a universal configuration space). Even the topology of this succession (its continuity and order up to a reversal) is determined by *intrinsic* properties of these states. No external parameter is required for this purpose. In general relativity, the global states are ambiguous (they depend on the chosen spacetime foliation), while the resulting spacetime geometry is invariantly defined. The latter determines many-fingered time (that is, invariant *proper times*) for all world lines of local objects (such as 'test clocks' or observers). In contrast to Newtonian action-at-a-distance gravity, proper time as a controller of motion is defined as a *local* concept. However, the dynamical laws define *relative* motion, since the metric is itself a dynamical object. A preferred global time parameter that 'simplified' these dynamical laws would be regarded as *absolute time*.

In the Friedmann model (5.19), where $N \equiv 1$, the increment of the time coordinate t is identical (up to a sign) with the increment of proper times τ of 'comoving' matter (being at rest in the chosen coordinates with $N_i = 0$). In general, the time coordinate t corresponds to the physically meaningless parameter λ of (5.29), while τ is a local version of Newton's time t as the controller of motion. Since homogeneity is presumed on all chosen simultaneities in this simple model, the many fingers of time form one closed 'fist'.

Elimination of the global time parameter t would here merely reproduce the equation of state, $\varrho(a)$ as the corresponding 'trajectory', since ϱ is not a dynamical degree of freedom. There is evidently no intrinsic distinction between expansion and contraction. The single variable a would determine proper times τ for comoving matter up to this degeneracy, since $\dot{a}^2$ is given as a function of a by the energy constraint (5.20). (Because of similar ambiguities, the boundary condition by means of two three-geometries is usually formulated as a '*thin* sandwich problem'.)

Matter can also be described dynamically by means of a homogeneous scalar field $\Phi(t)$, for example with an energy density

$$\varrho = \frac{1}{2}(\dot{\Phi}^2 + m^2\Phi^2) \quad , \tag{5.33}$$

while keeping the symmetries of the Friedmann model intact. The Hamiltonian of this model (without cosmological constant) with respect to the variables $\alpha = \ln a$ and Φ reads

$$H = \frac{e^{-3\alpha}}{2}\left(p_\alpha^2 - p_\Phi^2 + ke^{4\alpha} - m^2\Phi^2 e^{6\alpha}\right) \quad , \tag{5.34}$$

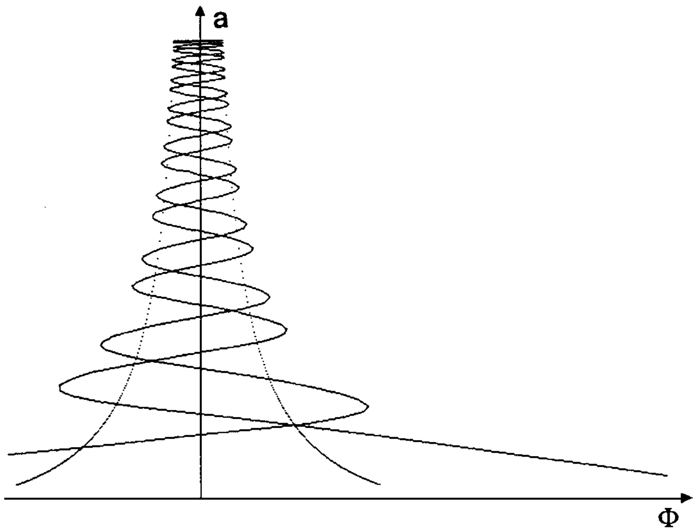


Fig. 5.7. Returning classical orbit representing the dynamics of a closed Friedmann universe, described by its expansion parameter a and a homogeneous massive scalar field Φ. The dotted curve represents vanishing Friedmann potential, $V = 0$. For slightly larger initial values $\Phi(a_0)$ than chosen in the example, the 'inflation era' (that is, the rising non-oscillating lower right part of the orbit) would extend over many orders of magnitude in a before the orbit entered the 'matter-dominated era', where it would then perform a huge number of oscillations before arriving at its turning point of the expansion, $a_{\max}$. (After Hawking and Wu 1985)

where k is again the sign of the spatial curvature, while the canonical momenta are $p_\alpha = e^{3\alpha}\dot{\alpha}/N$ and $p_\Phi = -e^{3\alpha}\dot{\Phi}/N$. A timeless trajectory for a closed universe ($k = 1$) in this model is depicted in Fig. 5.7.

The symmetry of this model can be relaxed by means of a multipole expansion on the Friedmann sphere,

$$\Phi(\chi, \theta, \phi) = \sum a_{nlm} Q^n_{lm}(\chi, \theta, \phi) \quad , \tag{5.35}$$

where $Q^n_{lm}(\chi, \theta, \phi)$ are three-dimensional spherical harmonics (Halliwell and Hawking 1985). The variable Φ in (5.33) represents the monopole component, $\Phi = a_{000}$, since $Q^0_{00} = 1$. A similar expansion of the metric tensor field g_{kl} requires vector and tensor harmonics in addition to the scalar harmonics Q^n_{lm}. Only the tensor harmonics turn out to represent physical (geometrical) properties, while the scalar and vector harmonics describe gauge degrees of freedom. In particular, the time parameter t does not exactly represent proper time of comoving matter any more in this 'perturbed Friedmann model'.

If there is a Hamiltonian constraint, $H(p, q) = 0$ in its general form, multiplication of the Hamiltonian with a function $f(p, q)$, $H \to H' = fH = 0$, would induce an orbit-dependent reparametrization $t \to t'(t)$, with $dt'/dt =$

$f(p(t), q(t))$, as can be seen by writing down the new Hamiltonian equations. For example, the choice $f \equiv -1$ would induce an inversion of the time parameter for all trajectories. Therefore, the factor $e^{-3\alpha}$ in (5.34) is irrelevant for the time-less trajectory.

In (5.34) and its generalizations, the kinetic energy of matter occurs with a negative sign (that is, with negative dynamical mass), since it entered the Hamiltonian as a source of gravity (representing negative potential energy). In Friedmann type models, all gauge-free geometric degrees of freedom but the global expansion parameter a (or its logarithm) share this property (Giulini and Kiefer 1994, Giulini 1995). The kinetic energy is thus not positive definite in cosmology, while the metric that it defines in infinite-dimensional superspace is *super-Lorentzian* (with signature $- + + + + + \ldots$) in this case.

In the familiar case of mechanics, vanishing kinetic energy, $E - V = 0$, describes turning points of the motion. However, since there are no forbidden regions for indefinite kinetic energy, the boundary $V = V - E = 0$ need here not force the trajectories to come to a halt and reverse direction. In general, this condition now describes a smooth transition between 'subluminal' and 'superluminal' directions in superspace (not in space!) — see Fig. 5.7. A trajectory would be reflected from an *infinite* potential 'barrier' only if this were either negative at a time-like boundary, or positive at a space-like one. Reversal of the cosmic expansion at $a_{\max}$ requires the vanishing of an appropriate $V_{\text{eff}}(\alpha)$ that includes the actual kinetic energy of the other degrees of freedom (similar to the effective radial potential in the Kepler problem).

In the Friedmann model, a point on the trajectory in configuration space determines Friedmann time t (that could be read from comoving test clocks) — except where the curve intersects itself. In a mini-superspace with more than two degrees of freedom (with a material clock, for example), physical time on a trajectory is generically determined *uniquely* by the dynamical state. This demonstrates that the essential requirement for the state to represent a carrier of information about time is reparametrization invariance of the dynamical laws — not its spacetime geometrical interpretation.

A drastic abuse of a time *parameter* is entertained in Veneziano's (1991) string model, based on a dilaton field Φ (with dynamics different from (5.34)). Its equations of motion lead to a time dependence of the form $f(t-t_0)$, with an integration constant t_0 that determines the value of the time parameter at the big bang (where $\alpha = -\infty$). A translation $t_0 \to t_0 + T$ would be meaningless (as pointed out already by Leibniz). The solution for $t < t_0$, where expansion *accelerates* exponentially in this model, has been interpreted as 'pre-big-bang', while the absence of a smooth connection between pre- and post-big-bang has then been regarded as a 'graceful exit problem' (Brustein and Veneziano 1994). However, this (in any case speculative) model has simply two different solutions, which could possibly be related through an infinite parameter time, $t = \pm\infty$ — similar to Schwarzschild time at a horizon. Coordinate times $t < t_0$ would then represent physical times *later* than $t > t_0$, while a continuation through t_0 is merely formal (Dabrowski and Kiefer 1997).

The shift functions N_i of (5.28) can be *chosen* to vanish even when symmetries are absent. The secondary *momentum constraints* $H_i := \partial\tilde{H}/\partial N_i = 0$ (which conserve vanishing canonical momenta p_{N_i}, and which are fulfilled automatically for the Friedmann solution because of its symmetry) have then to be solved explicitly. The lapse function $N(x, y, z, t)$ now determines truly *many-fingered* time (as a spatial *field* on the dynamically evolving hypersurface) with respect to the coordinate t. If, nonetheless, N is chosen as a function of t only, the foliation proceeds everywhere according to physical time (*normal* to the hypersurface) with fixed 'comoving' coordinates.

This is not always a convenient choice. For example, observers passing very close by a black hole horizon (without entering it) would experience an extreme time dilation (or observe the stars moving very fast through a little hole in the sky). In a recollapsing universe they could enter the expansion era within relatively short proper time. (This would render the immediate vicinity of horizons very sensitive to any conceivable cosmic *final* condition — see Zeh 1983 and Sect. 6.2.3.) In this case, a foliation according to *York time*, mentioned in Sect. 5.1, may be appropriate, since it arrives 'simultaneously' at all final singularities. Note, however, that the external curvature scalar K, which defines York time, is *not* a function of state, $f({}^{(3)}G)$.

Among the simplest inhomogeneous models are the spherically symmetric ones, with a metric

$$ds^2 = -N(\chi, t)^2 dt^2 + L(\chi, t)^2 d\chi^2 + R(\chi, t)^2 \left[d\theta^2 + \sin^2\theta d\phi^2\right] \quad . \tag{5.36}$$

They contain one remaining spatial gauge function, that has to be eliminated by means of the momentum constraint $H_\chi = 0$. It is analogous to Gauß' law in electrodynamics, as it refers to the radial coordinate.

Qadir (1988) proposed an illustrative toy model for such an inhomogeneous universe (Fig. 5.8). It forms a generalization of the Oppenheimer-Snyder model for the gravitational collapse of a homogeneous spherical dust cloud (see Misner, Thorne and Wheeler 1973, Chap. 32). The latter model pastes a spherical spatial boundary of a contracting, positive curvature Friedmann universe smoothly and consistently onto the spatial boundary of a Schwarzschild-Kruskal solution that extends to spatial infinity. Qadir connects the Schwarzschild solution in turn with another (much larger) partial Friedmann universe with much smaller mass density. This pasting at two spatial boundaries, with Friedmann coordinate values χ_1 and χ_2, say, is consistent if the masses of the two partial Friedmann universes are identical, and can thus be identified with the Schwarzschild mass M. The interpolating vacuum region is a strip from Fig. 5.2, spatially bounded by two non-intersecting geodesics which lead from the past to the future Kruskal singularity.

In order to comply with the Weyl tensor hypothesis as much as possible, Qadir assumed the two partial Friedmann universes to touch at the big bang. (In this deterministic model, an initial inhomogeneity is required as a seed.) Since the denser part of this toy universe feels stronger gravitational attraction, it contracts (or its expansion decelerates) faster. An empty

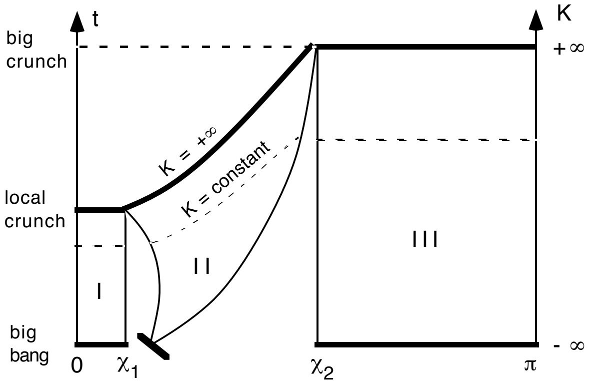


Fig. 5.8. Qadir's 'suture model' of a collapsing homogeneous dust cloud, I, in an expanding and recollapsing Friedmann universe, III. The Friedmann surfaces at χ_1 (left) and χ_2 (right) are initially identified. The Weyl tensor, representing the gravitational degrees of freedom, is chosen to vanish initially (except at the spatial boundary between the two regions), but will evolve and grow with the emerging interpolating 'Schwarzschild-Kruskal corridor', II, (a strip from Fig. 5.2). This corridor, with vanishing stress-energy tensor, must build up between the two regions with different positive spatial curvatures and energy densities. The spatial boundaries of the three spacetime regions have to be identified (including proper times on them, all chosen to start at the big bang). In a picture due to Penrose, the local crunch inside the black hole (region I) together with its attached Kruskal singularity (in region II) appears as a 'stalactite' hanging from the ceiling (representing the big crunch singularity — in region III). In contrast, there is only one (piecewise homogeneous) big bang singularity (a flat floor in Penrose's picture) at $K = -\infty$, that has been chosen as the first slice of the foliation (with $t = 0$)

'Schwarzschild corridor' must then form at the density discontinuity, and grow in size with increasing temporal distance from the big bang. As the energy-momentum tensor vanishes in the Schwarzschild-Kruskal region, its curvature is entirely due to the Weyl tensor, while the latter vanishes inside the two partial Friedmann universes. The time arrow of this process of 'gravitational monopole radiation' (the formation of the corridor with its non-zero gravitational degrees of freedom) is once more a consequence of the special initial condition.

This model is certainly interesting as an illustration of the Weyl tensor hypothesis, but it does not describe statistical (entropic) aspects. For this purpose, many degrees of freedom (such as transversal gravitational waves) would have to be taken into account. Qadir's cosmic evolution process simply describes an example of motion away from the chosen initial state; motion similar to that normally found in unbound mechanical systems (regardless of any statistical considerations). The term 'initial' refers here to the starting point of the computation, but not necessarily to a physical direction of time.

General literature: Chap. 21 of Misner, Thorne and Wheeler 1973

6. The Time Arrow in Quantum Cosmology

– Prelim. 4th edtn. (Feb 01): www.time-direction.de

> Our mistake is not that we take our theories too seriously, but that we do not take them seriously enough. Stephen Weinberg (1977)

The founders of quantum theory invented their theory as a theory of atoms, that was soon successfully applied also to other microscopic systems. Macroscopic objects were thought to require the established classical concepts. This point of view still seems to form the majority opinion among physicists *in spite* of their assertion that quantum theory be universally valid. We have seen in Sect. 4.3 that this schizophrenic position is not inevitable, since decoherence allows *quasi-classical concepts to emerge* from quantum mechanical ones within reasonable assumptions.

Everett (1957) seems to have first seriously considered a wave function of the universe (that must contain *internal observers*, required for its interpretation). Although he had in mind the quantization of general relativity with its cosmological aspects, Everett applied his ideas, which were based on a time-dependent Schrödinger equation, to nonrelativistic quantum theory. His main interpretational obstacle was the entanglement arising from measurements described by means of von Neumann's unitary interaction (4.30). This led him to his 'extravagant' interpretation (in Bell's words) in terms of *many* quasi-classical 'branches', which are separately experienced but are all assumed to exist in one superposition that defines the true and dynamically consistent quantum world. Beyond measurements proper and occasional interactions he does not seem to have regarded entanglement as particularly important (cf. Tegmark 1998).

The quantitative considerations reviewed in Sect. 4.3 have shown that *uncontrollable* ('measurement-like') interactions of von Neumann's type are essential and unavoidable for almost all systems under all realistic circumstances. Strong entanglement is, therefore, a *generic* aspect of quantum theory. The more macroscopic a system, the stronger its entanglement with its environment. The concept of a (pure) quantum state can be consistently applied only to the universe as a whole (Zeh 1970, Gell-Mann and Hartle 1990). This seems to be a far more powerful argument for quantum cosmology than merely an attempt to quantize general relativity or some unified field theory.

The quantum state of the universe must then *also* include quantum gravity (entangled with matter) with its novel conceptual consequences (see Sect. 6.2). However, many quantum cosmological aspects may be formulated on a quasi-classical background spacetime, using a fixed foliation, parametrized by a time coordinate t. Global states can then be dynamically described by means of a time-dependent Schrödinger equation with respect to this coordinate time t. This dynamics will be *derived* from quantum gravity (with its concept of an *intrinsic* time) in Sect. 6.2.2 as an approximation. Global states (such as those of quantum fields) depend on a foliation (or a reference frame) even on flat spacetime, while the density matrix of any *local* system should be invariant under a change of foliation that preserves the local rest frame — a requirement that does not seem to have attracted much attention.

If the quantum universe is thus conceptually regarded as a whole, it does not decohere, since there is no environment. Decoherence is meaningful only for *subsystems* of the universe, and with respect to observations by other subsystems (Wheeler's 'observer-participators'). If, furthermore, no objective collapse of the wave function is assumed to apply, one is forced to accept Everett's 'extravagant' wave function, which is a superposition of all 'possible' outcomes of measurements and measurement-like processes that ever occured in the universe. This global quantum state may always be assumed to be pure, since a global density matrix could be consistently interpreted as representing incomplete information about such a pure state. A measurement that merely selects a subset from those states which diagonalize this density matrix would be equivalent to a classical measurement (as depicted in Fig. 3.5 — in contrast to Fig. 4.3).

Initial *and* (nontrivial) final density matrices for the universe were suggested by Gell-Mann and Hartle (1994) to be used in Griffith's expression for probabilities of 'consistent histories'. Applied to *individual* measurements (cf. Sect. 4.6), final conditions would describe *postselection*, as discussed by Aharonov and Vaidman (1991) — see also Vaidman (1997). In contrast to Everett or collapse models, these interpretations do not assume the wave function to represent reality.

The decoherence of subsystems according to a *global* time-dependent Schrödinger equation leads dynamically to robust branches. They form approximately autonomous components that may factorize in the form $\phi_{obs1}\phi_{obs2}\ldots\psi_{rest}$ with respect to observer states describing robust (objectivizable) memory (cf. Sect. 4.3.2 and Tegmark 2000). This specific unitary evolution requires an arrow of time corresponding to a cosmic initial condition of type (4.56). Branching into components which contain definite observer states has to be *taken into account* in any *effective* dynamics that is to describe the history of the (quasi-classical) 'observed world' in quantum mechanical terms (cf. Fig. 4.3). It need *not* represent an objective (albeit objectivizable) fundamental dynamical process. The decrease of physical entropy, characterizing this 'apparent collapse' into a specific outcome, may be negligible on a thermodynamical scale in the usual situation of a measurement. Yet it may have dramatic consequences for phase transitions which describe a dynamical symmetry-breaking of the global vacuum. This will now be discussed:

6.1 Phase Transitions of the Vacuum

Heisenberg (1957), and Nambu and Jona-Lasinio (1961), invented the concept of a vacuum that breaks symmetries of a fundamental Hamiltonian 'spontaneously' (fact-like) — just as most physical states do. This suggestion is based on an analogy between the vacuum (the ground state of quantum field theory) and the phenomenological ground states of macroscopic systems, such as ferromagnets or solid bodies in general. Such ground states lead to specific modes of excitation, which in quantum theory define quasi-particles (phonons, for example). They form 'Fock spaces', spanned by their corresponding occupation number states. Similarly, a symmetry-violating vacuum may lead to *Goldstone bosons*, based on space-dependent oscillations about the macroscopic (collective) 'orientation' in the abstract space that is spanned by the generators of the presumed symmetry group. At their low initial excitation, the new effective degrees of freedom may then represent a very large entropy (or heat) capacity.

A symmetry-breaking (quasi-classical) 'ground state' is in general not an eigenstate of the symmetric (fundamental) Hamiltonian; it just forms the ground state of an effective (asymmetric) Fock space Hamiltonian. While those nondiagonal elements of the fundamental Hamiltonian which connect states of different collective orientation (lying in different Fock spaces) may become very small (or even vanish in a certain mathematical idealization), they remain essential for the definition of the exact eigenstates, since the coresponding diagonal elements must be equal under the assumed symmetry.

While the symmetry-breaking vacuum was initially assumed to be universally given as part of the theory (in order to define its 'absolute vacuum'), its excitations were used to form time-dependent states dynamically controlled by the Fock space Hamiltonian. The analogy was later generalized to describe a dynamical phase transition of the vacuum (in accordance with the exact Hamiltonian) during early stages of the universe. It may be caused, for example, by a varying external parameter (such as energy density, reflecting the expansion of the universe). The arising 'unitarily inequivalent' Fock spaces represent robust Everett branches (unless there is an objective collapse into one of them). Even the P or CP-violating terms of the weak-interaction Hamiltonian may have dynamically *emerged* in this way.

A popular model for describing symmetry-breaking in quantum field theory is the 'Mexican hat' or 'wine bottle potential' of the type $V(\Phi) = a|\Phi|^4 - b|\Phi|^2$ (with $a, b \geq 0$) for a fundamental complex matter field Φ (such as a *Higgs field*). It is assumed to possess a degenerate minimum on a circle in the complex plane with $|\Phi| = \Phi_0 > 0$. The classical field configurations of lowest energy may then be written as $\Phi \equiv \Phi_0 e^{i\alpha}$, with an arbitrary phase α. They break the Hamiltonian's symmetry under rotations in the complex Φ-plane. These classical ground states correspond to different quantum mechanical vacuum states $|\alpha\rangle$. One of them, $|\alpha_0\rangle$, say, is assumed to characterize our observed world (while the value of α_0 is physically meaningless).

