

# Forces and symmetries in the statistical mechanics of active and thermal many-body systems

Von der Universität Bayreuth  
zur Erlangung des Grades eines  
Doktors der Naturwissenschaften (Dr. rer. nat.)  
genehmigte Abhandlung

von

Sophie Hermann

aus Nabburg

1. Gutachter: Prof. Dr. Matthias Schmidt
2. Gutachter: Prof. Dr. Roland Roth
3. Gutachter: Prof. Dr. Klaus Kroy

Tag der Einreichung: 09.08.2022  
Tag des Kolloquiums: 28.11.2022



# Abstract

Noether's theorem of invariant variations is an important mathematical theorem in functional analysis that Emmy Noether derived in 1918 as part of her habilitation thesis. In the physics community her theorem is mainly known for linking the symmetries of a given system to corresponding exact conservation laws. In this thesis we use the invariance of statistical mechanics functionals with respect to several continuous symmetries to determine on the basis of Noether's theorem exact identities for many-body systems. These statistical mechanical identities are then exploited within various applications, in particular for active Brownian particles, which form a simple nonequilibrium model system of self-propelled entities. The self-propulsion leads to several interesting phenomena including a gas-liquid-like phase transition which even occurs for purely repulsive interparticle interactions and is hence motility-induced. Further application addresses the behaviour of active Brownian particles in a gravitational field confined by a lower bounding wall. Such active sedimentation is experimentally accessible.

We lay out Noether's theorem to statistical mechanics both in the grand canonical and in the canonical ensemble. Therefore, we consider the invariances of various important and fundamental statistical functionals such as the grand potential and the free energy under symmetry operations including spatial shifts and rotations. The argument rests on two facts. On the one hand the invariant functional does not change under the symmetry transformation. On the other hand one can still expand the transformed functional around the original (i.e. non-transformed) functional with respect to the variation parameter. Comparing the results of either perspective, it becomes apparent that the linear as well as all higher order contributions in the expansion have to vanish individually. In linear order the procedure yields sum rules that describe the vanishing of mean values such as both global and spatially resolved ("local") forces and torques. In quadratic order the cancellation relates variances and curvatures, e.g. the variance of the external force with a mean curvature of the external potential. Functional differentiation of global first order sum rules gives a full hierarchy of local sum rules which relate several correlation functions that are essential in liquid state theory. While some of the hierarchies are already known in the literature, the identification of the underlying Noether concept enables their systematic derivation and it provides a constructive way to obtain new sum rules. Those sum rules include exact memory identities of nonequilibrium time-correlation functions. The Noether concept generalizes from classical to quantum statistical mechanics as we demonstrate.

The sum rules that Noether's theorem generates hold quite generally in statistical mechanics as long as the system is closed by an impenetrable external potential and no boundary contributions arise. Considering boundary terms is necessary when we apply Noether's theorem to the thermal sedimentation-diffusion equilibrium, active sedimentation, and the phase separation of both active and thermal Brownian particles. For the later system we recover the well-known mechanical pressure balance at phase

coexistence and show that this relationship also holds in nonequilibrium for active Brownian particles. Boundary contributions are also crucial for the presented proof of the virial hard wall contact theorem, which states that the density at a hard wall is determined by the virial bulk pressure. The proof itself is based on the global total force balance, which is a direct consequence of the Noether invariance under a global displacement.

We take the continuity equation as the direct origin of an exact sum rule which relates the global polarization at the interface with the current at the system boundaries far away from the interface. This has since been verified experimentally and numerically. In systems that are bounded by bulk states such as sedimentation and motility-induced phase separation of active Brownian particles, the global polarization is solely determined by bulk values and hence constitutes a state function. In both examples we give explicit expressions for this state function. In combination with Noether's theorem the polarization sum rule is then applied to global force balance equations. This combined use of sum rules yields deeper insights into the dynamics of the center of mass motion as we demonstrated for the example of active Brownian particles.

We demonstrate that all applicable sum rules are satisfied within a previously developed theoretical power functional description of the bulk and the interface at motility-induced phase separation of active Brownian particles. The variational theory works on the basis of forces and rests on the force density balance and the continuity equation. We consider the validity of the sum rules as a strong support of the theory and determine on the basis of this description the free interfacial tension. Using a square gradient approximation for the interfacial force contributions we obtain positive results for the nonequilibrium tension, which is in accordance with the observed mechanical stability of the interface in both simulations and experiments.

# Zusammenfassung

Das Noether-Theorem über invariante Variationsprobleme ist ein wichtiges Theorem der Funktionalanalysis, das 1918 von der Mathematikerin Emmy Noether im Rahmen ihrer Habilitationsschrift beschrieben wurde. In der Physik ist ihr Theorem vor allem für die Verbindung zwischen den vorliegenden Symmetrien in einem System und den zugehörigen exakten Erhaltungssätzen bekannt. Innerhalb der vorliegenden Dissertation wird die Invarianz von Funktionalen der statistischen Mechanik unter verschiedenen kontinuierlichen Symmetrien verwendet, um auf Basis des Noetherschen Theorems exakte Identitäten für Vielteilchensysteme zu bestimmen. Diese statistischen Identitäten werden für vielfältige Anwendungen verwendet, insbesondere für aktive Brownsche Teilchen, einem einfachen Modellsystem von selbstangetriebenen Objekten im Nichtgleichgewicht. Der Selbstantrieb führt zu mehreren interessanten Phänomenen wie beispielsweise der Phasenseparation in eine gas- und flüssig-ähnliche Phase. Diese Separation tritt selbst bei einem rein repulsiven Wechselwirkungspotential zwischen den Teilchen auf und ist daher durch die Beweglichkeit der Teilchen verursacht. Eine weitere Anwendung beinhaltet das Verhalten aktiver Brownscher Teilchen in einem Gravitationsfeld an einer unteren begrenzenden Wand. Aktive Sedimentation und Phasenseparation sind experimentell zugänglich.

In der Dissertation wird die Anwendung des Noether Theorems für das großkanonische und kanonische Ensemble in der statistischen Mechanik demonstriert. Aus diesem Grund wird die Invarianz von verschiedenen, wichtigen statistischen Funktionalen betrachtet, wie dem großkanonischen Potential oder der freien Energie unter Symmetrievaryationen, die räumliche Verschiebungen und Rotationen beinhalten. Die Argumentation stützt sich auf zwei Fakten. Auf der einen Seite ändert sich ein invariantes Funktional nicht unter seiner zugehörigen Symmetrietransformation. Auf der anderen Seite kann man das transformierte Funktional dennoch formal um das ursprüngliche (also das nicht transformierte) Funktional in Abhängigkeit eines Variationsparameters entwickeln. Vergleicht man beide Betrachtungsweisen miteinander, stellt man fest, dass alle Beiträge linear sowie höherer Ordnung im Transformationsparameter jeweils einzeln Null ergeben müssen. In linearer Ordnung folgen aus dieser Bedingung Summenregeln, die das Verschwinden von Erwartungswerten für globale und räumlich aufgelöste („lokale“) Kräfte und Drehmomente beschreiben. In quadratischer Ordnung verknüpft dieses Wegheben Varianzen mit Krümmungsbeiträgen, beispielsweise die Varianz der externen Kraft mit dem Mittelwert der Krümmung des externen Potentials. Die Funktionalableitung von globalen Summenregeln gibt eine vollständige Hierarchie von lokalen Summenregeln, die verschiedene fundamentale Korrelationsfunktionen aus der Flüssigkeitstheorie miteinander in Verbindung setzt. Einige dieser Hierarchien sind bereits in der Literatur bekannt, doch die Identifikation des zugrundeliegenden Noetherschen Konzeptes ermöglicht nun eine systematische Konstruktion neuer Summenregeln. Diese Summenregeln schließen exakte Gedächtnis-Identitäten für Nichtgleichgewichts-Zeitkorrelationsfunktionen mit

ein. Das Noether Konzept lässt sich von der klassischen auf die quantenmechanische statistische Mechanik verallgemeinern, wie ebenfalls in der Schrift gezeigt wird.

Die vom Noether Theorem erzeugten Zusammenhänge gelten allgemein in der statistischen Mechanik, solange das betrachtete System durch ein undurchdringliches externes Potential begrenzt wird und entsprechend keine Flussterme am Rand auftreten können. Die Betrachtung von Randbeiträgen wird beispielsweise dann nötig, wenn man das Noether-Theorem auf thermische und aktive Sedimentation oder auf die Phasenseparation von aktiven und thermischen Brownschen Teilchen anwendet. Für Letzteres wird die bekannte mechanische Druck-Bilanz bei Phasenkoexistenz wiedergefunden. Dieses Druckgleichgewicht gilt für aktive Brownsche Schwimmer auch noch im Nichtgleichgewicht. Randbeiträge tragen entscheidend bei zum Beweis des virialen Kontakttheorems der harten Wand, welches die Kontaktdichte an einer harten Wand durch den Virialdruck der zugehörigen Bulk-Phase ausdrückt. Der Beweis selbst basiert auf dem globalen Gleichgewicht der totalen Kraft, das wiederum selbst eine direkte Konsequenz der Noether Invarianz unter einer globalen Verschiebung ist.

Weiterhin verwende ich die Kontinuitätsgleichung als direkten Ursprung für eine weitere exakte Summenregel. Diese setzt die globale Polarisierung an der Grenzfläche mit dem Bulkstrom an den Systemgrenzen, weit weg von der Grenzfläche, in Beziehung. Diese Summenregel wurde mittlerweile experimentell und numerisch bestätigt. Für Systeme, die von Bulkphasen begrenzt werden, wie Sedimentation oder bewegungsinduzierte Phasenseparation von aktiven Brownschen Teilchen, ist die globale Polarisierung dann nur über Bulkgrößen bestimmt und stellt daher eine Zustandsfunktion dar. In den beiden genannten Beispielfällen werden explizite Ausdrücke für diese Zustandsfunktion angegeben. Man kann sowohl die Noether-Summenregel als auch die Polarisierungssummenregel auf die globale Kraftbilanz anwenden. Die kombinierte Verwendung beider Relationen erlaubt es, tiefere Einsichten in die Dynamik der Schwerpunktsbewegung zu gewinnen, wie am Beispiel der aktiven Teilchen explizit gezeigt wird.

Die Gültigkeit der Summenregeln wird demonstriert für eine zuvor entwickelte approximative Theorie für aktive Phasenseparation zur Validierung der Approximation. Diese Variationstheorie arbeitet auf der Basis von Kräften und beruht auf der Kraftdichtebilanz und der Kontinuitätsgleichung. Die nachgewiesene Gültigkeit der Summenregeln dient zur Validierung der Theorie. Auf Grundlage dieser Beschreibung wird die Oberflächenspannung der freien Grenzfläche bestimmt. Unter Verwendung einer quadratischen Gradientennäherung für die nichtlokalen internen Kraftbeiträge wird eine positive Nichtgleichgewichts-Grenzflächenspannung bestimmt, in Übereinstimmung mit der beobachteten mechanischen Stabilität der Grenzfläche.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>A reminder: Noether's theorem in classical mechanics</b>	<b>4</b>
2.1	Conservation laws in classical mechanics . . . . .	4
2.2	Connecting symmetries and conservation laws . . . . .	6
<b>3</b>	<b>Noether's theorem in statistical mechanics</b>	<b>11</b>
3.1	Density functional background and sum rules in statistical mechanics .	12
3.2	Noether's theorem for the translational invariance of the grand potential	15
3.3	Statistical mechanical applications and generalizations of Noether's theorem . . . . .	21
<b>4</b>	<b>Interface polarization sum rule</b>	<b>33</b>
4.1	Derivation of an interface polarization sum rule . . . . .	33
4.2	Interface polarization of active Brownian particles . . . . .	35
4.3	Combining polarization and Noether sum rules . . . . .	38
<b>5</b>	<b>Free interfacial tension of active Brownian particles</b>	<b>41</b>
5.1	Literature overview over the free interfacial tension of active systems .	42
5.2	Theoretical determination of the free interfacial tension . . . . .	44
<b>6</b>	<b>Conclusions and outlook</b>	<b>49</b>
	<b>Bibliography</b>	<b>55</b>
<b>7</b>	<b>Publications</b>	<b>67</b>
7.1	List of publications . . . . .	67
7.2	Author's contributions . . . . .	68
7.3	Noether's Theorem in Statistical Mechanics . . . . .	71
7.4	Why Noether's Theorem applies to Statistical Mechanics . . . . .	85
7.5	Force density functional theory in- and out-of-equilibrium . . . . .	99
7.6	Force balance in thermal quantum many-body systems from Noether's theorem . . . . .	115
7.7	Variance of fluctuations from Noether invariance . . . . .	133
7.8	Active interface polarization as a state function . . . . .	139
7.9	Non-negative interfacial tension in phase-separated active Brownian particles . . . . .	145
<b>A</b>	<b>Bulk and interfacial theory of motility-induced phase separation</b>	<b>151</b>
	<b>Acknowledgements</b>	<b>155</b>
	<b>Eidesstattliche Versicherung</b>	<b>157</b>





# 1 Introduction

Emmy Noether's theorem for invariant variational problems [8] is a well-known and important relation in a wide range of physics. Students are usually introduced to it early on in their undergraduate studies within teaching of the Lagrangian mechanics section of a theoretical classical mechanics lecture course [9]. The theorem reveals the deep connection between symmetries and associated conservation laws and it shows the very significant value of symmetries beyond mere geometric simplification. Arguably the most prominent example is that of a translational invariant system, i.e. one which does not depend on its position in space along a given direction and as a consequence conserves momentum along this direction. Invariance with respect to time translations, which originates from the homogeneity of time, similarly causes conservation of energy. However, Emmy Noether's work is neither of mere didactic nor of only historic interest. Rather the theorem has innumerable applications in a broad range of physical fields such as quantum mechanics, high energy particle physics and (quantum) gauge-field theory [10, 11]. There is current research on generalizations of the theorem including the extension to systems with stochastic forces [12, 13], Markov processes [14] and thermodynamic systems [15–17]. Although often referred to in singular, there are in fact two types of the theorem. Noether's first theorem applies to invariances with respect to rigid transformations where the dependence is only on a uniform transformation parameter. The corresponding conservation laws represent (exact) statements about global quantities such as the conservation of the global momentum or of the energy. Noether's second theorem applies to functionals that are symmetric under transformations that depend on varying an entire function. The identities constitute and yield nontrivial locally resolved relations.

Both intrinsic symmetries and spontaneously broken symmetries are highly relevant in soft matter not least due to the prominence of microscopy techniques and visualization in computer simulations. Soft condensed matter encompasses systems such as colloids, liquid crystals and active matter, with prominent phenomena such as drying, wetting [18, 19], bubble formation [20] and not least bulk phase coexistence [21]. Therefore predicting the behavior of such complex systems is usually challenging in particular when starting with a microscopic description. The presence of thermal fluctuations is non-negligible and it leads to the important effect of Brownian motion which underlies and drives much of the dynamics on the mesoscale.

One popular example of a nonequilibrium system are active Brownian particles (ABPs). ABPs show interesting self-aligning behaviour [21–27]. These particles belong to the more general material class of active matter which encompasses further artificial and natural systems that are intrinsically out of equilibrium by construction. One important feature of active motion is the self-propelled movement of particles. For ABPs this swimming or self-motility is due to force along an intrinsic orientation, given e.g. by the orientation of Janus particles. Active Brownian particles also undergo Brownian diffusion as their name already suggests. Diffusion of the particle positions

plays a minor role for high particle swim speeds, but the rotational diffusion of the particle orientations is of significant importance for the collective particle dynamics. Although ABPs are typically exposed to no external torques they are also known to display spontaneous polarization effects.

Experimentally prepared hemispherical realizations of active particles include Janus particles [28, 29], which are spheres with two different surfaces such as latex colloids where on half of the surface is coated with platinum [29]. The motion of such particles is induced by the solvent itself, e.g. by the dissociation of hydrogen peroxide [29] or by imposing external radiation with techniques such as photon nudging [30–33]. Examples for active particles are bacteria such as *E. coli* that propel with their flagella [34, 35], but the concept is also widened beyond the soft matter realm to birds and fish and addresses respectively their gathering in flocks and schools [25–27, 36, 37]. Theoretical models to capture and describe the experimentally (and naturally) observed phenomena are the aforementioned active Brownian particles, but also run-and-tumble particles [38, 39] and active Ornstein-Uhlenbeck particles [40, 41]. In a similar way to ABPs, these particles self-propel by the action of a swim force, but they differ in the type of rotational motion. Such models then serve as the basis for developing theories and carrying out direct simulations. In contrast, field theories such as the active model B [42] and the active model B+ [20] form a different basis to model active matter and are also taken to represent ABPs. Those theories are based on identifying relevant coarse-grained fields such as the orientationally-averaged density field and sometimes also include higher orientational orders of the density such as the polarization field or the nematic order parameter tensor. Lee [43] investigated the polar, nematic and higher order parameters at the interface of MIPS based on integral formulas that he derived. The large number of models and the breadth of different and mutually competing theories for their description alone can be taken as an indication for the current popularity of this topic.

Active matter in general tends to phase separate into a dense and a dilute phase due to the intrinsic particle propulsion. This effect is therefore both a relevant and also useful testbed for the study and fundamental modeling of nonequilibrium phase transitions. The description of the motility-induced phase separation and especially the determination of the coexistence densities turned out to be a particularly challenging and controversial problem. Attempts to transfer Maxwell’s equal area construction to nonequilibrium did not yield satisfactory coexistence densities [44, 45]. In contrast the treatment based on power functional theory rests on the Maxwell construction and the results do agree well with simulation data [46]. A natural next question is to hence address the interfacial tension of the “free” interface between the phase separated bulk states. Bialké *et al.* [47] were first in obtaining the tension from simulations. They determined a large negative value (several hundred  $k_B T$ ). The negative sign seemingly strictly contradicts the observed stability of the interface. This finding started a controversy about a meaningful definition and the resulting sign of the tension in nonequilibrium [45, 47–59]. We put our own work [SH7] in perspective below.

For descriptions of phase transitions that occur in soft condensed matter one usual considers the density profile, which serves as a locally resolved order parameter. (Fluctuation profiles can be an equally well or an even more suitable choice in some situations like for critical drying [18, 19, 60].) Besides the density profile, correlation functions are important quantities for the characterization of fluids. Correlation

functions help to obtain a systematic understanding of the structure and dynamics of the system under investigation. Therefore exact identities, so-called sum rules, that relate different types of correlation functions to each other (typically via integration) are often invaluable for carrying out practical work. Lovett, Mou and Buff [61] and Wertheim [62] derived independently one relation that connects the one-body density distribution with the two-body direct correlation function, known by the acronym LMBW equation. Besides its use in carrying out consistency checks this equation is useful for the development of integral equation theories [63–66] and density functional theories [67–69]. Baus [70] determined a range of further sum rules related to the LMBW equation. His article also gives a nice and well structured overview of the field.

Density functional theory (DFT) is a powerful framework which is based on a formally exact variational principle and the existence of the generating grand potential functional [71,72], which e.g. in the case of hard spheres can be well approximated with fundamental measure theory [73,74]. References [75,76] provide textbook presentations of the theory and reference [77] gives a review of newer developments. In a nutshell DFT is a method to consider equilibrium many-body systems with inhomogeneous density distributions using statistical mechanics. Its nonequilibrium generalization is the power functional theory (PFT) [78], see reference [79] for a recent review. Here the generating power functional depends functionally not only on the density but also on the current. This allows to predict and also to rule out certain dependencies. Kinematic contributions to the functional thus only depend on the density and the current, but not on the external potential which itself only occurs in the external contributions to the power functional. The structure of dependencies can then be exploited in Taylor expansions in velocities or velocity gradients [80,81]. Power functional theory allows to theoretically describe nonequilibrium phenomena such as the occurrence of viscous and structural forces [81–83] and deconfinement [84] in sheared fluids, viscoelasticity and memory [85–87] and nonequilibrium fluid demixing as it occurs in lane formation [88]. The application of PFT to phase separation of active Brownian particles is central and has proven to be very fruitful [46,89,90]. The theory was developed for overdamped systems but it can be generalized to molecular dynamics and hence systems with inertia including quantum systems [91,92].

In the wide gamut of soft and active matter, fluctuations and correlations, density functional and power functional methods, we seek for common ground on the basis of symmetries and from the perspective of Noether’s theorem. In the application of Noether’s theorem one exploits the continuous symmetry of a functional. Such functionals of course appear naturally within functional theories such as DFT or PFT. The sum rules that result from the functional invariances relate distinct correlation functions via integration with each other. In the context of soft matter Noether’s theorem proves the validity of the LMBW equation, which relates the one-body and the two-body direct correlation function to each other. For active systems the pressure balance for phase separated ABPs is obtained in nonequilibrium.

In this thesis we exemplify and emphasize the importance of the connecting bridge that Noether’s theorem forms between symmetries and sum rules in these different areas. The thesis is structured as follows. We first recall elementary conservation laws and the application of Noether’s theorem in classical mechanics in chapter 2. In chapter 3 we show how Noether’s theorem applies to statistical mechanics using the explicit example of translational invariance of the grand potential. We proceed

with an overview over the possible applications and generalizations of Noether's theorem in statistical mechanics. This material has originally appeared in publications [SH1,SH2,SH3,SH4,SH5]. The selection of the here presented material includes the invariance of the free energy, rotational invariances, nonlinear order contributions, boundary effects and the application to active and passive Brownian particles. In chapter 4 we introduce a sum rule for the global interfacial polarization which relates this quantity with the boundary values of the current. The derivation is shown from a slightly different perspective than that adopted in publication [SH6], as our restriction to special system geometries is applied at a later stage of the derivation. The polarization sum rule is then applied to active Brownian particles that are exposed to a gravitational potential and undergo motility-induced phase separation. Our treatment yields the magnitude and the time evolution of the global interfacial polarization. We also report new insights from the combination of the interface polarization sum rule with the sum rules [SH1,SH2] resulting from Noether's theorem. This gives the time evolution of the center of mass velocity for sedimenting ABPs [SH1]. Finally in chapter 5 we give a short overview of the calculation of the interfacial tension of active Brownian particles which is the central result of publication [SH7]. The tension is determined via the van der Waals route on the basis of our theory of motility-induced phase separation. The formulation of the theory itself was part of the author's prior work [93,94]. Main ideas and concepts are described in appendix A for clarity and for providing a self-contained presentation for the benefit of the reader. A detailed description of the content of the papers that contribute to this cumulative thesis along with a declaration of the present author's contributions is given in section 7.1 on page 67.

## 2 A reminder: Noether's theorem in classical mechanics

### 2.1 Conservation laws in classical mechanics

In this introductory chapter we recapitulate the significance of conservation laws and of Noether's theorem in classical mechanics. Undoubtedly the study of conserved quantities is an important tool to describe physical phenomena. A conserved quantity is one that is constant in time. The time derivative hence vanishes and the conserved quantity satisfies a conservation law. Famous examples include the conservation of the total momentum as a consequence of the homogeneity of space which expresses the equivalence of all points along a given direction in space and implies the invariance of the Lagrangian under spatial translations in that direction.

Let us start with a simple example and consider the throw of an apple. The system, which we take to consist of both the apple and Earth, and hence also its Lagrangian are independent of the overall position in space, so the total momentum is conserved. Effects of friction of the apple against air are neglected here for simplicity. In a partial

system which includes only the apple but not Earth, the momentum of the apple is not a constant of motion. Instead the momentum increases as the apple accelerates (during its fall) perpendicular towards the surface of the earth. This is no contradiction to momentum conservation as the Lagrangian of the apple depends via the external gravitational potential on the height and hence on its spatial position. There is no spatial homogeneity for the apple alone. [In the combined system of apple and Earth not only the apple but also Earth are accelerated and both (anti-parallel) momentum changes sum to zero.] However, for the contribution of the movement parallel to the surface of Earth (disregarding its curvature) the momentum is conserved, even if one considers solely the apple. This parallel component of the of motion differs from the perpendicular motion. A spatial shift in the parallel direction would not change the movement itself. It is independent of the absolute position in space. The trajectory of the apple stays the same independent of whether it is thrown at the given or at any other starting place. The only difference is the trivial general offset. The trajectory itself changes qualitatively only if the initial height or the initial velocity of the throw is varied. So there appears to be a connection between the invariance under a spatial displacement and the conservation of momentum.

A more formal and entirely mathematical reason for the conservation of momentum parallel to Earth's surface is that the corresponding position coordinates are cyclic. Recall that within classical mechanics a generalized coordinate  $\mathbf{q}_i(t)$  is called cyclic, when the Lagrangian  $\mathcal{L}$  is independent of it. The general dependence  $\mathcal{L}(\mathbf{q}_1, \dots, \mathbf{q}_N, \dot{\mathbf{q}}_1, \dots, \dot{\mathbf{q}}_N, t)$  on the coordinates  $\mathbf{q}_i(t)$  of all  $N$  particles, the corresponding velocities  $\dot{\mathbf{q}}_i(t)$  and time  $t$  simplifies to  $\mathcal{L}(\mathbf{q}_1, \dots, \mathbf{q}_{i-1}, \mathbf{q}_{i+1}, \dots, \mathbf{q}_N, \dot{\mathbf{q}}_1, \dots, \dot{\mathbf{q}}_N, t)$ . (In the following we suppress the dependences of  $\mathcal{L}$  on its variables for clarity of notation.) So the corresponding derivative of the Lagrangian vanishes, hence

$$0 = \frac{\partial \mathcal{L}}{\partial \mathbf{q}_i} = \frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}_i} \right), \quad (2.1)$$

where we used the Euler-Lagrange equation to obtain the second equality in equation (2.1). The derivative  $\partial \mathcal{L} / \partial \dot{\mathbf{q}}_i = \mathbf{p}_i$  can be interpreted as a generalized momentum conjugate to  $\mathbf{q}_i$ . Integration in time of equation (2.1) in the form of  $0 = d\mathbf{p}_i / dt$  shows that the generalized momentum is a constant in time. So from the invariance  $\partial \mathcal{L} / \partial \mathbf{q}_i = 0$  one can conclude that the momentum  $\mathbf{p}_i$  is a temporally conserved quantity in this system, see footnote [95]. Hence cyclic coordinates state a more general and widely known connection between invariances and conservation laws in classical mechanics [9].

In case of the example of the throw of an apple, the Lagrangian of the apple does only depend on its velocity (due to the kinetic energy contribution) and on its height (because of the potential energy). We take the generalized coordinate to be just a (Cartesian) vectorial position variable. Its components parallel to the earth surface are cyclic as they neither affect potential nor kinetic energy. This yields the conservation of momentum in the two directions perpendicular to the gravitational force field.

These basic considerations demonstrate that conservation laws seem to be connected to the invariances and the symmetries of a system. This observation raises general questions under which circumstances conservation laws are valid in systems with external forces, and how or if the form of a conservation law connects to invariances. Both questions are directly related to the more general question of why conservation laws hold in the first place.



**Figure 2.1:** Emmy Noether (front, center) and some of her colleagues and students during a trip to the *Gasthof Vollbrecht* in the district Nikolausberg, Göttingen. The mathematicians Hermann Weyl (4th from left) and Emil Artin (behind Emmy Noether) are notable. The photograph was taken in 1932 by Natascha Artin [98].

## 2.2 Connecting symmetries and conservation laws

Emmy Noether (1882-1935) was one of the foremost mathematicians of the 20th century. Today she is one of the most well-known female mathematicians if not scientists of all times. Several schools, a lecture hall at the University of Erlangen-Nuremberg, the physics and math building of the University of Siegen and a funding program for young scientist of the German Research Foundation (*Deutsche Forschungsgemeinschaft*), were named in her honor.

In the following we give some interesting biographical facts and several anecdotes about Emmy Noether as a person [96, 97]. Her collaborators describe her as a very obliging and friendly person. One thing that characterizes her quite well is her ardor and love for mathematics. An anecdote that captures this passion very nicely is a conversation with fellow mathematician Emil Artin in 1934, after the seizure of control of Hitler. Both were discussing very fast, intensely and loudly about *Links- und Rechtsideale* (left and right ideals), *Gruppen* (groups), and *Untergruppen* (subgroups) in the subway of Hamburg. In German all these words have a strong political connotation besides the mathematical meaning. Given the historical period the chance of getting arrested for these statements was quite high, but she was completely oblivious to the danger [96]. Another habit of her was to go on long walks to the country side and discuss about mathematics with her colleagues and students. At one of these trips they took a group picture in front of the *Gasthof Vollbrecht*, see figure 2.1. Emmy Noether stands confidently and relaxed in the center of the group. Directly behind

her is her subway companion, friend and colleague Emil Artin, whose wife took the photograph.

Being a very passionate person, Emmy Noether's enthusiasm might have helped her to overcome several difficulties and resistances during her career. One reason for her troubles is politics since she was Jewish and expressed Marxism-inclined opinions. The second reason is simply her being a woman. Women have been given the right to vote in Germany only since 1918. So when young Emmy started to study in Erlangen in 1903 she had to ask the professors if she was allowed to sit and follow the lecture. Emmy Noether's first application for habilitation 1915 in Göttingen was rejected although she made already important contributions to the theory of differential and algebraic invariants. (Reasons were the fear of stealing jobs from the soldiers who returned from world war I and the question if the female brain is even capable to do proper mathematics although Emmy Noether was seen as an exception [97].)

Her habilitation thesis, which was finally accepted in 1918, addressed *Invariant Variation Problems* [8]. Within this work she was able to find the connection between invariances and divergence relationships, which reduce to conservation laws for dynamical variational problems. She further translated this finding into a formal mathematical description and fathomed the deeper mathematical origin of conservation laws. These insights in invariants are referred to as the so-called Noether's theorem which states informally:

*To each continuous symmetry corresponds a related conservation law.*      (2.2)

The theorem links a physical property (the conservation law) with a continuous symmetry (of the invariant quantity). This connection is very useful across a range of fields in physics, from classical mechanics to field theory and high-energy particle physics. One reason is that symmetries are usually easier to obtain than conservation laws which then of course greatly simplifies the physical description of a system. Furthermore Noether's theorem is often viewed to answer the questions posed above of why conservation laws hold.

To understand theorem (2.2) we first clarify the meaning of the terminology. As already stated above a conservation law is a statement that expresses the fact that a given physical observable does not change in time. A symmetry occurs when a functional, such as the action or the Lagrangian, does not change under a transformation, which can be a transformation of coordinates such as shifts and rotations in space. An object that resides on an (infinite) table is symmetric under spatial displacements in all directions parallel to the table surface. One would not be able to realize that the object was moved without reference to a coordinate system. There are two different kinds of symmetry: discrete and continuous ones. A continuous symmetry implies invariance under a symmetry operation even in the infinitesimally small limit of this operation. For a discrete symmetry the invariance only holds for special values of the transformation and it hence describes non-continuous changes, such as reflections, rotations by a certain fixed angle or translations by a certain fixed vector. (Each of these discrete operations is of course highly relevant in crystallography.) In contrast a continuous symmetry has an infinite number of symmetry operations, whereas discrete symmetries only have a finite or countably infinite number of operations. An equilateral triangle has a discrete rotational symmetry as it is only identical to itself when rotated around its center for multiples of  $120^\circ$ . An example for a continuous

symmetry is the rotation of a circle (or a sphere) around its center. The object is invariant under this symmetry operation for rotations either with an infinitesimally small angle or with any arbitrary finite angle.

Let us now be more specific and turn to classical mechanics. Applications of Noether's theorem in the form (2.2) usually exploit the continuous symmetries of the action functional

$$\mathcal{S} = \int_{t_1}^{t_2} dt \mathcal{L}. \quad (2.3)$$

The time integral runs between the initial time  $t_1$  and the final time  $t_2$  and symmetries of the Lagrangian itself are relevant. In the following we determine the conservation law in classical mechanics for transformations  $\mathbf{q}_i(t) \rightarrow \mathbf{q}_i^\epsilon(t) = \mathbf{q}_i(\epsilon, t)$ , where  $\epsilon$  denotes the constant parameter of the transformation performed on each generalized coordinate  $\mathbf{q}_i(t)$  for all  $i$ . In the limit of vanishing  $\epsilon$  the original (i.e. non-transformed) coordinate is regained,  $\mathbf{q}_i^{\epsilon \rightarrow 0}(t) = \mathbf{q}_i(t)$ . (Explicit examples of such transformations are given below and include a uniform shift and a uniform rotation.)

To be able to apply Noether's theorem the action functional has to be invariant under the given transformation, hence

$$\mathcal{S} = \mathcal{S}^\epsilon, \quad (2.4)$$

where  $\mathcal{S}^\epsilon$  indicates the action which results from applying the transformation. Note that the symmetry has to be continuous, such that equation (2.4) needs to remain valid particularly for infinitesimal values of  $\epsilon$ , so one can perform a Taylor expansion. The expansion of  $\mathcal{S}^\epsilon$  around  $\epsilon = 0$  up to linear order yields

$$\mathcal{S}^\epsilon = \mathcal{S} + \left. \frac{d\mathcal{S}^\epsilon}{d\epsilon} \right|_{\epsilon=0} \epsilon + \mathcal{O}(\epsilon^2). \quad (2.5)$$

Exploiting the symmetry of the action (2.4) allows to cancel the left-hand side of equation (2.5) with the first term on the right-hand side. As the transformation parameter  $\epsilon$  is arbitrary, we conclude that the derivative in the second term on the right-hand side of equation (2.5) must vanish:

$$0 = \left. \frac{d\mathcal{S}^\epsilon}{d\epsilon} \right|_{\epsilon=0} \quad (2.6)$$

$$= \int_{t_1}^{t_2} dt \left. \frac{d\mathcal{L}^\epsilon}{d\epsilon} \right|_{\epsilon=0} \quad (2.7)$$

$$= \int_{t_1}^{t_2} dt \sum_{i=1}^N \left( \frac{\partial \mathcal{L}}{\partial \mathbf{q}_i} \cdot \frac{d\mathbf{q}_i^\epsilon}{d\epsilon} + \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}_i} \cdot \frac{d\dot{\mathbf{q}}_i^\epsilon}{d\epsilon} \right)_{\epsilon=0}, \quad (2.8)$$

where we have inserted in equation (2.7) the definition (2.3) of the action  $\mathcal{S}$  and  $\mathcal{L}^\epsilon$  indicates the transformed Lagrangian. As the transformation parameter  $\epsilon$  is constant, the phase space derivatives trivially commute with the time integral. Using the chain rule we rewrite equation (2.7) to (2.8) exploiting that  $\partial \mathcal{L} / \partial \mathbf{q}_i = \partial \mathcal{L}^\epsilon / \partial \mathbf{q}_i^\epsilon$  and  $\partial \mathcal{L} / \partial \dot{\mathbf{q}}_i = \partial \mathcal{L}^\epsilon / \partial \dot{\mathbf{q}}_i^\epsilon$  hold in case of  $\epsilon = 0$  for each  $i$ , such that the derivatives do no longer depend on  $\epsilon$ . To proceed we interchange the order of the derivatives  $d/d\epsilon$  and



$d/dt$  in the factor  $d\dot{\mathbf{q}}_i^\epsilon/d\epsilon$  of equation (2.8) and perform an integration by parts on this term, which yields

$$0 = \int_{t_1}^{t_2} dt \sum_{i=1}^N \left[ \frac{\partial \mathcal{L}}{\partial \mathbf{q}_i} - \frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}_i} \right) \right] \cdot \frac{d\mathbf{q}_i^\epsilon}{d\epsilon} \Big|_{\epsilon=0} + \sum_{i=1}^N \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}_i} \cdot \frac{d\mathbf{q}_i^\epsilon}{d\epsilon} \Big|_{\epsilon=0} \Big|_{t_1}^{t_2}. \quad (2.9)$$

The first contribution in equation (2.9) vanishes due to the Euler-Lagrange equation

$$\frac{\partial \mathcal{L}}{\partial \mathbf{q}_i} - \frac{d}{dt} \left( \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}_i} \right) = 0, \quad (2.10)$$

which ultimately describes the force balance of the generalized forces with the temporal change of the generalized momentum for the  $i$ th particle. One can hence conclude that the remaining boundary term (that results from integration by parts) in equation (2.9) has the same value at the initial time  $t_1$  and the final time  $t_2$ . Recall, that both times  $t_1$  and  $t_2$  are arbitrary. (We consider cases where the symmetry for the action is valid for all times  $t$ .) Therefore, we have shown that

$$\sum_{i=1}^N \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}_i} \cdot \left( \frac{d\mathbf{q}_i^\epsilon}{d\epsilon} \right)_{\epsilon=0} = \text{const.} \quad (2.11)$$

Hence the left-hand side of equation (2.11) is a constant of motion, i.e. its value is independent of time. We conclude that equation (2.11) constitutes a conservation law.

We have thus explicitly shown that the invariance of the action (2.4) under transformations of the form  $\mathbf{q}_i(\epsilon, t)$  leads to the conservation law (2.11). The derivation exploits the invariance of the action under Taylor expansion. (A similar structure as in the determination of conserved quantities in publication [SH2] where we used Noether's theorem in statistical mechanics within the canonical ensemble. We will turn to statistical mechanics below in chapter 3.)

As the above considered transformations are formal we give some concrete examples of possible transformations. One simple transformation is a mere uniform shift of all generalized coordinates,  $\mathbf{q}_i^\epsilon = \mathbf{q}_i + \epsilon \mathbf{e}_\zeta$ , where  $\mathbf{e}_\zeta$  denotes the unit vector in the direction of the displacement and the index  $\zeta$  enumerates this direction, e.g.  $\zeta = x, y, z$  for the Cartesian axes. The magnitude of the invariant shift is arbitrary but their direction is fixed along  $\mathbf{e}_\zeta$ . Hence the derivative of  $\mathbf{q}_i^\epsilon(t)$  simplifies to  $d\mathbf{q}_i^\epsilon/d\epsilon = \mathbf{e}_\zeta$  and the conserved quantity (2.11) immediately becomes the  $\mathbf{e}_\zeta$ -component of global generalized momentum  $\mathbf{p}$ ,

$$\sum_i \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{q}}_i} \cdot \mathbf{e}_\zeta = \text{const.} \quad (2.12)$$

Exploiting the fact that the global momentum  $\mathbf{p} = \sum_i \mathbf{p}_i$  corresponds to a global force  $\mathbf{F}$  via Newton's second law,  $\dot{\mathbf{p}} = \mathbf{F}$ , allows to rephrase the conservation law (2.12) after differentiation with respect to time as the vanishing of the global force,  $\mathbf{F} = 0$ . When applied to the concrete example of the apple throw we have  $N = 1$  and take the generalized coordinate  $\mathbf{q}(t)$  to be given in the Cartesian coordinate system. Due to the gravitational force the invariance of the system and hence of the action (2.4) only holds for displacements  $\mathbf{e}_\zeta$  perpendicular to gravity. So the sum in the conservation law

(2.11) that is implied in the dot product only holds for components in that directions and reduces to one summand. Identifying  $\partial\mathcal{L}/\partial\dot{\mathbf{q}}$  as the momentum of the apple, we have shown the momentum conservation for its components parallel to the earth surface. This conservation law was expected from the previous force and momentum considerations in section 2.1. The uniform shift is one of the symmetry transformation which we also apply later in the context of statistical mechanics (see subsections 3.2.1, 3.2.2 and 3.3.2).

We give a further example to illustrate the general conservation law (2.11) and hence consider the uniform rotation of the system as the relevant symmetry transformation. An infinitesimal rotation around the  $\mathbf{e}_\zeta$ -axis can be expressed as a transformation of all generalized coordinates as  $\mathbf{q}_i^\epsilon = \mathbf{q}_i + \epsilon\mathbf{e}_\zeta \times \mathbf{q}_i$ , where  $\epsilon$  here indicates the angle of rotation. The change of this transformation is  $d\mathbf{q}_i^\epsilon/d\epsilon = \mathbf{e}_\zeta \times \mathbf{q}_i$ . Hence the action symmetric under rotation yields the conservation of equation (2.11)

$$\frac{\partial\mathcal{L}}{\partial\dot{\mathbf{q}}_i} \cdot (\mathbf{e}_\zeta \times \mathbf{q}_i) = \mathbf{e}_\zeta \cdot (\mathbf{q}_i \times \mathbf{p}_i) = \text{const}, \quad (2.13)$$

where we exploited the cyclic permutability of the triple product and the definition of the generalized momentum  $\mathbf{p}_i = \partial\mathcal{L}/\partial\dot{\mathbf{q}}_i$  in the first equality. The conserved quantity in (2.13) can be identified as the  $e_\zeta$ -component of the angular momentum  $\mathbf{L} = \mathbf{q}_i \times \mathbf{p}_i$ . As for the above case of a uniform shift (2.12) one can here differentiate the conservation law (2.13) with respect to time and find vanishing of the global torque  $\mathcal{T} = \dot{\mathbf{L}} = 0$ .

For cases where equation (2.13) is valid for all Cartesian axes,  $\zeta = x, y, z$ , the system is isotropic in space which results in the conservation of angular momentum or equivalently in the vanishing of global torque. (The sum rules that follow from rotation using Noether's theorem in statistical mechanics are shown in subsection 3.3.3.)

One of the advantages of using Noether's theorem is its universality, which is a characteristic feature for the work. Emmy Noether's original work focused on the invariance of a *general* functional with respect to *any* continuous symmetry transformation. In practice, most physicists usually consider the symmetry of the action functional or the corresponding Lagrangian. Even with this apparently restricted type of application her theorem applies and is useful for many theories in physics, which includes classical mechanics, quantum mechanics, quantum field theory, high-energy particle physics and general relativity [10, 11]. Noether's theorem [8] solved the back then important problem of energy conservation in general relativity, see reference [99] for a description. In classical mechanics energy conservation results from homogeneity in time, i.e. the absence of an explicit time dependence of the Lagrangian,  $\partial L/\partial t = 0$  [9]. It is also possible to derive Maxwell's equations of classical electromagnetism from a gauge symmetry [100]. Applications of Noether's theorem are not restricted to such historically important works. Rather current research addresses generalizations of the theorem, including the treatment of stochastic forces [12–14] and drawing conclusions for thermodynamic properties such as the entropy being a Noether invariant [15–17].

In this thesis we apply Noether's theorem to statistical mechanics problems, as we lay out in detail in the following chapter 3. In contrast to the usual applications this one does not consider symmetries of the action integral nor of the underlying

Lagrangian. Instead we focus on equally central objects from statistical mechanics such as the grand potential or the free energy. These thermodynamic potentials depend in general on the external potential function and can hence be expressed as a functional of this functions. Additionally one can look at these thermodynamic potentials from the viewpoint of classical density functional theory where they are phrased with a functional dependence on the density. A key difference to the previously shown usage of Noether’s theorem in classical mechanics is the level of corresponding physical description. Classical mechanical systems are entirely deterministic, so one given state is (in principle) sufficient to predict both the past and the future states of the system based on solving the equations of motion. In statistical mechanics the relevant physical problems depend on probabilities and fluctuations they exhibit. Our description hence focuses on probability statements and on averages. The difference also influences the exact identities that result from Noether’s theorem. The dynamical description in classical mechanics induced time as a variable and hence yields conservation laws, whereas in statistical mechanics the theorem determines sum rules. Here sum rules are relations which contain integrals (“sums”) that occur due to the statistical averages and phase space integrals.

Independent of the different structure we will show in the following chapter 3 that Noether’s theorem in statistical mechanics enables one to obtain comparably deep insights as in classical mechanics. We derive simple but essential identities such as the vanishing of forces and torques as well as more complex relations such as between time direct correlation functions that all result from fundamental symmetries. Hence this arguably elementary concept allows to gain important insights into the structure of the physical world.

### 3 Noether’s theorem in statistical mechanics

In this chapter we demonstrate how Noether’s theorem can be applied in statistical mechanics. Noether’s theorem is shown to generate exact statements for correlation functions from continuous symmetries of functionals to which we refer as Noether sum rules. Therefore we give in section 3.1 a short overview of classical density functional theory, where statistical mechanical functionals feature naturally and prominently sum rules are valuable e.g. in the development of new approximations. We also introduce more general readers to direct correlation functions and show the spirit of the original derivation of the inverse LMBW sum rule, as determined by Lovett, Mou and Buff [61] and Wertheim [62] (section 3.1). The derivation of these authors makes no reference to Noether’s theorem but it exploits the consequences of a uniform and global shift on the system. In the section 3.2 we exemplify Noether’s theorem of the grand potential (as a functional of the external potential) under translational invariance, which we developed in publication [SH1]. We derive the resulting sum rule of vanishing of the global external force and show further equations related to this sum rule, such as the

inverse LMBW equation which can be interpreted as being a direct consequence of the underlying system symmetries. Section 3.3 contains an overview of the possible applications and generalization of the developed concept to determine sum rules. Explicit applications of Noether's theorem in statistical mechanics include systems of active and passive particles, where we focus on the phenomena of thermal and active sedimentation (subsection 3.2.3).

### 3.1 Density functional background and sum rules in statistical mechanics

The physical description of many-body systems such as the currently popular model of active particles is a challenging task due to the mutual interactions between the particles. The only available possibility to handle and describe these complex classical systems is statistical mechanics. One specific formalism of *equilibrium* statistical mechanics is the classical density functional theory (DFT).

We give an introduction to classical DFT, which was established by Evans [72], and covered in several sources [72, 76, 77, 79]. The review [79] gives a nicely written, low-level introduction to the topic and also presents important concepts from statistical mechanics (for completeness). Many of the presented relations and proofs, such as the Mervin-Evans theorem [72, 101], are proven or derived explicitly. The paper [79] also includes a review of the historical evolution of the field and famous landmark contributions and an introduction to the dynamical generalization of DFT, the so-called power functional theory. For a more condensed overview of classical DFT one can read the corresponding chapter within the book [76], the standard reference within (and far beyond) the field of simple liquids. *Theory of Simple Liquids* [76] presents an accessible and self-contained description of liquid state theory on the basis of classical statistical mechanics. Going beyond the basics of DFT, the review [77] covers more recent developments of the field. Here we just give a short sketch of the main concepts of DFT.

Density functional theory is based on treating the density profile  $\rho(\mathbf{r}) = \langle \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \rangle$  as a variable, where  $\delta(\cdot)$  denotes the Dirac delta distribution and  $\mathbf{r}_i$  indicates the position of the  $i = 1, \dots, N$  particle, where  $N$  is the total number of particles. The term within the brackets is the density operator,  $\hat{\rho}(\mathbf{r}; \mathbf{r}_1, \dots, \mathbf{r}_N) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$  and the (angular) brackets indicate a statistical average. The DFT is usually formulated in the grand canonical ensemble. As a short reminder, grand-canonically the statistical average is defined as  $\langle \cdot \rangle = \text{Tr} e^{-\beta(H - \mu N)} / \Xi$ , where the trace is given as a series of phase space integrals  $\text{Tr} = \sum_N (h^{3N} N!)^{-1} \int d\mathbf{r}^N d\mathbf{p}^N$  and  $e^{-\beta(H - \mu N)} / \Xi$  denotes the probability distribution for microstates. Here  $\mu$  is the chemical potential,  $\beta = 1/k_B T$  indicates the inverse temperature with Boltzmann constant  $k_B$  and temperature  $T$ , and  $h$  denotes the Planck constant. We used the shorthand  $\mathbf{r}^N = \mathbf{r}_1, \dots, \mathbf{r}_N$  and  $\mathbf{p}^N = \mathbf{p}_1 \dots \mathbf{p}_N$ , where  $\mathbf{r}_i$  indicates the position of the  $i$ th particle and  $\mathbf{p}_i$  is its momentum. The partition function  $\Xi$  is given as the trace of the Boltzmann factor, i.e.  $\Xi = \text{Tr} e^{-\beta(H - \mu N)}$ , with the Hamiltonian

$$H = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + u(\mathbf{r}^N) + \sum_{i=1}^N V_{\text{ext}}(\mathbf{r}_i), \quad (3.1)$$

which consists of a kinetic part, an interparticle interaction term  $u(\mathbf{r}^N)$  and an external contribution. So the one-body density distribution  $\rho(\mathbf{r})$ , which only depends on one position coordinate, follows from the average of the  $N$ -body density operator  $\hat{\rho}(\mathbf{r}; \mathbf{r}^N)$ , which depends on all particle positions  $\mathbf{r}_i$ .

Within DFT one exploits that the “real” physical density profile corresponds to the minimum of the grand potential  $\Omega[\rho]$  expressed as a functional of the density. Throughout the square brackets indicate a functional dependence. To determine the minimum in practice one performs the functional differentiation of the grand potential with respect to a trial density  $\tilde{\rho}(\mathbf{r})$ ,

$$\left. \frac{\delta \Omega[\tilde{\rho}]}{\delta \tilde{\rho}(\mathbf{r})} \right|_{\tilde{\rho}(\mathbf{r})=\rho(\mathbf{r})} = 0, \quad (3.2)$$

which is after insertion of  $\Omega[\rho]$  the Euler-Lagrange equation of the system. Solving this equation (3.2) allows to determine the equilibrium density  $\rho(\mathbf{r})$  and the thermodynamic properties of the system, because  $\Omega[\rho]$  is the “real” value of the grand potential. Despite the fact that they are referred by the same term this is different from the Euler-Lagrange equation of motion (2.10) in classical mechanics. Here the Euler-Lagrange equation expresses the (one-body) balance of locally resolved chemical potentials.

The grand potential consists of four distinct contributions,

$$\Omega[\rho] = F_{\text{id}}[\rho] + F_{\text{exc}}[\rho] + \int d\mathbf{r} \rho(\mathbf{r})(V_{\text{ext}}(\mathbf{r}) - \mu), \quad (3.3)$$

where in the third term  $V_{\text{ext}}(\mathbf{r})$  indicates the external potential and in the fourth term  $\mu$  is the chemical potential as before. The ideal gas free energy density functional  $F_{\text{id}}[\rho]$  is known exactly and it has the following form:

$$F_{\text{id}}[\rho] = k_{\text{B}}T \int d\mathbf{r} \rho(\mathbf{r}) \left[ \ln(\rho(\mathbf{r})\Lambda^d) - 1 \right], \quad (3.4)$$

where  $\Lambda$  is the thermal de Broglie wavelength and  $d$  indicates the spatial dimension. After differentiation with respect to the density according to equation (3.2), the locally resolved intrinsic ideal gas free energy (3.4) yields the ideal gas chemical potential. The excess free energy density functional  $F_{\text{exc}}[\rho]$  results from the mutual interactions between the particles and therefore it has to be approximated in practice. The knowledge of an exact expression of  $F_{\text{exc}}[\rho]$  would imply the solution of the many-body problem for systems under the influence of arbitrary external potentials, which surely will remain an insurmountable problem as concerns analytic work.

The functional derivative of the excess free energy functional with respect to the density gives up to inverse temperature  $\beta$  the (negative) one-body direct correlation function,  $c_1(\mathbf{r}) = -\delta\beta F_{\text{exc}}[\rho]/\delta\rho(\mathbf{r})$ , whose gradient determines the internal force density field  $\mathbf{f}_{\text{int}}(\mathbf{r}) = k_{\text{B}}T\nabla c_1(\mathbf{r})$ . Using this definition of the direct correlation function is sometimes referred to as potential-DFT [SH3]. This DFT based on potentials is complemented by force-DFT, based on forces, see footnote [102] and reference [SH3]. Further differentiation of  $c_1(\mathbf{r})$  defines the two-body direct correlation function  $c_2(\mathbf{r}, \mathbf{r}') = \delta c_1(\mathbf{r})/\delta\rho(\mathbf{r}')$  as is central in liquid state theory [76].

In order to find practical approximations for  $F_{\text{exc}}[\rho]$  it is helpful to be aware of sum rules, which are exact equations that relate different physical (correlation) functions

to each other. One famous sum rule relates the one-body direct correlation function to an integral (“sum”) over the two-body direct correlation function via [72]

$$\nabla \ln \rho(\mathbf{r}) + \beta \nabla V_{\text{ext}}(\mathbf{r}) = \int d\mathbf{r}' c_2(\mathbf{r}, \mathbf{r}') \nabla' \rho(\mathbf{r}'), \quad (3.5)$$

where  $\nabla$  indicates differentiation with respect to  $\mathbf{r}$  and  $\nabla'$  is differentiation with respect to  $\mathbf{r}'$ . This so-called LMBW equation (3.5) was derived originally independently in 1976 by Lovett, Mou and Buff [61] and in 1976 by Wertheim [62]. The LMBW equation has found numerous applications in integral equation theory [63–66] and in classical density functional theory [68, 69]. It is often considered to be an equally valuable alternative to the Triezenberg-Zwanzig equation [103] or to the first member of the Yvon-Born-Green (YBG) hierarchy [104, 105]. As it relates the density with the two-body direct correlation function equation (3.5) can complement the Ornstein-Zernike (OZ) equation together with a further closure relation [63] in the construction of inhomogeneous liquid state theories. The LMBW equation was used to consider freezing and its structural precursors [63, 67], nonideal crystals [67–69], fluids at semipermeable walls [64], viscoelastic surface waves of liquids [106], and liquid-vapour as well as liquid-liquid interfaces [65, 66].

As our considerations are within the grand canonical ensemble the LMBW relation (3.5) has a formal inverse,

$$-\nabla \rho(\mathbf{r}) = \int d\mathbf{r}' \beta H_2(\mathbf{r}, \mathbf{r}') \nabla' V_{\text{ext}}(\mathbf{r}'), \quad (3.6)$$

where  $H_2(\mathbf{r}, \mathbf{r}')$  is the density-density correlation function that denotes the covariance of the local density operator,  $H_2(\mathbf{r}, \mathbf{r}') = \langle \hat{\rho}(\mathbf{r}) \hat{\rho}(\mathbf{r}') \rangle - \rho(\mathbf{r}) \rho(\mathbf{r}')$ . It has been shown [61, 62, 72] that the inverse LMBW equation (3.6) follows from the effects of a uniform coordinate transformation,  $\mathbf{r} \rightarrow \mathbf{r} + \boldsymbol{\epsilon}$ . This is equivalent to section 2.2, where the scalar transformation parameter  $\epsilon$  and the direction of the displacement  $\mathbf{e}_\zeta$  are combined within the vector of the uniform spatial shift  $\boldsymbol{\epsilon}$ .

In the remainder of this section we sketch the derivation of the inverse LMBW equation (3.6). Due to the global displacement the external potential shifts accordingly as  $V_{\text{ext}}(\mathbf{r}) \rightarrow V_{\text{ext}}(\mathbf{r} + \boldsymbol{\epsilon})$ . The density distribution is a unique functional of the external potential,  $\rho(\mathbf{r}, [V_{\text{ext}}])$ , such that density is (unambiguously) for the given interparticle interaction potential determined by the external potential via this functional dependence. Therefore the coordinate transformation might as well be interpreted as a shift of the external potential. Regardless of the chosen interpretation the density itself will in general change under the applied global displacement. As the only object that fixes the position of the equilibrium density profile in space is the external potential, the density will shift in the same way as the external potential,  $\rho(\mathbf{r}) \rightarrow \rho(\mathbf{r}, [V_{\text{ext}}(\mathbf{r} + \boldsymbol{\epsilon})]) = \rho(\mathbf{r} + \boldsymbol{\epsilon})$ . Recall that the interparticle interaction contribution solely depends on relative particle positions and that it is hence unaffected by the shift. The determined change of the density,  $\rho(\mathbf{r} + \boldsymbol{\epsilon}) - \rho(\mathbf{r})$ , implies that the functional  $\rho[V_{\text{ext}}]$  is not invariant under the considered global transformation. Hence Noether's theorem (see chapter 2) does not apply, as the theorem generates identities from the *invariances* of functionals.

To proceed one (functionally) Taylor expands the shifted density profile around the unshifted external potential  $V_{\text{ext}}(\mathbf{r})$  up to linear order [61, 62, 72], which yields:

$$\rho(\mathbf{r} + \boldsymbol{\epsilon}) = \rho(\mathbf{r}) + \int d\mathbf{r}' \frac{\delta \rho(\mathbf{r}, [V_{\text{ext}}])}{\delta V_{\text{ext}}(\mathbf{r}')} (V_{\text{ext}}(\mathbf{r}' + \boldsymbol{\epsilon}) - V_{\text{ext}}(\mathbf{r}')), \quad (3.7)$$

where the response of the density due to changes in the external potential can be rewritten as the correlation function of the density fluctuations,  $H_2(\mathbf{r}, \mathbf{r}') = -k_B T \delta\rho(\mathbf{r})/\delta V_{\text{ext}}(\mathbf{r}')$  [76]. In the limit of small displacements,  $\epsilon \rightarrow 0$ , the changes in the density  $\rho(\mathbf{r} + \epsilon) - \rho(\mathbf{r})$  and in the external potential  $V_{\text{ext}}(\mathbf{r} + \epsilon) - V_{\text{ext}}(\mathbf{r})$  in equation (3.7) can be rewritten as gradient expressions, which leads to and constructively proves equation (3.6). As already mentioned earlier, the LMBW equation (3.5) is the inverse relation of (3.6). It can be hence derived from a density-shift and the fact that the external potential is a unique functional of the density [72],  $V_{\text{ext}}(\mathbf{r}, [\rho])$ , using the identical argumentation as above. The existence and uniqueness of  $V_{\text{ext}}(\mathbf{r}, [\rho])$  was shown within the Mermin-Evans theorem [72, 101], see e.g. reference [76] for an accessible account of the proof of that theorem.

## 3.2 Noether's theorem for the translational invariance of the grand potential

Although the above derivations of the sum rules (3.5) and (3.6) are closely related to Noether's theorem, they are not yet direct applications of her theorem. The structure of the derivation is based on the consequences of a (coordinate) transformation as is Noether's theorem. However, in the above example there is no explicit exploitation of the symmetry as the density functional was not invariant with respect to the considered uniform shift. The exploitation of symmetries and invariances and their consequences is an important characteristic and the basis of Noether's theorem as we have seen for deterministic systems within classical mechanics in chapter 2. We have shown that the conservation law (2.12) holds in the case of classical mechanics in section 2.2 above. Knowing that Emmy Noether herself was a mathematician (see section 2.2) it is natural that she aimed to formulate and to prove her theorem as generally as possible. So her theorem is not only valid for the action functional but for any functionals that is invariant under a symmetry transformation. Notably this generality was already honored by Albert Einstein, who wrote in a letter to Hilbert [107]:

“Yesterday I received from Miss Noether a very interesting paper on invariant forms. I am impressed that one can comprehend these matters from so general a viewpoint. It would not have done the old guard at Göttingen any harm had they picked up a thing or two from her. [...]”

While standard applications of Emmy Noether's theorem are typically restricted to invariances of the action functional (which for instance solved the problem of energy conservation in general relativity [99]), here we intend to push these boundaries and test the generality that Einstein had immediately recognized. We therefore apply the theorem to the invariance of several different functionals from statistical mechanics [SH1, SH2].

### 3.2.1 Grand canonical ensemble

We start with a consideration in the grand canonical ensemble. To be applicable the theorem needs the identification of a continuous symmetry operation and of a specific functional which is invariant under this transformation. As the symmetry

transformation we choose a displacement of the external potential,  $V_{\text{ext}}(\mathbf{r}) \rightarrow V_{\text{ext}}(\mathbf{r} + \boldsymbol{\epsilon})$ , as in the above case to determine the inverse LMBW equation (3.6). Similarly, this shift is equivalent to a uniform global transformation of the coordinates, as we considered in section 2.2.

