



The five problems of irreversibility

Michael te Vrugt*

Institut für Theoretische Physik, Westfälische Wilhelms-Universität Münster, D-48149 Münster, Germany
 Center for Soft Nanoscience, Westfälische Wilhelms-Universität Münster, D-48149 Münster, Germany
 Philosophisches Seminar, Westfälische Wilhelms-Universität Münster, D-48143 Münster, Germany

ARTICLE INFO

Keywords:

Statistical mechanics
 Thermodynamic irreversibility
 Coarse-graining
 Arrow of time
 Dynamical density functional theory
 Soft condensed matter

ABSTRACT

Thermodynamics has a clear arrow of time, characterized by the irreversible approach to equilibrium. This stands in contrast to the laws of microscopic theories, which are invariant under time-reversal. Foundational discussions of this “problem of irreversibility” often focus on historical considerations, and do therefore not take results of modern physical research on this topic into account. In this article, I will close this gap by studying the implications of *dynamical density functional theory* (DDFT), a central method of modern nonequilibrium statistical mechanics not previously considered in philosophy of physics, for this debate. For this purpose, the philosophical discussion of irreversibility is structured into five problems, concerned with the source of irreversibility in thermodynamics, the definition of equilibrium and entropy, the justification of coarse-graining, the approach to equilibrium and the arrow of time. For each of these problems, it is shown that DDFT provides novel insights that are of importance for both physicists and philosophers of physics.

1. Introduction

The temporal asymmetry of thermodynamics is one of the central problems in philosophy of physics. If a cup of hot coffee stands in a room, it will cool down until it has room temperature, but it will not spontaneously heat up by extracting heat from its environment. This is commonly seen as a consequence of the second law of thermodynamics, which assigns to each of these systems a quantity known as “entropy” that increases in these processes and that, most importantly, cannot decrease. Often, this is considered one of the most fundamental laws of physics (Callender, 2001, p. 540).

However, as is also well-known, there is a conflict with the microscopic laws governing the motion of the individual particles that a macroscopic system consists of. These laws are invariant under time-reversal,¹ which means that if a process can occur in one direction of time, it can also occur in the other direction. Thus, a cup of coffee at room temperature that spontaneously heats up would be in perfect agreement with the microscopic laws of physics, which makes it very difficult to explain why such a behavior is never observed.

An intense discussion on this problem has emerged in philosophy of physics. It has evolved into a variety of sub-debates concerned with different explananda that, in this article will be classified into five problems. Often, foundational discussions of statistical mechanics focus on

historical aspects. Moreover, there has been a growth of interest in formal approaches to coarse-graining based on the projection operator formalism (Mori, 1965; Zwanzig, 1960). There is, however, a lack of work that considers the implications of modern research on nonequilibrium statistical mechanics for foundational problems (Wallace, 2015).

In this work, I will close this gap by providing the first philosophical discussion of *dynamical density functional theory* (DDFT) (Evans, 1979; Marini Bettolo Marconi & Tarazona, 1999; te Vrugt et al., 2020b), which is one of the cornerstones of modern statistical physics. DDFT has originally been developed for modeling simple and complex fluids. It is now a central method of theoretical soft matter physics and has, moreover, found applications in other fields such as biology (Angioletti-Uberti et al., 2018), chemistry (Liu & Liu, 2020), epidemiology (te Vrugt et al., 2020a), or plasma physics (Diaw & Murillo, 2015). Moreover, it is intimately connected to the projection operator formalism (Español & Löwen, 2009; Wittkowski et al., 2012), such that it provides a natural link between foundational debates in philosophy of physics and practical work in condensed matter physics. For each of the five sub-problems discussed in this work, it is shown that DDFT can make interesting contributions to the debate in philosophy.

This article is structured as follows: In Section 2, I explain how “the” problem of irreversibility arises, and structure it into five different sub-problems. An introduction to DDFT can be found in Section 3. In

* Correspondence to: Institut für Theoretische Physik, Center for Soft Nanoscience, Busso-Peus-Straße 10, 48149 Münster, Germany.

E-mail address: michael.tevrugt@uni-muenster.de.

¹ When it comes to the most fundamental laws, this is not strictly true. The standard model of particle physics is invariant under CPT, which is a combination of charge-conjugation, mirror reflection and time-reversal. However, this effect is too small to account for the temporal asymmetry of thermodynamics (Wallace, 2013, p. 270).

Sections 4 to 8, I explain in detail each problem and the contribution that DDFT can make to its solution. I conclude in Section 9.

2. The five problems of irreversibility

In this section, I will briefly recapitulate what is commonly referred to as “the problem of irreversibility”. This is a long withstanding debate in the foundations of statistical mechanics, which has evolved into a variety of sub-debates that are dealing with different aspects. A careful distinction of the different explananda is required both to ensure a conceptually precise treatment and to appreciate that studying DDFT is important for a variety of foundational debates. For this purpose, I will introduce here a scheme of “problems of irreversibility”, which is used to provide structure both for the overview over the general debate and for the discussion of DDFT in the subsequent chapters. These will also present the five problems in more detail. Given that DDFT is a classical theory (although it can be connected to quantum-mechanical methods, see [te Vrugt et al. \(2020b\)](#) for a discussion), I will restrict myself to classical statistical mechanics here.

The first theory that is relevant is thermodynamics, which is a phenomenological theory that describes macroscopic systems² A central observation in thermodynamics is that isolated systems tend to approach a state of *thermodynamic equilibrium*, in which its macroscopic properties are approximately constant. This will, in practice, be the case only on certain observational timescales ([Callender, 2001](#), p. 545). For example, a cup of hot coffee that stands in a cold room will cool down until it has room temperature (equilibrium state).

This is associated with a quantity known as *entropy*, which according to (a common interpretation of) the *second law of thermodynamics* can never decrease in an isolated system. For example, it is not possible that the coffee spontaneously heats up by absorbing heat from the cold room, since in this case entropy would decrease. The tendency to approach equilibrium is often associated with the second law of thermodynamics, although [Brown and Uffink \(2001\)](#) have argued that it has a more fundamental status and should be viewed as a *minus first law*. This leads us to the first problem (Section 4), which is concerned with how the observation that macroscopic systems exhibit irreversible behavior is built into the axioms of thermodynamics.

The second theory we require is classical mechanics, which describes the microscopic dynamics of the individual particles. These are described by *Hamilton’s equations*, which are the fundamental laws of classical mechanics and have the important property of being *time-reversal invariant*: If we record a movie of a process allowed by classical mechanics and play the movie backwards, then what we see will also be a process allowed by classical mechanics. The microscopic laws know no preferred direction of time. The microstate of a system is a point in the so-called *phase space*, which is the central playground of classical mechanics.

Thus, there seems to be a contradiction between the microscopic laws of classical mechanics and the macroscopic laws of thermodynamics. In thermodynamics, systems are expected to evolve towards an equilibrium state, accompanied by an increase of entropy. This process is *irreversible*, since its time-reversal would involve a decrease of entropy and is therefore forbidden by the second law. The laws governing the microscopic constituents of the system, however, are time-reversal invariant and would therefore allow for such a process.

The connection between the behavior of the individual particles and the dynamics of the macroscopic system it consists of is studied in *statistical mechanics*. It can be constructed in two frameworks, which are referred to as the *Boltzmann approach* and the *Gibbs approach*. I present

² By a *macroscopic system*, I mean a system that contains about 10^{23} particles. Strictly speaking, the thermodynamic limit requires that a macroscopic system has infinitely many particles ([Thiele et al., 2019](#)). This idealization, however, leads to additional philosophical problems ([Menon & Callender, 2013](#)) that are not relevant here.

them here following [Frigg \(2008\)](#). In the Boltzmann approach, a system is, on the macroscopic level, characterized by a small number of variables (such as temperature and volume), which define the *macrostate* of the system. Its *microstate*, on the other hand, is characterized by the phase-space coordinates of all particles the system consists of. The microstate uniquely determines the macrostate, while the converse is not true. The *entropy* of a macrostate is then introduced as a measure for the volume of the phase-space region corresponding to this macrostate. One can then introduce the equilibrium state as the state with maximal Boltzmann entropy.

In the framework of Gibbs, one studies many-particle systems using *ensembles*. These are hypothetical sets of infinitely many copies of the system which evolve according to the same laws, but with different initial conditions. One then introduces a probability density that for each point in phase space, i.e., for each microscopic state, gives the probability that a system that is randomly chosen from the ensemble is in this state (see [Frigg \(2008\)](#) and [Frigg and Werndl \(2021, p. 107\)](#)). It is helpful to think of this probability density as a fluid. In this framework, the entropy can then be given a microscopic definition as a function of the density — intuitively, it measures the volume of the fluid. The problem is now that one can easily prove using Hamilton’s equations that the volume of this “fluid” is constant, a result known as *Liouville’s theorem*. This implies, of course, that the entropy is also constant — in conflict with thermodynamics, which demands that it increases during the approach to equilibrium. In the equilibrium state, one has a stationary phase-space distribution that maximizes the Gibbs entropy.