A *dynamical* phase transition of the vacuum is now described by assuming that the universe was initially in the symmetric vacuum $|\Phi \equiv 0\rangle$ (a *symmetric initial condition*, or *SIC* — cf. Conradi and Zeh 1991 in a related context). This became a 'false' vacuum (relative minimum) by means of a changing parameter. It is then assumed to undergo a transition into a Fock space coresponding to one of the 'physical vacua' $|\alpha\rangle$. If potential energy is thereby released in a 'slow role' (similar to latent heat), it must be transformed into excitations (particle creation). However, just as in a normal phase transition, a symmetric state cannot evolve unitarily into an asymmetric one in accordance with a symmetric Hamiltonian. If the initial state is assumed to be pure, a unitary evolution can only lead to a symmetric superposition of all symmetry-breaking states. For example, the superposition of all physical vacua,

$$|0_{sym}\rangle = \frac{1}{N}\int |\alpha\rangle \, d\alpha \quad \neq \quad |\Phi \equiv 0\rangle \quad , \tag{6.1}$$

may possess even lower energy than $|\alpha\rangle$, and thus define the ground state of the exact theory. In view of the success of the Schrödinger equation, this consequence should not simply be regarded as 'formal', in particular in the absence of any external observer of the quantum universe. A globally symmetric superposition of type (6.1) would remain intact even when each of its component on the RHS contains or develops uncontrollable excitations in its own Fock space. A similar *intrinsic symmetry breaking* is known from deformed nuclei when forming angular momentum eigenstates by means of superpositions of different orientations in space, or from BCS states of superconductivity when forming particle number eigenstates in an analogous way (Zeh 1967). It describes intrinsic complexity, but not a global asymmetry. If $\pi_\alpha := i\partial/\partial\alpha$ generates a gauge transformation, equation (6.1) describes a state obeying a gauge constraint, $\pi_\alpha|\psi\rangle = 0$ (see Sect. 6.2).

An objective transition into *one* of the components would require a collapse of the state vector — similar to measurements or phase transitions in conventional description. Later emerging and evolving excitations may then lead to an observer. In the Everett interpretation, one obtains a superposition of different α-'worlds' (branches) with their corresponding α-observers. The question whether they all 'exist' could be answered operationally only by manipulating them to interfere locally (to 'recohere' — cf. (4.55)). While the superposition of different orientations, say, of a solid body would be docohered by its environment, a global vacuum can only be entangled with its *intrinsic* excitations. The 'order parameter' appears as a classical quantity (a fact) whenever a robust memory state of an observer is (or can only become) quantum correlated to it rather than to its superpositions.

As a consequence of the degeneracy, each homogeneous *classical* state α_0 permits small space-dependent oscillations, $\alpha_0 + \Delta\alpha(\boldsymbol{r}, t)$. Quantum mechanically they describe the massless *Goldstone bosons* (excitations with vanishing energy in the limit of infinite wave length). Their thermodynamically relevant degrees of freedom are thus *created* by the apparent indeterminism of the intrinsic symmetry breaking. It is remarkable that the most important cosmic entropy capacities are represented by zero-mass fields, *viz.*, electromagnetic and gravitational fields (Zeh 1986a, Joos 1987). If, for example, the electromagnetic degrees of freedom had emerged by means of symmetry-breaking, this would dynamically *explain* the Ritz conjecture of Chap. 2.

In contrast to the false vacuum, the symmetric superposition (6.1) describes a nonlocal (entangled) state — even in the absence of any excitations. If, for simplicity, one neglects the relativistic correlations of the vacuum discussed in Sect. 5.3, each vacuum $|\alpha\rangle$ may, just as the false vacuum, be written as a direct product of vacua on volume elements ΔV_k,

$$|\alpha\rangle \approx \prod_k |\alpha_{\Delta V_k}\rangle \quad . \tag{6.2}$$

This non-relativistic approximation describes a pure vacuum state on all volume elements (local subsystems). The superposition (6.1) would instead lead to a mixed state,

$$\varrho_{\Delta V_k} \propto \int |\alpha_{\Delta V_k}\rangle\langle\alpha_{\Delta V_k}|\, d\alpha \quad , \tag{6.3}$$

when subjected to a Zwanzig projection $\hat{P}_{\text{sub}}$, This density matrix, meaningful only to an imaginary *external* observer (who does not 'live' in one of the Fock spaces), describes a canonical distribution of Goldstone bosons with infinite temperature (where $e^{-E/kT} \rightarrow 1$) with respect to each of the vacua. Therefore, only a (genuine or apparent) collapse into *one* component gives rise to the pure (cold) vacuum experienced by the local observer.

Degeneracy parameters such as α (or other *macroscopic* field quantities arising from a phase transition) need not have the same value throughout the whole universe, but may be different in various spatial regions (similar to a ferromagnet again). If these regions are macroscopic, and thus decohere and become 'real', they break translational symmetry (Calzetta and Hu 1995, Kiefer, Polarski and Starobinsky 1998). Any homegeneous superposition of *microscopic* inhomogeneities would represent 'virtual' symmetry breaking ('vacuum fluctuations') — related to virtual decoherence (Sect. 4.3).

If a phase transition of the vacuum were able in a similar way to explain the Weyl tensor hypothesis in a causal manner (as discussed in Sect. 5.3), it would establish a master arrow of time. If this required a genuine collapse of the wave function as a fundamentally *asymmetric* process, there would indeed be no reason to expect the reverse situation after the 'turning of the tide' in an imagined recontracting universe. If it is instead described by means of decoherence on the basis of the Everett interpretation (presuming T- or CPT-symmetric dynamics), one has to assume fact-like 'quantum causality' (as described (4.56)), which may or may not be reversed, depending on the boundary conditions at the big bang and the big crunch. (See the following section for a time-less description in quantum gravity, however.)

6.2 Quantum Gravity and the Quantization of Time

Many physicists seem to believe that quantum gravity is still elusive, or even that quantum theory and general relativity are incompatible with one another. This prejudice derives from a number of quite different roots:

Einstein's attitude regarding quantum theory is well known. In particular, he is said to have remarked that a quantization of general relativity would be 'childish'. Another objection is based on the argument that gravitons may be unobservable, and the quantization of gravity hence *not required* (von Borzeszkowski and Treder 1988). However, a classical gravitational field or spacetime metric is inconsistent with quantum mechanics, since it would always allow one to determine the exact energy of a quantum object — in conflict with the uncertainty relations. This has been known since the early Bohr-Einstein debate (see Jammer 1974, for example), while other consistency questions were raised by Page and Geilker (1982). Concepts of quantum gravity are essential in cosmology and for a master arrow of time (as will be discussed). The classical *appearance* of spacetime cannot be regarded as an argument against its quantization, since these classical aspects may be understood within a universal quantum theory in a similar way as other quasi-classical properties have been (cf. Sect. 4.3).

The most popular objection against quantum gravity holds that straightforward quantization of general relativity does not lead to a renormalizable theory, and is therefore inconsistent. This argument would be valid if quantum gravity were understood as an exact theory. However, it can only be expected to represent an 'effective theory' that describes specific low energy aspects of an elusive unified field theory. We know that QED, too, has to be modified and replaced with electro-weak theory at high energies, while it represents an excellent and consistent description of all relevant phenomena at low energies. Its most general quantum aspects (described in terms of QED wave functionals that are not restricted to perturbation theory and Feynman diagrams) are studied with atomic and laser physics (cf. Sect. 4.3.4). An analogous (though technically and conceptually more demanding) canonical method of quantizing general relativity leads to the Wheeler-DeWitt equation (6.4) below (DeWitt 1967). Why should the Einstein equations be saved from quantization, while the Maxwell equations are not? The mere combination of two well-established theories (quantum theory and general relativity) can hardly be regarded as speculative as, for example, a theory of superstrings. (The latter would have to include quantum gravity as an effective low energy theory.) Nonetheless, the conceptual consequences of quantum gravity are profound.

Of course, the construction of a unified theory represents the major challenge to quantum field theory at its fundamental level. While its consequences must become essential in the vicinity of spacetime singularities (inside black holes or close to the big bang), they fortunately seem to be irrelevant for most *conceptual* quantum cosmological questions. The latter may thus be discussed by using effective quantum gravity. In contrast, models of unified field theories have so far mostly been studied at their classical level (perhaps including effective quantum correction terms), even though they are often *motivated* by quantum aspects (such as the renormalizability of perturbation theory). The suggestion that *M-theory* may eventually lead to a *derivation* of quantum theory (Witten 1997) seems to be based on a misunderstanding of quantum mechanics.

Much progress in canonical gravity has been achieved during recent years by using loop integrals (based on Ashtekhar variables) instead of local metric fields (Ashtekhar 1987, 1991, Rovelli and Smolin 1990). This approach offers fundamental insights into the consequences of general diffeomorphism invariance, since only the knot structure of these loops remains invariantly meaningful. Knots can be classified by discrete numbers, which are then often misinterpreted as quantum aspects, although they merely characterize certain classical solutions (or 'modes' — similar to multipoles in electrodynamics). Quantization (that is, application of the superposition principle) is rarely more than mentioned in this approach. Therefore, the Wheeler-DeWitt equation in its *field* representation appears to be the method of choice in quantum cosmology.

A major problem that nonetheless prevents many physicists from accepting the Wheeler-DeWitt equation as correctly describing quantum general relativity is the absence of any time parameter in the case of a closed universe (see Isham 1992). According to the Hamiltonian formulation, indicated in Sect. 5.4, one would naively expect free gravity to be described by a time-dependent wave functional on the configuration space of three-geometries, $\Psi[^{(3)}G, t]$, dynamically governed by a Schrödinger equation, $i\partial\Psi/\partial t = H\Psi$. However, there is neither an absolute time any more, nor are there trajectories of appropriate physical clock variables which could give this equation a physical interpretation. Different three-geometries $^{(3)}G$ (classically the carriers of 'information' about *many-fingered* time) occur instead as arguments of these wave functionals. In the absence of trajectories $^{(3)}G(t)$ (that is, spacetime geometries $^{(4)}G$), neither proper times nor time coordinates are available. Therefore, it appears quite consistent (cf. Zeh 1984, 1986b) that the quantized form of the Hamiltonian constraint, $H\Psi = 0$, completely removes any time parameter t from the wave function of any closed (though not necessarily finite) universe. [1]

If matter is now represented by a single scalar field Φ on space-like hypersurfaces that are defined by their three-geometries $^{(3)}G$ — leading to a

---

[1] Only because of the Heisenberg picture is the equation $H\psi = E\psi$ usually called a *stationary* Schrödinger equation, while in wave mechanical terms it describes *static* solutions. Even 'vacuum fluctuations' represent static quantum correlations (entanglement). Similarly, eigenvalues of the momentum operator are no more than *formally* analogous to classical momenta (which are defined as time derivatives). These conceptual subtleties will turn out to be essential for a consistent interpretation of the Wheeler-DeWitt equation.

Hamiltonian as in (5.34) —, the Wheeler-DeWitt equation assumes the form

$$H\Psi[\Phi,^{(3)}G] = 0 \quad . \tag{6.4}$$

Even though it represents dynamics, this does not describe a one-dimensional succession of states (or a history that may be labelled by a parameter t). This is the natural quantum consequence of a classically missing *absolute time* (the absence of any *preferred* time parameter). In contrast, *classical* general relativity still describes parametrizable trajectories in its configuration space, which allow the unique (one-dimensional) ordering of motion by comparison with physical clocks ('physical time'). While their reparametrization invariance leads to the Hamiltonian constraint, the latter's quantized form requires the wave function of the closed universe to be static. Although quasi-trajectories may then still be justified by means of intrinsic decoherence within a WKB approximation (see Sect. 6.2.2), the absence of any fundamental trajectories in configuration space *or* in Hilbert space leads to the new problem of how to pose a fundamental 'initial' condition of low entropy that would be able to explain the arrow of time in the conventional way.

In (6.4), the matter field Φ represents symbolically *all* matter, including material clocks that are dynamically controlled by the metric. A 'dust of test clocks' (as classically always conceivable without affecting the geometry) would in quantum theory decohere a superposition of different three-geometries, since these clocks would 'measure' proper time along their world lines. (This does indeed come close to what really happens in our universe.)

An uncertainty relation between time and energy has always been known in *quantum mechanics*. It forms an *incomplete* analogy to that between position and momentum (cf. Aharonov and Bohm 1961), since no time operator, canonically conjugate to the Hamiltonian, can be defined in a natural way, while a classical time *parameter* t is used in the Schrödinger equation. In quantum field theory, the coordinates x^k are parameters, too, rather than dynamical variables that have to be replaced by operators (see Hilgevoord 1998). All uncertainty relations reflect the Fourier theorem for wave functions depending on *configuration* space *and* time (even though t does not label any representation in Hilbert space). A fixed energy (frequency) for a closed system thus rules out any time dependence.

If time were consistently defined by motion (as discussed in Chap. 1), there could indeed be no time dependence in the Schrödinger equation for a closed quantum universe (such as the one governed by (5.32)). Instead of an absolute time parameter t, one would have to refer to a physical clock variable u, say, that then has to be quantized, too. The time-dependent wave function $\psi(x,t)$ is thus replaced with an entangled wave function $\psi(x,u)$ (Peres 1980b, Page and Wootters 1983, Wootters 1984). A quasi-classical time-less trajectory could then only be represented by a narrow 'standing wave tube' in this configuration space. In the conventional probability interpretation, $\psi(x,u)$ would describe a probability amplitude *for* 'time' u — not *at* u. Only wave functions for *given* values u_0, $\psi(x, u_0)$, could be interpreted as 'time'-dependent quantum states for x. Why do we always observe states 'at' such a definite time rather than their superpositions? We have to expect the relevant clock variables u (such as those contained in the three-geometry ${}^{(3)}G$) to be quasi-classical for reasons discussed in Sect. 4.3.

The Schrödinger equation of quantum mechanics is applied to *local* systems with respect to classical *proper* times. Spatially extended relativistic systems can be dynamically described only with respect to a foliation in terms of three-geometries. The corresponding global Hamiltonian will in general depend on this foliation in a complicated manner. If the spacetime metric is classically given, the dynamics with respect to a foliation defines the dynamics with respect to many-fingered physical time, which is determined by the change of the classical ${}^{(3)}G$. In *quantum* gravity with matter, $\Psi[\Phi, {}^{(3)}G]$ is the formal 'timeless' probability amplitude for configurations of matter *and* geometry as they may occur on space-like slices for *all* foliations.

Equation (6.4) does not yet represent the Wheeler-DeWitt equation in a form that can be used. Practically, one has to represent the three-geometry ${}^{(3)}G$ by a metric $h_{kl} = g_{kl}(x^1, x^2, x^3)$ with respect to a certain choice of coordinates (cf. Sect. 5.4). The wave functional $\Psi[h_{kl}]$ must then be invariant under spatial coordinate transformations. This is guaranteed by the three secondary momentum constraints, classically described as $H_i = 0$ (with $i = 1, 2, 3$). In their quantum mechanical form they have again to be imposed as constraints on the wave function, $H_i\Psi[h_{kl}] = 0$, similar to the Hamiltonian constraint. The wave functional $\Psi[h_{kl}]$ represents $\Psi[{}^{(3)}G]$ if the momentum constraints are fulfilled.

General literature: DeWitt 1967, Wheeler 1979

6.2.1 Quantization of the Friedmann Universe

A toy model of a quantum universe can be defined in analogy to the classical Friedmann universe described in Sect. 5.3 by presuming homogeneity and isotropy (Kaup and Vitello 1974, Blyth and Isham 1975). While the corresponding Wheeler-DeWitt equation may then simply be obtained by quantizing the classical model, it does not represent an approximation to reality. Symmetry requirements are much stronger in their quantum mechanical form than they are classically. For example, the rotation of a macroscopically spherical body reproduces a macroscopically equivalent state, whereas a spherically symmetric *quantum* state would remain invariant under this symmetry transformation. Translations and rotations are thus *identity* operations in a quantum Friedmann universe. For similar reasons, an intrinsically spherical quantum object does not possess any rotational degrees of freedom. Rotational spectra are restricted to situations of 'strong intrinsic symmetry-breaking' (Zeh 1967). The quantum Friedmann model therefore has to be regarded as merely a very first step in constructing a realistic model of the quantum universe (except close to the big bang). The configuration space based on this simple model is usually called a 'mini-superspace' (cf. Sect. 5.4).

If the Hamiltonian (5.34) is quantized in the conventional way in its field representation, canonical momenta have to be replaced with the corresponding differential operators. In general this leads to a factor-ordering problem for the Hamiltonian, since commutators $[p, q]$ vanish classically, and terms proportional to them could thus be arbitrarily added before quantization. Although (5.34) looks quite 'normal', its straight-forward translation into a Wheeler-DeWitt equation,

$$\frac{e^{-3\alpha}}{2}\left(\frac{\partial^2}{\partial\alpha^2} - \frac{\partial^2}{\partial\Phi^2} - ke^{4\alpha} + m^2 e^{6\alpha}\Phi^2\right)\Psi(\alpha,\Phi) = 0 \quad , \tag{6.5}$$

is far from being trivial. For example, the result would have been different for quantization in terms of the expansion parameter a instead of its logarithm α. The differential operator used in (6.5) is the invariant d'Alembertian with respect to DeWitt's *superspace metric* that is defined by the quadratic form representing the kinetic energy. This factor-ordering is analogous to that for a point mass in flat space when described in terms of non-cartesian coordinates. The pre-factor $e^{-3\alpha}/2$ can then be omitted from the Wheeler-DeWitt equation (6.5).

Even though the Wheeler-DeWitt equation is a stationary Schrödinger equation, with kinetic energy represented by a Laplace type operator, (6.5) and its higher-dimensional generalizations to more complex Friedmann type universes are hyperbolic (similar to a Klein-Gordon equation with variable mass) for reasons explained in Sect. 5.4. This situation offers the surprising possibility of formulating an *intrinsic initial value problem* (cf. Sect. 2.1) in spite of the absence of any time parameter. The logarithmic expansion parameter α may be regarded as a time-like variable with respect to this intrinsic dynamics. The wave function on a 'space-like' hypersurface in superspace (for example at a fixed value of α) then defines an *intrinsic dynamical state.* It is essentially the overall attractive nature of gravity that is responsible for this property of the superspace metric.[2] However, a time-like variable in superspace that controls this intrinsic quantum dynamics has clearly to be distinguished from many-fingered time intself, or from a time coordinate t; these are both functions on spacetime, and therefore applicable only in a quasi-classical approximation (see Sect. 6.2.2).

The big bang and a conceivable big crunch would 'coincide' with respect to this intrinsic 'quantum dynamical time' α, whereas the expansion of the universe would be a tautology. However, all correlations of the expansion

[2] The possibility and physical relevance of a signature change *in superspace* (Kiefer 1989, Giulini and Kiefer 1994) does not appear to be well understood. It may perhaps be circumvented from a local point of view by an appropriate choice of spacetime foliation.

parameter α with physical clocks (including physiological ones) remain meaningful. (External test clocks are excluded from a closed quantum universe.) The concept of a reversal of the cosmic expansion remains an artifact of the classical description in terms of trajectories. But what would 'happen' to a quasi-classical universe that were classically bound to recontract at a later time?

One may study the corresponding quantum state (its Wheeler-DeWitt wave function) in analogy to the 'stationary' wave function (that is, a superposition of incoming and outgoing waves) of a quantum particle with fixed energy E, reflected from a spatial potential barrier. A concept of time is then often *introduced* by means of 'corresponding' trajectories, although this situation is more correctly described dynamically in terms of time-dependent wave packets. However, in *fundamentally static* quantum gravity there is no reference phase $e^{-i\omega t}$ that would allow one to define moving wave packets by means of a phase velocity $e^{ik(x-vt)}$ for any dynamical variable x. Hence, there is no distinction between expanding and contracting wave components of the universe according to the sign of $e^{\pm ik\alpha}$.

Intrinsic quantum dynamics with respect to the 'variable' α may also be written by means of a Klein-Gordon type *reduced Hamiltonian.* For example, in mini-superspace it assumes the form

$$-\frac{\partial^2}{\partial\alpha^2}\Psi(\alpha,\Phi) = \left(-\frac{\partial^2}{\partial\Phi^2} + V(\alpha,\Phi)\right)\Psi(\alpha,\Phi) =: H^2_{\mathrm{red}}\Psi(\alpha,\Phi) \quad . \tag{6.6}$$

This intrinsic dynamics is non-unitary (though still deterministic) not only because of the second 'time' derivative, but mainly since H^2_{red} is not a non-negative operator. The 'reduced norm', $\int |\Psi(\alpha,\Phi)|^2\, d\Phi$, is thus not generally conserved with respect to the intrinsic dynamics. This is consistent with the interpretation of α as a physical 'variable', and with the interpretation of Ψ as a probability amplitude *for* 'time' (rather than *depending* on time). Although there is a conserved formal 'relativistic' two-current density in this mini-superspace,

$$j := \mathrm{Im}\,(\Psi^*\nabla\Psi) \quad , \tag{6.7}$$

its direction depends on the sign of $i = \pm\sqrt{-1}$ again, which no longer has physical meaning in the absence of a time derivative $\mathrm{i}\partial\Psi/\partial t$ from the Wheeler-DeWitt equation. Since there is no good reason to consider complex boundary conditions in superspace, this current density must even be expected to vanish for the total (Everett) wave function. In analogy to a timeless trajectory, one now obtains a 'standing wave' representing timeless reality.

This intrinsic dynamics cannot be expected to be symmetric under the new 'time reversal' $\alpha \to -\alpha$ because of the α-dependence of the potential V (although even this symmetry under 'duality transformations' has been proposed in certain string models). This *intrinsic dynamical* (law-like) asymmetry would then open up the possibility of *deriving an arrow of time with respect to the expansion of the universe* without imposing asymmetric boundary conditions by hand.

Because of the analogy between the Wheeler-DeWitt equation and the stationary description of quantum particles, it is suggestive to represent the individual history of our observed universe in quantum mechanical terms by a narrow 'wave tube' that approximately follows a classical trajectory in superspace. Again, the case of an expanding and recontracting universe, with positive spatial curvature, $k = +1$, (cf. Fig. 5.7) is particularly illustrative. Its trajectories in mini superspace would reverse direction with respect to α at some $\alpha_{\max}$. According to *classical* determinism, half of the trajectory (if *defined* to represent the collapsing universe) would be regarded as the successor of the other half (representing the 'expansion' era). However, this deterministic relation is symmetric, since there is no absolute dynamical direction. The wave determinism described by the hyperbolic equation (6.6), on the other hand, proceeds always with growing α, and permits one to freely choose the whole initial condition, consisting of Ψ and $\partial\Psi/\partial\alpha$, on any 'space-like' hypersurface in superspace (e. g. at a small value of α). One may thus exclude precisely that part of the wave tube which would be required by classical determinism (Zeh 1988). How can these two forms of determinism be reconciled?