For small  $\boldsymbol{\epsilon}$  one can expand the displaced external potential around  $\boldsymbol{\epsilon} = 0$  up to linear order:  $V_{\text{ext}}(\mathbf{r} + \boldsymbol{\epsilon}) = V_{\text{ext}}(\mathbf{r}) + \boldsymbol{\epsilon} \cdot \nabla V_{\text{ext}}(\mathbf{r})$ . Naively one would expect this change in  $V_{\text{ext}}$  to influence the grand potential  $\Omega[V_{\text{ext}}]$ , due to the functional dependence of the latter on the external potential. Functional Taylor expansion of the shifted grand potential functional with respect to  $V_{\text{ext}}$  up to linear order gives

$$\Omega[V_{\text{ext}}(\mathbf{r} + \boldsymbol{\epsilon})] = \Omega[V_{\text{ext}}(\mathbf{r})] + \int d\mathbf{r}' \frac{\delta\Omega[V_{\text{ext}}]}{\delta V_{\text{ext}}(\mathbf{r}')} \nabla V_{\text{ext}}(\mathbf{r}') \cdot \boldsymbol{\epsilon}. \quad (3.8)$$

The main requirement to apply Noether's theorem is the identification of a functional that is invariant under the given transformation, here the shift of the external potential. As can be easily seen, the grand potential is invariant under this transformation,  $\Omega[V_{\text{ext}}(\mathbf{r})] = \Omega[V_{\text{ext}}(\mathbf{r} + \boldsymbol{\epsilon})]$ . This invariance can be determined from explicitly evaluating the effects of displacing the statistical mechanical definition  $\Omega[V_{\text{ext}}] = -k_{\text{B}}T \ln \Xi[V_{\text{ext}}]$  [76, 79], where  $\Xi = \text{Tr} e^{-\beta(H - \mu N)}$  is the grand canonical partition function. Therefore the left-hand side of equation (3.8) cancels with the unshifted grand potential  $\Omega[V_{\text{ext}}]$ , i.e. the first term on the right-hand side. As the magnitude and direction of the shift  $\boldsymbol{\epsilon}$  are arbitrary, we can conclude that

$$\int d\mathbf{r} \frac{\delta\Omega[V_{\text{ext}}]}{\delta V_{\text{ext}}(\mathbf{r})} \nabla V_{\text{ext}}(\mathbf{r}) = \int d\mathbf{r} \rho(\mathbf{r}) \nabla V_{\text{ext}}(\mathbf{r}) = 0, \quad (3.9)$$

where we have evaluated the functional derivative  $\delta\Omega[V_{\text{ext}}]/\delta V_{\text{ext}}(\mathbf{r}) = \rho(\mathbf{r})$ , see e.g. reference [79] for an explicit derivation of this identity. One can identify in equation (3.9) the external force,  $\mathbf{f}_{\text{ext}}(\mathbf{r}) = -\nabla V_{\text{ext}}(\mathbf{r})$  [which only contains the conservative contribution due to the external potential  $V_{\text{ext}}(\mathbf{r})$ , as we are in equilibrium]. Equation (3.9) states that the global external force  $\mathbf{F}_{\text{ext}}^{\circ} = \int d\mathbf{r} \rho(\mathbf{r}) \mathbf{f}_{\text{ext}}(\mathbf{r})$  vanishes in equilibrium [SH1, SH2],

$$-\int d\mathbf{r} \rho(\mathbf{r}) \nabla V_{\text{ext}}(\mathbf{r}) = \mathbf{F}_{\text{ext}}^{\circ} = 0. \quad (3.10)$$

This exact relation (3.10) can be interpreted as the ‘‘conservation law’’ resulting from Noether's theorem applied to the global translational invariance of the grand canonical, although we usually refer to such equations as *sum rules* due to the occurrence of integrals (‘‘sums’’). The result (3.10) is similar to the conservation of the global momentum (2.12) or vanishing global force determined in section 2.2 from invariance of the action function under a uniform displacement. However, here the sum rule holds for the external contribution (3.10) and the shifting operation acts on the external potential (instead of on the general position  $\mathbf{r}$ ).

Equation (3.10) and its derivation do not require the use of the functional Taylor expansion. An ordinary Taylor expansion with respect to the displacement  $\boldsymbol{\epsilon}$  is sufficient as we briefly sketch in footnote [108]. The grand potential does satisfy the translational invariance for all closed systems such that equation (3.10) holds in every equilibrium system which is confined by an external potential. The confinement also ensure the vanishing of all boundary terms as required by the argumentation. The



case of open boundaries and the consideration of subvolumes of the entire (infinite) system are treated in the subsection 3.3.8 below.

The vanishing of the global external force (3.10) is also starting point for the derivation of further sum rules. We show in the following direct consequences of equation (3.10). Because of its generality, equation (3.10) is satisfied independently of the exact form of the external potential. This allows to functionally differentiate (3.10) with respect to  $V_{\text{ext}}$  and keep the equality satisfied. The differentiation gives two contributions due to the product rule. The first term contains the functional derivative of the density, which can be again identified as the correlation function of density fluctuations  $H_2(\mathbf{r}, \mathbf{r}') = -k_B T \delta\rho(\mathbf{r})/\delta V_{\text{ext}}(\mathbf{r}')$ . It forms the right-hand side of (3.6). The second term results from the functional differentiation of the external potential gradient,  $\delta\nabla V_{\text{ext}}(\mathbf{r})/\delta V_{\text{ext}}(\mathbf{r}') = \nabla\delta(\mathbf{r} - \mathbf{r}')$ . Integration by parts leads to the left-hand side of (3.6) and thus the inverse LMBW equation (3.6) is determined as a local external Noether sum rule [SH1].

One could stop at this point, as one has regained a known sum rule. But one can also go further and realize that equation (3.6) still holds for a considerable variety of forms of the external potential. As first realized and carried out by Baus [70] it is hence possible to continue differentiating the sum rule (3.6). The  $n$ th functional derivative of equation (3.10) is then [SH1, 70]

$$-\sum_{\alpha=1}^n \nabla_{\alpha} H_n = \int d\mathbf{r}_{n+1} \beta V_{\text{ext}}(\mathbf{r}_{n+1}) \nabla_{n+1} H_{n+1}, \quad (3.11)$$

which relates the  $(n+1)$ -body density-density correlation function  $H_{n+1} = \delta H_n / \delta\beta V_{\text{ext}}$  to its  $n$ -body version  $H_n = H_n(\mathbf{r}_1, \dots, \mathbf{r}_n)$  [72, 76]. The position arguments in equation (3.11) have been omitted for clarity. The inverse LMBW equation (3.6) and its higher-order generalizations (3.11) are well-known results, see e.g. reference [70]. Here we have found a new derivation of these sum rules. The method allows us to derive the relations with great ease and to ultimately identify them as consequences of Noether's theorem and hence of the symmetry properties of the system.

We now multiply equation (3.6) by  $V_{\text{ext}}(\mathbf{r})$  and integrate over position. The left-hand side then becomes equal to the global external force (3.10) and hence vanishes. We have thus shown that after integration by parts the following global sum rule ensues [SH1]:

$$\int d\mathbf{r} V_{\text{ext}}(\mathbf{r}) \int d\mathbf{r}' [\nabla' V_{\text{ext}}(\mathbf{r}')] H_2(\mathbf{r}, \mathbf{r}') = 0, \quad (3.12)$$

where the derivative  $\nabla'$  only acts on the external potential as indicated by the brackets. The  $n$ -body version of equation (3.12) is determined similarly by multiple position integration of the hierarchy (3.11) as

$$\int d\mathbf{r}_1 V_{\text{ext}}(\mathbf{r}_1) \dots \int d\mathbf{r}_n V_{\text{ext}}(\mathbf{r}_n) \nabla_{\alpha} H_n(\mathbf{r}_1, \dots, \mathbf{r}_n) = 0, \quad (3.13)$$

with  $\alpha = 1, \dots, n$  and  $n \geq 2$ . For  $n = 2$  equation (3.13) reduces again to the global Noether sum rule (3.12) upon integration by parts. The sum rule hierarchies (3.12) and (3.13) are new results to the best of our knowledge [SH1]. To find a physical interpretation of these results it is useful to rewrite each equation by exploiting the fact

that  $H_2$  is the covariance of the density operator,  $H_2(\mathbf{r}, \mathbf{r}') = \langle \hat{\rho}(\mathbf{r})\hat{\rho}(\mathbf{r}') \rangle - \rho(\mathbf{r})\rho(\mathbf{r}')$ . Using the definition of the grand canonical ensemble average and the vanishing of the external force (3.10), the sum rule (3.12) becomes after integration by parts [SH2]

$$-\left\langle \sum_{j=1}^N V_{\text{ext}}(\mathbf{r}_j) \sum_{i=1}^N \nabla_i V_{\text{ext}}(\mathbf{r}_i) \right\rangle = 0, \quad (3.14)$$

which states that there is no correlation between the global external potential and the global external force. The higher orders (3.13) mean that the correlation between the global external force and higher moments (powers) of the global external potential also vanishes. We will see below in subsection 3.3.7 that not all such correlators vanish when addressing the autocorrelation of the external force.

### 3.2.2 Canonical ensemble

Although the above derivations were performed in the grand canonical ensemble, the resulting identities remain valid in the canonical ensemble [SH2]. Baus [70] was first in demonstrating the universality of the sum rules in the canonical and the grand canonical ensemble. We re-derived his results and made the connection to Noether's theorem applied canonically [SH2]. The representation in the canonical ensemble is relevant for simulations with fixed particle numbers where there can be pronounced differences between the results of these two ensembles when the particle number is small. (Also for undergraduate students the canonical ensemble might be more accessible and intuitive as it does not require to engage with the arguably more abstract chemical potential.)

We again consider the shift of the external potential,  $V_{\text{ext}}(\mathbf{r}) \rightarrow V_{\text{ext}}(\mathbf{r} + \boldsymbol{\epsilon})$ , but now applied to a system that is described on the basis of the canonical ensemble. Here the invariant quantity is the (canonical) free energy  $F_N[V_{\text{ext}}]$ , which functionally depends on the external potential  $V_{\text{ext}}(\mathbf{r})$ . (The consequences of the translational invariance of the free energy density functional  $F[\rho]$  are also relevant and yield different sum rules as shown in subsection 3.3.2.) In (elementary) statistical mechanics the free energy is  $F_N[V_{\text{ext}}] = -k_B T \ln Z_N[V_{\text{ext}}]$ , i.e. the negative logarithm of the partition function  $Z_N = \text{Tr}_N e^{-\beta H}$  times thermal energy. The canonical trace is given as  $\text{Tr}_N = (h^{3N} N!)^{-1} \int d\mathbf{r}^N d\mathbf{p}^N$ , where the phase space integrals run over all particle positions  $\mathbf{r}_i$  and momenta  $\mathbf{p}_i$  and  $i = 1, \dots, N$  indicates the  $i$ th particle. The Hamiltonian is independent of the chosen ensemble and is hence again taken to have the standard form defined by equation (3.1). The ensemble average in the canonical average is  $\langle \cdot \rangle_N = \text{Tr}_N e^{-\beta H} \cdot / Z_N$ .

As the expansion of the canonical free energy viewed as a functional of the external potential is analogous to the grand canonical expansion of the grand potential (see subsection 3.2.1), we just give a brief overview here. A detailed derivation is contained in reference [SH2]. The displaced free energy expanded up to linear order in the displacement parameter  $\boldsymbol{\epsilon}$  is given as

$$F_N[V_{\text{ext}}(\mathbf{r} + \boldsymbol{\epsilon})] = F_N[V_{\text{ext}}(\mathbf{r})] + \int d\mathbf{r} \frac{\delta F_N[V_{\text{ext}}]}{\delta V_{\text{ext}}(\mathbf{r})} \nabla V_{\text{ext}}(\mathbf{r}) \cdot \boldsymbol{\epsilon}. \quad (3.15)$$

### 3.2 Noether's theorem for the translational invariance of the grand potential

The functional derivative of the free energy with respect to the external potential can be evaluated as

$$\frac{\delta F_N[V_{\text{ext}}]}{\delta V_{\text{ext}}(\mathbf{r})} = -\frac{k_B T}{Z_N} \frac{\delta Z_N[V_{\text{ext}}]}{\delta V_{\text{ext}}(\mathbf{r})} \quad (3.16)$$

$$= \frac{1}{Z_N} \text{Tr}_N e^{-\beta H} \frac{\delta H[V_{\text{ext}}]}{\delta V_{\text{ext}}(\mathbf{r})} \quad (3.17)$$

$$= \frac{1}{Z_N} \text{Tr}_N e^{-\beta H} \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i), \quad (3.18)$$

where we have used the definition of the free energy  $F_N[V_{\text{ext}}]$  in equation (3.16), of the partition function  $Z_N[V_{\text{ext}}]$  in equation (3.17) and of the Hamiltonian  $H$  (3.1) in equation (3.18). As the kinetic and the internal contribution of the Hamiltonian are independent of the external potential, only the external term contributes. Note that the functional derivative of a function with respect to itself gives the Dirac distribution, i.e.  $\delta V_{\text{ext}}(\mathbf{r}')/\delta V_{\text{ext}}(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}')$ . Insertion of the derivative (3.18) in the expansion (3.15) and exploiting the invariance  $F_N = F_N(\epsilon)$  yields

$$0 = \int d\mathbf{r} \text{Tr} \frac{e^{-\beta H[V_{\text{ext}}]}}{Z_N[V_{\text{ext}}]} \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \nabla V_{\text{ext}}(\mathbf{r}) \quad (3.19)$$

$$= \int d\mathbf{r} \rho(\mathbf{r}) \nabla V_{\text{ext}}(\mathbf{r}) \quad (3.20)$$

where we have used the canonical expression of the density distribution  $\rho(\mathbf{r}) = \langle \hat{\rho} \rangle_N = \text{Tr} e^{-\beta H} \sum_i \delta(\mathbf{r} - \mathbf{r}_i)/Z_N$  to determine equation (3.20) which constitutes the (negative) averaged global external force  $\mathbf{F}_{\text{ext}}^o$ . After an integration by parts we can identify,

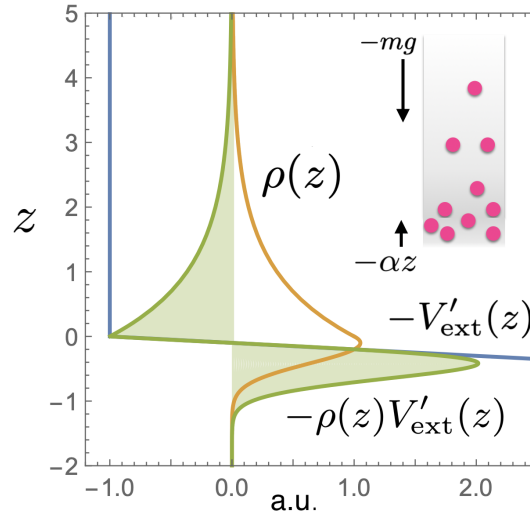
$$-\mathbf{F}_{\text{ext}}^o = - \int d\mathbf{r} V_{\text{ext}}(\mathbf{r}) \nabla \rho(\mathbf{r}) = 0. \quad (3.21)$$

Hence in the canonical ensemble the averaged global external force  $\mathbf{F}_{\text{ext}}^o$  indeed vanishes, as it does in the grand-canonical ensemble, see (3.10). Analogous versions of the sum rules (3.6) and (3.11)–(3.13) are also valid in the canonical ensemble [70].

#### 3.2.3 Application to thermal and active sedimentation

While the above derived exact sum rules are both important and valuable for equilibrium density functional theory, they are also useful in nonequilibrium topics such as active matter. To give a concrete application of the vanishing external force (3.10) and (3.21) we examine the influence of gravity in the following. We consider two cases. First we treat the sedimentation-diffusion equilibrium and secondly we address the active sedimentation of active Brownian particles.

We start with thermal sedimentation. We consider  $N$  Brownian particles in an external potential, which contains the gravitational contribution and a lower confining wall,  $V_{\text{ext}}(z) = mgz + V_{\text{wall}}(z)$ . Here the wall is represented by the external potential contribution  $V_{\text{wall}}(z)$ . The  $z$ -coordinate measures the height. For a sketch of the system see the inset in figure 3.1. When the system has equilibrated one can apply the Noether sum rule (3.10) as boundary terms vanish. The external wall confines the particles at the bottom of the sample. The gravitational potential prevents escape



**Figure 3.1:** Schematic of sedimentation-diffusion equilibrium with a lower bounding wall,  $V_{\text{ext}}(z) = mgz - \frac{\alpha}{2}z^2\theta(-z)$ . The corresponding local external forces  $-V'_{\text{ext}}(z) = -dV_{\text{ext}}(z)/dz$  (blue line) are indicated by arrows in the inset. There the pink discs represent the particles and the grey-shaded area denotes the density distribution. This density distribution is according to the barometric law (orange line). Noether's theorem states that the global external force, i.e. the integral over the external force density  $-\rho(z)V'_{\text{ext}}(z)$ , vanishes. So the area of both green shaded regions are equal as expected. The figure is taken from [SH2], ©IOP Publishing (2019). Reproduced with permission. All rights reserved.

on the top such that the density vanishes at both boundaries. Contributions of the lateral system boundaries are irrelevant as we assume the system for simplicity to be quasi-one dimensional and hence to only depend on the  $z$ -coordinate. This reduces equation (3.10) to

$$0 = \int d\mathbf{r} \rho(z) \frac{dV_{\text{ext}}(z)}{dz} = -mgN + F_{\text{wall}}^{\circ}, \quad (3.22)$$

where we identified the total number of particles  $N = \int d\mathbf{r} \rho(\mathbf{r})$  and the global vertical force on the wall is given by  $F_{\text{wall}}^{\circ} = - \int d\mathbf{r} \rho(z) dV_{\text{wall}}(z)/dz$ . So as the total external force has to vanish, the gravitational force of all particles exactly cancels the force that the particles exert on the wall,  $mgN = F_{\text{wall}}^{\circ}$ .

We restrict ourselves to the case of an ideal gas, so there are no interactions between the particles,  $u(\mathbf{r}^N) = 0$ . The density distribution is then given according to the generalized barometric law,  $\rho(z) = \Lambda^3 e^{-\beta(V_{\text{ext}}(z) - \mu)}$ , where  $\mu$  indicates the chemical potential and  $\Lambda$  denotes the thermal wavelength. Hence the density decays exponentially with respect to the external potential as is visualized in figure 3.1. To be able to explicitly plot this distribution we consider a truncated harmonic wall, so  $V_{\text{wall}}(z) = -\alpha z^2 \theta(-z)/2$ , where  $\theta(z)$  indicates the Heaviside step function and  $\alpha$  is the spring constant. Then the derivative of the external potential is constant for  $z > 0$  and linear for  $z < 0$ , given by:  $V'_{\text{ext}}(z) = dV_{\text{ext}}(z)/dz = mg - \alpha z \theta(-z)$ . So the magnitude of the external force density,  $-\rho(z)dV_{\text{ext}}(z)/dz$ , is negative outside wall and mainly positive inside the wall. Both the negative and the positive contributions cancel each other (indicated by the green shaded areas in Fig. 3.1), such that the global external force vanishes as we had expected on the basis of equation (3.10).

Going beyond of the above quite well-known sedimentation of thermal (“passive”) Brownian particles, we next consider active Brownian particles under gravity [29, 109–111]. Active Brownian particles move similar to Brownian particles, but they are driven out of equilibrium as they undergo self-propelled motion [23, 25, 30, 31, 44]. Their swimming is modeled as an additional force on the particle and its direction is along the particle orientation  $\omega$ . The orientation is a unit vector and it represents an additional degree of freedom that the particles possess. Hence the Brownian diffusion does not only affect the spatial position but also the (angular) orientation of the particles. The motion of active Brownian particles is taken to be overdamped and we focus here on spherical particles. For ideal gas-like and weakly interacting particles there are explicit analytic expressions for the density distribution that describe the system [109, 110, 112]. We have shown that the balance of gravitational forces and wall forces (3.22) holds in the steady sedimentation state for both ideal and mutually interacting active Brownian particles [SH1, SH2]. This may be somewhat surprisingly as the derivation of (3.10) was restricted to equilibrium systems. More details on the sum rule and its derivation in nonequilibrium steady states are given later in section 4.3. Note that the sum rule (3.22) holds independently of the presence and type of interparticle interactions, independently of the form of the particle-wall interactions and also independently of the magnitude of the swim speed and hence particle activity in steady states.

## 3.3 Statistical mechanical applications and generalizations of Noether's theorem

### 3.3.1 Overview

At this point one still might wonder about the usefulness of the above (section 3.2) determined sum rules. Baus already gathered these sum rules in his overview 1984 [70]. Except of equation (3.12) and (3.13) all relations have been previously derived in the literature albeit with a different strategy. One important advantage of our alternative derivation based on Noether's theorem is to be able to generalize. The connection to the Noether theorem makes it easy to apply the concept to further functionals and their invariances. Identifying the sum rules as a consequences of the system symmetry allows to gain a deeper understanding of their fundamental physical meaning. Furthermore, we reemphasize that the given sum rule are exact relations. Hence they are useful for carrying out consistency checks both in computer simulations and in theoretical descriptions and they might build the basis for the development of new theories.

In this section we specify further applications and generalizations of Noether's theorem in statistical mechanics. So far we have investigated the effects of the theorem considering the invariance of the grand canonical potential (subsection 3.2.1) and the invariance of the excess free energy in the canonical ensemble (subsection 3.2.2) with respect to a uniform spatial shift. This shows the vanishing of the global external force (3.10) in equilibrium and can be sharpened to further position dependence using functional differentiation. It is interesting to transfer the above considerations not only to a different (invariant) functional, but also to use further symmetry operations [SH1].

As it turns out the choice of the functional and of the symmetry operation makes it possible to guess the resulting sum rules, in particular the first order expansion

term [SH1,SH2]. For invariance under a spatial displacement one gains information about forces (see section 3.2) whereas invariance under rotation gives sum rules for the torques (see subsection 3.3.3). The functional specifies the force or torque contribution that vanishes globally. Invariance of the grand potential (or the free energy in the canonical ensemble) corresponds to external contributions as we had seen in section 3.2. Internal interaction terms are determined from the symmetry of the excess (above ideal gas) free energy density functional as we show below in subsection 3.3.2. In case of the superadiabatic excess free power functional  $P_t^{\text{exc}}[\rho, \mathbf{J}]$  one gains information about superadiabatic forces and their correlation functions, see subsection 3.3.4. Superadiabatic forces are genuine nonequilibrium contributions that arise in driven systems [78, 79].

The superadiabatic excess free power functional is a functional that appears in the power functional theory (PFT), the nonequilibrium generalization of the density functional theory (DFT). Similar to DFT, PFT is in principle an exact description based on a one-body variational principle [78]. As we only consider invariances of  $P_t^{\text{exc}}[\rho, \mathbf{J}]$  in this thesis we will not describe the power functional theory itself in detail here. For an accessible overview and introduction to the theory, please see the review [79]. The review also contains explicit approximations, and their applications, of the superadiabatic excess free power functional. The term superadiabatic indicates contributions beyond the adiabatic approximation. The adiabatic terms are independent of the flow and they can be determined as a functional of the instantaneous density profile on the basis of the adiabatic construction [78, 79]. The construction relates the real system to a hypothetical equilibrium system with the same one-body density distribution and identical interparticle interaction potential (as the nonequilibrium system) and at a fixed time  $t$ . For readers that are not familiar with power functional theory it might help to consider  $P_t^{\text{ext}}[\rho, \mathbf{J}]$  as the power of those interparticle interaction effects that result from pure nonequilibrium. This adiabatic-superadiabatic splitting is similar in spirit but different in practice to equilibrium ideal-excess splitting where the excess free energy contains all contributions that go beyond the interactions of a (non-interacting) ideal gas.

Due to the nonequilibrium nature of the superadiabatic excess free power this functional depends implicitly (but non explicitly) on time via its functional dependence on the “kinematic” current and density fields. Therefore there are three distinct possibilities to apply a spatial symmetry operation (such as a uniform shift or a rotation). The first and arguably the simplest operation is a displacement that only affects the current at time  $t$ , but not at previous times  $\tau < t$ . We refer to this operation as *instantaneous shifting*. Second, the *uniform shifting* constitutes the same uniform displacement  $\epsilon$  at all times and thus also affects previous times  $\tau$ . The third possibility is the *time-dependent shifting* where the displacement parameter  $\epsilon(t)$  itself changes with time. We only require the parameter to be a continuous function and to vanish for the starting time  $\tau = 0$  and the current time  $t$  for simplicity, i.e.  $\epsilon(0) = \epsilon(t) = 0$ . We give an overview over these time-dependent versions of the spatial shift and the resulting sum rules in subsection 3.3.4.

We consider the invariance with respect to spatially-dependent translations  $\epsilon(\mathbf{r})$  in classical statistical mechanics in subsection 3.3.5 and in quantum statistical mechanics in subsection 3.3.6. Generalizations to nonlinear order contributions and open systems with non-vanishing boundary conditions are treated in subsections 3.3.7 and 3.3.8

respectively. We have also generalized from spherical to uniaxial and biaxial particles, which is in principle straightforward generalization [SH1]. Therefore one adds one or two orientational degrees of freedom that describe the particle orientation besides its spatial position coordinate  $\mathbf{r}$ . Sum rules based on translational invariance stay (basically) valid, with just an additional integration over the orientations to be added. In case of rotational symmetries torque terms appear, as we show in subsection 3.3.3.

In the following subsections 3.3.2 to 3.3.8 we show a summary of a selection of interesting and important results from these different possible combinations and applications. Further sum rules which result from invariant functionals in and out of equilibrium due to Noether's theorem have been presented in references [SH1, SH2, SH3, SH4, SH5].

### 3.3.2 Translational invariance of the free energy

We start the overview of exemplary application of Noether's theorem in statistical mechanics by considering the spatial shift of the intrinsic excess free energy density functional  $F_{\text{exc}}[\rho]$ . This functional arises from the interparticle interaction potential  $u(\mathbf{r}^N)$  and hence describes effects above the ideal gas behaviour, where  $u(\mathbf{r}^N) = 0$ . The excess free energy functional is thus independent of the external potential and only depends on the density distribution. The excess free energy is invariant under a global shift of the density,  $F_{\text{exc}}[\rho(\mathbf{r} + \boldsymbol{\epsilon})] = F_{\text{exc}}[\rho(\mathbf{r})]$ . This holds for all enclosed systems which are confined by impenetrable boundary walls (such that the density distribution vanishes upon penetrating the walls). Functional differentiation of the free energy functional gives the one-body direct correlation function via  $c_1(\mathbf{r}) = -\delta\beta F_{\text{exc}}[\rho]/\delta\rho(\mathbf{r})$ . Alternatively in equilibrium the one-body direct correlation function  $c^{(1)}(\mathbf{r})$  can be expressed with the internal force density distribution [SH3, 79]  $\mathbf{f}_{\text{int}}(\mathbf{r}) = k_{\text{B}}T\nabla c_1(\mathbf{r})$ .

Starting from the invariance of  $F_{\text{exc}}[\rho]$ , functionally Taylor expanding with respect to the density and integrating by parts yields a Noether's sum rule. The technicalities of the derivation are very similar to equation (3.8) and (3.9), so we skip the details of the derivation here. For a detailed determination of the sum rule (3.24) please see reference [SH1]. After some straightforward algebra the sum rule states that

$$0 = k_{\text{B}}T \int d\mathbf{r} \rho(\mathbf{r})\nabla c_1(\mathbf{r}) = \int d\mathbf{r} \rho(\mathbf{r})\mathbf{f}_{\text{int}}(\mathbf{r}), \quad (3.23)$$

where the integral on the right hand side of equation (3.23) over the force density distribution  $\rho(\mathbf{r})\mathbf{f}_{\text{int}}(\mathbf{r})$  can be identified as the global internal force  $\mathbf{F}_{\text{int}}^{\circ} = -\langle\sum_i \nabla_i u(\mathbf{r}^N)\rangle$ . We can hence conclude that there is no global internal force in equilibrium,

$$\mathbf{F}_{\text{int}}^{\circ} = 0. \quad (3.24)$$

The vanishing of the internal force (3.24) is equivalent to Newton's third law, *actio est reactio*. Therefore equation (3.24) is always satisfied for enclosed systems. This property even holds separately for the adiabatic and for the full interaction contribution in nonequilibrium systems. The interaction potential  $u(\mathbf{r}^N)$  depends on all particle positions  $\mathbf{r}^N \equiv \mathbf{r}_1, \dots, \mathbf{r}_N$ , but only on the differences between the particles. (The absolute position in space has no influence on this term.) A straightforward calculation shows (see e.g. [SH1]) that already the global internal force vanishes,  $-\sum_i \nabla_i u(\mathbf{r}^N) = 0$  irrespective of the values of the  $\mathbf{r}^N$  and hence also the average of the total force vanishes.

We show the vanishing of the global internal force in nonequilibrium using Noether's theorem below in subsection 3.3.4.

A vivid application of the determined sum rule (3.24) is provided by one of the stories of the famous Baron of Munchausen. He claimed that he rescued himself (and his horse) out of a swamp by only pulling on his own hair (or on his bootstraps in the English version of the story, where he often rather is a cowboy). Common sense already tells us the story reports obviously an impossibility and the tale is fiction. With the above relation one can immediately identify the truth. Applying the force on himself constitutes an internal contribution. As the total internal force has to vanish, see equation (3.24), there must be another internal contribution that cancels his force. So in the end no net force remains which could lead to any motion.

The sum rule (3.24) is general in the sense that it remains valid upon functional differentiation with respect to the density. This differentiation yields [SH1]:

$$\nabla c_1(\mathbf{r}) = \int d\mathbf{r}' c^{(2)}(\mathbf{r}, \mathbf{r}') \nabla' \rho(\mathbf{r}'), \quad (3.25)$$

where the two-body direct correlation function is defined via  $c^{(2)}(\mathbf{r}, \mathbf{r}') = \delta c^{(1)}(\mathbf{r}) / \delta \rho(\mathbf{r}')$ . Alternatively using Noether's theorem and functional differentiation, equation (3.25) can be also derived from a shift of the density as a functional of the external potential [61, 72], similar to the calculation that determined equation (3.6) in section 3.1. Equation (3.25) is directly connected and hence equation (3.24) is indirectly connected to the LMBW equation (3.5). To show this connection consider the one-body direct correlation function on the left-hand side of equation (3.25) as being expressed with the Euler Lagrange equation,  $c_1(\mathbf{r}) = \ln \rho(\mathbf{r}) \Lambda^d + \beta V_{\text{ext}}(\mathbf{r}) - \beta \mu$ , where  $\Lambda$  is the thermal wavelength and  $d$  denotes spatial dimensionality. Evaluation of the derivative regains the LMBW equation (3.5) if one takes into account that the chemical potential  $\mu$  is a constant.

As an aside one can also consider the consequences of the invariance of the ideal free energy  $F_{\text{id}}[\rho]$  (3.4). One result is a vanishing global ideal diffusive force  $\mathbf{F}_{\text{id}}^{\circ}$  in equilibrium,

$$\mathbf{F}_{\text{id}}^{\circ} = k_{\text{B}} T \int_V d\mathbf{r} \nabla \rho(\mathbf{r}) = 0. \quad (3.26)$$

However, using the divergence theorem to rewrite the definition of the free energy (3.26) yields

$$\mathbf{F}_{\text{id}}^{\circ} = k_{\text{B}} T \int_{\partial V} d\mathbf{s} \rho(\mathbf{r}), \quad (3.27)$$

where  $\partial V$  indicates the boundary of that volume  $V$  and  $d\mathbf{s}$  denotes the vectorial surface element of  $\partial V$ . So expression (3.27) and hence equation (3.26) vanish trivially and do not contain any new information as we have assumed enclosed systems with vanishing density at the boundaries.

### 3.3.3 Rotational invariances of the grand potential and the free energy

We next consider Noether sum rules for rotational symmetry operations. The invariances are with respect to system rotations around an axis  $\mathbf{n}$ , where the absolute



value  $|\mathbf{n}|$  indicates the angle of rotation. This symmetry operation changes the position vector according to  $\mathbf{r} \rightarrow \mathbf{r} + \mathbf{n} \times \mathbf{r}$ . The invariance of the grand potential  $\Omega[V_{\text{ext}}(\mathbf{r} + \mathbf{n} \times \mathbf{r})] = \Omega[V_{\text{ext}}(\mathbf{r})]$  then implies a vanishing global external torque [SH1]:

$$\mathcal{T}_{\text{ext}}^{\circ} = - \int d\mathbf{r} \rho(\mathbf{r}) (\mathbf{r} \times \nabla V_{\text{ext}}(\mathbf{r})) = 0. \quad (3.28)$$

The sum rule (3.28) has some similarities to the conservation law (2.13) in classical mechanics which results from the invariance of the action functional under uniform rotations. These similarities become especially apparent when the latter conservation law (2.13) is differentiated with respect to time and then expresses the vanishing of the global total torque. However, in classical mechanics this relation holds for the total torque, whereas in statistical mechanics the external contribution (3.28) of the torque vanishes separately. This observation is analogous to the case of the uniform shift, considered in section 3.2.1.

For uniaxial particles not only the positional coordinate but also the orientation  $\boldsymbol{\omega}$  of the particles changes under rotation,  $\boldsymbol{\omega} \rightarrow \boldsymbol{\omega} + \mathbf{n} \times \boldsymbol{\omega}$ . It is important that both, the position and the orientation of the particles are rotated, because otherwise the grand potential would not be invariant under the operation in general. This can be seen as follows: Consider two parallel hard rods and only rotate the particle orientations, i.e. rotate each particle around its center but do not rotate the particle positions themselves in space. Then the minimal distance between the rod surfaces will change in general and the particles might even begin to overlap. Hence this rotation clearly affects both the excess free energy and the grand potential. Tarazona and Evans [113] corrected this problem which they identified in the work of Gubbins [114]. For global rotations the grand potential is invariant in systems confined by an external potential and so the global external torque vanishes [SH1, 113],

$$\mathcal{T}_{\text{ext}}^{\circ} = - \int d\mathbf{r} d\boldsymbol{\omega} \rho(\mathbf{r}, \boldsymbol{\omega}) (\mathbf{r} \times \nabla V_{\text{ext}}(\mathbf{r}, \boldsymbol{\omega})) - \int d\mathbf{r} d\boldsymbol{\omega} \rho(\mathbf{r}, \boldsymbol{\omega}) (\boldsymbol{\omega} \times \nabla_{\boldsymbol{\omega}} V_{\text{ext}}(\mathbf{r}, \boldsymbol{\omega})) = 0, \quad (3.29)$$

where  $\nabla_{\boldsymbol{\omega}}$  denotes the orientational derivative with respect to  $\boldsymbol{\omega}$ . Note that for uniaxial particles the external torque (3.29) now consists of two contributions. The first term can be identified as an orbital contribution that originates from the particle rotation around the axis  $\mathbf{n}$ . This term is similar to the torque that acts on spherical particles (3.28). The second term only occurs for anisotropic particles. We refer to this contribution as the spin torque as it originates from each particle rotating around its center, as is induced by carrying out the global system rotation.

Of course not only the grand potential but also the free energy of an enclosed system is an invariant under the rotation symmetry operation. A short derivation in reference [SH1] shows that there is also no global internal torque  $\mathcal{T}_{\text{int}}^{\circ}$  in the system [SH1, 113],

$$\beta \mathcal{T}_{\text{int}}^{\circ} = - \int d\mathbf{r} d\boldsymbol{\omega} \rho(\mathbf{r}, \boldsymbol{\omega}) (\mathbf{r} \times \nabla c^{(1)}(\mathbf{r}, \boldsymbol{\omega})) - \int d\mathbf{r} d\boldsymbol{\omega} \rho(\mathbf{r}, \boldsymbol{\omega}) (\boldsymbol{\omega} \times \nabla_{\boldsymbol{\omega}} c^{(1)}(\mathbf{r}, \boldsymbol{\omega})) = 0. \quad (3.30)$$

This torque also consists of two separate parts, an orbital and a spin torque, which cancel each other. For spherical particles the torque (3.30) only contains the first orbital

term; the second term vanishes trivially. The vanishing of the torque for spherical particles (3.30) and its higher order versions obtained from functional differentiation were determined previously by Baus [70] from the fact that the free energy does not change under rotations. The orientationally dependent equations (3.29) and (3.30) were first derived by Tarazona and Evans [113]. Our new Noetherian derivation [SH1] allows to identify these global relations for the torques as consequences of the rotational system invariances. The details of the explicit application of Noether's theorem to determine the internal (3.30) and the above external torque sum rule (3.29) for uniaxial and also biaxial particles can be found in reference [SH1].

### 3.3.4 Time-dependent translational invariance of the free power functional

In the following we turn from equilibrium to the dynamics of nonequilibrium overdamped systems. In contrast to the above results [such as (3.5) or (3.10), which have been determined before] the following nonequilibrium sum rules are new results [SH1]. To determine these nonequilibrium relations we consider the symmetries of the superadiabatic excess free power functional  $P_t^{\text{exc}}[\rho, \mathbf{J}]$  as is central in power functional theory (PFT) [78, 79]. This functional generates the nonequilibrium forces that result from the interparticle interactions and that are beyond equilibrium accessibility [78, 79]. The functional not only depends on the time-dependent density profile but also on the current distribution  $\mathbf{J}(\mathbf{r}, t)$  of the system. Similar to DFT, the PFT is based on a one-body variational principle and the theory describes overdamped many-body systems out of equilibrium.

As mentioned in the subsection 3.3.1 there are three different kinds of relevant nonequilibrium shifting operations. We consider the instantaneous shift (at the current time  $t$ ) which is like a “kick” of the whole system and hence it does not affect the system at previous times  $\tau < t$ . The transformation leaves internal interactions unchanged. The density remains unchanged, although it will be affected at future times  $\tau > t$ . However, the shift influences the current via generating an additive contribution  $\mathbf{J}(\mathbf{r}, t) \rightarrow \mathbf{J}(\mathbf{r}, t) - \dot{\epsilon}\rho(\mathbf{r}, t) \equiv \tilde{\mathbf{J}}(\mathbf{r}, t)$ . This change in the current has no influence on the functional,  $P_t^{\text{ext}}[\rho, \mathbf{J}] = P_t^{\text{ext}}[\rho, \tilde{\mathbf{J}}]$ , as the chosen description is restricted to overdamped systems where inertial effects are absent. Using Noether's theorem we conclude [SH1] that the global superadiabatic force vanishes,

$$\mathbf{F}_{\text{sup}}^{\text{o}}(t) = \int d\mathbf{r} \rho(\mathbf{r}, t) \mathbf{f}_{\text{sup}}(\mathbf{r}, t) = 0. \quad (3.31)$$

In general the superadiabatic force field  $\mathbf{f}_{\text{sup}}(\mathbf{r}, t)$  as it occurs in equation (3.31) is determined from the (negative) functional derivative of the excess free power functional with respect to the current,  $\mathbf{f}_{\text{sup}}(\mathbf{r}, t) = -\delta P_t^{\text{exc}}[\rho, \mathbf{J}]/\delta \mathbf{J}(\mathbf{r}, t)$ . This force field occurs because of the functional Taylor expansion from applying Noether's theorem. In power functional theory the internal force distribution splits into two contributions,  $\mathbf{f}_{\text{int}}(\mathbf{r}, t) = \mathbf{f}_{\text{ad}}(\mathbf{r}, t) + \mathbf{f}_{\text{sup}}(\mathbf{r}, t)$ . The adiabatic force field  $\mathbf{f}_{\text{ad}}(\mathbf{r}, t)$  is a quasi-equilibrium contribution, so the above equilibrium sum rule (3.24) still holds, i.e.  $\mathbf{F}_{\text{ad}}^{\text{o}}(t) = \int d\mathbf{r} \rho(\mathbf{r}, t) \mathbf{f}_{\text{ad}}(\mathbf{r}, t) = 0$  for a fixed but arbitrary time  $t$ . We conclude that the global internal force has to vanish not only in equilibrium but also in nonequilibrium for any time  $t$  in enclosed systems,  $\mathbf{F}_{\text{int}}^{\text{o}}(t) = \mathbf{F}_{\text{ad}}^{\text{o}}(t) + \mathbf{F}_{\text{sup}}^{\text{o}}(t) = 0$ . This conclusion is

in accordance with Newton's third law which also holds in nonequilibrium and can be also proven on a more fundamental level, see reference [SH1]. The fact that not only the total internal force  $\mathbf{F}_{\text{int}}^{\text{O}}(t)$  vanishes, but also its two separate adiabatic and superadiabatic contributions can be seen as a support of the chosen splitting in power functional theory.

As a second example of a nonequilibrium shift symmetry operation we consider a temporally constant uniform displacement. In contrast to the nonequilibrium instantaneous shifting this transformation does not only affect the current distribution but it also affects the density profile via the change in the position coordinate,  $\mathbf{r} \rightarrow \mathbf{r} + \boldsymbol{\epsilon}$ , at all times  $\tau$ . (Here  $\tau$  describes all previous times and must not be mistaken for the Brownian time scale which does not appear here but is often denoted  $\tau$  as well in the literature.) The time  $t$  indicates the current time. The invariance of the superadiabatic excess free power functional under this shifting,  $P_t^{\text{exc}}[\rho(\mathbf{r}, \tau), \mathbf{J}(\mathbf{r}, \tau)] = P_t^{\text{exc}}[\rho(\mathbf{r} + \boldsymbol{\epsilon}, \tau), \mathbf{J}(\mathbf{r} + \boldsymbol{\epsilon}, \tau)]$ , yields the sum rule [SH1]

$$0 = \int_0^t d\tau \int d\mathbf{r} \left( m_1(\mathbf{r}, \tau, t) \nabla \rho(\mathbf{r}, \tau) + \mathbf{m}_1(\mathbf{r}, \tau, t) \cdot \nabla \mathbf{J}(\mathbf{r}, \tau)^\top \right). \quad (3.32)$$

Here  $m_1(\mathbf{r}, \tau, t) = -\beta \delta P_t^{\text{exc}} / \delta \rho(\mathbf{r}, \tau)$  and  $\mathbf{m}_1(\mathbf{r}, \tau, t) = -\beta \delta P_t^{\text{exc}} / \delta \mathbf{J}(\mathbf{r}, \tau)$  denote respectively a scalar and a vectorial time correlation function [115, 116] and the superscript  $\top$  indicates the transpose of a matrix. Although equation (3.32) appears rather complicated and we still lack an intuitive understanding of  $m_1(\mathbf{r}, \tau, t)$  and  $\mathbf{m}_1(\mathbf{r}, \tau, t)$ , it is an interesting and potentially important relation. It relates the two time correlation functions with the time-dependent density and current. Thus equation (3.32) represents an exact relation for the memory structure of the system, as is manifest via the integration over the past time  $\tau$ . Formulating a feasible approximate description of time-dependent systems with memory is a significant challenge and of potential interest to current research of memory kernels [117–124], see [85–87] for recent work. Therefore, it is useful to be aware of exact expressions such as equation (3.32) which may form the basis of possible approximations for memory kernels.

The nonequilibrium time-dependent shift is a symmetry operation that may be seen as a generalization of the uniform shift applied to the nonequilibrium situation. However, for simplicity we assume that the boundary values of the time-dependent displacement  $\boldsymbol{\epsilon}(\tau)$  vanish, i.e.  $\boldsymbol{\epsilon}(0) = \boldsymbol{\epsilon}(t) = 0$ . At the starting time  $\tau = 0$  the system is assumed to be in equilibrium and  $\tau = t$  indicates the current time. As the superadiabatic excess free power functional for enclosed systems is invariant under this symmetry operation, a Noetherian sum rule ensues,

$$0 = \frac{d}{d\tau} \int d\mathbf{r} \rho(\mathbf{r}, \tau) \mathbf{m}_1(\mathbf{r}, \tau, t) + \int d\mathbf{r} \left( m_1(\mathbf{r}, \tau, t) \nabla \rho(\mathbf{r}, \tau) + \mathbf{m}_1(\mathbf{r}, \tau, t) \cdot \nabla \mathbf{J}(\mathbf{r}, \tau)^\top \right). \quad (3.33)$$

As for the constant in time displacement this relation (3.33) is an exact consequence of the translational symmetry of the system and characterizes the memory of the system via relating different time correlation functions to each other. Note that the second contribution is the integrand of the time integral in equation (3.32).

### 3.3.5 Space-dependent translational invariance of the grand potential

A natural and significant generalization of global symmetry operations as considered above are local transformations such as a local space-dependent shift,  $\boldsymbol{\epsilon}(\mathbf{r})$ . When the functional is invariant with respect to such symmetry operation, Noether's second theorem applies. So far our considerations only exploited Noether's first theorem, as we looked at invariances of functionals under a uniform shift or rotation. These "rigid-body" transformations determined global sum rules such as the vanishing of the global external (3.10) and internal (3.24) force. The application of Noether's second theorem in statistical mechanics is similar to the first theorem, but the more general transformation yields directly locally resolved sum rules. (The time-dependent global shifting of section 3.3.4 can be viewed as application of Noether's second theorem.) We restrict ourselves to equilibrium in the following.

A local displacement  $\boldsymbol{\epsilon}(\mathbf{r})$  in general does not only affect the absolute particle positions, but it also affects particle positions relative to each other and hence the distance vectors between particles change. This in turn influences the interparticle interaction energy, and therefore the grand potential, such that in general the grand potential is not an invariant under the transformation. However, a corresponding (locally resolved) change of the particle momenta can compensate the occurring variations of the functional and restore the symmetry. A detailed description and derivation is given in reference [SH3]. The theory is based on the fact that the grand potential in equilibrium does not change under canonical transformations of the phase space variables [9], as such transformations leave the phase space volume unchanged and they leave, when appropriately chosen, the Hamilton also invariant. Under the combined symmetry operations,  $\mathbf{r} \rightarrow \mathbf{r} - \boldsymbol{\epsilon}(\mathbf{r})$  and  $\mathbf{p} \rightarrow \mathbf{p} + \nabla\boldsymbol{\epsilon}(\mathbf{r}) \cdot \mathbf{p}$ , the grand potential is invariant,  $\Omega = \Omega[\boldsymbol{\epsilon}(\mathbf{r})]$ . (Here  $\boldsymbol{\epsilon}(\mathbf{r})$  is considered to be a small parameter.) Applying Noether's theorem then yields the local force density balance in equilibrium,

$$\mathbf{F}(\mathbf{r}) = -k_{\text{B}}T\nabla\rho(\mathbf{r}) + \rho(\mathbf{r})\mathbf{f}_{\text{int}}(\mathbf{r}) - \rho(\mathbf{r})\nabla V_{\text{ext}}(\mathbf{r}) = 0, \quad (3.34)$$

which consists of three contributions: The first term is the diffusive force density of the ideal gas, the second term indicates the internal force density (due to the interparticle pair potential  $\phi(r)$ ) and the last term represents the external contribution. As expected from Noether's second theorem the force density balance (3.34) depends on the position  $\mathbf{r}$  and is not a global position-independent but a locally resolved sum rule. When one restricts the interparticle interactions to pair interactions with the potential  $\phi(r)$  the internal force density can be rewritten as  $\rho(\mathbf{r})\mathbf{f}_{\text{int}}(\mathbf{r}) = \int d\mathbf{r}' \rho_2(\mathbf{r}, \mathbf{r}')\nabla'\phi(|\mathbf{r} - \mathbf{r}'|)$ , where  $\rho_2(\mathbf{r}, \mathbf{r}')$  denotes the two-body density. This rewriting allows to identify equation (3.34) as the first member of the YBG hierarchy [72, 76], as originally derived by Yvon [104] and Born and Green [105]. The YBG equation was taken as a starting point to develop a new scheme for carrying out classical DFT based on forces [SH3].

As an aside the consideration of rotations of spheres does not gain any new information for a locally dependent rotation axis  $\mathbf{n}(\mathbf{r})$ . This symmetry operation would yield the transformation  $\mathbf{r} \rightarrow \mathbf{r} + \mathbf{n}(\mathbf{r}) \times \mathbf{r} \equiv \mathbf{r} + \boldsymbol{\epsilon}(\mathbf{r})$  which can be seen as a special case of a local spatial displacement  $\boldsymbol{\epsilon}(\mathbf{r})$ .

### 3.3.6 Quantum statistical force balance

Our findings for the application of Noether's theorem in classical statistical mechanics can be straightforwardly generalized to quantum statistical mechanics. We focus here on invariances of the free energy  $F = -k_B T \ln Z$  in the canonical ensemble with the partition sum  $Z = \text{Tr} e^{-\beta H}$ . The Hamiltonian  $H$  of the considered quantum many-body systems is still given by equation (3.1), but the positions and momenta are now taken to be operators. The trace is defined as a sum of expectation values  $\text{Tr} \cdot = \sum_n \langle n | \cdot | n \rangle$ , where the  $|n\rangle$  label a complete orthonormal set of eigenvectors of  $H$ . In position representation the explicit expression of the expectation value is  $\langle n | \cdot | n \rangle = \int d\mathbf{r}^N \phi_n^*(\mathbf{r}^N) \cdot \phi_n(\mathbf{r}^N)$ , where  $|n\rangle = \phi_n(\mathbf{r}^N)$  and the asterisk denotes a complex conjugation. Recall that we use  $\mathbf{r}^N = \mathbf{r}, \dots, \mathbf{r}_N$  as a shorthand. The equilibrium ensemble average is then given by  $\langle \cdot \rangle_{\text{eq}} = \sum_n \langle n | e^{-\beta H} \cdot | n \rangle / Z$ .

As in subsection 3.3.5 we consider invariances with respect to a local space-dependent shift  $\boldsymbol{\epsilon}(\mathbf{r})$ . The new position of each particle is then given by  $\mathbf{r}_i \rightarrow \mathbf{r}_i + \boldsymbol{\epsilon}(\mathbf{r}_i)$ . Additionally also the corresponding momentum changes according to  $\mathbf{p}_i \rightarrow \mathbf{p}_i - \{(\nabla_i \boldsymbol{\epsilon}_i) \cdot \mathbf{p}_i + \mathbf{p}_i \cdot (\nabla_i \boldsymbol{\epsilon}_i)^\top\} / 2$ , where  $\nabla_i$  the differentiation with respect to  $\mathbf{r}_i$  only acts on  $\boldsymbol{\epsilon}_i = \boldsymbol{\epsilon}(\mathbf{r}_i)$  as indicated by the parentheses and the superscript  $\top$  denotes matrix transposition. Note that the momentum transformation is the self-adjoint version of the transformation in classical mechanics, see subsection 3.3.5. The symmetry operation ensures that (in linear order) the commutator relations of position and momentum are preserved (see reference [SH4]), such that the transformation is unitary on Hilbert space and hence denotes a quantum canonical transformation [125]. The free energy is invariant with respect to canonical transformations in general and in particular with respect to the given local transformation depending on the displacement field  $\boldsymbol{\epsilon}(\mathbf{r})$ ,  $F[\boldsymbol{\epsilon}] = F$ . We determined that using Noether's theorem this invariance yields the local quantum force density balance [SH4, 79, 126]

$$\mathbf{F}(\mathbf{r}) = \nabla \cdot \boldsymbol{\tau}(\mathbf{r}) + \mathbf{F}_{\text{int}}(\mathbf{r}) - \rho(\mathbf{r}) \nabla V_{\text{ext}}(\mathbf{r}) = 0. \quad (3.35)$$

The total force  $\mathbf{F}(\mathbf{r})$  (3.35) vanishes locally and contains three contributions. The first term is a consequence of kinetic energy and is given as the divergence of kinematic stress,  $\boldsymbol{\tau}(\mathbf{r}) = \langle \hat{\boldsymbol{\tau}}(\mathbf{r}) \rangle_{\text{eq}}$ . The corresponding locally resolved one-body kinematic stress operator is  $\hat{\boldsymbol{\tau}}(\mathbf{r}) = -\sum_i (\mathbf{p}_i \delta_i \mathbf{p}_i + \mathbf{p}_i \delta_i \mathbf{p}_i^\top) / 2m + \hbar^2 \nabla \nabla^2 \sum_i \delta_i / 4m$ , where  $\delta_i = \delta(\mathbf{r} - \mathbf{r}_i)$  is a shorthand. The second contribution in equation (3.35) is the internal force density  $\mathbf{F}_{\text{int}}(\mathbf{r}) = \langle \hat{\mathbf{F}}_{\text{int}}(\mathbf{r}) \rangle_{\text{eq}}$ , where the one-body interparticle force operator  $\hat{\mathbf{F}}_{\text{int}}(\mathbf{r}) = -\sum_i (\nabla_i u(\mathbf{r}^N)) \delta(\mathbf{r} - \mathbf{r}_i)$  accounts for all interparticle interactions resulting from the interaction potential  $u(\mathbf{r}^N)$ . The third term of the total force (3.35) is the external force density, which originates from the external potential  $V_{\text{ext}}(\mathbf{r})$ . Here the density is  $\rho(\mathbf{r}) = \langle \hat{\rho}(\mathbf{r}) \rangle_{\text{eq}}$  with the density operator  $\hat{\rho}(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$  being identical to its classical definition.

A comparison with the analogous classical transformation (subsection 3.3.5) shows many similarities such as that the corresponding Noether sum rule holds. In both cases it states the local force density balance (3.34) and (3.35). One might have even guessed the quantal identity from the classical analogue beforehand. In the limit of a uniform displacement  $\boldsymbol{\epsilon}$  the Noether sum rule (3.35) reduces to vanishing of the global external force [SH4],  $-\int d\mathbf{r} \rho(\mathbf{r}) \nabla V_{\text{ext}}(\mathbf{r}) = 0$ , which also results classically from the invariance of the canonical free energy under a uniform shift, see equation (3.21). Nevertheless the details in the classical and quantal derivations differ [SH3, SH4].

Although the concept of forces was rarely exploited in quantum mechanics so far, there has recently been drawn some attention and interest to the topic [79, 126, 127]. Tchenkoue *et al.* [126] suggested the construction of exchange-correlation potentials on the basis of force balance equations which they argue to be a systematic approach and avoid problems with differentiability and causality. Tarantino and Ullrich were able to formulate the Kohn-Sham approach to the time-dependent DFT using the time-dependent force balance. The authors suggest that this approach has a more transparent structure than the usual formulation. The application of locally resolved forces also is important within the quantum version of PFT [79, 128]. Forces are also essential to Tokatly's approach [129–132].

### 3.3.7 Nonlinear contributions

We return to our starting case of uniform shifts in classical grand canonical equilibrium. As the considered symmetries are valid for arbitrary values of  $\epsilon$  ( $\mathbf{n}$  for rotations) not only the linear but also all higher orders of the expansion around  $\epsilon = 0$  ( $\mathbf{n} = 0$ ) vanish. These higher nonlinear order sum rules increase in tensor rank by one unit due to the additional tensor contraction with  $\epsilon$  ( $\mathbf{n} \times \mathbf{r}$ ) for every additional order that is taken into account. Therefore the resulting relations become increasingly complex and difficult to interpret. However, considering the second order explicitly is feasible and has been presented in reference [SH5]. From the invariance of the grand potential under a spatial uniform displacement  $\epsilon$  one gets as a second order contribution [SH5]

$$\int d\mathbf{r}d\mathbf{r}' H_2(\mathbf{r}, \mathbf{r}') \nabla V_{\text{ext}}(\mathbf{r}) \nabla' V_{\text{ext}}(\mathbf{r}') = k_B T \int d\mathbf{r} \rho(\mathbf{r}) \nabla \nabla V_{\text{ext}}(\mathbf{r}). \quad (3.36)$$

The left-hand side of equation (3.36) is the twofold integral over the correlation functions of density fluctuations  $H_2(\mathbf{r}, \mathbf{r}')$  multiplied by the gradient of the external potential. This integral balances the average tensorial curvature (i.e. the Hessian) of the external potential on the right-hand side of equation (3.36). Recalling the global external force operator as  $\hat{\mathbf{F}}_{\text{ext}}^{\text{o}} = -\sum_i \nabla_i V_{\text{ext}}(\mathbf{r}_i) = -\int d\mathbf{r} \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \nabla V_{\text{ext}}(\mathbf{r})$  the left-hand side of equation (3.36) can be interpreted as the variance of this force operator,  $\langle \hat{\mathbf{F}}_{\text{ext}}^{\text{o}} \hat{\mathbf{F}}_{\text{ext}}^{\text{o}} \rangle$  [SH5]. Here the angular brackets indicate the grand canonical statistical average. Note that the left hand side of equation (3.36) is structurally similar to equations (3.12) and (3.14), but has an additional spatial derivative (acting on  $V_{\text{ext}}(\mathbf{r})$ ). Strikingly the correlation in equation (3.36) gives a curvature contribution (right hand side), whereas the correlator of external potential and the global external force (3.14) vanishes.

We have gained an alternative expression for the variance of the external force operator (3.36) from an invariance of the system, so apparently there is a deep connection between both quantities. As we have shown in equation (3.10) the average global external force  $\mathbf{F}_{\text{ext}}^{\text{o}} = \langle \hat{\mathbf{F}}_{\text{ext}}^{\text{o}} \rangle$  vanishes, hence the variance of the external force operator is equal to its covariance,  $\langle \hat{\mathbf{F}}_{\text{ext}}^{\text{o}} \hat{\mathbf{F}}_{\text{ext}}^{\text{o}} \rangle = \langle \hat{\mathbf{F}}_{\text{ext}}^{\text{o}} \hat{\mathbf{F}}_{\text{ext}}^{\text{o}} \rangle - \mathbf{F}_{\text{ext}}^{\text{o}} \mathbf{F}_{\text{ext}}^{\text{o}}$ . Alternatively equation (3.36) can be obtained by multiplying the inverse LMBW equation (3.6) with  $\nabla V_{\text{ext}}(\mathbf{r})$  and integrating over space. However, the above derivation from the second order displacement has the clear advantage that the connection to Noether's theorem becomes apparent and that there is no need to determine multiplication factors and the necessary integrals that lead to the meaningful result. In short, Noether's route

is explicit and *constructive*. Translational sum rules similar to equation (3.36) are presented for the internal and the (trivial) ideal force contributions in reference [SH5].

The consideration of invariances has also proven useful to determine the dependence of physical quantities such as the grand potential on the shape of the system [133]. Roth and his coworkers [133] restricted themselves to grand potentials which are invariant with respect to shifts and rotations, continuous and additive. This restriction allowed them to state an expression for the general structure of the shape dependence. In particular these authors investigated the shape dependence of the density, the surface tension [134], and of thermodynamic potentials [133] in the case of hard spheres in contact with curved walls. The analytic expressions these authors determined reveal a strong dependence on the curvature of the potential, in their case the external wall. This potentially could be related to the mean curvature of the external potential has appeared above (3.36).

### 3.3.8 Boundary contributions

So far our considerations of Noether's theorem apply to enclosed systems (i.e. confined by impenetrable external walls) and the respective symmetry operation acts on the whole system. Both restrictions are implicit assumptions in all previous derivations and consequently in the presented results. Hence so far possible boundary terms vanish as the particle density vanishes outside of the system itself.

There are of course interesting physical systems which do not satisfy these (boundary) restrictions. One might wonder whether and to what extent our considerations are valid or whether they can be transferred to more general settings. It turns out that the previously determined results still hold for systems with vanishing *net* boundary conditions. Then all appearing boundary contributions cancel each other and leave the respective sum rule itself unchanged. Examples are systems with periodic boundary conditions, an important case to e.g. represent bulk in computer simulations. Because of the periodically continuation the "left" and "right" boundary terms are exactly equal to each other up to a minus sign. A similar effect cancels the boundary terms of translational invariant directions as often occurring in theoretical descriptions. We have exploited this fact in the above description of the thermal and of active phase separation in section 3.2.3.

However, for systems with boundary conditions that give non-vanishing contributions the determined sum rules need in general to be modified. These modifications are relevant in systems with two different coexisting bulk states as they occur in motility-induced phase separation. The internal Noether sum rule (3.24) applied to this case (in steady state) has the physical interpretation of the pressure balance. Boundary terms are included as we show in subsection 4.3.2.

In the following, we demonstrate the inclusion of boundary contributions with an explicit example: For the proof of the virial version of the well-known hard wall contact theorem boundary terms play a crucial role [SH3, 135]. The theorem can be determined on basis of the force density balance [SH3]

$$-k_B T \nabla \rho(\mathbf{r}) - \rho(\mathbf{r}) \nabla V_{\text{ext}}(\mathbf{r}) - \int d\mathbf{r}' \rho_2(\mathbf{r}, \mathbf{r}') \nabla \phi(|\mathbf{r} - \mathbf{r}'|) = 0, \quad (3.37)$$

where  $\phi(r)$  denotes the pair potential of the particle interactions. In relation (3.37) the force density contributions from the ideal gas (first term), the external potential (second term) and the interparticle interactions (third term) balance each other. Equation (3.37) results from equation (3.34) upon restricting the internal force density  $\rho(\mathbf{r})\mathbf{f}_{\text{int}}(\mathbf{r})$  to pair interaction form. We continue by spatial integration of the force density balance (3.37). The external potential term vanishes up to its contribution at the hard wall,  $k_{\text{B}}T\rho_{\text{w}}\mathbf{e}_{\text{w}}$ , where  $\rho_{\text{w}}$  is the contact density at the hard wall and  $\mathbf{e}_{\text{w}}$  is a unit vector normal to the hard wall. The ideal force density yields after integration the ideal gas pressure  $p_{\text{id}} = k_{\text{B}}T\rho^{\text{b}}$ , where  $\rho^{\text{b}}$  indicates the density of the bulk phase that corresponds to the hard wall. The direction of this contribution is again along  $\mathbf{e}_{\text{w}}$ , normal to the hard wall. Because of the Noether sum rule (3.24) the internal interaction force density in (3.37) vanishes up to boundary terms after integration. As there is no density and no current within the hard wall and as the boundary contributions perpendicular to the hard wall cancel each other due to the planar symmetry, the bulk contribution at large distance away from the wall is the only one remaining. We show that this term is equal to the internal boundary term within the present phase separation example and hence it gives the bulk virial pressure without the ideal gas contribution [SH3, 76]

$$p_{\text{int}} = -\frac{\pi}{2} (\rho^{\text{b}})^2 \int_0^\infty dr r^2 g(r) \frac{d\phi(r)}{dr}, \quad (3.38)$$

where  $g(r)$  is the bulk pair distribution function. A more detailed derivation of this boundary force is given in the section *Phase coexistence* in reference [SH1] as well as section II.D of reference [SH3].