In physics, this problem has not gone unnoticed, and is routinely solved through a procedure known as *coarse-graining*, which has originally been suggested by [Gibbs \(1902\)](#) (see also [Robertson \(2020, pp. 550 - 551\)](#)). The starting point – in usual treatments – is that we are trying to describe macroscopic observations. Such observations do not allow to distinguish between certain microstates — if we shift a certain fluid molecule by a few nanometers, this will change the microstate of a bucket of full of water, but it will not make an observable difference. Therefore, we can simply group together the macroscopically indistinguishable microstates and average over them. Notably, this was done with an epistemic – and not ontological – justification. For the coarse-grained distribution, Liouville’s theorem does not hold. Hence, we can define, from the coarse-grained distribution, a coarse-grained entropy which measures its volume. This entropy is then allowed to increase. The coarse-graining can be mathematically implemented in various ways. I will denote by *coarse-graining* any procedure that involves replacing the microscopic (*fine-grained*) distribution function of the system by an averaged one.

Procedures of this form are routinely and successfully used in physics to derive irreversible macroscopic transport equations from the underlying Hamiltonian dynamics. Nevertheless, they are discussed very controversially in philosophy. The objection is (roughly) that we have, through the coarse-graining, artificially introduced an asymmetry that has not been there originally. This has been justified by an appeal to our limited capability of observation, but the increase of entropy is a physical effect that should not depend on how good our microscopes are ([Robertson, 2020](#), pp. 563 - 565). In fact, this discussion is – as we will see below – concerned with two different problems. On the one hand, we can ask whether the fine-grained or the coarse-grained definition of entropy and equilibrium is the “correct” one (Section 5), on the other hand, we can ask how the coarse-graining procedure can be justified (Section 6).

Finally, we need to find a way to derive – with or without coarse-graining – the irreversible macroscopic dynamics from the reversible microscopic dynamics. In statistical mechanics, there are well-established procedures for this purpose, which are often based on the argument that, for a nonequilibrium initial condition, it is far more likely to move to equilibrium than away from it, since the equilibrium state is the one with the largest phase-space volume ([North, 2011](#),

pp. 321–323). The problem is that such arguments can – because of the time-reversal invariance of the underlying laws – also be applied to the past (Callender, 1999, p. 363). This is a problem because our records of the past tell us that entropy used to be lower. A solution that is very popular in philosophy is the *past hypothesis*. The idea is that, since the asymmetry between past and future cannot be a consequence of the dynamical laws of the universe, it has to be a consequence of the boundary conditions. In particular, if we assume that the entropy of the early universe was very low, we could have an explanation for why it increases afterwards (Callender, 2021).

It is an interesting observation that, despite the fact that the entropy of the early universe plays such a prominent role in foundational discussions of statistical mechanics, it is typically of no importance in the everyday life of a statistical physicist who derives irreversible transport equations. The reason is that we are, once again, dealing with different questions. On the one hand, we can ask ourselves why systems, given an initial nonequilibrium state, tend to approach equilibrium irreversibly (Section 7). On the other hand, we should ask ourselves why the explanation we have found for this effect does not apply to the past (Section 8). Distinctions of this form have been expressed by Boltzmann (see Brown and Uffink (2001, p. 530)), Penrose (1994, p. 218) and Price (1996, pp. 47–48).

In summary, the five problems are:

- **What is the source of irreversibility in thermodynamics? (Q1)**
This question asks for the source of irreversibility *within thermodynamics*, i.e., purely on the level of the macroscopic theory. It is controversial which of the laws of thermodynamics is actually responsible here.
- **How should “equilibrium” and “entropy” be defined? (Q2)**
Here, we look for a definition of the explanandum. The questions “Why do systems approach equilibrium?” and “Why does entropy increase?” are answered differently by persons who have different ideas on what equilibrium and entropy are, which is a common source of confusion.
- **(How) Can coarse-graining be justified? (Q3)**
Since coarse-graining, which is frequently used in explanations of the approach to equilibrium, is a controversial procedure, it needs to be clarified whether and how it can be justified. This is a *separate* problem, since coarse-graining is also used in situations without a relation to thermodynamic irreversibility.
- **Why do systems approach equilibrium? (Q4)**
This question asks why systems, given an initial nonequilibrium state, (irreversibly) approach equilibrium.
- **Why does the arrow of time have the direction it has? (Q5)**
Here, we look for an explanation for why our answer to Q4 does not apply to the past, i.e., we wish to find out why the entropy always increases from past to future.

A popular framework for discussing the microscopic origin of irreversible dynamics is the projection operator method. In this work, I will refer to this approach as *Mori–Zwanzig formalism*, as is common in physics³ (te Vrugt & Wittkowski, 2019). The general idea is to introduce a projection operator to project the full dynamics onto the part that only depends on the “relevant” degrees of freedom of the system. One then derives a formally exact transport equation that contains a term depending on the irrelevant degrees of freedom at

³ Approaches based on projection operators are known by a variety of names. In addition to “Mori–Zwanzig formalism”, one also finds, e.g., “Zwanzig–Zeh–Wallace framework” (Robertson, 2020) – referring to the conceptual discussions of this formalism by Wallace (2015) and Zeh (2007) – “Kawasaki–Gunton operator method” (Yoshimori, 2005) – referring to Kawasaki and Gunton (1973) – or “Mori–Zwanzig–Forster technique” (Wittkowski et al., 2012, 2013) – referring also to Forster (1974). While these names sometimes refer to slightly different forms, the general idea is always the same.

the initial time t_0 and a term depending on the state of the system at previous times. To obtain a closed memoryless transport equation – that is irreversible! – it is then assumed that the irrelevant part of the density vanishes at t_0 and that memory effects can be ignored (Markovian approximation) (Robertson, 2020, pp. 553–556). The original formalism was developed by Nakajima (1958), Zwanzig (1960), and Mori (1965), and has subsequently been extended to incorporate, for example, the dynamics of fluctuations (Grabert, 1978) and time-dependent Hamiltonians (Meyer et al., 2019; te Vrugt & Wittkowski, 2019). An introduction to this method can be found in te Vrugt and Wittkowski (2020a), a general overview in Grabert (1982).

3. Dynamical density functional theory

After having presented and analyzed the discussion of thermodynamic irreversibility in philosophy of physics, we can now turn to an example of an irreversible theory used in modern physics, namely *dynamical density functional theory* (DDFT). Here, I give a brief overview. A detailed review can be found in te Vrugt et al. (2020b).

DDFT is a theory for the nonequilibrium dynamics of classical many-body systems. It is an extension of classical *density functional theory* (DFT), which is a highly successful and formally exact theory (based on the quantum-mechanical DFT by Hohenberg and Kohn (1964)) that allows to find the equilibrium state of a many-particle system. This is done by minimizing a free energy functional that depends on the one-body density $\rho(\vec{r})$ (which gives the probability of finding a particle at position \vec{r}). In DDFT (which is an approximate theory), this is extended to the out-of-equilibrium case by assuming that the nonequilibrium system is driven towards the state in which the free energy functional F is minimized. This leads to the governing equation of deterministic DDFT, given by

$$\frac{\partial}{\partial t} \rho(\vec{r}, t) = \Gamma \bar{\nabla} \cdot \left(\rho(\vec{r}, t) \bar{\nabla} \frac{\delta F[\rho]}{\delta \rho(\vec{r}, t)} \right) \quad (1)$$

with time t and mobility Γ . Here, δ denotes a functional derivative and $\bar{\nabla}$ is the del operator.

The DDFT equation (1) was originally proposed by Evans (1979) based on phenomenological arguments, and later derived from microscopic dynamics by Marini Bettolo Marconi and Tarazona (1999, 2000) and Archer and Evans (2004). DDFT also exists in the stochastic form

$$\frac{\partial}{\partial t} \rho(\vec{r}, t) = \Gamma \bar{\nabla} \cdot \left(\rho(\vec{r}, t) \bar{\nabla} \frac{\delta F[\rho]}{\delta \rho(\vec{r}, t)} \right) + \bar{\nabla} \cdot \left(\sqrt{2\Gamma k_B T \rho(\vec{r}, t)} \bar{\eta}(\vec{r}, t) \right), \quad (2)$$

where k_B is the Boltzmann constant, T is the temperature, and $\bar{\eta}(\vec{r}, t)$ is a multiplicative noise. Stochastic DDFT was, in various forms, developed by Dean (1996), Fraaije (1993), Kawasaki (1994) and Munakata (1989). (The relation between stochastic and deterministic DDFT is discussed in Sections 5 and 6.) Today, DDFT is among the most widely used methods in nonequilibrium statistical mechanics, with applications including biological swimmers (Menzel et al., 2016), disease spreading (te Vrugt et al., 2020a), ions in capacitors (Babel et al., 2018), plasmas (Diaw & Murillo, 2016), thin films (Robbins et al., 2011), tumor growth (Al-Saedi et al., 2018), and much more. Moreover, a variety of extensions have been developed that allow to model, e.g., flow fields (Rauscher et al., 2007), hydrodynamic interactions (Rex & Löwen, 2008), nonspherical and active particles (Wittkowski & Löwen, 2011), nonisothermal systems (Wittkowski et al., 2012), or systems with strict particle order (Wittmann et al., 2021).

Among the main fields of application of DDFT are colloidal and atomic fluids. Colloidal fluids consist of large particles (colloids) that are immersed in a fluid consisting of many small particles (solvent). Due to the presence of the solvent, the motion of the particles is overdamped (i.e., momentum degrees of freedom can be neglected), and subject to noise. The motion of the colloids is described by the Langevin equations

$$\frac{d\vec{r}_i(t)}{dt} = \Gamma \vec{F}_i(t) + \vec{\chi}_i(t), \quad (3)$$

where \vec{r}_i is the position of the i th particle, \vec{F}_i the force acting on it, and $\vec{x}_i(t)$ an additive white noise term with zero mean. As shown by [Marini Bettolo Marconi and Tarazona \(1999\)](#), the DDFT equation (1) can be derived from the Langevin equations (3) based on the assumption that the pair correlation of the nonequilibrium system is identical to that of an equilibrium system with the same one-body density (adiabatic approximation). Atomic fluids, on the other hand, consist of atoms, i.e., of only one type of particle. Their microscopic dynamics is given by Hamilton's equations, which are undamped, rather than by the overdamped Langevin equations. Deriving a DDFT for one-component fluids therefore requires an elimination of the momentum degrees of freedom, which is possible by assuming that the momentum relaxes much faster than the density ([Kawasaki, 1994](#), p. 47).