This problem can again be solved in analogy to conventional stationary states of quantum mechanics — now assumed to possess negative mass with respect to one variable in a formal configuration space. The simplest model of a recollapsing (and rebouncing) universe is a free 'particle' reflected between the walls of a rectangular box, where solutions (such as real-valued wave tubes) can be formed as superpositions of products of trigonometric functions matching the boundary conditions. (The boundaries would have to be represented by positive infinite space-like and negative infinite time-like potential barriers.) In order to allow nontrivial zero-energy solutions, $H\Psi = 0$, the box lengths L_i have to be commensurable when taking into account the mass ratio, that is, $L_\alpha/L_\Phi = \sqrt{-m_\Phi/m_\alpha}\, k/l$, with integers k and l.

Similar stationary wave tubes may be constructed analytically for an indefinite harmonic oscillator Hamiltonian with subtracted zero point energy, that is, for a two-dimensional 'particle' in the potential $2V(\Phi,\alpha) = (\omega_\Phi^2\Phi^2 - \omega_\Phi) - (\omega_\alpha^2\alpha^2 - \omega_\alpha)$. In the commensurable case, $\omega_\alpha/\omega_\Phi = l/k$, solutions to the constraint $H\Psi = 0$ may be expanded in terms of the factorizing eigensolutions $\Theta_{n_\alpha}(\sqrt{\omega_\alpha}\alpha)\Theta_{n_\Phi}(\sqrt{\omega_\Phi}\Phi)$ of H, with eigenvalues $E = E_\alpha + E_\Phi = -n_\alpha\omega_\alpha + n_\Phi\omega_\Phi = 0$, in the form

$$\Psi(\alpha,\Phi) = \sum_n c_n \Theta_{nk}(\sqrt{\omega_\alpha}\alpha)\Theta_{nl}(\sqrt{\omega_\Phi}\Phi) \quad . \tag{6.8}$$

They represent propagating wave *tubes* rather than packets as long as the reduced energy defined by H^2_{red} of (6.6) is positive (see Fig. 6.1). If the coefficients c_n are chosen to define an 'initial' Gaussian wave packet in Φ at 'time' $\alpha = 0$, centered at some $\Phi_0 \neq 0$ and with $\partial\Psi/\partial\alpha = 0$, say, the resulting tube-like solutions follow Lissajous figures, as discussed by DeWitt (1967) for classical orbits in mini-superspace. In the anti-isotropic case ($\omega_\alpha = \omega_\Phi$,

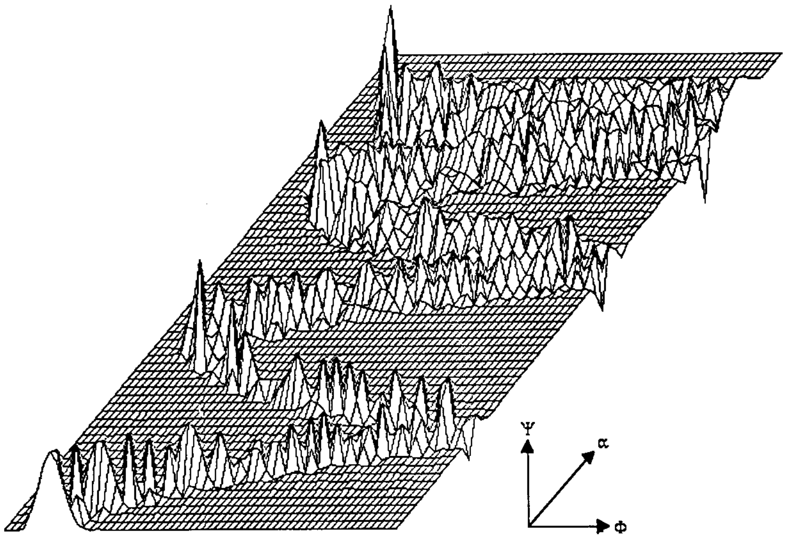


Fig. 6.1. Coherent 'wave tube' $\Psi(\alpha, \Phi)$ for the anisotropic indefinite harmonic oscillator (with $\omega_\Phi : \omega_\alpha = 7 : 1$) as a toy model of a periodically recollapsing and rebouncing quantum universe. It is plotted for the sector $\alpha > 0$ and $\Phi < 0$ only, since the potential and the chosen solutions are symmetric under reflections at both the α and the Φ axis, in contrast to Figs. 5.7 and 6.2. Interference occurs in regions where different parts of a 'wave tube' overlap. It would be suppressed by taking into account additional variables (see Sect. 6.2.2), since decoherence corresponds to the projection of the density matrix from high-dimensional superspace (where the wave tubes have little chance to come close to each other) on mini-superspace. (The intrinsic structure of the wave tube is here not completely resolved by the chosen grid size.)

that is, $k = l = 1$) they form ellipses. If one requires $\partial\Psi/\partial\alpha = 0$ at $\alpha = 0$, only the even quantum numbers n_Φ remain available for an expansion of initial states $\Psi(0, \Phi)$.

In contrast to Schrödinger's time-dependent coherent states for the conventional harmonic oscillator, which follow classical trajectories without changing their shape, these 'upside-down oscillators' display a rich intrinsic structure that represents relativistic 'rest energy oscillations' while following their quasi-classical trajectories. Wave packets on different parts of a trajectory must interfere whenever they come closer than the widths of their wave tubes. This happens particularly in the region of the classical turning point.

The wave functions (6.8) are chosen to satisfy the boundary condition of normalizability in Φ and α. This choice is responsible for the reflection of wave tubes at the potential barriers. It would be an unusual condition for a conventional time parameter, since in quantum mechanics there are no turning points in time. In quantum gravity, the 'final' condition restricts the freedom of choice of initial conditions, but this is consistent with the interpretation of α as a dynamical *variable*. On the other hand, any degeneracy of the solution

(6.8) would demonstrate that this boundary condition of normalizability at $\alpha \to \pm\infty$ does not completely determine the state, but offers *some* freedom for 'initial' conditions. In particular, a 'final' condition of normalizability in α (such as for a Friedmann universe that is bound to recontract) would eliminate just half of a complete set of solutions of this second order equation (the exponentially increasing ones) — certainly a much weaker restriction than a low entropy final condition would be (cf. (5.27)). Tube-like solutions may then mimic the classical determinism of returning orbits, since they are forced to contain the returning wave tubes from the 'beginning' (or, as in the right half plane of Fig. 6.1, at $\alpha = 0$).

A similar treatment can be applied, with some further approximations, to the Wheeler-DeWitt equation with a Friedmann Hamiltonian (6.5) (Kiefer 1988). The oscillator potential with respect to Φ may here be assumed to be weakly α-dependent over many classical oscillations in Φ (shown in Fig. 5.7) except for small values of α. If α-dependent oscillator wave functions $\Theta_n(\Phi)$ are now defined by the eigenvalue equation

$$\left(-\frac{\partial^2}{\partial\Phi^2} + m^2 \mathrm{e}^{6\alpha}\Phi^2\right)\Theta_n\left(\sqrt{me^{3\alpha}}\Phi\right) = (2n+1)me^{3\alpha}\Theta_n\left(\sqrt{me^{3\alpha}}\Phi\right), \quad (6.9)$$

one may expand

$$\Psi(\alpha,\Phi) = \sum_n c_n(\alpha)\Theta_n(\sqrt{me^{3\alpha}}\Phi) \quad . \quad (6.10)$$

This leads, in the adiabatic approximation with respect to α (that may here be based on the Born-Oppenheimer expansion in terms of the inverse Planck mass – see Banks 1985), to decoupled equations for the coefficients $c_n(\alpha)$,

$$\left(+\frac{\partial^2}{\partial\alpha^2} + 2E_n(\alpha)\right)c_n(\alpha) = 0 \quad . \quad (6.11)$$

For positive spatial curvature, $k = 1$, the 'effective potentials' $2E_n(\alpha) = (2n+1)me^{3\alpha} - k\mathrm{e}^{4\alpha}$ become negative for $\alpha \to +\infty$. Even though $V(\alpha,\Phi)$ is positive almost everywhere in this limit (cf. (5.34)), the Φ-oscillations are finally drawn into the narrow region $V < 0$ (in the vicinity of the α-axis — see Fig. 5.7) by their damping due to the α-dependence of E_n.

If one requires square integrability for $\alpha \to \infty$ again, only those partial wave solutions $c_n(\alpha)$ which exponentially die out in this limit are admitted. Wave packets consisting of oscillator states with quantum numbers $n \approx n_0$, say, (see Fig. 6.2) are reflected from the repulsive potential at their corresponding classical turning points — very much like fixed energy solutions of the Schrödinger equation that are reflected from a potential barrier in space.

In the Friedmann model, however, wave tubes cannot remain narrow wave packets in Φ during reflection, since the turning point of the n-th partial wave, $a_n = \mathrm{e}^{\alpha_n} = (2n+1)m$, depends strongly on n. For $a < a_n$ the coefficients $c_n(\alpha)$ can, according to Kiefer, be written by means of a 'scattering' phase shift as a sum of incoming and outgoing waves. In lowest WKB

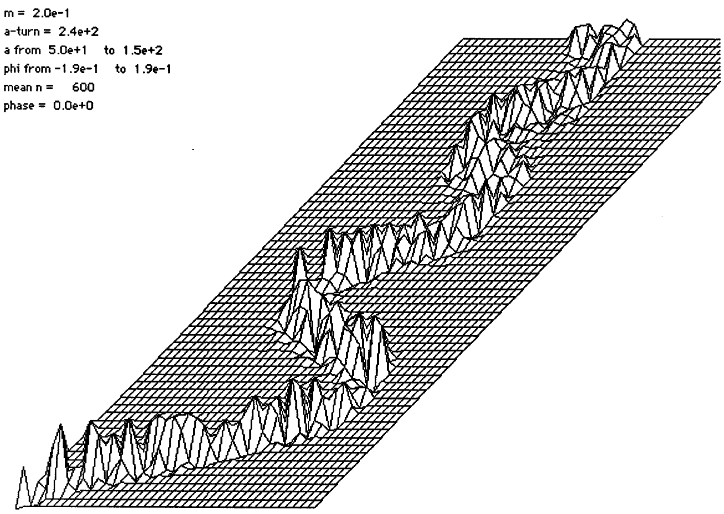


Fig. 6.2. Real-valued wave tube of the 'time'-dependent (damped) oscillator (6.5) in adiabatic approximation *without* taking into account the reflection at $a_{\max}$ (that is, the first term of (6.12a) is used only). Expansion parameter $a = \mathrm{e}^{\alpha}$ is plotted upward, the amplitude of a massive scalar field Φ from left to right

approximation one has

$$c_n(\alpha) \propto \cos\left[\phi_n(\alpha) + n\Delta\phi\right] + \cos\left[\phi_n(\alpha) - n\Delta\phi + \frac{\pi}{4}{a_n}^2\right] \quad . \qquad (6.12a)$$

Here,

$$\phi_n(\alpha) := \left(\frac{a_n}{4} - \frac{a}{2}\right)\sqrt{a(a_n - a)} + \left[\arcsin\left(1 - \frac{2a}{a_n}\right) - \frac{\pi}{2}\right]\frac{a_n^2}{8} - \frac{\pi}{4} \qquad (6.12b)$$

is a function of α and α_n, while the integration constant $\Delta\phi$ is the phase of the corresponding classical Φ-oscillation at its turning point in α. If complex factors c_n^0 are chosen in (6.12a) before substitution into (6.10) in such a way that the first cosines describe a coherent wave packet, the 'scattering phase shifts' $\pi {a_n}^2/4$ of the second cosine terms cause the reflected wave to spread widely (see Fig. 6.3). Only for pathological potentials, such as the indefinite harmonic oscillator (6.8), or for integer values of $m^2/2$ in the specific model (6.4), could the phase shift differences be omitted as multiples of π.

Therefore, even this WKB approximation can in general not describe an expanding and recollapsing quasi-classical universe by means of ('initially' prepared) wave packets that propagate as narrow tubes. The concept of a universe expanding and recollapsing along a certain trajectory in superspace (hence of the corresponding quasi-classical spacetime) is as incompatible with quantum cosmology as the concept of an electron orbit in the hydrogen atom is with quantum mechanics. Many 'expanding' quasi-trajectories (wave tubes)

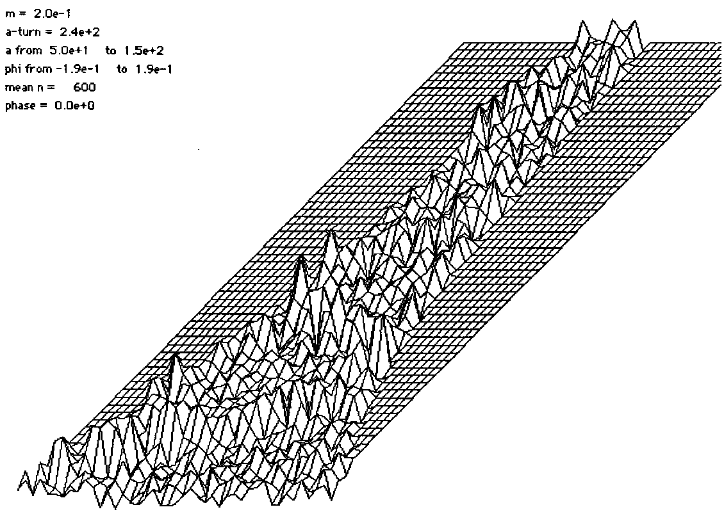


Fig. 6.3. Same as Fig. 6.2, but with contribution of reflected part (second term of (6.12a)) included. The coherent wave tube that is assumed to represent the expanding universe is here hardly recognizable against the background of the diffuse contribution representing the recontracting universe(s)

have to be superposed in order to describe *one* quasi-trajectory representing the collapse era (and vice versa).

These compatibility problems for boundary conditions do not occur for the (non-normalizable) Wheeler-DeWitt wave function that represents forever expanding universes (k= -1, for example). On the other hand, an additional boundary condition (selecting wave functions which vanish exponentially for $\alpha \to -\infty$) would further restrict the solutions, and could permit eigensolutions with eigenvalue zero (required by the Hamiltonian constraint) only in singular cases. The commensurable anisotropic oscillator (6.8) is a toy model. Solutions of (6.5) with $k = 0$ or -1 are not square integrable for $\alpha \to +\infty$. Such models without any repulsive potential for large a would have to be excluded if normalizability were required in order to obtain a Hilbert space for wave functions on superspace. Normalization in $\Phi \to \pm\infty$ has already been taken into account by the *ansatz* (6.10).

These qualitative properties of the Wheeler-DeWitt wave function in mini-superspace must in principle be expected to remain valid when further 'space-like' degrees of freedom are added in order to form a more realistic model (Sect. 6.2.2). They should similarly apply to unified field theories that contain gravity. Regarded as a toy model for a closed universe, $\Psi(\alpha, \Phi)$ obeys the Wheeler-DeWitt equation, while in the presence of additional variables it would be subject to strong decoherence (Zeh 1986b, Kiefer 1987).

General literature: Ryan 1972, Kiefer 1988

6.2.2 The Emergence of Classical Time

If classical time emerges, it cannot emerge in classical time

According to Sect. 5.4, in general relativity or other reparametrization invariant *classical* theories there is no *absolute* time defined as a dynamically preferred global time parameter. The dynamical succession of global states, which may be conveniently described by means of a time coordinate, depends furthermore on the choice of a foliation, while the invariant spacetime geometry defines many-fingered *physical* time, that is, invariant proper times as local controllers of motion *along all world lines*.

In *quantum* gravity, no global time parameter is available at all, since there are no trajectories in superspace (that is, no one-dimensional successions of classical states). Two given three-geometries can then not be dynamically connected in a way that would determine a classical four-geometry in between. Therefore, no world lines *in spacetime* are defined (hence no many-fingered proper times which could control Schrödinger equations for local systems). Three-geometry is no 'carrier of information about time' any more, while the foliation of a pure manifold (without a metric) would lack physical or geometrical (hence operational) meaning.

We know from Chap. 4 that the Schrödinger equation can be exact only as a *global* equation. Its application to *all matter* in the whole universe would require the global foliation of a *classical* spacetime. While the foliation can nonetheless be chosen to proceed only locally (defining a 'finger of time'), the Wheeler-DeWitt equation in its static form defines 'entangled dynamics' for states of matter *and* three-geometry.

How can classical time (either in the form of many-fingered time, or as a parameter for the dynamics of global states) be recovered from the Wheeler-DeWitt equation? This would require concepts and methods as discussed in Sect. 4.3, where quasi-classical quantities emerged dyamically and irreversibly by means of decoherence. This dynamics of decoherence has to be modified in a way that is appropriate for the timeless Wheeler-DeWitt equation: for if classical time *emerges*, it cannot emerge *in* classical time. Similarly, classical spacetime cannot itself have entered existence in a global 'quantum event' (cf. Sect. 4.3.5).

In the local field representation, the general Wheeler-DeWitt equation can be explicitly written in its gauge-dependent form (DeWitt 1967) as

$$-\frac{1}{2M}\sum_{klk'l'} G_{klk'l'}\frac{\delta^2\Psi}{\delta h_{kl}\delta h_{k'l'}} - 2M\sqrt{h}(R-2\Lambda)\Psi + H_{\mathrm{matter}}\Psi = 0 \quad (6.13)$$

(disregarding factor ordering, for simplicity). Here, Ψ is a functional of the *six* independent functions $h_{kl} = g_{kl}$ $(k,l = 1,2,3)$, which represent the spatial metric on the hypersurface. h is their determinant, R the corresponding spatial Riemann curvature scalar, and Λ the cosmological constant.

$M^{1/2} := (32\pi G)^{-1/2}$ is the Planck mass, and H_{matter} the Hamiltonian density of matter, implicitly depending on the metric through kinetic energy terms. $G_{klk'l'} := (h_{kk'}h_{ll'} + h_{kl'}h_{lk'} - 2h_{kl}h_{k'l'})/2\sqrt{h}$ is DeWitt's superspace metric with respect to the six symmetric *pairs* of indices kl. It has the *locally* hyperbolic signature $- + + + + +$, hence one minus sign at each space point. The global Hamiltonian H is obtained by integrating the density (6.13) over space. This integral is required in addition to the sum over vector indices in order to define the complete Laplacean in configuration space as a sum over *all* dynamical variables.

The Wheeler-DeWitt equation (6.13) is local at the cost of gauge degrees of freedom. Their elimination requires the wave functional to obey the three momentum constraints (cf. Sect. 5.4), in their quantum mechanical form written as $H_k\Psi = (\delta\Psi/\delta h_{kl})|_l = 0$, where $|_l$ is the covariant derivative with respect to the spatial metric h_{kl}. They represent *three* functional differential equations for the functional $\Psi[h_{kl}]$. Functional integration of the (*one*) Wheeler-DeWitt equation (6.13) then leaves *two* functions as integration constants, even though there is no explicit time dependence. These two degrees of freedom at each space point may be regarded as representing the two components of the graviton field. Only in special situations (and by means of nonlocal variables) can they be analytically separated from the gauge variables and intrinsic time. In the superspace region that describes Friedmann type universes, all but one of the negative kinetic energy terms of the Wheeler-DeWitt equation (one at each space point) represent gauge degrees of freedom (cf. Footnote 2).

The value of the Planck mass that appears in (6.13) is large compared to elementary particle masses. This suggests a Born-Oppenheimer expansion with respect to $1/M$ in analogy to molecular physics (Banks 1985). The wave function will then adiabatically depend on the massive variables. However, since there is no environment to the universe, the wave function Ψ cannot decohere into narrow wave tubes. This situation resembles *small* molecules (or deformed nuclei), which are usually found in their energy eigenstates rather than in states describing quasi-classical motion — in spite of the validity of the same Born-Oppenheimer approximation as used for large molecules. However, as a *novel* aspect of quantum cosmology, the universe contains its own observers, while a molecule or a nucleus can only be observed from outside. Because of the adiabatic correlation of the observer with the rest of the world, the quantum state of geometry then *appears* classical to him (cf. Sect. 4.6 and see below).

In order to arrive at a semiclassical dynamical description of spacetime geometry,[3] one has to use the *ansatz*

$$\Psi[h_{kl}, \Phi] = \exp[iS_0(h_{kl})]\,\chi(h_{kl}, \Phi) \quad , \tag{6.14}$$

[3] Occasionally the term 'semiclassical quantum gravity' is used to characterize quantized fields on a classical spacetime background. However, this would represent a *new* (probably inconsistent) theory rather than an approximation to quantum gravity.

where again matter is represented by a spatial scalar field Φ. S_0 is defined as a solution of the Hamilton-Jacobi equation of geometro-dynamics (Peres 1962), with a *self-consistent source* (here often called the 'backreaction') derived as an expectation value $\langle\chi|T_{\mu\nu}|\chi\rangle$ from the resulting matter states χ. This is similar to the treatment of classical *orbits* for the atomic nuclei in large molecules, calculated by means of an effective potential obtained from the adiabatic electron state. If the exact boundary conditions are appropriate (and approximations to them appropriately *chosen* for the Hamilton-Jacobi function), the matter wave function χ may indeed depend weakly ('adiabatically') on the massive variables h_{kl} in certain regions of configuration space. As this spatial metric describes the three-geometry as the carrier of information about time along every WKB trajectory, *the dependence on* ${}^{(3)}G(h_{kl})$ *may be regarded as a generalized physical time dependence* of the matter states $\chi(\Phi)$.