So the full contact theorem states that the density at a hard wall is proportional to the bulk virial pressure  $p_{\text{v}}$  in two dimensions,

$$p_{\text{v}} \equiv p_{\text{id}} + p_{\text{int}} = k_{\text{B}}T\rho_{\text{w}}. \quad (3.39)$$

In contrast to the derivation of Lovett and Baus [135] our route discloses the underlying invariance due to Noether's theorem. The Noether derivation holds for the approximations made in the force-DFT [SH3], which are based on fundamental measure theory (FMT) [73, 136] in the applications of reference [SH3]. The exact theoretical result (3.39) can be verified within numerical evaluation of the force-DFT [SH3, 102]. This implementation of DFT works on the basis of forces and it complements the usual implementation founded on one-body potentials, the so-called potential-DFT (recall the introduction at the beginning of section 3.1). In the potential-based case the contact density is related to the bulk compressibility pressure [SH3, 76]. Note that both, the virial and the compressibility pressure coincide formally one the basis of the (unknown) exact density functional, but do not so, in general, within an approximation for  $F_{\text{exc}}[\rho]$ .

In conclusion, in the present chapter 3 we have shown that Noether's theorem is applicable to statistical mechanics. The probability-based description and the prevalence of fluctuations constitutes no hindrances. Noether's mechanism allows to derive known sum rules such as the LMBW equation (3.5) and it reveals that such relations are consequences of system symmetries. We have shown that Noether's theorem holds in both the canonical and the grand canonical ensemble. We were successful in generalizing the theorem to nonequilibrium applications as the superadiabatic excess



free power functional which yields new memory constraints. With the given examples we hope to have convinced the reader that our considerations are actually useful for problems, such as the presented sedimentation and phase separation of active particles as well as for situations beyond those.

## 4 Interface polarization sum rule

Before we turn from the more general and exact results shown in chapters 2 and 3 to the more concrete but approximate applications, the present section introduces an exact polarization sum rule (section 4.1) and shows its consequences for active Brownian particles. As these particles have an intrinsic orientation that indicates the direction of their self-propulsion and as the sum rule determines the global polarization within a system, it is possible to identify several interesting relations that bear relevance for numerical [137] and experimental work [138]. Such relations include the vanishing of the global polarization up to boundary terms, which we will lay out in section 4.2. In section 4.3 we determine the center of mass velocity for sedimenting active Brownian particles and validate the mechanical pressure balance for motility-induced phase separation in nonequilibrium.

### 4.1 Derivation of an interface polarization sum rule

The interface polarization sum rule is a direct consequence of the continuity equation. Therefore it can be interpreted as an indirect Noether sum rule, following the argumentation of Revzen [139]: According to his considerations the continuity equation is the statistical mechanics analogue of Noether's theorem. In his work Revzen derives the continuity equation as the differential conservation law corresponding to the invariance of the partition function. As a symmetry transformation he used an arbitrary change in the phase factor of the wave function which leaves the partition function, as a physical quantity, unchanged [139]. The invariant quantity is the Lagrangian and not a statistical mechanics functional as considered in this thesis.

The continuity equation for particles with an intrinsic orientation expressed by the unit vector  $\boldsymbol{\omega}$  is given as

$$\frac{\partial}{\partial t}\rho(\mathbf{r}, \boldsymbol{\omega}, t) = -\nabla \cdot \mathbf{J}(\mathbf{r}, \boldsymbol{\omega}, t) - \nabla_{\boldsymbol{\omega}} \cdot \mathbf{J}^{\boldsymbol{\omega}}(\mathbf{r}, \boldsymbol{\omega}, t). \quad (4.1)$$

The temporal change of the density  $\rho(\mathbf{r}, \boldsymbol{\omega}, t)$  is hence compensated with the (negative) divergence of the current. The latter contains the spatial divergence (contradiction with  $\nabla$ ) of the translational current  $\mathbf{J}(\mathbf{r}, \boldsymbol{\omega}, t)$  and the orientational divergence ( $\nabla_{\boldsymbol{\omega}}$ ) of the rotational current  $\mathbf{J}^{\boldsymbol{\omega}}(\mathbf{r}, \boldsymbol{\omega}, t)$ . The continuity equation (4.1) hence states a local particle conservation law.

We assume that the rotational current in equation (4.1) is purely diffusive, such that there act no explicit torques in the system and the rotational current hence originates purely from an orientationally inhomogeneous density distribution,  $\mathbf{J}^{\boldsymbol{\omega}}(\mathbf{r}, \boldsymbol{\omega}, t) =$

#### 4 Interface polarization sum rule

$-D_{\text{rot}}\nabla_{\omega}\rho(\mathbf{r}, \boldsymbol{\omega}, t)$ , where  $D_{\text{rot}}$  indicates the rotational diffusion constant. For simplicity and to focus on the main concept we restrict ourselves to two-dimensional systems with planar geometry. The relevant axis along which the system is nonuniform we denote as the  $x$ -axis. (In general it is possible to consider more complex geometries, e.g. spherical setups [140].) Applying the planar restriction to the continuity equation (4.1) yields

$$\frac{\partial\rho(x, \varphi, t)}{\partial t} = -\frac{\partial J^x(x, \varphi, t)}{\partial x} + D_{\text{rot}}\frac{\partial^2\rho(x, \varphi, t)}{\partial\varphi^2}. \quad (4.2)$$

In two dimensions the orientation  $\boldsymbol{\omega} = (\cos\varphi, \sin\varphi)$  simplifies to a dependence on the angle  $\varphi$  which is measured against the positive  $x$ -axis and angular Laplace operator  $\nabla_{\boldsymbol{\omega}}^2$ , is simply a second derivative with respect to the angle,  $\partial^2/\partial\varphi^2$ .

It is useful to expand the density in spherical moments as this allows to directly evaluate the orientational derivatives. In the two-dimensional case such an expansion is the angular Fourier series as a function of the angle  $\varphi$ . The Fourier expansion of the density is then given as

$$\rho(x, \varphi, t) = \sum_{n=0}^{\infty} \left[ \rho_n^c(x, t) \cos(n\varphi) + \rho_n^s(x, t) \sin(n\varphi) \right], \quad (4.3)$$

where  $\rho_n^c(x, t)$  indicates the  $n$ th cosine Fourier coefficient and  $\rho_n^s(x, t)$  denotes the  $n$ th sine Fourier coefficient. Similarly the angular Fourier expansion of the  $x$ -component of the current is

$$J^x(x, \boldsymbol{\omega}, t) = \sum_{n=0}^{\infty} \left[ J_n^{x,c}(x, t) \cos(n\varphi) + J_n^{x,s}(x, t) \sin(n\varphi) \right], \quad (4.4)$$

where  $J_n^{x,c}(x, t)$  [ $J_n^{x,s}(x, t)$ ] denotes the corresponding  $n$ th cosine [sine] Fourier coefficient [SH6].

Inserting these expansions into the continuity equation (4.2) and evaluating the angular derivatives gives

$$\begin{aligned} & \sum_{n=0}^{\infty} \left[ \frac{\partial\rho_n^c(x, t)}{\partial t} + \frac{\partial J_n^{x,c}(x, t)}{\partial x} + D_{\text{rot}}n^2\rho_n^c(x, t) \right] \cos(n\varphi) \\ & + \sum_{n=0}^{\infty} \left[ \frac{\partial\rho_n^s(x, t)}{\partial t} + \frac{\partial J_n^{x,s}(x, t)}{\partial x} + D_{\text{rot}}n^2\rho_n^s(x, t) \right] \sin(n\varphi) = 0. \end{aligned} \quad (4.5)$$

Because of the independence of the Fourier basis functions  $\cos(n\varphi)$  and  $\sin(n\varphi)$  each other the expressions in the square brackets have to vanish separately for each  $n$  in order that equation (4.5) is satisfied in general. Here we focus on the second contributions to the sum,  $n = 1$ , although analog considerations apply straightforwardly to all higher order Fourier contributions,  $n > 1$ . The first Fourier coefficient of the density is interesting as it is proportional to the polarization  $\mathbf{M}(\mathbf{r}, t)$ , i.e. the first momentum

of the density distribution with respect to the particle orientation,

$$\mathbf{M}(\mathbf{r}, t) = \int d\boldsymbol{\omega} \boldsymbol{\omega} \rho(\mathbf{r}, \boldsymbol{\omega}, t) \quad (4.6)$$

$$= \int_0^{2\pi} d\varphi \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \sum_{n=0}^{\infty} \left[ \rho_n^c(x, t) \cos(n\varphi) + \rho_n^s(x, t) \sin(n\varphi) \right] \quad (4.7)$$

$$= \pi \begin{pmatrix} \rho_1^c(x, t) \\ \rho_1^s(x, t) \end{pmatrix}, \quad (4.8)$$

where we have inserted the orientation as  $\boldsymbol{\omega} = (\cos \varphi, \sin \varphi)$  and the Fourier series expression of the density (4.3) to determine (4.7). Evaluation of the angular integral shows that the polarization corresponds to the first Fourier components (4.8). Note that sometimes the polarization is instead defined as a quantity per unit radiant,  $\mathbf{M}(\mathbf{r}, t) = \int d\boldsymbol{\omega} \boldsymbol{\omega} \rho / \int d\boldsymbol{\omega}$ , see e.g. reference [137, 138]. To compare one just needs to divide our polarization by  $\int d\boldsymbol{\omega} = 2\pi$ .

Inserting the definition of the polarization (4.8) into the  $n = 1$  contribution of equation (4.5) gives

$$\frac{\partial \mathbf{M}(x, t)}{\partial t} = -D_{\text{rot}} \mathbf{M}(x, t) - \pi \frac{\partial}{\partial x} \begin{pmatrix} J_1^{x,c}(x, t) \\ J_1^{x,s}(x, t) \end{pmatrix}, \quad (4.9)$$

where the cosine (sine) contribution of the continuity equation (4.5) corresponds to the  $x(y)$ -component of the polarization and we have exploited that the temporal derivative and the orientational integration commute with each other. The spatial distribution and temporal evolution of the first Fourier components of the current  $J_1^{x,c}$  and  $J_1^{x,s}$  are usually unknown and they are influenced by the particle polarization. Thus it is useful to further integrate equation (4.9) over the whole system volume,  $L \int_{x_1}^{x_2} dx$ , where  $x_1$  indicates the left and  $x_2$  the right boundary of the system. The translational invariance of the system along the  $y$ -axis allows to directly evaluate the  $y$ -integral, which results in the system length  $L$  in this direction. The locally resolved polarization hence becomes the global polarization  $\mathbf{M}^o(t) = L \int_{x_1}^{x_2} dx \mathbf{M}(x, t)$  of the entire system. Integrating in this way the entire equation (4.9) determines the global polarization by the resulting differential equation:

$$\frac{d\mathbf{M}^o(t)}{dt} = -D_{\text{rot}} \mathbf{M}^o(t) - \pi L \left[ \begin{pmatrix} J_1^{x,c}(x_2, t) \\ J_1^{x,s}(x_2, t) \end{pmatrix} - \begin{pmatrix} J_1^{x,c}(x_1, t) \\ J_1^{x,s}(x_1, t) \end{pmatrix} \right], \quad (4.10)$$

which now only depends on  $D_{\text{rot}}$ ,  $L$  and the first Fourier moments of the current evaluated at the boundaries. Considering the trivial case in the limit of spherical particles which have no preferred orientation, the local and hence also the global orientation vanishes. In that case equation (4.9) will reduce to the trivial identities  $\mathbf{M}(x, t) = 0$  and hence  $\mathbf{M}^o(t) = 0$ .

## 4.2 Interface polarization of active Brownian particles

It is instructive to evaluate equation (4.10) for different boundary situations. One scenario is the vanishing of the occurring boundary currents, which we consider

exemplary for the case of active sedimentation below in subsection 4.2.1. Another case is the presence of bulk states at the boundaries, which occurs for example in MIPS of active Brownian particles (see subsection 4.2.2). Additionally to the self-propelled motion along their particle orientation, the swimmers undergo both translational and rotational diffusion. As we assume the absence of external torques the rotational motion of the active particles is purely diffusive with the rotational diffusion constant  $D_{\text{rot}}$  and hence the differential equation (4.10) is indeed applicable.

#### 4.2.1 Active sedimentation

We first consider sedimentation of active Brownian particles that are bounded by a lower wall with gravity acting along the  $x$ -axis [109–112]. (Hence  $x$  is taken as the vertical coordinate.) The particle density distribution vanishes for both limits:  $\rho(x \rightarrow -\infty) = 0$  due to the lower confining wall and  $\rho(x \rightarrow \infty) = 0$  due to the counter-acting gravitational force. Hence not only the density but also the current vanishes in these limits. The solution of the differential equation for the global polarization (4.10) in case of vanishing boundary currents is

$$\mathbf{M}^o(t) = \mathbf{M}^o(0)e^{-D_{\text{rot}}t}, \quad (4.11)$$

where  $\mathbf{M}^o(0)$  indicates an initial global polarization at time  $t = 0$ . According to equation (4.11) the global polarization decays exponentially with the inverse decay time  $D_{\text{rot}}$ . This behaviour is independent of the presence of interparticle forces and independent of the type of particle-wall interactions. Therefore the sedimenting active particles cannot develop a global polarization in the long time steady state limit, where we obtain from equation (4.11) and  $t \rightarrow \infty$  simply:

$$\mathbf{M}^o = 0. \quad (4.12)$$

This result is entirely plausible as there is no reason why a global polarization should develop in view of the absence of any acting torques and the overdamped character of the dynamics. However, active particles tend to develop a local polarization in the proximity of a wall pointing typically towards the wall. This effect has been observed in simulations and experiments [29, 109–111]. Using this observation we can draw the important conclusion that in these systems there must be also a region (in further distance of the boundary) where the particles are oriented away from the wall on average. Then both local polarizations need to compensate each other in steady state to satisfy equation (4.12).

#### 4.2.2 Motility-induced phase separation

In our second polarization example we consider the formation of two different and coexisting bulk states of active particles, where one phase forms at the left and the other phase forms at the right boundary. We assume a steady state, where there are no temporal changes of the global polarization,  $d\mathbf{M}^o/dt = 0$ , which simplifies the differential equation (4.10) to an algebraic equation. Furthermore in steady state the system becomes symmetric with respect to reflection at the  $x$ -axis, i.e.  $y \rightarrow -y$  and  $\varphi \rightarrow -\varphi$ , due to the two-dimensional planar geometry. Thus the  $x$ -component of the

current becomes an even function in the angle  $\varphi$  and only even terms contribute to the angular Fourier expansion,  $J^x(x, \boldsymbol{\omega}, t) = \sum_n J_n^{x,c}(x, t) \cos(n\varphi)$ . All sine contributions vanish,  $J_n^{x,s}(x, t) = 0$ . Inserting into equation (4.10) yields that in steady state there is no  $y$ -component of the global polarization,  $M_y^o = 0$ . The absolute value of the global polarization consists solely of its  $x$ -component,  $M^o = |\mathbf{M}^o| = M_x^o$ .

Motility-induced phase separation (MIPS) of active Brownian particles (in steady state), although being a genuine nonequilibrium effect, shares many similarities with equilibrium vapour-liquid phase separation. However, the phenomenon can occur for sufficiently high density and self-propulsion speed of the swimmers even for purely repulsive interparticle interactions. Cates and Tailleur [38] suggest an explanation by the effective slowing down and hence accumulation of the particles in denser regions. When the swimmers accidentally bump into each other, their direction first has to change via diffusive rotation before they can swim on. For sufficient large densities meanwhile other particles get stuck and hinder them from escaping, which these authors [38] interpret as the origin of the phase separation.

We choose the planar geometry such that the system forms the more dilute gaseous-like bulk phase for  $x \rightarrow x^g = x_1$  and the dense liquid-like bulk phase for  $x \rightarrow x^l = x_2$  with a free interface in between. Because of the homogeneous density distribution in bulk the ideal and internal force contributions only affect the magnitude of the swim current but not its direction along the internal particle orientation  $\boldsymbol{\omega}$ . Therefore the bulk current only consists of its first Fourier component,  $J_1^{x,c}(x_1)\boldsymbol{\omega}$  or  $J_1^{x,c}(x_2)\boldsymbol{\omega}$ . We implicitly assume here that there is no orientation-independent current contribution  $J_0^{x,c}(x)$ , which would also contribute to the bulk current. For phase-separated active Brownian particles in steady states the global polarization (4.10) is hence [SH6]:

$$M^o = \frac{\pi L}{D_{\text{rot}}}(J^g - J^l), \quad (4.13)$$

where  $J^g = J_1^{x,c}(x^g)$  ( $J^l = J_1^{x,c}(x^l)$ ) is the magnitude of the current of the gaseous (liquid) bulk phase. Thus the magnitude of the global polarization across the interface is different of the two bulk values of the current. Hence the value of  $M^o$  can be seen as a consequence of the bulk states alone and hence it can be interpreted as a state function. This finding seriously questions the claims of Solon *et al.* [45, 48] that the polarization influences and controls the bulk states at coexistence. Given that the global polarization is a state function, the suggested coupling of interface to bulk seems implausible. The direction of  $M^o$  is oriented towards that bulk phase that possesses the higher current. (The bulk states themselves do not develop a local polarization as these phases are translationally and rotationally symmetric and have no preferred direction.) Particles interacting via the Weeks-Chandler-Anderson potential, which is a truncated and shifted Lennard-Jones potential that is hence purely repulsive, are oriented on average towards the denser phase [SH6, 46].

### 4.2.3 Experimental and numerical verification: the Leipzig system

The exact sum rule (4.13) has been tested and verified in several different ways. As an example we have successfully checked its validity in our Brownian dynamics simulations [SH6, 46] for the case of motility-induced phase separation. Other systems that possess bounding bulk states and hence satisfy the polarization sum rule (4.13)

include the motion of a single active particle either in a ratchet potential [141] or under a local abrupt activity step [137, 138, 142]. In the latter system the change of the local activity causes the formation of two regions, where the particle is found to have a higher probability for being in the low activity region. The phenomenology of the resulting density distribution is comparable to that of MIPS, although the underlying physical mechanisms that yield the respective distribution are clearly very different. The system was investigated in the groups of Cichos and Kroy [137, 138] at the University of Leipzig and these authors confirmed experimentally and numerically the validity of our expression for the global polarization (4.13). We give a brief account of their work in the following.

The experimental active particle was a Janus particle realized as a polystyrene spherical colloid half coated with gold. Radiation of a laser locally heats the particle and propels it forward [32, 143, 144]. The free swim speed depends on the laser intensity. The activity step and the system boundaries are realized by the so-called photon nudging [30–33] which is based on a feedback mechanism. The particle position is tracked and the laser is switched on if the particle is located in the active region of the volume and switched off when the colloid is located within the passive region. Different levels of activities were realized by changing the laser intensity [138]. The boundaries are implemented in a similar way but with an additional tracking of the particle orientation. When the particle moves outside of the actual system volume the laser and hence the activity is only switched on when the particle points towards the system. The process successively nudges the particle back. In this system, Söker *et al.* [138] find that the global polarization vanishes. This finding is in accordance with our equation (4.13) as in the experimental setup the current also vanishes beyond the photon nudging boundary. Hence the measured polarization is an experimental confirmation of the polarization sum rule (4.13). Their result allows the authors in *Physical Review Letters* [138, p. 1] “to experimentally confirm, on the single-particle level, that the interfacial polarization is emerging from unbalanced hidden bulk currents” and they cite [SH6] as their source. This is a further indication that the suggested influence of the interfacial polarization on the bulk states [45, 48] does not exist in the present system [SH6].

The authors also performed numerical calculations on that system [137]. The corresponding numerical results of the Smoluchowski (Fokker-Planck) equation were in good agreement with their theoretical truncated moment expansion. Auschra *et al.* [137] were able to determine the expression (4.13) for the global polarization within their theoretical treatments and verify its validity for their exact numerical and approximate theoretical considerations, i.e. a truncated orientational momentum expansion of the dynamic probability density which evolves in time according to a corresponding Fokker-Planck equation. Hence these authors could show that the polarization sum rule (4.13) is “verified within [their] approximate theory” [137, p. 6].

### 4.3 Combining polarization and Noether sum rules

The above section 4.2 shows that the polarization sum rules indeed constitute useful relations. They can serve as a fast consistency check for results from experiments or from simulations. Furthermore these relations allow to make statements about the

validity of or formulate restrictions on approximations in theories. Here we demonstrate that it can be very beneficial to combine the polarization sum rule with the Noether sum rules. For convenience we use the same exemplary systems as above, i.e. active sedimentation and active phase separation in two-dimensional planar geometry.

We start with the general orientation- and position-resolved one-body force density balance of active Brownian particles [90],

$$\begin{aligned} \gamma \mathbf{J}(\mathbf{r}, \boldsymbol{\omega}, t) = & -k_{\text{B}}T \nabla \rho(\mathbf{r}, \boldsymbol{\omega}, t) + \rho(\mathbf{r}, \boldsymbol{\omega}, t) \mathbf{f}_{\text{int}}(\mathbf{r}, \boldsymbol{\omega}, t) + \gamma s \rho(\mathbf{r}, \boldsymbol{\omega}, t) \boldsymbol{\omega} \\ & + \rho(\mathbf{r}, \boldsymbol{\omega}, t) \mathbf{f}_{\text{ext}}(\mathbf{r}, \boldsymbol{\omega}, t), \end{aligned} \quad (4.14)$$

where  $s$  denotes the free swim speed of the active particle. The (negative) friction force density on the left-hand side is balanced by several contributions on the right hand side: The ideal contribution (first term), the internal contribution originating from the interparticle interactions (second term), the swim force density (third term) and the external force density (fourth term).

### 4.3.1 Active sedimentation

In case of active sedimentation the external force consists of the force that the lower confining wall exerts on the particles and the gravitational force, i.e.  $\mathbf{f}_{\text{ext}}(x, \varphi, t) = (f_{\text{wall}}(x, \varphi, t) - mg) \mathbf{e}_x$ . Here  $f_{\text{wall}}(x, \varphi, t)$  is the magnitude of the force that the wall exerts on a particle,  $m$  denotes the particle mass,  $g$  is the gravitational acceleration and  $\mathbf{e}_x$  indicates the unit vector in the  $x$ -direction. The restriction to two-dimensional planar geometry again simplifies the position dependence on  $\mathbf{r}$  to an  $x$ -dependence along the axis of density inhomogeneity. The orientation  $\boldsymbol{\omega}$  reduces to the angle  $\varphi$ , which measures the particle orientation against the  $x$ -axis as before. Inserting the form of the external force in the force density balance (4.14) and exploiting the two-dimensional planar geometry yields

$$\begin{aligned} \gamma \mathbf{J}(x, \varphi, t) = & -k_{\text{B}}T \frac{\partial \rho(x, \varphi, t)}{\partial x} \mathbf{e}_x + \rho(x, \varphi, t) \mathbf{f}_{\text{int}}(x, \varphi, t) + \gamma s \rho(x, \varphi, t) \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix} \\ & + \rho(x, \varphi, t) f_{\text{wall}}(x, \varphi, t) \mathbf{e}_x - mg \rho(x, \varphi, t) \mathbf{e}_x. \end{aligned} \quad (4.15)$$

To proceed we integrate equation (4.15) over all possible positions and orientations. As before the density vanishes at the limits of the  $x$ -integration such that there are no boundary terms. Therefore both the global ideal and the global internal force density contributions vanish due to the Noether sum rules (3.24) and (3.26) at all times  $t$ . One can relate the integrated swim force density to the global polarization (4.8) and thus the corresponding polarization sum rule (4.11) for the case of vanishing boundary currents applies. We define the global external force exerted from the particles on the wall as  $F_{\text{wall}}^{\text{o}}(t) = \int d\mathbf{r} d\boldsymbol{\omega} \rho(\mathbf{r}, \boldsymbol{\omega}, t) f_{\text{wall}}(\mathbf{r}, \boldsymbol{\omega}, t)$  and identify the total number of particles as  $N = \int d\mathbf{r} d\boldsymbol{\omega} \rho(\mathbf{r}, \boldsymbol{\omega}, t)$ . With these considerations the integral over equation (4.15) becomes

$$\gamma N \mathbf{v}_{\text{cm}}(t) = \gamma s \mathbf{M}^{\text{o}}(0) e^{-D_{\text{rot}} t} + F_{\text{wall}}^{\text{o}}(t) \mathbf{e}_x - mg N \mathbf{e}_x \quad (4.16)$$

where  $\mathbf{v}_{\text{cm}}(t)$  is the center of mass velocity  $\mathbf{v}_{\text{cm}}(t) = \int d\mathbf{r} d\boldsymbol{\omega} \mathbf{J} / \int d\mathbf{r} d\boldsymbol{\omega} \rho$ . Equation (4.16) describes the global motion of the system via the time evolution of its center of

mass. Solely the motion of the  $x$ -component is nontrivial as both the wall force and gravity only act along this direction.

In the long-time limit the sedimentation process evolves to a steady state where the general motion, as characterized by the center-of-mass velocity  $\mathbf{v}_{\text{cm}}(t)$ , and the global polarization (4.11) both vanish. One has hence shown that equation (4.16) constitutes

$$F_{\text{wall}}^{\circ} = mgN \quad (4.17)$$

which is remarkably the same external Noether sum rule as for thermal sedimentation in equilibrium (3.22). Hence the gravitational force balances the global force on the wall for sedimenting active Brownian particles in steady state. This relation (4.17) allows to determine the weight of the active particles by simply measuring the force on the lower confining wall after the system reached a steady state. For doing so, it is not necessary that the particles are gathered at the bottom of the system, but they can rather swim freely, which might be surprising at the first glance.

### 4.3.2 Motility-induced phase separation

For motility-induced phase separation the force density balance (4.14) reduces in steady state to

$$\gamma \mathbf{J}(\mathbf{r}, \boldsymbol{\omega}) = -k_{\text{B}}T \nabla \rho(\mathbf{r}, \boldsymbol{\omega}) + \rho(\mathbf{r}, \boldsymbol{\omega}) \mathbf{f}_{\text{int}}(\mathbf{r}, \boldsymbol{\omega}) + \gamma s \rho(\mathbf{r}, \boldsymbol{\omega}) \boldsymbol{\omega}, \quad (4.18)$$

as no external force is present. We integrate again over the whole system volume and all orientations in order to be able to apply the previously determined global sum rules. Here the  $x$ -integral reaches from the gaseous to the liquid bulk phase. The integrated current is assumed to vanish,  $\int d\mathbf{r} d\boldsymbol{\omega} \mathbf{J}(\mathbf{r}, \boldsymbol{\omega}) = 0$ , so the interface between the two phase-separated bulk states is fixed in space. Nevertheless the integrated right-hand side of equation (4.18) has non-vanishing contributions, i.e. boundary terms from the bulk states, which have to be taken into account. The detailed derivation of these contributions is given in reference [SH1]. Here we only give an overview over the derivation of the mechanical pressure balance in nonequilibrium.

The spatial and orientational integral over the diffusive force is  $(p_{\text{id}}^{\text{l}} - p_{\text{id}}^{\text{g}})L\mathbf{e}_x$ , where  $L$  denotes again the interfacial length and we identified the ideal pressure  $p_{\text{id}}^{\text{b}} = k_{\text{B}}T\rho^{\text{b}}$ . The bulk density  $\rho^{\text{b}}$  of the liquid or gas bulk state is indicated by the index  $\text{b} \in \{\text{g}, \text{l}\}$  and is equal to the rotationally integrated zeroth Fourier mode of the density:  $\rho^{\text{b}} = 2\pi\rho_0$ . Using the Noether sum rule (3.24) and taking boundary terms into account the global internal force is given by  $\int d\mathbf{r} d\boldsymbol{\omega} \rho(\mathbf{r}, \boldsymbol{\omega}) \mathbf{f}_{\text{int}}(\mathbf{r}, \boldsymbol{\omega}) = (p_{\text{int}}^{\text{l}} - p_{\text{int}}^{\text{g}})L\mathbf{e}_x$ . It contains the internal pressure in two dimensions given in equation (3.38). Here the virial pressure  $p_{\text{v}} = p_{\text{id}} + p_{\text{int}}$  results from the ideal and internal contributions as a boundary contribution, see subsection 3.3.8 and references [SH1, SH3]. We use the polarization sum rule (4.13) to evaluate the integrated self-propulsion term, which yields  $(p_{\text{swim}}^{\text{l}} - p_{\text{swim}}^{\text{g}})L\mathbf{e}_x$ , where the swim pressure is defined as  $p_{\text{swim}}^{\text{b}} = \gamma s \pi J^{\text{b}} / D_{\text{rot}}$  [39]. The current in the bulk does only consist of the first orientational Fourier contribution  $J^{\text{b}} = J_1^{x,c}(x^{\text{b}})$ . Note that there are various definitions of the swim pressure in the literature [39, 44, 145–148], including the virial expression [47, 146–148]. In bulk these pressures agree with each other and with our swim pressure, which we define by the fact that its negative gradient yields the corresponding swim force density [46, 94].



However, the local active pressures across the interface usually differ and no universal consensus has been reached over this rather fundamental physical quantity.

The integral of the global force density balance (4.18) can be hence expressed as a pressure balance,

$$p^g \equiv p_{\text{id}}^g + p_{\text{int}}^g + p_{\text{swim}}^g = p_{\text{id}}^l + p_{\text{int}}^l + p_{\text{swim}}^l \equiv p^l, \quad (4.19)$$

where the total pressure includes the ideal, internal and swim contribution,  $p^b = p_{\text{id}}^b + p_{\text{int}}^b + p_{\text{swim}}^b$ . For each of these contributions there are explicit derivations of the well-known expressions [SH1]. So the well-known mechanical pressure balance at phase coexistence [76],  $p^g = p^l$ , still holds in this nonequilibrium active system.

Note that the pressure balance of the equilibrium vapour-liquid phase separation in absence of an external field can be derived similarly from the force density balance (4.18) and the internal Noether sum rule (3.24) [SH1]. The main difference is the absence of a free swim speed (i.e.  $s = 0$ ), so application of the polarization sum rule is not necessary and no swim pressure contributes to the pressure balance (4.19),  $p^b = p_{\text{id}}^b + p_{\text{int}}^b$ .

We emphasize that our theoretical description of the phase coexistence of active Brownian particles [46, 94] does indeed satisfy the pressure balance (4.19) and both the underlying Noether (3.24) and polarization sum rules (4.13). The accordance with these exact relations supports our theory, especially as a wider community has not yet agreed on a common theoretical description of MIPS [44, 45, 48]. This disagreement becomes particularly apparent in the consideration of the interfacial tension where there is not even consensus on neither its sign nor on the magnitude. For convenience we summarize the main ideas of the theory in appendix A and determine the interfacial tension on the basis of our description in the next chapter 5.

## 5 Free interfacial tension of active Brownian particles

In this chapter we present our results for the interfacial tension of phase-separated two-dimensional active Brownian particles. Therefore we first specify the system that we are interested in and give a short overview of the existing literature about the free interfacial tension in these systems (section 5.1). In section 5.2 we give a short description how we determined the interfacial tension by transferring and applying the van der Waals-route to nonequilibrium. The interfacial tension we calculate is positive [SH7], which is in accordance with our physical intuition but contradicts many of the previous findings [47, 49–58].

## 5.1 Literature overview over the free interfacial tension of active systems

We consider systems of  $N$  active Brownian particles which satisfy the following well-known Langevin equations of motion

$$\gamma \dot{\mathbf{r}}_i(t) = -\nabla_i u(\mathbf{r}^N) + s\gamma \boldsymbol{\omega}_i(t) + \boldsymbol{\chi}_i(t), \quad (5.1)$$

where  $\gamma$  denotes the friction coefficient,  $\dot{\mathbf{r}}_i(t) = d\mathbf{r}_i(t)/dt$  indicates the velocity of the  $i$ th particle,  $\nabla_i$  is the divergence with respect to the position coordinate  $\mathbf{r}_i(t)$  of the  $i$ th particle at time  $t$  and  $u(\mathbf{r}^N)$  represents the internal interaction potential, which depends on the positions of all particles  $\mathbf{r}^N \equiv \mathbf{r}_1, \dots, \mathbf{r}_N$ . For pair particle interaction potentials  $\phi(r)$  we can rewrite  $u(\mathbf{r}^N) = \sum_{i < j} \phi(|\mathbf{r}_i - \mathbf{r}_j|)$  and in the following we focus on the purely repulsive Weeks-Chandler-Anderson interparticle interaction potential  $\phi(r)$  with the particle size  $\sigma$ . The stochastic force  $\boldsymbol{\chi}_i(t)$  models the Brownian diffusive motion. Here it represents a Gaussian white noise, with vanishing average  $\langle \boldsymbol{\chi}_i(t) \rangle = 0$  and a delta-correlated variance  $\langle \boldsymbol{\chi}_i(t) \boldsymbol{\chi}_j(t') \rangle = 2k_B T \gamma \delta_{ij} \delta(t - t') \mathbf{1}$ , where  $k_B$  is the Boltzmann constant,  $T$  is the temperature,  $\delta_{ij}$  denotes a Kronecker delta,  $\delta$  indicates the Dirac delta distribution, and  $\mathbf{1}$  represents the  $2 \times 2$  unit matrix. The second term on the right-hand side of equation (5.1) generates the self-propulsion of the active particle. This force is proportional to the free swim speed  $s$  and points along the particle orientation  $\boldsymbol{\omega}_i(t)$ . In two dimensions the orientation can be written as  $\boldsymbol{\omega}_i(t) = (\cos \varphi_i(t), \sin \varphi_i(t))$ , where the particle angle  $\varphi_i$  is measured against the  $x$ -axis and undergoes free rotational diffusion,

$$\dot{\varphi}_i(t) = \chi_i^{\text{rot}}(t), \quad (5.2)$$

where the random torque has again a vanishing mean,  $\langle \chi_i^{\text{rot}}(t) \rangle = 0$  and a delta correlated variance,  $\langle \chi_i^{\text{rot}}(t) \chi_j^{\text{rot}}(t') \rangle = 2D_{\text{rot}} \delta_{ij} \delta(t - t')$  with the rotational diffusion constant  $D_{\text{rot}}$ .

Somewhat surprisingly at first glance, the active Brownian particles tend to phase separate into an dense and a dilute phase, if the particle density and the free swim speed are high enough, even though the interparticle interaction potential is purely repulsive. This phase separation was already mentioned and described in the above applications of the Noether and polarization sum rules, see the subsections 4.2.2 and 4.3.2. The phenomena occurs due to the swimming or motility of the active particles and is therefore called motility-induced phase separation (MIPS), which most readers familiar with active matter will know. The reviews [21–27] give much background for those who are unfamiliar with the topic or those who are interested in further details in the behaviour of active particles and their theoretical description.

Important properties, which are interesting to investigate and useful to gain new insights, are interfacial quantities of these systems, such as the interfacial tension considered here. The first study that actually measured in simulation of equations (5.1) and (5.2) the tension of the free interface for phase separated active Brownian particles is that by Bialké *et al.* [47]. These authors surprisingly determined within their simulations a huge negative value of  $\gamma_{\text{gl}} = -475k_B T / \sigma^2$  for the tension. This result lead to a controversy in this field concerning the magnitude and sign of the surface tension in active systems. The discussion is still ongoing as can be seen in the

recent work by Chacón *et al.* [56] where they comment that for nonequilibrium phase separations “a clear agreement has not been reached on the physical properties such as the surface tension.” [56, p. 2647]. While the controversy is apparent in direct personal exchanges with researchers at conferences [149] and elsewhere, very little criticism has been published, with our publication [SH7] being an exception. In the following we hence give a short overview over the main theoretical concepts to determine this interfacial tension.

In equilibrium there are several different routes to determine the interfacial tension (e.g. via the capillary wave spectrum, the integral over the difference in the normal and transversal components of the pressure tensor or via the free energy cost of forming the interface), which all lead to the same result, i.e. to the same value of the interfacial tension [76]. It has been argued that it is a priori not clear whether this equivalence of the mechanical and thermodynamic routes also holds in nonequilibrium, especially because it is unclear how a thermodynamic quantity such as the free energy is defined in nonequilibrium and whether such a definition would be useful. It turned out that both routes indeed do not agree, which was interpreted as an indication that the free energy does not exist for active particles [51,52]. However, mechanically the interfacial tension should be a unique quantity determined by its physical meaning. One might define the free interfacial tension by its mechanical interpretation and hence consider the thermodynamic route just as a rewriting, which is entirely possible in equilibrium. Following the equilibrium interpretation a negative tension would mean a negative energetic cost for the system to form an interface. Therefore a maximal length of the interface would be most beneficial which is equivalent to a homogeneous state. In direct interpretation the phase separation would not be stable with a negative interfacial tension in equilibrium. (A number of stability mechanisms were considered in the literature [47,51].)

There are other theoretical and simulation studies [49–52] that seemingly support the determined negative value obtained by Bialké *et al.* [47]. Some theories [45,48] support both positive and negative values of the interfacial tension. In other studies the determined tension is approximately zero [53,54] or exactly zero [59]. Some publications [57,58] only describe methods how the interfacial tension could be determined in principle, but make no comment about the resulting sign or magnitude of the tension. There are also works [SH7,55] that report the surface tension of the free interface to be positive in nonequilibrium. The first researcher that derived a positive surface tension was Lee [55]. Very recently Chacón *et al.* [56] used a capillary wave analysis and obtained a small positive tension which remarkably seems to be (almost) independent of the self-propulsion strength. (Bulk coexistence densities depend very strongly on this parameter.) Values for the interfacial tension that were based on the capillary wave theory yield by construction positive results [47]. The positive sign is in accordance with our findings of the tension which are based on our theoretical description developed earlier [46,94]. The key ideas and the comparison with numerical and computer simulation results are given in reference [46] whereas reference [94] gives much background for the analytic calculations that underlie the bulk phase diagram and the density and current distributions across the interface. The theory is based on the continuity equation and the force balance which are both resolved locally in position and orientation. Preferably one would have some direct experimental data for the free interfacial tension as a comparison or for validation. This might help to resolve

the confusion and to solve the problems in the theoretical descriptions. Unfortunately no such results are available at present.

As the topic of the interfacial tension in active systems is so controversial, it is clearly very useful to have at our disposal the two previously (in subsection 3.3.2 and (4.2.2)) determined fundamental exact identities. Let us give a short reminder and summary of these sum rules in the present context (of active Brownian particles undergoing motility-induced phase separation). The first relation, derived in subsection 4.2.2, connects the global interface polarization  $\mathbf{M}^o$  in the system to the bulk values of the current. Due to the quasi one-dimensional geometry of the system and the steady state that we consider, the direction of the global polarization is perpendicular to the interface. This alignment of the active particles in the interfacial region is indeed found in simulations and experiments [44, 46, 137, 138]. Active particles that interact via the purely repulsive WCA-potential point on average towards the denser phase. Its magnitude of the total particle polarization across the interface in steady state is a consequence of the continuity equation and can be expressed by equation (4.13) as  $M^o = \pi L(J^g - J^l)/D_{\text{rot}}$ , where  $J^b$  denotes the strength of the bulk current at the system boundaries,  $L$  is the system length parallel to the interface and  $D_{\text{rot}}$  denotes the rotational diffusion constant. The global polarization is thus determined by bulk values, i.e. the difference between the bulk currents in the dilute and in the dense phase. Hence the total polarization constitutes a state function, which is a direct consequence of the continuity equation and of the free rotational diffusion of the particle orientations [SH6]. The sum rule (4.13) does not only hold for MIPS but for more general systems of active Brownian particles and as laid out in detail in subsection 4.2.3 it was verified numerically [137] and in experiments [138]. The vectorial total polarization  $\mathbf{M}^o$  also leads to a global swim force  $\mathbf{F}_{\text{swim}}^o = \gamma s \mathbf{M}^o$  which pushes the interface against the denser phase.

The second exact relation states that the global internal force vanishes up to trivial boundary terms,  $\mathbf{F}_{\text{int}}^o = 0$ , see equation (3.24) in subsection 3.3.2. This is a conclusion from Noether's theorem resulting from the invariance of the system and of the excess superadiabatic free power functional  $P_t^{\text{exc}}[\rho, \mathbf{J}]$  with respect to a uniform spatial shift at the current time [SH1]. Alternatively the interparticle interaction force can be seen from the viewpoint of Newton's third law, so there is no global contribution besides boundary effects. This boundary term is in case of motility induced phase separation the virial pressure difference (above the ideal gas pressure) between both bulk phases as we showed in subsection 4.3.2.

Both the interface polarization sum rule as well as the internal force sum rule are satisfied within our theory of references [46, 94] which we summarize in appendix A. The combination of both exact identities yields the mechanical pressure balance (4.19) at phase coexistence, see section 4.3. We exploited this balance in the calculations to determine the bulk coexistence densities (see appendix A). This material forms the background and basis for the next step.

## 5.2 Theoretical determination of the free interfacial tension

Forces are undoubtedly prominent and eminently useful quantities to describe many-body systems in statistical mechanics. In contrast to thermodynamic objects such as

the free energy or the grand potential the concept of forces generalizes naturally to nonequilibrium and it has a clear physical interpretation and purpose. This is one of the reasons why our theory for the description of phase separation of active Brownian particles is ultimately based on the force density balance. We use a microscopically sharp one-body picture where all relevant quantities are locally resolved in position and orientation. In the appendix A we give a brief summary of the most important ideas and concepts of the theory [46, 93, 94] (which itself is not part of this thesis). Within this theory the force density balance splits into two equations of which one determines the flow as well as the higher order orientational Fourier components of the density distribution and the second equation is used to identify the coexistence densities and the phase diagram. We identify the former equation as the force balance of flow contributions and the later equation as the structural force balance. The equations are coupled but they can be analyzed separately. This splitting of the force balance in flow and structural contributions is unique, general and it turned out to be an insightful decomposition before [81] in the context of driven (but passive) Brownian dynamics.

For generating the interfacial tension only the structural forces contribute, whereas flow forces do not. Therefore we focus on the structural force balance,

$$\mathbf{f}_{\text{id}}(\mathbf{r}) + \mathbf{f}_{\text{ad}}(\mathbf{r}) + \mathbf{f}_{\text{struc}}(\mathbf{r}) = 0, \quad (5.3)$$

which consists of a sum of the ideal gas contribution, the adiabatic term and the superadiabatic structural force. In general the dependence of the forces would be on both position and orientation. However, the ideal gas contribution is usually small compared to the other forces and we hence approximate the orientation-resolved density by its orientational average. The adiabatic term does not depend on the interparticle orientation by construction. We hence conclude that the intrinsic structural chemical potential  $\mu_{\text{struc}}$  within this approximation cannot depend on the orientation either. We identify this chemical potential as that specific scalar function whose negative gradient gives the corresponding force, i.e.  $\mathbf{f}_{\text{struc}}(\mathbf{r}) = -\nabla\mu_{\text{struc}}(\mathbf{r})$ . This concept is the same that we apply to the ideal gas and adiabatic contributions. Spatial integration over the structural force balance (which is now a sum of gradient terms) hence yields

$$\mu(\mathbf{r}) = \mu_{\text{id}}(\mathbf{r}) + \mu_{\text{ad}}(\mathbf{r}) + \mu_{\text{struc}}(\mathbf{r}) = \text{const}, \quad (5.4)$$

where the total chemical potential  $\mu(\mathbf{r})$  consists of ideal, adiabatic and structural parts and it would be an exact constant if no approximations were made. The ideal gas chemical potential is  $\mu_{\text{id}}(\mathbf{r}) = -k_{\text{B}}T \ln \rho_0(\mathbf{r})$  [76], where  $\rho_0(\mathbf{r})$  denotes the orientational averaged density distribution. The adiabatic contribution is that specific chemical potential that results from the interparticle interactions in the corresponding equilibrium system, i.e. in a system with the same one-body density distribution and the same interparticle interaction potential. (For a detailed description of the adiabatic construction see [79].) One can hence exploit all equilibrium density functional approximations for the adiabatic term and we chose to use the expression from scaled particle theory for simplicity [76],

$$\mu_{\text{ad}}(\mathbf{r}) = k_{\text{B}}T \left[ -\ln(1 - \eta(\mathbf{r})) + \eta(\mathbf{r}) \frac{3 - 2\eta(\mathbf{r})}{(1 - \eta(\mathbf{r}))^2} \right], \quad (5.5)$$

with the rescaled (to take particle softness into account) packing fraction  $\eta(\mathbf{r}) = 0.8\rho_0(\mathbf{r})/\rho_{\text{jam}}$  and the jamming density  $\rho_{\text{jam}}$ , which denotes the density at which the system would come to arrest due to sterical hindrance. Treating the adiabatic repulsion in a local density approximation is a frequent strategy. The ideal and the adiabatic chemical potential only consist of the given local contributions. However, the structural chemical potential contains both a local and a non-local, interfacial term [46, 94],

$$\mu_{\text{struc}}(\mathbf{r}) = \mu_{\text{struc}}^{\text{loc}}(\mathbf{r}) + \mu_{\text{struc}}^{\text{nlloc}}(\mathbf{r}) \quad (5.6)$$

$$= e_1 \frac{\gamma}{2D_{\text{rot}}} v_{\text{f}}^2(\mathbf{r}) \frac{\rho_0(\mathbf{r})}{\rho_{\text{jam}}} + \nabla m \cdot \nabla \rho_0(\mathbf{r}). \quad (5.7)$$

The first, local contribution  $\mu_{\text{struc}}^{\text{loc}}(\mathbf{r})$  originates from an expansion in the forward velocity  $v_{\text{f}}(\mathbf{r})$  and the constant  $e_1$  is the fitting parameter for the amplitude. We know that the local nonequilibrium chemical potential depends on the velocity as it is a kinematic contribution within power functional theory [46, 78, 94]. Structural terms are described by even powers of the velocity [81]. For the nonlocal structural chemical potential  $\mu_{\text{struc}}^{\text{nlloc}}(\mathbf{r})$  we use a square gradient term according to the van der Waals route [150] and for simplicity we set  $m = e_2 \gamma s^2 / 2D_{\text{rot}} \rho_{\text{jam}}^2$  to a constant, which contains the fit parameter  $e_2$ . In general  $m$  can depend on the density and the interfacial contribution will be in reality more complex.

As the assumed interfacial contribution  $\mu_{\text{struc}}^{\text{nlloc}}(\mathbf{r})$  has a gradient structure, one can translate the van der Waals route to the tension [150] to nonequilibrium. Hence the gaseous-liquid interfacial tension  $\gamma_{\text{gl}}$  within this treatment is unique and given by an integral over the interfacial density profile [SH7, 150],

$$\gamma_{\text{gl}} = \int_{-\infty}^{\infty} dx \left[ -W + \frac{1}{2} m (\nabla \rho_0(x))^2 \right]. \quad (5.8)$$

Equation (5.8) makes the square gradient structure resulting from the above nonlocal structural chemical potential contribution, the second term in equation (5.8), more apparent. The local contributions are included in the term  $-W(\rho_0) = (\mu^{\text{loc}} - \mu^{\text{b}})\rho_0 - (p^{\text{loc}} - p^{\text{b}})$ , where the local chemical potential  $\mu^{\text{loc}}$  is given by the total chemical potential  $\mu(\mathbf{r})$  (5.4) without the nonlocal contribution  $\mu_{\text{struc}}^{\text{nlloc}}(\mathbf{r})$ . The corresponding local pressure is determined with an equivalent of the Gibbs-Duhem relation [46, 94] as we show in appendix A. The index b indicates the bulk values of the pressure and of the chemical potential which are numerically identical in the two coexisting phases due to the mechanical and chemical stability (see equations (A.7) and (A.8) in appendix A). The mechanical balance is indeed a consequence of a deep system symmetry as we have shown using Noether's theorem in subsection 4.3.2.

We restrict ourselves to phase separations with quasi one-dimensional geometry, which simplifies  $\mathbf{r}$  to the nontrivial position coordinate  $x$  across the interface. Further we assume  $\rho_0(x)$  to be a hyperbolic tangent interpolating between the coexistence densities, which is a common assumption and in accordance with simulation results, see e.g. reference [44].

Following the mechanical analogy of Rowlinson and Widom [150] one can interpret  $W(\rho_0)$  in equation (5.8) as a (negative) double well energy potential as sketched in

figure 5.1(a). The comparison results from the mathematical similarity of the interfacial equations with a mechanical problem. This problem has the same description as a ball (indicated by the orange disc in figure 5.1(a)) with mass  $m$  that moves in the potential  $W(\rho_0)$ . It starts at time  $x = -\infty$  at the left maximum and rolls to the right maximum of  $W(\rho_0)$  in the limit of  $x = \infty$ . Here  $\rho_0$  indicates the position coordinate. The analogue of the interfacial tension is then the action  $\gamma_{\text{gl}}$ , which is an integral over the Lagrangian consisting of the kinetic energy (nonlocal term) minus potential energy  $W(\rho_0)$  (local terms). Following Noether's theorem in classical mechanics (chapter 2) one can now consider invariances of the action  $\gamma_{\text{gl}}$ . As the mechanical system is homogeneous in time (the integrand in equation (5.8) does not explicitly depend on  $x$ ) Noether's theorem implies the conservation and vanishing of the energy, see equation (5.9) below.

Kerins and Boiteux [151] further generalized the mechanical analogy to the van der Waals theory of nonuniform fluids, in particular the three-phase line contact. There the application of Noether's theorem on the translational invariance of the excess free energy functional determines the force balance around the contact line [151]. The authors were also able to generalize the concept to more general density functionals [152]. Similarly the rotational invariance of the free energy functional yields with Noether's theorem the balance of the global torques around the three-phase contact line [153].

Using the chemical potential balance in the van der Waals theory [SH7, 150] or alternatively expressing the energy balance within the classical mechanics analogy [150] one finds

$$W + \frac{1}{2}m(\nabla\rho_0(\mathbf{r}))^2 = 0. \quad (5.9)$$

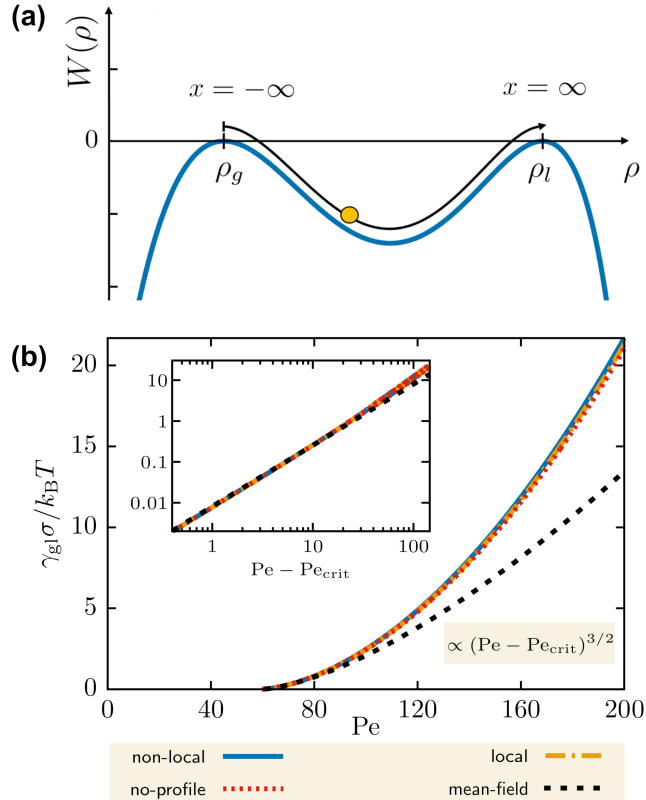
Equation (5.9) states a connection between the local and nonlocal contributions and it can be identified as the first integral of the force balance around the interface [SH7, 150]. Using this equation (5.9) we can re-express the tension  $\gamma_{\text{gl}}$  (5.8) by three different equivalent expressions to calculate [SH7], similar as in equilibrium [150]:

$$\gamma_{\text{gl}} = \int_{-\infty}^{\infty} dx m(\nabla\rho_0(x))^2 \quad (5.10)$$

$$= -2 \int_{-\infty}^{\infty} dx W(\rho_0(x)) \quad (5.11)$$

$$= \int_{\rho^{\text{g}}}^{\rho^{\text{l}}} d\rho_0 \sqrt{-2mW(\rho_0)}. \quad (5.12)$$

All three routes yield the same unique interfacial tension when  $\rho_0(x)$ ,  $m$  and  $W(\rho_0)$  are chosen such that they yield a constant total chemical potential  $\mu(\mathbf{r})$ . We have shown that within our theory the interfacial tensions determined from the three equations (5.10)-(5.12) agree on a very satisfactory level, consider figure 5.1(b). The first nonlocal route given by equation (5.10) does only depend on the interfacial contribution but it is independent of the local function  $W(\rho)$ . This result is indicated by the blue solid line in figure 5.1(b). Equation (5.11) is exactly the opposite in spirit since it only depends on local contributions but it is independent of the nonlocal terms and of the



**Figure 5.1:** (a) Sketch of the classical mechanics analogy of the van der Waals theory. The orange circle indicates a particle with mass  $m$  that moves in the (negative) double well potential energy landscape  $W(\rho_0)$ . It starts at time  $x = -\infty$  at position  $\rho = \rho^g$  and rolls to the second maximum of  $W(\rho)$  at  $\rho^l$  for time  $x = \infty$ . (b) Plot of the gas-liquid interfacial tension  $\gamma_{gl}$  in dependence of the Péclet-number  $Pe$  (main panel) and of the Péclet-number above the critical point  $Pe - Pe_{crit}$  (inset). The three theoretical routes that determine the tension, the non-local method from equation (5.10) (blue line), the local (yellow dash-dotted line, equation (5.11)) and the no-profile method (red dotted line, equation (5.12)), agree with each other. Close to the critical point the tension increases with the mean-field critical exponent  $3/2$  (black dashed line). Panel (b) taken from [SH7], ©American Physical Society (2011). All rights reserved.

constant  $m$ . We therefore refer to it as the local route. In figure 5.1(b) this is indicated by the dash-dotted yellow line. The third relation (5.12) is called the no-profile route (red dashed line in figure 5.1(b)) as it is independent of the density distribution. This property distinguishes this route from the two others and it is useful in practice when  $\rho_0(x)$  across the interface is not known explicitly.

In general our result for the interfacial tension is unique and positive,  $\gamma_{gl} > 0$ , which can be seen for example from equation (5.10) as both the constant  $m$  and the square of the density gradient are positive. We find that the surface tension increases with increasing Péclet-number  $Pe = 3s/\sigma D_{rot}$ . This control parameter relates the active free swim speed  $s$  against the rotational diffusion with the diffusion constant  $D_{rot}$ . Recall that the length scale  $\sigma$  indicates the particle size as it is determined by the Weeks-Chandler-Anderson potential. Close to the critical point the increase of the surface tension is, as expected within the present approximations, similar to the mean



field exponent  $3/2$  (dashed black line in figure 5.1(b)),  $\gamma_{\text{gl}} \propto (\text{Pe} - \text{Pe}_{\text{crit}})^{3/2}$ , where  $\text{Pe}_{\text{crit}}$  is the Péclet-number at the critical point.

Our positive result for the interfacial tension leads us back to the disagreement about this quantity in the literature. It would be beneficial to identify the origin of the differences. For carrying out such work the determined Noether sum rules can be a highly useful resource. The sum rules state nontrivial exact identities and testing whether they are satisfied within a given approach should be computationally feasible. Recall that only structural terms (5.3) contributed in the above calculation of the phase behaviour and the interfacial tension. The flow contributions are important for the general description of coupled motion and dynamics, but apparently not for the spatial structures of the system. This conceptual idea was elaborated for general Brownian many-body systems in nonequilibrium in reference [81]. These authors analyzed the splitting using computer simulations and determined quantitative expressions for the splitting within power functional theory [78, 79]. For these approximations one can expand in a power series of the velocity field [80, 82]. Their results are consistent with our observations and seems to support the utility of separating different contributions, such as structural and flow parts, of the motion.

## 6 Conclusions and outlook

In this thesis we studied the consequences of symmetries in many-body problems on the basis of Noether's theorem. Hereby we exploited the generality of Noether's theorem and investigated the invariances of thermodynamic functionals such as the grand potential and the free energy rather than symmetries of the action functional or the corresponding Lagrangian. Our analytic calculations give exact information about forces, torques, their variances as well as the polarization. We explicitly apply these sum rules to active and thermal Brownian particles, e.g. in a gravitational field or under phase separation. The considerations yield new insights into the importance of symmetries in the treatment of both equilibrium and nonequilibrium many-body systems. We are able to systematically derive or re-derive sum rules such as the relation between the external force on an external wall and the mass of all particles for thermal and active sedimentation and the connection of the global polarization in active particle system with the boundary currents.

Noether's theorem is usually applied to invariances of the action functional or the Lagrangian to determine the corresponding conserved quantities. In order to apply the theorem to statistical mechanics we considered invariances of classical and quantum statistical mechanical functionals [SH1, SH2, SH3, SH4, SH5]. This method is demonstrated for the translational invariance of the grand potential from which one obtains the vanishing of the global external force in equilibrium systems. This and the other determined sum rules are consequences of symmetries and they hold in the grand canonical ensemble [SH1]. The vanishing of both the global external and global internal force is also valid in the canonical ensemble [SH2]. Some of the obtained relations, such as the LMBW equation and its inverse, are well-known and have proven

useful [70], while others, such as the vanishing of the global superadiabatic force or the dynamical sum rules including memory, constitute (to the best of our knowledge) new results. From linear order terms of the expansion in the symmetry parameters we deduce the vanishing of global force and torque terms such as the ideal, the external, the internal and the superadiabatic contributions. Contributions quadratic in the transformation parameter relate the variance of forces and torques to the curvature of their corresponding potentials [SH5].

Locally resolved relations, such as the LMBW equation, can be obtained by functional differentiation of global sum rules, such as the vanishing of the global external force. Another possibility to obtain local sum rules is to consider invariant canonical transformations on phase space. An example of this strategy is the derivation of the force density balance in many-body quantum mechanics [SH4] and the derivation of the first equation of the YBG hierarchy in classical statistical mechanics [SH3]. Both are consequences of an invariance against a spatial dependent distortion of positions and the corresponding changes in momenta. The latter forms the starting point of a novel implementation of density functional theory (DFT) which works on the basis of forces instead of potentials and is therefore referred to as force-DFT [SH3]. We refer the reader to the corresponding paper [SH3] for analytic proofs of the hard wall contact theorem within the realizations of the potential-DFT and the force-DFT. This theorem expresses the equality of the bulk pressure with the particle density at a hard wall up to a factor  $k_{\text{B}}T$ . For the potential-DFT the wall contact density is given by the bulk pressure determined from the compressibility route. In contrast the contact density determined with the force-DFT equals the virial pressure as we show in reference [SH3], which deviates from the compressibility pressure where using an approximative free energy functional.

We have demonstrated that the interfacial polarization sum rule ensues from exploiting the orientation-resolved continuity equation [SH6]. The sum rule relates the global polarization of particle orientations at the interface of a system to the boundary values of the swim current. In case of a confined system with vanishing net-flux through its boundaries the sum rule clearly states that no global polarization persists in the long time limit. To make further progress we inserted the polarization relation together with the Noether sum rule of vanishing global internal force in the force density balance, and then integrated in space and orientation [SH1,SH6]. In case of active sedimentation this procedure determines the center of mass velocity. In the steady state limit this yields the equivalence of the gravitational force with the force exerted on the wall. For motility-induced phase separation it confirms the validity of the pressure balance at phase coexistence.

We have calculated from the pressure balance the interfacial tension for the planar free interface between coexisting bulk phases in motility-induced phase separation of two-dimensional active Brownian particles [SH7]. The treatment works on the basis of a splitting of the forces balance equation into two different categories of flow forces and structural forces. Within this theory we have shown that the three ways of the van der Waals route to determine the interfacial tension, which are consistent with each other in equilibrium, also agree in nonequilibrium. The value of the tension itself turned out to be positive and be of order of unity in natural units. Both findings, the sign and the magnitude, oppose most results from the literature which rather suggest a very large negative or approximately zero value of the surface tension. As

our detailed microscopical description satisfies exact Noether identities and determines a physically reasonable tension the present work casts doubts on the validity and the interpretation of the findings in the literature.