The Mori–Zwanzig formalism (projection operator method) is an important tool in DDFT, as it allows for a systematic derivation of the DDFT equations (1) and (2) from the microscopic dynamics of the individual particles. (This is discussed in detail in [te Vrugt et al. \(2020b\)](#).) Consequently, the projection operator method allows to relate the physical research on DDFT to the discussion of irreversibility in philosophy of physics. Microscopic derivations of DDFT from the reversible Hamiltonian dynamics were presented by [Español and Löwen \(2009\)](#) and [Yoshimori \(2005\)](#). Moreover, one can obtain DDFT from a stochastic underlying dynamics through additional projections ([Español & Vázquez, 2002](#); [Kawasaki, 1994](#)). An important application of projection operators is the derivation of generalizations of DDFT that incorporate additional order parameter fields. For example, the energy density is used as a relevant variable in addition to the number density in “extended dynamical density functional theory” (EDDFT), which was derived using the projection operator formalism by [Wittkowski et al. \(2012, 2013\)](#).

4. First problem: The source of irreversibility in thermodynamics

As discussed in Section 2, entropy and irreversibility are studied in two fields of physics: Thermodynamics, which is a phenomenological theory of macroscopic states, and statistical mechanics, which describes macroscopic systems by developing a statistical description of their microscopic constituents. One of the aims of statistical mechanics – though by no means the only one – is to provide a microscopic justification for the principles that are introduced as axioms on the level of thermodynamics. Here, we are interested in a particular property of thermodynamics, namely the existence of irreversible processes. If we wish to figure out where in the transition from microscopic to macroscopic physics these come into play, a good way to start is therefore to first figure out in which of the axioms of thermodynamics irreversibility can be found, and then to have a look at the microscopic foundations of that particular axiom.

[Uffink \(2001, pp. 313–315\)](#) attributes the difficulty of locating the temporal asymmetry in thermodynamics to the absence of equations of motion. A typical understanding of time-reversal symmetry is that a theory has this symmetry if, when we time-reverse a temporal evolution that is allowed by the theory, the resulting (backwards) evolution would also be allowed by the theory. It is not immediately clear how to apply this definition to a theory that is not primarily concerned with time evolutions.

Thermodynamics is based on four axioms (“Hauptsätze”), namely the zeroth law (transitivity of thermal equilibrium), the first law (conservation of energy), the second law (entropy cannot decrease) and the third law (a system cannot reach zero temperature). In physics textbooks, the second law typically gets credit for irreversibility. This law was stated in Section 2 as “Entropy cannot decrease in a closed system”. Although this is a very common understanding, it is very difficult to state the content of the second law in a precise way since it exists in many different forms that are not fully equivalent and have differing connections to the arrow of time. [Uffink \(2001, p. 306\)](#), who has analyzed this point in detail, has compared the interpretation of the second law to the interpretation of a work from Shakespeare.

The relation of the second law to irreversibility depends, as shown by [Uffink \(2001\)](#) in a detailed historical analysis, on its precise formulation. In an influential article, [Brown and Uffink \(2001\)](#) have argued that the temporal asymmetry of thermodynamics is more fundamental than and logically prior to the second law. Therefore, they propose a “minus first law” that should be added to the standard set of laws. It states that for each isolated system there exists a unique state of equilibrium that it will spontaneously enter. The time-asymmetry of *thermodynamics* (in contrast to statistical mechanics) then arises from the notion of equilibrium states, since these are spontaneously reached but not left without external intervention.

The distinction between the second and the minus first law, and the question where exactly irreversibility is to be located, is very important for providing a microscopic foundation of irreversible processes. [Uffink \(2001, p. 316\)](#) distinguishes between different meanings of “irreversibility”, namely “time-reversal-noninvariance” (there is a process allowed by a theory whose time-reversal is not allowed) and “irreversibility” (a transition from a state A to a state B cannot be fully undone). This distinction has also been employed by [Luczak \(2017\)](#). He claims that minus first and the second law are logically independent statements. The minus first law predicts that systems initially in a nonequilibrium state will irreversibly approach an equilibrium state, where “irreversible” means “time-reversal-noninvariant”. The second law, on the other hand, is concerned with transitions between equilibrium states, which can be irreversible in the sense of “making the initial state irrecoverable”. Statistical mechanics, he concludes, should aim at giving a microscopic underpinning to both the minus first and the second law, which requires two different solutions given that we have two different problems.

Clarifying the location of irreversibility within thermodynamics is thus important in order to decide what precisely we seek an explanation for. Since in this work we are interested in what DDFT can offer, we should therefore ask whether DDFT is related to the second or to the minus first law. As shown by [Munakata \(1994, p. 2348\)](#), the deterministic DDFT (1) possesses the H-theorem

$$\frac{dF}{dt} = \int d^3r \frac{\delta F}{\delta \rho(\vec{r}, t)} \frac{\partial}{\partial t} \rho(\vec{r}, t) = - \int d^3r \Gamma \rho(\vec{r}, t) \left(\vec{\nabla} \cdot \frac{\delta F}{\delta \rho(\vec{r}, t)} \right)^2 \leq 0. \quad (4)$$

In the last step, we have used Eq. (1) and integration by parts. A minimum is reached for $\delta F / \delta \rho = \mu$ (with chemical potential μ), which is the equilibrium state. Thus, DDFT predicts that nonequilibrium systems spontaneously and monotonously approach an equilibrium state. Moreover, the (local and global) minimum is unique, although local minima can arise in practical applications as a consequence of approximations ([Marini Bettolo Marconi & Tarazona, 1999](#), p. 8036). Consequently, DDFT provides a microscopic foundation for the minus first law (for fluids).

5. Second Problem: The Definition of “Equilibrium” and “Entropy”

If we wish to explain something, it is, in general, important to have a precise idea about what exactly the thing we need to explain is. In the case of thermodynamics, explaining the increase of entropy or the approach to equilibrium therefore requires that we have an idea about what precisely entropy and equilibrium are. As it turns out, this is not at all clear, which is problematic because different opinions on what equilibrium is lead to different opinions on why and whether equilibrium is approached. Here, I will consider two aspects of the debate concerning thermodynamic equilibrium, which can both be understood better when analyzing the example of DDFT. These are the question whether “entropy” (and, consequently, “equilibrium”) is an approximate notion, i.e., whether it should be understood in a fine-grained or a coarse-grained way, and the question whether it is a time-asymmetric notion, i.e., whether fluctuations away from equilibrium are possible.

For the first of these aspects, we consider the argument formulated by Blatt (1959) and Ridderbos and Redhead (1998) against the coarse-graining procedure, which is based on the famous spin echo experiment performed by Hahn (1950).⁴ Spins in a magnetic field are aligned by a strong pulse. The magnetic field causes precession of the spins, and because of small inhomogeneities in the field, the spins precess with different frequencies and are oriented “randomly” after some time. Hence, they are in a state of coarse-grained equilibrium. Now, a second pulse is applied, which reverses the direction of precession. In consequence, the spins return, after some time, to their initial aligned state. Since the spins were isolated after the second pulse, we appear to have a case of an isolated system spontaneously evolving away from equilibrium, which could be interpreted as a violation of the second law. The explanation for why this happens and why this is *not* a violation of the second law, Ridderbos and Redhead (1998, p. 1251) argue, is that the system was not actually in fine-grained equilibrium. It was only in coarse-grained equilibrium. But since the typical argument for coarse-graining is that we cannot distinguish between fine-grained and coarse-grained equilibrium, and since we here have found a way in which we can, coarse-graining is not a reasonable way of explaining the approach to equilibrium. In particular, if we think of “equilibrium” as “coarse-grained equilibrium”, the second part of the spin-echo experiment would constitute a violation of the second law of thermodynamics (which, of course, it is not).

I will get back to the justification of coarse-graining in Section 6. What is interesting here is that Ridderbos and Redhead defend a certain characterization of equilibrium as “fine-grained equilibrium”. They contrast their position with that of Sklar (1995, p. 253), who claims that the first part of the spin-echo experiment constitutes “normal” thermodynamic behavior. Ridderbos and Redhead (1998, pp. 1254–1257), on the other hand, argue that thermodynamic behavior would be an approach to “true” equilibrium, including the loss of correlations. On this perspective, an isolated gas that spreads out in a container would not exhibit an approach to equilibrium, since its distribution would not turn into the uniform equilibrium distribution. This, they conclude, requires external interventions.