Since the Hamilton-Jacobi equation describes an ensemble of dynamically independent trajectories in the configuration space of the three-geometries, the remaining equation for the matter states χ can be integrated along those trajectories. This procedure becomes particularly convenient after the exponential function $\exp(iS_0)$ has been raised to the usual second order WKB approximation (that includes a 'pre-factor' which warrants the conservation of stationary probability currents). In its local form, (6.13), the Wheeler-DeWitt equation can then be written as a *Tomonaga-Schwinger equation* (Lapchinsky and Rubakov 1979, Banks 1985):

$$i\sum_{klk'l'} G_{klk'l'}\frac{\delta S_0}{\delta h_{kl}}\frac{\delta\chi}{\delta h_{k'l'}} = H_{\mathrm{matter}}\chi \quad . \tag{6.15}$$

Its LHS has the form of a functional derivative of χ in the direction of a *local component* of the functional gradient ∇S_0 in the configuration space of three-geometries. Written as $i\nabla S_0 \cdot \nabla\chi =: id\chi/d\tau$, it defines local (many-fingered) physical time τ *for each trajectory* in superspace (hence on each classical spacetime described by S_0). Because of its limited validity it may be called a *WKB time*. In contrast to the original Tomonaga-Schwinger equation, (6.15) is here applied to a broad coherent wave function on superspace; but neither to an individual spacetime nor to a Hamilton-Jacobi *ensemble* of them.

The local equation (6.15) represents a *global* Schrödinger equation for matter with respect to all fingers of time (since τ is a function on space), and for all quasi-classical spacetimes contained in this superposition. Higher orders of the WKB approximation lead to corrections to this derived Schrödinger dynamics (Kiefer and Singh 1991). The local fingers of time may also be combined by integrating (6.15) over three-space (Giulini 1995). This integration elevates the local inner product $\nabla S_0 \cdot \nabla\chi$ in (6.15), that is the sum over $klk'l'$ by means of the Wheeler-DeWitt metric, to a global one. It defines (regardless of any choice of spacetime coordinates) the 'progression' of a *global* dynamical time on all those spacetimes (that is, trajectories in superspace) which are described by the Hamilton-Jacobi function S_0. The definition of coordinates requires lapse and shift functions N and N_k (Sect. 5.4), which would then also lead to formal velocities $\dot{h}_{kl}$ according to

$$\dot{h}_{kl} = -NG_{klk'l'} \frac{\delta S_0}{\delta h_{k'l'}} + N_{k|l} + N_{l|k} \quad . \tag{6.16}$$

It may be convenient to identify the global time τ (a *dynamical* foliation) with the time coordinate resulting from the lapse function N (for vanishing N_i, as done for the metrically meaningful Friedmann time t in (5.19)). Since τ is defined for all WKB-trajectories (thus 'simultaneously' characterizing their corresponding variable three-geometries), it also defines a *foliation of superspace* (Giulini and Kiefer 1994). While spacetime foliations are meaningful only in the WKB approximation, the dynamical foliation of superspace may represent a *general* concept of quantum gravity for Friedmann-type universes.

Essential for the result (6.15) was the *ansatz* (6.14) with a *complex* exponential. This form (with real S_0) will in general not be compatible with the physical boundary conditions, since there is no reason to presume a complex wave function of the universe. This defect of the *ansatz* can be overcome by replacing (6.14) with its real part, $\Psi \rightarrow \Psi + \Psi^*$. The two terms may then *separately* obey (6.15) and its complex conjugate to an excellent approximation, similar to the various WKB trajectories (or wave tubes) which *all together* form the wave front of geometrical optics described by S. This robust separation of Ψ and Ψ^* may be referred to as intrinsic decoherence (Halliwell 1989, Kiefer 1992), although it has always been implicitly presumed when the Born-Oppenheimer approximation was applied to calculating energy eigenstates of small molecules. It may also be regarded as an example of strong intrinsic symmetry breaking (cf. Sect. 6.1, or Sect. 9.6 of Giulini et al. 1996).

If the matter states χ in (6.14) contain observer systems, they have to assume the form of superpositions of product states $\chi^{\mathrm{env}}\chi^{\mathrm{obs}}$, where the environment χ^{env} represents the rest of the matter world. According to (6.15), these states 'evolve' approximately along dynamical foliations of quasi-classical spacetimes. Realistically, the environment as well as the observer must thereby always 'measure' (that is, dynamically depend on) the three-geometry. While the environment leads to the latter's decoherence, and thus to an apparent ensemble of spatial geometries, a definite (subjective) state of the observer selects a *specific* geometry (as indicated by Hawking's remark quoted at the end of Sect. 5.2). These conclusions are based on an appropriate form of *quantum causality* in order to avoid recoherence (4.55).

The Born-Oppenheimer approximation with respect to the Planck mass (only) is not always quite appropriate. Macroscopic matter variables may also be described by a WKB approximation, while certain geometric modes (such as gravitational waves) may not. These variables could then be shifted in (6.14) between S and χ (cf. Vilenkin 1989). For example, Halliwell and Hawking (1985) applied the WKB approximation to the monopoles α and Φ which characterize the quantum Friedmann universe (6.5). The corresponding trajectories $a(t)$, $\Phi(t)$ in mini-superspace are classical approximations to the wave tubes discussed in Sect. 6.2.1, while t is the Friedmann time. Wave functions for the (nonlocal) higher multipole amplitudes of matter and geometry can then be described by a derived Schrödinger equation, similar to (6.15). This choice of *nonlocal* variables offers the advantage of separating physical degrees of freedom (in this case the pure tensor modes) from the remaining pure gauge modes. As mentioned in Sect. 5.4, the physical tensor modes all occur with a dynamical mass of the same sign (here chosen positive) as the matter modes, while only the monopole variable a (or $\alpha = \ln a$) has negative dynamical mass. This result is the basis of the conjecture that the expansion parameter may serve as a dynamical time variable in superspace. In the simple dynamical model proposed by Halliwell and Hawking, the monopoles a and Φ are then very efficiently decohered by the tensor modes (Zeh 1986b, Kiefer 1987) — see also (6.17) below.

Solutions of the Tomonaga-Schwinger equation (6.15) require initial conditions with respect to the global time parameter τ. In order to describe our observed time-asymmetric universe, these initial conditions must include those which are responsible for the arrow(s) of time (such as quantum causality, essential for decoherence). However, these initial conditions cannot be freely chosen any more, since this equation has been *derived* from the Wheeler-DeWitt equation. Therefore, the initial conditions have similarly to be derived from the boundary conditions of the Wheeler-DeWitt wave function. For these boundary conditions, the hyperbolic nature of the Wheeler-DeWitt equation becomes essential. As discussed in Sect. 6.2.1, it means that *intrinsically*-initial conditions may be posed at any fixed value of a time-like variable in superspace (such as α). 'Before' anything may evolve with respect to classical time, classical time must itself emerge from an appropriate boundary condition according to the 'intrinsic dynamics' (6.13).

If the approximately-derived Schrödinger (or Tomonaga-Schwinger) equation along a WKB trajectory in superspace of geometry, as applied to the derived initial condition, describes *measurements*, it is practically useless for calculating 'backwards' in global time τ. One would have to know all (observed and unobserved) final *branches* of the total matter wave function χ (such as those that have unitarily arisen during measurements in the past). In the 'forward' direction, even this global Schrödinger equation for matter has to be replaced with a master equation if decoherence of the quantum state of matter by gravity is phenomenologically taken into account (just as matter decoheres geometry in accordance with the Born-Oppenheimer approximation — cf. Sect. 4.3.4). This provides an explanation in terms of the Wheeler-DeWitt equation for the oft-proposed gravity-induced collapse of the wave function. Most proposals of this kind of collapse do indeed owe their motivation to the decoherence of matter by geometry (while some of them then claim to solve the problem of measurement by once more confusing the genuine and apparent ensembles distinguished in Sect. 4.2).

The asymmetric time evolution of the observed world thus requires a specific (low entropy) structure of the Wheeler-DeWitt wave function (of type (4.56), for example) in a region of superspace where the relevant WKB trajectories have one of their ends (such as at small values of a). This suggests an appropriate boundary condition for Ψ at $\alpha \to -\infty$, in accordance with the hyperbolic nature of the Wheeler-DeWitt equation. If this boundary condition also justifies a WKB approximation according to the *ansatz* (6.14), the Schrödinger equation may be integrated along trajectories in superspace over cosmological times τ. However, because of these rather stringent conditions, the validity of a Schrödinger or master equation for matter and other appropriate variables cannot always and everywhere be taken for granted. It would be mistaken to integrate these equations while losing *dynamically relevant* information in the corresponding direction of calculation.

This limited validity of the Schrödinger or master equation would become particularly important in a recontracting universe (cf. Sect. 5.3). If a master equation can be derived for all or most WKB trajectories with respect to that direction of τ that coincides with increasing a, it cannot remain valid along a classical spacetime history that turns into recontraction. One would then expect effects of an apparent two-time boundary condition, similar to the true two-time boundary conditions discussed by Davies and Twamley (1993) — cf. Fig. 6.3.

'Simplicity' of the initial wave function has often been suggested in the form of a symmetry (e.g. Conradi and Zeh 1991). An example is the false vacuum of the Higgs field (Sect. 6.1). Another model may be defined by means of an appropriate set of variables x_i (such as higher multipoles on the Friedmann sphere) in addition to the time-like variable α, with a Hamiltonian (cf. (6.6))

$$\begin{aligned} H &= \frac{\partial^2}{\partial\alpha^2} + H^2_{\mathrm{red}} \\ &= \frac{\partial^2}{\partial\alpha^2} + \sum_i \left(-\frac{\partial^2}{\partial x_i^2} + V_i(\alpha, x_i) \right) + V_{\mathrm{int}}(\alpha, \{x_i\}) \quad , \end{aligned} \tag{6.17}$$

and $V_{\mathrm{int}} \to 0$ for $\alpha \to -\infty$, for example.[4] This would be compatible with an initial condition (in α) of the kind $\Psi \to \prod_i \chi_i(x_i)$, as conjectured in (4.56) for *local* variables. If the equation $H^2_{\mathrm{red}}\psi_n = \epsilon_n^2 \psi_n$ defines the α-dependent eigenstates $\psi_n(\alpha, \{x_i\})$ of H^2_{red}, solutions to $H\Psi = 0$ can be written in analogy to (6.10) as $\Psi(\alpha, \{x_i\}) = \sum_n c_n(\alpha)\psi_n(\alpha, \{x_i\})$. The eigenstates ψ_n must then adiabatically develop correlations induced by the interaction V_{int}. In

[4] If the large Planck mass is absorbed into the variables of geometry (such as $\alpha \to \sqrt{M}\alpha$ in (6.17) when compared with (6.13)), the Born-Oppenheimer approximation may be based on the resulting weak dependence of the potential on the new variable rather than on an expansion in terms of $1/\sqrt{M}$. However, the validity of the approximation depends invariantly on the actual boundary conditions.

an adiabatic WKB approximation one obtains for $c_n(\alpha)$ the exponentials $\exp(\pm i \int \epsilon_n(\alpha) d\alpha)$. (As explained in Sect. 6.2.1, the sign of the exponent does here *not* distinguish between expansion and contraction.) A similar model has recently been studied by Halliwell (2000).

These simple models are sufficient in principle to describe a recollapsing quantum universe by means of a potential barrier at $\alpha = \alpha_{\max}$, say, (with $\epsilon_n^2(\alpha) < 0$ for $\alpha > \alpha_{max}$) and with a boundary condition of low entropy at $\alpha \to -\infty$ (or at some $\alpha_{\min}$). (For a two-dimensional model see Fukuyama and Morikawa 1989.) The boundary condition at α_{max} selects 'standing waves' $c_n \propto \sin(\int_\alpha^{\alpha_{\max}} \epsilon_n d\alpha)$, which force stationary wave tubes to 'turn around' and continue 'backwards' in α (Zeh 1988). What will then happen to their entropy and the arrow of time?

Formal entropy would trivially depend on α only (regardless of any WKB trajectories or wave tubes) if it were calculated from the complete Wheeler-DeWitt wave function Ψ at fixed α in the form $S(\alpha) = -k\mathrm{Trace}_{x_i}[\hat{P}\varrho \ln(\hat{P}\varrho)]$, with $\varrho = |\Psi\rangle\langle\Psi|$, and, for example, $\hat{P} = \hat{P}_{\mathrm{sep}}$ or $\hat{P}_{\mathrm{local}}$. However, for a quasi-classical universe one has to define entropy as a function of macroscopic variables, $S(\alpha, y_i)$, along trajectories in a *midi*-superspace spanned by $\alpha(\tau)$ and $y_i(\tau)$, with trace over the remaining 'microscopic' or 'irrelevant' variables only. If the macroscopic variables justify a WKB approximation all the way from $\alpha = -\infty$ to $\alpha = -\infty$ (from big bang to big crunch), one may suggest postulating low entropy conditions at $\alpha \to -\infty$ in a certain *region* of this hypersurface of midi-superspace, and assume that the quasi-classical history of our observed universe, approximately following a curve $\alpha(\tau), y_i(\tau)$, had *one* end in this region.

However, neither would the suggestion of such a *specific* region appear reasonably motivated, nor is this description compatible with *quantum* gravity. In realistic models of a quantum Friedmann universe, the wave tubes spread widely and overlap each other for $\alpha \to -\infty$. Boundary conditions can then not be appropriately localized in different regions, and low entropy cannot be expected at but one end of a turning quasi-trajectory. The semi-realistic quantum Friedmann model of Sect. 6.2.1 demonstrates furthermore that quasi-classical spacetimes cannot individually be continued beyond their turning point, but are subject to drastic coherent 'quantum scattering' over all of mini-superspace (cf. Fig. 6.3). This phenomenon would be crucial for the arrow of time, as it is incompatible with an evolving entropy, $S(\tau) := S(\alpha(\tau), y_i(\tau))$, close to the turning point.

Page (1985) and Hawking (1985) arrived at the opposite conclusion regarding the arrow of time in a recontracting universe when insisting on semi-classical or classical approximations and methods (see the discussion at the end of Zeh 1994). This is not surprising, since they applied their specific initial condition to just one end of their Feynman paths in mini-superspace (which become *trajectories* in their subsequent WKB approximation), without consistently relating them to a static Wheeler-DeWitt wave function. Their *ad hoc* choice of a classical time-direction is present already in the

Hartle-Hawking ('no boundary') boundary condition (Hartle and Hawking 1983), without any semiclassical approximation. Feynman path integrals are defined as a tool for calculating the *propagation* of wave functionals with respect to an external time parameter t. Propagation is thereby assumed to proceed in a given direction of this parameter, even though Feynman diagrams may contain formal *particle* lines going back and forth in time. Parametrizable paths in superspace with starting points on a boundary (at the parameter value $t = 0$, say) would be *required* in order to distinguish between their 'first' and 'second' intersection with a time-like hypersurface (such as $a = constant$) in an expanding and recontracting universe. Similar concepts cannot be reasonably defined for an absolutely stationary wave function (see Kiefer and Zeh 1995). In particular, quasi-classical wave tubes can only be justified by an 'irreversible' process of decoherence with respect to an *intrinsic* dynamical parameter, such as α.

A static wave function of quantum mechanics is often *interpreted* in a time-dependent way. This may have its roots either in the time-dependent theory, where moving wave packets may be formed as superpositions, or in the Heisenberg picture (cf. Sect. 4.6), where definite values of 'observables' are assumed to 'occur' irreversibly (see Hartle 1998). This second possibility would require an *external* concept of time (such as that defined by the external observer or classical apparatus in the Copenhagen interpretation).

A remarkable time-less probability interpretation of the wave function in terms of presumed *classical* configurations has been proposed by Barbour (1994a,c). Similar to Bohm's model, Barbour's classical configurations include all (spatial) fields, hence also three-geometries. In contrast to conventional quantum theory, this superspace (of geometry *and* matter) is not only assumed to support the wave function(al), but also an *ensemble* of classical states (which are tacitly assumed to describe states of local observers). This ensemble is postulated to possess 'statistical weights' according to $|\Psi|^2$. However, unlike Bohm's model, this is not an initial condition that has to be dynamically preserved by means of (unobservable) dynamics. There are no 'surrealistic' trajectories any more, since there is no fundamental time in this model. Reality is absolutely time-less in Barbour's theory.

Now, the structure of the Wheeler-DeWitt wave function (arising from the WKB approximation) favors statistical weights for those classical states which lie on *apparent histories*. This result is analogous to Mott's (1929) description of α-particle tracks in a cloud chamber, where the (static) wave function suppresses configurations which represent droplets not approximately lying along particle tracks (even though *many* different quasi-tracks contribute to the whole quantum state). Barbour calls such states 'time capsules', since they possess consistent memories, here without a corresponding history. 'Time is in the instant' (in the state), as he would say — the instant is not in time (in a history). All classical states in the ensemble are regarded as 'real' (similar to the objective reality of all past and future states on a one-dimensional history in the block universe description) — although with different 'statistical weights'. The weights are here postulated to describe probabilities as to where *we* (in the sense of objectivized subjectivity) may find our universe in superspace; they do neither describe a stochastic time evolution nor incomplete knowledge (about one real state). Their sole purpose is to exclude states at small values of $|\Psi|$ from representing our *actual* (observed) universe. This is similar to the Everett interpretation (Sect. 4.6), where states of reality need not be classical configurations, however.

One may characterize this interpretation by saying that the traditional concept of global physical time (a one-dimensional succession of global states) is replaced with the multi-dimensional concept of a (real) wave function. This entirely novel conception of the *wave function as a generalization of time* is remarkable, even though Julian Barbour (1997) rejects it (mainly — I think — because he insists on regarding *classical* general relativity as already timeless).

According to this interpretation, the electron in a specified hydrogen atom, for example, is defined to have a position (though not any velocity or momentum). Since this position is *not* part of a memorized or documented (real or apparent) history, we are led to believe that it 'actually' exists simultaneously at *all* its positions which lie in the support of the wave function. Its effects on measurement devices are in fact dynamically 'caused' by *all* these positions — according to the intrinsic dynamics of the static wave function, but not by a one-dimensional history. Barbour's interpretation would explain why the arena for the wave function is a classical configuration space, while removing the most disturbing aspect of the Bohm model, *viz.*, unobservable classical trajectories. However, other disadvantages persist — for example the fundamental role of just those classical concepts which could be *derived* from wave mechanical ones by means of decoherence in Sect. 4.3. While the presumed classical states would thus provide an answer to the pointer basis problem, this very answer seems to be quite artificial and based on prejudice (traditionalism) — similar to Bohm's choice of classical concepts.

General literature: Kiefer 1994

6.2.3 Black Holes in Quantum Cosmology

During the early days of general relativity, the spacetime region behind a black hole horizon was regarded as meaningless, since it is inaccessible to observers in the external region. From their positivistic point of view, it would 'not exist'. Later one realized that world lines, including those of observers, can smoothly be continued beyond the horizon towards the singularity within finite physical time. The new conclusion, that the internal region of black holes be physically 'normal' (although for a limited future only), remained essentially applicable to Bekenstein-Hawking black holes with their thermodynamical aspects (Sect. 5.1). However, arguments indicating a genuine (possibly dramatic) quantum nature of the event horizon have also been raised ('t Hooft 1990, Keski-Vakkuri et al. 1995, Li and Gott 1998).

While a consistent quantum description of black holes has not yet been given (for attempts see Kiefer and Louko 1998, or Vaz and Witten 1998), semiclassical black holes have been discussed by means of available concepts (Kiefer 1997). When combined with quantum cosmology, they lead to important novel consequences, which seem to revive the initial doubts about the meaning and existence of black hole interiors.

Consider the Schwarzschild metric (Fig. 5.1) as far as it is relevant for a black hole formed by collapsing matter (Fig. 5.3a), where the Kruskal regions III and IV do not occur. Its *dynamical* (3+1) description in terms of three-geometries depends on the choice of a foliation (cf. the Oppenheimer-Snyder model described in Box 32.1 of Misner, Thorne and Wheeler 1973). Three-geometries which intersect the event horizon may spatially extend to the singularity at $r = 0$, and thus render the global quantum states that they carry prone to consequences of a yet elusive unified field theory. However, a foliation according to Schwarzschild time t would describe regular three-geometries for $t < \infty$, and they could be continued *in time* beyond $t = \infty$ by means of the new time coordinate r (with physical time growing with decreasing r from $r = 2M$ to $r = 0$). The black hole interior with its singularity would always remain in our formal future according to this regular foliation. Therefore, the latter seems to be appropriate for the formulation of a cosmological boundary condition (in superspace), that may determine an arrow of time.

For a discussion assume again that the expansion of this classical universe is reversed at a finite Schwarzschild time $t = t_{\mathrm{turn}}$ much bigger than the formation time of the black hole (Fig. 6.4). No horizon exists on this Schwarzschild simultaneity $t = t_{\mathrm{turn}}$. If the cosmic time arrow does change direction (while the universe passes through an era of thermodynamical indefiniteness), the gravitationally collapsing matter close to the expected horizon will very soon (in terms of its own proper time) enter the era where radiation is advanced in the sense of Chap. 2. The black hole can then no longer 'lose hair' by emitting retarded radiation: it must instead 'grow hair' in an anti-causal manner (Fig. 6.5). According to a 'time-reversed no-hair theorem' it has to re-expand when the universe starts recontracting (Zeh 1994, Kiefer and Zeh 1995).