On the basis of this thesis a number of direct questions and corresponding possible future projects ensue of which we describe some briefly below. The symmetry operations we have used so far for the statistical mechanics applications Noether's theorem have only been spatial (namely translations and rotations). It might be worthwhile to investigate the consequences of a temporal symmetry operation [139], such as a time shift. The comparison with Noether's theorem in classical mechanics suggests that this consideration would determine the statistical mechanic analogue to the conservation of energy which would be certainly a useful sum rule.

The local canonical symmetry transformations could also be extended to beyond linear order terms, with the quadratic order contributions being especially interesting. The previously determined force variance relations from second order contributions correspond to the vanishing of the global force contributions in linear order [SH1,SH5]. In analogy we expect a relation for the locally resolved variance of the total force from Noether's theorem for local canonical transformations; we recall that the linear contribution yields the locally resolved force balance [SH3]. This sum rule might be beneficial especially when it can be further generalized to nonequilibrium, where this and related variances appear, such as the van Hove function and van Hove current as recently investigated in a hard sphere fluid [86,87] or the stress-stress auto-correlator in glass-forming fluids [154,155]. In principle it should also be possible to extend our considerations of the Noether theorem in statistical mechanics from the overdamped Brownian dynamics to molecular dynamics by considering the invariances of the power rate functional  $G_t$  [79]. Additionally to the density and the current this functional also depends on the temporal change of the current  $\dot{\mathbf{J}}(\mathbf{r}, t)$ , i.e. the acceleration density. It is interesting to further explore Noether sum rules in the context of quantum statistical mechanics in particular on the basis of quantum density functional theory and quantum power functional theory.

Next to these generalizations of Noether's theorem an investigation of its relation to Nambu-Goldstone modes might be fruitful. A deep connection would be natural as Nambu-Goldstone modes result from spontaneous breaking of a continuous symmetry and the emerge of corresponding excitations. Nambu-Goldstone modes have been recently studied at fluid interfaces [156,157] and in colloidal dispersions with a glass transition [158].

Using Noether's theorem we have shown that the mechanical stability, expressed by the pressure balance, still holds in nonequilibrium for the simple model system of phase-separated active Brownian particles. An obvious question that arises is whether the chemical stability at phase coexistence can be proven similarly to give an exact chemical potential balance. Further investigations may show the validity of these balances for different types of particles and further systems in nonequilibrium. Both the mechanical and the chemical stability balance can contribute to the development of a general nonequilibrium theory for phase separation and interfacial phenomena beyond the special case of active Brownian particles.

Actually active Brownian particle as considered here are not the only simple model system to mimic the behavior of active matter [21–27]. Other popular particle-based models include run-and-tumble particles (RTPs) [38,39] and active Ornstein-Uhlenbeck

particles (AOUPs) [40, 41]. These particles do not undergo free rotational diffusion but rather have a fixed orientation which they change randomly with a given tumble rate (RTPs) or change their self-propulsion velocity according to an Ornstein-Uhlenbeck process (AOUPs). Nevertheless these different particle types behave similarly to ABPs and they all tend to phase separate induced by their motility. Considering RTPs and AOUPs within our theory for phase separations [46, 94] and also investigating their polarization behaviour [SH6] would allow to identify and quantify differences and similarities between the different models. Such work could supplement previous comparisons of active Brownian particles and RTPs [38, 39]. It would be interesting to generalize the considered theory [46, 94] of phase coexistence for ABPs to RTPs and AOUPs. Of course it is a priori not clear whether the chosen approximations, especially for the structural chemical potential, are still valid for these different types of particles. Therefore it would be a test and if successful an indication of the strength of the description if the resulting phase diagrams and results for the tension are also in good agreement with simulations.

Additionally to the investigation of active Brownian particles and their phase coexistence in two dimensions one could also consider the analog in three dimensions. Active Brownian particles also phase separate in three dimensions but show marked quantitative differences in their phase behaviour in dependence of the dimensionality [159]. From a technical point of view one has to change the Fourier series expansion in the flow force balance equation to an expansion in spherical harmonics. This change has effects on the entire calculation but should leave the concepts and the main steps of the derivation unchanged (up to some dimension dependent constant prefactors).

An interesting effect that occurs in three dimensions is the formation of an ordered dense phase, i.e. an active crystal. This phase has been observed in simulations by Omar *et al.* [160] and by Turci and Wilding [161]. Both works report the fluid-fluid motility-induced phase separation but also realize that this transition is only metastable with respect to the gaseous-crystal phase separation for a large range of Péclet-numbers. To be able to describe this effect one could adapt the adiabatic chemical potential used in the theory to an expression that is valid for crystals (like from cell theory). If it were possible to predict the crystallization phenomenon one could also observe whether spontaneous active crystallization occurs in two dimensions as well albeit one faces difficulties due to hexatic ordering. Such a resulting phase diagram could be compared with simulation data of references [160, 161]. Additionally one could look for connections to the phase separation of polymer-colloid mixtures in the theory. Turci and Wilding [161] have identified similarities to such mixtures in the topology of their phase diagrams.

The isothermal compressibility of systems of active Brownian particles within our description is also a worthwhile quantity to investigate. In analogy to equilibrium there are different theoretical methods to determine the compressibility. These were studied and compared to each other for active Brownian particles by Dulaney *et al.* [162]. Turci and Wilding [161] established that the isothermal compressibility is a useful quantity to characterize phase behaviour as it measures the proximity to the nonequilibrium critical point at which it diverges as it does in equilibrium. These authors investigated the compressibility from particle fluctuations in subareas of their simulations. Investigations of this (and other) fluctuations have shown that those can form the basis for a description of inhomogeneous fluids in equilibrium [60] and

are relevant for the observation of hydrophobicity and drying as they can be better indicators for transitions than the bare density profile [163–165].

To further test the validity of the chosen approximation to the structural contribution to the chemical potential one could examine phase separation of active particles with interparticle attraction, e.g. particles interacting with a Lennard-Jones potential. Clearly the adiabatic chemical potential has to include an additional contribution which models the attraction [76]. From simulations one knows that the activity weakens the phase separation when the system is already passively phase-separated [166–168]. Only with increasing swim speed the motility-induced phase separation enters, see reference [166] for a phase diagram consisting of simulation snapshots. Depending on the overall density and strength of particle attractions the transition between both phase separation mechanisms can be with or without a homogeneous phase in between.

Finally it might be interesting to try to include additional external potentials in the description of phase coexistence [46, 94], which then also split into their flow and structural contribution as in [81]. This would extend the theory to various new applications as wetting phenomena [18, 19] or active sedimentation [109, 110].

The suggested generalizations on the description of MIPS and ABPs such as active crystallization, adding of interparticle attraction and external potentials are also interesting to be investigated in the context of Noether’s theorem.



# Bibliography

- [SH1] S. Hermann and M. Schmidt, *Noether's theorem in statistical mechanics*, Commun. Phys. **4**, 176 (2021).
- [SH2] S. Hermann and M. Schmidt, *Why Noether's theorem applies to statistical mechanics*, J. Phys.: Condens. Matter **34**, 213001 (2022).
- [SH3] S. Tschopp, F. Sammüller, S. Hermann, M. Schmidt, and J. M. Brader, *Force density functional theory in- and out-of-equilibrium*, Phys. Rev. E **106**, 014115 (2022).
- [SH4] S. Hermann and Schmidt, *Force balance in thermal quantum many-body systems from Noether's theorem*, arXiv: 2208.00473 (submitted to J. Phys. A Math. Theor.).
- [SH5] S. Hermann and M. Schmidt, *Variance of fluctuations from Noether invariance*, arXiv: 2203.15654 (submitted to Commun. Phys.).
- [SH6] S. Hermann and M. Schmidt, *Active interface polarization as a state function*, Phys. Rev. Research **2**, 022003(R) (2020).
- [SH7] S. Hermann, D. de las Heras, and M. Schmidt, *Non-negative interfacial tension in phase-separated active Brownian particles*, Phys. Rev. Lett. **123**, 268002 (2019).
- [8] E. Noether, *Invariante Variationsprobleme*, Nachr. d. König. Gesellsch. d. Wiss. zu Göttingen, Math.-Phys. Klasse **235** (1918). English translation by M. A. Tavel, *Invariant variation problems*, Transp. Theo. Stat. Phys. **1**, 186 (1971); for a version in modern typesetting see: F. Y. Wang, arXiv:physics/0503066v3 (2018).
- [9] H. Goldstein, C. Poole, and J. Safko, *Classical mechanics*, 3rd ed. (Addison-Wesley, Boston, 2001).
- [10] Y. Kosmann-Schwarzbach, *The Noether theorems - Invariance and conservation laws in the twentieth century* (Springer, New York, 2018).
- [11] D. E. Neuenschwander, *Emmy Noether's wonderful theorem* (Johns Hopkins University Press, Baltimore, 2011).
- [12] A. G. Lezcano and A. C. M. de Oca, *A stochastic version of the Noether theorem*, Found. Phys. **48**, 726 (2018).
- [13] I. Marvian and R. W. Spekkens, *Extending Noether's theorem by quantifying the asymmetry of quantum states*, Nat. Commun. **5**, 3821 (2014).
- [14] J. C. Baez and B. Fong, *A Noether theorem for Markov processes*, J. Math. Phys. **54**, 013301 (2013).
- [15] S. Sasa and Y. Yokokura, *Thermodynamic entropy as a Noether invariant*, Phys. Rev. Lett. **116**, 140601 (2016).
- [16] S. Sasa, S. Sugiura, and Y. Yokokura, *Thermodynamical path integral and emergent symmetry*, Phys. Rev. E **99**, 022109 (2019).

- [17] Y. Minami and S. Sasa, *Thermodynamic entropy as a Noether invariant in a Langevin equation*, J. Stat. Mech. **2020**, 013213 (2020).
- [18] R. Evans, M. C. Stewart, and N. B. Wilding, *Drying and wetting transitions of a Lennard-Jones fluid: Simulations and density functional theory*, J. Chem. Phys. **147**, 044701 (2017).
- [19] M. K. Coe, R. Evans, and N. B. Wilding, *Density depletion and enhanced fluctuations in water near hydrophobic solutes: Identifying the underlying physics*, Phys. Rev. Lett. **128**, 045501 (2022).
- [20] E. Tjhung, C. Nardini, and M. E. Cates, *Cluster phases and bubbly phase separation in active fluids: Reversal of the Ostwald process*, Phys. Rev. X **8**, 031080 (2018).
- [21] P. Romanczuk, M. Bär, W. Ebeling, B. Lindner, and L. Schimansky-Geier, *Active Brownian Particles - From Individual to Collective Stochastic Dynamics*, Eur. Phys. J. Special Topics **202**, 1 (2012).
- [22] C. Bechinger, R. Di Leonardo, H. Löwen, C. Reichhardt, G. Volpe, and G. Volpe, *Active particles in complex and crowded environments*, Rev. Mod. Phys. **88**, 045006 (2016).
- [23] S. H. L. Klapp, *Collective dynamics of dipolar and multipolar colloids: From passive to active systems*, Curr. Opin. Colloid Interface Sci. **21**, 76 (2016).
- [24] A. Zöttl and H. Stark, *Emergent behavior in active colloids*, J. Phys.: Condens. Matter **28**, 253001 (2016).
- [25] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha, *Hydrodynamics of soft active matter*, Rev. Mod. Phys. **85**, 1143 (2013).
- [26] É. Fodor and M. C. Marchetti, *The statistical physics of active matter: From self-catalytic colloids to living cells*, Physica A **504**, 106 (2018).
- [27] G. Gompper, R. G. Winkler, T. Speck, A. Solon, C. Nardini, F. Peruani, H. Löwen, R. Golestanian, U. B. Kaupp, L. Alvarez, T. Kiørboe, E. Lauga, W. C. K. Poon, A. DeSimone, S. Muiños-Landin, A. Fischer, N. A. Söker, F. Cichos, R. Kapral, P. Gaspard, M. Ripoll, F. Sagues, A. Doostmohammadi, J. M. Yeomans, I. S. Aranson, C. Bechinger, H. Stark, C. K. Hemelrijk, F. J. Nedelec, T. Sarkar, T. Aryaksama, M. Lacroix, G. Duclos, V. Yashunsky, P. Silberzan, M. Arroyo, and S. Kale, *The 2020 motile active matter roadmap*, J. Phys.: Condens. Matter **32**, 193001 (2020).
- [28] J. R. Howse, R. A. L. Jones, A. J. Ryan, T. Gough, R. Vafabakhsh, and R. Golestanian, *Self-motile colloidal particles: From directed propulsion to random walk*, Phys. Rev. Lett. **99**, 048102 (2007).
- [29] J. Palacci, C. Cottin-Bizonne, C. Ybert, and L. Bocquet, *Sedimentation and effective temperature of active colloidal suspensions*, Phys. Rev. Lett. **105**, 088304 (2010).
- [30] M. Selmke, U. Khadka, A. P. Bregulla, F. Cichos, and H. Yang, *Theory for controlling individual self-propelled micro-swimmers by photon nudging I: directed transport*, Phys. Chem. Chem. Phys. **20**, 10502 (2018).



- [31] M. Selmke, U. Khadka, A. P. Bregulla, F. Cichos, and H. Yang, *Theory for controlling individual self-propelled micro-swimmers by photon nudging II: confinement*, Phys. Chem. Chem. Phys. **20**, 10521 (2018).
- [32] B. Qian, D. Montiel, A. Bregulla, F. Cichos, and H. Yang, *Harnessing thermal fluctuations for purposeful activities: the manipulation of single micro-swimmers by adaptive photon nudging*, Chem. Sci. **4**, 1420 (2013).
- [33] A. P. Bregulla, H. Yang, and F. Cichos, *Stochastic localization of microswimmers by photon nudging*, ACS Nano **8**, 6542 (2014).
- [34] M. J. Schnitzer, *Theory of continuum random walks and application to chemotaxis*, Phys. Rev. E **48**, 2553 (1993).
- [35] J. Schwarz-Linek, J. Arlt, A. Jepsen, A. Dawson, T. Vissers, D. Miroli, T. Pilizota, V. A. Martinez, and W. C. K. Poon, *Escherichia coli as a model active colloid: A practical introduction*, Colloids Surf. B **137**, 2 (2016).
- [36] A. Cavagna and I. Giardina, *Bird flocks as condensed matter*, Annu. Rev. Condens. Matter Phys. **5**, 183 (2014).
- [37] D. S. Cambu and A. Rosas, *Density induced transition in a school of fish*, Physica A **391**, 3908 (2012).
- [38] M. E. Cates and J. Tailleur, *When are active Brownian particles and run-and-tumble particles equivalent? Consequences for motility-induced phase separation*, EPL **101**, 20010 (2013).
- [39] A. P. Solon, J. Stenhammar, R. Wittkowski, M. Kardar, Y. Kafri, M. E. Cates, and J. Tailleur, *Pressure and phase equilibria in interacting active Brownian spheres*, Phys. Rev. Lett. **114**, 198301 (2015).
- [40] L. L. Bonilla, *Active Ornstein-Uhlenbeck particles*, Phys. Rev. E **100**, 022601 (2019).
- [41] D. Martin, J. O’Byrne, M. E. Cates, . Fodor, C. Nardini, J. Tailleur, and F. van Wijland, *Statistical mechanics of active Ornstein-Uhlenbeck particles*, Phys. Rev. E **103**, 032607 (2021).
- [42] R. Wittkowski, A. Tiribocchi, J. Stenhammar, R. J. Allen, D. Marenduzzo, and M. E. Cates, *Scalar  $\varphi^4$  field theory for active-particle phase separation*, Nat. Commun. **5**, 4351 (2014).
- [43] C. F. Lee, *An infinite set of integral formulae for polar, nematic, and higher order structures at the interface of motility-induced phase separation*, New J. Phys. **24**, 043010 (2022).
- [44] S. Paliwal, J. Rodenburg, R. van Roij, and M. Dijkstra, *Chemical potential in active systems: predicting phase equilibrium from bulk equations of state?*, New J. Phys. **20**, 015003 (2018).
- [45] A. P. Solon, J. Stenhammar, M. E. Cates, Y. Kafri, and J. Tailleur, *Generalized thermodynamics of motility-induced phase separation: phase equilibria, Laplace pressure, and change of ensembles*, New J. Phys. **20**, 075001 (2018).
- [46] S. Hermann, P. Krinninger, D. de las Heras, and M. Schmidt, *Phase coexistence of active Brownian particles*, Phys. Rev. E **100**, 052604 (2019).
- [47] J. Bialke, H. Lowen, and T. Speck, *Negative interfacial tension in phase-separated active Brownian particles*, Phys. Rev. Lett. **115**, 098301 (2015).

- [48] A. P. Solon, J. Stenhammar, M. E. Cates, Y. Kafri, and J. Tailleur, *Generalized thermodynamics of phase equilibria in scalar active matter*, Phys. Rev. E **97**, 020602(R) (2018).
- [49] U. M. B. Marconi and C. Maggi, *Towards a statistical mechanical theory of active fluids*, Soft Matter **11**, 8768 (2015).
- [50] A. Patch, D. M. Sussman, D. Yllanes, and M. C. Marchetti, *Curvature-dependent tension and tangential flows at the interface of motility-induced phases*, Soft Matter **14**, 7435 (2018).
- [51] T. Speck, *Collective forces in scalar active matter*, Soft Matter **16**, 2652 (2020).
- [52] T. Speck, *Coexistence of active Brownian disks: van der Waals theory and analytical results*, Phys. Rev. E **103**, 012607 (2021).
- [53] A. K. Omar, Z.-G. Wang, and J. F. Brady, *Microscopic origins of the swim pressure and the anomalous surface tension of active matter*, Phys. Rev. E **101**, 012604 (2020).
- [54] N. Lauersdorf, T. Kolb, M. Moradi, E. Nazockdast, and D. Klotsa, *Phase behavior and surface tension of soft active Brownian particles*, Soft Matter **17**, 6337 (2021).
- [55] C. F. Lee, *Interface stability, interface fluctuations, and the Gibbs–Thomson relationship in motility-induced phase separations*, Soft Matter **13**, 376 (2017).
- [56] E. Chacón, F. Alarcón, J. Ramírez, P. Tarazona, and C. Valeriani, *Intrinsic structure perspective for MIPS interfaces in two-dimensional systems of active Brownian particles*, Soft Matter **18**, 2646 (2022).
- [57] U. M. B. Marconi, C. Maggi, and S. Melchionna, *Pressure and surface tension of an active simple liquid: a comparison between kinetic, mechanical and free-energy based approaches*, Soft Matter **12**, 5727 (2016).
- [58] S. Das, G. Gompper, R. G. Winkler, *Local stress and pressure in an inhomogeneous system of spherical active Brownian particles*, Sci. Rep. **9**, 6608 (2019).
- [59] J. Rodenburg, *Thermodynamic Variables for Active Brownian Particles : Pressure, Surface Tension, and Chemical Potential*, PhD thesis, Utrecht (2020).
- [60] T. Eckert, N. C. X. Stuhlmüller, F. Sammüller, and M. Schmidt, *Fluctuation profiles in inhomogeneous fluids*, Phys. Rev. Lett. **125**, 268004 (2020).
- [61] R. A. Lovett, C. Y. Mou, and F. P. Buff, *The structure of the liquid–vapor interface*, J. Chem. Phys. **65**, 570 (1976).
- [62] M. S. Wertheim, *Correlations in the liquid–vapor interface*, J. Chem. Phys. **65**, 2377 (1976).
- [63] J. M. Brader, *Structural precursor to freezing: An integral equation study*, J. Chem. Phys. **128**, 104503 (2008).
- [64] P. Bryk, D. Henderson, and S. Sokolowski, *A fluid in contact with a semipermeable surface: Second-order integral equation approach*, J. Chem. Phys. **107**, 3333 (1997).
- [65] M. Kasch and F. Forstmann, *An orientational instability and the liquid–vapor interface of a dipolar hard sphere fluid*, J. Chem. Phys. **99**, 3037 (1993).
- [66] S. Iatsevitch and F. Forstmann, *Density profiles at liquid–vapor and liquid–liquid interfaces: An integral equation study*, J. Chem. Phys. **107**, 6825 (1997).

- [67] X. Xu and S. A. Rice *A density functional theory of one- and two-layer freezing in a confined colloid system*, Proc. R. Soc. A **464**, 65 (2008).
- [68] C. Walz and M. Fuchs, *Displacement field and elastic constants in nonideal crystals*, Phys. Rev. B **81**, 134110 (2010).
- [69] J. M. Häring, C. Walz, G. Szamel, and M. Fuchs, *Coarse-grained density and compressibility of nonideal crystals: General theory and an application to cluster crystals*, Phys. Rev. B **92**, 184103 (2015).
- [70] M. Baus, *Broken symmetry and invariance properties of classical fluids*, Mol. Phys. **51**, 211 (1984).
- [71] N. D. Mermin, *Thermal properties of the inhomogeneous electron gas*, Phys. Rev. **137**, A1441 (1965).
- [72] R. Evans, *The nature of the liquid-vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids*, Adv. Phys. **28**, 143 (1979).
- [73] R. Roth, *Fundamental measure theory for hard-sphere mixtures: a review*, J. Phys.: Condens. Matter **22**, 063102 (2010).
- [74] R. L. Davidchack, B. B. Laird, and R. Roth, *Hard spheres at a planar hard wall: simulations and density functional theory*, Condens. Matter Phys. **19**, 23001 (2016).
- [75] R. Evans, *Density functionals in the theory nonuniform fluids*, in “Fundamentals of Inhomogeneous Fluids”, ed. by D. Henderson (Dekker, New York, 1992).
- [76] J.-P. Hansen and I. R. McDonald, *Theory of simple liquids*, 4th ed. (Academic Press, London, 2013).
- [77] R. Evans, M. Oettel, R. Roth, and G. Kahl, *New developments in classical density functional theory*, J. Phys.: Condens. Matter **28**, 240401 (2016).
- [78] M. Schmidt and J. M. Brader, *Power functional theory for Brownian dynamics*, J. Chem. Phys. **138**, 214101 (2013).
- [79] M. Schmidt, *Power functional theory for many-body dynamics*, Rev. Mod. Phys. **94**, 015007 (2022).
- [80] D. de las Heras and M. Schmidt, *Velocity gradient power functional for Brownian dynamics*, Phys. Rev. Lett. **120**, 028001 (2018).
- [81] D. de las Heras and M. Schmidt, *Flow and structure in nonequilibrium Brownian many-body systems*, Phys. Rev. Lett. **125**, 018001 (2020).
- [82] N. C. X. Stuhlmüller, T. Eckert, D. de las Heras, and M. Schmidt, *Structural nonequilibrium forces in driven colloidal systems*, Phys. Rev. Lett. **121**, 098002 (2018).
- [83] J. Renner, M. Schmidt, and D. de las Heras, *Shear and bulk acceleration viscosities in simple fluids*, Phys. Rev. Lett. **128**, 094502 (2022).
- [84] N. Jahreis and M. Schmidt, *Shear-induced deconfinement of hard disks*, Col. Pol. Sci. **298**, 895 (2020).
- [85] L. L. Treffenstädt and M. Schmidt, *Memory-induced motion reversal in Brownian liquids*, Soft Matter **16**, 1518 (2020).
- [86] L. L. Treffenstädt and M. Schmidt, *Universality in driven and equilibrium hard sphere liquid dynamics*, Phys. Rev. Lett. **126**, 058002 (2021).

- [87] L. L. Treffenstädt, T. Schindler, and M. Schmidt, *Dynamic decay and superadiabatic forces in the van Hove dynamics of bulk hard sphere fluids*, SciPost Phys. **12**, 133 (2022).
- [88] T. Geigenfeind, D. de las Heras, and M. Schmidt, *Superadiabatic demixing in nonequilibrium colloids*, Commun. Phys. **3**, 23 (2020).
- [89] P. Krinninger, M. Schmidt, and J. M. Brader, *Nonequilibrium phase behaviour from minimization of free power dissipation*, Phys. Rev. Lett. **117**, 208003 (2016).
- [90] P. Krinninger and M. Schmidt, *Power functional theory for active Brownian particles: general formulation and power sum rules*, J. Chem. Phys. **150**, 074112 (2019).
- [91] M. Schmidt, *Power functional theory for Newtonian many-body dynamics*, J. Chem. Phys. **148**, 044502 (2018).
- [92] M. Schmidt, *Quantum power functional theory for many-body dynamics*, J. Chem. Phys. **143**, 174108 (2015).
- [93] S. Hermann, *Superadiabatic forces and phase separation of active Brownian particles*, Master thesis, Bayreuth (2019).
- [94] S. Hermann, D. de las Heras, and M. Schmidt, *Phase separation of active Brownian particles in two dimensions: Anything for a quiet life*, Mol. Phys. **119**, e1902585 (2021).
- [95] A similar concept to the cyclic coordinates also holds in Hamiltonian Mechanics. If the Hamiltonian is independent of the generalized coordinate  $\mathbf{q}_i$  one can conclude using the equation of motion,  $\partial H/\partial \mathbf{q}_i = \dot{\mathbf{p}}_i = 0$ , such that the corresponding momentum  $\mathbf{p}_i$  is conserved.
- [96] C. H. Kimberling, *Emmy Noether, greatest woman mathematician*, Math. Teach. **75**, 246 (1982).
- [97] C. Tollmien, *“Sind wir doch der Meinung, daß ein weiblicher Kopf nur ganz ausnahmsweise in der Mathematik schöpferisch tätig sein kann...” - eine Biographie der Mathematikerin Emmy Noether (1882 - 1935) und zugleich ein Beitrag zur Geschichte der Habilitation von Frauen an der Universität Göttingen*, Göttinger Jahrbuch **38**, 153 (1990).
- [98] Image from Natascha Artin (1932). Source: Archive of P. Roquette (Heidelberg) and C. Kimberling (Evansville), Oberwolfach Photo Collection, [https://opc.mfo.de/detail?photo\\_id=9265](https://opc.mfo.de/detail?photo_id=9265). Copyright: Creative Commons License Attribution-Share Alike 2.0 Germany, <https://creativecommons.org/licenses/by-sa/2.0/de/deed.en>
- [99] N. Byers, *E. Noether’s discovery of the deep connection between symmetries and conservation laws*, arXiv:physics/9807044 (1998).
- [100] D. H. Kobe, *Derivation of Maxwell’s equations from the gauge invariance of classical mechanics*, Am. J. Phys. **48**, 348 (1980).
- [101] N. D. Mermin, *Thermal properties of the inhomogeneous electron gas*, Phys. Rev. **137**, A1441 (1965).
- [102] Complementary to the standard potential-DFT (see section 3.1) an alternative implementation of DFT has been recently developed [SH3]. We only give a very rough sketch of the main ideas of the force-DFT here. The theory directly defines

the approximate one-body direct correlation function as a consequence of the internal forces and not as the derivative of an approximated  $F_{\text{exc}}[\rho]$  functional. In case of a pair interaction potential  $\phi(r)$ , where  $r$  denotes the particle distance, the direct correlation function  $c_1(\mathbf{r})$  is given by  $\rho(\mathbf{r})\nabla c_1(\mathbf{r}) = -\int d\mathbf{r}'\rho_2(\mathbf{r},\mathbf{r}')\nabla\beta\phi(r)$ , where  $\rho_2(\mathbf{r},\mathbf{r}')$  denotes the two-body density and  $\beta = 1/k_{\text{B}}T$  is the inverse temperature. The difficulty of the determination of the two body density is overcome by solving the inhomogeneous Ornstein-Zernike equation [136]. If an exact expression for the free energy functional were known, the above definitions of  $c_1(\mathbf{r})$  agrees with the standard expression  $c_1(\mathbf{r}) = -\delta\beta F_{\text{exc}}[\rho]/\delta\rho(\mathbf{r})$ . See reference [SH3] for details.

- [103] D. G. Triezenberg and R. Zwanzig, *Fluctuation theory of surface tension*, Phys. Rev. Lett. **28**, 1183 (1972).
- [104] J. Yvon, *Actualités Scientifiques et Industrielles*, (Hermann & Cie., 1935).
- [105] M. Born and H. S. Green, *A general kinetic theory of liquids I. The molecular distribution functions*, Proc. R. Soc. London, Ser. A **188**, 10 (1946).
- [106] C. F. Tejero and M. Baus, *Viscoelastic surface waves and the surface structure of liquids*, Mol. Phys. **54**, 1307 (1985).
- [107] C. H. Kimberling, *Emmy Noether and her influence*, in *Emmy Noether: A Tribute to Her Life and Work*, ed. by J. W. Brewer and M. K. Smith, (Marcel Dekker Inc., New York, 1981).
- [108] It is possible to derive the vanishing of the global external force (3.10) without the usage of a functional Taylor expansion. Therefore one starts as before by considering the effects of a global external potential shift,  $V_{\text{ext}}(\mathbf{r}) \rightarrow V_{\text{ext}}(\mathbf{r} + \boldsymbol{\epsilon})$ . This influences the grand potential in general and we Taylor expand the shifted grand potential  $\Omega^\epsilon$  with respect to the displacement  $\boldsymbol{\epsilon}$  up to linear order,  $\Omega^\epsilon = \Omega + \partial\Omega^\epsilon/\partial\boldsymbol{\epsilon}|_{\boldsymbol{\epsilon}=0} \cdot \boldsymbol{\epsilon}$ . Exploiting that the grand potential is invariant under an arbitrary spatial displacement yields that the derivative of  $\Omega$  with respect to  $\boldsymbol{\epsilon}$  has to vanish. The ideal and the excess free energy do not depend on the external potential and hence not on  $\boldsymbol{\epsilon}$  as they are both intrinsic contributions. This can be seen from the internal-external splitting as shown in reference [79]. So using equation (3.3) we obtain

$$0 = \frac{\partial\Omega}{\partial\boldsymbol{\epsilon}} = \int d\mathbf{r} \frac{\partial}{\partial\boldsymbol{\epsilon}} (V_{\text{ext}}(\mathbf{r} + \boldsymbol{\epsilon}) - \mu)\rho(\mathbf{r}) = \int d\mathbf{r} \rho(\mathbf{r})\nabla V_{\text{ext}}(\mathbf{r} + \boldsymbol{\epsilon}),$$

where we have used that the chemical potential is a constant in the last step. Inserting the limit  $\boldsymbol{\epsilon} \rightarrow 0$  yields the cancellation of the external force (3.10).

- [109] J. Vachier and M. G. Mazza, *Dynamics of sedimenting active Brownian particles*, Eur. Phys. J. E **42**, 11 (2019).
- [110] S. Hermann and M. Schmidt, *Active ideal sedimentation: Exact two-dimensional steady states*, Soft Matter **14**, 1614 (2018).
- [111] M. Enculescu and H. Stark, *Active colloidal suspensions exhibit polar order under gravity*, Phys. Rev. Lett. **107**, 058301 (2011).
- [112] A. P. Solon, M. E. Cates, and J. Tailleur, *Active Brownian particles and run-and-tumble particles: A comparative study*, Eur. Phys. J. Spec. Top. **224**, 1231 (2015).

- [113] P. Tarazona and R. Evans, *On the validity of certain integro-differential equations for the density-orientation profile of molecular fluid interfaces*, Chem. Phys. Lett. **97**, 279 (1983).
- [114] K. E. Gubbins, *Structure of nonuniform molecular fluids – integrodifferential equations for the density-orientation profile*, Chem. Phys. Lett. **76**, 329 (1980).
- [115] J. M. Brader and M. Schmidt, *Nonequilibrium Ornstein-Zernike relation for Brownian many-body dynamics*, J. Chem. Phys. **139**, 104108 (2013).
- [116] J. M. Brader and M. Schmidt, *Dynamic correlations in Brownian many-body systems*, J. Chem. Phys. **140**, 034104 (2014).
- [117] D. Lesnicki, R. Vuilleumier, A. Carof, and B. Rotenberg, *Molecular hydrodynamics from memory kernels*, Phys. Rev. Lett. **116**, 147804 (2016).
- [118] D. Lesnicki and R. Vuilleumier, *Microscopic flow around a diffusing particle*, J. Chem. Phys. **147**, 094502 (2017).
- [119] G. Jung and F. Schmid, *Computing bulk and shear viscosities from simulations of fluids with dissipative and stochastic interactions*, J. Chem. Phys. **144**, 204104 (2016).
- [120] G. Jung, M. Hanke, and F. Schmid, *Iterative reconstruction of memory kernels*, J. Chem. Theo. Comput. **13**, 2481 (2017).
- [121] G. Jung, M. Hanke, and F. Schmid, *Generalized Langevin dynamics: construction and numerical integration of non-Markovian particle-based models*, Soft Matter **14**, 9368 (2018).
- [122] L. Yeomans-Reyna, and M. Medina-Noyola, *Overdamped van Hove function of colloidal suspensions*, Phys. Rev. E **62**, 3382 (2000).
- [123] M. A. Chávez-Rojo, and M. Medina-Noyola, *Van Hove function of colloidal mixtures: Exact results*, Physica A **366**, 55 (2006).
- [124] E. Lázaro-Lázaro, P. Mendoza-Méndez, L. F. Elizondo-Aguilera, J. A. Perera-Burgos, P. E. Ramirez-González, G. Pérez-Angel, R. Castaneda-Priego, and M. Medina-Noyola, *Self-consistent generalized Langevin equation theory of the dynamics of multicomponent atomic liquids*, J. Chem. Phys. **146**, 184506 (2017).
- [125] A. Anderson, *Canonical transformations in quantum mechanics*, Ann. Phys. **232**, 292 (1994).
- [126] M.-L. M. Tchenkoue, M. Penz, I. Theophilou, M. Ruggenthaler, and A. Rubio, *Force balance approach for advanced approximations in density functional theories*, J. Chem. Phys. **151**, 154107 (2019).
- [127] W. Tarantino and C. A. Ullrich, *A reformulation of time-dependent Kohn-Sham theory in terms of the second time derivative of the density*, J. Chem. Phys. **154**, 204112 (2021).
- [128] M. Schmidt, *Quantum power functional theory for many-body dynamics*, J. Chem. Phys. **143**, 174108 (2015).
- [129] I. V. Tokatly, *Time-dependent deformation functional theory*, Phys. Rev. B **75**, 125105 (2007).
- [130] I. V. Tokatly, *Quantum many-body dynamics in a Lagrangian frame: I. Equations of motion and conservation laws*, Phys. Rev. B **71**, 165104 (2005).

- [131] I. V. Tokatly, *Quantum many-body dynamics in a Lagrangian frame: II. Geometric formulation of time-dependent density functional theory*, Phys. Rev. B **71**, 165105 (2005).
- [132] C. A. Ullrich and I. V. Tokatly, *Nonadiabatic electron dynamics in time-dependent density-functional theory*, Phys. Rev. B **73**, 235102 (2006).
- [133] P.-M. König, R. Roth, and K. R. Mecke, *Morphological thermodynamics of fluids: shape dependence of free energies*, Phys. Rev. Lett. **93**, 160601 (2004).
- [134] P. Bryk, R. Roth, K. R. Mecke, and S. Dietrich, *Hard-sphere fluids in contact with curved substrates*, Phys. Rev. E **68**, 031602 (2003).
- [135] R. Lovett and M. Baus, *A family of equivalent expressions for the pressure of a fluid adjacent to a wall*, J. Chem. Phys. **95**, 1991 (1991).
- [136] S. M. Tschopp and J. M. Brader, *Fundamental measure theory of inhomogeneous two-body correlation functions*, Phys. Rev. E **103**, 042103 (2021).
- [137] S. Auschra, V. Holubec, N. A. Söker, F. Cichos, and K. Kroy, *Polarization-density patterns of active particles in motility gradients*, Phys. Rev. E **103**, 062601 (2021).
- [138] N. A. Söker, S. Auschra, V. Holubec, K. Kroy, and F. Cichos, *How Activity landscapes polarize microswimmers without alignment forces*, Phys. Rev. Lett. **126**, 228001 (2021).
- [139] M. Revzen, *Functional integrals in statistical physics*, Am. J. Phys. **28**, 611 (1970).
- [140] S. Auschra and V. Holubec, *Density and polarization of active Brownian particles in curved activity landscapes*, Phys. Rev. E **103**, 062604 (2021).
- [141] J. Rodenburg, S. Paliwal, M. de Jager, P. G. Bolhuis, M. Dijkstra, and R. van Roij, *Ratchet-induced variations in bulk states of an active ideal gas*, J. Chem. Phys. **149**, 174910 (2018).
- [142] A. Fischer, F. Schmid, and T. Speck, *Quorum-sensing active particles with discontinuous motility*, Phys. Rev. E **101**, 012601 (2020).
- [143] H.-R. Jiang, N. Yoshinaga, and M. Sano, *Active motion of a Janus particle by self-thermophoresis in a defocused laser beam*, Phys. Rev. Lett. **105**, 268302 (2010).
- [144] S. Auschra, A. Bregulla, K. Kroy, and F. Cichos, *Thermotaxis of Janus particles*, Eur. Phys. J. E **44**, 90 (2021).
- [145] T. Speck and R. L. Jack, *Ideal bulk pressure of active Brownian particles*, Phys. Rev. E **93**, 062605 (2016).
- [146] S. C. Takatori, W. Yan, and J. F. Brady, *Swim pressure: stress generation in active matter*, Phys. Rev. Lett. **113**, 028103 (2014).
- [147] X. Yang, M. L. Manning, and M. C. Marchetti, *Aggregation and segregation of confined active particles*, Soft Matter **10**, 6477 (2014).
- [148] R. G. Winkler, A. Wysocki, and G. Gompper, *Virial pressure in systems of spherical active Brownian particles*, Soft Matter **11**, 6680 (2015).
- [149] *11th Liquid matter conference*, Prague, Czech Republic, July 18-23 (2021).

- [150] J. S. Rowlinson and B. Widom, *Molecular theory of Capillarity* (Dover, New York, 2002).
- [151] J. Kerins and M. Boiteux, *Applications of Noether's theorem to inhomogeneous fluids*, Phys. A **117**, 575 (1983).
- [152] M. Boiteux and J. Kerins, *Thermodynamic properties of inhomogeneous fluids*, Phys. A **121**, 399 (1983).
- [153] D. J. Bukman, *Torque balance at a line of contact*, Phys. A **319**, 151 (2003).
- [154] M. Maier, A. Zippelius, and M. Fuchs, *Emergence of long-ranged stress correlations at the liquid to glass transition*, Phys. Rev. Lett. **119**, 265701 (2017).
- [155] M. Maier, A. Zippelius, and M. Fuchs, *Stress auto-correlation tensor in glass-forming isothermal fluids: From viscous to elastic response*, J. Chem. Phys. **149**, 084502 (2018).
- [156] A. O. Parry, C. Rascón and R. Evans, *The local structure factor near an interface; beyond extended capillary-wave models*, J. Phys.: Condens. Matter **28**, 244013 (2016).
- [157] A. O. Parry and C. Rascón, *The Goldstone mode and resonances in the fluid interfacial region*, Nat. Phys. **15**, 287 (2019).
- [158] F. Vogel, A. Zippelius, and M. Fuchs, *Emergence of Goldstone excitations in stress correlations of glass-forming colloidal dispersions*, EPL **125**, 68003 (2019).
- [159] J. Stenhammar, D. Marenduzzo, R. J. Allen, and M. E. Cates, *Phase behaviour of active Brownian particles: the role of dimensionality*, Soft matter **10**, 1489 (2014).
- [160] A. K. Omar, K. Klymko, T. GrandPre, and P. L. Geissler, *Phase diagram of active Brownian spheres: crystallization and the metastability of motility-induced phase separation*, Phys. Rev. Lett. **126**, 188002 (2021).
- [161] F. Turci and N. B. Wilding, *Phase separation and multibody effects in three-dimensional active Brownian particles*, Phys. Rev. Lett. **126**, 038002 (2021).
- [162] A. R. Dulaney, S. A. Mallory, and J. F. Brady, *The "isothermal" compressibility of active matter*, J. Chem. Phys. **154**, 014902 (2021).
- [163] R. Evans and M. C. Stewart, *The local compressibility of liquids near non-adsorbing substrates: a useful measure of solvophobicity and hydrophobicity?*, J. Phys.: Condens. Matter **27**, 194111 (2015).
- [164] R. Evans and N. B. Wilding, *Quantifying density fluctuations in water at a hydrophobic surface: Evidence for critical drying*, Phys. Rev. Lett. **115**, 016103 (2015).
- [165] M. K. Coe, R. Evans, and N. B. Wilding, *Measures of fluctuations for a liquid near critical drying*, Phys. Rev. E **105**, 044801 (2022).
- [166] G. S. Redner, A. Baskaran, and M. F. Hagan, *Reentrant phase behavior in active colloids with attraction*, Phys. Rev. E **88**, 012305 (2013).
- [167] V. Prymidis, S. Paliwal, M. Dijkstra, and L. Fillion, *Vapour-liquid coexistence of an active Lennard-Jones fluid*, J. Chem. Phys. **145**, 124904 (2016).
- [168] S. Paliwal, V. Prymidis, L. Fillion, and M. Dijkstra, *Non-equilibrium surface tension of the vapour-liquid interface of active Lennard-Jones particles*, J. Chem. Phys. **147**, 084902 (2017).



- [169] J. Stenhammar, A. Tiribocchi, R. J. Allen, D. Marenduzzo, and M. E. Cates, *Continuum theory of phase separation kinetics for active Brownian particles*, Phys. Rev. Lett. **111**, 145702 (2013).



# 7 Publications

## 7.1 List of publications

Overview of full versions of all seven publications contributing to this cumulative thesis as listed in the table below. All papers have been published [SH1, SH2, SH3, SH4, SH5, SH6, SH7].

In these original papers we aimed to investigate symmetries and their consequences on a variety of many-body systems such as active Brownian particles. The resulting conclusions are useful for both theoretical and physically intuitive descriptions based on the central concept of forces in statistical mechanics. Specific applications include the sedimentation and phase behavior of active Brownian particles.

In publication [SH1] we studied how to conceptually transfer and apply Noether's theorem to statistical mechanics systems, in particular by exploiting the invariances

List of publications	Ref.	Pages
Noether's Theorem in Statistical Mechanics S. Hermann and M. Schmidt, <i>Commun. Phys.</i> <b>4</b> , 176 (2021)	[SH1]	71 – 83
Why Noether's Theorem applies to Statistical Mechanics S. Hermann and M. Schmidt, <i>J. Phys.: Condens. Matter</i> <b>34</b> , 213001 (2022) (Invited Topical Review)	[SH2]	85 – 98
Force density functional theory in- and out-of-equilibrium S. M. Tschopp, F. Sammüller, S. Hermann, M. Schmidt, and J. M. Brader, <i>Phys. Rev. E</i> <b>106</b> , 014115 (2022)	[SH3]	99 – 114
Force balance in thermal quantum many-body systems from Noether's theorem S. Hermann and M. Schmidt, <i>J. Phys. A: Math. Theor.</i> <b>55</b> , 464003 (2022)	[SH4]	115 – 132
Variance of fluctuations from Noether invariance S. Hermann and M. Schmidt, <i>Commun. Phys.</i> <b>5</b> , 276 (2022)	[SH5]	133 – 137
Active interface polarization as a state function S. Hermann and M. Schmidt, <i>Phys. Rev. Research</i> <b>2</b> , 022003 (2020) (Rapid Communication)	[SH6]	139 – 144
Non-negative interfacial tension in phase-separated active Brownian particles S. Hermann, D. de las Heras and M. Schmidt, <i>Phys. Rev. Lett.</i> <b>123</b> , 268002 (2019)	[SH7]	145 – 150

of statistical mechanics functionals such as the grand potential instead of the action functional as is conventionally done. These concepts were illustrated in several equilibrium and non-equilibrium applications for various different functionals and different types of inherent invariances. Explicit applications include the sedimenting and phase separating of active Brownian particles. A pedagogical and detailed introduction to this topic is given in publication [SH2] and it is therefore a good starting point for reading. Here we work on the basis of the canonical ensemble rather than in the grand canonical ensemble [SH1]. Our considerations based on Noether's theorem also generalize from classical statistical mechanics to quantum statistical mechanics [SH4]. In all cases we obtained statements about global forces and torques. The underlying basic concepts were subsequently generalized in publication [SH5] from linear to non-linear order terms. This work revealed a connection between the invariance of the system with respect to symmetry operations and the variance (auto-correlation) of global forces.

To go beyond global relations, which hold for the whole system, we considered local symmetries in publications [SH3,SH4]. The invariance of the grand potential under local canonical transformations generates the local force balance which constitutes the first order member of the YBG-hierarchy [SH3]. This equation forms the basis of a force-based density functional theory which was developed in [SH3]. Hence this approach constitutes an alternative implementation of classical density functional theory. The work [SH3] also includes proofs for the virial and compressibility contact theorem.

Related to the sum rules that follow from applying Noether's theorem we have also obtained exact statements for the global polarization of particles as a result of the continuity equation. We developed this expression for particles whose orientations change by rotational diffusion in publication [SH6]. The global polarization depends on the boundary values of the current and we applied the sum rule to active Brownian particles. Those particles are intrinsically non-symmetric and can hence develop a polarization as they have an assigned swim direction. The polarization sum rule was verified experimentally [138] and numerically [137] for active Brownian particles under a local abrupt activity step.

At sufficiently high densities and swim speeds active Brownian particles tend to phase separate into a dense and a dilute phase. We derived in publication [SH7] results for the tension of the corresponding interface based on the previously determined theoretical description of the phase coexistence [46,94]. We obtain positive values for the tension in contrast to several claims in the literature [47,49–56]. All determined sum rules [SH1,SH2,SH3,SH5,SH6] are tested for the theoretical description [46,94] and turn out to be satisfied.

## 7.2 Author's contributions

The publications [SH1,SH2,SH4,SH5,SH6] were designed, carried out and written in close cooperation with the author's supervisor M. Schmidt. In [SH6] the author identified the research question and independently wrote the first draft of the manuscript.

In publication [SH3] the author derived the local force density balance using Noether's theorem (section III.A) together with M. Schmidt. She developed in collaboration with

J. M. Brader and S. M. Tschopp the proofs of the virial and compressibility contact theorem (sections III.D and III.E in [SH3]). In publication [SH7] the author performed the computer algebra calculations and generated the theoretical data. Furthermore the author contributed to the analytic calculations and to the preparation of the manuscript.

The author was significantly involved in all work on the revision steps in response to the referee comments for all the papers.



# communications physics

ARTICLE

<https://doi.org/10.1038/s42005-021-00669-2>

OPEN

## Noether's theorem in statistical mechanics

Sophie Hermann <sup>1</sup>✉ & Matthias Schmidt <sup>1</sup>✉

Noether's calculus of invariant variations yields exact identities from functional symmetries. The standard application to an action integral allows to identify conservation laws. Here we rather consider generating functionals, such as the free energy and the power functional, for equilibrium and driven many-body systems. Translational and rotational symmetry operations yield mechanical laws. These global identities express vanishing of total internal and total external forces and torques. We show that functional differentiation then leads to hierarchies of local sum rules that interrelate density correlators as well as static and time direct correlation functions, including memory. For anisotropic particles, orbital and spin motion become systematically coupled. The theory allows us to shed new light on the spatio-temporal coupling of correlations in complex systems. As applications we consider active Brownian particles, where the theory clarifies the role of interfacial forces in motility-induced phase separation. For active sedimentation, the center-of-mass motion is constrained by an internal Noether sum rule.

<sup>1</sup>Theoretische Physik II, Physikalisches Institut, Universität Bayreuth, Bayreuth, Germany. ✉email: [Sophie.Hermann@uni-bayreuth.de](mailto:Sophie.Hermann@uni-bayreuth.de); [Matthias.Schmidt@uni-bayreuth.de](mailto:Matthias.Schmidt@uni-bayreuth.de)

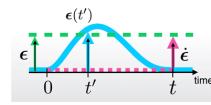
Emmy Noether's 1918 Theorems for *Invariant Variation Problems*<sup>1,2</sup>, as applied to action functionals both in particle-based and field-theoretic contexts, form a staple of our fundamental description of nature. The formulation of energy conservation in general relativity had been the then open and vexing problem, that triggered Hilbert and Klein to draw Noether into their circle, and she ultimately solved the problem<sup>3</sup>. Her deep insights into the relationship of the emergence and validity of conservation laws with the underlying local and global symmetries of the system has been exploited for over a century.

While Noether's work has been motivated by the then ongoing developments in general relativity, being a mathematician, she has formulated her theory in a much broader setting than given by the specific structure of the action as a space-time integral over a Lagrangian density, as formulated by Hilbert in 1916 for Einstein's field equations. Her work rather applies to functionals of a much more general nature, with only mild assumptions of analyticity and careful treatment of boundary conditions of integration domains.

In Statistical Physics, the use of Noether's theorems is significantly more scarce, as opposed to both classical mechanics and high energy physics. Notable exceptions include the square-gradient treatment of the free gas-liquid interface, cf. Rowlinson and Widom's enlightening description<sup>4</sup> of van der Waals' prototypical solution<sup>5</sup>. In a striking analogy, the square gradient contribution to the free energy is mapped onto kinetic energy of an effective particle that traverses in time between two potential energy maxima of equal height. Exploiting energy conservation in the effective system yields a first integral, which constitutes a nontrivial identity in the statistical problem. This reasoning has been generalized to the delicate problem of the three-phase contact line that occurs at a triple point of a fluid mixture<sup>6,7</sup>. While these treatments strongly rely on the square-gradient approximation, Boiteux and Kerins also developed a method that they refer to as variation under extension, which permitted them to treat more general cases<sup>8</sup>.

Evans has derived a number of exact sum rules for inhomogeneous fluids in his pivotal treatment of the field<sup>9</sup>. While not spelling out any connection to Noether's work, he carefully examines the effects of spatial displacements on distribution functions. This shifting enables him, as well as Lovett et al.<sup>10</sup> and Wertheim<sup>11</sup> in earlier work, to identify systematically the effects that result from the displacement and formulate these as highly nontrivial interrelations ("sum rules") between correlation functions. This approach was subsequently generalized to higher than two-body direct<sup>12</sup> and density<sup>13</sup> correlation functions and the relationship to integral equation theory was addressed<sup>14,15</sup>. Considering also rotations Tarazona and Evans<sup>16</sup> have addressed the case of anisotropic particles, where their sum rules correct earlier results by Gubbins<sup>17</sup>. The exploitation of the fundamental spatial symmetries<sup>9-16</sup> appears to be intimately related to Noether's thinking. This is no coincidence, as Evans' classical density functional approach (DFT) is variational as is the general problem that she addresses.

DFT constitutes a powerful modern framework for the description of a broad range of interfacial, adsorption, solvation, and phase phenomenology in complex systems<sup>9,18-20</sup>. Examples of recent pivotal applications include the treatments of hydrophobicity<sup>21-27</sup> and of drying<sup>23,24,26</sup>, electrolytes near surfaces<sup>28</sup>, dense fluid structuring as revealed in atomic force microscopy<sup>29</sup>, thermal resistance of liquid-vapor interfaces<sup>30</sup>, and layered freezing in confined colloids<sup>31</sup>. Xu and Rice<sup>31</sup> have used the sum rules of Lovett et al.<sup>10</sup> and Wertheim<sup>11</sup> (LMBW) to carry out a bifurcation analysis of the confined fluid state. The sum rules were instrumental for investigating a range of topics, such as precursors to freezing<sup>32</sup>, nonideal<sup>33</sup> and cluster crystals<sup>34</sup>, liquid



**Fig. 1 Illustration of the three types of dynamical transformations considered.** The system is spatially displaced by  $\epsilon = \text{const}$  at all times (green dashed), analogously to the operation in equilibrium. The system is dynamically displaced by  $\epsilon(t')$ , such that the spatial displacement vanishes at the boundaries of the considered time interval,  $\epsilon(0) = \epsilon(t) = 0$  (cyan solid). The system is displaced instantaneously only at the latest time  $t$ , such that the differential displacement is  $\epsilon \dot{d}t$  (purple dotted).

crystal deformations<sup>35</sup>, and –prominently– interfaces of liquids<sup>36-40</sup>. A range of further techniques besides DFT was used in this context, including integral equation theory<sup>32,36,39</sup>, mode-coupling theory<sup>41</sup>, and Mori-Zwanzig equations<sup>33,35</sup>.

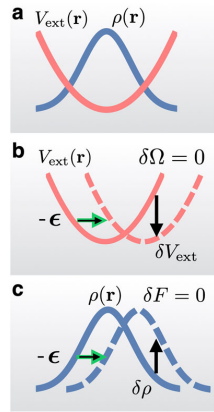
Much of very current attention in Statistical Physics is devoted to nonequilibrium and active systems that are driven in a controlled way out of equilibrium, such as e.g. active Brownian particles<sup>42-44</sup> and magnetically controlled topological transport of colloids<sup>45-47</sup>. The power functional (variational) theory<sup>48</sup> (PFT) offers to obtain a unifying perspective on nonequilibrium problems such as the above. In PFT the (time-dependent) density distribution is complemented by the (time-dependent) current distribution as a further variational field. A rigorous extremal principle determines the motion of the system, on the one-body level of correlation functions. The concept enabled to obtain a fundamental understanding and quantitative description of a significant array of nonequilibrium phenomena, such as the identification of superadiabatic forces<sup>49</sup>, the treatment of active Brownian particles<sup>50-53</sup>, of viscous<sup>54</sup>, structural<sup>55,56</sup> and flow forces<sup>56</sup>. Crucially, the DFT remains relevant for the description of nonequilibrium situations, via the adiabatic construction<sup>48,49</sup>, which captures those parts of the dynamics that functionally depend on the density distribution alone, and do so instantaneously. Both equilibrium DFT and nonequilibrium PFT provide formally exact variational descriptions of their respective realm of Statistical Physics. While action integrals feature in neither formulation, the relevant functionals do fall into the general class of functionals that Noether considered in her work.

Here we apply Noether's theorem to Statistical Physics. We first introduce the basic concepts via treating spatial translations for both the partition sum and for the free energy density functional. Considering the symmetries of the partition sum does not require to engage with density functional concepts; the elementary definition suffices. We demonstrate that this approach is consistent with the earlier work in equilibrium<sup>9-16</sup>, and that it enables one to go, with relative ease, beyond the sum rules that these authors formulated. In nonequilibrium, we apply the same symmetry operations to the time-dependent case and obtain novel exact and nontrivial identities that apply for driven and active fluids. The three different types of time-dependent shifting are illustrated in Fig. 1. The resulting sum rules are different from the nonequilibrium Ornstein-Zernike (NOZ) relations<sup>57,58</sup>, but they possess an equally fundamental status. We also consider the more general case of anisotropic interparticle interactions and treat rotational invariance both in and out of equilibrium. To illustrate the theory we apply it to both passive and active phase coexistence as well as to active sedimentation under gravity.

## Results and discussion

**Adiabatic state.** We start with an initial illustration of Noether's concept as applied to the grand potential  $\Omega$ . We consider spatial translations of the position coordinate  $\mathbf{r}$  at fixed chemical





**Fig. 2 Illustrations of the effects induced by shifting in equilibrium.** **a** In the presence of external potential  $V_{\text{ext}}(\mathbf{r})$ , the system develops an inhomogeneous density profile  $\rho(\mathbf{r})$ , where  $\mathbf{r}$  denotes the position coordinate. **b** Shifting the external potential by a displacement vector  $-\epsilon$  (green arrow) induces a local change in external potential  $\delta V_{\text{ext}}(\mathbf{r})$  (black arrow) between the original (solid line) and the shifted external potential (dashed line); the grand potential is invariant,  $\delta\Omega = 0$ . **c** The displaced density profile (dashed line) implies a local change  $\delta\rho(\mathbf{r})$  (black arrow) in comparison to the initial density profile (solid line), which leaves the intrinsic free energy unchanged,  $\delta F = 0$ .

potential  $\mu$  and fixed temperature  $T$ . The system is under the influence of a one-body external potential  $V_{\text{ext}}(\mathbf{r})$ , cf. Fig. 2a. We take  $V_{\text{ext}}(\mathbf{r})$  to also describe container walls, such that there is no need for the system volume as a further thermodynamic variable. For the moment we only examine systems completely bounded by external walls. Systems with open boundaries are considered below. Clearly the value of the grand potential  $\Omega$  is independent of the global location of a system. Hence spatial shifting by a (global) displacement vector  $\epsilon$  leaves the value of  $\Omega$  invariant. To exploit this symmetry in a variational setting, note that the value of  $\Omega$  depends on the function  $V_{\text{ext}}(\mathbf{r})$ , hence  $V_{\text{ext}}(\mathbf{r}) \rightarrow \Omega$  constitutes a functional map, at given  $\mu$  and  $T$ . Here the grand potential is defined by its elementary Statistical Mechanics form  $\Omega[V_{\text{ext}}] = -k_B T \ln \Xi$ , with the grand partition sum  $\Xi$  depending functionally via the Boltzmann factor on  $V_{\text{ext}}(\mathbf{r})$ . The spatial displacement amounts to the operation  $V_{\text{ext}}(\mathbf{r}) \rightarrow V_{\text{ext}}(\mathbf{r} + \epsilon)$ , cf. Fig. 2b. For small  $\epsilon$  we can Taylor expand to linear order:  $V_{\text{ext}}(\mathbf{r} + \epsilon) = V_{\text{ext}}(\mathbf{r}) + \delta V_{\text{ext}}(\mathbf{r})$ , where  $\delta V_{\text{ext}}(\mathbf{r}) = \epsilon \cdot \nabla V_{\text{ext}}(\mathbf{r})$  indicates the local change of the external potential that is induced by the shift. As  $\Omega[V_{\text{ext}}]$  is invariant under the shift (which can be shown by translating all particle coordinates in  $\Xi$  accordingly), we have

$$\Omega[V_{\text{ext}}] = \Omega[V_{\text{ext}} + \delta V_{\text{ext}}] = \Omega[V_{\text{ext}}] + \int d\mathbf{r} \frac{\delta\Omega[V_{\text{ext}}]}{\delta V_{\text{ext}}(\mathbf{r})} \epsilon \cdot \nabla V_{\text{ext}}(\mathbf{r}). \quad (1)$$

Here the second equality constitutes a functional Taylor expansion in  $\delta V_{\text{ext}}(\mathbf{r})$  to linear order, and  $\delta\Omega[V_{\text{ext}}]/\delta V_{\text{ext}}(\mathbf{r})$  indicates the functional derivative of  $\Omega[V_{\text{ext}}]$  with respect to its argument, evaluated here at the unshifted function  $V_{\text{ext}}(\mathbf{r})$ , i.e.  $\epsilon = 0$ . It is a straightforward elementary exercise<sup>9,18</sup> to show via explicit calculation that  $\delta\Omega[V_{\text{ext}}]/\delta V_{\text{ext}}(\mathbf{r}) = \rho(\mathbf{r})$ , where  $\rho(\mathbf{r}) = \langle \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \rangle_{\text{eq}}$  is the microscopically resolved one-body density profile. Here  $\mathbf{r}_i$  indicates the position of particle  $i = 1 \dots N$ , with  $N$  being the total number of particles,  $\delta(\cdot)$  indicates the Dirac

distribution, and the average is over the equilibrium distribution at fixed  $\mu$  and  $T$ ; the sum runs over all particles  $i = 1 \dots N$ .

Comparing the left and right hand sides of (1) and noticing that  $\epsilon$  is arbitrary, we conclude

$$\mathbf{F}_{\text{ext}}^{\text{tot}} = - \int d\mathbf{r} \rho(\mathbf{r}) \nabla V_{\text{ext}}(\mathbf{r}) = 0, \quad (2)$$

where we have defined the total external force  $\mathbf{F}_{\text{ext}}^{\text{tot}}$  using the one-body fields  $\rho(\mathbf{r})$  and  $V_{\text{ext}}(\mathbf{r})$ . It is straightforward to show the equivalence with the more elementary form  $\mathbf{F}_{\text{ext}}^{\text{tot}} = - \langle \sum_i \nabla_i V_{\text{ext}}(\mathbf{r}_i) \rangle_{\text{eq}}$ , where  $\nabla_i$  indicates the derivative with respect to  $\mathbf{r}_i$ . Clearly, (2) expresses the vanishing of the total external force (consider e.g. the gravitational weight of an equilibrium colloidal sediment being balanced by the force that the lower container wall exerts on the particles).