But we do not necessarily need to use this definition of equilibrium. Maybe, we are perfectly happy with the weaker characterization of “equilibrium” as a state in which the values of certain macroscopic quantities are approximately constant (Callender, 2001, p. 547). For example, we macroscopically observe that hot coffee cools down. When we, starting from a microscopic theory, derive a prediction for how the average temperature of the coffee evolves in time, we get, after some approximations, just that: The temperature of the coffee will approach room temperature. These two definitions thus lead to very different meanings of Q4 (“Why do systems approach equilibrium?”). If we use fine-grained equilibrium as a definition, combined with the very strict reading proposed by Ridderbos and Redhead, then the explanation for why gas expands in an isolated container, or why an isolated cup of coffee cools down, would not be a part of an answer to Q4. Using a definition based on constant macroscopic quantities, on the other hand, we would seek for an explanation of precisely this behavior.

DDFT is an interesting example for the problem of the approximate nature of thermodynamic equilibrium: As discussed in Section 3, static DFT allows to find equilibrium states of a many-particle system by minimizing its free energy functional. It is based on a theorem by Mermin (1965) stating that *in equilibrium*, the one-body density uniquely determines the phase-space distribution and thus the free energy. Thereby, DFT provides a significant computational advantage, since the one-body is only a function of the position and not of the phase-space coordinates of all particles. This theorem does, however, no longer hold out of equilibrium. An extension was proven by Runge and Gross

⁴ I explain it here in a simplified form, following Ridderbos and Redhead (1998, pp. 1242–1243).

(1984) for quantum and by Chan and Finken (2005) for classical time-dependent systems, which shows that in the nonequilibrium case, there is also a functional dependence on the initial state of the system. Consequently, the one-body density does not fully determine the fine-grained free energy for a system that is initially out of equilibrium. Therefore, DDFT describes – strictly speaking – only an approach to coarse-grained equilibrium. If we are in a state that has the same one-body distribution as the equilibrium state, but a different correlation function – i.e., in a state that corresponds to coarse-grained, but not fine-grained equilibrium – then DDFT will not predict that it relaxes, since the right-hand side of Eq. (1) is zero in this case. Problems of this form are relevant, e.g., for the glass transition (Heinrichs et al., 2004, p. 1124).

Given the approximate nature of DDFT, a significant amount of research has been devoted to analyzing its shortcomings and potential improvements in the past years. Forces that are not incorporated in DDFT are known as *superadiabatic forces*.⁵ As shown by Fortini et al. (2014), superadiabatic forces can have important effects in systems of Brownian particles. A systematic extension of DDFT that allows to describe them is *power functional theory* (PFT), which was developed by Schmidt and Brader (2013). This exact formalism is based on a variational principle for the power dissipation, where contributions that go beyond DDFT are incorporated in an “excess power functional”. An important superadiabatic effect is memory: The Mori–Zwanzig formalism shows that, when the full microscopic description of a system is reduced to the subdynamics of the relevant degrees of freedom, one obtains a dependence on the history of the system (Treffenstädt & Schmidt, 2020, p. 1518). This dependence is usually neglected in derivations of irreversible dynamical theories, this holds both for the Mori–Zwanzig formalism and for DDFT.

The argument from interventionists outlined above is based on the objection that the history dependence cannot be neglected, since it is responsible for the motion reversal observed in the second part of the spin-echo experiment. Similar effects have also been found in PFT: Treffenstädt and Schmidt (2020) have recently studied a system of Brownian particles under shear using PFT and Brownian dynamics simulations. Starting from equilibrium, an inhomogeneous shear force field is switched on, such that the system settles to a nonequilibrium steady state. When the shear is switched off again, the system exhibits a global current reversal before approaching equilibrium. This is a consequence of the memory-induced superadiabatic forces, which balance the external forces in steady state and become driving forces after the external forces are switched off. This effect is reminiscent of the spin-echo effect,⁶ which indicates that the large amount of work on superadiabatic forces that has emerged in the past years is also relevant for foundational research.

The comparison of fine-grained and coarse-grained equilibrium and entropy leads us to the second aspect if we consider the connection to thermodynamics (as discussed in Section 4): If we define “equilibrium” as “the state with maximal fine-grained entropy”, then the minus first law is, due to Liouville’s theorem stating that the fine-grained entropy is constant, simply wrong (recall that the minus first law is a statement about closed systems). For approaches based on coarse-graining, on the other hand, the minus first law does not pose a problem. Brown and Uffink (2001, p. 530) therefore distinguish in their discussion two different notions of equilibrium: In thermodynamics, the concept

⁵ This name is motivated by the fact that DDFT is based on the adiabatic approximation, see Section 3.

⁶ I do not claim here that this is the same effect — there are important differences, such as the fact that the spins are a closed quantum system, while the Brownian particles are a dissipative classical system. Nevertheless, the general effect (a system exhibits a motion reversal rather than the monotonous approach to equilibrium simpler models would predict) is similar, and given that philosophers of physics tend to be more interested in closed quantum systems, dissipative classical systems offer a lot of unexplored potential for them.

of equilibrium is introduced through the minus first law. Here, it is a state that the system spontaneously enters and then remains in. Therefore, equilibrium is by its nature a time-asymmetric notion. In (Boltzmannian) statistical mechanics, on the other hand, equilibrium is the macrostate with the largest entropy or phase-space volume, i.e., the state that can be realized with the largest number of microscopic configurations. This definition is time-symmetric, since deviations are, in principle, possible in both directions of time.

Deterministic DDFT does, as shown in Section 4, predict a monotonous approach to an equilibrium state, and thus provides a microscopic basis for the minus first law (for soft matter systems). Once the system has reached a minimum of the free energy, it will stay there forever. This is not the case in the stochastic DDFT (2): Here, due to the presence of the noise term, there will still be fluctuations once an equilibrium state is reached. This is more reminiscent of the time-symmetric idea from statistical mechanics, where deviations from equilibrium are allowed. However, both stochastic and deterministic DDFT are microscopic, statistical–mechanical theories. The question whether noise terms should be included in DDFT has led to an intense debate in statistical physics. An important argument against the presence of noise terms, presented by Marini Bettolo Marconi and Tarazona (1999, p. 8034), is that the free energy functional of DFT already includes all fluctuations such that one would overcount them if they are added to the evolution equation.

An explanation was provided by Archer and Rauscher (2004), who showed that the difference between Eqs. (1) and (2) arises because the density ρ has a different meaning in both theories. As an example, let us compare the derivation of stochastic DDFT by Dean (1996) and the derivation of deterministic DDFT by Marini Bettolo Marconi and Tarazona (1999). Dean (1996) considered the microscopic density

$$\hat{\rho}(\vec{r}, t) = \sum_{i=1}^N \delta(\vec{r} - \vec{r}_i(t)), \tag{5}$$

where $\delta(\vec{r})$ is the Dirac delta distribution and N is the number of particles. (The density $\hat{\rho}$ is often denoted “density operator”, and is written with a hat to distinguish it from averaged densities.) He then showed that, if the individual particles obey the Langevin equations (3), the density (5) satisfies the exact evolution equation

$$\frac{\partial}{\partial t} \hat{\rho}(\vec{r}, t) = \Gamma \vec{\nabla} \cdot \left(\hat{\rho}(\vec{r}, t) \vec{\nabla} \frac{\delta F[\hat{\rho}]}{\delta \hat{\rho}(\vec{r}, t)} \right) + \vec{\nabla} \cdot \left(\sqrt{2\Gamma k_B T} \hat{\rho}(\vec{r}, t) \vec{\eta}(\vec{r}, t) \right), \tag{6}$$

with the free energy

$$F[\hat{\rho}] = \frac{1}{2} \int d^3r \int d^3r' \hat{\rho}(\vec{r}, t) U_2(\vec{r} - \vec{r}') \hat{\rho}(\vec{r}', t) + k_B T \int d^3r \hat{\rho}(\vec{r}, t) (\ln(\Lambda^3 \hat{\rho}(\vec{r}, t)) - 1), \tag{7}$$

Here, $U_2(\vec{r})$ is a two-body interaction potential and Λ is the (irrelevant) thermal de Broglie wavelength. Notably, the free energy (7) is *not* the free energy of DFT, i.e., of equilibrium statistical mechanics. In particular, the interaction term is very complicated and not known in general in DFT, whereas it can be constructed straightforwardly from the interaction potential in Eq. (7). Marini Bettolo Marconi and Tarazona (1999), on the other hand, studied the ensemble-averaged density

$$\rho(\vec{r}, t) = \left\langle \sum_{i=1}^N \delta(\vec{r} - \vec{r}_i(t)) \right\rangle, \tag{8}$$

where $\langle \cdot \rangle$ is an average over the microscopic noise. This density can be shown to follow (after an adiabatic approximation, see Section 3) the deterministic equation (1). (Another form of stochastic DDFT was obtained by Kawasaki (1994), this will be discussed in Section 6.)

This distinction is of central importance for theoretical soft matter physics, but it is also interesting for the discussion in philosophy of physics: Here, it is (as shown above) a matter of debate to which extent the monotonous approach to equilibrium predicted by thermodynamics

is a consequence of approximations, and whether systems will be found to fluctuate out of equilibrium. Comparing the two forms of DDFT discussed here shows that this depends, to a very large extent, on the way we define our observables (here ρ) and our thermodynamic functionals (here F), in particular on whether they are understood as ensemble averages. This point deserves much more attention in foundational discussions than it typically gets.

It also relates in an interesting way to the distinction between the Gibbs and the Boltzmann approach explained in Section 2: As discussed by Uffink (2007), these frameworks differ in their conceptions of equilibrium: For Boltzmann, it is a property of a single system. A system that is in equilibrium is not guaranteed to stay there forever. For Gibbs, on the other hand, the idea of equilibrium applies to ensembles, and an ensemble that is in equilibrium will remain there forever (since equilibrium corresponds to a stationary distribution). In DDFT, one has an equilibrium state which the density remains in forever if and only one considers the ensemble-averaged density — this is the case of deterministic DDFT, which is then analogous to the Gibbsian approach. In stochastic DDFT, on the other hand, the system may always fluctuate away from the state that minimizes the free energy, and the density is not to be understood as an ensemble average.