This scenario does not contradict the geometro-dynamical theorems about a monotonic growth of black hole areas, since no horizons ever form. A classical spacetime will not even exist close to the 'turning of the tide'. Here decoherence is replaced with recoherence. Only region I of Fig. 5.1 then applies. However, events which appear 'later' than t_{turn} in the classical picture are 'earlier' in the sense of the fundamental dynamics of the Wheeler-DeWitt equation (6.6) — and therefore also in the thermodynamical sense. This quantum cosmological model describes an *apparent* (quasi-classical) two-time Weyl

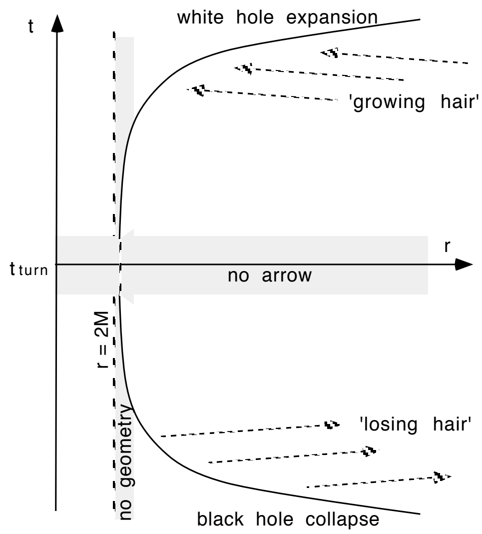


Fig. 6.4. Classical trajectory of a collapsing dust shell (or the surface of a collapsing star) in a thermodynamically symmetric universe, represented in compressed Schwarzschild coordinates (with Schwarzschild metric valid only outside the shell). Due to scale compression, light rays appear almost horizontally in the graphic. For $t > t_{\rm turn}$, advanced radiation from the (then) formal past (in our future) would focus onto the black hole, and 'cause' it to re-expand and grow hair in this scenario, while contemporary observers would experience time in the opposite direction. No horizon ever forms. Because of the drastic quantum effects close to the turning point of a Friedmann universe (cf. Fig. 6.3), there will in general be no more than a 'probabilistic' connection between quasi-classical expansion and contraction eras of the universe. (From Kiefer and Zeh 1995)

tensor (or similar) condition (cf. Fig. 6.5). In quantized general relativity, the two apparently different boundaries are identical, and thus represent *one and the same* condition. The problem of their consistency (Sect. 5.4) is considerably relaxed, since it is reduced to the 'final' condition of normalizability for $a \to \infty$.

The description used so far in this section does not directly apply to a forever-expanding universe, where the arrow would preserve its direction along a complete quasi-trajectory from $a = 0$ to $a = +\infty$. The Wheeler-DeWitt wave function is then *not* normalizable for $a \to \infty$. However, as a symmetric generalization of the Weyl tensor hypothesis, one may require this wave function to vanish on *all* somewhere-singular three-geometries. This would again lead to important thermodynamical and quantum effects close to event horizons (Zeh 1983), and drastically affect the possibility of continuing a quasi-classical spacetime beyond them. These consequences would be unobservable in practice by external observers, since the immediate vicinity of a future horizon remains outside their backward light cones for any finite future. In order to receive information from the vicinity of a future horizon,

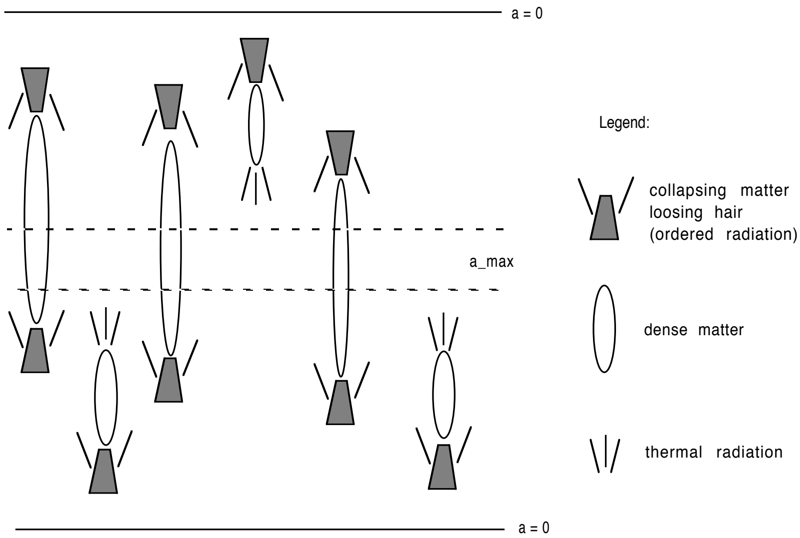


Fig. 6.5. Quasi-classical picture of a thermodynamically T-symmetric quantum universe with 'black-and-white holes' that re-expand by anti-causal effects. Instead of horizons and singularities there are merely spacetime regions of large curvature in this scenario. They may serve as a short-cut in proper time between big bang and big crunch (or between the presumed eras of opposite arrows of time). 'Information-gaining systems' could not survive as such. In quantum cosmology there is no unique connection between quasi-classical histories (Everett branches) represented by the upper and lower halves of the Figure

one has to come dangerously close to it (see Fig. 5.2, where the light cone structure is made evident).

These conclusions do indeed seem to throw serious doubts on the validity of a classical continuation of spacetime into black hole interiors. Event horizons in classical general relativity may signal the presence of drastic thermodynamical effects rather than representing 'physically normal' regions of spacetime. Even though their observable consequences depend on the world lines (accelerations, in particular) of the required detectors (Sect. 5.2), global quantum states, such as a specific 'vacuum', are invariantly defined (though not invariantly observed). These global states may define an objective arrow of time, including 'quantum causality', by means of a fundamental boundary condition for the Wheeler-DeWitt wave function.

Epilog

– Prelim. 4th edtn. (Aug 00): www.time-direction.de

Four weeks before his death, Albert Einstein wrote in a letter of condolence to the family of his life-long friend Michael Besso (Dukas and Hoffman 1979):[1] "For us believing physicists, the division into past, present and future has merely the meaning of an albeit obstinate illusion." There is no doubt that Einstein meant this remark seriously. Evidently, it refers to the four-dimensional ('static') spacetime picture of a 'block universe' that his theory of relativity uses so efficiently. This picture seems to be at variance with the experience of a *present passing through time* (the 'flow' or 'passage of time'). In contrast, the objective (classical) spacetime framework contains only a concept of *local events*, which may be regarded as a set of dynamically related *here-and-nows*. Because of these empirical dynamical relations, characterized by time-symmetric *laws*, a *local* present can be viewed as formally 'moving' along the world line of an observer (a succession of appropriate local groups of events), dynamically controlled by proper time, while the *global* dynamical state depends on an arbitrarily chosen foliation of the objective and invariant spacetime (Sect. 5.4).

In Hermann Weyl's words: "The objective world simply *is*; it does not happen. Only to the gaze of my consciousness, crawling upward along the life line of my body, does a section of this world come to life as a fleeting image in space that continuously changes in time." Any objectivization of subjective empirical evidence can thus not be restricted to space, as it used to be in the pre-relativistic concept of an *objective global present*. For this reason, Price (1996) chose the subtitle 'A View from Nowhen' to his book on time's arrow. However, whether the objective world 'simply is', or rather 'comes into being', seems to be a pure matter of words (or description). Weyl's 'is' does not exclude *dynamical* time. The four- (or more-) dimensional 'static' *view* is by no means specific to the theory of relativity. It does not even require deterministic laws, as it was already referred to by St. Augustine in his *Confessiones*. He regarded it as a divine view of the world — presumably as he understood it as a *complete* description, not merely as a conceptual framework. A mortal physicist may at least conceive of the future history of the world (though with less confidence than about the past). Even Laplace could not expect his model world as being *practically* determinable (cf. Sect. 3.3);

[1] "*Für uns gläubige Physiker hat die Scheidung zwischen Vergangenheit, Gegenwart und Zukunft nur die Bedeutung einer wenn auch hartnäckigen Illusion.*"

he had to assume an extra-physical demon for this purpose. The argument that the future cannot be exactly known to physical systems, such as humans or computers, should not be confused with an indeterminism of the empirically founded dynamical laws (as often done in the popular theory of chaos — cf. Bricmont 1996).

The peculiarity of the subjective present, often mistakenly regarded as part of an objective 'structure of time', was emphasized by Einstein in a conversation with Carnap. According to Carnap (1963), "Einstein said that the problem of the Now worried him seriously. He explained that the experience of the Now means something special for man, something essentially different from the past and the future, but that this important difference does not and cannot occur within physics. That this experience cannot be grasped by science seemed to him a matter of painful but inevitable resignation." So he concluded "that there is something essential about the Now which is just outside the realm of science."

Carnap emphasized, however, that Einstein agreed with him (in contrast to Bergson and other philosophers) that this situation does *not* indicate a *defect of the physical concept of time*. (The non-confirmation of a prejudice is easily viewed as a defect.) It should rather be understood as reflecting the missing definition of a psycho-physical parallelism that would then have to take into account the fundamental and underivable empirical here-and-now character of subjective awareness, which forms the basis of any observation of reality.

The 'divine world picture' of a block universe, such as that based on relativistic spacetime, does not contain any concept of a present, and can therefore not describe any *flow of time* that we (seem to) experience as we 'live in time', that is, in the local instant. Does this subjective (and precisely therefore real) experience now fundamentally *induce* a time asymmetry that would otherwise be absent from the observed world? Einstein's above-quoted remark is indeed often misinterpreted for supporting an arrow of time that is an illusion, based on the subjective observer who 'moves' in a certain direction of time. It is, therefore, frequently mentioned by the information-theoretical school of statistical mechanics (Sect. 3.3.1), or in favor of an *extra-physical* concept of (growing) 'human knowledge' (Heisenberg's 'idealistic' interpretation of the wave function). The increase of entropy would then only be meaningful with respect to information-gaining human observers whose extra-physical growth of knowledge would determine in which direction of time statistical reasoning has to be applied.

In various chapters of the book I have tried to explain that the concept of entropy, which is crucial for most time asymmetric phenomena, is indeed observer-related by means of a *relevance concept* on that it has to be based. This observer-relatedness of the macroscopic description of our physical world is particularly important in quantum mechanics, and — as it turned out — even for the emergence of a classical concept of time from a timeless quantum world. However, the time-asymmetry of the phenomena could always be traced back (at least in principle) to an asymmetric structure of physical reality, and perhaps ultimately to that of the wave function of the universe. The concept of an objective physical reality could thereby be upheld as a heuristic fiction [2] (that is all it can ever mean) — for example by means of a collapse of the wave function, or in the Everett interpretation (cf. Sect. 4.6).

Memories must then also be asymmetric (in relation to their sources) because of the asymmetric structure of the physical world. This may be sufficient to explain the apparent flow of time once there is a psycho-physical parallelism based on the *moment of awareness*. Only this (not necessarily asymmetric) concept of a local present is fundamentally subjective — not the asymmetry between past and future. What we usually call the *identity* of a person (who changes considerably during his lifetime) is 'in reality' nothing but a particularly strong causal (fact-like asymmetric) connection between *different* local physical states which represent the individual carriers of a subjective present. The *here-and-now* subjectivity as the center of all awareness is a concept that goes beyond objective reality, although it must severely affect our perception of the 'real world'.

These latter remarks are essentially based on a classical description. As is well known at least since von Neumann's (1932) book, the formulation of a psycho-physical parallelism based on instantaneous states of local systems encounters severe problems in quantum theory because of the latter's kinematical nonlocality. Generically, there are no local physical states any more. I have therefore tried to explain in Sects. 4.6 and 6.2.2 that the enforced reformulation of this parallelism (or epi-phenomenalism) in terms of the nonlocal wave function of the universe does in my opinion offer the chance of simultaneously solving the quantum measurement problem and justifying the (objectivizable) observer-relatedness of entropy. This universal quantum theory (based on the Everett interpretation) describes the empirical quantum indeterminism in a deterministic or even time-less quantum world by means of *observers with an indeterministic identity*. Within the conceptual framework of quantum theory, this unconventional interpretation seems to be enforced by the profound contrast between *quantum nonlocality* (non-trivial 'wholeness') and the also empirically founded *locality of the observer* in spacetime.

The canonical quantization of general relativity (or of any unified field theories containing it) led in Sect. 6.2 to the problem of wave functionals (su-

---

[2] The notion of a 'fiction' thus does here not contradict that of reality. Instead, it emphasizes the impossibility of *proving* a certain reality. For example, Einstein's *gläubiger Physiker* (believing physicist) may express the *belief* in an objective reality (preferentially a local one) to be described by the physical formalism. As an admirer of Hume (but not of Kant — see Franck 1949) Einstein is obviously aware of this hypothetical character of reality. The evidence that nature appears comprehensible was regarded by Einstein as "the most incomprehensible thing about nature". One should here recall that Descartes and Hume raised their doubts and criticism, which essentially forces us into fictionalism, against any attempts of obtaining absolute *certainty* in the empirical sciences — not against any more or less *hidden* reality — see d'Espagnat (1995).

perpositions) of fields defined on *global* three-geometries (global simultaneities). They do *not* form a one-dimensional succession (a 'history'). Just as nonlocal quantum mechanical states in general, these global quantum states define states of local subsystems only in certain components (branches). While they must be defined on the space of states for all foliations of all possible spacetimes, the local density matrix of the observer system should be invariant under re-foliations which preserve its local inertial spacetime frame. This situation may perhaps suggest a reformulation (or modification) of quantum theory in a way that does not require the concept of global states (on artificial simultaneities). However, the fact that entangled global superpositions can and do lead to observable effects (such as in EPR/Bell type experiments and decoherence) must have profound consequences on any possible future theory.

The combination of our two most general and successful theories also seems to indicate that the concept of a black hole interior becomes doubtful (Sect. 6.3). Although we cannot perform experiments to confirm this conclusion, we may be able to find out what we *would* observe according to our theory in such extreme situations. Nonetheless, the question why the beautiful conceptual spacetime symmetry of Maxwell's and Einstein's theories is of limited validity represents the greatest mystery of present day theoretical physics. It is, therefore, not too surprising — but utterly discomforting — that some relativists and modern field theorists seem to be adopting forms and interpretations of quantum theory that begin to deviate from those which had to be invented by low energy physicists on empirical grounds.

Our physical world models (or heuristic fictions) must, of course, become more and more speculative with increasing extrapolation away from those realms of knowledge that are 'directly' accessible to our senses and observing instruments. Since their novel conceptions seem to be essential for explaining the observed 'direction of time', we have to trust in the steadily growing wealth of experience. This (apparent?) historical process, that is itself based on the phenomenological arrow of time, has often revealed the limited validity of our 'everyday concepts' that were derived from our direct sense impressions. While this situation emphasizes the provisional status of our knowledge, it should also motivate us to base cosmological considerations on concepts of presently widest applicability (although not on *merely* speculative ones) — concepts which therefore have to be 'abstract'. Physical models have to be tried to their intrinsically defined limits if they are to be regarded as candidates for a description of reality.

Appendix: A Simple Numerical Toy Model

Various elementary models, intended to illustrate statistical methods in physics, can be found in the literature. Best known may be the *urn model* (Ehrenfest and Ehrenfest 1911) and the *ring model* (Kac 1959). Both are based on stochastic dynamics, applied in a given time direction. The model discussed below is deterministic, with specific assumptions applied to initial conditions. Its main intention is to illustrate the concept of a Zwanzig projection with the simple example of coarse-graining in space. This is used for the definition of entropy *and* in a master equation. However, it fails to represent a good approximation in the master equation, thereby demonstrating the importance of *dynamical* requirements for a useful Zwanzig projection.

The model is defined by a swarm of n_P free particles ($i = 1, \ldots, n_P$), moving with fixed velocities $v_i = v_0 + \Delta v_i$ on a 'ring' (a periodic interval) with a length of n_S 'units cells' ($j = 1, \ldots, n_S$), whereby the Δv_i's form a homogeneous random distribution within boundaries $\pm\Delta v/2$. Coarse-graining is defined over each cell. All particles are assumed to start in cell $j = 1$.

Our Zwanzig concept of relevance is also assumed to imply the identification of all (distinguishable) particles, thus leading to an 'occupation number representation on cells'. Since the momenta are assumed constant, only the spatial distribution of particles has to be considered. If n_j particles are in the j-th cell, the entropy according to this Zwanzig projection is given by

$$S = -\sum_{j=1}^{n_S} \frac{n_j}{n_P} \ln \frac{n_j}{n_P} \quad . \tag{A.1}$$

It vanishes initially under the condition stated above, while its maximum, $S^{\max} = \ln n_S$, holds for equipartition, $n_j = n_P/n_S$. With respect to this coarse-grained entropy, the model has a statistical Poincaré recurrence time $t_{\text{Poincaré}} = n_S^{n_P - 1}$, while the relaxation time, required for the approach to equipartition, is of the order $n_S/\Delta v$. It is very short, since this coarse-graining is not robust. Other equilibrium distributions could easily be produced by using intervals of different lengths.

The first plot of the *Mathematica* notebook [1] below (plot1) shows the entropy evolution during the first 2000 units of time in the case of a relaxation time scale of about 1000 units (with $n_P = 100$ and $n_S = 20$). Only integer

[1] Programming help by E. Joos is acknowledged.

numbers of time units are used in the plot in order to eliminate disturbing lattice effects. At some later time (plot2), the coarse-grained representation does not reveal any information about the existence of a low-entropy state in the recent past any more (although this information must still exist, since the motion could in principle be reversed).

One can now simply enforce a two-time boundary condition (Sect. 5.3) by restricting all velocity variables Δv_i to integer multiples of n_S divided by a large integer (e. g., by rounding them off at a certain figure). The changes induced thereby are negligible for medium times, although the evolution is now exactly periodic over a large time that is nonetheless a very small fraction of the statistical Poincaré recurrence time. *Relative* entropy minima may occur at rational fractions of this period (such as 2/5, considered in plot3).

If the coarse-graining is applied dynamically, one obtains a master equation, valid for *mean* occupation numbers $\bar{n}_j(t)$ in an ensemble of individual solutions. For $v_0 = 0$ it would read (cf. (4.42))

$$\frac{d\bar{n}_j}{dt} = \lambda(\bar{n}_{j-1} + \bar{n}_{j+1} - 2\bar{n}_j) \quad , \tag{A.2}$$

where λ (in this case given by $\Delta v/8$) is the mean rate for a particle to move by one cell in one or the other direction. The same result is obtained for $v_0 \neq 0$ in the center of mass frame, that is, when the mean motion (according to v_0) is treated exactly. (Otherwise, Galileo invariance would be lost.) The lower curve of plot4 shows the resulting monotonic increase of ensemble entropy, compared with the individual solution of plot1.

Evidently, the two curves agree only for the first $2/\Delta v$ units of time. This demonstrates that this Zwanzig projection is not very appropriate for *dynamical* purposes, since some fine-grained information remains dynamically relevant. Here, the *individually* conserved velocities lead to relevant correlations between position and velocity. The master equation, which does not distinguish between individual particles, allows them even to change their direction of motion. A slightly improved long-time approximation can therefore be obtained by assuming all particles to diffuse in the same direction — equivalent to using a reference frame that moves with a velocity $-\Delta v/2$ in the center-of-mass frame (upper curve of plot4). However, this procedure still neglects dynamically relevant correlations representing the fact that the most distant particles continue travelling fastest. These correlations remain dynamically relevant until most particles have travelled the whole interval (the relaxation time), as can be recognized in plot4.

Stochastic *individual* histories $n_j(t)$ representing the master equation (A.2), that describes an ensemble, can be produced from a Langevin equation (dynamically using a random number generator). Their histories are *not determined* by initial states. Similar arguments apply to a quantum mechanical Itô process (Sect. 4.4), which can be used to describe an *apparent ensemble* that would represent an open system in a closed universe even in the deterministic Everett interpretation.