Equation (2) was previously obtained by Baus<sup>13</sup>. Here we have identified it as a Noether sum rule for the case of spatial displacement of  $\Omega[V_{\text{ext}}]$ . We can generate local sum rules by observing that (2) holds for any form of  $V_{\text{ext}}(\mathbf{r})$  and that hence  $V_{\text{ext}}(\mathbf{r}) \rightarrow \rho(\mathbf{r})$  constitutes a functional map (defined by the grand canonical average  $\langle \hat{\rho}(\mathbf{r}) \rangle_{\text{eq}}$ , which features  $V_{\text{ext}}(\mathbf{r})$  in the equilibrium many-body probability distribution). We hence functionally differentiate (2) by  $V_{\text{ext}}(\mathbf{r}')$ , where  $\mathbf{r}'$  is a new position variable. The first and the  $n$ th functional derivatives yield, respectively, the identities

$$\nabla \rho(\mathbf{r}) = - \int d\mathbf{r}' \beta H_2(\mathbf{r}, \mathbf{r}') \nabla' V_{\text{ext}}(\mathbf{r}'), \quad (3)$$

$$\sum_{\alpha=1}^n \nabla_{\alpha} H_n = - \int d\mathbf{r}_{n+1} \beta V_{\text{ext}}(\mathbf{r}_{n+1}) \nabla_{n+1} H_{n+1}, \quad (4)$$

where  $\beta = 1/(k_B T)$ , with  $k_B$  indicating the Boltzmann constant,  $H_2(\mathbf{r}, \mathbf{r}') = -\delta\rho(\mathbf{r})/\delta\beta V_{\text{ext}}(\mathbf{r}')$  is the two-body correlation function of density fluctuations, and  $H_n = \delta H_{n-1}/\delta\beta V_{\text{ext}}(\mathbf{r}_n)$  is its  $n$ -body version<sup>9,18</sup>. Here position arguments have been omitted for clarity:  $H_n \equiv H_n(\mathbf{r}_1 \dots \mathbf{r}_n)$ , and  $\nabla_{\alpha}$  indicates the derivative with respect to  $\mathbf{r}_{\alpha}$ . The variable names  $\mathbf{r}$  and  $\mathbf{r}'$  have been interchanged in (3) and  $\nabla'$  indicates the derivative with respect to  $\mathbf{r}'$ . The derivation of (3) and (4) requires spatial integration by parts. Recall that boundary terms vanish as we only consider systems with impenetrable bounding walls.

The sum rule (3) has been obtained by LMBW<sup>10,11</sup> and by Evans<sup>9</sup> on the basis of shifting considerations. The present formulation based on Noether's more general perspective allows to reproduce (3) with great ease and to generalize to the hierarchy (4), as previously obtained by Baus<sup>13</sup>. Equation (3) has the interpretation of the density gradient  $\nabla \rho(\mathbf{r})$  being stabilized by the action of the external force field,  $-\nabla V_{\text{ext}}(\mathbf{r})$ . The effect is mediated by  $\beta H_2(\mathbf{r}, \mathbf{r}')$ , where the correlation of the density fluctuations is due to the coupled nature of the interparticle interactions. Equation (4) is the multi-body generalization of this mechanism. Via multiplying (3) by  $V_{\text{ext}}(\mathbf{r})$ , integrating over  $\mathbf{r}$ , and using (2), and iteratively repeating this process for all orders, one obtains a multi-body analog of the vanishing external force (2):

$$\int d\mathbf{r}_1 V_{\text{ext}}(\mathbf{r}_1) \dots \int d\mathbf{r}_n V_{\text{ext}}(\mathbf{r}_n) \nabla_{\alpha} H_n = 0, \quad (5)$$

for  $\alpha = 1 \dots n$ .

We turn to intrinsic contributions. As Noether's theorem poses no restriction on the type of physical functional, we consider the intrinsic Helmholtz free energy  $F[\rho]$  as a functional of the density profile as its natural argument. Here a functional Legendre transform<sup>9,18</sup> yields  $F[\rho] = \Omega[V_{\text{ext}}] - \int d\mathbf{r} (V_{\text{ext}}(\mathbf{r}) - \mu) \rho(\mathbf{r})$ . Crucially,  $F[\rho]$  is independent of  $V_{\text{ext}}(\mathbf{r})$ , and its excess (over ideal gas) contribution  $F_{\text{exc}}[\rho]$  is specific to the form of the interparticle interaction potential  $u$

( $\mathbf{r}^N$ ); here we use the shorthand  $\mathbf{r}_1 \dots \mathbf{r}_N \equiv \mathbf{r}^N$ . The full intrinsic free energy functional consists of a sum of ideal gas and excess contributions, i.e.  $F[\rho] = k_B T \int d\mathbf{r} \rho(\mathbf{r}) [\ln(\rho(\mathbf{r}) \Lambda^D) - 1] + F_{\text{exc}}[\rho]$ , where  $\Lambda$  is the (irrelevant) thermal de Broglie wavelength and  $D$  is the dimensionality of space.

As  $u(\mathbf{r}^N)$  is globally translationally invariant,  $F_{\text{exc}}[\rho]$  will not change its value when evaluated at a spatially displaced density,  $\rho(\mathbf{r} + \boldsymbol{\epsilon}) = \rho(\mathbf{r}) + \delta\rho(\mathbf{r})$ , where  $\delta\rho(\mathbf{r}) = \boldsymbol{\epsilon} \cdot \nabla \rho(\mathbf{r})$ , cf. Fig. 2(c). Hence in analogy to (1), we obtain  $F_{\text{exc}}[\rho] = F_{\text{exc}}[\rho + \delta\rho] = F_{\text{exc}}[\rho] + \int d\mathbf{r} (\delta F_{\text{exc}}[\rho] / \delta\rho(\mathbf{r})) \boldsymbol{\epsilon} \cdot \nabla \rho(\mathbf{r})$ . Again  $\boldsymbol{\epsilon}$  is arbitrary. As boundary terms vanish in the considered systems, integration by parts yields

$$\mathbf{F}_{\text{ad}}^{\text{tot}} = - \int d\mathbf{r} \rho(\mathbf{r}) \nabla \frac{\delta F_{\text{exc}}[\rho]}{\delta\rho(\mathbf{r})} = \int d\mathbf{r} \rho(\mathbf{r}) \mathbf{f}_{\text{ad}}(\mathbf{r}) = 0, \quad (6)$$

where the first equality expresses the total internal force  $\mathbf{F}_{\text{ad}}^{\text{tot}} = -(\sum_i \nabla_i u(\mathbf{r}^N))_{\text{eq}}$  in DFT language. Hence (6) expresses the fact that the total internal force vanishes in equilibrium; the more general time-dependent case is treated below. The functional derivatives of  $F_{\text{exc}}[\rho]$  constitute direct correlation functions<sup>9,18,59</sup>, with the lowest order being the one-body direct correlation function  $c_1(\mathbf{r}) = -\delta F_{\text{exc}}[\rho] / \delta\rho(\mathbf{r})$ . The equilibrium ("adiabatic") force field is simply  $\mathbf{f}_{\text{ad}}(\mathbf{r}) = k_B T \nabla c_1(\mathbf{r})$ . This one-body force field arises from the interparticle forces that all other particles exert on the particle that resides at position  $\mathbf{r}$ .

From the global internal Noether sum rule (6), we can obtain local sum rules by observing that (6) holds for all  $\rho(\mathbf{r})$  and hence that its functional derivative with respect to  $\rho(\mathbf{r})$  vanishes identically, i.e.,

$$\nabla c_1(\mathbf{r}) = \int d\mathbf{r}' c_2(\mathbf{r}, \mathbf{r}') \nabla' \rho(\mathbf{r}'), \quad (7)$$

$$\sum_{\alpha=1}^n \nabla_{\alpha} c_n = - \int d\mathbf{r}_{n+1} \rho(\mathbf{r}_{n+1}) \nabla_{n+1} c_{n+1}, \quad (8)$$

where  $c_2(\mathbf{r}, \mathbf{r}')$  is the (inhomogeneous) two-body direct correlation function of liquid state theory<sup>18</sup>;  $c_n \equiv c_n(\mathbf{r}_1 \dots \mathbf{r}_n)$  is the  $n$ -body direct correlation function, defined recursively via  $c_{n+1} = \delta c_n / \delta\rho(\mathbf{r}_{n+1})$ . As identified by LMBW<sup>10,11</sup> and Evans<sup>9</sup>, (7) expresses the conversion of the density gradient, via the two-body direct correlations, to the locally resolved intrinsic force field; recall that  $\mathbf{f}_{\text{ad}}(\mathbf{r}) = k_B T \nabla c_1(\mathbf{r})$ . Via the Noether formalism the corresponding hierarchy (8) is obtained straightforwardly from repeated functional differentiation<sup>12-14</sup> with respect to  $\rho(\mathbf{r})$ . Note that similar to the structure of (4), only consecutive terms of order  $n$  and  $n+1$  are directly coupled in (8). A multi-body version of (6) is obtained by multiplying (7) with  $\rho(\mathbf{r})$ , integrating over  $\mathbf{r}$ , exploiting (6), and iterating for all orders. The result is:

$$\int d\mathbf{r}_1 \rho(\mathbf{r}_1) \dots \int d\mathbf{r}_n \rho(\mathbf{r}_n) \nabla_{\alpha} c_n = 0, \quad (9)$$

for  $\alpha = 1 \dots n$ . In the case  $\alpha = n = 1$  we recover (6).

The global sum rule (6) of vanishing total internal force can be straightforwardly obtained by more elementary analysis. We exploit translation invariance in this non-functional setting:  $u(\mathbf{r}^N) \equiv u(\mathbf{r}_1 + \boldsymbol{\epsilon} \dots \mathbf{r}_N + \boldsymbol{\epsilon})$ . Then the derivative with respect to  $\boldsymbol{\epsilon}$  vanishes,  $0 = \partial u(\mathbf{r}_1 + \boldsymbol{\epsilon} \dots \mathbf{r}_N + \boldsymbol{\epsilon}) / \partial \boldsymbol{\epsilon} = \sum_i \nabla_i u(\mathbf{r}^N)$ . The latter expression follows from the chain rule and constitutes the total internal force (up to a minus sign), which hence vanishes for each microstate  $\mathbf{r}^N$ . The connection to (the many-body version of) Newton's third law *actio equals reactio* becomes apparent in the rewritten form  $-\nabla_{\alpha} u(\mathbf{r}^N) = \sum_{i \neq \alpha} \nabla_i u(\mathbf{r}^N)$ , for  $\alpha = 1 \dots N$ . The thermal equilibrium average is then trivial and on average  $\mathbf{F}_{\text{ad}}^{\text{tot}} = 0$ . This argument is very general and it remains true if the average is taken over a nonequilibrium many-body

distribution function. The total internal force in such a general situation is

$$\mathbf{F}_{\text{int}}^{\text{tot}} = - \langle \sum_i \nabla_i u(\mathbf{r}^N) \rangle = 0, \quad (10)$$

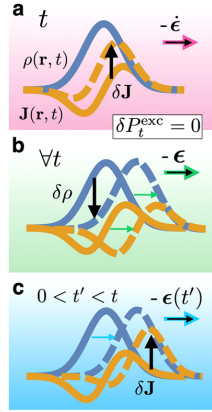
where the average is taken over the nonequilibrium many-body probability distribution at time  $t$ . We have hence proven that the total internal force vanishes for all times  $t$ . In addition, the particles can possess additional degrees of freedom  $\omega_i$ ,  $i = 1 \dots N$ , as is the case for the orientation vectors of active Brownian particles, to which we return after first laying out the setup in nonequilibrium.

**Nonequilibrium states.** To be specific, we consider overdamped Brownian motion, at constant temperature  $T$  and with no hydrodynamic interactions present<sup>18</sup>, as described by the Smoluchowski (Fokker-Planck) equation. The microscopically resolved local internal force field is  $\mathbf{f}_{\text{int}}(\mathbf{r}, t) = - \langle \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \nabla_i u(\mathbf{r}^N) \rangle / \rho(\mathbf{r}, t)$ , where the average is over the nonequilibrium distribution (which evolves in time according to the Smoluchowski equation) at time  $t$ . The total internal force is then the spatial integral  $\mathbf{F}_{\text{int}}^{\text{tot}} = \int d\mathbf{r} \rho(\mathbf{r}, t) \mathbf{f}_{\text{int}}(\mathbf{r}, t)$ . Applying Noether's theorem to the nonequilibrium case requires to have a variational description, as is provided by PFT<sup>48</sup>. Here the variational fields are the time-dependent density profile  $\rho(\mathbf{r}, t)$  and the time-dependence one-body current  $\mathbf{J}(\mathbf{r}, t) = \langle \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \mathbf{v}_i \rangle$ , where  $\mathbf{v}_i(\mathbf{r}^N, t)$  is the configurational velocity of particle  $i$ . The microscopically resolved average velocity profile is  $\mathbf{v}(\mathbf{r}, t) = \mathbf{J}(\mathbf{r}, t) / \rho(\mathbf{r}, t)$ . PFT ascertains the splitting  $\mathbf{f}_{\text{int}}(\mathbf{r}, t) = \mathbf{f}_{\text{ad}}(\mathbf{r}, t) + \mathbf{f}_{\text{sup}}(\mathbf{r}, t)$ , where the adiabatic force field is that in a corresponding equilibrium ("adiabatic") system with identical instantaneous density profile,  $\mathbf{f}_{\text{ad}}(\mathbf{r}, t) = - \nabla \delta F_{\text{exc}}[\rho] / \delta\rho(\mathbf{r}, t)$  and  $\mathbf{f}_{\text{sup}}(\mathbf{r}, t)$  is the superadiabatic internal force field, obtained as  $\mathbf{f}_{\text{sup}}(\mathbf{r}, t) = - \delta P_t^{\text{exc}}[\rho, \mathbf{J}] / \delta\mathbf{J}(\mathbf{r}, t)$ , where  $P_t^{\text{exc}}[\rho, \mathbf{J}]$  is the superadiabatic excess free power functional<sup>48</sup>.

Crucially,  $\mathbf{f}_{\text{ad}}(\mathbf{r}, t)$  is a density functional, independent of the flow in the system, while  $\mathbf{f}_{\text{sup}}(\mathbf{r}, t)$  is a kinematic functional, i.e. with dependence on both  $\rho(\mathbf{r}, t)$  and  $\mathbf{J}(\mathbf{r}, t)$ , including memory, i.e. dependence on the value of the fields at times  $< t$ . As the local force fields split into adiabatic and superadiabatic contributions, so do the total forces:  $\mathbf{F}_{\text{int}}^{\text{tot}} = \int d\mathbf{r} \rho \mathbf{f}_{\text{int}} = \int d\mathbf{r} \rho \mathbf{f}_{\text{ad}} + \int d\mathbf{r} \rho \mathbf{f}_{\text{sup}} \equiv \mathbf{F}_{\text{ad}}^{\text{tot}} + \mathbf{F}_{\text{sup}}^{\text{tot}}$ . We have seen above that  $\mathbf{F}_{\text{int}}^{\text{tot}} = \mathbf{F}_{\text{ad}}^{\text{tot}} = 0$ . Hence also

$$\mathbf{F}_{\text{sup}}^{\text{tot}} = - \int d\mathbf{r} \rho(\mathbf{r}, t) \frac{\delta P_t^{\text{exc}}[\rho, \mathbf{J}]}{\delta\mathbf{J}(\mathbf{r}, t)} \equiv \int d\mathbf{r} \rho \mathbf{f}_{\text{sup}} = 0. \quad (11)$$

While the above reasoning required to rely on the many-body level, the same result (11) can be straightforwardly obtained in a pure Noetherian way, by considering an instantaneous shift of coordinates at time  $t$ , i.e.  $\mathbf{J}(\mathbf{r}, t) \rightarrow \mathbf{J}(\mathbf{r}, t) - \dot{\boldsymbol{\epsilon}}\rho(\mathbf{r}, t)$ , cf. Fig. 3a (Fig. 3 gives an overview of the three different types of shifting). Here  $\dot{\boldsymbol{\epsilon}}$  is the corresponding instantaneous change in velocity with  $\mathbf{v}(\mathbf{r}, t) \rightarrow \mathbf{v}(\mathbf{r}, t) - \dot{\boldsymbol{\epsilon}}$ , as obtained by dividing the current by the density profile. Due to the overdamped nature of the Smoluchowski dynamics, the internal interactions are unaffected and the shift constitutes a symmetry operation for the generator of the superadiabatic forces,  $P_t^{\text{exc}}[\rho, \mathbf{J}]$ . Hence the instantaneous current perturbation  $\delta\mathbf{J}(\mathbf{r}, t) = -\dot{\boldsymbol{\epsilon}}\rho(\mathbf{r}, t)$  that is generated by the invariance transformation leads to  $P_t^{\text{exc}}[\rho, \mathbf{J}] = P_t^{\text{exc}}[\rho, \mathbf{J} + \delta\mathbf{J}] = P_t^{\text{exc}}[\rho, \mathbf{J}] - \int d\mathbf{r} (\delta P_t^{\text{exc}}[\rho, \mathbf{J}] / \delta\mathbf{J}(\mathbf{r}, t)) \cdot \dot{\boldsymbol{\epsilon}}\rho(\mathbf{r}, t)$ . As  $\dot{\boldsymbol{\epsilon}}$  is arbitrary, we obtain (11). Treating the dynamical adiabatic contribution  $\dot{F}[\rho] = \int d\mathbf{r} \mathbf{J}(\mathbf{r}, t) \cdot \nabla \delta F_{\text{exc}}[\rho] / \delta\rho(\mathbf{r}, t)$  in the same way, we re-obtain (6). As Noether's theorem is converse, it allows for alternative reasoning: the invariance of  $P_t^{\text{exc}}[\rho, \mathbf{J}]$  to the instantaneous shift of the current (or analogously of the velocity) can hence be derived from (11) by simply reversing the above chain of arguments.



**Fig. 3** Illustrations of the effects induced by shifting in nonequilibrium.

All transformations affect both the density profile  $\rho(\mathbf{r}, t)$  (blue) and the current profile  $\mathbf{J}(\mathbf{r}, t)$  (yellow), while the superadiabatic excess power functional is invariant,  $\delta P_t^{\text{exc}} = 0$ . Here  $t$  indicates the time and  $\mathbf{r}$  is the position coordinate. **a** An instantaneous spatial shift by  $\epsilon$  at time  $t$  induces a current change  $\delta \mathbf{J} = -\epsilon \rho$ . **b** A static shift by  $\epsilon = \text{const}$  is applied at all times  $t$ . **c** A time-dependent shift  $\epsilon(t')$  is applied between initial time 0 and final time  $t$  of the considered time interval.

We can generate nonequilibrium sum rules by differentiating the Noether identity (11) with respect to  $\mathbf{J}(\mathbf{r}, t)$ , which yields

$$\int d\mathbf{r}' M_2(\mathbf{r}, \mathbf{r}', t) \rho(\mathbf{r}', t) = 0, \quad (12)$$

where  $M_2(\mathbf{r}, \mathbf{r}', t) = -\beta \delta^2 P_t^{\text{exc}}[\rho, \mathbf{J}] / \delta \mathbf{J}(\mathbf{r}, t) \delta \mathbf{J}(\mathbf{r}', t)$  is the tensorial two-body equal-time direct correlation function<sup>57,58</sup>. Its  $n$ -body version is obtained from  $M_{n+1}(\mathbf{r}_1 \dots \mathbf{r}_{n+1}, t) = \delta M_n(\mathbf{r}_1 \dots \mathbf{r}_n, t) / \delta \mathbf{J}(\mathbf{r}_{n+1}, t)$ , and it satisfies the hierarchy

$$\int d\mathbf{r}_{n+1} M_{n+1}(\mathbf{r}_{n+1}, t) = 0, \quad (13)$$

as obtained by differentiating (12) repeatedly with respect to the current.

**Open boundaries.** All our considerations have been based on applying the symmetry operation to the entire system confined by external walls. The effects of these system walls are modeled by a suitable form of  $V_{\text{ext}}(\mathbf{r})$ . The position integrals formally run over all space, with the cutoff provided by hard (or steeply rising) external wall potentials. In many practical and relevant situations, it is more useful to consider a system with open boundaries. Alternatively one can consider only a subvolume  $V$  of the entire system, and restrict the accounting of force contributions to those particles that reside inside of  $V$  at a given time. In doing so, one needs to take account of boundary effects<sup>60</sup>, as the boundaries of  $V$  are open, such that interparticle forces can be transmitted, and flow can occur.

In case that there are no net boundary contributions, all previous derived sum rules still hold. This includes e.g. an effectively one-dimensional system in planar geometry that evolves to the same bulk state at the left and right boundaries or if the boundary conditions are periodic. In both cases left and right boundary terms are equal up to a minus sign and hence cancel each other. This example can be generalized straightforwardly to more complex geometries.

For nonvanishing net boundary terms additional contributions arise in the above sum rules. These contributions occur if the

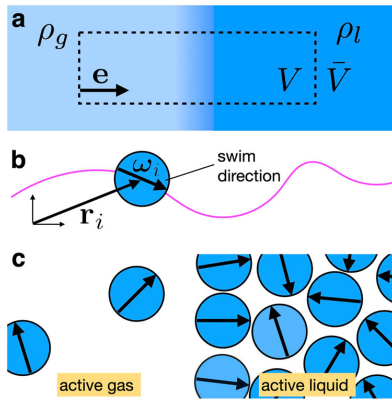
system develops different (bulk) states, e.g. for  $x \rightarrow \pm\infty$  as is relevant for bulk phase separation (see the section below). We demonstrate that such cases can be systematically treated in the current framework, by exemplary considering the total internal force. Then boundary force contributions arise due to an imbalance of “outside” particles that exert forces on “inside” particles. The outside particles are per definition excluded from the accounting of the total internal force exerted by all particles inside of  $V$ . The sum of all interactions between inside particles vanishes due to the global internal Noether sum rule (10). For simplicity we restrict ourselves to systems that interact via short-ranged pairwise central forces, where  $\mathbf{F}_{ij}$  indicates the force on particle  $i$  exerted by particle  $j$ . So only forces exerted from an inside to an outside particle contribute. The total internal force that acts on  $V$  is hence  $\mathbf{F}_{\text{int}}^{\text{tot}} = \langle \sum'_{ij} \mathbf{F}_{ij} \rangle$ , where the restricted sum (prime) runs only over those  $i \in V$  and  $j \in \bar{V}$ , where  $\bar{V}$  indicates the complement of  $V$ . The total internal force between particles inside of  $V$ , i.e.  $i \in V$  and  $j \in V$ , vanishes due to (10). We then rewrite  $\mathbf{F}_{\text{int}}^{\text{tot}}$  via inserting the identity  $\int d\mathbf{r} \delta(\mathbf{r}) = 1$  twice into the average. Then the restrictions of the sums can be transferred to restrictions on the spatial integration domains. As a result the total internal force acting on  $V$  can be expressed via correlation functions as

$$\mathbf{F}_{\text{int}}^{\text{tot}} = \int_V d\mathbf{r} \int_{\bar{V}} d\mathbf{r}' \langle \sum_{i,j \neq i} \delta(\mathbf{r} - \mathbf{r}_i) \delta(\mathbf{r}' - \mathbf{r}_j) \mathbf{F}_{ij} \rangle \quad (14)$$

$$= - \int_V d\mathbf{r} \int_{\bar{V}} d\mathbf{r}' \rho(\mathbf{r}) \rho(\mathbf{r}') g(\mathbf{r}, \mathbf{r}') \nabla \phi(|\mathbf{r} - \mathbf{r}'|), \quad (15)$$

where  $\phi(r)$  indicates the interparticle pair potential as a function of interparticle distance  $r$ . In order to obtain the form (15) we have identified the many-body definition of the radial distribution function  $g(\mathbf{r}, \mathbf{r}') = \langle \sum_{i,j \neq i} \delta(\mathbf{r} - \mathbf{r}_i) \delta(\mathbf{r}' - \mathbf{r}_j) \rangle / (\rho(\mathbf{r}) \rho(\mathbf{r}'))$ . Recall that the pair distribution function  $g$  and the density-density correlation function  $H_2$ , as used in (3)–(5), are related via  $H_2(\mathbf{r}, \mathbf{r}') = (g(\mathbf{r}, \mathbf{r}') - 1) \rho(\mathbf{r}) \rho(\mathbf{r}') + \delta(\mathbf{r} - \mathbf{r}') \rho(\mathbf{r})$ . Equation (15) still holds for non-conservative interparticle forces, when  $-\nabla \phi(|\mathbf{r} - \mathbf{r}'|)$  is replaced by the (nongradient) interparticle force field. We demonstrate in the following section the practical relevance of these considerations.

**Phase coexistence.** We turn to situations of phase coexistence. As we demonstrate, considering a large, but finite subvolume  $V$  of the entire system is useful but it also requires to take boundary terms into account. Here we take  $V$  to be cuboidal and to contain the free (planar) interface between two coexisting phases, see Fig. 4a for a graphical illustration. The volume boundaries parallel to the interface are taken to be seated deep inside either bulk phase. The internal force contributions on those faces of  $V$  that “cut through” the interface, i.e. have a normal that is perpendicular to the interface normal, vanish by symmetry. It remains to evaluate (15) over each of the two faces in the respective bulk region. Therefore the position dependences simplify to  $\rho(\mathbf{r}) = \rho_b = \text{const}$ , where  $\rho_b$  indicates the bulk number density, and the inhomogeneous pair distribution function simplifies as  $g(\mathbf{r}, \mathbf{r}') = g(|\mathbf{r} - \mathbf{r}'|)$ . Furthermore only force contributions colinear with  $\mathbf{e}$ , the outer interface normal of the considered bulk phase  $b$ , contribute. For a single face in bulk phase  $b$ , it is straightforward to show that the result is the virial pressure multiplied by the interface area  $A$ , i.e. the force  $A p_{\text{int}}^b \mathbf{e}$ , where the internal interaction pressure  $p_{\text{int}}^b$  in phase  $b$  is e.g. given via the Clausius virial<sup>18</sup>,  $p_{\text{int}}^b = -\frac{\pi}{2} \rho_b^2 \int_0^{\infty} dr r^2 g(r) \frac{d\phi}{dr}$  in two spatial



**Fig. 4 Phase separation into macroscopically distinct phases.**

**a** Illustration of the geometry of phase separation of passive or active particles. Shown are the gaseous  $\rho_g$  and liquid  $\rho_l$  plateau values of the density profile (indicated by the color gradient) and the direction vector  $\mathbf{e}$  normal to the interface. The total system volume  $V + \bar{V}$  consists of the subvolume  $V$  and its complement  $\bar{V}$ . **b** Illustration of an active Brownian particle (blue disc) with position  $\mathbf{r}_i$  and orientation  $\omega_i$ , undergoing translational and rotational diffusion. The self-propulsion along  $\omega_i$  creates directed motion as indicated by the trajectory (magenta line). **c** Schematics of motility-induced phase separation into active gas (left) and active liquid phases (right). The arrows indicate the orientations  $\omega_i$  of the active particles. The interface is polarized.

dimensions (the argument remains general). The constant  $r_0$  denotes the range of the interparticle interactions.

The total force density balance for equilibrium phase separation contains thermal diffusion and the internal force density, which cancel each other,

$$0 = -k_B T \nabla \rho(\mathbf{r}) + \rho(\mathbf{r}) \mathbf{f}_{\text{int}}(\mathbf{r}). \quad (16)$$

Integration over the volume  $V$  yields the total force which is proportional to the pressure. Hence the total internal force  $\mathbf{F}_{\text{int}}^{\text{tot}} = \int_V d\mathbf{r} \rho \mathbf{f}_{\text{int}}$  on  $V$ , cf. (15), amounts to the pressure difference  $(p_{\text{int}}^g - p_{\text{int}}^l) \mathbf{A} \mathbf{e}$ , where  $\mathbf{e}$  is the (unit vector) normal of the interface, pointing from, say, the gas (index  $g$ ) to the liquid phase (index  $l$ ). The total diffusive force is  $(p_{\text{id}}^g - p_{\text{id}}^l) \mathbf{A} \mathbf{e}$  with the pressure of the ideal gas  $p_{\text{id}} = k_B T \rho$ , evaluated at the gas ( $\rho_g$ ) and liquid bulk density ( $\rho_l$ ). As there are no external forces ( $V_{\text{ext}} \equiv 0$ ) then requesting the volume  $V$  to be forcefree amounts to  $p_{\text{tot}}^g = p_{\text{tot}}^l$ , where the total pressure is the sum  $p_{\text{tot}} = p_{\text{id}} + p_{\text{int}}$ . Hence the boundary consideration yields the mechanical equilibrium condition of equality of pressure in the coexisting phases.

We conclude that the internal interactions that occur across the free interface do not influence the (bulk) balance of the pressure at phase coexistence, as the net effect of these interactions vanishes. At the heart of this argument lies Noether's theorem for invariance against spatial displacements.

**Anisotropic particles.** We turn to anisotropic interparticle interactions, where  $\omega_i, \hat{\omega}_i$  are two perpendicular unit vectors that describes the particle orientation in space. Such systems are described by an interparticle interaction potential  $u(\mathbf{r}^N, \omega^N, \hat{\omega}^N)$ , which is assumed a priori to be invariant under spatial translations. Similarly one-body fields in general depend on position  $\mathbf{r}$  and orientations  $\omega$  and  $\hat{\omega}$  of the particles, e.g.  $V_{\text{ext}}(\mathbf{r}, \omega, \hat{\omega})$  for the

external field. The fully resolved one-body density distribution is  $\rho(\mathbf{r}, \omega, \hat{\omega}, t) = \langle \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \delta(\omega - \omega_i) \delta(\hat{\omega} - \hat{\omega}_i) \rangle$ .

It is straightforward to ascertain that all the above (force) sum rules for translation remain valid, as the orientations are unaffected by translations, upon trivially generalizing from position-only to position-orientation integration,  $\int d\mathbf{r} \rightarrow \int d\mathbf{r} d\omega d\hat{\omega}$  etc.

In the following for simplicity of notation we first consider uniaxial particles. Uniaxial particles depend only on one single orientation  $\omega_i$ , as the particles are rotationally invariant around this vector. Hence the  $\hat{\omega}$ -dependence of both the one- and many-body quantities vanish and the total integral simplifies to  $\int d\mathbf{r} d\omega$ .

**Motility-induced phase separation.** We use active Brownian particles as an example for uniaxial particles. For simplicity we consider spherical particles (discs) in two dimensions. The particles repel each other and they undergo self-propelled motion along their orientation vector  $\omega$ . Hence an additional one-body force  $\gamma s \omega$  acts on each swimmer, with  $\gamma$  the friction constant and  $s$  the speed of free swimming. This self-propulsion creates characteristic trajectories (see Fig. 4b for a schematic) which are also affected by thermal diffusion (omitted in the schematic) of the particle position  $\mathbf{r}_i$  and orientation  $\omega_i$ . Experimental realizations of active Brownian particles include e.g. Janus colloids driven by photon nudging<sup>61–65</sup>. If the density is high enough, motility-induced phase separation (MIPS) into an active gas and active liquid phase occurs for high enough values of the swim speed  $s$ , cf. Fig. 4c.

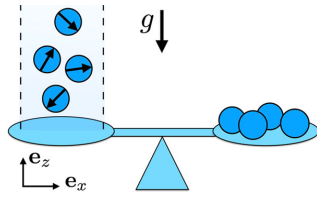
The force density balance (see e.g. the work of Hermann et al.<sup>52</sup>) of such a system in steady state (no time dependence) is

$$\begin{aligned} \gamma \mathbf{J}(\mathbf{r}, \omega) = & -k_B T \nabla \rho(\mathbf{r}, \omega) + \rho(\mathbf{r}, \omega) \mathbf{f}_{\text{int}}(\mathbf{r}, \omega) + \gamma s \rho(\mathbf{r}, \omega) \omega \\ & + \rho(\mathbf{r}, \omega) \mathbf{f}_{\text{ext}}(\mathbf{r}, \omega). \end{aligned} \quad (17)$$

The (negative) frictional force density on the left hand side is balanced with the ideal gas contribution (first term), the interparticle interactions (second term), the self-propulsion (third term) and the external contribution (fourth term) on the right hand side. As in case of the equilibrium phase separation we integrate (17) over the volume  $V$  (see Fig. 4(a) for an illustration) and over all orientations  $\omega$ . In the following we discuss each term separately. For simplicity we assume planar geometry of the system and we assume a vanishing external force,  $\mathbf{f}_{\text{ext}}(\mathbf{r}, \omega) = 0$ . Here, the interaction contribution  $p_{\text{int}} \mathbf{A} \mathbf{e}$  is obtained as above in equilibrium via (15), but the virial is averaged over the nonequilibrium steady state many-body probability distribution. The integral over the current  $\int d\mathbf{r} d\omega \mathbf{J}$  is assumed to vanish in steady state.

The total swim force that acts on  $V$  contributes to the total force. In the considered situation, the swim force is entirely due to the polarization  $\mathbf{M}_{\text{tot}} = \int d\mathbf{r} d\omega \omega \rho$  of the free interface in MIPS. Particles at the interface tend to align against the dense phase<sup>43</sup> if they interact purely repulsively (cf. Fig. 4(c)), and they align against the dilute phase if interparticle attraction is present<sup>44</sup>. No such spontaneous polarization occurs in bulk. The interface polarization is a state function of the coexisting phases<sup>53</sup> as verified both experimentally<sup>66</sup> and numerically<sup>67</sup>. The total swim force that acts on  $V$  is  $(p_{\text{swim}}^g - p_{\text{swim}}^l) \mathbf{A} \mathbf{e}$ , where  $p_{\text{swim}}^b = \gamma s J_b / (2D_{\text{rot}})$ , with  $J_b$  the bulk current in the forward direction  $\omega$  and  $D_{\text{rot}}$  indicating the rotational diffusion constant. Apart from the ideal term no further forces act, cf. the force density balance (17). The integral over the ideal term is similar to the total equilibrium ideal contribution,  $(p_{\text{id}}^g - p_{\text{id}}^l) \mathbf{A} \mathbf{e}$ . Combination of all results from integration yields the total force balance.

As the negative integral over the force density defines the nonequilibrium pressure, the volume  $V$  being force free amounts to  $p_{\text{tot}}^g = p_{\text{tot}}^l$ , where  $p_{\text{tot}} = p_{\text{id}} + p_{\text{int}} + p_{\text{swim}}$ . Hermann et al.<sup>52</sup>



**Fig. 5** Illustration of sedimentation of active Brownian particles under gravity  $g$ . The active particles with orientation  $\omega$  (black arrows) are confined by a lower wall and periodic boundary conditions on the sides (dashed lines). The total force that the swimming particles exert on the bottom wall (left scale pan) is equal to their weight (right scale pan) in steady states of the system.

demonstrate the splitting of  $p_{\text{int}}$  into adiabatic and superadiabatic contributions and present results for the phase diagram based on approximate forms for the interparticle interaction contributions. The pressure, especially its swim contribution is defined in various different ways in the literature<sup>68–73</sup>.

We conclude that the internal interactions that occur across the free interface do not influence the (bulk) balance of the pressure at phase coexistence, as the net effect of these interactions vanishes. In essence this argument follows from Noether's theorem for invariance against spatial displacements.

**Active sedimentation.** Sedimentation under the influence of gravity is a ubiquitous phenomenon in soft matter that has attracted considerable interest, e.g. for colloidal mixtures<sup>74–77</sup> and for active systems<sup>78–81</sup>. As sedimentation is a force driven phenomenon the Noether sum rules apply directly, as we show in the following.

We assume that the system is translationally invariant in the  $x$ -direction and that an impenetrable wall at  $z=0$  acts as a lower boundary of the system<sup>82</sup> (cf. Fig. 5). We assume the wall-particle interaction potential to be short-ranged. Its precise form is irrelevant for the following considerations. The force density balance for such a system is (17) with the external force field chosen as

$$\mathbf{f}_{\text{ext}}(\mathbf{r}, \omega) = -mg\mathbf{e}_z + \mathbf{f}_{\text{wall}}(\mathbf{r}, \omega), \quad (18)$$

where  $m$  denotes the mass of a particle,  $g$  is the gravitational acceleration and  $\mathbf{e}_z$  indicates the unit vector in  $z$ -direction. Hence the external force field  $\mathbf{f}_{\text{ext}}$  consists of gravity and the wall contribution  $\mathbf{f}_{\text{wall}}$ .

To proceed we integrate the force density balance over all positions  $\mathbf{r}$  in the volume  $V$  and over all orientations  $\omega$ . The total integral of the density distribution (per radian)  $\int d\mathbf{r}d\omega\rho$ , as appears in the gravitational term, gives the total number of particles  $N$ . The integral over the total current  $\int d\mathbf{r}d\omega\mathbf{J}$  is proportional to the center of mass velocity  $\mathbf{v}_{\text{cm}}(t) = \int d\mathbf{r}d\omega\rho\mathbf{v} / \int d\mathbf{r}d\omega\rho = \frac{1}{N} \int d\mathbf{r}d\omega\mathbf{J}$ . Here  $\mathbf{v}_{\text{cm}}(t)$  is a global quantity and hence it is independent of both position and orientation. We first only consider steady states, so the center of mass velocity  $\mathbf{v}_{\text{cm}}(t)=0$  and hence the total current vanishes. The total thermal diffusion term vanishes because  $\int d\mathbf{r}\nabla\rho = \partial_V dS\rho = 0$  as there is no contribution of  $\rho$  from the boundaries  $\partial V$  of the integration volume  $V$ . At the upper and lower boundary the density is zero as it vanishes in the wall and also for  $z \rightarrow \infty$ . The left and the right boundary contributions cancel each other as the density is independent of  $x$  due to translational invariance. The total internal interaction force density vanishes,  $\mathbf{F}_{\text{int}}^{\text{tot}} = \int d\mathbf{r}d\omega\rho\mathbf{f}_{\text{int}} = 0$ , using the global

internal Noether sum rule (10). The integrated swim force density is proportional to the total polarization  $\mathbf{M}_{\text{tot}} = \int d\mathbf{r}d\omega\omega\rho$ . This quantity vanishes,  $\mathbf{M}_{\text{tot}}=0$ , as there is no net flux through the boundaries in steady state (see Eq. (10) by Hermann et al.<sup>53</sup>). Combination of all integrals yields the relation

$$\mathbf{F}_{\text{wall}}^{\text{tot}} = \int d\mathbf{r}d\omega\rho\mathbf{f}_{\text{wall}} = mgN\mathbf{e}_z. \quad (19)$$

Hence the  $z$ -component of the total force on the wall  $F_{\text{wall}}^{\text{tot}}$  is equal to the total gravitational force acting on all particles (see Fig. 5 for a graphical representation). Equation (19) of course also holds for passive colloids ( $s=0$ ).

Keeping the translational invariance in the  $x$ -direction, we next turn to time-dependent systems. Therefore all one-body field in (17) additionally depend on the time  $t$ . Integration of the force density balance (17) gives identical results for the thermal diffusion, the internal force density and for the gravitational contribution as in the above case of steady state. Even for the time-dependent dynamics these integrals are independent of time  $t$ . The total wall force density is given as  $\mathbf{F}_{\text{wall}}^{\text{tot}} = F_{\text{wall}}^{\text{tot}}\mathbf{e}_z$  and it only acts along the unit vector in the  $z$ -direction  $\mathbf{e}_z$ , due to the symmetry of the system. The integral of the self-propulsion term is still proportional to the total polarization. However, the total polarization does not vanish in general but decays exponentially (see Eq. (21) by Hermann et al.<sup>53</sup>),

$$\mathbf{M}_{\text{tot}}(t) = \mathbf{M}_{\text{tot}}(0)e^{-D_{\text{rot}}t}, \quad (20)$$

where  $\mathbf{M}_{\text{tot}}(0)$  indicates the initial polarization at time  $t=0$  and the time constant  $1/D_{\text{rot}}$  is the inverse rotational diffusion constant. Similarly, integration of the current still gives the (time-dependent) center of mass velocity  $\mathbf{v}_{\text{cm}}(t)$ .

Insertion of these results into the spatial and orientational integration of (17) leads to the total friction force

$$\gamma N\mathbf{v}_{\text{cm}}(t) = s\gamma\mathbf{M}_{\text{tot}}(0)e^{-D_{\text{rot}}t} + F_{\text{wall}}^{\text{tot}}(t)\mathbf{e}_z - mgN\mathbf{e}_z, \quad (21)$$

which is hence a direct consequence of the Noether sum rule (10). We find that the  $x$ -component of the center of mass velocity decays simultaneously with the total polarization, cf. the first term on the right hand side of (21). The  $z$ -component of  $\mathbf{v}_{\text{cm}}(t)$  depends on  $\mathbf{M}_{\text{tot}}(t)$  and additionally on the time-dependent total force exerted by the wall and the total gravitational force. Hence measuring the total force on the wall (i.e. by weighing, cf. Fig. 5) and knowledge of the total initial polarization and the total particle number allows one to determine the center of mass velocity. Note that in the limit of long times,  $t \rightarrow \infty$ , the total polarization vanishes and this system evolves to a steady state. Hence the center of mass velocity vanishes and (19) is recovered. As we have demonstrated both statements (19) and (21) ultimately follow from the global Noether identity (10).

**Rotational invariance.** We return to the general case and initially consider spatial rotations in systems of spheres, i.e. systems where  $u(\mathbf{r}^N)$  depends solely on (relative) particle positions, and where it is invariant under global rotation of all  $\mathbf{r}^N$  around the origin. We parameterize the rotation by a vector  $\mathbf{n}$ . The direction of  $\mathbf{n}$  indicates the rotation axis and the modulus  $|\mathbf{n}|$  is the angle of rotation. To lowest nonvanishing order, the rotation amounts to  $\mathbf{r} \rightarrow \mathbf{r} + \mathbf{n} \times \mathbf{r}$ . One-body functions change accordingly: the external potential undergoes  $V_{\text{ext}}(\mathbf{r}) \rightarrow V_{\text{ext}}(\mathbf{r}) + \delta V_{\text{ext}}(\mathbf{r})$ , with  $\delta V_{\text{ext}}(\mathbf{r}) = (\mathbf{n} \times \mathbf{r}) \cdot \nabla V_{\text{ext}}(\mathbf{r})$  and the density profile  $\rho(\mathbf{r}) \rightarrow \rho(\mathbf{r}) + \delta\rho(\mathbf{r})$  with  $\delta\rho(\mathbf{r}) = (\mathbf{n} \times \mathbf{r}) \cdot \nabla\rho(\mathbf{r})$ . Much of the reasoning of the above case of spatial displacement can be applied readily:  $\Omega[V_{\text{ext}}]$  is invariant under the rotation, and  $\delta\Omega = \int d\mathbf{r}(\delta\Omega/\delta V_{\text{ext}}(\mathbf{r}))\delta V_{\text{ext}}(\mathbf{r}) = \int d\mathbf{r}\rho(\mathbf{r})(\mathbf{n} \times \mathbf{r}) \cdot \nabla V_{\text{ext}}(\mathbf{r}) = 0$ . As the rotation vector  $\mathbf{n}$  is arbitrary, we can conclude that the total external torque  $\mathcal{T}_{\text{ext}}^{\text{tot}}$

vanishes in equilibrium<sup>13</sup>,

$$\mathcal{T}_{\text{ext}}^{\text{tot}} \equiv - \int d\mathbf{r} \rho(\mathbf{r}) (\mathbf{r} \times \nabla V_{\text{ext}}(\mathbf{r})) = 0. \quad (22)$$

As this holds true for any form of the applied  $V_{\text{ext}}(\mathbf{r})$ , we can differentiate with respect to  $V_{\text{ext}}(\mathbf{r}')$ , and obtain<sup>13</sup>

$$\mathbf{r} \times \nabla \rho(\mathbf{r}) = - \int d\mathbf{r}' \beta H_2(\mathbf{r}, \mathbf{r}') (\mathbf{r}' \times \nabla' V_{\text{ext}}(\mathbf{r}')), \quad (23)$$

$$\sum_{\alpha=1}^n (\mathbf{r}_{\alpha} \times \nabla_{\alpha} H_n) = - \int d\mathbf{r}_{n+1} \beta V_{\text{ext}}(\mathbf{r}_{n+1}) (\mathbf{r}_{n+1} \times \nabla_{n+1} H_{n+1}). \quad (24)$$

The excess free energy density functional  $F_{\text{exc}}[\rho]$  can be treated accordingly. It is invariant under rotation, as its sole dependence is on  $u(\mathbf{r}^N)$ , which by assumption is rotationally invariant. Analogous to this reasoning, we obtain the result that the total interparticle adiabatic torque  $\mathcal{T}_{\text{ad}}^{\text{tot}}$  vanishes,

$$\mathcal{T}_{\text{ad}}^{\text{tot}} = \int d\mathbf{r} \rho(\mathbf{r}) (\mathbf{r} \times \mathbf{f}_{\text{ad}}(\mathbf{r})) = 0. \quad (25)$$

Differentiation with respect to the independent field  $\rho(\mathbf{r})$  once and  $n$  times yields the respective identities<sup>13,14</sup>:

$$\mathbf{r} \times \nabla c_1(\mathbf{r}) = \int d\mathbf{r}' c_2(\mathbf{r}, \mathbf{r}') (\mathbf{r}' \times \nabla' \rho(\mathbf{r}')), \quad (26)$$

$$\sum_{\alpha=1}^n (\mathbf{r}_{\alpha} \times \nabla_{\alpha} c_n) = - \int d\mathbf{r}_{n+1} \rho(\mathbf{r}_{n+1}) (\mathbf{r}_{n+1} \times \nabla_{n+1} c_{n+1}). \quad (27)$$

The multi-body versions of the theorems of vanishing total external (22) and adiabatic internal (25) torques are, respectively,

$$\int d\mathbf{r}_1 V_{\text{ext}}(\mathbf{r}_1) \dots \int d\mathbf{r}_n V_{\text{ext}}(\mathbf{r}_n) (\mathbf{r}_{\alpha} \times \nabla_{\alpha} H_n) = 0, \quad (28)$$

and

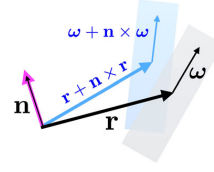
$$\int d\mathbf{r}_1 \rho(\mathbf{r}_1) \dots \int d\mathbf{r}_n \rho(\mathbf{r}_n) (\mathbf{r}_{\alpha} \times \nabla_{\alpha} c_n) = 0, \quad (29)$$

for  $\alpha = 1 \dots n$ . These identities are respectively derived from (23) by multiplying with  $V_{\text{ext}}(\mathbf{r})$ , integrating over  $\mathbf{r}$  and exploiting (22), and from (26) by multiplying with  $\rho(\mathbf{r})$ , integrating over  $\mathbf{r}$  and exploiting (25), and iteratively repeating for each order  $n$ .

On the many-body level, it is straightforward to see that the total internal torque  $-\sum_i (\mathbf{r}_i \times \nabla_i u(\mathbf{r}^N)) = 0$ . Hence, as this identity holds for each microstate, its general, nonequilibrium average vanishes,  $\mathcal{T}_{\text{int}}^{\text{tot}} = 0$ . The force field splitting  $\mathbf{f}_{\text{int}} = \mathbf{f}_{\text{ad}} + \mathbf{f}_{\text{sup}}$  induces corresponding additive structure for the internal total torque:  $\mathcal{T}_{\text{int}}^{\text{tot}} = \mathcal{T}_{\text{ad}}^{\text{tot}} + \mathcal{T}_{\text{sup}}^{\text{tot}}$ , with the total superadiabatic (internal) torque  $\mathcal{T}_{\text{sup}}^{\text{tot}} = \int d\mathbf{r} \rho(\mathbf{r}) [\mathbf{r} \times \mathbf{f}_{\text{sup}}(\mathbf{r}, t)]$ . As  $\mathcal{T}_{\text{int}}^{\text{tot}} = \mathcal{T}_{\text{ad}}^{\text{tot}} = 0$ , we conclude  $\mathcal{T}_{\text{sup}}^{\text{tot}} = 0, \forall t$ .

To apply Noether's theorem to the power functional, we consider an instantaneous rotation, with infinitesimal angular velocity  $\dot{\mathbf{n}}$  at time  $t$ . The effect is a change in current  $\mathbf{J} \rightarrow \mathbf{J} + \delta\mathbf{J}$  with  $\delta\mathbf{J} = (\dot{\mathbf{n}} \times \mathbf{r}) \rho(\mathbf{r})$ . Correspondingly, the velocity field acquires an instantaneous global rotational contribution, according to  $\mathbf{v}(\mathbf{r}, t) \rightarrow \mathbf{v}(\mathbf{r}, t) + \dot{\mathbf{n}} \times \mathbf{r}$ . The superadiabatic excess power functional is invariant under this operation and hence  $P_t^{\text{exc}}[\rho, \mathbf{J}] = P_t^{\text{exc}}[\rho, \mathbf{J} + \delta\mathbf{J}] = P_t^{\text{exc}}[\rho, \mathbf{J}] + \int d\mathbf{r} (\delta P_t^{\text{exc}}[\rho, \mathbf{J}] / \delta \mathbf{J})(\mathbf{r}, t) \cdot (\dot{\mathbf{n}} \times \mathbf{r}) \rho(\mathbf{r}, t)$ . As  $\dot{\mathbf{n}}$  is arbitrary, we can conclude

$$\mathcal{T}_{\text{sup}}^{\text{tot}} = \int d\mathbf{r} \rho(\mathbf{r}, t) (\mathbf{r} \times \mathbf{f}_{\text{sup}}(\mathbf{r}, t)) = 0, \quad \forall t, \quad (30)$$



**Fig. 6 Illustration of the rotation operation of uniaxial particles.** The particles (rectangular shapes) at position  $\mathbf{r}$  and with orientation  $\boldsymbol{\omega}$  are shown in original (black) and rotated (blue) configuration, where  $\mathbf{n}$  indicates the rotation axis (direction) and angle (length).

as is consistent with the result of the above many-body derivation. As (30) holds for any (trial)  $\mathbf{J}(\mathbf{r}, t)$ , the derivative of (30) with respect to  $\mathbf{J}(\mathbf{r}, t)$  vanishes. Hence

$$\int d\mathbf{r} \rho(\mathbf{r}, t) (\mathbf{r} \times \mathbf{M}_2(\mathbf{r}, \mathbf{r}', t)) = 0, \quad (31)$$

where the cross product with a tensor is defined via contraction with the Levi-Civita tensor. At  $n$ th order we obtain

$$\int d\mathbf{r}_n \rho(\mathbf{r}_n, t) (\mathbf{r}_n \times \mathbf{M}_n) = 0. \quad (32)$$

This identity and (31) express the vanishing of the total superadiabatic torque, when resolved on the  $n$ -body level of (time direct) correlation functions.

**Orbital and spin coupling.** The case of rotational symmetry of uniaxial particles is clearly more complex, as both particle coordinates and particle orientations are affected by a global (“rigid”) operation on the entire system, i.e. both positions and orientations are rotated consistently. Noether's theorem ensures though that this operation is indeed the fundamental one, and that the physically expected coupling of orbital and spinning effects will naturally and systematically emerge. Here we use (common) terminology for referring to spin as orientation vector rotation (i.e. particle rotation around its center), as opposed to orbital rotation (of position vector) around the origin of position space. Hence all the above considered torques that already occur in systems of spheres are of orbital nature. These of course remain relevant for anisotropic particles, but the nontrivial orientational behavior of the latter will generate additional spin torques. The global rotation consists of an orbital part,  $\mathbf{r} \rightarrow \mathbf{r} + \mathbf{n} \times \mathbf{r}$ , and a spin part,  $\boldsymbol{\omega} \rightarrow \boldsymbol{\omega} + \mathbf{n} \times \boldsymbol{\omega}$ ; see Fig. 6 for a graphical representation.

For anisotropic systems the external field naturally acquires dependence on position  $\mathbf{r}$  and orientation  $\boldsymbol{\omega}$ , i.e.  $V_{\text{ext}}(\mathbf{r}, \boldsymbol{\omega})$ , where  $-\nabla V_{\text{ext}}(\mathbf{r}, \boldsymbol{\omega})$  is the external force field as before, and  $-\boldsymbol{\omega} \times \nabla^{\boldsymbol{\omega}} V_{\text{ext}}(\mathbf{r}, \boldsymbol{\omega})$  is the external torque field, where  $\nabla^{\boldsymbol{\omega}}$  is the derivative with respect to  $\boldsymbol{\omega}$  in orientation space. Hence the induced change of external potential is  $\delta V_{\text{ext}}(\mathbf{r}, \boldsymbol{\omega}) = (\mathbf{n} \times \mathbf{r}) \cdot \nabla V_{\text{ext}} + (\mathbf{n} \times \boldsymbol{\omega}) \cdot \nabla^{\boldsymbol{\omega}} V_{\text{ext}}$ . This change leaves the grand potential  $\Omega[V_{\text{ext}}]$  invariant, hence  $\delta\Omega = \int d\mathbf{r} d\boldsymbol{\omega} (\delta\Omega / \delta V_{\text{ext}}(\mathbf{r}, \boldsymbol{\omega})) \mathbf{n} \cdot (\mathbf{r} \times \nabla V_{\text{ext}} + \boldsymbol{\omega} \times \nabla^{\boldsymbol{\omega}} V_{\text{ext}}) = 0$ , from which we identify the rotational Noether theorem for the total external torque of anisotropic particles:

$$\mathcal{T}_{\text{ext}}^{\text{tot}} = - \int d\mathbf{r} d\boldsymbol{\omega} \rho(\mathbf{r}, \boldsymbol{\omega}) (\mathbf{r} \times \nabla V_{\text{ext}} + \boldsymbol{\omega} \times \nabla^{\boldsymbol{\omega}} V_{\text{ext}}) = 0. \quad (33)$$

Differentiation with respect to  $V_{\text{ext}}(\mathbf{r}, \boldsymbol{\omega})$  yields

$$\mathbf{r} \times \nabla \rho + \boldsymbol{\omega} \times \nabla^\omega \rho = - \int d\mathbf{r}' d\boldsymbol{\omega}' \beta H_2(\mathbf{r}, \boldsymbol{\omega}, \mathbf{r}', \boldsymbol{\omega}') (\mathbf{r}' \times \nabla' V'_{\text{ext}} + \boldsymbol{\omega}' \times \nabla'^\omega V'_{\text{ext}}), \quad (34)$$

$$\begin{aligned} & \sum_{\alpha=1}^n (\mathbf{r}_\alpha \times \nabla_\alpha H_n + \boldsymbol{\omega}_\alpha \times \nabla_\alpha^\omega H_n) \\ &= - \int d\mathbf{r}' d\boldsymbol{\omega}' \beta V'_{\text{ext}} (\mathbf{r}' \times \nabla' H_{n+1} + \boldsymbol{\omega}' \times \nabla'^\omega H_{n+1}), \end{aligned} \quad (35)$$

where we have used  $V'_{\text{ext}} = V_{\text{ext}}(\mathbf{r}', \boldsymbol{\omega}')$  as a shorthand,  $H_2(\mathbf{r}, \boldsymbol{\omega}, \mathbf{r}', \boldsymbol{\omega}') = -\delta\rho(\mathbf{r}, \boldsymbol{\omega})/\delta\beta V_{\text{ext}}(\mathbf{r}', \boldsymbol{\omega}')$  is the density-density correlation function, and its  $n$ -body version  $H_n = H_n(\mathbf{r}_1, \boldsymbol{\omega}_1, \dots, \mathbf{r}_n, \boldsymbol{\omega}_n)$  with  $H_{n+1} = \delta H_n/\delta\beta V_{\text{ext}}(\mathbf{r}_{n+1}, \boldsymbol{\omega}_{n+1})$ ; in (35) the prime refers to the  $n+1$ th degrees of freedom. On all levels of  $n$ -body correlation functions, the spin and orbital torques remain coupled.

Turning to internal torques, the change of density upon global rotation is  $\delta\rho(\mathbf{r}, \boldsymbol{\omega}) = \mathbf{n} \cdot (\mathbf{r} \times \nabla \rho + \boldsymbol{\omega} \times \nabla^\omega \rho)$  and the net effect on the excess free energy is  $\delta F_{\text{exc}} = \int d\mathbf{r} d\boldsymbol{\omega} (\delta F_{\text{exc}}[\rho]/\delta\rho(\mathbf{r}, \boldsymbol{\omega})) \mathbf{n} \cdot (\mathbf{r} \times \nabla \rho + \boldsymbol{\omega} \times \nabla^\omega \rho) = 0$ . We hence obtain

$$\mathcal{T}_{\text{ad}}^{\text{tot}} = \int d\mathbf{r} d\boldsymbol{\omega} \rho(\mathbf{r}, \boldsymbol{\omega}) (\mathbf{r} \times \mathbf{f}_{\text{ad}} + \boldsymbol{\tau}_{\text{ad}}) = 0, \quad (36)$$

where the adiabatic spin torque field is  $\boldsymbol{\tau}_{\text{ad}}(\mathbf{r}, \boldsymbol{\omega}) = -\boldsymbol{\omega} \times \nabla^\omega \delta F_{\text{exc}}[\rho]/\delta\rho(\mathbf{r}, \boldsymbol{\omega})$ . From differentiation with respect to  $\rho(\mathbf{r}, \boldsymbol{\omega})$  we obtain

$$\mathbf{r} \times \mathbf{f}_{\text{ad}} + \boldsymbol{\tau}_{\text{ad}} = \int d\mathbf{r}' d\boldsymbol{\omega}' c_2(\mathbf{r}, \boldsymbol{\omega}, \mathbf{r}', \boldsymbol{\omega}') (\mathbf{r}' \times \nabla' \rho' + \boldsymbol{\omega}' \times \nabla'^\omega \rho'), \quad (37)$$

$$\begin{aligned} & \sum_{\alpha=1}^n (\mathbf{r}_\alpha \times \nabla_\alpha c_n + \boldsymbol{\omega}_\alpha \times \nabla_\alpha^\omega c_n) \\ &= - \int d\mathbf{r}' d\boldsymbol{\omega}' \rho' (\mathbf{r}' \times \nabla' c_{n+1} + \boldsymbol{\omega}' \times \nabla'^\omega c_{n+1}), \end{aligned} \quad (38)$$

where  $\rho' = \rho(\mathbf{r}', \boldsymbol{\omega}')$ . Multi-body versions of (33) and (36) read as

$$\int d1 V_{\text{ext}}(1) \dots \int dn V_{\text{ext}}(n) (\mathbf{r}_\alpha \times \nabla_\alpha + \boldsymbol{\omega}_\alpha \times \nabla_\alpha^\omega) H_n = 0 \quad (39)$$

and

$$\int d1 \rho(1) \dots \int dn \rho(n) (\mathbf{r}_\alpha \times \nabla_\alpha + \boldsymbol{\omega}_\alpha \times \nabla_\alpha^\omega) c_n = 0. \quad (40)$$

Here we have used the shorthand notation  $1 \equiv \mathbf{r}_1, \boldsymbol{\omega}_1$  etc., and the derivation is analogous to the above rotational case of spherical particles.

The two-body sum rules (34) and (37) are identical to those obtained by Tarazona and Evans<sup>16</sup> using rotational invariance arguments applied directly to correlation functions. Our methodology not only allows to naturally re-derive their results, but also to identify the full gamut of adiabatic rotational sum rules, from the global statements (33) and (36) to the infinite hierarchies (35) and (38) (we use notational convention different from Tarazona and Evans<sup>16</sup>: our  $\nabla^\omega$  is (only) a partial derivative with respect to  $\boldsymbol{\omega}$ , i.e.  $\nabla^\omega \equiv \partial/\partial\boldsymbol{\omega}$ , whereas their  $\nabla_\omega \equiv \boldsymbol{\omega} \times \partial/\partial\boldsymbol{\omega}$ . The modulus is fixed,  $|\boldsymbol{\omega}| = 1$ , in both versions. Tarazona and Evans<sup>16</sup> notate  $H_2$  as  $G$  in their (18) and (19)).