6. Third problem: The justification of coarse-graining

The third problem is the question whether and how we can justify coarse-graining. It is an important point here that “Can we use coarse-graining to explain the approach to equilibrium?” and “How can we use coarse-graining to explain the approach to equilibrium?” are two very different questions. While the second one requires, of course, that the first one has been answered with “yes”, it has a different content.

A first point we need to make here is that, to not get things mixed up, it is important to distinguish “coarse-graining”, which is a mathematical technique, from the position of “the coarse-grainers”, which is the aim of criticism from interventionists. If, in the philosophical literature, someone criticizes “the coarse-grainers”, the criticism is usually aimed at people who use a coarse-grained definition of equilibrium, typically combined with an epistemic justification based on finite measurement resolutions (Ridderbos & Redhead, 1998). This criticism, however, is *not* directed at the procedure of coarse-graining as such (at least it should not be), it is directed at a particular justification of coarse-graining or at a particular definition of “equilibrium”. The method of coarse-graining, in particular in the form of projection operators, can also be used to formally describe the influence of external interventions, and thus form a basis for interventionism (Robertson, 2020, p. 556). Since “the position of the coarse-grainers” is just one possible justification of the mathematical procedure of coarse-graining, and since this mathematical procedure is used very successfully throughout physics, one should be very careful with rejecting it altogether, and if one does, this requires a very good justification.

To see why Q3 is a problem on its own, it is helpful to compare the way coarse-graining is used and discussed in physics and in philosophy. Philosophers of physics are interested in coarse-graining because of the role it plays in the explanation of the approach to equilibrium.⁷ Physicists, on the other hand, also use it in situations where the approach to equilibrium is of no interest or not even expected. DDFT is an excellent case study for this point, which receives little attention in the foundations of physics. It is a coarse-grained theory that, in its standard form, describes the irreversible approach to thermodynamic equilibrium. However, modified forms of DDFT that are derived by

⁷ I do not claim here that philosophers always believe that coarse-graining is only applicable to thermodynamic irreversibility, the relation of different levels of description is also studied in a more general context (List, 2019). Nevertheless, in the philosophy of statistical mechanics, coarse-graining is typically discussed as a possible origin of thermodynamic irreversibility, which differs from the way it is used, e.g., in active matter physics.

similar coarse-graining procedures do *not* describe the approach to thermodynamic equilibrium. Hence, interesting insights can be gained by comparing these different forms of DDFT.

A field of application of DDFT where equilibration is not observed is *active matter*. Active particles are particles which use energy in order to create directed motion. A typical example would be a swimming bacterium. Systems consisting of active particles are permanently out of equilibrium, i.e., they do (as long as there is enough energy available) never reach or approach a state of thermodynamic equilibrium. In the description of active particles in statistical mechanics, coarse-graining methods are frequently used to derive a macroscopic field theory based on a microscopic theory of the individual active particles (see, e.g., Bickmann and Wittkowski (2020a, 2020b) for such a derivation). The theories derived in this way provide good descriptions for the dynamics of active matter systems.

A DDFT for active particles can be constructed in various ways. Wensink and Löwen (2008) described uniaxial⁸ active particles by adding to a DDFT for passive (non-active) nonspherical particles a term accounting for self-propulsion. This was generalized to particles with arbitrary shape by Wittkowski and Löwen (2011). An alternative approach, considered, e.g., by Enculescu and Stark (2011) and Wittmann and Brader (2016), is to construct a DDFT equation of the form

$$\frac{\partial}{\partial t} \rho(\vec{r}, t) = \Gamma \vec{\nabla} \cdot \left(\rho(\vec{r}, t) \vec{\nabla} \frac{\delta F_{\text{eff}}[\rho]}{\delta \rho(\vec{r}, t)} \right). \quad (9)$$

Although Eq. (9) looks exactly like Eq. (1), it uses an “effective” free energy F_{eff} rather than the equilibrium free energy F . In the construction of F_{eff} , the external potential acting on the system is modified by the steady-state active force. Since active systems are not in equilibrium, the DDFT (9) shows that a coarse-grained theory that describes the approach to a stationary state does not necessarily also describe the approach to thermodynamic equilibrium.

The justification of coarse-graining can then proceed in various ways. One option is to justify the replacement of the fine-grained by a coarse-grained density by our ignorance about microscopic details. Robertson (2020, pp. 565–570) justifies coarse-graining by its ability to reveal autonomous dynamics on a higher level of description. Notably, both justifications do not require a connection between coarse-graining and equilibrium and are thus applicable more generally. We can also be indifferent about microscopic details of an active system, and active systems also show autonomous dynamics on higher levels of description.

Robertson (2020, p. 561) has suggested to split the project of justifying coarse-graining into two further sub-problems: We need to justify both why we coarse-grain *at all* and why we coarse-grain *in a particular way*. The answers to these questions might be linked, but do not have to be identical. Up to now, we have been concerned with the first sub-problem. We now turn to problem of why one should coarse-grain *in a particular way*. Again, it is useful to study DDFT.

As discussed in Section 3, DDFT exists in stochastic and deterministic forms. In Section 5, we have compared the deterministic DDFT by Marini Bettolo Marconi and Tarazona (1999) to the stochastic DDFT by Dean (1996), and noted that the former describes the ensemble-averaged density, while the latter is concerned with the microscopic density operator. Another form of stochastic DDFT was derived by Kawasaki (1994). In this theory, the variable $\rho(\vec{r}, t)$ denotes a density after spatial coarse-graining, which is done by dividing space into many small cells. As discussed by Archer and Rauscher (2004), the different definition of the density in Kawasaki’s theory is the reason why it is governed by a stochastic rather than a deterministic theory (since the average is taken over space and not over realizations of the noise, the noise does not vanish after averaging). Since the presence or absence of noise terms does, as emphasized in Section 5, make a

difference for the way equilibrium is approached and for the resulting equilibrium state, the precise way in which coarse-graining is done should be paid attention to when discussing irreversibility.

More generally, the relation between stochastic and deterministic approaches can also be understood from the Mori–Zwanzig formalism, which, as mentioned in Section 3, allows to derive DDFT. The Mori–Zwanzig formalism exists in various forms, which differ in the form of the projection operator (Kawasaki, 2000, p. 6345): In a *microcanonical* projection operator formalism, the values of the macroscopic variables are specified exactly. When using a *canonical* projection operator, on the other hand, one only specifies the average values of the relevant variables. A canonical projection operator allows to derive deterministic DDFT, with the free energy F being given by the equilibrium free energy of DFT. With a microcanonical projection operator, one can derive the stochastic form of DDFT. The resulting free energy is not identical to the DFT free energy, a connection can be established using a fluctuation renormalization (Kawasaki, 2006; te Vrugt et al., 2020b). Philosophical discussions of the Mori–Zwanzig formalism do usually not distinguish between different types of projection operators. However, the example of DDFT shows that this is more than a technical difference, since different projection operators lead to different forms of the dynamic equations that, as shown in Section 5, describe the approach to equilibrium in a different way.

7. Fourth problem: The (irreversible) approach to equilibrium

While the first three problems have been setting the stage, we now dive more deeply into what is usually thought of as “the problem of irreversibility”: We wish to explain why systems that are initially out of equilibrium tend to move towards an equilibrium state.

In physics, derivations of irreversible equations from the reversible microscopic physics have a long tradition, starting from Boltzmann’s H-theorem (see Boltzmann (1872), Brown et al. (2009)). “H-theorem” is, in fact, now a standard name for a proof that a certain theory leads to a monotonous behavior of (for example) the entropy (Anero & Español, 2007; Español & Löwen, 2009). Here is a typical idea of how this can work (North, 2011, pp. 321–323): As discussed above, the entropy is, in statistical mechanics, introduced as a measure for the volume in phase space. Equilibrium is the macrostate with the largest phase space volume.⁹ In fact, the phase space volume of the equilibrium state is overwhelmingly larger than that of other states. Hence, if we move into some direction in phase space, it is far more likely that we are going towards equilibrium than away from it (the latter result would correspond to extremely special initial conditions). The approach to equilibrium is therefore explained by statistical considerations. This argument goes back to Boltzmann, and notably, it is a *probabilistic* argument (Brown et al., 2009, pp. 185–187). In the Gibbsian framework, it can be shown that the *coarse-grained* distribution function tends towards a uniform distribution on phase space.

Today, a popular framework for the description of irreversible processes is *stochastic dynamics* (Uffink, 2007, pp. 1038–1063). Here, the laws governing the behavior of the constituents of a system are assumed not to be deterministic, but stochastic. In particular, a specific type of stochastic processes known as *Markov processes* is employed (Sober, 2020). A process is Markovian if it has no memory. This means that (the probability of) what happens in the future does not depend on what has happened in the past. For example, the probability for getting “heads” in a fair coin toss is always 0.5, no matter how often the coin landed on “heads” in previous coin tosses. The Markov processes employed in stochastic dynamics have the attractive feature that they tend towards an equilibrium distribution after a while, so if we describe a system in statistical mechanics by Markovian dynamical equations, then chances are good that we will obtain a description of the irreversible approach

⁸ A particle is uniaxial if it has an axis of continuous rotational symmetry (te Vrugt & Wittkowski, 2020b).