Notebook: A Toy Model

1. Definitions

```
nS = 20;
nP = 100;
v0 = 1.;
Δv = .02;

shannon[x_] := If[x == 0, 0., N[x Log[x]]]

entropMax = Log[nS] // N

2.995732274

poincareTime = nS ^ (nP - 1) // N

6.338253001 × 10^128

relaxTime = nS / Δv

1000.
```

2. Initial Values

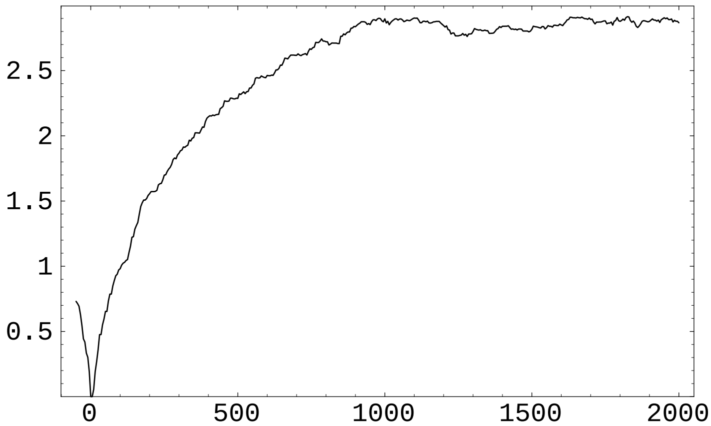
```
SeedRandom[2]

v = Table[v0 + Δv (Random[] - .5), {nP}];

x0 = Table[Random[], { nP}];

<< "Statistics`DescriptiveStatistics`"

{Mean[x0], Mean[v], Variance[v]}

{0.5156437577, 0.9997682434, 0.00003356193642}
```

3. Exact Model

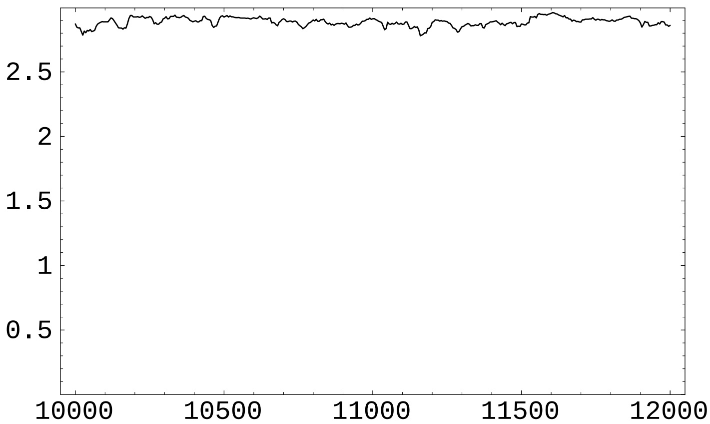
```
entropy[t_] := Module[{cellnb, n, e},
  (cellnb = Ceiling[Mod[v t + x0, nS]];
  Do[n[i] = Count[cellnb, i], {i, nS}];
  e = N[Log[nP] - (Sum[shannon[n[j]], {j, 1, nS}])/nP])]
```

$$e = N\left[\mathrm{Log}[nP] - \frac{\sum_{j=1}^{nS} \mathrm{shannon}[n[j]]}{nP}\right]$$

```
plot1 = ListPlot[Table[{t, entropy[t]}, {t, -50, 2000, 5}],
  PlotJoined → True, Frame → True, PlotRange → {0, entropMax}];
```

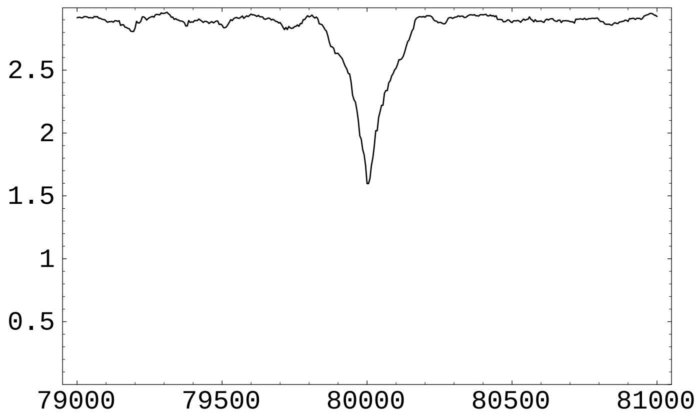



```
plot2 = ListPlot[Table[{t, entropy[t]}, {t, 10000, 12000, 5}],
  PlotJoined → True, Frame → True, PlotRange → {0, entropMax}];
```



4. Two-Time Boundary Condition

```
v = Floor[10000 v] / 10000 ;

period = 10000 nS

200000

plot3 = ListPlot[Table[{t, entropy[t]}, {t, 79000, 81000, 5}],
  PlotJoined → True, Frame → True, PlotRange → {0, entropMax}];
```



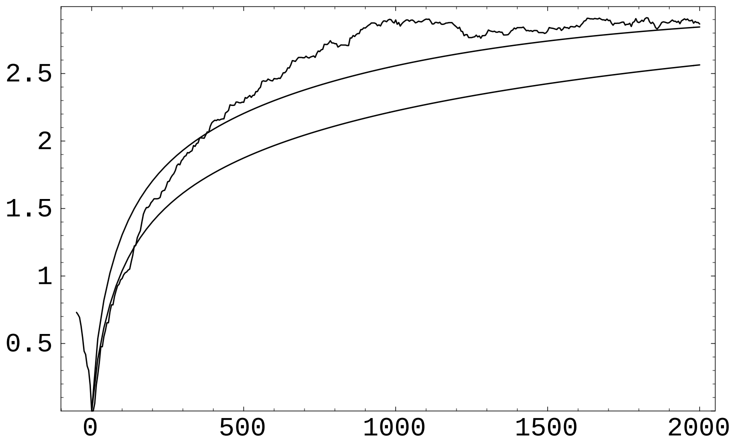
5. Master Equation

```
λ = Δv / 8

0.0025

equList = Table[n[j]'[t] == λ (-2 n[j][t]
      + n[Mod[j, nS] + 1][t] + n[Mod[j - 2, nS] + 1][t]), {j, nS} ];

iniList = Join[ {n[1][0] == nP}, Table[n[j][0] == 0., {j, 2, nS}] ];
```

```
nSolution = NDSolve[Join[equList, iniList], Table[n[j],
   {j, nS}], {t, 0, 2000} ];
```

$$\texttt{entropyMaster[t\_] := N[Log[nP]] - }\frac{\sum_{j=1}^{nS}\texttt{shannon[n[j][t]]}}{\texttt{nP}}$$

```
 /. nSolution  // First

plotMasterShort = ListPlot[Table[{x, entropyMaster[x] },
   {x, 0, 2000, 20}] , PlotJoined -> True, PlotRange ->
   {0, entropMax}, DisplayFunction -> Identity];

equList  = Table[n[j]'[t] == 4 λ (-n[j][t]
      + n[Mod[j - 2, nS] + 1][t]), {j, nS} ];


nSolution = NDSolve[Join[equList, iniList], Table[n[j],
   {j, nS}], {t, 0, 2000} ];

plotMasterLong = ListPlot[Table[{x, entropyMaster[x] },
   {x, 0, 2000, 20}] , PlotJoined -> True, PlotRange ->
   {0, entropMax}, DisplayFunction -> Identity];

plot4 = Show[plot1, plotMasterShort, plotMasterLong];
```



References

Aharonov, Y., Albert, D., Casher, A., and Vaidman, L. (1987): Surprising quantum effects. Phys. Lett. **A124**, 199 – [129]

Aharonov, Y., Bergmann, P.G., and Lebowitz, J.L. (1964): Time Symmetry in the Quantum Process of Measurement. Phys. Rev. **134**, B1410 – **(WZ)** – [128]

Aharonov, Y., and Bohm, D. (1961): Time in the Quantum Theory and the Uncertainty Relation for Time and Energy. Phys. Rev. **122**, 1649 – **(WZ)** – [176]

Aharonov, Y., and Vaidman, L. (1991): Complete description of a quantum system at a given time. J. Phys. **A24**, 2315 – [170]

Aharonov, Y., and Vaidman, L. (1996): "About position measurements which do not show the Bohmian particle positions." In *Bohmian Mechanics and Quantum Theory: An Appraisal,* ed. by Cushing, J.T., Fine, A., and Goldstein, S. (Kluwer) – [129]

Albrecht, A., (1992): Investigating decoherence in a simple system. Phys. Rev. **D46**, 5504 – [131]

Albrecht, A. (1993): Following a "Collapsing" Wave Function. Phys. Rev. **D48**, 3768 – [131]

Alicki, R., and Lendi, K. (1987): *Quantum Dynamical Semigroups and Applications* (Springer) – [116]

Anglin, J.R., and Zurek, W.H. (1996): A Precision Test of Decoherence. (e-print quant-ph/9611049) – [106]

Arndt, M., Nairz, O., Vos-Andreae, J., Keler, C., van der Zouw, G., and Zeilinger, A. (1999): Wave-particle duality of C_{60} molecules. Nature **401**, 680 – [101]

Arnol'd, V.I., and Avez, A. (1968): *Ergodic Problems of Classical Mechanics* (Benjamin) – [52]

Arnowitt, R., Deser, S., and Misner, C.W. (1962): "The dynamics of general relativity." In *Gravitation: An Introduction to Current Research*, ed. by Witten, L. (Wiley) – [159]

References marked with '**(WZ)**' can also be found (in English translation if applicable) in Wheeler and Zurek 1983, those with '**(LR)**' in Leff and Rex 1990. Some reprints are available through www.zeh-hd.de. Pages of quotation are added in square brackets.

Ashtekhar, A. (1987): New Hamiltonian formulation of general relativity. Phys. Rev. **D36**, 1587 – [175]

Ashtekhar, A. (1991): *Lectures on non-perturbative canonical quantum gravity* (World Scientific) – [175]

Baierlein, R.F., Sharp, D.H., and Wheeler, J.A. (1962): Three-Dimensional Geometry as Carrier of Information about Time. Phys. Rev. **126**, 1864 – [161]

Balian, R. (1991): *From Microphysics to Macrophysics* (Springer) – [65]

Banks, T. (1985): TCP, Quantum Gravity, the Cosmological Constant, and all that ... . Nucl. Physics **B249**, 332 – [182,186,187]

Barbour, J.B. (1986): "Leibnizian time, Machian dynamics and quantum gravity." In *Quantum Concepts in Space and Time*, ed. by Penrose, R., and Isham, C.J. (Clarendon Press) – [10]

Barbour, J.B. (1989): *Absolute or Relative Motion?* (Cambridge University Press) – [14]

Barbour, J.B. (1994a): "The Emergence of Time and its Arrow from Timelessness." In *Physical Origins of the Asymmetry of Time*, ed. by Halliwell, J.J., Perez-Mercader, J., and Zurek, W.H. (Cambridge University Press) – [10,81,163,192]

Barbour, J.B. (1994b): The Timelessness of Quantum Gravity: I. The Evidence from the Classical Theory. Class. Quantum Grav. **11**, 2853 – [163]

Barbour, J.B. (1994c): The Timelessness of Quantum Gravity: II. The Appearance of Dynamics in Static Configurations. Class. Quantum Grav. **11**, 2875 – [192]

Barbour, J.B. (1997): private communication – [193]

Barbour, J.B. (1999): *The End of Time* (Weidenfeld & Nicolson) – [163]

Barbour, J.B., and Bertotti, B. (1982): Mach's Principle and the Structure of Dynamical Theories. Proc. R. Soc. London **A382**, 295 – [163]

Barbour, J.B., and Pfister, H., edts. (1995): *Mach's Principle: From Newton's Bucket to Quantum Gravity* (Birkhäuser, Basel) – [11,161]

Bardeen, J.M., Carter, B., and Hawking, S.W. (1973): The Four Laws of Black Hole Mechanics. Comm. Math. Phys. **31**, 161 – [143]

Bardeen, J., Cooper, L.N., and Schrieffer, J.R. (1957): Theory of superconductivity. Phys. Rev. **108**, 1175 – [149]

Barrabès, C., Brady, P.R., and Poisson, E. (1993): Death of white holes. Phys. Rev. **D47**, 2383 – [141]

Barrow, J.D., and Tipler, F.J. (1986): *The Anthropic Cosmological Principle* (Oxford University Press) – [155]

Bartnik, R., and Fodor, G. (1993): On the restricted validity of the thin sandwich conjecture. Phys. Rev. **D48**, 3596 – [160]

Bekenstein, J.D. (1973): Black Holes and Entropy. Phys. Rev. **D7**, 2333 – [143]

Bekenstein, J.D. (1980): Black-hole Thermodynamics. Physics Today **33**, 24 – [146]

Belavkin, V.P. (1988): “Nondemolition measurements, nonlinear filtering and dynamic programming of quantum stochastic processes.” In *Modeling and Control of Systems*, ed. by Blanquière, A. (Springer) – [114]

Bell, J.S. (1964): On the Einstein Podolsky Rosen paradox. Physics **1**, 195 – **(WZ)** – [93]

Bell, J.S. (1981): “Quantum mechanics for cosmologists”, In *Quantum Gravity II*, ed. by Isham, C.J., Penrose, R., and Sciama, D.W. (Clarendon Press) – [123]

Beller, M. (1996): The Conceptual and the Anecdotal History of Quantum Mechanics. Found. Phys. **26**, 545 – [122]

Beller, M. (1999): *Quantum Dialogue* (University of Chicago Press) – [122]

Bennett, C.H. (1973): Logical reversibility of computation. IBM J. Res. Dev. **17**, 525 – **(LR)** – [72,74]

Bennett, C.H. (1987): Demons, Engines, and the Second Law. Sci. Amer. **257**, 108 – [74,75]

Bennett, C.H., and Landauer, R. (1985): The fundamental physical limits of computation. Sci. Amer. **253**, 38 – [74]

Bertalanffi, L. von (1953): *Biophysik des Fließgleichgewichts* (Vieweg) – [75]

Birrell, N.D., and Davies, P.C.W. (1983): *Quantum fields in curved space* (Cambridge University Press) – [151]

Bläsi, B., and Hardy, L. (1995): Reality and time symmetry in quantum mechanics. Phys. Lett. **A207**, 119 – [128]

Blyth, W.F., and Isham, C.J. (1975): Quantization of a Friedmann universe filled with a scalar field. Phys. Rev. **D11**, 768 – [177]

Böhm, A. (1978): *The Rigged Hilbert Spaces and Quantum Mechanics* (Springer) – [116]

Bohm, A., Doebner, H.-D., and Kielanowski, P. (1998): *Irreversibility and Causality* (Springer) – [120]

Bohm, D. (1952): A suggested Interpretation of the Quantum Theory in Terms of “Hidden” Variables. Phys.Rev. **85**, 166; dto. **85**, 180 – [122]

Bohm, D., and Bub, J. (1966): A proposed solution of the measurement problem in quantum mechanics by a hidden variable theory. Rev. Mod. Phys. **38**, 453 – [114]

Bohm, D., and Hiley, B. (1993): *The Undivided Universe: an Ontological Interpretation of Quantum Mechanics* (Routledge and Kegan) – [122]

Bohr, N. (1928): The Quantum Postulate and the Recent Development of Atomic Theory. Nature **121**, 580 – **(WZ)** – [110]

Boltzmann, L. (1866): Über die mechanische Bedeutung des zweiten Hauptsatzes der Wärmetheorie. Wien. Ber. **53**, 195 – [41]

Boltzmann, L. (1896): *Vorlesungen über Gastheorie* (Barth, Leipzig); Engl. transl. by Brush, S.G. (Univ. Cal. Press, 1964) – [41,154]

Bondi, H. (1961): *Cosmology* (Cambridge University Press) – [152]

Borel, E. (1924): *Le hasard* (Alcan, Paris) – [53]

Born, M. (1926): Das Adiabatenprinzip in der Quantenmechanik. Z. Physik **40**, 167 – [89]

Born, M. (1955): Continuity, Determinism, and Reality. Kgl. Danske Videnskab. Mat. Fys. Medd. **30**, 2 – [71]

Borzeskowski, H.-H. von, and Treder, H.-J. (1988): *The Meaning of Quantum Gravity* (Reidel) – [174]

Boulware, D.G. (1980): Radiation from a Uniformly Accelerated Charge. Ann. Phys. (N.Y.) **124**, 169 – [29,32,148]

Brady, P.R., Moss, I.G., and Myers, R.C. (1998): Cosmic Censorship: As Strong As Ever. Phys. Rev. Lett. **80**, 3432 – [140]

Bricmont, J. (1996): "Science of chaos or chaos in science?" in *The Flight from Science and Reason*, ed. by Gross, P.R., Levitt, N., and Levis, M.W. (New York Acad. Sci., Vol. 755) – [70,198]

Brillouin, L. (1962): *Science and Information Theory* (Academic Press) – [71,74,158]

Bronstein, M., and Landau, L. (1933): Über den zweiten Wärmesatz und die Zusammenhangsverhältnisse der Welt im Großen. Sowj. Physik **4**, 114 – [81]

Brown, J.D., and Kuchař, K.V. (1995): Dust as a standard of space and time. Phys. Rev. **D51**, 5600 – [160]

Brun, T.A. (2000): Continuous measurement, quantum trajectories, and decoherent histories. Phys. Rev. **A61**, 042107 – [124]

Brune, M., Hagley, Dreyer, J., Maître, X., Maali, Y., Wunderlich, C., Raimond, J.M., and Haroche, S. (1996): Observing the Progressive Decoherence of the "Meter" in a Quantum Measurement. Phys. Rev. Lett. **77**, 4887 – [108]

Brustein, R., and Veneziano, G. (1994): The graceful exit problem in string cosmology. Phys. Lett. **324**, 429 – [166]

Busch, P., Lahti, P.J., and Mittelstaedt, P. (1991): *The Quantum Theory of Measurement* (Springer) – [132]

Butterfield, J., and Isham, C.J. (1999): On the Emergence of Time in Quantum Gravity. (e-print gr-qc/9901024) – [8]

Caldeira, A.O., and Leggett, A.J. (1983): Path Integral Approach to Quantum Brownian Motion. Physica **121A**, 587 – [102]

Caldirola P. (1956): A new model of the classical electron. Nuovo Cim. **3**, Suppl. 2, 297 – [31]

Calzetta, E., and Hu, B.L. (1995): Quantum fluctuations, decoherence of the mean field, and structure formation of the early Universe. Phys. Rev. **D52**, 6770 – [154,173]

Carnap, R. (1963): "Autobiography", In *The Philosophy of Rudolf Carnap*, ed. by Schilpp, P.A. (Library of Living Philosophers) p. 37 – [198]

Carr, B.J., and Rees, M.J. (1979): The anthropic principle and the structure of the physical world. Nature **278**, 605 – [155]

Casati, G., Chirikov, B.V., and Zhirov, O.V. (2000): How Many "Arrows of Time" Do We Really Need to Comprehend Statistical Laws? Phys. Rev. Lett. **85**, 896 – [54]

Casper, B.B., and Freier, S. (1973): 'Gibbs Paradox' Paradox. Am. J. Phys. **41**, 509 – [69]

Cassirer, E. (1977): *Zur modernen Physik* (Wiss. Buchges. Darmstadt) – [71]

Chiao, R. Y. (1998): Tunneling Times and Superluminality: a Tutorial. (e-print quant-ph/9811019) – [102]

Christodoulou, D. (1970): Reversible and Irreversible Transformations in Black-Hole Physics. Phys. Rev. Lett. **25**, 1596 – [143]

Cocke, W.J. (1967): Statistical Time Symmetry and Two-Time Boundary Conditions in Physics and Cosmology. Phys. Rev. **160**, 1165 – [158]

Conradi, H.D., and Zeh, H.D. (1991): Quantum cosmology as an initial value problem. Phys. Lett. **A154**, 321 – [172,190]

Craig, D.A. (1996): Observation of the Final Boundary Condition: Extra-galactic Background Radiation and the Time Symmetry of the Universe. Ann. Phys. (N.Y.) **251**, 384 – [158]

Cramer, J.G. (1986): Generalized Absorber Theory and the Einstein-Podolsky-Rosen Paradox. Phys. Rev. **D22**, 362 – [97]

Cross, M.C., and Hohenberg, P.C. (1993): Pattern formation outside of equilibrium. Rev. Mod. Phys. **65**, 851 – [81]

Dabrowski, M.P., and Kiefer, C. (1997): Boundary Conditions in Quantum String Cosmology. Phys. Lett. **B397**, 185 – [166]

Davidovich, L., Brune, M., Raimond, J.M., and Haroche, S. (1996): Mesoscopic quantum decoherences in cavity QED: Preparation and decoherence monitoring schemes. Phys. Rev. **A53**, 1295 – [108]

Davies, E.B. (1976): *Quantum Theory of Open Systems* (Academic Press) - [96]

Davies, P.C.W. (1983): Inflation and Time Asymmetry of the Universe. Nature **301**, 398 – [156]

Davies, P.C.W. (1984): Inflation in the Universe and time asymmetry. Nature **312**, 524 – [156,157]

Davies, P.C.W., and Fulling, S.A. (1977): Radiation from Moving Mirrors and from Black Holes. Proc. R. Soc. London **A356**, 237 – [150]

Davies, P.C.W., and Twamley, J. (1993): Time-Symmetric Cosmology and the Opacity of the Future Light Cone. Class. Quantum Grav. **10**, 931 – [35,158,190]

Demers, J.-G., and Kiefer, C. (1996): Decoherence of Black Holes by Hawking Radiation. Phys. Rev. **D53**, 7050 – [146]

Denbigh, K.G. (1981):*Three Concepts of Time* (Springer-Verlag) – [14,76]

Denbigh, K.G., and Denbigh, J.S. (1985): *Entropy in Relation to Incomplete Knowledge* (Cambridge University Press) – [75]

d'Espagnat, B. (1966): "An elementary note about mixtures" In *Preludes in theoretical physics*, ed. by De-Shalit, A., Feshbach, H., and van Hove, L. (North Holland) – [94]

d'Espagnat, B. (1976): *Conceptual Foundations of Quantum Mechanics* (Benjamin) – [99]

d'Espagnat, B. (1983): *In Search of Reality* (Springer-Verlag) – [99]

d'Espagnat, B. (1995): *Veiled Reality* (Addison-Wesley) – [132,199]

Deutsch, D. (1997): *The Fabric of Reality* (Penguin) – [130]

DeWitt, B.S. (1967): Quantum Theory of Gravity. I. The Canonical Theory. Phys. Rev. **160**, 1113 – [174,177,180,185]

DeWitt, B.S. (1971): "The many-universes interpretation of quantum mechanics." In *Foundations of Quantum Mechanics* (Fermi School of Physics **IL**), ed. by d'Espagnat, B. (Academic Press) – [130]

Diósi, L. (1987): A Universal Master Equation for the Gravitational Violation of Quantum Mechanics. Phys. Lett. **A120**, 377 – [115,141]

Diósi, L. (1988): Continuous quantum measurement and Itô formalism. Phys. Lett. **A129**, 419 – [114]

Diósi, L., Gisin, N., Halliwell, J., and Percival, I.C. (1995): Decoherent Histories and Quantum State Diffusion. Phys. Rev. Lett. **74**, 203 – [124]

Diósi, L., and Kiefer, C. (2000): Robustness and diffusion of pointer states. Phys. Rev. Lett. **85**, 3552 – [114]

Diósi, L., and Lukácz, B. (1994): *Stochastic Evolution of Quantum States in Opoen Systems and in Measurement Processes* (World Scientific) – [115]

Dirac, P.A.M. (1938): Classical Theory of Radiating Electrons. Proc. Roy. Soc. **A167**, 148 – [29]

Dirac, P.A.M. (1947): *The Principles of Quantum Mechanics.* (Clarendon Press) – [107]

Dukas, H., and Hoffman, B. (1979): *Albert Einstein: The Human Side* (Princeton University Press) – [197]

Dyson, F. (1979): Time without End: Physics and Biology in an Open Universe. Rev. Mod. Phys. **51**, 447 – [82]

Eardley, D.M. (1974): Death of White Holes in the Early Universe. Phys. Rev. Lett. **33**, 442 – [141]

Eddington, A.S. (1928): *The Nature of the Physical World* (Cambridge University Press) – [5]

Ehrenfest, P., and Ehrenfest, T. (1911): Begriffliche Grundlagen der statistischen Auffassung in der Mechanik. Encycl. Math. Wiss. **4**, No. 32; English translation: The Conceptual Foundations of the Statistical Approach in Mechanics. (Cornell University Press, 1959) – [50,201]

Einstein, A. (1949): Autobiographical Notes, in: *Albert Einstein, Philosopher-Scientist*, ed. by Schilpp, P.A. (The Library of Living Philosophers) – [6]

Einstein, A., and Born, M. (1969): *Briefwechsel 1916–1955* (Ullstein) – [8]

Einstein, A., and Ritz, W. (1909): Zum gegenwärtigen Stand des Strahlungsproblems. Phys. Z. **10**, 323 – [17]

Ekert, A., and Jozsa, R. (1996): Quantum computation and Shor's factoring algorithm. Rev. Mod. Phys. **68**, 733 – [104]

Ellis, J., Mohanty, S., and Nanopoulos, D.V. (1989): Quantum Gravity and the Collapse of the Wave Function. Phys. Lett. **B221**, 113 – [115,141]