Again for each microstate  $\sum_i (\mathbf{r}_i \times \nabla_i u + \boldsymbol{\omega}_i \times \nabla_i^\omega u) = 0$  and hence on average  $\mathcal{T}_{\text{int}}^{\text{tot}} = 0$ . From the splitting  $\mathcal{T}_{\text{int}}^{\text{tot}} = \mathcal{T}_{\text{ad}}^{\text{tot}} + \mathcal{T}_{\text{sup}}^{\text{tot}}$ , we conclude  $\mathcal{T}_{\text{sup}}^{\text{tot}} = 0$ . From rotational invariance of  $P_i^{\text{exc}}[\rho, \mathbf{J}, \mathbf{J}^\omega]$ , where  $\mathbf{J}^\omega(\mathbf{r}, \boldsymbol{\omega}, t)$  is the rotational current<sup>50</sup>, against an instantaneous angular “kick”  $\delta\mathbf{J} = (\hat{\mathbf{n}} \times \mathbf{r})\rho$

and  $\delta\mathbf{J}^\omega = (\hat{\mathbf{n}} \times \boldsymbol{\omega})\rho$ , we find

$$\mathcal{T}_{\text{sup}}^{\text{tot}} = - \int d\mathbf{r} d\boldsymbol{\omega} \rho(\mathbf{r}, \boldsymbol{\omega}, t) \left( \mathbf{r} \times \frac{\delta P_i^{\text{exc}}}{\delta\mathbf{J}} + \boldsymbol{\omega} \times \frac{\delta P_i^{\text{exc}}}{\delta\mathbf{J}^\omega} \right) = 0. \quad (41)$$

Local sum rules can be obtained straightforwardly by building derivatives with respect to  $\mathbf{J}(\mathbf{r}, \boldsymbol{\omega}, t)$  and  $\mathbf{J}^\omega(\mathbf{r}, \boldsymbol{\omega}, t)$ . The result is:

$$\int d\mathbf{r}_1 d\boldsymbol{\omega}_1 \rho(1) [\mathbf{r}_1 \times \delta/\delta\mathbf{J}(1) + \boldsymbol{\omega}_1 \times \delta/\delta\mathbf{J}^\omega(1)] M_{n,m} = 0. \quad (42)$$

Here the tensorial equal-time direct correlation functions are defined as  $M_{n,m} = -\beta \delta^{n+m} P_i^{\text{exc}}/\delta\mathbf{J}(1) \dots \delta\mathbf{J}(n) \delta\mathbf{J}^\omega(n+1) \dots \mathbf{J}^\omega(n+m)$ , where the roman numerals refer to position, orientation and time  $t$  (no index), e.g.  $1 \equiv \mathbf{r}_1, \boldsymbol{\omega}_1, t$ . We have assumed that all functional derivatives commute.

So far we have restricted ourselves to uniaxial particles. The derived rotational sum rules can be analogously determined for general anisotropic particles with an additional orientation vector  $\hat{\boldsymbol{\omega}}$ . This can be done by simply replacing  $\boldsymbol{\omega} \times \nabla^\omega \rightarrow \boldsymbol{\omega} \times \nabla^\omega + \hat{\boldsymbol{\omega}} \times \nabla^{\hat{\boldsymbol{\omega}}}$ ,  $\boldsymbol{\omega} \times \partial/\partial\mathbf{J}^\omega \rightarrow \boldsymbol{\omega} \times \partial/\partial\mathbf{J}^\omega + \hat{\boldsymbol{\omega}} \times \partial/\partial\mathbf{J}^\omega$  and  $\int d\mathbf{r} d\boldsymbol{\omega} \rightarrow \int d\mathbf{r} d\boldsymbol{\omega} d\hat{\boldsymbol{\omega}}$  in (33)–(42). This replacement also affects the adiabatic spin torque field  $\boldsymbol{\tau}_{\text{ad}}$  and all one-body field depend additionally to  $\mathbf{r}$  and  $\boldsymbol{\omega}$  on the orientation  $\hat{\boldsymbol{\omega}}$ . For anisotropic particles there exists a more general version of (42) by exchange of  $M_{n,m} \rightarrow M_{n,m,l}$  where  $l$  denotes the number of functional derivations with respect to the rotational current  $\mathbf{J}^\omega$ . The indices  $n$  and  $m$  belong to the number of functional derivatives with respect to  $\mathbf{J}$  and  $\mathbf{J}^\omega$  as before.

**Memory invariance.** In the above treatment of nonequilibrium situations we have exploited invariance against an instantaneous transformation applied to the system. As we have shown, the corresponding Noether identities carry imminent physical meaning. These sum rules hold for the nonequilibrium effects that arise from the interparticle interaction, i.e. they constrain superadiabatic forces and torques, as obtained from translation and rotation, respectively. Here we exploit that the corresponding nonequilibrium functional generator  $P_i^{\text{exc}}$  carries further invariances, once one allows the transformation to act also on the history of the system. As we demonstrate in the following, the resulting identities constitute exact constraints on the memory structure that are induced by the coupled interparticle interactions. Recall that a reduced one-body description of a many-body system is generically non-Markovian (i.e. nonlocal in time)<sup>83</sup>. The study of memory kernels, often carried out in the framework of generalized Langevin equations, is a topic of significant current research activity<sup>84–91</sup>.

Our approach differs from these efforts in that no a priori generic form of a reduced equation of motion is assumed. Rather our considerations are formally exact and interrelate (and hence constrain) time correlation functions, which are generated from the central nonequilibrium object  $P_i^{\text{exc}}$  via functional differentiation. Very little is known about the memory structure of superadiabatic forces, with exceptions being the NOZ framework<sup>57,58</sup> and the demonstration of the relevance of memory for the observed viscoelasticity of hard sphere liquids<sup>92</sup>. Both the Ornstein–Zernike (OZ) and NOZ relations are different from the Noether identities. The former relations are a direct consequence of the generality of the variational principle. Per se, neither the OZ nor the NOZ relations reflect the Noether symmetries.

Recalling the illustrated overview of the different types of shifting in Fig. 1, in the following we treat two further types of invariance transformations: One is the static transformation. This operation is formally analogous to the above equilibrium treatment of the adiabatic state, but it is here carried out in the

same way at all times. This static transformation contrasts (and complements) the instantaneous transformation used above for the time-dependent case. The corresponding changes to density and current are graphically illustrated in Fig. 3b. The second invariance operation is that of memory shifting, where the transformation parameter is taken to be time-dependent, cf. Fig. 3c for a graphical representation. For simplicity we restrict ourselves to cases where at both ends of the considered time interval, no shifting occurs (i.e. such that the transformation is the identity at the limiting times).

The static spatial shift consists of  $\rho(\mathbf{r}, t') \rightarrow \rho(\mathbf{r} + \boldsymbol{\epsilon}, t')$  and  $\mathbf{J}(\mathbf{r}, t') \rightarrow \mathbf{J}(\mathbf{r} + \boldsymbol{\epsilon}, t')$ , where the time argument  $t'$  is arbitrary, and  $\boldsymbol{\epsilon} = \text{const}$  characterizes magnitude and direction of the translation, see the illustration in Fig. 3b. Hence the time derivative  $\dot{\boldsymbol{\epsilon}} = 0$  such that the current does not acquire any displacement contribution, as also  $-\dot{\boldsymbol{\epsilon}}\rho = 0$  at all times. The spatial shift applies to all times  $t'$  considered, and we can restrict ourselves to  $0 \leq t' \leq t$ . Hence the changes in kinematic fields are to first order given by  $\delta\rho(\mathbf{r}, t') = \boldsymbol{\epsilon} \cdot \nabla\rho(\mathbf{r}, t')$  and  $\delta\mathbf{J}(\mathbf{r}, t') = \boldsymbol{\epsilon} \cdot \nabla\mathbf{J}(\mathbf{r}, t')$ . As  $P_t^{\text{exc}}$  originates solely from the interparticle interaction potential, the invariance of  $u(\mathbf{r}^N)$  against the global displacement at all times induces invariance of  $P_t^{\text{exc}}$ . Hence  $P_t^{\text{exc}}[\rho, \mathbf{J}] = P_t^{\text{exc}}[\rho + \delta\rho, \mathbf{J} + \delta\mathbf{J}] = P_t^{\text{exc}}[\rho, \mathbf{J}] + \delta P_t^{\text{exc}}$ . Here  $\delta P_t^{\text{exc}}$  indicates the change in superadiabatic free power and due to the invariance  $\delta P_t^{\text{exc}} = 0$ . On the other hand we can express  $\delta P_t^{\text{exc}}$  via the functional Taylor expansion. To linear order the result consists of two integrals. One integral comes from the time-slice functional derivative at fixed (end) time  $t$  and is given by  $\int d\mathbf{r}[(\delta P_t^{\text{exc}}/\delta\rho)\delta\rho + (\delta P_t^{\text{exc}}/\delta\mathbf{J}) \cdot \delta\mathbf{J}]$ . The second integral is from a functional derivative at (variable) time  $t'$ , given by  $\int_0^t dt' \int d\mathbf{r}'[(\delta P_t^{\text{exc}}/\delta\rho')\delta\rho' + (\delta P_t^{\text{exc}}/\delta\mathbf{J}') \cdot \delta\mathbf{J}']$ , where the prime indicates dependence on arguments  $\mathbf{r}'$  and  $t'$ . We then exploit that the displacement  $\boldsymbol{\epsilon}$  of the static shift (which parametrizes the changes in density and in current) is arbitrary. The result is a global nonequilibrium Noether theorem, given by

$$\int d\mathbf{r} \left( \frac{\delta P_t^{\text{exc}}}{\delta\rho} \nabla\rho - \mathbf{f}_{\text{sup}} \cdot \nabla\mathbf{J}^\top \right) + \int_0^t dt' \int d\mathbf{r}' \left( \frac{\delta P_t^{\text{exc}}}{\delta\rho'} \nabla'\rho' + \frac{\delta P_t^{\text{exc}}}{\delta\mathbf{J}'} \cdot \nabla'\mathbf{J}'^\top \right) = 0, \tag{43}$$

where we have used the relationship of the superadiabatic force field to its generator,  $\mathbf{f}_{\text{sup}}(\mathbf{r}, t) = -\delta P_t^{\text{exc}}[\rho, \mathbf{J}]/\delta\mathbf{J}(\mathbf{r}, t)$ , the primed symbol  $\nabla'$  indicates the derivative with respect to  $\mathbf{r}'$ , and the superscript  $\top$  denotes the matrix transpose (in index notation the  $k$ -component of the vector  $\mathbf{a} \cdot \nabla\mathbf{b}^\top$  is  $\sum_k a_k b_k$ ).

Equation (43) constitutes a global identity that links density, current, and superadiabatic force field in a nontrivial spatial and temporal form. As the central variation principle<sup>48</sup> allows to vary  $\mathbf{J}(\mathbf{r}, t)$  freely, (43) remains true upon building the functional derivative with respect to  $\mathbf{J}(\mathbf{r}, t)$ . The result is a local identity

$$\beta\nabla\mathbf{f}_{\text{sup}} = \int d\mathbf{r}'(\mathbf{m}_2(\mathbf{r}, \mathbf{r}', t)\nabla'\rho(\mathbf{r}', t) + \mathbf{M}_2(\mathbf{r}, \mathbf{r}', t) \cdot \nabla'\mathbf{J}'^\top(\mathbf{r}', t)) + \int_0^t dt' \int d\mathbf{r}'(\mathbf{m}_2(\mathbf{r}, t, \mathbf{r}', t')\nabla'\rho' + \mathbf{M}_2(\mathbf{r}, t, \mathbf{r}', t') \cdot \nabla'\mathbf{J}'^\top), \tag{44}$$

where two-body time direct correlation functions occur in vectorial form:  $\mathbf{m}_2(\mathbf{r}, \mathbf{r}', t) = -\beta\delta^2 P_t^{\text{exc}}/\delta\mathbf{J}(\mathbf{r}, t)\delta\rho(\mathbf{r}', t)$ ,  $\mathbf{m}_2(\mathbf{r}, t, \mathbf{r}', t') = -\beta\delta^2 P_t^{\text{exc}}/\delta\mathbf{J}(\mathbf{r}, t)\delta\rho(\mathbf{r}', t')$ , as well as in tensorial form:  $\mathbf{M}_2(\mathbf{r}, \mathbf{r}', t) = -\beta\delta^2 P_t^{\text{exc}}/\delta\mathbf{J}(\mathbf{r}, t)\delta\mathbf{J}(\mathbf{r}', t)$ ,  $\mathbf{M}_2(\mathbf{r}, t, \mathbf{r}', t') = -\beta\delta^2 P_t^{\text{exc}}/\delta\mathbf{J}(\mathbf{r}, t)\delta\mathbf{J}(\mathbf{r}', t')$ . Here we have made the (common) assumption that the second derivatives can be interchanged.

Repeated differentiation of (44) with respect to  $\mathbf{J}(\mathbf{r}, t)$  generates a hierarchy,

$$\sum_{a=1}^{n-1} \nabla_a \mathbf{M}_{n-1}(\mathbf{r}^{n-1}, t) = \int d\mathbf{r}_n(\mathbf{m}_n(\mathbf{r}^n, t)\nabla_n\rho(\mathbf{r}_n, t) + \mathbf{M}_n(\mathbf{r}^n, t) \cdot \nabla_n\mathbf{J}(\mathbf{r}_n, t)^\top) + \int_0^t dt' \int d\mathbf{r}_n(\mathbf{m}_n(\mathbf{r}^n, t', \mathbf{r}_n, t')\nabla_n\rho(\mathbf{r}_n, t') + \mathbf{M}_n(\mathbf{r}^n, t', \mathbf{r}_n, t') \cdot \nabla_n\mathbf{J}(\mathbf{r}_n, t')^\top), \tag{45}$$

where the  $n$ -body equal-time direct correlation functions of rank  $n$  are  $\mathbf{M}_n(\mathbf{r}^n, t) = -\beta\delta^n P_t^{\text{exc}}/\delta\mathbf{J}(\mathbf{r}_1, t)\dots\delta\mathbf{J}(\mathbf{r}_n, t)$  and  $\mathbf{m}_n(\mathbf{r}^n, t) = -\beta\delta^n P_t^{\text{exc}}/\delta\mathbf{J}(\mathbf{r}_1, t)\dots\delta\mathbf{J}(\mathbf{r}_{n-1}, t)\delta\rho(\mathbf{r}_n, t)$ , where we have used the shorthand  $\mathbf{r}^n = \mathbf{r}_1\dots\mathbf{r}_n$ . Furthermore at unequal times we have:  $\mathbf{M}_n(\mathbf{r}^{n-1}, t, \mathbf{r}_n, t') = -\beta\delta^n P_t^{\text{exc}}/\delta\mathbf{J}(\mathbf{r}_1, t)\dots\delta\mathbf{J}(\mathbf{r}_{n-1}, t)\delta\mathbf{J}(\mathbf{r}_n, t')$  as a rank  $n$  tensor, and also a rank  $n-1$  tensor  $\mathbf{m}_n(\mathbf{r}^{n-1}, t, \mathbf{r}_n, t') = -\beta\delta^n P_t^{\text{exc}}/\delta\mathbf{J}(\mathbf{r}_1, t)\dots\delta\mathbf{J}(\mathbf{r}_{n-1}, t)\delta\rho(\mathbf{r}_n, t')$ .

In the second case, we consider a more general invariance transformation that is obtained by letting the transformation parameter be time-dependent. In this case of time-dependent shifting, we prescribe a displacement vector  $\boldsymbol{\epsilon}(t')$  for times  $0 \leq t' \leq t$ , i.e. between the initial time, throughout the past and up to the "current" time  $t$ . We restrict ourselves to vanishing shift at the boundaries of the considered time interval, i.e.  $\boldsymbol{\epsilon}(0) = \boldsymbol{\epsilon}(t) = 0$ . Due to the overdamped character of the dynamics, its interparticle contributions are unaffected by this transformation, and hence  $P_t^{\text{exc}}$  is invariant. The induced changes that the density and the current acquire arise from shifting their position argument, but the current also acquires an additive shifting current contribution. The latter contribution is analogous to the (sole) effect that is present in the instantaneous shifting, but here applicable at all times (in the considered time interval).

Hence the time-dependent shifting, as illustrated in Fig. 3c, induces the following changes to the density and the current:  $\delta\rho(\mathbf{r}', t') = \boldsymbol{\epsilon}(t') \cdot \nabla'\rho(\mathbf{r}', t')$  and  $\delta\mathbf{J}(\mathbf{r}', t') = \boldsymbol{\epsilon}(t') \cdot \nabla\mathbf{J}(\mathbf{r}', t') - \dot{\boldsymbol{\epsilon}}(t')\rho(\mathbf{r}', t')$ , where  $\dot{\boldsymbol{\epsilon}}(t') = d\boldsymbol{\epsilon}(t')/dt'$ . Next we can regard  $P_t^{\text{exc}}[\rho + \delta\rho, \mathbf{J} + \delta\mathbf{J}]$  as a functional of  $\boldsymbol{\epsilon}(t')$  and  $\dot{\boldsymbol{\epsilon}}(t')$ . Its invariance amounts to stationarity, i.e. vanishing first functional derivative, with respect to the displacement. This problem, in particular for the present case of fixed boundary values, amounts to one of the most basic problems in the calculus of variations. It is realized, e.g., in the determination of catenary curves and indeed, in Hamilton's principle of classical mechanics. Exploiting the corresponding Euler-Lagrange equation leads to

$$\frac{d}{dt'} \int d\mathbf{r}' \mathbf{m}'_1 \rho' + \int d\mathbf{r}' \mathbf{m}'_1 \cdot \nabla'\mathbf{J}'^\top + \int d\mathbf{r}' \mathbf{m}'_1 \nabla'\rho' = 0, \tag{46}$$

where the one-body time direct correlation functions are  $\mathbf{m}'_1(\mathbf{r}', t', t) = -\beta\delta P_t^{\text{exc}}/\delta\mathbf{J}(\mathbf{r}', t')$  and  $m'_1(\mathbf{r}', t', t) = -\beta\delta P_t^{\text{exc}}/\delta\rho(\mathbf{r}', t')$ . Differentiation with respect to  $\mathbf{J}(\mathbf{r}, t)$  yields again a local memory identity.

**Conclusions.** We have demonstrated that Noether's theorem for exploiting symmetry in a variational context has profound implications for Statistical Physics. Known sum rules can be derived with ease and powerfully generalized to full infinite hierarchies, to the rotational case, and to time-dependence in nonequilibrium. Recall the selected applications<sup>31-41</sup> of the equilibrium sum rules, as we have laid out in the introduction. For the time-dependent case, we envisage similar insights from using the newly formulated nonequilibrium sum rules in investigations of e.g., the dynamics of freezing, of liquid crystal flow, and of driven fluid interfaces. On the conceptual level, Noether's



theorem assigns a clear meaning and physical interpretation to all resulting identities, as being generated from an invariance property of an underlying functional generator. Although the symmetry operation that we considered are simplistic, and only their lowest order in a power series expansion needs to be taken into account, a significantly complex body of sum rules naturally emerges. Hence the application of Noether's theorem is significantly deeper than mere exploiting of symmetries of arguments of correlation functions, i.e. that the direct correlation function  $c_2(\mathbf{r}, \mathbf{r}')$  in bulk fluids depends solely on  $|\mathbf{r} - \mathbf{r}'|$ . Rather, as governed by functional calculus, coupling of different levels of correlation functions occurs.

The Noether sum rules are different from the variational principle, as embodied in the Euler–Lagrange equation. On the formal level, the difference is that the Euler–Lagrange equation (both of DFT and of PFT) is a formally closed equation on the one-body level. In contrast, the Noether rules couple  $n$ - and  $(n + 1)$ -body correlation functions, hence they are of genuine hierarchical nature. They also describe different physics, as the Euler–Lagrange equation expresses a chemical potential equilibrium in DFT and the local force balance relationship in PFT. In contrast, the Noether identities stem from the symmetry properties of the respective underlying physical system.

The standard DFT approximations, ranging from simple local, square-gradient, and mean-field functional to more sophisticated weighted-density-schemes including fundamental measure theory satisfy the internal force relationships. This can be seen straightforwardly by observing that these functionals do satisfy global translation invariance (the value of the free energy is independent of the choice of coordinate origin). All higher-order Noether identities are then automatically satisfied, as these inherit the correct symmetry properties from the generating (excess free energy) functional. Our formalism hence provides a concrete reason, over mere empirical experience, why the practitioners' choices for approximate functionals are sound. The situation for more complex DFT schemes could potentially be different though. As soon as, say, self-consistency of some form is imposed, or coupling to auxiliary field comes into play, it is easy to imagine that the Noether identities help in restricting choices in the construction of such approximation schemes.

The sum rules imposed by the three types of dynamical displacements are satisfied within the velocity gradient form of the power functional<sup>54,56</sup>. It is straightforward to see that the functional is independent of the coordinate origin (static shifting). For the cases of dynamical shifting, the invariance of the functional stems from invariance of the velocity field against shifting. For both instantaneous and memory shifting, the velocity gradient remains invariant under the displacement.

We envisage that the higher than two-body Noether identities can facilitate the construction of advanced liquid state/density functional approximations. Such work should surely be highly challenging. In the context of fundamental measure theory (see e.g. the work by Roth<sup>20</sup> for an enlightening review) it is worth recalling that in Rosenfeld's original 1989 paper<sup>93</sup>, he calculated the three-body direct correlation function from his then newly proposed functional. The result for the corresponding three-body pair correlations compared favorably against simulation data. Furthermore, the recent insights into two-body correlations in inhomogeneous liquids<sup>94</sup> and crystals<sup>95</sup> demonstrates that working with higher-body correlation functions is feasible.

In future work it would be very beneficial to bring together the Noether identities with the nonequilibrium Ornstein–Zernike relations<sup>57,58</sup>, in order to aid construction of new dynamical approximations. One could exploit the rotational invariance of the superadiabatic excess power functional  $P_t^{\text{exc}}$  to gain deeper

insights in its memory structure and also generalize the translational memory relations (43)–(46) to anisotropic particles. It would be highly interesting to apply (49) to the recently obtained direct correlation function of the hard sphere crystal. This would allow to investigate whether Triezenberg and Zwanzig's concept that they originally developed for the free gas–liquid interface applies to the also self-sustained density inhomogeneity in a solid. Furthermore, addressing further cases of self motility<sup>42–44</sup>, including active freezing<sup>96,97</sup>, as well as further types of time evolution, such as molecular dynamics or quantum mechanics should be interesting. This is feasible, as the Noether considerations are not restricted to overdamped classical systems, as (formal) power functional generators exist for quantum<sup>98</sup> and classical Hamiltonian<sup>99</sup> many-body systems. On the methodological side, besides power functional theory, our framework could be complemented by e.g. mode-coupling theory and Mori-Zwanzig techniques<sup>100</sup>, as well as approaches beyond that<sup>101</sup>. Given that the equilibrium force sum rules are crucial in the description of crystal<sup>33,34</sup> and liquid crystal<sup>35</sup> excitations, the study of such systems under drive is a further exciting prospect.

## Methods

**Relationship to classical results.** We give an overview of how the Noether sum rules relate to previously known results. The famous LMBW-equation was derived independently by Lovett et al.<sup>10</sup> and by Wertheim<sup>11</sup> and reads

$$\nabla \ln \rho(\mathbf{r}) + \beta \nabla V_{\text{ext}}(\mathbf{r}) = \int d\mathbf{r}' c_2(\mathbf{r}, \mathbf{r}') \nabla' \rho(\mathbf{r}'). \quad (47)$$

We can conclude that (47) is a combination of the local internal Noether sum rule (7) for translational symmetry and the equilibrium Euler–Lagrange equation  $c_1(\mathbf{r}) = \ln \rho(\mathbf{r}) + \beta \nabla V_{\text{ext}}(\mathbf{r}) - \beta \mu$ , where  $\mu$  indicates the chemical potential. LMBW also derived a lesser known external relation, which is equivalent to (47) and reads

$$\nabla \rho(\mathbf{r}) + \beta \rho(\mathbf{r}) \nabla V_{\text{ext}}(\mathbf{r}) = \int d\mathbf{r}' (g(\mathbf{r}, \mathbf{r}') - 1) \rho(\mathbf{r}') \nabla' V_{\text{ext}}(\mathbf{r}'). \quad (48)$$

We find that (48) contains the local external Noether sum rule (3) along with the relation  $H_2(\mathbf{r}, \mathbf{r}') = (g(\mathbf{r}, \mathbf{r}') - 1) \rho(\mathbf{r}) \rho(\mathbf{r}') + \rho(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}')$ <sup>18</sup>.

The Triezenberg–Zwanzig equation<sup>102</sup> holds for vanishing external potential  $V_{\text{ext}}(\mathbf{r}) = 0$  and is given by

$$\nabla \ln \rho(\mathbf{r}) = \int d\mathbf{r}' c_2(\mathbf{r}, \mathbf{r}') \nabla' \rho(\mathbf{r}'). \quad (49)$$

Originally (49) was derived for the free liquid–vapor interface in a parallel geometry. We find that the relation consists of the local internal Noether sum rule (7) and the equilibrium Euler–Lagrange equation. The LMBW equation (48) reduces to the Triezenberg–Zwanzig equation (49) for cases of  $V_{\text{ext}}(\mathbf{r}) = 0$ .

Another related equation is the first member of the Yvon–Born–Green (YBG) hierarchy<sup>103,104</sup>,

$$\rho(\nabla \ln \rho(\mathbf{r}) + \beta \nabla V_{\text{ext}}(\mathbf{r})) = -\beta \int d\mathbf{r}' g(\mathbf{r}, \mathbf{r}') \rho(\mathbf{r}') \nabla \phi(|\mathbf{r} - \mathbf{r}'|), \quad (50)$$

where  $\phi(r)$  denotes the interparticle pair potential as before. Although (50) has a similar structure as the LMBW equation, it is not based on symmetry or Noether arguments but arises from integration out of degrees of freedom. If one would like to include the translational symmetry one can simply replace the left hand side of (50) with the right hand side of the LMBW equation (47), which leads to

$$\int d\mathbf{r}' c_2(\mathbf{r}, \mathbf{r}') \nabla' \rho(\mathbf{r}') = -\beta \int d\mathbf{r}' g(\mathbf{r}, \mathbf{r}') \rho(\mathbf{r}') \nabla \phi(|\mathbf{r} - \mathbf{r}'|). \quad (51)$$

Some of the here derived sum rules are rederivations of known relations. We reiterate the relationships. In his overview Baus<sup>13</sup> showed that in equilibrium the total external (2) and internal (6) force vanish and derived the corresponding local hierarchies (3), (4) and (7), (8). Similar sum rules<sup>13</sup> hold for the external (22) and internal (25) total torques and their corresponding hierarchies (23), (24) and (26), (27). Tarazona and Evans<sup>16</sup> generalized these equations for uniaxial particles and derived the first order of the external (34) and internal (37) hierarchies due to rotations.

To the best of our knowledge the hierarchies of global Noether sum rules, such as (5), (9), (28), and (29), have not been determined previously. Furthermore the global external (33) and internal (36) Noether sum rules for uniaxial colloids and their corresponding hierarchies (35) and (38) (with exception of the first order) are reported here for the first time. As the considerations in the literature focused on equilibrium, all our nonequilibrium relations, as e.g. (11)–(13) and especially the ones including memory (43)–(46) have not been found before.

**Data availability**

Data sharing is not applicable to this study as no datasets were generated or analyzed during the current study.

Received: 2 March 2021; Accepted: 1 July 2021;

Published online: 05 August 2021

**References**

- Noether, E. Invariante Variationsprobleme. *Nachr. d. König. Gesellsch. d. Wiss. zu Göttingen, Math.-Phys. Klasse* **235** (1918). English translation by Tavel, M. A. Invariant variation problems. *Transp. Theo. Stat. Phys.* **1**, 186 (1971); for a version in modern typesetting see: Wang, F.Y. [arXiv:physics/0503066v3](https://arxiv.org/abs/physics/0503066v3) (2018).
- Neuenschwander, D. E. *Emmy Noether's Wonderful Theorem* (Johns Hopkins University Press, 2011). For a description of many insightful and pedagogical examples and applications.
- Byers, N. E. Noether's discovery of the deep connection between symmetries and conservation laws. Preprint at <https://arxiv.org/abs/physics/9807044> (1998).
- Rowlinson, J. S. & Widom, B. *Molecular theory of capillarity* (Dover, New York, 2002).
- van der Waals, J. D. The thermodynamic theory of capillarity under the hypothesis of a continuous variation of density. *Z. Phys. Chem.* **13**, 657 (1894); English translation by J. S. Rowlinson. *J. Stat. Phys.* **20**, 197 (1979).
- Kerins, J. & Boiteux, M. Applications of noether's theorem to inhomogeneous fluids. *Phys. A* **117**, 575 (1983).
- Bukman, D. J. Torque balance at a line of contact. *Phys. A* **319**, 151 (2003).
- Boiteux, M. & Kerins, J. Thermodynamic properties of inhomogeneous fluids. *Phys. A* **121**, 399 (1983).
- Evans, R. The nature of the liquid-vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids. *Adv. Phys.* **28**, 143 (1979).
- Lovett, R. A., Mou, C. Y. & Buff, F. P. The structure of the liquid-vapor interface. *J. Chem. Phys.* **65**, 570 (1976).
- Wertheim, M. S. Correlations in the liquid-vapor interface. *J. Chem. Phys.* **65**, 2377 (1976).
- Kayser, R. F. & Raveché, H. J. Emergence of periodic density patterns. *Phys. Rev. B* **22**, 424 (1980).
- Baus, M. Broken symmetry and invariance properties of classical fluids. *Mol. Phys.* **51**, 211 (1984).
- Lovett, R. & Buff, F. P. Examples of the construction of integral equations in equilibrium statistical mechanics from invariance principles. *Physica A* **172**, 147 (1991).
- Baus, M. & Lovett, R. A direct derivation of the profile equations of Buff-Lovett-Mou-Wertheim from the Born-Green-Yvon equations for a non-uniform equilibrium fluid. *Physica A* **181**, 329 (1992).
- Tarazona, P. & Evans, R. On the validity of certain integro-differential equations for the density-orientation profile of molecular fluid interfaces. *Chem. Phys. Lett.* **97**, 279 (1983).
- Gubbins, K. E. Structure of nonuniform molecular fluids – integrodifferential equations for the density-orientation profile. *Chem. Phys. Lett.* **76**, 329 (1980).
- Hansen, J. P. & McDonald, I. R. *Theory of Simple Liquids*, 4th ed. (Academic Press, London, 2013).
- Evans, R., Oettel, M., Roth, R. & Kahl, G. New developments in classical density functional theory. *J. Phys.: Condens. Matter* **28**, 240401 (2016).
- Roth, R. Fundamental measure theory for hard-sphere mixtures: a review. *J. Phys. Condens. Matter* **22**, 063102 (2010).
- Levesque, M., Vuilleumier, R. & Borgis, D. Scalar fundamental measure theory for hard spheres in three dimensions: Application to hydrophobic solvation. *J. Chem. Phys.* **137**, 034115 (2012).
- Jeanmairat, G., Levesque, M. & Borgis, D. Molecular density functional theory of water describing hydrophobicity at short and long length scales. *J. Chem. Phys.* **139**, 154101 (2013).
- Evans, R. & Wilding, N. B. Quantifying density fluctuations in water at a hydrophobic surface: evidence for critical drying. *Phys. Rev. Lett.* **115**, 016103 (2015).
- Evans, R., Stewart, M. C. & Wilding, N. B. A unified description of hydrophilic and superhydrophobic surfaces in terms of the wetting and drying transitions of liquids. *Proc. Nat. Acad. Sci.* **116**, 23901 (2019).
- Remsing, R. C. Commentary: playing the long game wins the cohesion-adhesion rivalry. *Proc. Nat. Acad. Sci.* **116**, 23874 (2019).
- Evans, R., Stewart, M. C. & Wilding, N. B. Critical drying of liquids. *Phys. Rev. Lett.* **117**, 176102 (2016).
- Chacko, B., Evans, R. & Archer, A. J. Solvent fluctuations around solvophobic, solvophilic, and patchy nanostructures and the accompanying solvent mediated interactions. *J. Chem. Phys.* **146**, 124703 (2017).
- Martin-Jimenez, D., Chacón, E., Tarazona, P. & Garcia, R. Atomically resolved three-dimensional structures of electrolyte aqueous solutions near a solid surface. *Nat. Comm.* **7**, 12164 (2016).
- Hernández-Muñoz, J., Chacón, E. & Tarazona, P. Density functional analysis of atomic force microscopy in a dense fluid. *J. Chem. Phys.* **151**, 034701 (2019).
- Muscattello, J., Chacón, E., Tarazona, P. & Bresme, F. Deconstructing temperature gradients across fluid interfaces: the structural origin of the thermal resistance of liquid-vapor interfaces. *Phys. Rev. Lett.* **119**, 045901 (2017).
- Xu, X. & Rice, S. A. A density functional theory of one- and two-layer freezing in a confined colloid system. *Proc. R. Soc. A* **464**, 65 (2008).
- Brader, J. M. Structural precursor to freezing: an integral equation study. *J. Chem. Phys.* **128**, 104503 (2008).
- Walz, C. & Fuchs, M. Displacement field and elastic constants in nonideal crystals. *Phys. Rev. B* **81**, 134110 (2010).
- Häring, J. M., Walz, C., Szamel, G. & Fuchs, M. Coarse-grained density and compressibility of nonideal crystals: General theory and an application to cluster crystals. *Phys. Rev. B* **92**, 184103 (2015).
- Häring, J. M. Microscopically founded elasticity theory for defect-rich systems of anisotropic particles. Ph.D. Thesis, Universität Konstanz (2020).
- Bryk, P., Henderson, D. & Sokolowski, S. A fluid in contact with a semipermeable surface: second-order integral equation approach. *J. Chem. Phys.* **107**, 3333 (1997).
- Henderson, J. R. & van Swol, F. On the interface between a fluid and a planar wall. *Mol. Phys.* **51**, 991 (1984).
- Tejero, C. F. & Baus, M. Viscoelastic surface waves and the surface structure of liquids. *Mol. Phys.* **54**, 1307 (1985).
- Iatsevitch, S. & Forstmann, F. Density profiles at liquid-vapor and liquid-liquid interfaces: an integral equation study. *J. Chem. Phys.* **107**, 6925 (1997).
- Kasch, M. & Forstmann, F. An orientational instability and the liquid-vapor interface of a dipolar hard sphere fluid. *J. Chem. Phys.* **99**, 3037 (1993).
- Mandal, S., Lang, S., Botan, V. & Franosch, T. Nonergodicity parameters of confined hard-sphere glasses. *Soft. Matter* **13**, 6167 (2017).
- Farage, T. F. F., Krinninger, P. & Brader, J. M. Effective interactions in active Brownian suspensions. *Phys. Rev. E* **91**, 042310 (2015).
- Paliwal, S., Rodenburg, J., van Roij, R. & Dijkstra, M. Chemical potential in active systems: predicting phase equilibrium from bulk equations of state? *New J. Phys.* **20**, 015003 (2018).
- Paliwal, S., Prymidis, V., Filion, L. & Dijkstra, M. Non-equilibrium surface tension of the vapour-liquid interface of active Lennard-Jones particles. *J. Chem. Phys.* **147**, 084902 (2017).
- Loehr, J., Loenne, M., Ernst, A., de las Heras, D. & Fischer, T. M. Topological protection of multiparticle dissipative transport. *Nat. Commun.* **7**, 11745 (2016).
- Loehr, J. et al. Colloidal topological insulators. *Comms. Phys.* **1**, 4 (2018).
- Rossi, A. M. E. B. et al. Hard topological versus soft geometrical magnetic particle transport. *Soft Matter* **15**, 8543 (2019).
- Schmidt, M. & Brader, J. M. Power functional theory for Brownian dynamics. *J. Chem. Phys.* **138**, 214101 (2013).
- Fortini, A., de las Heras, D., Brader, J. M. & Schmidt, M. Superadiabatic forces in Brownian many-body dynamics. *Phys. Rev. Lett.* **113**, 167801 (2014).
- Krinninger, P., Schmidt, M. & Brader, J. M. Nonequilibrium phase behaviour from minimization of free power dissipation. *Phys. Rev. Lett.* **117**, 208003 (2016).
- Hermann, S., de las Heras, D. & Schmidt, M. Non-negative interfacial tension in phase-separated active Brownian particles. *Phys. Rev. Lett.* **123**, 268002 (2019).
- Hermann, S., Krinninger, P., de las Heras, D. & Schmidt, M. Phase coexistence of active Brownian particles. *Phys. Rev. E* **100**, 052604 (2019).
- Hermann, S. & Schmidt, M. Active interface polarization as a state function. *Phys. Rev. Research* **2**, 022003(R) (2020).
- de las Heras, D. & Schmidt, M. Velocity gradient power functional for Brownian dynamics. *Phys. Rev. Lett.* **120**, 028001 (2018).
- Stuhlmüller, N. C. X., Eckert, T., de las Heras, D. & Schmidt, M. Structural nonequilibrium forces in driven colloidal systems. *Phys. Rev. Lett.* **121**, 098002 (2018).
- de las Heras, D. & Schmidt, M. Flow and structure in nonequilibrium Brownian many-body systems. *Phys. Rev. Lett.* **125**, 018001 (2020).
- Brader, J. M. & Schmidt, M. Nonequilibrium Ornstein-Zernike relation for Brownian many-body dynamics. *J. Chem. Phys.* **139**, 104108 (2013).
- Brader, J. M. & Schmidt, M. Dynamic correlations in Brownian many-body systems. *J. Chem. Phys.* **140**, 034104 (2014).
- Ornstein, L. S. & Zernike, F. The influence of accidental deviations of density on the equation of state. *Proc. Acad. Sci. Amsterdam* **17**, 793 (1914); this article is reprinted in Frisch, H. & Lebowitz, J. L. *The Equilibrium Theory of Classical Fluids* (Benjamin, New York, 1964).

60. Requardt, M. & Wagner, H. J. (Infinite) Boundary corrections for the LMBW-Equations and the TZ-Formula of surface tension in the presence of spontaneous symmetry breaking. *Phys. A* **154**, 183 (1988).
61. Khadka, U., Holubec, V., Yang, H. & Cichos, F. Active particles bound by information flows. *Nat. Commun.* **9**, 3864 (2018).
62. Selmke, M., Khadka, U., Bregulla, A. P., Cichos, F. & Yang, H. Theory for controlling individual self-propelled micro-swimmers by photon nudging I: directed transport. *Phys. Chem. Chem. Phys.* **20**, 10502 (2018).
63. Selmke, M., Khadka, U., Bregulla, A. P., Cichos, F. & Yang, H. Theory for controlling individual self-propelled micro-swimmers by photon nudging II: confinement. *Phys. Chem. Chem. Phys.* **20**, 10521 (2018).
64. Bregulla, A. P., Yang, H. & Cichos, F. Stochastic localization of microswimmers by photon nudging. *ACS Nano* **8**, 6542 (2014).
65. Qian, B., Montiel, D., Bregulla, A., Cichos, F. & Yang, H. Harnessing thermal fluctuations for purposeful activities: the manipulation of single micro-swimmers by adaptive photon nudging. *Chem. Sci.* **4**, 1420 (2013).
66. Söker, N. A., Auschra, S., Holubec, V., Kroy, K. & Cichos, F. How Activity Landscapes Polarize Microswimmers without Alignment Forces. *Phys. Rev. Lett.* **126**, 228001 (2021).
67. Auschra, S., Holubec, V., Söker, N. A., Cichos, F. & Kroy, K. Polarization-density patterns of active particles in motility gradients. *Phys. Rev. E* **103**, 062601 (2021).
68. Speck, T. & Jack, R. L. Ideal bulk pressure of active Brownian particles. *Phys. Rev. E* **93**, 062605 (2016).
69. Takatori, S. C., Yan, W. & Brady, J. F. Swim pressure: stress generation in active matter. *Phys. Rev. Lett.* **113**, 028103 (2014).
70. Yang, X., Manning, M. L. & Marchetti, M. C. Aggregation and segregation of confined active particles. *Soft. Matter* **10**, 6477 (2014).
71. Winkler, R. G., Wysocki, A. & Gompper, G. Virial pressure in systems of spherical active Brownian particles. *Soft. Matter* **11**, 6680 (2015).
72. Solon, A. P. et al. Pressure and phase equilibria in interacting active Brownian spheres. *Phys. Rev. Lett.* **114**, 198301 (2015).
73. Paliwal, S., Rodenburg, J., van Roij, R. & Dijkstra, M. Chemical potential in active systems: predicting phase equilibrium from bulk equations of state? *New J. Phys.* **20**, 015003 (2018).
74. de las Heras, D. et al. Floating nematic phase in colloidal platelet-sphere mixtures. *Sci. Rep.* **2**, 789 (2012).
75. de las Heras, D. & Schmidt, M. The phase stacking diagram of colloidal mixtures under gravity. *Soft Matter* **9**, 98636 (2013).
76. de las Heras, D. & Schmidt, M. Sedimentation stacking diagram of binary colloidal mixtures and bulk phases in the plane of chemical potentials. *J. Phys.: Condens. Matter* **27**, 194115 (2015).
77. Geigenfeind, T. & de las Heras, D. The role of sample height in the stacking diagram of colloidal mixtures under gravity. *J. Phys.: Condens. Matter* **29**, 064006 (2017).
78. Enculescu, M. & Stark, H. Active colloidal suspensions exhibit polar order under gravity. *Phys. Rev. Lett.* **107**, 058301 (2011).
79. Hermann, S. & Schmidt, M. Active ideal sedimentation: exact two-dimensional steady states. *Soft Matter* **14**, 1614 (2018).
80. Vachier, J. & Mazza, M. G. Dynamics of sedimenting active Brownian particles. *Eur. Phys. J. E* **42**, 11 (2019).
81. Solon, A. P., Cates, M. E. & Tailleur, J. Active Brownian particles and run-and-tumble particles: A comparative study. *Eur. Phys. J. Spec. Top.* **224**, 1231 (2015).
82. Lee, C. F. Active particles under confinement: aggregation at the wall and gradient formation inside a channel. *New J. Phys.* **15**, 055007 (2013).
83. Zwanzig, R. *Nonequilibrium statistical mechanics* (Oxford University Press, 2001).
84. Lesnicki, D., Vuilleumier, R., Carof, A. & Rotenberg, B. Molecular hydrodynamics from memory kernels. *Phys. Rev. Lett.* **116**, 147804 (2016).
85. Lesnicki, D. & Vuilleumier, R. Microscopic flow around a diffusing particle. *J. Chem. Phys.* **147**, 094502 (2017).
86. Jung, G. & Schmid, F. Computing bulk and shear viscosities from simulations of fluids with dissipative and stochastic interactions. *J. Chem. Phys.* **144**, 204104 (2016).
87. Jung, G., Hanke, M. & Schmid, F. Iterative reconstruction of memory kernels. *J. Chem. Theo. Comput.* **13**, 2481 (2017).
88. Jung, G., Hanke, M. & Schmid, F. Generalized Langevin dynamics: construction and numerical integration of non-Markovian particle-based models. *Soft Matter* **14**, 9368 (2018).
89. Yeomans-Reyna, L. & Medina-Noyola, M. Overdamped van Hove function of colloidal suspensions. *Phys. Rev. E* **62**, 3382 (2000).
90. Chávez-Rojó, M. A. & Medina-Noyola, M. Van Hove function of colloidal mixtures: Exact results. *Physica A* **366**, 55 (2006).
91. Lázaro-Lázaro, E. et al. Self-consistent generalized Langevin equation theory of the dynamics of multicomponent atomic liquids. *J. Chem. Phys.* **146**, 184506 (2017).
92. Treffendstätt, L. L. & Schmidt, M. Memory-induced motion reversal in Brownian liquids. *Soft Matter* **16**, 1518 (2020).
93. Rosenfeld, Y. Free-energy model for the inhomogeneous hard-sphere fluid mixture and density-functional theory of freezing. *Phys. Rev. Lett.* **63**, 9 (1989).
94. Tschopp, S. M. & Brader, J. M. Fundamental measure theory of inhomogeneous two-body correlation functions. *Phys. Rev. E* **103**, 042103 (2021).
95. Lin, S.-C., Oettel, M., Häring, J. M., Haussmann, R., Fuchs, M., & Kahl, G. Direct correlation function of a crystalline solid. *Phys. Rev. Lett.* (to appear). Preprint at <https://arxiv.org/abs/2104.11558> (2021).
96. Turci, F. & Wilding, N. B. Phase separation and multibody effects in three-dimensional active Brownian particles. *Phys. Rev. Lett.* **126**, 038002 (2021).
97. Omar, A. K., Klymko, K., GrandPre, T. & Geissler, P. L. Phase diagram of active Brownian spheres: crystallization and the metastability of motility-induced phase separation. *Phys. Rev. Lett.* **126**, 188002 (2021).
98. Schmidt, M. Quantum power functional theory for many-body dynamics. *J. Chem. Phys.* **143**, 174108 (2015).
99. Schmidt, M. Power functional theory for Newtonian many-body dynamics. *J. Chem. Phys.* **148**, 044502 (2018).
100. Janssen, L. M. C. Mode-coupling theory of the glass transition: A primer. *Front. Phys.* **6**, 97 (2018).
101. Vogel, F. & Fuchs, M. Stress correlation function and linear response of Brownian particles. *Eur. Phys. J. E* **43**, 70 (2020).
102. Triezenberg, D. G. & Zwanzig, R. Fluctuation theory of surface tension. *Phys. Rev. Lett.* **28**, 1183 (1972).
103. Yvon, J. *Actualités Scientifiques et Industrielles* (Hermann & Cie., 1935).
104. Born, M. & Green, H. S. A general kinetic theory of liquids I. The molecular distribution functions. *Proc. R. Soc. London, Ser. A* **188**, 10 (1946).

### Acknowledgements

This work is supported by the German Research Foundation (DFG) via project number 436306241. Open Access funding is provided by project DEAL. We would like to thank Matthias Fuchs, Johannes Häring, Bob Evans, Pedro Tarazona, Daniel de las Heras, Matthias Krüger, Jim Lutsko, Hartmut Löwen, and Thomas Fischer for useful discussions and comments.

### Author contributions

S.H. and M.S. have jointly carried out the work and written the paper.

### Funding

Open Access funding enabled and organized by Projekt DEAL.

### Competing interests

The authors declare no competing interests.

### Additional information


**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42005-021-00669-2>.

**Correspondence** and requests for materials should be addressed to S.H. or M.S.

**Peer review information** *Communications Physics* thanks Benjamin Goddard, Gyula Toth and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021



## Topical Review

# Why Noether's theorem applies to statistical mechanics

Sophie Hermann\*<sup>1</sup> and Matthias Schmidt<sup>1,\*</sup>

Theoretische Physik II, Physikalisches Institut, Universität Bayreuth, D-95447 Bayreuth, Germany

E-mail: [Sophie.Hermann@uni-bayreuth.de](mailto:Sophie.Hermann@uni-bayreuth.de) and [Matthias.Schmidt@uni-bayreuth.de](mailto:Matthias.Schmidt@uni-bayreuth.de)

Received 8 November 2021, revised 3 March 2022

Accepted for publication 7 March 2022

Published 21 April 2022



CrossMark

**Abstract**

Noether's theorem is familiar to most physicists due its fundamental role in linking the existence of conservation laws to the underlying symmetries of a physical system. Typically the systems are described in the particle-based context of classical mechanics or on the basis of field theory. We have recently shown (2021 *Commun. Phys.* 4 176) that Noether's reasoning also applies to thermal systems, where fluctuations are paramount and one aims for a statistical mechanical description. Here we give a pedagogical introduction based on the canonical ensemble and apply it explicitly to ideal sedimentation. The relevant mathematical objects, such as the free energy, are viewed as functionals. This vantage point allows for systematic functional differentiation and the resulting identities express properties of both macroscopic average forces and molecularly resolved correlations in many-body systems, both in and out-of-equilibrium, and for active Brownian particles. To provide further background, we briefly describe the variational principles of classical density functional theory, of power functional theory, and of classical mechanics.

Keywords: statistical mechanics, density functional theory, power functional theory, invariance, Noether's theorem, liquid state theory, sum rules

(Some figures may appear in colour only in the online journal)

**1. Introduction**

Symmetries and their breaking in often stunningly beautiful ways are at the core of a broad range of phenomena in physics, from phase transitions in condensed matter to mass generation via the Higgs mechanism. Most readers will be very familiar with the importance of symmetry operations, including complex operations such as CPT-invariance in high energy physics as well as the simple challenge of centering the webcam while having mirroring switched off in a video call.

<sup>1</sup> [www.mschmidt.uni-bayreuth.de](http://www.mschmidt.uni-bayreuth.de)

\* Authors to whom any correspondence should be addressed.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

The exploitation of the underlying symmetries of a physical system is an important and central concept that allows to simplify the mathematical description and arguably more importantly to gain physical insights and achieve an understanding of the true mechanisms at play. This is what the mathematician Emmy Noether did in her groundbreaking work in functional analysis early in the twentieth century [1].

Noether analyzed carefully the changes that occur upon performing a symmetry operation on a system. Her work solved the then open deep problems of energy conservation in general relativity, as the new theory of gravity that Einstein had just formed. Noether considered the formulation of general relativity via Hilbert's action integral, which is a formal object—a functional—that generates Einstein's field equations. Nowadays Noether's theorems [1–3] are widely known and used to

connect each continuous symmetry of a system with a corresponding conservation law. Noether's work therefore forms a staple of physics, relevant from introductory classical mechanics to advanced theories such as the standard model of high energy particle physics.

In practice the theorems are usually applied to the action functional in a Lagrangian or Hamiltonian theory. This strategy is not of mere historic interest, as much active current research is being carried out, see e.g. recent developments that addressed the action functional for systems that include random forces [4–6] and work that shows, starting from the symmetry of an action functional, that the thermodynamic entropy can be viewed as a Noether invariant [7–9]. However, from a mathematical point of view, Noether's theorem is actually not restricted to the specific case of the action integral. The theorem rather applies to a much more general class of functionals, where it specifies general consequences of invariance under continuous symmetry transformations.

We recall some basics of functional calculus. A functional is a mathematical object that maps an entire function, i.e. the function values together with the corresponding values of the argument, to a single number. A popular introductory example of a functional is the definite integral, say over the unit interval from 0 to 1. When viewed as a functional, the definite integral accepts the integrand (a function) and it returns a number (the area under the curve that the function represents). Although the functional point of view might appear slightly uncommon (or even trivial in this case), the inherent abstract concept allows to formulate very significant insights and use powerful mathematical techniques of variational calculus which can be straightforwardly and widely applied.

The occurrence of functionals in physics is not restricted to the study of behaviour at very large length scales, such as that of the cosmos in the case of general relativity, or to very high energies, as is the case for fundamental theories of elementary particles. In fact the mathematical concept of a functional dependence is very general. Hence there is an according wide variety of objects in physics, such as e.g. the partition sum and the free energy in statistical mechanics that can be viewed as being a functional [10–13]; we give an introduction below. As soon as one is willing to accept this notion, making much headway is possible by analyzing physical properties of the considered system from this formal point of view.

To perform the transfer and use Noether's theorem for thermal systems, from a formal point of view one would need both to identify a suitable functional as well as a symmetry transformation under which this functional is invariant. One primary candidate for the choice of the functional is the partition function, which constitutes an integral over the high-dimensional phase space of classical mechanics. Within this context, phase space describes all degrees of freedom, i.e. the positions and momenta of all particles in the system. The partition sum itself is hence an integral over all these variables. Its integrand is, up to a constant, the Boltzmann factor of the energy function that characterizes the system. So the partition sum actually complies with the nature of a functional as it maps this function to just a number, i.e. the value of the partition sum. (As

detailed below the interesting functional dependence is that on the external potential.) The partition sum is arguably the most fundamental object in statistical physics, as all thermodynamic quantities, such as thermodynamic potentials including the free energy, the equation of state, but also position-resolved correlation functions can be obtained from it, at least in principle.

Within statistical mechanics, where one identifies the free energy with the negative logarithm of the partition sum, ordinary (parametric) derivatives of the free energy with respect to e.g. temperature and other thermodynamic variables generate thermodynamic quantities [11–13]. While the familiar process of building the derivative of a function, as giving a measure of the local slope, is a concept that dates back to Newton and Leibniz, functional differentiation is slightly less common. However, functionals can be differentiated in much the same way that functions can be differentiated. In case of the free energy, functional derivatives give microscopically resolved correlation functions [11–13]. These are quantities, such as the structure factor of a liquid, that are measurable in a lab, say with a scattering apparatus or even with a microscope upon further data processing.

When applying Emmy Noether's thinking to the free energy, one could expect mere abstraction to result, but that is not the case [14]. Consider the invariance under a spatial shift. This classical application of Noether's theorem to the action functional yields the well-known result of momentum conservation. When rather exploiting the invariance of the partition function and hence of the free energy with respect to shifting, what follows are fundamental statements about forces that act in the system [14]. One of them states that the total internal force vanishes. Here the total internal force is that which arises from the interactions only *between* the constituents of the system. The famous Baron Munchausen tale of bootstrapping himself out of the swamp by pulling on his own hair is identified as a fairytale by the Noetherian argument. The impossibility of this feat holds on the scale of his entire body, but also when locally resolving his structure on the molecular scale.

In addition to shifting, one can also consider rotations. In case of the action functional being invariant under rotations Noether's theorem implies that the angular momentum around the rotation axis is conserved. If the free energy has rotational symmetry, fundamental statements about torques emerge [14, 15]. These sum rules express inherent coupling of spin and orbital degrees of freedom. Figuratively speaking, the identities state that a bolt cannot screw itself into the wall and that a Baron Munchausen stuck in mud cannot spontaneously start to rotate by twisting his head.

Recognizing the functional dependence of the free energy allows to build a theory fully founded on a variational principle of thermal systems, as formulated by Mermin [10] and Evans [11, 12]. Their so-called density functional theory is a well-accepted and widely used theory, see reference [13] for a textbook presentation and reference [16] for an overview of recent work. Excellent approximations are available for relevant model fluids, such as for hard spheres [17, 18] (see

reference [19] for recent work addressing hard sphere crystal properties). The density functional approach hence allows explicit calculations to be carried out to predict the behavior of a wide range of physical systems, including solvation [20], hydrophobicity [21–23], critical drying of liquids [24], solvent fluctuations [25], electrolyte solutions near surfaces [26], interpretation of atomic force microscopy data [27], temperature gradients at fluid interfaces [28], and local fluctuations [22–25, 29]. In reference [14] we also apply Noether's thinking to a very recent variational approach for dynamics, called power functional theory [30, 31], which propels the functional concepts from equilibrium to nonequilibrium [30–45], including the recently popular active Brownian particles [46–52]. The generalization is important, as it shows that not only a dead Munchausen cannot bootstrap himself out of his misery, but that being alive does not help (in this particular case).

In the present contribution we demonstrate that the concepts of reference [14] apply to the canonical ensemble, as is relevant for confined systems [53–55] and for the dynamics [56–58]. Hence having an open system with respect to particle exchange is not necessary for the Noetherian arguments to apply. We give a detailed and somewhat pedagogical derivation of the fundamental concepts and also make much relevant background explicit, which has not been spelled out in reference [14].

The paper is organized as follows. In section 2 we go into some detail and we present in the following the arguably simplest example of the application of Noether's theorem to statistical mechanics. We expect the reader to be familiar with Newtonian mechanics and to ideally know about classical mechanics formulated in a more formal setting (we supply some basic notions thereof below). We lay out the canonical ensemble and averages in section 2.1. Forces and their relation to symmetries are addressed in section 2.2. Statistical functionals and their invariances are described in section 2.3. As an example we describe the application to sedimentation in section 2.4. The relationship of the Noether invariance to correlation functions is laid out in section 2.5. We give further background that is relevant for reference [14], such as the details of the grand canonical treatment and the variational principles of density functional theory and of power functional theory, in section 2.6. We present our conclusions in section 3.

## 2. Theory

### 2.1. Canonical ensemble and averages

We consider a system with fixed number of particles  $N$ . The state of the system is characterized by all positions  $\mathbf{r}_1, \dots, \mathbf{r}_N$  and momenta  $\mathbf{p}_1, \dots, \mathbf{p}_N$ , where the subscript labels the  $N$  particles, which we take to all have identical properties. We assume that the total energy consists of kinetic and potential energy contribution, according to

$$H = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + u(\mathbf{r}_1, \dots, \mathbf{r}_N) + \sum_{i=1}^N V_{\text{ext}}(\mathbf{r}_i). \quad (1)$$

Here  $H$  is the Hamiltonian of the system, with the interparticle interaction potential  $u(\mathbf{r}_1, \dots, \mathbf{r}_N)$  and the external one-

body potential  $V_{\text{ext}}(\mathbf{r}_i)$  acting on particle  $i$ . The equations of motion are generated via  $\dot{\mathbf{r}}_i = \partial H / \partial \mathbf{p}_i$  and  $\dot{\mathbf{p}}_i = -\partial H / \partial \mathbf{r}_i$ , where the overdot indicates a time derivative,  $m$  is the particle mass, and the index  $i = 1, \dots, N$ . Using the explicit form (1) of the Hamiltonian then leads to the equations of motion in the familiar form

$$\dot{\mathbf{r}}_i = \frac{\mathbf{p}_i}{m}, \quad (2)$$

$$\dot{\mathbf{p}}_i = \mathbf{f}_i, \quad (3)$$

where  $\mathbf{f}_i$  indicates the force on particle  $i$ , which consists of a contribution from all other particles as well as the external force. Explicitly, the force on particle  $i$  is given by

$$\mathbf{f}_i = -\nabla_i u(\mathbf{r}_1, \dots, \mathbf{r}_N) - \nabla_i V_{\text{ext}}(\mathbf{r}_i), \quad (4)$$

where  $\nabla_i$  denotes the derivative with respect to  $\mathbf{r}_i$ . (Building the derivative by a vector implies building the derivative with respect to each component of the vector, hence  $\nabla_i$  can be viewed as building the gradient with respect to  $\mathbf{r}_i$ .) Certainly we could have written down the equations of motion (2) and (3) *a priori*. Equation (2) expresses the standard relation of velocity  $\dot{\mathbf{r}}_i$  with momentum  $\mathbf{p}_i$ , and (3) is Newton's second law. Hence we have reproduced the Newtonian theory within the Hamiltonian formalism.

So far everything has been deterministic and we were concerned with obtaining a description on the level of individual particles. As our aim is to describe very large systems, we wish to 'zoom out' and investigate and describe the macroscopic properties of the system, as they result from the above formulated microscopic picture. Statistical mechanics provides the means for doing so. We will not attempt to give a comprehensive description of the concepts of this theory. Rather we will guide the reader through some essential steps, including in particular how thermal averages are built, to see how Noether's theorem applies in this context. As we will see, both the physical concept and the outcome are different from the standard application of Noether's theorem based on the action expressed as a time integral over a Lagrangian that corresponds to (1); we give a brief description of this standard argument at the end of section 2.2.

Statistical mechanics rests on the concept of having a statistical ensemble, in the sense of the collection of microstates  $\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N$ , i.e. all phase space points. These are transcended beyond classical mechanics by each being assigned a probability for its occurrence. (There is much discussion about who throws the dice here; we recommend Zwanzig's cool-headed account [59].) The microstate probability distribution is given by a standard Boltzmann form,

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N) = \frac{e^{-\beta H}}{Z_N}, \quad (5)$$

where the inverse temperature is  $\beta = 1/(k_B T)$ , with the Boltzmann constant  $k_B$  and absolute temperature  $T$ . Here  $Z_N$  is the partition sum, and it acts to normalize the probability distribution to unity, when summed up over all microstates. The sum over microstates is in practice a high-dimensional integral over phase space, explicitly given as

$$Z_N = \frac{1}{N!h^{3N}} \int d\mathbf{r}_1 \dots d\mathbf{r}_N d\mathbf{p}_1 \dots d\mathbf{p}_N e^{-\beta H}, \quad (6)$$

where  $h$  indicates the Planck constant. Here each position integral and each momentum integral runs over  $\mathbb{R}^3$ . (We are considering systems in three spatial dimensions.) The system volume is rendered finite by confining walls that are modelled by a suitable form of the external potential  $V_{\text{ext}}(\mathbf{r})$ . As a note on units, recall that  $h$  carries energy multiplied by time, i.e. Js, such that the partition sum (6) carries no units.

The purpose of the probability distribution (5) is to build averages. Taking the Hamiltonian (1) as an example, we can express the total energy, averaged over the statistical ensemble, as

$$E = \frac{1}{N!h^{3N}} \int d\mathbf{r}_1 \dots d\mathbf{r}_N d\mathbf{p}_1 \dots d\mathbf{p}_N H \Psi. \quad (7)$$

Here we recall the dependence of the Hamiltonian (1) on the phase space point, and in the notation we have left away the arguments  $\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N$  of both  $H$  and  $\Psi$ .