⁹ I assume here, for the sake of the argument, that these are acceptable definitions of “entropy” and “equilibrium”.

to equilibrium. (Mathematically, the description in terms of Markov processes corresponds to using so-called “Master equations”, which can be obtained as an approximation of the actual microscopic dynamics within the Mori–Zwanzig formalism (Zeh, 2007, pp. 57–68).)

The question is, of course, why we should be justified to model a system using stochastic Markov processes given that the actual dynamics is governed by Hamilton’s equations, which are deterministic. Three main viewpoints can be found in the literature (Uffink, 2007, pp. 1038–1039). The first one is an appeal to coarse-graining (van Kampen, 2002): We ignore microscopic details of the system that we cannot measure anyway and focus on macrostates. Then, we can find transition probabilities for the change from one macrostate to another one and thus obtain a stochastic process. Another option is interventionism (Blatt, 1959; Ridderbos & Redhead, 1998): The system we describe is not actually closed, but subject to external influences. Since these are not known, our description of the system is stochastic. Finally, some authors have advocated being agnostic about the origin of stochasticity and simply treating stochastic dynamics as fundamental (Streater, 2009).

Importantly, these approaches are not necessarily in contradiction: Recall that on the interventionist’s definition, the facts that gas in an isolated container spreads out, that a hot cup of coffee in an isolated room acquires room temperature, or that the spins go out of phase in the first part of the spin echo experiment do not constitute instances of an approach to equilibrium. Nevertheless, they do not deny that these things happen. Now let us have a look at how these effects are explained in the physics literature, where coarse-graining is employed. An illustration used by Donev et al. (2014, p. 14) is a system consisting of many oscillators that, after some time, will go out of phase. Every individual oscillator follows a time-reversible law, yet the average amplitude will decay due to dephasing. This, remarkably, is exactly the explanation that Ridderbos and Redhead (1998, p. 1242) use for what happens in the first part of the spin-echo experiment — the spins oscillate and go out of phase. Thus, there is no difference between the physical mechanisms by which interventionists explain the spread out of an isolated gas and the mechanisms by which a physicist using coarse-graining would explain the same effect. The major difference is that for an interventionist, “approach to equilibrium” means something different, such that different mechanisms (namely external interventions) are required to explain it. Therefore, interventionists can still agree that the coarse-grained density of a closed system will tend to a uniform distribution, they will just not classify this as an “approach to equilibrium”.

Again, DDFT provides an instructive example. As discussed in Section 3, one can distinguish between DDFT for colloidal fluids and for atomic fluids. In the case of atomic fluids, the microscopic dynamics is given by Hamilton’s equations, and is therefore reversible. Hence, a derivation of the irreversible DDFT equation (1) requires a coarse-graining procedure. A good example is the derivation of DDFT from the Mori–Zwanzig formalism by Español and Löwen (2009): Starting from the reversible Hamiltonian dynamics, one obtains an exact equation of motion for the relevant variable (here the one-body density) that involves a memory term. It is then assumed that the one-body density is a slow variable (more precisely: that it changes in time much more slowly than the current correlation function), such that memory effects can be neglected. This leads to the irreversible DDFT equation¹⁰ (1). Thus, it is a typical example for the type of derivation “coarse-grainers” have in mind when discussing irreversibility: We start from a reversible system, coarse-grain it, make the approximation that the relevant part of the dynamics is Markovian, and arrive at an irreversible law.

¹⁰ Strictly speaking, it leads to a non-local generalization of Eq. (1), that has to be made local by the additional assumption that the system is a dilute suspension. This, however, is a technical problem that is not relevant for the present discussion.

However, DDFT is also (in fact: more often) applied to colloidal fluids. Here, the microscopic dynamics is given by the Langevin equations (3). Since these are already asymmetric in time (Luczak, 2016, p. 405), coarse-graining is not required to obtain irreversibility. Stochasticity here comes into play through the noise term in Eq. (3). As shown by Dean (1996), one can derive from Eq. (3) an exact irreversible DDFT equation of the form (2), which describes an approach to equilibrium. (The additional approximations employed by Marini Bettolo Marconi and Tarazona (1999) are required to obtain a closed equation for the ensemble-averaged one-body density.) This is a typical example of a derivation that interventionists think of when discussing irreversibility: The colloidal system approaches equilibrium because it is – due to collisions of the colloids with the solvent particles, which act as a heat bath and provide noise – not a closed system. Consequently, both main approaches to stochastic dynamics – coarse-graining and interventionism – are required to understand modern nonequilibrium statistical mechanics, since both play a role in DDFT.

What is, moreover, interesting is that there is not necessarily a disagreement between these two positions: The irreversible Langevin equations (3) can be derived by coarse-graining the full (reversible) microscopic equations of motion for the system consisting of both colloidal and solvent particles. What interventionists and coarse-grainers would disagree on is whether an atomic fluid or a closed system consisting of colloids and solvents can be said to approach equilibrium. However, as shown above, this is merely a consequence of the way “equilibrium” is defined in these two approaches.

8. Fifth problem: The arrow of time

We have now reached the final problem, which is concerned with the *arrow of time*. Let us assume we have found some explanation for why, if we go forwards in time, the systems are observed to approach equilibrium, i.e., that we have answered question four. We now need to find out why, despite the symmetry of the fundamental laws of physics, this answer cannot be applied to the past.

The distinction between Q4 and Q5 allows us to understand why, as mentioned in Section 2, philosophers and physicists concerned with the origin of irreversibility are discussing it based on considerations about the big bang and cosmology, while researchers working on nonequilibrium statistical mechanics are successful without ever spending a thought on these issues (Penrose, 1994). The methods employed in statistical physics – coarse-graining and Markovian approximations – allow for a solution of Q4. This is sufficient if one takes the temporal asymmetry as given, which is why the early universe plays no role in the everyday life of a statistical physicist. Philosophers (and physicists) who are interested in the source of this asymmetry, on the other hand, need to address Q5, which requires a different type of answer.

Recall how the standard solution discussed in Section 2 works: We start from a reversible theory describing the dynamics of the density and then perform a coarse-graining, in which we replace the fine-grained by the coarse-grained density, allowing to define a coarse-grained entropy. The dynamics of the coarse-grained density then leads to an increase of the coarse-grained entropy. Unfortunately, the original microdynamics we started with was time-reversal symmetric, i.e., it does not know a difference between past and future. Hence, we could have simply decided to perform the coarse-graining in the other direction of time, which would have been mathematically just as well justified as the original derivation. This “backwards coarse-graining” would then lead to a theory in which the entropy decreases in the future and increases in the past, in strong contradiction to every observation (Wallace, 2011, p. 15).

Albert (2000, pp. 115–119) has explained this problem as a problem of inference: Q4 is, as discussed in Section 7, frequently solved by arguing that an evolution towards equilibrium is more likely on statistical grounds. Albert approaches this from an epistemological perspective: In principle, he argues, we have two ways of obtaining information. *Prediction* means that we, from the known laws of physics and the present

state of the world as an initial condition, infer what will happen in the future. Doing the same for the past is called *retrodiction*. The alternative is to use *records*, from which we obtain most of our knowledge about the past. The question is now, Albert argues, why we should have sources of information other than prediction or retrodiction given that we only have direct reliable empirical information about the present state of the world. Unfortunately, if we use classical mechanics combined with statistical considerations about phase space to retrodict what happened in the past based on our knowledge of the present, we will conclude – by the line of argument employed in Section 7 – that the past had a higher entropy than the present. Hence, most of our records would most likely be wrong. Even worse, given that we only believe in the laws of mechanics because of experiments that we have records of, we could infer from classical mechanics that classical mechanics is, presumably, also wrong. Thus, statistical mechanics has brought us into a position of skepticism. To avoid this, he introduces an additional postulate, the *past hypothesis*: The entropy of the initial state of the universe was very low. If we then use, as a basis for our inference, not only the laws of mechanics and the statistics of phase space, but in addition also the past hypothesis, then we can believe that our records of the past are, in general true. Thus, we have avoided Albert's problem of skepticism.

Wallace (2011) has suggested an alternative formulation,¹¹ the *simple past hypothesis*: He defends the position that coarse-graining is justified if the density is forwards compatible, i.e., if it does not affect the predictions for the macroscopic variables. This is satisfied if the density is “simple”.¹² Unfortunately, a simple distribution is also backwards compatible, i.e., we can perform a coarse-graining in the other direction of time, predicting an (unphysical) increase of entropy in the past. To avoid this, and therefore to answer Q5, we assume that the initial state of the universe was simple. In this case, we cannot perform a problematic backwards coarse-graining, since the initial state of the universe has no past.

DDFT is an interesting example also for this problem. As mentioned in Section 7, it is possible in the Mori–Zwanzig formalism to derive the irreversible DDFT equation from the reversible laws of classical mechanics, a theory where the entropy always increases. However, this only answers Q4 — depending on the choice of initial conditions, it is also possible in the Mori–Zwanzig formalism to derive a theory in which entropy always *decreases* (Zeh, 2007, p. 67). This depends on the choice of initial conditions: To obtain a closed equation of motion in which the entropy increases, we have to assume that the irrelevant part of the phase-space density vanishes at the initial time (Wallace, 2015, p. 292). This assumption is therefore also required for derivations of DDFT in the Mori–Zwanzig formalism. Unfortunately, this is not always made clear in such derivations. An explicit discussion of this hypothesis and its role in DDFT was given by Yoshimori (2005).