Everett, H. (1957): 'Relative State' Formulation of Quantum Mechanics. Rev. Mod. Phys. **29**, 454 – **(WZ)** – [130,169]

Favre, C., and Martin, P.A. (1968): Dynamique quantique des systèmes amortis non-markovien. Helv. Phys. Acta **41**, 333 – [96]

Feshbach, H. (1962): A Unified Theory of Nuclear Reactions. Ann. Phys. (N.Y.) **19**, 287 – [61]

Feynman, R.P., Leighton, R.B., and Sands, M. (1963): *Lectures on Physics* (Addison-Wesley) – [73]

Feynman, R.P., and Vernon Jr., F.L. (1963): The Theory of a General Quantum System Interacting with a Linear Dissipative System. Ann. Phys. (N.Y.), **24**, 118 – [102,111]

Flasch, K. (1993): *Was ist Zeit?* (Klostermann, Frankfurt) – [9]

Fokker, A.D. (1929): Ein invarianter Variationssatz für die Bewegung mehrerer elektrischer Masseteilchen. Z. Phys. **58**, 386 – [33]

Foong, S.K., and Kanno, S. (1994): Proof of Page's Conjecture on the Average Entropy of a Subsystem. Phys. Rev. Lett. **72**, 1148 – [96]

Ford, J. (1989): "What is chaos, that we should be mindful of it?" In *The New Physics*, ed. by P.C.W. Davies (Cambridge University Press) – [76]

Franck, Ph. (1949): *Einstein* (Vieweg) – [199]

Friedman, J., Morris, M.S., Novikov, I.D., Echeverria, F., Klinkhammer, G., and Thorne, K.S. (1990): Cauchy problem in spacetime with closed timelike curves. Phys. Rev. **D42**, 1915 – [13]

Fröhlich, H. (1973): The Connection between Macro- and Microphysics. Riv. del Nuovo Cim. **3**, 490 – [48]

Frolov, V.P., and Novikov, I.D. (1990): Physical effects in wormholes and time machines. Phys. Rev. **D42**, 1057 – [13]

Fugmann, W., and Kretzschmar, M. (1991): Classical Electromagnetic Radiation in Noninertial Reference Frames. Nuovo Cim. **106 B**, 351 – [28]

Fukuyama, T., and Morikawa, M. (1989): Two-dimensional quantum cosmology: Directions of dynamical and thermodynamical arrows of time. Phys. Rev. **D39**, 462 – [191]

Gabor, D. (1964): Light and Information. Progr. in Optics **1**, 111 – **(LR)** – [74]

Gal-Or, B. (1974): Modern Developments in Thermodynamics (Wiley) – [156]

Gardner, M. (1967): Can time go backwards? Sci. Am. **216**, 98 – [71]

Gell-Mann, M., and Hartle, J.B. (1990): "Quantum Mechanics in the Light of Quantum Cosmology", In *Complexity, Entropy and the Physics of Information*, ed. by W.H. Zurek (Addison Wesley) – [124,169]

Gell-Mann, M., and Hartle, J.B. (1993): Classical Equations for Quantum Systems. Phys. Rev. **D47**, 3345 – [99,124]

Gell-Mann, M., and Hartle, J.B. (1994): "Time Symmetry and Asymmetry in Quantum Mechanics and Quantum Cosmology." In *Physical Origins of the Asymmetry of Time*, ed. by Halliwell, J.J., Perez-Mercader, J., and Zurek, W.H. (Cambridge University Press) – [170]

Gerlach, U.H. (1988): Dynamics and symmetries of a field partitioned by an accelerated frame. Phys. Rev. **D38**, 522 – [149]

Ghirardi, G.C., Rimini, A., and Weber, T. (1986): Unified dynamics for microscopic and macroscopic systems. Phys. Rev. **D34**, 470 – [115]

Ghirardi, G.C., Pearle, Ph., and Rimini, A. (1990): Markov processes in Hilbert space and continuous spontaneous localization of systems of identical particles. Phys. Rev. **A42**, 78 – [115]

Gibbs, J.W., (1902): *Elementary Principles in Statistical Mechanics* (Yale University Press) – [46]

Gisin, N. (1984): Quantum Measurement and Stochastic Processes. Phys. Rev. Lett. **52**, 1657 – [114]

Gisin, N., and Percival, I. (1992): The quantum-state diffusion model applied to open systems. J. Phys. **A25**, 5677 – [114]

Giulini, D. (1995): What is the geometry of superspace? Phys. Rev. **D51**, 5630 – [166,187]

Giulini, D. (1998): The generalized thin-sandwich problem and its local solvability. (e-print gr-qc/9805065) – [160]

Giulini, D., Joos, E., Kiefer, C., Kupsch, J., Stamatescu, I.-O., and Zeh, H.D. (1996): *Decoherence and the Appearance of a Classical World in Quantum Theory* (Springer) – [72,100,102,104,106,109,114,116,123,124,132,188]

Giulini, D., and Kiefer, C. (1994): Wheeler-DeWitt metric and the attractivity of gravity. Phys. Lett. **A193**, 21 – [166,178,188]

Giulini, D., Kiefer, C., and Zeh, H.D. (1995): Symmetries, superselection rules, and decoherence. Phys. Lett. **A199**, 291 – [105,106]

Glansdorff, P., and Prigogine, I. (1971): *Thermodynamic Theory of Structure, Stability and Fluctuation* (Wiley) – [57,81]

Gold, T. (1962): The Arrow of Time. Am. J. Phys. **30**, 403 – [135,155]

Goldwirth, D.S., and Piran, T. (1991): Entropy, inflation and the arrow of time. Class. Quantum Grav. **8**, L155 – [156]

Gorini, V., Kossakowski, A., and Sudarshan, E.C.G. (1976): Completely positive dynamical semigroups of N-level systems. J. Math. Phys. **17**, 821 – [112]

Gott, J.R., and Li, L.-X. (1997): Can the Universe Create Itself? Phys. Rev. **D58**, 3501 – [145]

Gottfried, K. (1966): *Quantum Mechanics* (Addison-Wesley). – [99]

Gould, A. (1987): Classical derivation of black hole entropy. Phys. Rev. **D35**, 449 – [145]

Gourgoulkon, E., and Bonazzola, S. (1994): A foundation of the virial theorem in general relativity. Class. Quantum Grav. **11**, 443 – [133]

Grad, H. (1961): The Many Faces of Entropy. Comm. Pure a. Appl. Math. **XIV**, 323 – [58]

Graham, N. (1970): The Everett interpretation of quantum mechanics. (Thesis: Univ. North Carolina, Chapel Hill) – [132]

Greenberger, D.M., Horne, M.A., Shimony, A., and Zeilinger, A. (1990): Bell's theorem without inequalities. Am. J. Phys. **58**, 1131 – [93]

Griffiths, R.B. (1984): Consistent Histories and the Interpretation of Quantum Mechanics. J. Stat. Phys. **36**, 219 – [99,124]

Groot, S.R. de, and Mazur, P. (1962): *Non-Equilibrium Thermodynamics* (North-Holland) – [64]

Habib, S., Kluger, Y., Mottola, and Paz, J.P. (1996): Dissipation and Decoherence in Mean Field Theory. Phys. Rev. Lett. **76**, 4660 – [107]

Habib, S. Shizume, K., and Zurek, W.H. (1998): Decoherence, Chaos, and the Correspondence Principle. Phys. Rev. Lett. **80**, 4361 – [88]

Haken, H. (1978): *Synergetics* (Springer-Verlag) – [81]

Halliwell, J.J. (1989): Decoherence in quantum cosmology. Phys. Rev. **D39**, 2912 – [188]

Halliwell, J.J. (2000): Trajectories for the Wave Function of the Universe from a Simple Detector Model. e-print gr-qc/0008046 – [191]

Halliwell, J.J., Perez-Mercader, J., and Zurek, W.H. (1994): *Physical Origins of the Asymmetry of Time.* (Cambridge University Press) – [208]

Halliwell, J.J., and Hawking, S.W. (1985): Origin of structure in the Universe. Phys. Rev. **D31**, 1777 – [152,165,188]

Hao Bai-Lin (1987): *Direction in Chaos* (World Scientific) – [70]

Haroche, S., and Kleppner, D. (1989): Cavity Quantum Electrodynamics. Physics Today **42**, 24 – [119]

Haroche, S., and Raimond, J.M. (1996): Quantum Computing: Dream or Nightmare? Phys. Today **49**, (August) 51 – [104]

Harrison, E.R. (1977): The dark night sky paradox. Am. J. Phys. **45**, 119 – [24]

Hartle, J.B. (1998): Quantum Pasts and the Utility of History. Phys. Scripta **T76**, 159 – [124,192]

Hartle, J.B., and Hawking, S.W. (1983): Wave Function of the Universe. Phys. Rev. **D28**, 2960 – [191]

Hawking, S.W. (1975): Particle Creation by Black Holes. Comm. Math. Phys. **43**, 199 – [144]

Hawking, S.W. (1976): Black Holes and Thermodynamics. Phys. Rev. **D13**, 191 – [145,146,151]

Hawking, S.W. (1985): Arrow of Time in Cosmology. Phys. Rev. **D32**, 2489 – [191]

Hawking, S.W. (1987): Quantum Coherence down the Wormhole. Phys. Lett. **B195**, 337 – [115]

Hawking, S.W. (1992): Chronology protection conjecture. Phys. Rev. **D46**, 603 – [13]

Hawking, S.W., and Ellis, G.F.R. (1973): *The Large Scale Structure of Spacetime* (Cambridge University Press) – [141,143]

Hawking, S.W., and Moss, I.G. (1982): Supercooled Phase Transitions in the Very Early Universe. Phys. Lett. **110B**, 35 – [156]

Hawking, S.W., and Ross, S.F. (1997): Loss of quantum coherence through scattering off virtual black holes. Phys. Rev. **D56**, 6403 – [141]

Hawking, S.W., and Wu, Z.C. (1985): Numerical Calculations of Minisuperspace Cosmological Models. Phys. Lett. **151B**, 15 – [165]

Hegerfeldt, G. (1994): Causality Problems for Fermi's Two-Atom System. Phys. Rev. Lett. **72**, 596 – [120]

Heisenberg, W. (1956): Die Entwicklung der Deutung der Quantentheorie. Phys. Blätter **12**, 289 – [83]

Heisenberg, W. (1957): Quantum Theory of Fields and Elementary Particles. Rev. Mod. Phys. **29**, 269 – [171]

Hepp, K. (1972): Macroscopic Theory of Measurement and Macroscopic Observables. Helv. Phys. Acta **45**, 237 – [86,94]

Hilgevoord, J. (1998): The uncertainty principle for energy and time. II. Am. J. Phys. **66**, 396 – [176]

Hogarth, J.E. (1962): Cosmological Considerations of the Absorber Theory of Radiation. Proc. Royal Soc. **A267**, 365 – [25]

Holland, P.R. (1993): The de Broglie-Bohm Theory of Motion and Quantum Field Theory. Phys. Rep. **224**, 95 – [124]

't Hooft, G. (1990): The Black Hole Interpretation of String Theory. Nucl. Phys. **B335**, 138 – [142,194]

Horwich, P. (1987): *Asymmetries in Time* (MIT Press) – [16]

Hoyle, F., and Narlikar, J.V. (1995): Cosmology and action-at-a-distance electrodynamics. Rev. Mod. Phys. **67**, 113 – [36]

Hu, B.L., Paz, J.P., and Zhang, Y. (1992): Quantum Brownian Motion in a General Environment: A model for system-field interactions. Phys. Rev. **D45**, 2843 – [102]

Imry, Y. (1997): *Mesoscopic Physics* (Oxford UP) – [106]

Isham, C.J. (1992): "Canonical Quantum Gravity and the Problem of Time." In *Integrable Systems, Quantum Groups, and Quantum Field Theory*, ed. by Ibart, L.A., and Rodrigues, M.A. (Kluwer Academic) – [175]

Israel, W. (1986): Third Law of Black Hole Dynamics: A Formulation and Proof. Phys. Rev. Letters **57**, 397 – [143]

Jadczyk, A. (1995): On Quantum Jumps, Events and Spontaneous Localization Models. Found. Phys. **25**, 743 – [110]

Jammer, M. (1974): *The Philosophy of Quantum Mechanics* (Wiley) – [132, 174]

Jancel, R. (1963): *Foundations of Quantum and Classical Statistical Mechanics* (Pergamon Press) – [65,88,91]

Jauch, J.M. (1968): *Foundations of Quantum Mechanics* (Addison-Wesley) – [86]

Jaynes, E.T. (1957): Information Theory and Statistical Mechanics. Phys. Rev. **106**, 620; **108**, 171 – [66]

Joos, E. (1984): Continuous Measurement: Watchdog Effect versus Golden Rule. Phys. Rev. **D29**, 1626 – [60,89,104,110]

Joos, E. (1986): Why do we observe a classical spacetime? Phys. Lett. **A116**, 6 – [108,115]

Joos, E. (1987): Time Arrow in Quantum Theory. J. Non-Equil. Thermodyn. **12**, 27 – [173]

Joos, E., and Zeh, H.D. (1985): The Emergence of Classical Properties through Interaction with the Environment. Z. Phys. **B59**, 223 – [100,101,103,104]

Kac, M. (1959): *Probability and Related Topics in Physical Sciences* (Interscience) – [65,201]

Károlyházy, F., Frenkel, A., and Lukácz, B. (1986): “On the possible role of gravity in the reduction of the wave function”, In *Quantum Concepts in Space and Time*, ed. by Penrose, R., and Isham, C.J. (Clarendon Press) – [115,141]

Kastner, R.E. (1998): Time-Symmetrized Quantum Theory, Counterfactuals, and ‘Advanced Action’. (e-print quant-ph/9806002) – [129]

Kauffman, S.A. (1991): *Origin of Order: Self-Organization and Selection in Evolution* (Oxford University Press) – [79]

Kaup, D.J., and Vitello, A.P. (1974): Solvable quantum cosmological model and the importance of quantizing in a special canonical frame. Phys. Rev. **D9**, 1648 – [177]

Kent, A. (1997): Quantum histories and their implications. Phys. Rev. Lett. **78**, 2874 – [124]

Kent, A., and McElwaine, J. (1997): Quantum prediction algorithms. Phys. Rev. **A55**, 1703 – [132]

Keski-Vakkuro, E., Lifschytz, G., Mathur, S.D., and Ortiz, M.E. (1995): Breakdown of the semiclassical approximation at the black hole horizon. Phys. Rev. **D51**, 1764 – [194]

Khalfin, L.A. (1958): Contribution to the Theory of a Quasi-Stationary State. Sov. Phys. JETP **6 (33)**, 1053 – [117]

Kiefer, C. (1987): Continuous measurement of mini-superspace variables by higher multipoles. Class. Qu. Gravity **4**, 1369 – [184,189]

Kiefer, C. (1988): Wave packets in mini-superspace. Phys. Rev. **D38**, 1761 – [182,184]

Kiefer, C. (1989): Non-minimally coupled scalar fields and the initial value problem in quantum gravity. Phys. Lett. **B225**, 227 – [178]

Kiefer, C. (1992): Decoherence in quantum electrodynamics and quantum cosmology. Phys. Rev. **D46**, 1658 – [104,107,188]

Kiefer, C. (1994): “The semiclassical approximation to quantum gravity.” In *Canonical gravity: From classical to quantum*, ed. by Ehlers, J., and Friedrich, H. (Springer) – [193]

Kiefer, C. (1997): Semiclassical quantum states for black holes. Nucl. Phys. **B57**, (Proc. Suppl.) 173 – [194]

Kiefer, C. (1998): Interference of two independent sources. Am. J. Phys. **66**, 661 – [107]

Kiefer, C. (1999): “Decoherence in Situations Involving the Gravitational Field.” In *Decoherence: Theoretical, Experimental, and Conceptual Problems*, ed. by Blanchard, Ph., Giulini, D., Joos, E., Kiefer, C., and Stamatescu, I.-O. (Springer) – [109]

Kiefer, C., and Louko, J. (1998): Hamilton evolution and quantization for extremal black holes. (e-print gr-qc/9809005) - [194]

Kiefer, C., Müller, R., and Singh, T.P. (1994): Quantum Gravity and Non-Unitarity in Black Hole Evaporation. Mod. Phys. Lett. **A9** 2661 – [108]

Kiefer, C., Polarski, D., and Starobinsky, A. A. (1998): Quantum-to-classical transition for fluctuations in the early universe. Int. J. Mod. Phys. **D7**, 455 – [154,173]

Kiefer, C., and Singh T.P. (1991): Quantum Gravitational Corrections to the Functional Schrödinger equation. Phys. Rev. **D44**, 1067 – [187]

Kiefer, C., and Zeh, H.D. (1995): Arrow of time in a recollapsing quantum universe. Phys. Rev. **D51**, 4145 – [192,194,195]

Knight, P. (1998): Where the weirdness comes from. Nature **395**, 12 – [125]

Kolmogorov, A.N. (1954): On Conservation of Conditionally-Periodic Motions for a Small Charge in Hamilton's Function. Dokl. Akad. Nauk USSR **98**, 525 – [70]

Kraus, K. (1971): General State Changes in Quantum Theory. Ann. Phys. (N.Y.) **64**, 311 – [111,112]

Kübler, O., and Zeh, H.D. (1973): Dynamics of Quantum Correlations. Ann. Phys. (N.Y.) **76**, 405 – [94,96,106,107,131]

Kwiat, P.G., Steinberg, A.M., and Chiao, R.Y. (1992): Observation of a 'quantum eraser': A revival of coherence in a two-photon interference experiment. Phys. Rev. **45**, 7729 – [100]

Lanczos, C. (1970): *The Variational Principles of Mechanics*, 4th ed. (University of Toronto Press) – [163]

Landau, L.D., and Lifschitz, E.M. (1959): *Fluid Mechanics* (Pergamon) – [57]

Landauer, R. (1961): Irreversibility and heat generation in the computing process. IBM J. Res. Dev. **3**, 183 **(LR)** – [74]

Landauer, R. (1996): The physical nature of information. Phys. Lett. **A217**, 188 – [75]

Lapchinsky, V.G., and Rubakov, V.A. (1979): Canonical Quantization of Gravity and Quantum Field Theory in Curved Space-Time. Acta Physica Polonica, **B10**, 1041 – [187]

Leff, H.S., and Rex, A.F. (1990): *Maxwell's Demon: Entropy, Information, Computing* (Princeton University Press) – [75,207]

Leggett, A.J. (1980): Macroscopic Quantum Systems and the Quantum Theory of Measurement. Suppl. Progr. Theor. Phys. **69**, 80 – [100]

Levine, H., Moniz, E., and Sharp, D.H. (1977): Motion of extended charges in classical electrodynemics. Am. J. Phys. **45**, 75 – [32]

Lewis, R.M. (1967): A Unifying Principle in Statistical Mechanics. J. Math. Physics **8**, 1449 – [51,56]

Lewis, D. (1986): *Philosophical Papers*, Vol. II (Oxford University Press) – [16]

Li, L.-X., and Gott III, J.R. (1998): Self-Consistent Vacuum for Misner Space and the Chronology Protection Conjecture. Phys. Rev. Lett. **80**, 2980 – [194]

Lindblad, G. (1976): On the Generators of Quantum Mechanical Semigroups. Comm. Math. Phys. **48**, 119 – [111,113]

Linde, A. (1979): Phase Transitions in Gauge Theories and Cosmology. Rep. Progr. Phys. **42**, 389 – [154]

Lloyd, S., and Pagels, H. (1988): Complexity as Thermodynamical Depth. Ann. Phys. (N.Y.) **188**, 186 – [82]

Lockwood, M. (1996): 'Many Minds' Interpretation of Quantum Mechanics. Brit. J. Phil. Sci. **47**, 159 – [132]

London, F., and Bauer, E. (1939): *La théorie de l'observation en Méchanique quantique* (Hermann, Paris) **(WZ)** – [127]

Lubkin, E. (1978): Entropy of an n-system from its correlation with a k-system. J. Math. Phys. **19**, 1028 – [96]

Lubkin, E. (1987): Keeping the Entropy of Measurement: Szilard Revisited. Intern. J. Theor. Phys. **26**, 523 **(LR)** – [127]

Ludwig, G. (1990): *Die Grundstrukturen einer physikalischen Theorie* (Springer-Verlag) – [111]

Mackey, M.C. (1989): The dynamical origin of increasing entropy. Rev. Mod. Phys. **61**, 981 – [52,58,78]

Maeda, K., Ishibashi, A., and Narita, M. (1998): Chronology Protection and Non-Naked Singularities. Class. Quantum Grav. **15**, 1637 – [13]

Mayer, J.E., and Mayer, M.G. (1940): *Statistical Mechanics* (Wiley) – [48]

Mensky, M.B. (1979): Quantum restrictions for continuous observations of an oscillator. Phys. Rev. **D20**, 384 – [102]

Mensky, M.B. (2000): *Quantum Measurements and Decoherence* (Kluwer Academic) – [102]

Mermin, D.N. (1998): Nonlocal character of quantum theory? Am. J. Phys. **66**, 920 – [8]

Misner, C.W. (1969): Absolute Zero of Time. Phys. Rev. *186*, 1328 – [178]

Misner, C.W., Thorne, K.S., and Wheeler, J.A. (1973): *Gravitation* (Freeman) – [159,167,168,194]

Misra, B. (1978): Nonequilibrium, Lyaponov variables, and ergodic properties of classical systems. Proc. Natl. Acad. USA **75**, 1627 – [58]

Misra, B., and Sudarshan, B.C.G. (1977): The Zeno's paradox in quantum theory. J. Math. Phys. **18**, 756 – [60]

Mittelstaedt, P. (1976): *Der Zeitbegriff in der Physik* (BI Wissenschaftsverlag) – [14]

Moniz, E., and Sharp, D.H. (1977): Radiation reaction in nonrelativistic quantum electrodynamics. Phys. Rev. **15**, 2850 – [31]

Monroe, C., Meekhof, King, B.E., and Wineland, D.J. (1996): A "Schrödinger Cat" Superposition State of an Atom. Science **272**, 1131 – [108]