It is useful to introduce more compact notation, as this reduces clutter and allows to express the structure of the theory more clearly. Let us denote the integral over phase space, together with its normalizing factor in (6) as the ‘classical trace’ operation, hence defined as

$$\text{Tr}_N = \frac{1}{N!h^{3N}} \int d\mathbf{r}_1 \dots d\mathbf{r}_N d\mathbf{p}_1 \dots d\mathbf{p}_N, \quad (8)$$

which is to be understood as acting on an integrand, such as on  $H\Psi$  in the example (7) above. Equation (7) can hence be expressed much more succinctly as

$$E = \text{Tr}_N H \Psi. \quad (9)$$

In a similar way we can express other averaged quantities, such as the average external (potential) energy,

$$U_{\text{ext}} = \text{Tr}_N \Psi \sum_{i=1}^N V_{\text{ext}}(\mathbf{r}_i). \quad (10)$$

In order to build some trust for the compact notation, we use (5) and (8) to re-write (10) explicitly as

$$U_{\text{ext}} = \frac{1}{N!h^{3N}} \int d\mathbf{r}_1 \dots d\mathbf{r}_N d\mathbf{p}_1 \dots d\mathbf{p}_N \times \frac{e^{-\beta H}}{Z_N} \sum_{i=1}^N V_{\text{ext}}(\mathbf{r}_i). \quad (11)$$

This allows to see explicitly that  $U_{\text{ext}}$  depends on the number of particle  $N$  and on the temperature  $T$  (via the Boltzmann factor and the partition sum). Surely (10) allows to see the physical content, that of an average being carried out, more clearly than (11) and we will continue to use the compact notation. (Readers who wish to familiarize themselves more intimately with these benefits are encouraged to put pen to scratch paper and re-write the following material in explicit notation.)

## 2.2. Forces and symmetries

Before continuing with thermal concepts, such as the free energy, we take a detour from standard paths in statistical mechanics, and return to forces. After all, it was the microscopically and particle-resolved forces  $\mathbf{f}_i$  in (4) that formed the starting point for the description of the coupled system. As an example, let us hence consider the total external force that acts on the system, in the sense that we sum up the external force that acts on each individual particle,  $-\nabla_i V_{\text{ext}}(\mathbf{r}_i)$ . This accounting results in  $-\sum_i \nabla_i V_{\text{ext}}(\mathbf{r}_i)$ . Note that this expression still applies per microstate, or in other words, the total external force varies in general across phase space. As a cautionary note on terminology, we use throughout the term ‘total’ in the above sense of denoting a global, macroscopic, extensive quantity. This usage is different from the also frequent meaning of total referring to the sum of intrinsic and external contributions.

In order to obtain the macroscopic description we need to trace over phase space and respect the probability for the occurrence of each given microstate. Hence the average total external force is given by

$$\mathbf{F}_{\text{ext}}^0 = -\text{Tr}_N \Psi \sum_{i=1}^N \nabla_i V_{\text{ext}}(\mathbf{r}_i). \quad (12)$$

Due to the structure of (12),  $\mathbf{F}_{\text{ext}}^0$  depends on the number of particles  $N$  (via the upper limit of the sum and the dimensionality of the phase space integrals), on temperature  $T$  (via the thermal distribution  $\Psi$ , cf (5) and (6)), and it of course also depends on the form of the function  $V_{\text{ext}}(\mathbf{r})$ . Note that the function  $V_{\text{ext}}(\mathbf{r})$  appears both explicitly in the gradient in (12) as well as in a more hidden form in the probability distribution  $\Psi$ , cf (5) and (1).

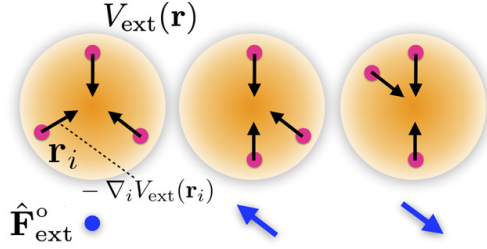
Let us halt for a moment and ponder the physics. Imagine having a vessel with impenetrable walls, such that the system stays confined inside of the vessel. Furthermore, to add some flavour, imagine an external field such as gravity acting on the system. Then the external potential consists of two contributions, i.e. the potential energy that the container walls exert on each given particle plus the gravitational energy. In an equilibrium situation, what would we expect the total external force to be like? Surely, it should not change in time. (Technically any time evolution had been superseded by the ensemble, which is a static one in the present case.) The reader might expect that

$$\mathbf{F}_{\text{ext}}^0 = 0, \quad (13)$$

because otherwise the system would surely start to move! However, as for any given microstate the total external force will in general *not* vanish, (13), if true, is a nontrivial property of thermal equilibrium. See figure 1 for an illustration of this concept, based on a system confined in a spherical cavity. Hence we wish to address carefully in the following whether we can prove (13) from first principles.

In the following we give two derivations of (13), which both rest on spatial translations of the system. The first derivation only requires vector calculus. The second derivation shows the Noetherian symmetry argument based on the functional





**Figure 1.** Three representative microstates  $\mathbf{r}^N$  for  $N = 3$  particles inside of a spherical cavity modelled by a confining external potential  $V_{\text{ext}}(\mathbf{r})$  (orange). Shown are the particle positions  $\mathbf{r}_i$  (pink dots) and the respective external force  $-\nabla_i V_{\text{ext}}(\mathbf{r}_i)$  acting on particle  $i$  (black arrows). The resulting total external force  $\hat{\mathbf{F}}_{\text{ext}}^o = -\sum_i \nabla_i V_{\text{ext}}(\mathbf{r}_i)$  is shown for each microstate (blue arrows and blue dot, the later indicating zero). Although  $\hat{\mathbf{F}}_{\text{ext}}^o$  for each microstate is in general nonzero, the average over the thermal ensemble vanishes,  $\mathbf{F}_{\text{ext}}^o = \langle \hat{\mathbf{F}}_{\text{ext}}^o \rangle = 0$ .

setting. This requires to adopt the notion of functional dependencies, which we have used only implicitly so far. In the following we make these relationships and dependencies explicit. We also supply the necessary methodology of functional differentiation and will attempt to convince the reader that their background in ordinary calculus can be flexed in order to follows these steps.

The fundamental ingredients to both derivations are identical though. We use the free energy and we monitor its changes upon spatial displacement of the system. The free energy, and more generally thermodynamic potentials, are central to thermal physics, and the following material can be viewed as a demonstration why this indeed is the case.

The free energy  $F_N$ , or more precisely: the total Helmholtz free energy is given by

$$F_N = -k_B T \ln Z_N, \quad (14)$$

where  $Z_N$  is the partition sum, as defined in (6). One can show that the relation of free energy and internal energy is given by the thermodynamic identity  $F_N = E - TS$ , where  $S$  is the entropy, here defined on a microscopic basis and the internal energy  $E$  is given by (9). One can surely be surprised by the promotion of the rather banal normalization factor  $Z_N$  to such a prominent and as we show decisive role. We demonstrate in the following that  $Z_N$  had been a dark horse, and that its status to generate the free energy via (14) is well-deserved.

Besides the free energy, the second ingredient that we require is a spatial shift of the entire system according to a displacement vector  $\epsilon$  of the system. We hence displace the external potential spatially by a constant vector  $\epsilon$ . (Although we Taylor expand in  $\epsilon$  below, the displacement  $\epsilon$  can be finite and arbitrary.) The displaced system is then under the influence of an external potential which has changed according to

$$V_{\text{ext}}(\mathbf{r}) \rightarrow V_{\text{ext}}(\mathbf{r} + \epsilon). \quad (15)$$

Formally, the free energy of the displaced system will depend on the displacement vector, i.e.

$$F_N \rightarrow F_N(\epsilon), \quad (16)$$

where  $F_N$  is the free energy (14) expressed in the original coordinates, and the new free energy is given by

$$F_N(\epsilon) = -k_B T \ln Z_N(\epsilon). \quad (17)$$

Here the partition sum of the shifted system is

$$Z_N(\epsilon) = \text{Tr}_N \exp \left[ -\beta \left( H_{\text{int}} + \sum_i V_{\text{ext}}(\mathbf{r}_i + \epsilon) \right) \right], \quad (18)$$

where the intrinsic part  $H_{\text{int}}$  of the Hamiltonian consists of kinetic energy and interparticle interaction potential energy only, i.e.  $H_{\text{int}} = \sum_i \mathbf{p}_i^2 / (2m) + u(\mathbf{r}_1, \dots, \mathbf{r}_N)$ .

We proceed by first recognizing that the shift does not change the value of the free energy (in other words, the choice of origin of the coordinate system does not matter). We can see this explicitly by performing a coordinate transformation  $\mathbf{r}_i \rightarrow \mathbf{r}_i - \epsilon$ . This leaves  $H_{\text{int}}$  invariant, as the momenta are unaffected and the internal interaction potential is unaffected. Recall that the interparticle energy only depends on relative particle positions, which remain invariant under the transformation:  $\mathbf{r}_i - \mathbf{r}_j \rightarrow (\mathbf{r}_i - \epsilon) - (\mathbf{r}_j - \epsilon) = \mathbf{r}_i - \mathbf{r}_j$ . Furthermore, due to the simplicity of the coordinate transformation that the shift represents, the phase space integral, of the classical trace (8), is unaffected as the Jacobian of the transformation is unity. Note that in the shifting operation, the momenta are unaffected and their behaviour remains governed by the Maxwell distribution throughout. Hence we have shown that the original free energy is identical to the free energy of the shifted system

$$F_N = F_N(\epsilon), \quad (19)$$

for any value of the displacement vector  $\epsilon$ .

At this point one could conclude mission accomplished. This is not what Emmy Noether did in her mathematical formulation of the problem—we hint at her variational techniques below. The way forward at this point is to rather ignore (19) and return to the explicit expression (17) for the free energy in the shifted system. We consider small displacements  $\epsilon$  and Taylor expand to first order,

$$F_N(\epsilon) = F_N + \left. \frac{\partial F_N(\epsilon)}{\partial \epsilon} \right|_{\epsilon=0} \cdot \epsilon, \quad (20)$$

where quadratic and higher order terms in  $\epsilon$  have been omitted. The partial derivative in (20) can be calculated explicitly:

$$\frac{\partial F_N(\epsilon)}{\partial \epsilon} = -\frac{k_B T}{Z_N(\epsilon)} \frac{\partial}{\partial \epsilon} Z_N(\epsilon) \quad (21)$$

$$= -\frac{k_B T}{Z_N(\epsilon)} \text{Tr}_N \frac{\partial}{\partial \epsilon} e^{-\beta H(\epsilon)} \quad (22)$$

$$= -\frac{k_B T}{Z_N(\epsilon)} \text{Tr}_N e^{-\beta H(\epsilon)} \frac{\partial}{\partial \epsilon} (-\beta) \sum_{i=1}^N V_{\text{ext}}(\mathbf{r}_i - \epsilon), \quad (23)$$

where in the first step (21) the partition sum in the denominator arises from the derivative of the logarithm in (17) and in the

second step (22) we have interchanged the phase space integration (as notated by  $\text{Tr}_N$ , cf (8)) and the  $\epsilon$ -derivative. The third step (23) follows directly from the structure of the Hamiltonian (1) and the fact that  $H_{\text{int}}$  is independent of  $\epsilon$ . We continue to obtain

$$\frac{\partial F_N(\epsilon)}{\partial \epsilon} = \text{Tr}_N \frac{e^{-\beta H(\epsilon)}}{Z_N(\epsilon)} \sum_{i=1}^N \frac{\partial}{\partial \epsilon} V_{\text{ext}}(\mathbf{r}_i - \epsilon) \quad (24)$$

$$= -\text{Tr}_N \Psi(\epsilon) \sum_{i=1}^N \frac{\partial}{\partial \mathbf{r}_i} V_{\text{ext}}(\mathbf{r}_i - \epsilon), \quad (25)$$

where in (24) we have pulled the partition sum as a constant inside of phase space integral and have moved the  $\epsilon$ -derivative inside the sum over all particles. In (25) we have combined the Boltzmann factor with the partition sum in order to express the many-body probability distribution function in the shifted system,  $\Psi(\epsilon) = \exp(-\beta H(\epsilon))/Z_N(\epsilon)$ , in generalization of (5). Furthermore the spatial derivative of the external potential is re-written via using  $\partial/\partial \epsilon = -\partial/\partial \mathbf{r}_i$  (which is valid due to the dependence on only the difference  $\mathbf{r}_i - \epsilon$ ). Considering the case  $\epsilon = 0$  allows us to conclude that

$$\left. \frac{\partial F_N(\epsilon)}{\partial \epsilon} \right|_{\epsilon=0} = -\text{Tr}_N \Psi \sum_{i=1}^N \frac{\partial}{\partial \mathbf{r}_i} V_{\text{ext}}(\mathbf{r}_i). \quad (26)$$

Remarkably the right-hand side is the average total external force as previously defined in (12). The left-hand side is identically zero, as  $\epsilon$  is arbitrary in (20) and the linear order (as well as all higher orders) need to vanish in the Taylor expansion (20) by virtue of the invariance (19) of the free energy upon spatial displacement. Hence

$$-\text{Tr}_N \Psi \sum_{i=1}^N \frac{\partial}{\partial \mathbf{r}_i} V_{\text{ext}}(\mathbf{r}_i) = 0, \quad (27)$$

which proves constructively the anticipated vanishing (13) of the average total external force (12).

As a preliminary summary, we have shown that the invariance of a global thermodynamic potential, the Helmholtz free energy expressed in the canonical ensemble, against spatial displacement (as generated by a shift of the external potential) leads to the non-trivial force identity of vanishing total external force. This identity holds true for any value of the number of particles in the system, at arbitrary temperature, and most notably irrespective of the precise form of the external potential. Hence we refer to statements such as  $\mathbf{F}_{\text{ext}}^0 = 0$ , cf (13), as a Noether identity or Noether sum rule. Clearly the concept is general, as both the symmetry operation can be altered (rotations are considered in reference [14]) as well as the type of thermodynamic object can be changed (the grand potential and the excess free energy density functional are considered in reference [14] and we shift the total external energy  $U_{\text{ext}}$  below in section 2.5).

We have presented here the shifting from the point of view that the actual physical system is moved to a different location. Alternatively, one could adopt a ‘passive’ point of view and displace only the origin of the coordinate system,

in the sense of using shifted coordinates that still describe an unchanged physical system. Then going through a chain of arguments analog to those given above yields identical results.

For completeness we contrast the present statistical mechanical treatment with the standard application of Noether’s theorem to deterministic dynamics. We keep the same  $N$ -body classical many-body system as before, i.e. with Hamiltonian  $H$  given by (1). The equations of motion (2) and (3) follow from the action integral  $S = \int_{t_1}^{t_2} dt L$ , where the Lagrangian  $L$  is obtained via  $L = \sum_i \mathbf{p}_i \dot{\mathbf{r}}_i - H$  and  $t_1$  and  $t_2$  are two fixed points in time. We apply the global shifting transformation  $\mathbf{r}_i \rightarrow \mathbf{r}_i - \epsilon$ , as before, to all particle coordinates in the system and at all times. As a consequence, the Lagrangian acquires a corresponding dependence on  $\epsilon$ . Taylor expanding the action to first order in  $\epsilon$  then yields

$$S(\epsilon) = S + \left. \frac{\partial S(\epsilon)}{\partial \epsilon} \right|_{\epsilon=0} \cdot \epsilon \quad (28)$$

$$= S + \int_{t_1}^{t_2} dt \left. \frac{\partial L}{\partial \epsilon} \right|_{\epsilon=0} \cdot \epsilon \quad (29)$$

$$= S - \int_{t_1}^{t_2} dt \sum_i \left. \frac{\partial L}{\partial \mathbf{r}_i} \right|_{\epsilon=0} \cdot \epsilon \quad (30)$$

$$= S - \int_{t_1}^{t_2} dt \sum_i \frac{d\mathbf{p}_i}{dt} \cdot \epsilon \quad (31)$$

$$= S - \sum_i \mathbf{p}_i \Big|_{t_1}^{t_2} \cdot \epsilon, \quad (32)$$

where  $S = S(\epsilon = 0)$  is the action in the original unshifted system; we have used the representation of  $S(\epsilon)$  as the time integral of the Lagrangian in the derivation of (29), the identity  $\partial L/\partial \epsilon = -\sum_i \partial L/\partial \mathbf{r}_i$  to obtain (30), the Lagrangian equations of motion  $d\mathbf{p}_i/dt = \partial L/\partial \mathbf{r}_i$  to derive (31), and the fact that the integrand of (31) is a total time differential to obtain (32).

Suppose now that the system is invariant under the displacement, such that  $S = S(\epsilon)$  for any value of  $\epsilon$  and the second term in (32) needs to vanish. This implies that the global momentum  $\mathbf{P}^0 = \sum_i \mathbf{p}_i$  is conserved, i.e.  $\mathbf{P}^0(t_2) = \mathbf{P}^0(t_1)$ .

### 2.3. Functionals and invariances

The abstraction that is yet to be performed and that allows to see the above statistical mechanical force result in an even wider setting, is based on functional methods. As we had hinted at in the introduction, integrals often allow for direct interpretation as functionals as they map their integrand (or part thereof) to the value of the quadrature. In the specific case at hand, we stay with the canonical free energy  $F_N$  and observe that its value certainly depends on the form of the external potential  $V_{\text{ext}}(\mathbf{r})$ , cf its occurrence in the Hamiltonian (1), which via the partition sum (6) enters the free energy (14). Hence we have  $F_N[V_{\text{ext}}]$ , where we indicate the functional dependence by square brackets (and leave away in the

notation the position argument  $\mathbf{r}$ , despite the fact that the functional depends on the entire function). In order to highlight this point of view, we rewrite (14) and (6), respectively, in the form

$$F_N[V_{\text{ext}}] = -k_B T \ln Z_N[V_{\text{ext}}], \quad (33)$$

$$Z_N[V_{\text{ext}}] = \text{Tr}_N \exp \left( -\beta H_{\text{int}} - \beta \sum_{i=1}^N V_{\text{ext}}(\mathbf{r}_i) \right), \quad (34)$$

where still the partition sum, viewed now as a functional of the external potential,  $Z_N[V_{\text{ext}}]$  is given by its elementary form, i.e. the right-hand side of (6). In a more compact form, eliminating  $Z_N[V_{\text{ext}}]$  as a standalone object, we have

$$F_N[V_{\text{ext}}] = -k_B T \ln \text{Tr}_N \exp \left( -\beta H_{\text{int}} - \beta \sum_{i=1}^N V_{\text{ext}}(\mathbf{r}_i) \right). \quad (35)$$

We dwell on the functional concept and demonstrate some practical consequences. As an analogy, viewing the functional dependence in (35) akin to the dependence of an ordinary function  $f(x)$  on its argument  $x$  brings concepts of calculus immediately to mind, such as building the derivative  $f'(x)$  and investigating its properties.

This analogy extends to functionals and their derivatives with respect to the argument function, in a process referred to as functional differentiation. For the present case, functionally deriving  $F_{\text{ext}}[V_{\text{ext}}]$  with respect to  $V_{\text{ext}}(\mathbf{r})$  can be viewed as monitoring the change of the value of the functional upon changing its argument function at position  $\mathbf{r}$ . The change will in general depend on position  $\mathbf{r}$ , hence building functional derivatives creates position dependence. (The result of the functional derivative is again a functional, as the dependence on the argument function persists.) Functional calculus is in many ways similar to ordinary multi-variable calculus. We do not attempt to give a tutorial here (see e.g. the appendix of reference [31] for a very brief one), but rather present a single example that is relevant for the present physics of invariance operations applied to many-body systems.

We use standard notation and denote the functional derivative with respect to the function  $V_{\text{ext}}(\mathbf{r})$  as  $\delta/\delta V_{\text{ext}}(\mathbf{r})$ . Applying this procedure to the free energy (33) yields

$$\frac{\delta F_N[V_{\text{ext}}]}{\delta V_{\text{ext}}(\mathbf{r})} = -k_B T \frac{\delta}{\delta V_{\text{ext}}(\mathbf{r})} \ln Z_N[V_{\text{ext}}] \quad (36)$$

$$= -\frac{k_B T}{Z_N[V_{\text{ext}}]} \frac{\delta}{\delta V_{\text{ext}}(\mathbf{r})} Z_N[V_{\text{ext}}], \quad (37)$$

where in the first step we have taken the multiplicative constant  $-k_B T$  out of the derivative and in the second step have used the ordinary chain rule, which also holds for functional differentiation. We next use the explicit form (34) to obtain

$$\frac{\delta F_N[V_{\text{ext}}]}{\delta V_{\text{ext}}(\mathbf{r})} = -\frac{k_B T}{Z_N[V_{\text{ext}}]} \text{Tr}_N \frac{\delta}{\delta V_{\text{ext}}(\mathbf{r})} e^{-\beta H} \quad (38)$$

$$= -\frac{k_B T}{Z_N[V_{\text{ext}}]} \text{Tr}_N e^{-\beta H} \frac{\delta}{\delta V_{\text{ext}}(\mathbf{r})} (-\beta H) \quad (39)$$

$$= \frac{1}{Z_N[V_{\text{ext}}]} \text{Tr}_N e^{-\beta H} \frac{\delta}{\delta V_{\text{ext}}(\mathbf{r})} \sum_{i=1}^N V_{\text{ext}}(\mathbf{r}_i), \quad (40)$$

where we have first exchanged the order of the functional derivative and the phase space integral, i.e. moved the derivative inside of the trace in (38), then in the second step (39) have used the chain rule to differentiate the exponential, and in the last step (40) have exploited the structure (1) of the Hamiltonian. Moving the derivative inside of the sum over  $i$  and identifying the many-body probability distribution function  $\Psi$  according to (5) yields the final result

$$\frac{\delta F_N[V_{\text{ext}}]}{\delta V_{\text{ext}}(\mathbf{r})} = \text{Tr}_N \Psi \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \equiv \rho(\mathbf{r}), \quad (41)$$

where we have used one central rule of functional differentiation: differentiating a function by itself gives  $\delta V_{\text{ext}}(\mathbf{r}_i)/\delta V_{\text{ext}}(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_i)$ , where the result  $\delta(\cdot)$  is the Dirac delta distribution (here in three dimensions, as its argument is a three-dimensional vector).

Notably in (41) we have arrived at the form of a thermal average over the statistical ensemble; recall the generic form exemplified by the average internal energy (9). Rather than the expectation value of the Hamiltonian, the present case represents the average of the microscopically resolved density operator  $\sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i)$ , which can be viewed as an indicator function that measures whether any particle resides at the given position  $\mathbf{r}$ . The result of the average is the one-body density distribution, or in short the density profile  $\rho(\mathbf{r})$ . That functional differentiation yields useful, spatially-resolved ('correlation') functions is a general mechanism. See e.g. [13] for much background on correlation functions and their generation via functional differentiation. Reference [14] carries this concept much further than we do here.

We return to the shifting symmetry operation of above, but now monitor the system response via tracking the changes in the function  $V_{\text{ext}}(\mathbf{r})$  that are induced by the spatial shifting. Recall the elementary Taylor expansion

$$V_{\text{ext}}(\mathbf{r} + \boldsymbol{\epsilon}) = V_{\text{ext}}(\mathbf{r}) + \boldsymbol{\epsilon} \cdot \nabla V_{\text{ext}}(\mathbf{r}), \quad (42)$$

where  $\nabla$  indicates the derivative (gradient) with respect to  $\mathbf{r}$  and we have truncated at linear order. See figure 2 for an illustration. The first order term in (42) can be viewed as a local change in the external potential,  $\delta V_{\text{ext}}(\mathbf{r})$ , which is given by

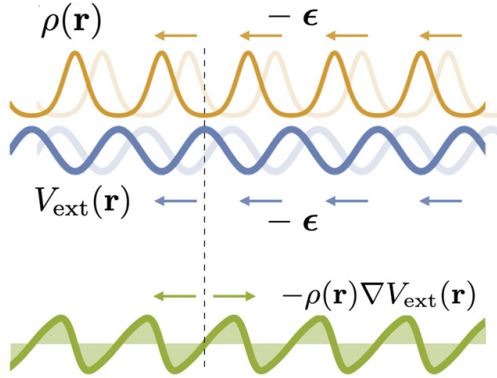
$$\delta V_{\text{ext}}(\mathbf{r}) \equiv \boldsymbol{\epsilon} \cdot \nabla V_{\text{ext}}(\mathbf{r}). \quad (43)$$

In order to capture the resulting effect on the functional, we can functionally Taylor expand the dependence of the free energy on  $V_{\text{ext}}(\mathbf{r}) + \delta V_{\text{ext}}(\mathbf{r})$  around the function  $V_{\text{ext}}(\mathbf{r})$ . To linear order in  $\delta V_{\text{ext}}(\mathbf{r})$  the functional Taylor expansion reads

$$F_N[V_{\text{ext}} + \delta V_{\text{ext}}] = F_N[V_{\text{ext}}] + \int d\mathbf{r} \frac{\delta F_N[V_{\text{ext}}]}{\delta V_{\text{ext}}(\mathbf{r})} \delta V_{\text{ext}}(\mathbf{r}) \quad (44)$$

$$= F_N[V_{\text{ext}}] + \int d\mathbf{r} \rho(\mathbf{r}) \boldsymbol{\epsilon} \cdot \nabla V_{\text{ext}}(\mathbf{r}), \quad (45)$$

where in (45) we have used the explicit form (43) of  $\delta V_{\text{ext}}(\mathbf{r})$  as it arises from the fact that the variation in the shape of the



**Figure 2.** Illustration of the shifting. A sinusoidal external potential  $V_{\text{ext}}(\mathbf{r})$  (blue lines) is spatially displaced by a displacement  $-\epsilon$  (blue arrows). The density profile  $\rho(\mathbf{r})$  (amber lines) measures the local probability to find a particle; it hence has e.g. peaks at the troughs of the external potential and it is shifted accordingly (amber arrows). Also shown is the magnitude of the external force density  $-\rho(\mathbf{r})\nabla V_{\text{ext}}(\mathbf{r})$  (green line); the green arrows represent its local direction. The horizontal dashed line is a guide that indicates the position of locally vanishing external force density.

external potential is specifically generated by a spatial displacement, cf (42). Furthermore we have used (41) to identify the functional derivative in (44) as the density profile.

The result (45) is based on the properties of functional calculus alone. Hence the identity is general and holds, to linear order in  $\epsilon$ , irrespective of any invariance properties. For the case of the total free energy, which as we have shown above in (19) is invariant under spatial displacement, we have

$$F_N[V_{\text{ext}}] = F_N[V_{\text{ext}} + \delta V_{\text{ext}}], \quad (46)$$

where  $\delta V_{\text{ext}}(\mathbf{r})$  is generated from the spatial displacement of the system, cf (43). Hence (43) together with (46) express in functional language the translational symmetry properties of the free energy.

From the identity (46) and the linear Taylor expansion (45) we can conclude that the correction term needs to vanish,

$$\int d\mathbf{r} \rho(\mathbf{r}) \epsilon \cdot \nabla V_{\text{ext}}(\mathbf{r}) = 0. \quad (47)$$

The displacement vector  $\epsilon$  is arbitrary, as there was no restriction on the direction of the shift. Hence the above expression can only identically vanish provided that [14, 66]

$$\mathbf{F}_{\text{ext}}^0 = - \int d\mathbf{r} \rho(\mathbf{r}) \nabla V_{\text{ext}}(\mathbf{r}) = 0, \quad (48)$$

where we have multiplied by  $-1$  in order to identify the one-body expression for the total external force  $\mathbf{F}_{\text{ext}}^0$ ; the equivalence with the many-body form (12) is straightforward to show upon using the definition of the density profile (41). See figure 2 for an illustration of the local force density profile, i.e. the integrand of (48).

#### 2.4. Application to sedimentation

We exemplify the general result (48) using the concrete example of a thermal system under gravity, such that sedimentation-diffusion equilibrium is reached. Recall that we consider systems at finite temperature, where entropic effects compete with ordering generated by the potential energy. We first omit the interparticle interactions, and hence consider the classical monatomic ideal gas. We assume that the external potential consists of a gravitational contribution,  $mgz$ , where  $g$  indicates the gravitational acceleration and  $z$  is the height variable. Furthermore due to the presence of a lower container wall, there is a repulsive contribution, which we take to be a harmonic potential with spring constant  $\alpha$  acting ‘inside’ the wall, i.e. at altitudes  $z < 0$ . Hence the specific form of the total external potential is

$$V_{\text{ext}}(z) = mgz + \frac{\alpha z^2}{2} \Theta(-z), \quad (49)$$

where  $\Theta(\cdot)$  indicates the Heaviside (unit step) function, which ensures that the parabolic potential only acts for  $z < 0$ . There is no need for the presence of an upper wall to close the system, as gravity alone already ensures that  $V_{\text{ext}} \rightarrow \infty$  for  $z \rightarrow \infty$ . The magnitude of the external force field is obtained as  $-V'_{\text{ext}}(z) = -mg - \alpha z \Theta(-z)$ , see figure 3 for an illustration (blue line).

The density distribution of the isothermal ideal gas is given by the generalized barometric law [13],

$$\rho(z) = \Lambda^{-3} e^{-\beta(V_{\text{ext}}(z) - \mu)}, \quad (50)$$

where  $\Lambda$  is the thermal de Broglie wavelength which arises from carrying out the momentum integrals in  $\text{Tr}_N$  (this is analytically possible due to the simple kinetic energy part of the Boltzmann factor). The chemical potential  $\mu$  in (50) is a constant that ensures the correct normalization,  $\int dz \rho(z) = N/A$ , where  $A$  is the lateral system size (i.e. the area perpendicular to the  $z$ -direction). That the value of the chemical potential  $\mu$  controls the number of particles in the system is universal. However, the mathematical formulation in the grand ensemble, where the particle number in the system can fluctuate, is very different from the present canonical treatment. (Some basics of the grand canonical description, as used in reference [14], are described below in section 2.6.)

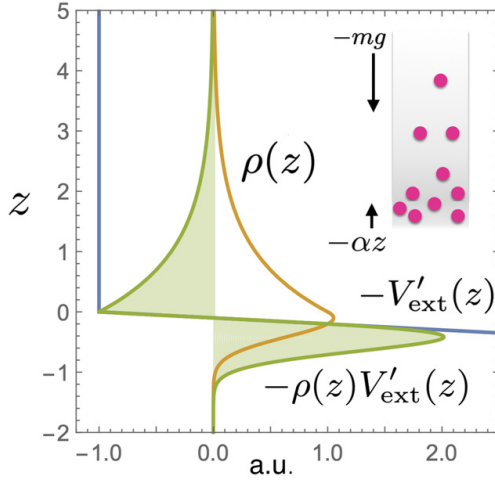
The general expression for the total external force (48) together with the specific density profile (50) gives

$$\mathbf{F}_{\text{ext}}^0 = - \frac{A \mathbf{e}_z}{\Lambda^3} \int_{-\infty}^{\infty} dz e^{-\beta(V_{\text{ext}}(z) - \mu)} V'_{\text{ext}}(z) \quad (51)$$

$$= \frac{A \mathbf{e}_z}{\Lambda^3 \beta} \left[ e^{-\beta(V_{\text{ext}}(z) - \mu)} \right]_{-\infty}^{\infty} \quad (52)$$

$$= 0, \quad (53)$$

where  $\mathbf{e}_z$  is the unit vector pointing into the positive  $z$ -direction and the prime denotes differentiation with respect to the argument, hence  $\nabla V_{\text{ext}}(\mathbf{r}) = V'_{\text{ext}}(z) \mathbf{e}_z$ . The integrand in (51) is a total differential,  $d e^{-\beta(V_{\text{ext}} - \mu)} / dz$ , which upon integration gives



**Figure 3.** Illustration of sedimentation of a fluid against a lower soft wall represented by a harmonic potential. The total external potential is  $V_{\text{ext}}(z) = mgz + \Theta(-z)\alpha z^2/2$ . The resulting external force field is  $-V'_{\text{ext}}(z) \equiv -\partial V_{\text{ext}}(z)/\partial z = -mg - \Theta(-z)\alpha z$  (blue line). The direction of both force contributions is indicated in the inset (arrows), where pink dots represent particles. In the main plot the density profile  $\rho(z)$  (amber line) decays for large and for small values of  $z$ . The external force density is the product  $-\rho(z)V'_{\text{ext}}(z)$  (green line). The total external force (per unit area) is the integral  $-\int dz V'_{\text{ext}}(z)\rho(z) = 0$ ; note that the shaded green areas cancel each other. Representative values of the parameters are chosen; the unit of length is the sedimentation length  $k_B T/(mg)$  and all energies are scaled with  $k_B T$ .

(52); for (53) we have exploited that for  $z \rightarrow \pm\infty$  the external potential  $V_{\text{ext}} \rightarrow \infty$ , leading to vanishing Boltzmann factor. We have hence shown explicitly the vanishing of the total external force acting on a bounded ideal gas in thermal equilibrium under gravity. Figure 3 illustrates the density profile  $\rho(z)$  and the force density profile  $-V'_{\text{ext}}(z)\rho(z)$  for representative values of the parameters.

We briefly sketch the effect of interparticle interactions. On a formal level, and returning to the general case of arbitrary form of  $V_{\text{ext}}(\mathbf{r})$ , the density profile is given by a modified form of (50), which reads

$$\rho(\mathbf{r}) = \Lambda^{-3} e^{-\beta(V_{\text{ext}}(\mathbf{r}) - \mu) + c_1(\mathbf{r})}, \quad (54)$$

where the so-called one-body direct correlation function [11, 13]  $c_1(\mathbf{r})$  contains the effects of the interparticle interactions. The total interparticle force density is then given by

$$\mathbf{F}_{\text{int}}^0 = k_B T \int d\mathbf{r} \rho(\mathbf{r}) \nabla c_1(\mathbf{r}) = 0. \quad (55)$$

Here the vanishing of the total internal force can be viewed as a consequence of Newton's third law *actio equals reactio*; see reference [14] for the derivation. Note the formal similarity of the total external and total intrinsic force Noether sum rules, cf (48) and (55). The no-bootstrap theorem (55) holds beyond equilibrium, as shown in reference [14], and it hence debunks any swamp escape myths.

An alternative derivation of (55) rests on the Noether invariance of the free energy, where the latter is constructed to be a functional of the density profile; we refer the reader to reference [14] for a description of these considerations and comment briefly on the embedding into the frameworks of classical density functional theory and power functional theory below in section 2.6.

### 2.5. Relationship to correlation functions

Global identities, such as the sum rules of vanishing external force (48) and of vanishing internal force (55), can be used as a starting point to obtain position-resolved identities. Functional differentiation with respect to an appropriate field creates dependence on position. Integrating over these additional variables (or 'root points' [13]) then yields novel global identities. While we refer the reader to reference [14] for this treatment, we wish to demonstrate here the direct derivation of such global identities.

We stick to the canonical ensemble and as a specific case return to our initial example of a thermal average, i.e. the global external potential energy  $U_{\text{ext}}$ , as equivalently expressed in compact notation (10) or the explicitly written out phase space integral (11). Let us shift! The external energy in the new system is then given by

$$U_{\text{ext}}(\epsilon) = \text{Tr}_N \frac{e^{-\beta H(\epsilon)}}{Z_N(\epsilon)} \sum_{i=1}^N V_{\text{ext}}(\mathbf{r}_i - \epsilon). \quad (56)$$

We Taylor expand to first order,

$$U_{\text{ext}}(\epsilon) = U_{\text{ext}} + \left. \frac{\partial U_{\text{ext}}(\epsilon)}{\partial \epsilon} \right|_{\epsilon=0} \cdot \epsilon. \quad (57)$$

Here the derivative of (56) can be calculated via the product rule as

$$\begin{aligned} \left. \frac{\partial U_{\text{ext}}(\epsilon)}{\partial \epsilon} \right|_{\epsilon=0} &= \text{Tr} \frac{\partial \Psi(\epsilon)}{\partial \epsilon} \sum_{i=1}^N V_{\text{ext}}(\mathbf{r}_i) \\ &\quad - \text{Tr} \Psi \sum_{i=1}^N \nabla_i V_{\text{ext}}(\mathbf{r}_i). \end{aligned} \quad (58)$$

We can recognize the second term as the average external force, which we have proven to vanish, cf (27). The first term in (58) requires carrying out the derivative of  $\Psi(\epsilon)$  with respect to the displacement  $\epsilon$ , which yields

$$\left. \frac{\partial U_{\text{ext}}(\epsilon)}{\partial \epsilon} \right|_{\epsilon=0} = -\text{Tr} \Psi \sum_{i=1}^N \beta V_{\text{ext}}(\mathbf{r}_i) \sum_{j=1}^N \nabla_j V_{\text{ext}}(\mathbf{r}_j). \quad (59)$$

Here an additional term, generated by the derivative, vanishes:  $-\beta U_{\text{ext}} \mathbf{F}_{\text{ext}}^0 = 0$ , again due to (27).

Clearly (59) is the correlator of the global external potential energy and the global external force. Using the by now familiar invariance argument, we argue that the value of  $U_{\text{ext}}(\epsilon)$  is an invariant under the displacement, and that hence the first order

term in (57) needs to be zero. As  $\epsilon$  is arbitrary, we conclude

$$-\text{Tr} \Psi \sum_{i=1}^N V_{\text{ext}}(\mathbf{r}_i) \sum_{j=1}^N \nabla_j V_{\text{ext}}(\mathbf{r}_j) = 0, \quad (60)$$

where we have divided by  $\beta$ . Hence the global external potential,  $\sum_i V_{\text{ext}}(\mathbf{r}_i)$ , and the global external force,  $-\sum_j \nabla_j V_{\text{ext}}(\mathbf{r}_j)$ , are uncorrelated with each other. The sum rule (60) is derived in reference [14] via the route of integration over free position variables (root points), cf (5) in reference [14] for the order  $n = 2$  of the sum rule hierarchy. An important distinction in the presentation though lies in the choice of ensemble, which is an issue to which we turn in the next subsection.

As a final comment, when applied to the above example of sedimentation against a lower harmonic wall, (60) can be explicitly verified by carrying out the  $z$ -integral, which yields  $-A \int_{-\infty}^{\infty} dz \rho(z) V_{\text{ext}}(z) V'_{\text{ext}}(z) = 0$ .

### 2.6. Density functional and power functional

In all of the above, we have described the thermal system on the basis of the canonical ensemble, as specified by the classical phase space, the probability distribution (5) and the canonical partition sum (6). Hence the system is coupled to a heat bath at temperature  $T$ , where the value of  $T$  determines the mean energy  $E$  in the system, cf the form of  $E$  as an expectation value (7). The system is thermally open, and hence energy fluctuations occur between system and bath.

Corresponding fluctuations in particle number  $N$  can be implemented in the grand canonical ensemble where the system is furthermore coupled to a particle bath. The particle bath sets the value of the chemical potential  $\mu$ , which then determines the average number of particles  $\bar{N}$  in the system. (This mechanism is analogous to the relationship of  $T$  and  $E$  described above.) Although the grand canonical formalism poses this additional level of abstraction, and the bare formulae increase somewhat in complexity due to the average over  $N$ , in typical theoretical developments this framework is significantly more powerful and more straightforward to use. (There is no need having to implement  $N = \text{const}$ , which in practice can be awkward.) We briefly sketch the essentials of the grand ensemble as they underlie reference [14].

The grand canonical ensemble consists of the microstates given by phase space points of  $N$  particles, with  $N$  being a non-negative integer, which is treated as a random variable. The corresponding probability distribution is

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N, N) = \frac{e^{-\beta(H - \mu N)}}{\Xi}, \quad (61)$$

where the grand partition sum is given by

$$\Xi = \text{Tr} e^{-\beta(H - \mu N)}, \quad (62)$$

with the grand canonical trace operation defined by

$$\text{Tr} = \sum_{N=0}^{\infty} \text{Tr}_N \quad (63)$$

$$= \sum_{N=0}^{\infty} \frac{1}{h^{3N} N!} \int d\mathbf{r}_1 \dots d\mathbf{r}_N d\mathbf{p}_1 \dots d\mathbf{p}_N, \quad (64)$$

where we have obtained (64) by using the explicit form (8) for the canonical trace. The thermodynamic potential which is fundamental for the grand ensemble is the grand potential (also referred to as the grand canonical free energy) and it is given by

$$\Omega = -k_B T \ln \Xi, \quad (65)$$

with the grand partition sum  $\Xi$  according to (62). Note the strong formal analogy with the corresponding canonical expressions for: the probability distribution (5) with (61); the partition sum (6), i.e.  $Z_N = \text{Tr}_N e^{-\beta H}$ , with (62); the trace (8) with (63); and the free energy (14) with (65).

Despite the system being open to particle exchange, Noether's reasoning continues to hold [14]. Briefly, the grand potential is a functional of the external potential,  $\Omega[V_{\text{ext}}]$  (we suppress the dependence on the thermodynamic parameters  $\mu, T$ ), and  $\Omega[V_{\text{ext}}]$  is invariant under spatial displacements according to (15). As a consequence, the sum rule of vanishing external force (13) emerges, expressed in the form (27) with  $\text{Tr}_N$  replaced by  $\text{Tr}$ , as is appropriate for the open system.

Why is the functional point of view important? In what we have presented above it had played the role of adding abstraction and re-deriving results that we could obtain via more elementary arguments. The importance of the variational formulation stems from two sources, one being that it provides a mechanism for the generation of correlation functions via functional differentiation, in extension of the generation of the density profile via (41), see e.g. references [13] for a comprehensive account. The second point lies in the variational principle itself which formulates the many-body problem in a way that allows to systematically introduce approximations and make much headway in identifying and studying physical mechanisms in complex, coupled many-body problems. While giving a self-contained overview of these concepts is beyond the scope of the present contribution (see reference [31] for a recent account), we wish to briefly describe certain central points, to—hopefully—provide motivation for further study.

We hence sketch the two variational principles as they are relevant for equilibrium (classical density functional theory) and for the dynamics (power functional theory); these form the basis of reference [14]. Classical density functional theory is based on treating the density profile  $\rho(\mathbf{r})$ , rather than the external potential  $V_{\text{ext}}(\mathbf{r})$ , as the fundamental variational field. The grand potential, when viewed as a density functional [11, 12], has the form

$$\Omega[\rho] = F[\rho] + \int d\mathbf{r} \rho(\mathbf{r}) (V_{\text{ext}}(\mathbf{r}) - \mu), \quad (66)$$

where  $F[\rho]$  is the intrinsic Helmholtz free energy functional. Crucially,  $F[\rho]$  is independent of the external potential, which

features solely in the second term in (66). Here  $\rho(\mathbf{r})$  is conceptually treated as a variable; its true form as the equilibrium density profile is that which minimizes  $\Omega[\rho]$  and for which hence the functional derivative vanishes,

$$\frac{\delta\Omega[\rho]}{\delta\rho(\mathbf{r})} = 0 \quad (\text{min}). \quad (67)$$

Inserting the split form (66) of the grand potential into the minimization condition (67) and using the splitting into ideal gas and excess (over ideal gas) free energy contributions,  $F[\rho] = k_B T \int d\mathbf{r} \rho(\mathbf{r}) [\ln(\rho(\mathbf{r})\Lambda^3) - 1] + F_{\text{exc}}[\rho]$ , yields upon exponentiating the modified barometric law (54). Here the one-body direct correlation function  $c_1(\mathbf{r})$  is identified as the functional derivative of the excess free energy functional, i.e.  $c_1(\mathbf{r}) = -\beta\delta F_{\text{exc}}[\rho]/\delta\rho(\mathbf{r})$ . As the functional dependence on the density profile persists upon building the derivative, i.e. in more explicit notation  $c_1(\mathbf{r}, [\rho])$ , equation (54) constitutes a self-consistency condition for the determination of the equilibrium density profile; determining the solution thereof requires to have an approximation for  $F_{\text{exc}}[\rho]$  and typically involves numerical work.

Power functional theory generalizes the variational concept of working on the level of one-body correlation functions to nonequilibrium. For overdamped Brownian motion, as is a simple model for the description for the temporal behaviour of mesoscopic particles that are suspended in a liquid, the free power is a functional of both the time-dependent density profile  $\rho(\mathbf{r}, t)$  and of the locally resolved current distribution  $\mathbf{J}(\mathbf{r}, t)$ , where  $t$  indicates time. The power functional has the form

$$\begin{aligned} R_t[\rho, \mathbf{J}] &= \dot{F}[\rho] + P_t[\rho, \mathbf{J}] \\ &- \int d\mathbf{r} (\mathbf{J}(\mathbf{r}, t) \cdot \mathbf{f}_{\text{ext}}(\mathbf{r}, t) - \rho(\mathbf{r}, t) \dot{V}_{\text{ext}}(\mathbf{r}, t)), \end{aligned} \quad (68)$$

where  $\dot{F}[\rho]$  is the time derivative of the intrinsic free energy functional,  $P_t[\rho, \mathbf{J}]$  consists of an ideal gas and a superadiabatic part, where the latter arises from the internal interactions in the nonequilibrium situation,  $\mathbf{f}_{\text{ext}}(\mathbf{r}, t)$  is a time-dependent external one-body force field, which in general consists of a (conservative) gradient term  $-\nabla V_{\text{ext}}(\mathbf{r}, t)$  and an additional rotational (non-gradient, non-conservative) contribution, and  $\dot{V}_{\text{ext}}(\mathbf{r}, t)$  is the time derivative of the external potential. The density profile and the current distribution are linked by the continuity equation,  $\partial\rho(\mathbf{r}, t) = -\nabla \cdot \mathbf{J}(\mathbf{r}, t)$ , which is sharply resolved on the microscopic scale. The dynamic variational principle states that  $R_t[\rho, \mathbf{J}]$  is minimized, at time  $t$ , by the physically realized current,

$$\frac{\delta R_t[\rho, \mathbf{J}]}{\delta \mathbf{J}(\mathbf{r}, t)} = 0 \quad (\text{min}). \quad (69)$$

Inserting the splitting (68) of the total free power into the minimization condition (69) yields the formally exact force density relationship,

$$\gamma \mathbf{J}(\mathbf{r}, t) = -k_B T \nabla \rho(\mathbf{r}, t) + \mathbf{F}_{\text{int}}(\mathbf{r}, t) + \rho(\mathbf{r}, t) \mathbf{f}_{\text{ext}}(\mathbf{r}, t), \quad (70)$$

where  $\gamma$  is the friction constant of the overdamped motion, such that the left-hand side constitutes the (negative) friction force density at position  $\mathbf{r}$  and time  $t$ . The right-hand side of (70) consists of an ideal, an internal and an external driving contribution, with  $\mathbf{F}_{\text{int}}(\mathbf{r}, t)$  being the internal force density distribution, as it arises from the effect of all interparticle interactions that act on a given particle at position  $\mathbf{r}$  and time  $t$ . The internal force density  $\mathbf{F}_{\text{int}}(\mathbf{r}, t)$  consists of an adiabatic contribution, which follows from the excess free energy functional via  $\mathbf{F}_{\text{ad}}(\mathbf{r}, t) = -\rho(\mathbf{r}, t) \nabla \delta F[\rho]/\delta\rho(\mathbf{r}, t)$  and an additional genuine nonequilibrium contribution, i.e. the superadiabatic force density,  $\mathbf{F}_{\text{sup}}(\mathbf{r}, t)$ . Honoring its functional dependence on the kinematic fields  $\rho(\mathbf{r}, t)$  and  $\mathbf{J}(\mathbf{r}, t)$  forms the basis for much recent work in nonequilibrium statistical mechanics based on the power functional concept. See reference [31] for an overview.

As a comment on terminology, we note that sometimes the term Euler–Lagrange equation is applied generically to refer to the vanishing of the first functional derivative of the given variational problem, i.e. equation (67) for the case of DFT and equation (69) for PFT, which respectively turn into the explicit forms (54) and (70). This terminology is different from the also frequent use of referring specifically to the Euler–Lagrange equations of motion of classical mechanics, as they result from Hamilton's principle, i.e. the stationarity of the action functional (see e.g. the appendix of reference [31] for a description of the functional methods involved).

### 3. Conclusions

In conclusion, we have demonstrated on an elementary level how fundamental symmetries in statistical mechanics lead to exact statements (sum rules) about average forces when considering translations. These considerations also apply to torques when considering rotations [14]. We have based our presentation on the canonical ensemble, as is relevant in a variety of contexts [53–58]. While the canonical ensemble avoids the complexity of particle number fluctuations that occur grand canonically, nevertheless an open system is retained with respect to energy exchange with a heat bath. As we have shown, treating such fluctuating systems is well permissible on the basis of Noetherian arguments. The arguably simplest Noether sum rule is that of vanishing average total external force in thermal equilibrium. As an application we have presented the case of a fluid confined inside of a container and subject to the effect of gravity. While we have selected this example for its relative simplicity, the influence of gravity on mesoscopic soft matter is also a topic of relevance for studying e.g. complex phase behaviour in colloidal mixtures; see e.g. reference [60] for recent work that addresses colloidal liquid crystals. Our derivations imply that the symmetry operation is applied to the entire system. Here the system must be enclosed by an external potential that represents confinement by e.g. walls. The shift then applies also to these walls. In cases where system boundaries are open (as can be suitable for a periodically repeated system like that shown in figure 2), Noether's theorem remains applicable upon taking account of additional

boundary terms, see reference [14] for a detailed discussion of such treatment.

In the presented considerations, we have started on the basis of arguably the most fundamental statistical mechanical object, i.e. the partition sum, as it enters the elementary definition of the (here canonical) free energy. Investigating invariance properties of further statistical objects, such as the global external energy, is also worthwhile, as then Noether's reasoning leads to the correlator identity (60) of vanishing correlation between global external force and global external potential. Investigating the outcome of invariance applied in this way constitutes an interesting task for future work.

Statistical mechanical derivations often rely on very similar reasoning; reference [14] gives an overview. A particularly insightful example is the work by Bryk *et al* on hard sphere fluids in contact with curved substrates [64]. These authors derive a contact sum rule of the hard sphere fluid against a hard curved wall. Their argumentation rests on the observation that the force that is necessary to move the wall by an amount  $\epsilon$  is balanced by the presence of the fluid. The authors then succeed in relating this force to the value of the density profile close to the wall. Closely related work was carried out for the shape dependence of free energies [65]. Further studies that are related to Noether's theorem were aimed at broken symmetries [66] and emerging Goldstone modes [67–69].

The general form of Noether's theorem applies to variational calculus, and statistical mechanics falls well into this realm. We have spelled out the connections explicitly, such as the canonical free energy being viewed as a functional of the external potential [13]. Notably only elementary statistical objects such as the partition sum are required. We have also described two more advanced variational theories. Classical density functional theory [11–13] allows to view the grand potential as a functional of the one-body density distribution. A formally exact minimization principle then reformulates the physics of system in thermal (and chemical) equilibrium. The dynamic variational principle of power functional theory [30, 31] consists of instantaneous minimization with respect to the time- and position-resolved current distribution. Together with the continuity equation, a formally closed one-body reformulation of the dynamics of the underlying many-body system is achieved.

Both density functional theory and power functional theory can be viewed as systematic approaches to coarse-graining the many-body problem to the level of one-body correlation functions. In the static case, the correlation functions hence depend on position alone, in the dynamics case the dependence is on position and on time. Crucially, a microscopically sharp description is formally retained, which is important for the description of correlations on the particle (i.e. molecular or colloidal) level. One of the most important features of these theories is the identification of a universal intrinsic functional that contains the coupled effects of the interparticle interac-

tions, but is independent of the external forces that act on the system.

A wealth of productive research has been devoted to constructing powerful approximations for free energy functionals for specific model systems. In the context of liquids the important case of the hard sphere fluid is treated with excellent accuracy within Rosenfeld's fundamental measure theory [17, 18], see e.g. reference [61] for a quantitative assessment of the quality of theoretical density profiles against simulation data. Notable recent progress to incorporate short-ranged attraction into density functional theory is due to Tschopp, Brader and their co-workers [62, 63], who systematically addressed and exploited two-body correlations.

Despite power functional theory [30, 31] being significantly younger than density functional theory, its usefulness has been amply demonstrated, both for formal work as well as for practical solution of physical problems and the discovery of novel fundamental mechanisms. The reformulation on the basis of the velocity gradient [39], instead of the current distribution, allowed to identify and to study structural forces [40, 41] in driven systems that are governed by overdamped Brownian dynamics. The splitting of the total internal force field into flow and structural contributions is fundamental to understanding the emerging effects in microscopically inhomogeneous flows [41]. Active Brownian particles, as a model for self-propelled colloids (see e.g. [46–48]), are well suited for the application of power functional theory. The general framework [33, 34] for active systems was shown to physically explain and quantitatively predict the motility-induced phase separation that occurs in such systems at high enough levels of driving [35–37]. Interfacial properties such as polarization [38] and surface tension [35] were systematically studied.

The dynamical sum rules for forces and correlation functions presented in reference [14] offer great potential for systematic progress in the description of complex temporal behaviour, including memory [44, 45]. The nonequilibrium rules play a similar role than fundamental equilibrium sum rules such e.g. the Lovett–Mou–Buff–Wertheim equation [70, 71]. The section on 'methods' in reference [14] gives a detailed description of the relationship of the equilibrium Noether sum rules to such classical results from the liquid state literature. Together with the nonequilibrium Ornstein–Zernike relations [42, 43] the dynamical sum rules provide fertile ground for making progress in nonequilibrium many-body physics; see also the recent study of the relevance of invariance in inhomogeneous dense liquids [72] and of the role of fluctuations when going to effects that are higher than linear in the displacement [73]. Hence the fundamental character of Emmy Noether's work will surely continue to prove its worth in the future.

## Acknowledgments

We thank Daniel de las Heras, Roland Roth, Gerhard Kahl, and Bob Evans for useful comments and discussions. This work is supported by the German Research Foundation (DFG) via Project Number 436306241.



**Data availability statement**

All data that support the findings of this study are included within the article.

**ORCID iDs**

Sophie Hermann  <https://orcid.org/0000-0002-4012-9170>

Matthias Schmidt  <https://orcid.org/0000-0002-5015-2972>

**References**

- [1] Noether E 1918 Invariante Variationsprobleme *Nachr. d. Königl. Gesellsch. d. Wiss. zu Göttingen Math. Phys. Klasse* p 235]
- Tavel M A 1971 Invariant variation problems *Transp. Theor. Stat. Phys.* **1** 186 (Engl. transl.)
- For a version in modern typesetting see Wang F Y 2018 arXiv:physics/0503066v3
- [2] For a description of many insightful and pedagogical examples and applications, see: Neuschwander D E 2011 *Emmy Noether's Wonderful Theorem* (Baltimore, MD: John Hopkins University Press)
- [3] Byers N 1998 E Noether's discovery of the deep connection between symmetries and conservation laws (arXiv:physics/9807044)
- [4] Lezcano A G and de Oca A C M 2018 A stochastic version of the Noether theorem *Found. Phys.* **48** 726
- [5] Baez J C and Fong B 2013 A Noether theorem for Markov processes *J. Math. Phys.* **54** 013301
- [6] Marvian I and Spekkens R W 2014 Extending Noether's theorem by quantifying the asymmetry of quantum states *Nat. Commun.* **5** 3821
- [7] Sasa S-i and Yokokura Y 2016 Thermodynamic entropy as a Noether invariant *Phys. Rev. Lett.* **116** 140601
- [8] Sasa S, Sugiura S and Yokokura Y 2019 Thermodynamical path integral and emergent symmetry *Phys. Rev. E* **99** 022109
- [9] Minami Y and Sasa S-I 2020 Thermodynamic entropy as a Noether invariant in a Langevin equation *J. Stat. Mech.* 013213
- [10] Mermin N D 1965 Thermal properties of the inhomogeneous electron gas *Phys. Rev.* **137** A1441
- [11] Evans R 1979 The nature of the liquid–vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids *Adv. Phys.* **28** 143
- [12] Evans R 1992 Density functionals in the theory nonuniform fluids *Fundamentals of Inhomogeneous Fluids* ed D Henderson (New York: Dekker)
- [13] Hansen J P and McDonald I R 2013 *Theory of Simple Liquids* 4th edn (London: Academic)
- [14] Hermann S and Schmidt M 2021 Noether's theorem in statistical mechanics *Commun. Phys.* **4** 176
- [15] Tarazona P and Evans R 1983 On the validity of certain integro-differential equations for the density-orientation profile of molecular fluid interfaces *Chem. Phys. Lett.* **97** 279
- [16] Evans R, Oettel M, Roth R and Kahl G 2016 New developments in classical density functional theory *J. Phys.: Condens. Matter* **28** 240401
- [17] Rosenfeld Y 1989 Free-energy model for the inhomogeneous hard-sphere fluid mixture and density-functional theory of freezing *Phys. Rev. Lett.* **63** 980
- [18] Roth R 2010 Fundamental measure theory for hard-sphere mixtures: a review *J. Phys.: Condens. Matter* **22** 063102
- [19] Lin S-C, Oettel M, Häring J M, Haussmann R, Fuchs M and Kahl G 2021 The direct correlation function of a crystalline solid *Phys. Rev. Lett.* **127** 085501
- [20] Levesque M, Vuilleumier R and Borgis D 2012 Scalar fundamental measure theory for hard spheres in three dimensions: application to hydrophobic solvation *J. Chem. Phys.* **137** 034115
- [21] Jeanmairet G, Levesque M and Borgis D 2013 Molecular density functional theory of water describing hydrophobicity at short and long length scales *J. Chem. Phys.* **139** 154101
- [22] Evans R, Stewart M C and Wilding N B 2019 A unified description of hydrophilic and superhydrophobic surfaces in terms of the wetting and drying transitions of liquids *Proc. Natl Acad. Sci. USA* **116** 23901
- [23] Evans R and Wilding N B 2015 Quantifying density fluctuations in water at a hydrophobic surface: evidence for critical drying *Phys. Rev. Lett.* **115** 016103
- [24] Evans R, Stewart M C and Wilding N B 2016 Critical drying of liquids *Phys. Rev. Lett.* **117** 176102
- [25] Chacko B, Evans R and Archer A J 2017 Solvent fluctuations around solvophobic, solvophilic, and patchy nanostructures and the accompanying solvent mediated interactions *J. Chem. Phys.* **146** 124703
- [26] Martín-Jiménez D, Chacón E, Tarazona P and García R 2016 Atomically resolved three-dimensional structures of electrolyte aqueous solutions near a solid surface *Nat. Commun.* **7** 12164
- [27] Hernández-Muñoz J, Chacón E and Tarazona P 2019 Density functional analysis of atomic force microscopy in a dense fluid *J. Chem. Phys.* **151** 034701
- [28] Muscatello J, Chacón E, Tarazona P and Bresme F 2017 Deconstructing temperature gradients across fluid interfaces: the structural origin of the thermal resistance of liquid–vapor interfaces *Phys. Rev. Lett.* **119** 045901
- [29] Eckert T, Stuhlmüller N C X, Sammüller F and Schmidt M 2020 Fluctuation profiles in inhomogeneous fluids *Phys. Rev. Lett.* **125** 268004
- [30] Schmidt M and Brader J M 2013 Power functional theory for Brownian dynamics *J. Chem. Phys.* **138** 214101
- [31] Schmidt M 2022 Power functional theory for many-body dynamics *Rev. Mod. Phys.* **94** 015007
- [32] Fortini A, de las Heras D, Brader J M and Schmidt M 2014 Superadiabatic forces in Brownian many-body dynamics *Phys. Rev. Lett.* **113** 167801
- [33] Krinninger P, Schmidt M and Brader J M 2016 Nonequilibrium phase behavior from minimization of free power dissipation *Phys. Rev. Lett.* **117** 208003
- [34] Krinninger P and Schmidt M 2019 Power functional theory for active Brownian particles: general formulation and power sum rules *J. Chem. Phys.* **150** 074112
- [35] Hermann S, de las Heras D and Schmidt M 2019 Non-negative interfacial tension in phase-separated active Brownian particles *Phys. Rev. Lett.* **123** 268002
- [36] Hermann S, Krinninger P, de las Heras D and Schmidt M 2019 Phase coexistence of active Brownian particles *Phys. Rev. E* **100** 052604
- [37] Hermann S, de las Heras D and Schmidt M 2020 Phase separation of active Brownian particles in two dimensions: anything for a quiet life *Mol. Phys.* **119** e1902585
- [38] Hermann S and Schmidt M 2020 Active interface polarization as a state function *Phys. Rev. Res.* **2** 022003
- [39] de las Heras D and Schmidt M 2018 Velocity gradient power functional for Brownian dynamics *Phys. Rev. Lett.* **120** 028001
- [40] Stuhlmüller N C X, Eckert T, de las Heras D and Schmidt M 2018 Structural nonequilibrium forces in driven colloidal systems *Phys. Rev. Lett.* **121** 098002

- [41] de las Heras D and Schmidt M 2020 Flow and structure in nonequilibrium Brownian many-body systems *Phys. Rev. Lett.* **125** 018001
- [42] Brader J M and Schmidt M 2013 Nonequilibrium Ornstein–Zernike relation for Brownian many-body dynamics *J. Chem. Phys.* **139** 104108
- [43] Brader J M and Schmidt M 2014 Dynamic correlations in Brownian many-body systems *J. Chem. Phys.* **140** 034104
- [44] Treffenstädt L L and Schmidt M 2020 Memory-induced motion reversal in Brownian liquids *Soft Matter* **16** 1518
- [45] Treffenstädt L L and Schmidt M 2021 Universality in driven and equilibrium hard sphere liquid dynamics *Phys. Rev. Lett.* **126** 058002
- [46] Farage T F F, Krinninger P and Brader J M 2015 Effective interactions in active Brownian suspensions *Phys. Rev. E* **91** 042310
- [47] Paliwal S, Rodenburg J, van Roij R and Dijkstra M 2018 Chemical potential in active systems: predicting phase equilibrium from bulk equations of state? *New J. Phys.* **20** 015003
- [48] Paliwal S, Prymidis V, Filion L and Dijkstra M 2017 Nonequilibrium surface tension of the vapour–liquid interface of active Lennard–Jones particles *J. Chem. Phys.* **147** 084902
- [49] Takatori S C, Yan W and Brady J F 2014 Swim pressure: stress generation in active matter *Phys. Rev. Lett.* **113** 028103
- [50] Hermann S and Schmidt M 2018 Active ideal sedimentation: exact two-dimensional steady states *Soft Matter* **14** 1614
- [51] Söker N A, Auschra S, Holubec V, Kroy K and Cichos F 2021 How activity landscapes polarize microswimmers without alignment forces *Phys. Rev. Lett.* **126** 228001
- [52] Auschra S, Holubec V, Söker N A, Cichos F and Kroy K 2021 Polarization-density patterns of active particles in motility gradients *Phys. Rev. E* **103** 062601
- [53] González A, White J A, Román F L, Velasco S and Evans R 1997 Density functional theory for small systems: hard spheres in a closed spherical cavity *Phys. Rev. Lett.* **79** 2466
- [54] González A, White J A, Román F L and Evans R 1998 How the structure of a confined fluid depends on the ensemble: hard spheres in a spherical cavity *J. Chem. Phys.* **109** 3637
- [55] de las Heras D and Schmidt M 2014 Full canonical information from grand-potential density-functional theory *Phys. Rev. Lett.* **113** 238304
- [56] de las Heras D, Brader J M, Fortini A and Schmidt M 2016 Particle conservation in dynamical density functional theory *J. Phys.: Condens. Matter* **28** 244024
- [57] Schindler T, Wittmann R and Brader J M 2019 Particle-conserving dynamics on the single-particle level *Phys. Rev. E* **99** 012605
- [58] Wittmann R, Löwen H and Brader J M 2021 Order-preserving dynamics in one dimension—single-file diffusion and caging from the perspective of dynamical density functional theory *Mol. Phys.* **115** e1867250
- [59] Zwanzig R 2001 *Nonequilibrium Statistical Mechanics* (Oxford: Oxford University Press) see ch 10
- [60] Eckert T, Schmidt M and de las Heras D 2021 Gravity-induced phase phenomena in plate-rod colloidal mixtures *Commun. Phys.* **4** 202
- [61] Davidchack R L, Laird B B and Roth R 2016 Hard spheres at a planar hard wall: simulations and density functional theory *Condens. Matter Phys.* **19** 23001
- [62] Tschopp S M, Vuijk H D, Sharma A and Brader J M 2020 Mean-field theory of inhomogeneous fluids *Phys. Rev. E* **102** 042140
- [63] Tschopp S M and Brader J M 2021 Fundamental measure theory of inhomogeneous two-body correlation functions *Phys. Rev. E* **103** 042103
- [64] Bryk P, Roth R, Mecke K R and Dietrich S 2003 Hard-sphere fluids in contact with curved substrates *Phys. Rev. E* **68** 031602
- [65] König P-M, Roth R and Mecke K R 2004 Morphological thermodynamics of fluids: shape dependence of free energies *Phys. Rev. Lett.* **93** 160601
- [66] Baus M 1984 Broken symmetry and invariance properties of classical fluids *Mol. Phys.* **51** 211
- [67] Walz C and Fuchs M 2010 Displacement field and elastic constants in nonideal crystals *Phys. Rev. B* **81** 134110
- [68] Parry A O, Rascón C and Evans R 2016 The local structure factor near an interface; beyond extended capillary-wave models *J. Phys.: Condens. Matter* **28** 244013
- [69] Parry A O and Rascón C 2019 The Goldstone mode and resonances in the fluid interfacial region *Nat. Phys.* **15** 287
- [70] Lovett R, Mou C Y and Buff F P 1976 The structure of the liquid–vapor interface *J. Chem. Phys.* **65** 570
- [71] Wertheim M S 1976 Correlations in the liquid–vapor interface *J. Chem. Phys.* **65** 2377
- [72] Tschopp S M, Sammüller F, Hermann S, Schmidt M and Brader J M 2022 Force density functional theory in- and out-of-equilibrium (arXiv:2203.01795) (unpublished)
- [73] Hermann S and Schmidt M 2022 Variance of fluctuations from Noether invariance (arXiv:2203.15654) (unpublished)

**Force density functional theory in- and out-of-equilibrium**Salomé M. Tschopp<sup>1,\*</sup>, Florian Sammüller<sup>2</sup>, Sophie Hermann<sup>2</sup>, Matthias Schmidt<sup>2,†</sup> and Joseph M. Brader<sup>1,‡</sup><sup>1</sup>*Department of Physics, University of Fribourg, CH-1700 Fribourg, Switzerland*<sup>2</sup>*Theoretische Physik II, Physikalisches Institut, Universität Bayreuth, D-95447 Bayreuth, Germany*

(Received 4 March 2022; accepted 6 June 2022; published 12 July 2022)

When a fluid is subject to an external field, as is the case near an interface or under spatial confinement, then the density becomes spatially inhomogeneous. Although the one-body density provides much useful information, a higher level of resolution is provided by the two-body correlations. These give a statistical description of the internal microstructure of the fluid and enable calculation of the average interparticle force, which plays an essential role in determining both the equilibrium and dynamic properties of interacting fluids. We present a theoretical framework for the description of inhomogeneous (classical) many-body systems, based explicitly on the two-body correlation functions. By consideration of local Noether-invariance against spatial distortion of the system we demonstrate the fundamental status of the Yvon-Born-Green (YBG) equation as a local force-balance within the fluid. Using the inhomogeneous Ornstein-Zernike equation we show that the two-body correlations are density functionals and, thus, that the average interparticle force entering the YBG equation is also a functional of the one-body density. The force-based theory we develop provides an alternative to standard density functional theory for the study of inhomogeneous systems both in- and out-of-equilibrium. We compare force-based density profiles to the results of the standard potential-based (dynamical) density functional theory. In-equilibrium, we confirm both analytically and numerically that the standard approach yields profiles that are consistent with the compressibility pressure, whereas the force-density functional gives profiles consistent with the virial pressure. For both approaches we explicitly prove the hard-wall contact theorem that connects the value of the density profile at the hard-wall with the bulk pressure. The structure of the theory offers deep insights into the nature of correlation in dense and inhomogeneous systems.