When setting up the initial condition in a different way, the irreversible dynamics would apply backwards in time. This possibility is, in the philosophy of statistical mechanics, usually dismissed as an unphysical artifact that one has to get rid of. However, it is notable that a “backwards coarse-grained theory” is sometimes actually employed in statistical physics (Lutsko, 2019): In the theory of nucleation, where particles in a fluid aggregate to a nucleus that then grows leading to solidification, one is interested in calculating the “most likely path” along which nucleation occurs. For nucleation, the system has to overcome an energy barrier. The most likely path can be obtained by starting at the top of the energy barrier, evolving the system forwards in time by the (coarse-grained and irreversible) DDFT equation towards the final state and backwards in time (!) towards the initial state which had a lower free energy. Lutsko (2011, p. 2), who introduced this procedure, emphasizes that it is only a mathematical trick. Nevertheless, it is an interesting observation that this is mathematically possible and occasionally useful, and it is an interesting question why this is only a mathematical trick.

¹¹ See Brown (2017) for a comparison of Albert's and Wallace's formulation.

¹² See Wallace (2011) for a more mathematical discussion.

9. Conclusion

In summary, I have shown that the study of dynamical density functional theory (DDFT) provides interesting and important insights for the philosophy of statistical mechanics. This has been demonstrated for each of the five “problems of irreversibility” discussed here: DDFT provides a microscopic basis for the minus first law (Q1), it reveals the importance of taking into account memory effects and of distinguishing between various types of free energy (Q2), it illustrates the subtle relation between irreversibility and coarse-graining as well as the fact that different forms of coarse-graining lead to different types of irreversible theories (Q3), it shows that both a coarse-grained description and external interventions are possible sources of irreversible behavior (Q4), and it reveals that it can be useful to apply coarse-grained theories backwards in time (Q5). Consequently, for philosophers of physics who wish their positions to incorporate results of modern research on nonequilibrium systems, DDFT is a useful starting point for connecting foundational work based on the Mori–Zwanzig formalism to applied research on soft condensed matter systems - a connection that can be expected to be beneficial for both fields.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

I am very grateful to my philosophy supervisor Paul M. Näger for many interesting discussions that helped me in developing the ideas presented here, and for continuous support in writing this article. In addition, I would like to thank Ulrich Krohs, Oliver Robert Scholz and the other participants of the thesis colloquium for useful comments on a preliminary version of this work. I further acknowledge interesting discussions with various members of the Münsterian “Zentrum für Wissenschaftstheorie” (Centre for Philosophy of Science). In particular, I am grateful for helpful feedback from Markus Seidel, Gernot Münster, Jörg Friedrich, Renzo Kapust and the other participants of the “Arbeitskreis Wissenschaftstheorie” (Philosophy of Science discussion group). Moreover, I thank Berthold te Vrugt for a critical reading of the manuscript. Finally, I thank my physics supervisor Raphael Wittkowski for supporting this project and for introducing me to the relevant physics.

Funding

This work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – grant number WI 4170/3-1. I also wish to thank the Studienstiftung des deutschen Volkes, Germany for financial support.

References

- Al-Saedi, H. M., Archer, A. J., & Ward, J. (2018). Dynamical density-functional-theory-based modeling of tissue dynamics: application to tumor growth. *Physical Review E*, 98(2), 022407.
- Albert, D. Z. (2000). *Time and chance*. Cambridge (Massachusetts): Harvard University Press.
- Anero, J. G., & Español, P. (2007). Dynamic Boltzmann free-energy functional theory. *Europhysics Letters*, 78(5), 50005.
- Angioletti-Uberti, S., Ballauff, M., & Dzubiel, J. (2018). Competitive adsorption of multiple proteins to nanoparticles: the vroman effect revisited. *Molecular Physics*, 116(21–22), 3154–3163.
- Archer, A. J., & Evans, R. (2004). Dynamical density functional theory and its application to spinodal decomposition. *Journal of Chemical Physics*, 121(9), 4246–4254.
- Archer, A. J., & Rauscher, M. (2004). Dynamical density functional theory for interacting Brownian particles: stochastic or deterministic?. *Journal of Physics A: Mathematical and General*, 37(40), 9325–9333.