Monteoliva, D., and Paz, J.P. (2000): Decoherence and the rate of entropy production in chaotic quantum systems. e-print quant-ph/0007052 – [103]

Morris, M.S., Thorne, K.S., and Yurtsever, U. (1988): Wormholes, Time Machines, and the Weak Energy Condition. Phys. Rev. Lett. **61**, 1446 – [13]

Mott, N.F. (1929): The wave mechanics of α-ray tracks. Proc. R. Soc. London **A126**, 79 – [101,192]

Mould, R.A. (1964): Internal Absorption Properties of Accelerated Detectors. Ann. Phys. **27**, 1 – [28,148]

Nagourney, W. Sandberg, J., and Dehmelt, H. (1986): Shelved Optical Electron Amplifier: Observation of Quantum Jumps. Phys. Rev. Lett. **56**, 2797 – [109]

Nambu, Y., and Jona-Lasinio, G. (1961): Dynamical Model of Elementary Particles Based on an Analogy with Superconductivity. Phys. Rev. **122**, 345 – [171]

Neumann, J. von (1932): *Mathematische Grundlagen der Quantenmechanik* (Springer-Verlag); Engl. transl. (1955): *Mathematical Foundations of Quantum Mechanics* (Princeton University Press) – [97,199]

Nicklaus, M., and Hasselbach, F. (1993): Wien Filter: A wave-packet-shifting device for restoring longitudinal coherence in in charged-matter-wave interferometer. Phys. Rev. **A48**, 152 – [105]

Norton, J. (1989): Coordinates and Covariance: Einstein's View of Space-Time and the Modern View. Found. Phys. **19**, 1215 – [11,161]

Norton, J. (1995): "Mach's Principle before Einstein" In *Mach's Principle: From Newton's Bucket to Quantum Gravity*, ed. by Barbour, J.B., and Pfister, H. (Birkhäuser, Basel) – [11]

Omnès, R. (1988): Logical Reformulation of Quantum Mechanics. J. Stat. Phys. **53**, 893 – [124]

Omnès, R. (1992): Consistent interpretation of quantum mechanics. Rev. Mod. Phys. **64**, 339 –[99,124]

Omnès, R. (1997): General theory of decoherence effect in quantum mechanics. Phys. Rev. **A56**, 3383 – [103]

Omnès, R. (1998): "Are there unsolved problems in the interpretation of Quantum Mechanics?" In *Open Systems and Measurements in Relativistic Quantum Theory*, ed. by Breuer, H.P., and Petruccione, F. – [124]

Omnès, R. (1999): *Understanding Quantum Mechanics* (Princeton UP) – [99]

Page, D.N. (1980): Is Black-Hole Evaporation Predictable? Phys. Rev. Lett. **44**, 301 – [145]

Page, D.N. (1984): Can Inflation Explain the Second Law of Thermodynamics? Int. J. Theor. Phys. **23**, 725 – [156]

Page, D.N. (1985): Will entropy increase if the Universe recollapses? Phys. Rev. **D32**, 2496 – [191]

Page, D.N. (1993): Average Entropy of a Subsystem. Phys. Rev. Lett. **71**, 1291 – [96]

Page, D.N., and Geilker, C.D. (1982): "An Experimental Test of Quantum Gravity", In *Quantum Structure of Space and Time*, ed. by Duff, M.J., and Isham, C.J. (Cambridge University Press) – [174]

Page, D.N., and Wootters, W.K. (1983): Evolution without Evolution: Dynamics Described by Stationary Observables. Phys. Rev. **D27**, 2885 – [176]

Partridge, R.B. (1973): Absorber Theory of Radiation and the Future of the Universe. Nature **244**, 263 – [36]

Pauli, W. (1928): “Über das H-Theorem vom Anwachsen der Entropie vom Standpunkt der neuen Quantenmechanik”, In *Probleme der Modernen Physik, Sommerfeld-Festschrift* (Hirzel) – [86,88]

Paz, J.P., and Zurek, W.H. (1999): Quantum Limit of Decoherence: Environment Induced Superselection of Energy Eigenstates. Phys. Rev. Lett. **82**, 5181 – [110]

Pearle, Ph. (1976): Reduction of the state vector by a non-linear Schrödinger equation. Phys. Rev. **D13**, 857 – [114]

Pearle, Ph. (1979): Toward Explaining Why Events Occur. Int. J. Theor. Phys. **48**, 489 – [96]

Pearle, Ph. (1989): Combining stochastic dynamical state-vector reduction with spontaneous localization. Phys. Rev. **A39**, 2277 – [115]

Pearle, Ph., and Squires, E.J. (1994): Bound State Excitation, Nucleon Decay Experiments, and Models of Wave Function Collapse. Phys. Rev. Lett. **73**, 1 – [115]

Peierls, R. (1985): “Observation in Quantum Mechanics and the collapse of the wave function”, In *Symposium on the Foundations of Modern Physics*, ed. by Lahti, P., and Mittelstaedt, P. (World Scientific) – [98]

Penrose, O., and Percival, I.C. (1962): The Direction of Time. Proc. Phys. Soc. London **79**, 605 – [56]

Penrose, R. (1969): Gravitational Collapse: The role of general relativity. Nuovo Cim. **1**, special number, 252 – [13,136,142]

Penrose, R. (1979): “Singularities and Time-Asymmetry”, In *General Relativity*, ed. by Hawking, S.W., and Israel, W. (Cambridge University Press) – [128]

Penrose, R. (1981): “Time Asymmetry and Quantum Gravity”, In *Quantum Gravity 2*, ed. by Isham, C.J., Penrose, R., and Sciama, D.W. (Clarendon Press) – [136,155]

Penrose, R. (1986): “Gravity and state vector reduction”, In *Quantum Concepts in Space and Time*, ed. by Penrose, R., and Isham, C.J. (Clarendon Press) – [115,141]

Percival, I.C. (1997): Atom interferometry, spacetime and reality. Phys. World **10**, (March) 43 – [141]

Peres, A. (1962): On the Cauchy problem in general relativity, II. Nuovo Cim. **26**, 53 – [187]

Peres, A. (1980a): Nonexponential Decay Law, Ann. Phys. (N.Y.) **129**, 33 – [117,119]

Peres, A. (1980b): Measurement of Time by Quantum Clocks. Am. J. Phys. **48**, 552 – [176]

Peres, A. (1994): Time asymmetry in quantum mechanics: a retrodiction paradox. Phys. Lett. **A194**, 21 – [129]

Peres, A. (1996): Separability Criterion for Density Matrices. Phys. Rev. Lett. **77**, 1413 – [95]

Petzold, J. (1959): Zum Anfangswertproblem zerfallender Zustände. Z. Phys. **157**, 122 – [117,119]

Poincaré, H. (1902): *La science et l'hypothèse* (Flammarion) – [10,82,129]

Pöppel, E., Schill, K., and von Steinbüchel, N. (1990): Sensory integration within temporally neutral system states. Naturwiss. **77**, 89 – [12]

Popper, K. (1950): Indeterminism in quantum physics and in classical physics, British J. Philos. of Science **1**, 117-173 – [71]

Popper, K. (1956): The Arrow of Time. Nature **177**, 538 – [36]

Price, H. (1996): *Time's Arrow & Archimedes' Point: A View from Nowhen* (Oxford University Press) – [II,6,8,16,36,39,57,97,155,197]

Prigogine, I. (1962): *Non-Equilibrium Statistical Mechanics* (Wiley) – [49]

Prigogine, I. (1980): *From Being to Becoming* (Freeman) – [6,71,116]

Primas, H. (1983): *Chemistry, Quantum Mechanics and Reductionism*, (Springer-Verlag) – [103]

Qadir, A., and Wheeler, J.A. (1985): "York's Cosmic Time Versus Proper Time as Relevant to Changes in the Dimensionless "Constants", K-Meson Decay, and the Unity of Black Hole and Big Crunch", In *From SU(3) to gravity*, ed. by Grothmann, E.S., and Tauber, G. (Cambridge University Press) – [142]

Qadir, A. (1988): "Black holes in closed universes", In *Proceedings of the Fifth Marcel Grossmann Meeting*, ed. by D. Blair and M. Buckingham (World Scientific) – [167]

Quinn, T.C., and Wald, R.M. (1997): Axiomatic approach to electromagnetic and gravitational radiation reaction of particles in curved spacetime. Phys. Rev. **D56**, 3381 – [32]

Reichenbach, H. (1956): *The Direction of Time* (University of California Press) – [14,16]

Rempe, G., Walther, H., and Klein, N. (1987): Observation of Quantum Collapse and Revival in a One-Atom Maser. Phys. Rev. Lett. **58**, 353 – [109,119]

Rhim, W.-K., Pines, A., and Waugh, J.S. (1971): Time-Reversal Experiments in Dipolar-Coupled Spin Systems. Phys. Rev. **B3**, 684 – [54]

Rohrlich, F. (1965): *Classical Charged Particles* (Addison-Wesley) – [26,29,32]

Rohrlich, F. (1997): The dynamics of a charged sphere and the electron. Am. J. Phys. **65**, 1051 – [32]

Rohrlich, F. (1998): The Arrow of Time in the Equations of Motion. Found. Phys. **28**, 1045 – [32]

Röpke, G. (1987): *Statistische Mechanik für das Nichtgleichgewicht* (Physik-Verlag, Weinheim) – [64]

Rovelli, C. (1995): Analysis of the dinstinct meanings of the notion of "Time" in different physical theories. Nuovo Cim **110B**, 81 – [163]

Rovelli, C., and Smolin, L. (1990): Loop representation for Quantum General Relativity. Nucl. Phys. **B133**, 80 – [175]

Ryan, M.P. (1972): *Hamiltonian Cosmology* (Springer) – [184]

Sachs, R.G. (1987): *The Physics of Time Reversal* (University of Chicago Press) – [3]

Sauter, Th., Neuhauser, W., Blatt, R., and Toschek, P.E. (1986): Observation of Quantum Jumps. Phys. Rev. Lett. **57**, 1696 – [109]

Schlögl, F. (1966): Zur statistischen Theorie der Entropieproduktion in nicht abgeschlossenen Systemen. Z. Phys. **191**, 81 – [68]

Schmidt, E. (1907): Zur Theorie der linearen und nichtlinearen Integralgleichungen. Math. Ann. **63**, 433 – [94]

Schrödinger, E. (1935): Discussion of probability relations between separated systems. Proc. Cambridge Phil. Soc. **31**, 555 – [94,125]

Schulman, L.S. (1986): Deterministic Quantum Evolution through Modification of the Hypothesis of Statistical Mechanics. J. Stat. Phys. **42**, 689 – [97]

Schulman, L.S. (1997): "Time's arrows and quantum mechanics" (Cambridge University Press) – [II,97,155,158]

Schulman, L.S. (1999): Opposite Thermodynamical Arrows of Time. Phys. Rev. Lett. **83**, 5419 – [54]

Schuster, H.G. (1984): *Deterministic Chaos* (Physik-Verlag, Weinheim) – [70]

Scully, M.O., Englert, B.-G., and Walther, H. (1991): Quantum optical tests of complementarity. Nature **351**, 111 – [102]

Shannon, C. (1948): A Mathematical Theory of Communication. Bell System Techn. J. **27**, 370 – [66]

Smart, J.J.C. (1967): "Time", In *Encyclopedia of Philosophy*, Vol. 8, ed. by Edwards, Th. (Free Press) – [8]

Smoluchowski, M. Ritter von (1912): Experimentell nachweisbare, der üblichen Thermodynamik widersprechende Molekularphänomene. Phys. Z. **13**, 1069 – [71]

Smullyan, R.M. (1981): "An Epistemological Nightmare", In *The Mind's I*, ed. by Hofstadter, D.R., and Dennett, D.C. (Basic Books) – [71]

Sokolovsky, D. (1998): From Feynman histories to observables. Phys. Rev. **A57**, R1469 – [129]

Squires, E.J. (1990): *Conscious Mind in the Physical World*, (Hilger) – [132]

Szilard, L. (1929): Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen. Z. Phys. **53**, 840 – **(WZ,LR)** – [72]

Taylor, J.H. (1994): Binary pulsars and relativistic gravity. Rev. Mod. Phys. **66**, 711 – [25]

Tegmark, M. (1993): Apparent Wave Function Collapse Caused by Scattering. Found. Phys. Lett. **6**, 571 – [115]

Tegmark, M. (1998): The Interpretation of Quantum Mechanics: Many Worlds or Many Words. Fortschr. Physik **45**, 855 – [131,169]

Tegmark, M. (2000): Importance of quantum decoherence in brain processes. Phys. Rev. **61**, 4194 – [104,126,127,170]

Teitelboim, C. (1970): Splitting of the Maxwell Tensor: Radiation Reaction without Advanced Fields. Phys. Rev. **D1**, 1572 – [28]

Tittel, W., Brendel, T., Zbinden, H., and Gisin, N. (1998): Violation of Bell Inequalities by Photons More Than 10 km Apart. Phys. Rev. Lett. **81**, 3563 – [125]

Tolman, R.C. (1938): *The Principles of Statistical Mechanics* (Oxford University Press) – [38,95]

Unruh, W.G. (1976): Notes on Black-Hole Evaporation. Phys. Rev. **D14**, 870 – [148,150]

Unruh, W.G., and Wald, R.M. (1982): Acceleration Radiation and the Generalized Second Law of Thermodynamics. Phys. Rev. **D25**, 942 – [146,151]

Unruh, W.G., and Zurek, W.H. (1989): Reduction of the wave packet in quantum Brownian motion. Phys. Rev. **D40**, 1071 – [104]

Vaidman, L. (1997): Time-symmetrized quantum theory. (e-print quant-ph/9710036) – [170]

Vaihinger, H. (1911): *Die Philosophie des Als Ob* (Scientia Verlag) – [82,129]

Van Hove, L. (1957): The Approach to Equilibrium in Quantum Statistics. Physica **23**, 441 – [90]

Van Kampen, N.G. (1954): Quantum Statistics of Irreversible Processes. Physica **20**, 603 – [89]

Vaz, C., and Witten, L. (1998): Mass Quantization of Schwarzschild Black Holes, (e-print gr-qc/9811062) – [194]

Veneziano, G. (1991): Scale factor duality for classical and quantum strings. Phys. Lett. **B265**, 287 – [166]

Vilenkin, A. (1989): Interpretation of the wave function of the universe. Phys. Rev. **D39**, 1116 – [188]

Wald, R.M. (1980): Quantum Gravity and Time Reversibility. Phys. Rev. **D21**, 2742 – [141]

Wald, R.M. (1997): Gravitational Collapse and Cosmic Censorship. (e-print gr-qc/9710068) – [140]

Wehrl, A. (1978): General Properties of Entropy. Rev. Mod. Phys. **50**, 221 – [50]

Weinberg, S. (1977): *The First Three Minutes* (Basic Books, New York) – [169]

Weizsäcker, C.F. von (1939): Der Zweite Hauptsatz und der Unterschied zwischen Vergangenheit und Zukunft. Ann. Physik **36**, 275 – [81]

Weizsäcker, C.F. von (1982): *Die Einheit der Natur* (Hanser, München) – [6]

Werner, R.F. (1989): Quantum states with Einstein-Rosen correlations admitting a hidden variable model. Phys. Rev. **A40**, 4277 – [95]

Wheeler, J.A. (1979): “Frontiers of Time.” In *Problems in the Foundations of Physics* (Fermi School of Physics **LXXII**), ed. by G.T. di Francia (North Holland) – [71,177]

Wheeler, J.A., and Feynman, R.P. (1945): Interaction with the Absorber as the Mechanism of Radiation. Rev. Mod. Phys. **17**, 157 – [22,34,36]

Wheeler, J.A., and Feynman, R.P. (1949): Classical Electrodynamics in Terms of Direct Interparticle Action. Rev. Mod. Phys. **21**, 425 – [22,36]

Wheeler, J.A., and Zurek, W.H. (1983): *Quantum Theory and Measurement* (Princeton University Press) – [207]

Whitrow, G.J. (1980): *The Natural Philosophy of Time* (Clarendon Press) – [14]

Wick, G.C., Wightman, A.S., and Wigner, E.P. (1952): The intrinsic parity of elementary particles. Phys. Rev. **88**, 101 – [86]

Wigner, E.P. (1972): Events, laws of nature, and invariance principles. In *Nobel lectures – Physics* (Elsevier) p. 248 – [1]

Wilkinson, S.R., Barucha, C.F., Fischer, M.C., Madison, K.W., Morrow, P.R., Niu, Q., Sundaram, B., and Raizen, M.G. (1997): Experimental evidence for non-expontial decay in quantum tunneling. Nature **387**, 575 – [119]

Williams, D.C. (1951): The myth of the passage. J. Philosophy **48**, 457 – [9]

Willis, C.R., and Picard, R.H. (1974): Time-dependent projection-operator approach to coupled systems. Phys. Rev. **A9**, 1343 – [56]

Winter, R.G. (1961): Evolution of quasi-stationary state. Phys. Rev. **123**, 1503 – [119]

Witten, E. (1997): Duality, Spacetime and Quantum Mechanics. Phys. Today **50**, (May) 28 – [175]

Woolley, R.G. (1986): Molecular Shapes and Molecular Structures. Chem. Phys. Lett. **125**, 200 – [103]

Wootters, W.K. (1984): "Time" Replaced by Quantum Correlations. Int. J. Theor. Phys. **23**, 701 – [176]

Wu, K.S., Lahav, O., and Reeves, M.M. (1999): The large-scale smoothness of the Universe. Nature **397**, 225 – [24]

Yaghjian, A.D. (1992): *Relativistic Dynamics of a Charged Sphere* (Springer), App. D – [31]

Yoccoz, J.-C. (1992): Travaux De Hermann Sur les Tores Invatiantes. Asterisque **206**, 311 – [53]

York, J.W. (1983): Dynamical Origin of Black-Hole Radiance. Phys. Rev. **D28**, 2929 – [144]

Zeh, H.D. (1965): Symmetry Violating Trial Wave Functions. Z. Phys. **188**, 361 – [99]

Zeh, H.D. (1967): Symmetrieverletzende Modellzustände und kollektive Bewegungen. Z. Phys. **202**, 38 – [99,172,178]

Zeh, H.D. (1970): On the Interpretation of Measurement in Quantum Theory. Found.Phys. **1**, 69 – **(WZ)** – [100,103,125,130,132,169]

Zeh, H.D. (1971): "On the irreversibility of time and observation in quantum theory." In *Foundations of Quantum Mechanics* (Fermi School of Physics **IL**), ed. by d'Espagnat, B. (Academic Press) – [100]

Zeh, H.D. (1973): Toward a Quantum Theory of Observation. Found. Phys. **3**, 109 – [100,131]

Zeh, H.D. (1979): Quantum Theory and Time Asymmetry. Found. Phys. **9**, 803 – [131,132]

Zeh, H.D. (1983): "Einstein Nonlocality, Space-Time Structure, and Thermodynamics", In *Old and New Questions in Physics, Cosmology, Philos-*

ophy, and Theoretical Biology, ed. by van der Merwe, A. (Plenum) – [167,196]

Zeh, H.D. (1984): *Die Physik der Zeitrichtung* (Springer) – [II,175]

Zeh, H.D. (1986a): Die Suche nach dem Urzeitpfeil. Phys. Blätter **42**, 80 – [173]

Zeh, H.D. (1986b): Emergence of Classical Time from a Universal Wave Function. Phys. Lett. **A116**, 9 – [175,184,189]

Zeh, H.D. (1988): Time in Quantum Gravity. Phys. Lett. **A126**, 311 – [180,191]

Zeh, H.D. (1993): There are no quantum jumps, nor are there particles! Phys. Lett. **A172**, 189 – [110]

Zeh, H.D. (1994): "Time-(A-)Symmetry in a Recollapsing Quantum Universe" In *Physical Origins of the Asymmetry of Time*, ed. by Halliwell, J.J., Perez-Mercader, J., and Zurek, W.H. (Cambridge University Press) – [191,195]

Zeh, H.D. (1999a): Note on the Time Reversal Asymmetry of Equations of Motion. Found. Phys. Lett. **12**, 193 – [32]

Zeh, H.D. (1999b): Why Bohm's quantum theory? Found. Phys. Lett. **12**, 197 –[124]

Zeh, H.D. (2000): Conscious Observation in Quantum Mechanical Description. Found. Phys. Lett. **13**, 221 – [126,132]

Zurek, W.H. (1981): Pointer basis of quantum apparatus: Into what mixture does the wave packet collapse? Phys. Rev. **D24**, 1516 – [100]

Zurek, W.H. (1982a): Environment-induced superselection rules. Phys. Rev. **D26**, 1862 – [100,106]

Zurek, W.H. (1982b): Entropy Evaporated by a Black Hole. Phys. Rev. Lett. **49**, 1683 – [145]

Zurek, W.H. (1986): "Reduction of the Wave Packet: How Long Does it Take?" In *Frontiers of Nonequilibrium Statistical Physics*, ed. by G.T. Moore and M.T. Scully (Plenum) – [100]

Zurek, W.H. (1989): Algorithmic randomness and physical entropy. Phys. Rev. **A40**, 4731 – [70]

Zurek, W.H. (1991): Decoherence and the Transition from Quantum to Classical. Phys. Today **44**, (Oct.) 36 – [100,102]

Zurek, W.H. (1998): Decoherence, einselection and the existential interpretation (the rough guide). Phil. Trans. R. Soc. Lond. **A356**, 1793 – [104]

Zurek, W.H., Habib, S., and Paz., J.P. (1993): Coherent States via Decoherence. Phys. Rev. Lett. **70**, 1187 – [103,107]

Zurek, W.H., and Page, D.N. (1984): Black-hole thermodynamics and singular solutions of the Tolman-Oppenheimer-Volkhoff equation. Phys. Rev. **D29**, 628 – [145]

Zurek, W.H: and Paz, J.P. (1994): Decoherence, Chaos, and the Second Law. Phys. Rev. Lett. **72**, 2508 – [103]

Zwanzig, R. (1960): Statistical Mechanics of Irreversibility. Boulder Lectures in Theoretical Physics **3**, 106 – [55]