- Babel, S., Eikerling, M., & Löwen, H. (2018). Impedance resonance in narrow confinement. *The Journal of Physical Chemistry C*, 122(38), 21724–21734.
- Bickmann, J., & Wittkowski, R. (2020a). Collective dynamics of active Brownian particles in three spatial dimensions: a predictive field theory. *Physical Review Research*, 2(3), 033241.
- Bickmann, J., & Wittkowski, R. (2020b). Predictive local field theory for interacting active Brownian spheres in two spatial dimensions. *Journal of Physics: Condensed Matter*, 32(21), 214001.
- Blatt, J. M. (1959). An alternative approach to the ergodic problem. *Progress of Theoretical Physics*, 22(6), 745–756.
- Boltzmann, L. (1872). Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen. *Sitzungsberichte der Akademie der Wissenschaften zu Wien, mathematisch-naturwissenschaftliche Klasse*, 66, 275–370.
- Brown, H. R. (2017). Once and for all: the curious role of probability in the past hypothesis. preprint, available at <http://philsci-archives.pitt.edu/13008/>.
- Brown, H. R., Myrvold, W., & Uffink, J. (2009). Boltzmann's H-theorem, its discontents, and the birth of statistical mechanics. *Studies in History and Philosophy of Modern Physics*, 40(2), 174–191.
- Brown, H. R., & Uffink, J. (2001). The origins of time-asymmetry in thermodynamics: The minus first law. *Studies in History and Philosophy of Modern Physics*, 32(4), 525–538.
- Callender, C. (1999). Reducing thermodynamics to statistical mechanics: The case of entropy. *Journal of Philosophy*, 96(7), 348–373.
- Callender, C. (2001). Taking thermodynamics too seriously. *Studies in History and Philosophy of Modern Physics*, 32(4), 539–553.
- Callender, C. (2021). Thermodynamic asymmetry in time. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2021 ed.). Metaphysics Research Lab, Stanford University.
- Chan, G. K.-L., & Finken, R. (2005). Time-dependent density functional theory of classical fluids. *Physical Review Letters*, 94(18), 183001.
- Dean, D. S. (1996). Langevin equation for the density of a system of interacting Langevin processes. *Journal of Physics A: Mathematical and General*, 29(24), L613–L617.
- Diaw, A., & Murillo, M. S. (2015). Generalized hydrodynamics model for strongly coupled plasmas. *Physical Review E*, 92(1), 013107.
- Diaw, A., & Murillo, M. S. (2016). A dynamic density functional theory approach to diffusion in white dwarfs and neutron star envelopes. *Astrophysical Journal*, 829(1), 16.
- Donev, A., Fai, T. G., & Vanden-Eijnden, E. (2014). A reversible mesoscopic model of diffusion in liquids: from giant fluctuations to Fick's law. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(4), P04004.
- Enculescu, M., & Stark, H. (2011). Active colloidal suspensions exhibit polar order under gravity. *Physical Review Letters*, 107(5), 058301.
- Español, P., & Vázquez, F. (2002). Coarse graining from coarse-grained descriptions. *Philosophical Transactions of the Royal Society, Series A*, 360(1792), 383–394.
- Español, P., & Löwen, H. (2009). Derivation of dynamical density functional theory using the projection operator technique. *Journal of Chemical Physics*, 131(24), 244101.
- Evans, R. (1979). The nature of the liquid-vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids. *Advances in Physics*, 28, 143–200.
- Forster, D. (1974). Hydrodynamics and correlation functions in ordered systems: nematic liquid crystals. *Annals of Physics*, 84(1), 505–534.
- Fortini, A., de las Heras, D., Brader, J. M., & Schmidt, M. (2014). Superadiabatic forces in Brownian many-body dynamics. *Physical Review Letters*, 113(16), 167801.
- Fraaije, J. G. E. M. (1993). Dynamic density functional theory for microphase separation kinetics of block copolymer melts. *Journal of Chemical Physics*, 99(11), 9202–9212.
- Frigg, R. (2008). A field guide to recent work on the foundations of statistical mechanics. In D. Rickles (Ed.), *The Ashgate Companion to Contemporary Philosophy of Physics* (pp. 99–196). Ashgate.
- Frigg, R., & Werndl, C. (2021). Can somebody please say what Gibbsian statistical mechanics says? *British Journal for the Philosophy of Science*, 72(1), 105–129.
- Gibbs, J. W. (1902). *Elementary principles in statistical mechanics: developed with especial reference to the rational foundation of thermodynamics*. Charles Scribner's Sons, available at <http://eremita.di.uminho.pt/gutenberg/5/0/9/9/50992/50992-pdf.pdf>.
- Grabert, H. (1978). Nonlinear transport and dynamics of fluctuations. *Journal of Statistical Physics*, 19(5), 479–497.
- Grabert, H. (1982). *Springer tracts in modern physics: vol. 95, Projection operator techniques in nonequilibrium statistical mechanics* (1st ed.). (p. 166). Berlin: Springer-Verlag.
- Hahn, E. L. (1950). Spin echoes. *Physical Review*, 80(4), 580.
- Heinrichs, S., Dieterich, W., Maass, P., & Frisch, H. (2004). Static and time dependent density functional theory with internal degrees of freedom: Merits and limitations demonstrated for the Potts model. *Journal of Statistical Physics*, 114(3–4), 1115–1125.
- Hohenberg, P., & Kohn, W. (1964). Inhomogeneous electron gas. *Physical Review*, 136(3B), 864–871.
- Kawasaki, K. (1994). Stochastic model of slow dynamics in supercooled liquids and dense colloidal suspensions. *Physica A. Statistical Mechanics and its Applications*, 208(1), 35–64.
- Kawasaki, K. (2000). Theoretical methods dealing with slow dynamics. *Journal of Physics: Condensed Matter*, 12(29), 6343–6351.
- Kawasaki, K. (2006). Interpolation of stochastic and deterministic reduced dynamics. *Physica A. Statistical Mechanics and its Applications*, 362(2), 249–260.
- Kawasaki, K., & Gunton, J. D. (1973). Theory of nonlinear transport processes: nonlinear shear viscosity and normal stress effects. *Physical Review A*, 8, 2048–2064.
- List, C. (2019). Levels: descriptive, explanatory, and ontological. *Noûs*, 53(4), 852–883.
- Liu, Y., & Liu, H. (2020). Development of reaction-diffusion DFT and its application to catalytic oxidation of NO in porous materials. *AIChE Journal*, 66(2), e16824.
- Luczak, J. (2016). On how to approach the approach to equilibrium. *Philosophy of Science*, 83(3), 393–411.
- Luczak, J. (2017). How many aims are we aiming at?. *Analysis*, 78(2), 244–254.
- Lutsko, J. F. (2011). Communication: A dynamical theory of homogeneous nucleation for colloids and macromolecules. *Journal of Chemical Physics*, 135(16), 161101.
- Lutsko, J. F. (2019). How crystals form: A theory of nucleation pathways. *Science Advances*, 5(4), eaav7399.
- Marini Bettolo Marconi, U., & Tarazona, P. (1999). Dynamic density functional theory of fluids. *Journal of Chemical Physics*, 110(16), 8032–8044.
- Marini Bettolo Marconi, U., & Tarazona, P. (2000). Dynamic density functional theory of fluids. *Journal of Physics: Condensed Matter*, 12(8A), 413–418.
- Menon, T., & Callender, C. (2013). Turn and face the strange ...ch-changes: Philosophical questions raised by phase transitions. In R. Batterman (Ed.), *The Oxford Handbook of Philosophy of Physics* (pp. 189–223). Oxford: Oxford University Press.
- Menzel, A. M., Saha, A., Hoell, C., & Löwen, H. (2016). Dynamical density functional theory for microswimmers. *Journal of Chemical Physics*, 144(2), 024115.
- Mermin, N. D. (1965). Thermal properties of the inhomogeneous electron gas. *Physical Review*, 137(5A), A1441–A1443.
- Meyer, H., Voigtmann, T., & Schilling, T. (2019). On the dynamics of reaction coordinates in classical, time-dependent, many-body processes. *Journal of Chemical Physics*, 150(17), 174118.
- Mori, H. (1965). Transport, collective motion, and Brownian motion. *Progress of Theoretical Physics*, 33(3), 423–455.
- Munakata, T. (1989). A dynamical extension of the density functional theory. *Journal of the Physical Society of Japan*, 58(7), 2434–2438.
- Munakata, T. (1994). Time-dependent density-functional theory with H theorems. *Physical Review E*, 50(3), 2347–2350.
- Nakajima, S. (1958). On quantum theory of transport phenomena: steady diffusion. *Progress of Theoretical Physics*, 20(6), 948–959.
- North, J. (2011). Time in thermodynamics. In C. Callender (Ed.), *The Oxford Handbook of Philosophy of Time* (pp. 312–350). Oxford University Press.
- Penrose, R. (1994). On the second law of thermodynamics. *Journal of Statistical Physics*, 77(1), 217–221.
- Price, H. (1996). *Time's Arrow and Archimede's Point*. Oxford: Oxford University Press.
- Rauscher, M., Domínguez, A., Krüger, M., & Penna, F. (2007). A dynamic density functional theory for particles in a flowing solvent. *Journal of Chemical Physics*, 127(24), 244906.
- Rex, M., & Löwen, H. (2008). Dynamical density functional theory with hydrodynamic interactions and colloids in unstable traps. *Physical Review Letters*, 101(14), 148302.
- Ridderbos, T. M., & Redhead, M. L. G. (1998). The spin-echo experiments and the second law of thermodynamics. *Foundations of Physics*, 28(8), 1237–1270.
- Robbins, M. J., Archer, A. J., & Thiele, U. (2011). Modelling the evaporation of thin films of colloidal suspensions using dynamical density functional theory. *Journal of Physics: Condensed Matter*, 23(41), 415102.
- Robertson, K. (2020). Asymmetry, abstraction, and autonomy: Justifying coarse-graining in statistical mechanics. *British Journal for the Philosophy of Science*, 71(2), 547–579.
- Runge, E., & Gross, E. K. U. (1984). Density-functional theory for time-dependent systems. *Physical Review Letters*, 52(12), 997–1000.
- Schmidt, M., & Brader, J. M. (2013). Power functional theory for Brownian dynamics. *Journal of Chemical Physics*, 138(21), 214101.
- Sklar, L. (1995). *Physics and chance: Philosophical issues in the foundations of statistical mechanics*. Cambridge University Press.
- Sober, E. (2020). Histories, dynamical laws, and initial conditions- invariance under time-reversibility and its failure in Markov processes, with application to the second law of thermodynamics and the past hypothesis. *Studies in History and Philosophy of Modern Physics*, 69, 26–31.
- Streater, R. (2009). *Statistical dynamics: A stochastic approach to nonequilibrium thermodynamics*. London: Imperial College Press.
- te Vrugt, M., Bickmann, J., & Wittkowski, R. (2020a). Effects of social distancing and isolation on epidemic spreading modeled via dynamical density functional theory. *Nature Communications*, 11, 5576.
- te Vrugt, M., Löwen, H., & Wittkowski, R. (2020b). Classical dynamical density functional theory: from fundamentals to applications. *Advances in Physics*, 69(2), 121–247.
- te Vrugt, M., & Wittkowski, R. (2019). Mori-Zwanzig projection operator formalism for far-from-equilibrium systems with time-dependent Hamiltonians. *Physical Review E*, 99, 062118.
- te Vrugt, M., & Wittkowski, R. (2020a). Projection operators in statistical mechanics: a pedagogical approach. *European Journal of Physics*, 41(4), 045101.
- te Vrugt, M., & Wittkowski, R. (2020b). Relations between angular and Cartesian orientational expansions. *AIP Advances*, 10(3), 035106.

- Thiele, U., Frohoff-Hülsmann, T., Engelnkemper, S., Knobloch, E., & Archer, A. J. (2019). First order phase transitions and the thermodynamic limit. *New Journal of Physics*, 21(12), 123021.
- Treffenstädt, L. L., & Schmidt, M. (2020). Memory-induced motion reversal in Brownian liquids. *Soft Matter*, 16, 1518–1526.
- Uffink, J. (2001). Bluff your way in the second law of thermodynamics. *Studies in History and Philosophy of Modern Physics*, 32(3), 305–394.
- Uffink, J. (2007). Compendium of the foundations of classical statistical physics. In J. Butterfield, & J. Earman (Eds.), *Philosophy of Physics* (pp. 923–1074). Netherlands: Elsevier.
- van Kampen, N. G. (2002). The road from molecules to Onsager. *Journal of Statistical Physics*, 109(3), 471–481.
- Wallace, D. (2011). The logic of the past hypothesis. preprint, available at <http://philsci-archiv.pitt.edu/8894/>.
- Wallace, D. (2013). The arrow of time in physics. In H. Dyke, & A. Bardon (Eds.), *A Companion to the Philosophy of Time* (pp. 262–281). John Wiley & Sons, (Chapter 16).
- Wallace, D. (2015). The quantitative content of statistical mechanics. *Studies in History and Philosophy of Modern Physics*, 52, 285–293.
- Wensink, H. H., & Löwen, H. (2008). Aggregation of self-propelled colloidal rods near confining walls. *Physical Review E*, 78(3), 031409.
- Wittkowski, R., & Löwen, H. (2011). Dynamical density functional theory for colloidal particles with arbitrary shape. *Molecular Physics*, 109(23–24), 2935–2943.
- Wittkowski, R., Löwen, H., & Brand, H. R. (2012). Extended dynamical density functional theory for colloidal mixtures with temperature gradients. *Journal of Chemical Physics*, 137(22), 224904.
- Wittkowski, R., Löwen, H., & Brand, H. R. (2013). Microscopic approach to entropy production. *Journal of Physics A: Mathematical and Theoretical*, 46(35), 355003.
- Wittmann, R., & Brader, J. M. (2016). Active Brownian particles at interfaces: an effective equilibrium approach. *Europhysics Letters*, 114(6), 68004.
- Wittmann, R., Löwen, H., & Brader, J. M. (2021). Order-preserving dynamics in one dimension – single-file diffusion and caging from the perspective of dynamical density functional theory. *Molecular Physics*, e1867250. <http://dx.doi.org/10.1080/00268976.2020.1867250>.
- Yoshimori, A. (2005). Microscopic derivation of time-dependent density functional methods. *Physical Review E*, 71(3), 031203.
- Zeh, H.-D. (2007). *The physical basis of the direction of time* (5th ed.). Berlin Heidelberg: Springer Verlag.
- Zwanzig, R. (1960). Ensemble method in the theory of irreversibility. *Journal of Chemical Physics*, 33(5), 1338–1341.