DOI: [10.1103/PhysRevE.106.014115](https://doi.org/10.1103/PhysRevE.106.014115)**I. INTRODUCTION**

The analysis of spatial inhomogeneity is a primary means to characterize a wide range of self-organized and complex states of matter [1]. Representative examples of systems and effects with inherent position-dependence cover a broad range of soft matter [2,3], including hydrophobic solvation in complex environments [4], desorption of water at short and long length scales [5], liquids at hydrophobic and hydrophilic substrates characterized by wetting and drying surface phase diagrams [6,7], critical drying of liquids [9], solvent-mediated forces between nanoscopic solutes [10], electrolyte aqueous solutions near a solid surface [11], layering in liquids [12], the structure of liquid-vapor interfaces [13,14], and locally resolved density fluctuations [6–10].

Obtaining a systematic understanding of the physics that emerges in such systems can be achieved by using microscopically resolved correlation functions. In particular the one-body density profile captures a broad spectrum of behaviours, from strong oscillations in dense liquids, where molecular packing effects dominate [4,8,11–13], to pronounced drying layers near hydrophobic substrates when

approaching bulk evaporation [5–7,9,10]. Effects such as these can be induced by walls or other external influence, which typically is modeled by a position-dependent external potential  $V_{\text{ext}}(\mathbf{r})$ . The physical relationship of the external potential with the density profile  $\rho(\mathbf{r}')$  is often viewed in a causal way, such that a change in the external potential at some position  $\mathbf{r}$  will create a density response in the system [1]. In general this response will not only occur at the same position, but also, mediated by the interparticle interactions, at positions  $\mathbf{r}'$  further away. Near a surface phase transition [6,7] the associated length-scale can become very large.

On a formal level,  $V_{\text{ext}}(\mathbf{r})$  and  $\rho(\mathbf{r}')$  form a pair of conjugate variables within the variational framework of classical density functional theory (DFT) [1,15–17]. DFT is based on the existence of a generating (free energy) functional. Its nontrivial contribution, the intrinsic excess free energy functional,  $F_{\text{exc}}[\rho]$ , originates from the interparticle interactions. Due to inherent coupling of the degrees of freedom of the many-body system, exact expressions for this quantity do not exist except for rare special cases. Approximations are thus required for most applications. Minimizing the grand potential functional,  $\Omega[\rho]$ , typically by numerically solving the associated Euler-Lagrange (EL) equation, then gives results for the spatial structure and the thermodynamics of the inhomogeneous system under consideration. The EL equation can be viewed as a condition of local chemical equilibrium throughout the system [15]. Here the local

\*salomee.tschopp@unifr.ch

†Matthias.Schmidt@uni-bayreuth.de

‡joseph.brader@unifr.ch

chemical potential consists of three physically distinct contributions: a trivial ideal gas term, an excess (over ideal) term which arises from the interparticle interactions and an external contribution.

An analogous point of view, which at first sight seems to be based on quite different physical intuition, is that of a force balance relationship. As an equilibrium system is on average at rest, the total local force must vanish at each point in space. This is a classical result obtained by Yvon [18], Born and Green [19] (YBG) and it forms an exact property (sum-rule). Within computer simulation methodology, working on the level of force distributions has recently received a boost through the introduction of smart sampling strategies. “Use the force” [20] constitutes a new paradigm for obtaining data with significantly reduced statistical noise [21–23], as compared to direct sampling via simple counting of events. Force distributions naturally generalize to nonequilibrium, where the equilibrium ensemble average is replaced by a dynamical average over the corresponding set of states, e.g., for overdamped Brownian dynamics [24,25]. For quantum systems, the locally resolved force-balance relationship was recently addressed for dynamical situations [26,27]. Furthermore, two-body correlation functions are central to the recently developed conditional probability DFT [28,29].

On a fundamental level it is apparent that out of equilibrium, it is forces, rather than potentials, that play the central role in determining the particle motion. A dynamical theory based on potentials will clearly be incapable of treating non-conservative forces and can also be expected to break down whenever the microstructure of the system deviates strongly from that of equilibrium. These difficulties present a fundamental limitation to the usefulness of existing dynamical density functional (DDFT) approaches [15,30,31] and have served to motivate development of the force-based power functional theory (PFT) [32,33].

Focusing on equilibrium, the YBG derivation conventionally rests on formally integrating the full  $N$ -body equilibrium distribution over  $N - 1$  spatial degrees of freedom [1,18,19]. The one-body density is thus expressed in terms of an integral of the two-body density. In contrast, DFT is closed on the one-body level (using the EL equation) and hence neither requires consideration of the two-body level, nor does it permit to systematically incorporate such information. In this paper we present a theoretical density functional approach, which accounts explicitly for the interparticle forces and enables calculation of the one-body density for inhomogeneous fluids both in- and out-of-equilibrium.

## II. ROADMAP

In the following we give an overview to guide the reader through the main results and concepts presented in this work. We begin, in Sec. III A, by developing the Noether theorem for the invariance of the grand potential under spatial distortions, as characterized by a vector displacement field  $\epsilon(\mathbf{r})$ . Expressing the grand potential as a functional of  $\epsilon(\mathbf{r})$  leads to the variational condition,

$$\left. \frac{\delta \Omega[\epsilon]}{\delta \epsilon(\mathbf{r})} \right|_{\epsilon(\mathbf{r})=0} = 0,$$

which generates the following force-balance (YBG) relation

$$-k_B T \nabla_{\mathbf{r}_1} \ln(\rho(\mathbf{r}_1)) - \nabla_{\mathbf{r}_1} V_{\text{ext}}(\mathbf{r}_1) - \int d\mathbf{r}_2 \frac{\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2)}{\rho(\mathbf{r}_1)} \nabla_{\mathbf{r}_1} \phi(|\mathbf{r}_1 - \mathbf{r}_2|) = 0,$$

where the subscripted position variables  $\mathbf{r}_1$  and  $\mathbf{r}_2$  play the role of a fixed point in space,  $\mathbf{r}_1$ , and a “field point” which is integrated over,  $\mathbf{r}_2$ . The two-body density and pair interaction potential are indicated by  $\rho^{(2)}$  and  $\phi$ , respectively ( $k_B$  denotes the Boltzmann constant and  $T$  is the absolute temperature). Our variational derivation highlights the fundamental status of the YBG equation, which we then take as a starting point for a self-consistent approach to determining the one-body density. It can be written without approximation in the following form:

$$\rho(\mathbf{r}_1) = e^{\beta[\mu - V_{\text{ext}}(\mathbf{r}_1)] + c_{\mathbf{r}_1}^{(1)}(\mathbf{r}_1)},$$

where  $\mu$  is the chemical potential,  $\beta = (k_B T)^{-1}$  and the contribution  $-k_B T c_{\mathbf{r}_1}^{(1)}(\mathbf{r}_1)$  acts as an effective external field arising from interparticle interactions. In Sec. III B, we introduce the concept that the two-body correlation functions are functionals of the one-body density and we then use this to reinterpret the YBG equation as a *closed* integral equation for the one-body density. This leads us to the definition

$$c_{\mathbf{r}_1}^{(1)}(\mathbf{r}_1) \equiv -\nabla_{\mathbf{r}_1}^{-1} \cdot \int d\mathbf{r}_2 \frac{\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2; [\rho])}{\rho(\mathbf{r}_1)} \nabla_{\mathbf{r}_1} \beta \phi(|\mathbf{r}_1 - \mathbf{r}_2|),$$

in which the interparticle forces appear explicitly via  $\nabla_{\mathbf{r}_1} \phi$ . The integral operator  $\nabla_{\mathbf{r}_1}^{-1}$  is defined in the main text and the square brackets indicate a functional dependence. An essential feature of our approach is that we have a computationally feasible scheme to evaluate the inhomogeneous density functional  $\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2; [\rho])$ , which then leads to a closed, self-consistent “force-DFT.”

In Sec. III C, our force-based approach is contrasted with the standard DFT methodology (referred to in this work as potential-DFT) in which the grand potential is expressed as a functional of the one-body density and satisfies the variational condition,

$$\left. \frac{\delta \Omega[\rho]}{\delta \rho(\mathbf{r})} \right|_{\rho(\mathbf{r})=\rho_0(\mathbf{r})} = 0,$$

where  $\rho_0(\mathbf{r})$  is the equilibrium density profile (the subscript will be omitted in the following). This leads to the well-known EL equation

$$\ln \rho(\mathbf{r}) - \beta(\mu - V_{\text{ext}}(\mathbf{r})) - c_{\mathbf{r}}^{(1)}(\mathbf{r}) = 0,$$

which can be expressed in the following alternative form

$$\rho(\mathbf{r}) = e^{\beta(\mu - V_{\text{ext}}(\mathbf{r})) + c_{\mathbf{r}}^{(1)}(\mathbf{r})},$$

where the function  $c_{\mathbf{r}}^{(1)}$  is defined as a functional derivative of the excess Helmholtz free energy,

$$c_{\mathbf{r}}^{(1)}(\mathbf{r}) = -\frac{\delta \beta F_{\text{exc}}[\rho]}{\delta \rho(\mathbf{r})}.$$

In contrast to the force-DFT, the potential-DFT involves only one-body functions. We thus require only a single vector position,  $\mathbf{r}$ , as an independent variable and there is no need to employ additional subscripts. The EL equation is

the potential-DFT analog of the YBG equation arising from invariance with respect to spatial distortions. If the free energy and the two-body density functionals are known only approximately, then the two approaches to DFT will lead in general to different density profiles for a given external field. This allows for deep insight into the inner workings of DFT. Section III C is intended primarily for readers who are less familiar with the details of potential-DFT.

In potential-DFT the well-used contact theorem predicts that the contact density at a hard planar wall is equal to  $\beta P^c$ , where  $P^c$  is the compressibility pressure, see Ref. [1] for its definition. For the force-DFT we find that the equivalent result links the contact density to  $\beta P^v$ , where  $P^v$  is the virial pressure. In Secs. III D and III E, we prove the corresponding contact theorem for both approaches. These sum-rules are exact and hold within any reasonable approximation scheme. If the reader is prepared to accept these assertions without proof, then both of these subsections can be passed-over on a first reading of the manuscript.

As mentioned previously, to implement the force-DFT, we require a feasible method to obtain the density functional  $\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2; [\rho])$ . Therefore, from that point onwards we will rely on approximation schemes. In Secs. III F and III G, we recall the fundamental measure theory (FMT) for hard-spheres and provide information about the numerical implementation. The FMT generates an explicit expression for the two-body direct correlation function,  $c^{(2)}(\mathbf{r}_1, \mathbf{r}_2; [\rho])$ , as a functional of the one-body density. The two-body direct correlation function, now uniquely determined by the one-body density, can be used as input to the inhomogeneous Ornstein-Zernike (OZ) equation,

$$h(\mathbf{r}_1, \mathbf{r}_2) = c^{(2)}(\mathbf{r}_1, \mathbf{r}_2) + \int d\mathbf{r}_3 h(\mathbf{r}_1, \mathbf{r}_3)\rho(\mathbf{r}_3)c^{(2)}(\mathbf{r}_3, \mathbf{r}_2),$$

which is then a linear integral equation for determination of the total correlation function,  $h$ . By self-consistent solution of the OZ equation we obtain  $h$  for any given one-body density;  $h$  is thus a density functional. It has been shown that the inhomogeneous OZ equation can be numerically solved to high accuracy both in planar and spherical geometry [14,34,35]. Using the relation

$$\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = \rho(\mathbf{r}_1)\rho(\mathbf{r}_2)(h(\mathbf{r}_1, \mathbf{r}_2) + 1)$$

then gives a clear self-consistent scheme to determine  $\rho^{(2)}$  as a functional of the one-body density. In Sec. III H, we show numerical results for the equilibrium density of hard-spheres at a hard-wall using both force- and potential-DFT, and validate the analytical predictions for the contact density. This demonstrates explicitly that the presented framework is not merely formal, but that it forms a concrete numerical scheme for the systematic study of inhomogeneous fluids.

In Sec. IV, we consider nonequilibrium systems subject to overdamped Brownian dynamics and show how the force-DFT allows calculation of the time-dependent density,  $\rho(\mathbf{r}, t)$ . We argue that the force-DFT provides the most natural starting point for the development of a dynamical theory for the density, as it is the forces which are responsible for moving the particles. Conservation of particle number dictates that the

density obeys the continuity equation

$$\frac{\partial \rho(\mathbf{r}_1, t)}{\partial t} = -\nabla_{\mathbf{r}_1} \cdot \mathbf{j}(\mathbf{r}_1, t),$$

where  $\mathbf{j}(\mathbf{r}_1, t)$  is the current, which needs to be specified to have a closed theory.

In Sec. IV A, we describe the force-DDFT, which is based on the following exact expression for the current

$$\mathbf{j}(\mathbf{r}_1, t) = -D_0 \rho(\mathbf{r}_1, t) \left( \nabla_{\mathbf{r}_1} \ln[\rho(\mathbf{r}_1, t)] + \nabla_{\mathbf{r}_1} \beta V_{\text{ext}}(\mathbf{r}_1) + \int d\mathbf{r}_2 \frac{\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2, t)}{\rho(\mathbf{r}_1, t)} \nabla_{\mathbf{r}_1} \beta \phi(|\mathbf{r}_1 - \mathbf{r}_2|) \right),$$

where  $D_0$  is the diffusion coefficient. Using the previously described equilibrium functional for  $\rho^{(2)}$  yields a closed adiabatic theory for the one-body density. At each time-step the integral term is explicitly evaluated to obtain the average force due to interparticle interactions. In contrast, the familiar potential-DDFT, recalled in Sec. IV B, employs only one-body functions. The current in this case is given by

$$\mathbf{j}(\mathbf{r}, t) = -D_0 \rho(\mathbf{r}, t) \nabla_{\mathbf{r}} (\ln(\rho(\mathbf{r}, t)) + \beta V_{\text{ext}}(\mathbf{r}) - c_p^{(1)}(\mathbf{r}, t)).$$

In Sec. IV C, we employ the FMT to generate numerical results for the density relaxation in a harmonic-trap and we compare the predictions of the force-DDFT with those of the potential-DDFT. This demonstrates that our force-based theory provides a firm basis for developing a systematic understanding of nonequilibrium phenomena. Finally, in Sec. V, we draw our conclusions and give an outlook for future work.

### III. EQUILIBRIUM THEORY

#### A. Force-balance generated by Noether's theorem

We begin our development of force-DFT by starting with the microscopic Hamiltonian and using invariance arguments. Let us consider a classical system of  $N$  particles described by position coordinates  $\mathbf{r}_1, \dots, \mathbf{r}_N \equiv \mathbf{r}^N$  and momenta  $\mathbf{p}_1, \dots, \mathbf{p}_N \equiv \mathbf{p}^N$ . The Hamiltonian  $H$  has the standard form consisting of kinetic, internal, and external potential energy contributions according to

$$H = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} + U_N(\mathbf{r}^N) + \sum_{i=1}^N V_{\text{ext}}(\mathbf{r}_i). \quad (1)$$

Here  $m$  indicates the particle mass,  $U_N$  denotes the total interparticle interaction potential and  $V_{\text{ext}}$  is an external one-body field.

We consider a canonical transformation on phase-space, parameterized by a vector field  $\boldsymbol{\epsilon}(\mathbf{r})$  that describes a spatial displacement ("distortion") at position  $\mathbf{r}$ . The transformation affects both coordinates and momenta and is given by

$$\mathbf{r}_i \rightarrow \mathbf{r}_i + \boldsymbol{\epsilon}(\mathbf{r}_i) \equiv \mathbf{r}'_i, \quad (2)$$

$$\mathbf{p}_i \rightarrow \mathbf{p}_i - \nabla_{\mathbf{r}_i} \boldsymbol{\epsilon}(\mathbf{r}_i) \cdot \mathbf{p}_i \equiv \mathbf{p}'_i, \quad (3)$$

where the primes indicate the new phase-space variables and  $\nabla_{\mathbf{r}_i}$  denotes differentiation with respect to  $\mathbf{r}_i$ . We consider the displacement field  $\boldsymbol{\epsilon}(\mathbf{r})$  and its gradient to be small.

The change in phase space variables affects the Hamiltonian and renders it functionally dependent on the displacement field,  $H \rightarrow H[\epsilon]$ . Inserting transformations (2) and (3) into Eq. (1) and expanding in the displacement field to linear order yields

$$H[\epsilon] = H_0 - \sum_{i=1}^N \frac{\mathbf{p}_i \mathbf{p}_i}{m} : \nabla_{\mathbf{r}_i} \epsilon(\mathbf{r}_i) + \sum_{i=1}^N \epsilon(\mathbf{r}_i) \cdot \nabla_{\mathbf{r}_i} (U_N(\mathbf{r}^N) + V_{\text{ext}}(\mathbf{r}_i)), \quad (4)$$

where  $H_0 = H[\epsilon = 0]$  is the original Hamiltonian as given in Eq. (1). The colon indicates the contraction  $\mathbf{p}_i \mathbf{p}_i : \nabla_{\mathbf{r}_i} \epsilon(\mathbf{r}_i) = \sum_{\alpha, \gamma} p_{i\alpha} p_{i\gamma} \nabla_{r_{i\gamma}} \epsilon_{\alpha}$ , where Greek indices indicate Cartesian components.

Turning to a statistical description, the grand potential,  $\Omega_0 = \Omega[\epsilon = 0]$ , and the grand partition sum,  $\Xi_0 = \Xi[\epsilon = 0]$ , of the original system are given, respectively, by

$$\Omega_0 = -k_B T \ln \Xi_0, \quad (5)$$

$$\Xi_0 = \text{Tr} e^{-\beta(H_0 - \mu N)}. \quad (6)$$

In the grand canonical ensemble the trace is defined as  $\text{Tr} = \sum_{N=0}^{\infty} (h^{3N} N!)^{-1} \int d\mathbf{r}_1 \dots d\mathbf{r}_N d\mathbf{p}_1 \dots d\mathbf{p}_N$ , with  $h$  indicating the Planck constant [1].

The transformed Hamiltonian,  $H[\epsilon]$ , can be used to define a correspondingly transformed grand potential functional,  $\Omega[\epsilon] = -k_B T \ln(\text{Tr} e^{-\beta(H[\epsilon] - \mu N)})$ . To linear order in  $\epsilon(\mathbf{r})$  the functional Taylor expansion of  $\Omega[\epsilon]$  is given by

$$\Omega[\epsilon] = \Omega_0 + \int d\mathbf{r} \left. \frac{\delta \Omega[\epsilon]}{\delta \epsilon(\mathbf{r})} \right|_{\epsilon(\mathbf{r})=0} \cdot \epsilon(\mathbf{r}). \quad (7)$$

The functional derivative in Eq. (7) can be calculated as follows:

$$\begin{aligned} \frac{\delta \Omega[\epsilon]}{\delta \epsilon(\mathbf{r})} &= -\frac{k_B T}{\Xi[\epsilon]} \text{Tr} \frac{\delta}{\delta \epsilon(\mathbf{r})} e^{-\beta(H[\epsilon] - \mu N)} \\ &= -\frac{k_B T}{\Xi[\epsilon]} \text{Tr} e^{-\beta(H[\epsilon] - \mu N)} \left( -\beta \frac{\delta H[\epsilon]}{\delta \epsilon(\mathbf{r})} \right) \\ &= \text{Tr} \Psi \frac{\delta H[\epsilon]}{\delta \epsilon(\mathbf{r})}, \end{aligned} \quad (8)$$

where we have identified  $\Psi = e^{-\beta(H[\epsilon] - \mu N)} / \Xi[\epsilon]$  as the grand ensemble probability distribution. Notably the form (8) constitutes a grand ensemble average of  $\delta H[\epsilon] / \delta \epsilon(\mathbf{r})$ . Formally, the average is taken in the displaced system, but we will find the form (8) to be sufficient to calculate averages with respect to the original, undisplaced distribution. Using Eq. (4) and thus retaining only the lowest relevant order in  $\epsilon(\mathbf{r})$  we obtain

$$\begin{aligned} \frac{\delta H[\epsilon]}{\delta \epsilon(\mathbf{r})} &= \sum_{i=1}^N \left( -\frac{\mathbf{p}_i \mathbf{p}_i}{m} \cdot \nabla_{\mathbf{r}_i} \delta(\mathbf{r} - \mathbf{r}_i) \right. \\ &\quad \left. + \delta(\mathbf{r} - \mathbf{r}_i) \nabla_{\mathbf{r}_i} [U_N(\mathbf{r}^N) + V_{\text{ext}}(\mathbf{r}_i)] \right). \end{aligned} \quad (9)$$

Here we have used the fundamental rule of functional differentiation  $\delta \epsilon(\mathbf{r}) / \delta \epsilon(\mathbf{r}') = \delta(\mathbf{r} - \mathbf{r}') \mathbb{1}$ , where  $\delta(\cdot)$  indicates the (three-dimensional) Dirac distribution, and  $\mathbb{1}$  denotes the

$3 \times 3$  unit matrix. Using Eq. (9) inside of the average (8), carrying out the phase-space integrals, evaluating at  $\epsilon(\mathbf{r}) = 0$  and multiplying by  $-1$  yields

$$-\left. \frac{\delta \Omega[\epsilon]}{\delta \epsilon(\mathbf{r})} \right|_{\epsilon(\mathbf{r})=0} = -k_B T \nabla_{\mathbf{r}} \rho(\mathbf{r}) + \mathbf{F}_{\text{int}}(\mathbf{r}) - \rho(\mathbf{r}) \nabla_{\mathbf{r}} V_{\text{ext}}(\mathbf{r}), \quad (10)$$

where the one-body density profile is defined as the average  $\rho(\mathbf{r}) = \text{Tr} \Psi \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$  and the internal force density is given by  $\mathbf{F}_{\text{int}}(\mathbf{r}) = -\text{Tr} \Psi \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \nabla_{\mathbf{r}_i} U_N(\mathbf{r}^N)$ . Furthermore, the ideal diffusion force density  $-k_B T \nabla_{\mathbf{r}} \rho(\mathbf{r})$  follows from carrying out the phase-space momentum integrals explicitly or, alternatively, using the equipartition theorem  $\text{Tr} \Psi \mathbf{p}_i \partial H / \partial \mathbf{p}_i = k_B T \mathbb{1}$ . The spatial gradients in Eq. (10) emerge by exploiting  $\nabla_{\mathbf{r}} \delta(\mathbf{r} - \mathbf{r}_i) = -\nabla_{\mathbf{r}_i} \delta(\mathbf{r} - \mathbf{r}_i)$ . The right-hand side of Eq. (10) represents the sum of the position-resolved average one-body force densities of ideal, interparticle and external origin.

Using the transformations (2) and (3) one can easily verify that the differential volume elements for coordinates and momenta follow to linear order in  $\epsilon$  as

$$d\mathbf{r}_i \rightarrow (1 + \nabla_{\mathbf{r}_i} \cdot \epsilon(\mathbf{r}_i)) d\mathbf{r}'_i \equiv d\mathbf{r}'_i, \quad (11)$$

$$d\mathbf{p}_i \rightarrow (1 - \nabla_{\mathbf{r}_i} \cdot \epsilon(\mathbf{r}_i)) d\mathbf{p}'_i \equiv d\mathbf{p}'_i. \quad (12)$$

We thus see that for each particle  $d\mathbf{r}_i d\mathbf{p}_i = d\mathbf{r}'_i d\mathbf{p}'_i$  holds to linear order in  $\epsilon$  and, therefore, for the entire phase-space  $\prod_{i=1}^N d\mathbf{r}_i d\mathbf{p}_i = \prod_{i=1}^N d\mathbf{r}'_i d\mathbf{p}'_i$ , as befits a canonical transformation (see Appendix A).

As the transformation is also time-independent, the Hamiltonian is an invariant (see again Appendix A). Then trivially the partition sum (6) and the grand potential (5) are also invariants. It follows that

$$\Omega[\epsilon] = \Omega_0. \quad (13)$$

The linear term in the functional Taylor expansion (7) thus vanishes and it does so irrespective of the form of  $\epsilon(\mathbf{r})$ . The functional derivative (10) itself must therefore vanish,

$$\left. \frac{\delta \Omega[\epsilon]}{\delta \epsilon(\mathbf{r})} \right|_{\epsilon(\mathbf{r})=0} = 0,$$

from which we can conclude that

$$-k_B T \nabla_{\mathbf{r}} \rho(\mathbf{r}) + \mathbf{F}_{\text{int}}(\mathbf{r}) - \rho(\mathbf{r}) \nabla_{\mathbf{r}} V_{\text{ext}}(\mathbf{r}) = 0, \quad (14)$$

which is the known equilibrium force density relationship [1,33].

When considering systems interacting via a pair potential  $\phi$ , the internal potential energy has the form  $U_N(\mathbf{r}^N) = \sum_{i < j} \phi(|\mathbf{r}_i - \mathbf{r}_j|)$ . The internal force density can then be written as

$$\mathbf{F}_{\text{int}}(\mathbf{r}_1) = - \int d\mathbf{r}_2 \rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2) \nabla_{\mathbf{r}_1} \phi_{12}, \quad (15)$$

where the two-body density is defined microscopically as  $\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = \text{Tr} \Psi \sum_{i \neq j} \delta(\mathbf{r}_1 - \mathbf{r}_i) \delta(\mathbf{r}_2 - \mathbf{r}_j)$ , with the prime on the summation indicating the omission of the terms with  $i = j$ , and we indicate the pair interaction potential by the shorthand  $\phi_{12} = \phi(|\mathbf{r}_1 - \mathbf{r}_2|)$ . We have relabeled  $\mathbf{r} \rightarrow \mathbf{r}_1$  to give clarity to equations involving two-body functions. Using

the explicit form (15) in the force density relationship (14) and rearranging yields

$$\begin{aligned} & -k_B T \nabla_{\mathbf{r}_1} (\ln \rho(\mathbf{r}_1)) - \nabla_{\mathbf{r}_1} V_{\text{ext}}(\mathbf{r}_1) \\ & - \int d\mathbf{r}_2 \frac{\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2)}{\rho(\mathbf{r}_1)} \nabla_{\mathbf{r}_1} \phi_{12} = 0, \end{aligned} \quad (16)$$

which is the explicit form of the first member of the YBG hierarchy [1]. We have thus shown that Eq. (16) arises from a variational principle on the grand potential. It is hence of no lesser status than the EL equation of potential-DFT, to be discussed in Sec. III C.

### B. Force-DFT

The YBG Eq. (16), which has been derived using Noether invariance in the previous subsection, has the appealing feature that it explicitly contains the interparticle pair interaction,  $\phi_{12}$ . We thus take the YBG Eq. (16) as a fundamental starting point for describing the equilibrium state. The third term in Eq. (16), which gives the mean interparticle interaction force at the point  $\mathbf{r}_1$ , is not written as the gradient of a potential. However, as an equilibrium system is conservative by construction we can formally rewrite it as a potential force using the inverse of the gradient

$$\begin{aligned} & \int d\mathbf{r}_2 \frac{\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2)}{\rho(\mathbf{r}_1)} \nabla_{\mathbf{r}_1} \phi_{12} \\ & = \nabla_{\mathbf{r}_1} \left( \nabla_{\mathbf{r}_1}^{-1} \cdot \int d\mathbf{r}_2 \frac{\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2)}{\rho(\mathbf{r}_1)} \nabla_{\mathbf{r}_1} \phi_{12} \right), \end{aligned} \quad (17)$$

where  $\nabla_{\mathbf{r}_1}^{-1} = \frac{1}{4\pi} \int d\mathbf{r}_2 \frac{(\mathbf{r}_1 - \mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|^3}$  is an integral operator (see, e.g., Refs. [20,21]). We thus define a scalar one-body function  $c_f^{(1)}$  according to

$$c_f^{(1)}(\mathbf{r}_1) \equiv -\nabla_{\mathbf{r}_1}^{-1} \cdot \int d\mathbf{r}_2 \frac{\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2)}{\rho(\mathbf{r}_1)} \nabla_{\mathbf{r}_1} \beta \phi_{12}. \quad (18)$$

Although we employ the notation usually reserved for the one-body direct correlation function, Eq. (18) originates here from a quite different, but arguably more intuitive and fundamental way of thinking. We can thus re-express the YBG Eq. (16) in the following form:

$$\nabla_{\mathbf{r}_1} (-k_B T \ln \rho(\mathbf{r}_1) - V_{\text{ext}}(\mathbf{r}_1) + k_B T c_f^{(1)}(\mathbf{r}_1)) = 0.$$

Equilibrium implies that the term in parentheses is equal to a constant  $\mu$ , which leads to

$$\rho(\mathbf{r}_1) = e^{\beta[\mu - V_{\text{ext}}(\mathbf{r}_1) + c_f^{(1)}(\mathbf{r}_1)].} \quad (19)$$

In contrast to standard potential-DFT here the function  $c_f^{(1)}$  is simply defined by Eq. (18) and is generated directly from an explicit integral over the pair interaction force. Combining Eqs. (19) and (18) yields

$$\begin{aligned} \rho(\mathbf{r}_1) = \exp \left( \beta(\mu - V_{\text{ext}}(\mathbf{r}_1)) \right. \\ \left. - \nabla_{\mathbf{r}_1}^{-1} \cdot \int d\mathbf{r}_2 \frac{\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2; [\rho])}{\rho(\mathbf{r}_1)} \nabla_{\mathbf{r}_1} \beta \phi_{12} \right), \end{aligned} \quad (20)$$

which is the central equation of force-DFT. Given an explicit expression for the two-body density as a functional of the

one-body density,  $\rho^{(2)} \equiv \rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2; [\rho])$ , Eq. (20) enables calculation of  $\rho(\mathbf{r}_1)$  for any given external potential. While one can argue on formal grounds that the force integral and its nontrivial essence, the two-body density distribution, are one-body density functionals, our current treatment makes this formal dependence both analytically explicit and computationally tractable.

### C. Potential-DFT

The standard implementation of DFT (referred to in this work as potential-DFT) is based on the grand potential density functional,  $\Omega[\rho]$ , given by

$$\Omega[\rho] = F_{\text{id}}[\rho] + F_{\text{exc}}[\rho] - \int d\mathbf{r} (\mu - V_{\text{ext}}(\mathbf{r})) \rho(\mathbf{r}), \quad (21)$$

where  $\beta F_{\text{id}}[\rho] = \int d\mathbf{r} \rho(\mathbf{r}) \{\ln[\rho(\mathbf{r})] - 1\}$  is the ideal gas contribution with the thermal wavelength set equal to unity. Variational minimization of the grand potential

$$\frac{\delta \Omega[\rho]}{\delta \rho(\mathbf{r})} = 0, \quad (22)$$

generates the EL equation,

$$\rho(\mathbf{r}) = e^{\beta[\mu - V_{\text{ext}}(\mathbf{r}) + c_p^{(1)}(\mathbf{r})]}. \quad (23)$$

The function  $c_p^{(1)}$  is defined to be the first functional derivative of the excess (over ideal) Helmholtz free energy with respect to the density [1,15],

$$c_p^{(1)}(\mathbf{r}) = -\frac{\delta \beta F_{\text{exc}}[\rho]}{\delta \rho(\mathbf{r})}. \quad (24)$$

This function  $c_p^{(1)}$ , which is now the familiar one-body direct correlation function, is the first member of a hierarchy of correlation functions generated by successive functional differentiation of  $F_{\text{exc}}$  with respect to the density. For situations in which all quantities are known exactly the definition given in Eq. (24) is equivalent to that of Eq. (18). Even though the EL Eq. (23) has the same structure as Eq. (19), these are conceptually different and have distinct origins. The potential-DFT is constructed using only one-body functions and the average interaction force is generated by taking the gradient of  $c_p^{(1)}$ . This should be contrasted with the force-DFT, which works on the two-body level, in which the interaction force is calculated by explicit spatial integration of the pair-interaction, see Eq. (17).

If both the one-body direct correlation function and the two-body density are generated from the same, exact free energy functional, then both the potential- and force-DFT implementations will yield the same average interaction force and thus the same density profiles. This will not be the case when using an approximate free energy functional and differences can be expected. The special case of a hard-wall substrate enables the degree of consistency between these two routes to be examined analytically and this will be the focus of the following two Secs. III D and III E (which can be skipped if the reader is more interested in the numerical predictions of potential- and force-DFT, which are presented in Sec. III H). Route-dependency will also turn out to be highly relevant for the dynamical versions of potential- and force-DFT, as shown later in Sec. IV.

#### D. Virial contact theorem

Before we proceed to investigate the virial route version of the contact theorem, we recall the very general and well-known version of it [1,36,37], namely, that  $\rho_w = \beta P$ , which relates the density of a fluid at a planar hard-wall,  $\rho_w$ , to the corresponding bulk pressure,  $P$ . This can be proven without the need to specify by which method the one-body density is obtained. For completeness we provide a general proof of this in Appendix B. Other general proofs of the contact theorem are based on the balance of forces [1,36,37], a linear displacement of the free energy [37] and the connection between the pressure and the mean kinetic energy density [36]. The contact theorem is satisfied within DFT for excess free energy functionals within the weighted density approximation [38,39], with FMT [40] being an important example.

There are several generalizations of the wall theorem, which include a version for higher-body densities [41] and extensions to hard-walls with additional soft particle-wall interactions [42,43] as well as to nonplanar locally curved hard-walls [44,45], for which one can also get a local version of the contact theorem [46]. Another important generalization is the extension to ionic liquids [42,47–49], where an additional term proportional to the squared surface charge arises in the contact theorem.

In the aforementioned derivations, the contact value of the density is only related to a general bulk pressure. Exceptions are the work of Lovett and Baus [37], where the authors identify the virial pressure and the study of Tarazona and Evans [39], where the contact theorem for the Percus-Yevick and the hypernetted chain approximation were determined. Due to approximations within theoretical descriptions the pressures from different routes do not necessarily agree with each other. When used in potential-DFT studies it is always implicitly assumed that the relevant pressure is that of the compressibility route [38], which we prove is indeed the case in the next subsection. For the force-DFT we prove here first that the relevant bulk pressure is that of the virial route. The ability to access these two routes for inhomogeneous systems offers both the possibility of new insight into the formal structure of DFT and a useful tool for constructing approximate functionals.

Let us focus now on the virial contact theorem. The force-DFT is generated by the YBG Eq. (16). We begin by spatially integrating it over the system volume  $V$  to obtain

$$\begin{aligned} & - \int d\mathbf{r}_1 \rho(\mathbf{r}_1) \nabla_{\mathbf{r}_1} \beta V_{\text{ext}}(\mathbf{r}_1) \\ & = \int d\mathbf{r}_1 \nabla_{\mathbf{r}_1} \rho(\mathbf{r}_1) + \int d\mathbf{r}_1 \int d\mathbf{r}_2 \rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2) \nabla_{\mathbf{r}_1} \beta \phi_{12}. \end{aligned} \quad (25)$$

Exploiting the planar symmetry imposed by the hard-wall allows to simplify the density,  $\rho(\mathbf{r}_1) = \rho(z_1)$ , the external potential  $V_{\text{ext}}(\mathbf{r}_1) = V_{\text{ext}}(z_1)$  and the spatial derivative  $\nabla_{\mathbf{r}_1} = \mathbf{e}_z d/dz_1$ , where  $\mathbf{e}_z$  is the unit vector normal to the wall and pointing away from it (see Fig. 1 for illustration). The ideal contribution, i.e., the first term on the right-hand side of Eq. (25), can then be rewritten as

$$\int d\mathbf{r}_1 \nabla_{\mathbf{r}_1} \rho(\mathbf{r}_1) = A \int_{-\infty}^{\infty} dz_1 \frac{d\rho(z_1)}{dz_1} \mathbf{e}_z = A \rho_b \mathbf{e}_z, \quad (26)$$

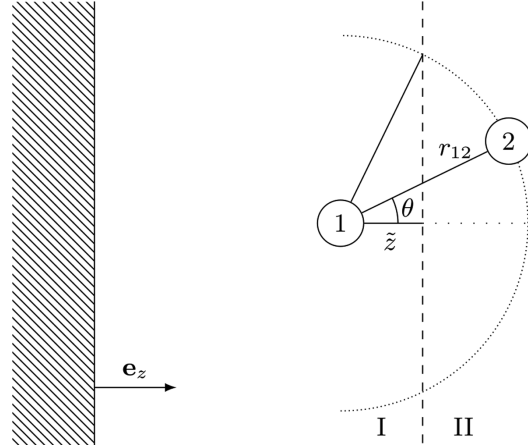


FIG. 1. Geometrical sketch of the planar geometry at a hard-wall. For the evaluation of the virial integral [Eqs. (28) and (29)] the space is divided into two subregions I and II. For a given value of coordinate 1 we integrate coordinate 2 over the angle  $\theta = 0 \rightarrow \arccos(z/r_{12})$ . The shaded region on the left indicates the wall,  $\mathbf{e}_z$  denotes a unit vector in the  $z$  direction and  $\tilde{z}$  is the  $z$  coordinate measured relative to coordinate 1.

where  $A = \int dx \int dy$  indicates the area of the hard-wall. In the second equality of (26) we used that the density reaches a bulk value for large values of  $z_1$ ,  $\rho(z_1 \rightarrow \infty) = \rho_b$ , and vanishes inside the hard-wall,  $\rho(z_1 \rightarrow -\infty) = 0$ . The external contribution, i.e., the left-hand side of Eq. (25), becomes

$$\begin{aligned} - \int d\mathbf{r}_1 \rho(\mathbf{r}_1) \nabla_{\mathbf{r}_1} \beta V_{\text{ext}}(\mathbf{r}_1) & = -A \int_{-\infty}^{\infty} dz_1 \rho(z_1) \frac{d\beta V_{\text{ext}}(z_1)}{dz_1} \mathbf{e}_z \\ & = A \rho_w \mathbf{e}_z. \end{aligned} \quad (27)$$

In deriving Eq. (27) we used that the derivative of the hard-wall external potential yields a (negative)  $\delta$  distribution at the wall. To obtain this result it is useful to rewrite the density as  $\rho(z_1) = n(z_1) \exp[-\beta V_{\text{ext}}(z_1)]$ , where  $n(z_1)$  is a continuous function of  $z_1$ .

The second term on the right-hand side of Eq. (25) arises from the internal interparticle interactions and it is related to the global internal force,  $\mathbf{F}_{\text{int}}^o$ . Noether's theorem [50] states that the global internal force vanishes in a closed system. As the semi-infinite system with a planar hard-wall is open, one has to take boundary contributions into account [50]. The boundary terms corresponding to the  $x$  and  $y$  axis cancel due to the planar symmetry. The wall contribution,  $z_1 \rightarrow -\infty$ , vanishes as there are no particles inside the hard-wall. (We refer the reader to Fig. 1 to help visualize the situation for the following analysis.)

The remaining bulk boundary term,  $z_1 \rightarrow \infty$ , can be treated by considering the force contributions between a particle inside a chosen integration volume and a particle outside of it. The force contributions where both particles are within the integration volume vanish for pair potential-type interparticle interactions, due to Newton's third law (*action equals reaction*). We choose the volume to be bounded from the right



(positive values of  $z$ ) by a virtual plane parallel to the hard-wall and deep inside the bulk phase. Therefore and because of the assumed finite interparticle interaction range the integrand and thus the two-body density,  $\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$ , reduces to its bulk expression,  $\rho_b^{(2)}(|\mathbf{r}_1 - \mathbf{r}_2|)$ , at locations beyond this virtual separation plane. These considerations yield the following simplifications of the global internal force

$$\begin{aligned} -\beta \mathbf{F}_{\text{int}}^0 &= \int d\mathbf{r}_1 \int d\mathbf{r}_2 \rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2) \nabla_{\mathbf{r}_1} \beta \phi_{12} \\ &= \int d\mathbf{r}_1 \int_{\text{II}} d\mathbf{r}_2 \rho_b^{(2)}(r_{12}) \nabla_{\mathbf{r}_1} \beta \phi_{12} \\ &= \rho_b^2 \int d\mathbf{r}_1 \int_{\text{II}} d\mathbf{r}_2 g(r_{12}) \frac{d\beta \phi_{12}}{dr_{12}} \cos \theta \mathbf{e}_z, \end{aligned} \quad (28)$$

where  $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$ . The subscript I on the integral denotes the integration volume with  $z$  coordinate reaching from minus infinity to the volume boundary and II indicates the outside region at large  $z$  values. To obtain the last relation in Eq. (28) we use the identity  $\rho_b^{(2)}(r_{12}) = \rho_b^2 g(r_{12})$ , where  $g$  is the pair correlation function. The interparticle interaction force simplifies to  $\nabla_{\mathbf{r}_1} \phi(|\mathbf{r}_1 - \mathbf{r}_2|) = \mathbf{e}_z \cos(\theta) d\phi(r_{12})/dr_{12}$  due to the planar symmetry, where  $\theta$  indicates the angle between the  $z$ -axis and the difference vector  $\mathbf{r}_1 - \mathbf{r}_2$ .

We employ two different coordinate systems for each of the integration regions. For region I we use Cartesian coordinates, where  $\tilde{z}$  measures the distance to the volume boundary. The integral over region II involves only the relative coordinate between the inside and outside regions. We express this integral in spherical coordinates, where the polar angle  $\theta$  varies only between 0 and  $\hat{\theta} = \arccos(\tilde{z}/r_{12})$  to ensure that the second coordinate remains within region II. The  $z$  component of Eq. (28) is given by

$$\begin{aligned} 2\pi A \rho_b^2 \int_0^\infty dr_{12} r_{12}^2 g(r_{12}) \frac{d\phi(r_{12})}{dr_{12}} \int_0^{r_{12}} d\tilde{z} \int_0^{\hat{\theta}} d\theta \sin \theta \cos \theta \\ = 2\pi A \rho_b^2 \int_0^\infty dr_{12} r_{12}^2 g(r_{12}) \frac{d\phi(r_{12})}{dr_{12}} \int_0^{r_{12}} d\tilde{z} \frac{1}{2} \left(1 - \frac{\tilde{z}^2}{r_{12}^2}\right) \\ = \frac{2\pi}{3} A \rho_b^2 \int_0^\infty dr_{12} r_{12}^3 g(r_{12}) \frac{d\phi(r_{12})}{dr_{12}}. \end{aligned} \quad (29)$$

Inserting Eqs. (26), (27), and (29) into the  $z$  component of Eq. (25) gives

$$\begin{aligned} \rho_w &= \rho_b - \frac{2\pi}{3} \rho_b^2 \int_0^\infty dr_{12} r_{12}^3 g(r_{12}) \frac{d\beta \phi(r_{12})}{dr_{12}} \\ &= \beta P_{\text{id}} + \beta P_{\text{exc}}^v = \beta P^v, \end{aligned} \quad (30)$$

where we have identified the standard expression [1] for the virial pressure  $P^v$ . We have thus proven the contact theorem relevant to the force-DFT, namely, that if one uses force-DFT to calculate the density profile at a hard-wall, then the contact density will correspond to the reduced virial pressure,  $\beta P^v$ . Note that the derivation of the corresponding contact theorem in two dimensions can be done similarly.

### E. Compressibility contact theorem

The virial contact theorem derived above follows naturally from the forces acting within the system. In contrast the

compressibility contact theorem, based on the one-body direct correlation function, is more formal and requires therefore more involved manipulations of the fundamental equations to arrive at the desired result. Although the contact theorem is a result frequently cited in the literature, there is to our knowledge no calculation which shows explicitly that the wall contact density from potential-DFT is given by the reduced pressure from the *compressibility* route. Since potential-DFT is generated by the EL Eq. (23), our proof begins by taking its gradient, which yields

$$\begin{aligned} \nabla_{\mathbf{r}_1} \rho(\mathbf{r}_1) + \rho(\mathbf{r}_1) \nabla_{\mathbf{r}_1} \beta V_{\text{ext}}(\mathbf{r}_1) \\ - \rho(\mathbf{r}_1) \nabla_{\mathbf{r}_1} c_p^{(1)}(\mathbf{r}_1) = 0. \end{aligned} \quad (31)$$

The following simple identity from the product rule of differentiation

$$\begin{aligned} \nabla_{\mathbf{r}_1} [\rho(\mathbf{r}_1) c_p^{(1)}(\mathbf{r}_1)] \\ = c_p^{(1)}(\mathbf{r}_1) \nabla_{\mathbf{r}_1} \rho(\mathbf{r}_1) + \rho(\mathbf{r}_1) \nabla_{\mathbf{r}_1} c_p^{(1)}(\mathbf{r}_1), \end{aligned}$$

allows us then to rewrite Eq. (31) in the following alternative form

$$\begin{aligned} \nabla_{\mathbf{r}_1} \rho(\mathbf{r}_1) + \rho(\mathbf{r}_1) \nabla_{\mathbf{r}_1} \beta V_{\text{ext}}(\mathbf{r}_1) \\ - \nabla_{\mathbf{r}_1} (\rho(\mathbf{r}_1) c_p^{(1)}(\mathbf{r}_1)) + c_p^{(1)}(\mathbf{r}_1) \nabla_{\mathbf{r}_1} \rho(\mathbf{r}_1) = 0. \end{aligned} \quad (32)$$

Equation (32) involves only the one-body direct correlation function,  $c_p^{(1)}$ , defined in its standard form by Eq. (24). However, the bulk compressibility pressure is typically expressed in terms of the two-body direct correlation function,  $c^{(2)}$ . Therefore, we seek to re-express  $\rho(\mathbf{r}_1) c_p^{(1)}(\mathbf{r}_1)$  using the method of ‘‘functional line integration’’ [51]. By reintegrating the functional derivative of  $\rho(\mathbf{r}_1) c_p^{(1)}(\mathbf{r}_1)$  with respect to the density we obtain the formal result

$$\begin{aligned} \rho(\mathbf{r}_1) c_p^{(1)}(\mathbf{r}_1) \\ = \int d\mathbf{r}_2 \int_0^{\rho(\mathbf{r}_2)} d\rho'(\mathbf{r}_2) \frac{\delta(\rho(\mathbf{r}_1) c_p^{(1)}(\mathbf{r}_1))}{\delta \rho(\mathbf{r}_2)} \Big|_{\rho(\mathbf{r}_2)=\rho'(\mathbf{r}_2)}, \end{aligned}$$

where at each spatial point  $\mathbf{r}_2$  we integrate from an empty system (zero density) up to the density of interest,  $\rho(\mathbf{r}_2)$ . Evaluation of the functional derivative then yields

$$\begin{aligned} \rho(\mathbf{r}_1) c_p^{(1)}(\mathbf{r}_1) &= \int d\mathbf{r}_2 \int_0^{\rho(\mathbf{r}_2)} d\rho'(\mathbf{r}_2) (\rho'(\mathbf{r}_1) c^{(2)}(\mathbf{r}_1, \mathbf{r}_2; [\rho'])) \\ &\quad + \delta(\mathbf{r}_1 - \mathbf{r}_2) c_p^{(1)}(\mathbf{r}_1; [\rho']) \\ &= \int d\mathbf{r}_2 \int_0^{\rho(\mathbf{r}_2)} d\rho'(\mathbf{r}_2) \rho'(\mathbf{r}_1) c^{(2)}(\mathbf{r}_1, \mathbf{r}_2; [\rho']) \\ &\quad + \int_0^{\rho(\mathbf{r}_1)} d\rho'(\mathbf{r}_1) c_p^{(1)}(\mathbf{r}_1; [\rho']). \end{aligned}$$

This result can be substituted into Eq. (32) to give

$$\begin{aligned} 0 &= \nabla_{\mathbf{r}_1} \rho(\mathbf{r}_1) + \rho(\mathbf{r}_1) \nabla_{\mathbf{r}_1} \beta V_{\text{ext}}(\mathbf{r}_1) + c_p^{(1)}(\mathbf{r}_1) \nabla_{\mathbf{r}_1} \rho(\mathbf{r}_1) \\ &\quad - \int d\mathbf{r}_2 \int_0^{\rho(\mathbf{r}_2)} d\rho'(\mathbf{r}_2) \nabla_{\mathbf{r}_1} (\rho'(\mathbf{r}_1) c^{(2)}(\mathbf{r}_1, \mathbf{r}_2; [\rho'])) \\ &\quad - \nabla_{\mathbf{r}_1} \int_0^{\rho(\mathbf{r}_1)} d\rho'(\mathbf{r}_1) c_p^{(1)}(\mathbf{r}_1; [\rho']). \end{aligned} \quad (33)$$

We henceforth specialize to external fields which impose a planar geometry, such that the density profile and the two-body correlation functions exhibit cylindrical symmetry. As pointed out above, for the case of a planar hard-wall located at  $z = 0$  the density only varies in the  $z$  direction,  $\rho(\mathbf{r}) = \rho(z)$ . Equation (33) can then be integrated to yield

$$\begin{aligned} \rho_b - \rho_w = & \overbrace{\int d\mathbf{r}_2 \int_0^{\rho(z_2)} d\rho'(z_2) \rho'_b c^{(2)}(z_1 = \infty, z_2, r_2; [\rho'])}^{\text{A}} \\ & + \underbrace{\int_0^{\rho_b} d\rho'_b c_{p,b}^{(1)}(\rho'_b)}_{\text{B}} - \underbrace{\int_{-\infty}^{\infty} dz_1 c_p^{(1)}(z_1) \frac{d}{dz_1} \rho(z_1)}_{\text{C}}, \end{aligned} \quad (34)$$

where we have used  $\rho(z_1 \rightarrow \infty) = \rho_b$  and  $\rho(z_1 \rightarrow -\infty) = 0$ , as in Eq. (26). To connect this expression with the bulk compressibility pressure we analyze each of the three terms labeled A, B, and C in Eq. (34) separately.

**Term A:** Having  $z_1 \rightarrow \infty$  as an argument of the two-body correlation function  $c^{(2)}$  (which is of finite range) has the consequence that only bulk values contribute to the integral over the coordinate labeled 2, thus

$$\begin{aligned} \text{Term A} &= \int_0^{\rho_b} d\rho'_b \rho'_b \int d\mathbf{r}_{12} c_b^{(2)}(r_{12}; [\rho'_b]) \\ &= \int_0^{\rho_b} d\rho'_b \rho'_b \tilde{c}_b^{(2)}(q=0; \rho'_b) \\ &= -\beta P_{\text{exc}}^c, \end{aligned}$$

where  $\tilde{c}_b^{(2)}(q=0)$  is the Fourier transform of the two-body direct correlation function in the zero wave vector limit. The second equality gives the well-known integral giving the excess (over ideal) pressure in the compressibility route,  $P_{\text{exc}}^c$ , see Ref. [52].

**Term B:** This term does not require further manipulation and it can be given a clear physical interpretation. By identifying the bulk one-body direct correlation function,  $c_{p,b}^{(1)}$ , with the excess reduced chemical potential,  $\mu_{\text{exc}}$ , it follows that

$$\begin{aligned} \int_0^{\rho_b} d\rho'_b c_{p,b}^{(1)}(\rho'_b) &= - \int_0^{\rho_b} d\rho'_b \beta \mu_{\text{exc}}(\rho'_b) \\ &= - \int_0^{\rho_b} d\rho'_b \beta \frac{\partial f_{\text{exc}}}{\partial \rho'_b} \\ &= -\beta f_{\text{exc}}(\rho_b), \end{aligned}$$

where  $f_{\text{exc}} = F_{\text{exc}}/V$  is the bulk excess Helmholtz free energy per unit volume.

**Term C:** Using that  $c_p^{(1)}$  evaluated at a bulk density becomes position independent,

$$\begin{aligned} - \int_{-\infty}^{\infty} dz_1 \frac{d\rho(z_1)}{dz_1} c_p^{(1)}(z_1) &= - \int_0^{\rho_b} d\rho'_b c_p^{(1)}(z_1; [\rho'_b]) \\ &= - \int_0^{\rho_b} d\rho'_b c_{p,b}^{(1)}(\rho'_b) \\ &= \text{Term B}. \end{aligned}$$

Terms B and C cancel out and we finally get

$$\begin{aligned} \rho_w = \rho_b - \int_0^{\rho_b} d\rho'_b \rho'_b \tilde{c}_b^{(2)}(q=0; \rho'_b) \\ = \beta P_{\text{id}} + \beta P_{\text{exc}}^c = \beta P^c. \end{aligned} \quad (35)$$

We have thus proven the contact theorem for the compressibility route, namely, that if one uses Eq. (31) to calculate the density profile at a hard-wall, then the contact density will correspond to the reduced compressibility pressure,  $\beta P^c$ .

So far all our definitions and analytical considerations were not constrained to any specific system. However, at this point, to implement these general frameworks and show numerical results, we will focus on a particular simple model.

### F. Hard-sphere FMT

We now specialize to the minimal fluid model, which we take to be hard-spheres of radius  $R$  in three dimensions. The force-DFT approach is in no way restricted to this particular system. Hard-spheres simply provide a convenient test-case for which FMT gives an accurate approximation to the excess Helmholtz free energy functional,

$$\beta F_{\text{exc}}[\rho] = \int d\mathbf{r}_1 \Phi(\{n_\alpha(\mathbf{r}_1)\}). \quad (36)$$

The original Rosenfeld formulation of FMT [53] employs the following reduced excess free energy density

$$\Phi = -n_0 \ln(1 - n_3) + \frac{n_1 n_2 - \mathbf{n}_1 \cdot \mathbf{n}_2}{1 - n_3} + \frac{n_2^3 - 3n_2 \mathbf{n}_2 \cdot \mathbf{n}_2}{24\pi(1 - n_3)^2}.$$

The weighted densities are generated by convolution

$$n_\alpha(\mathbf{r}_1) = \int d\mathbf{r}_2 \rho(\mathbf{r}_2) \omega_\alpha(\mathbf{r}_1 - \mathbf{r}_2), \quad (37)$$

where the weight functions,  $\omega_\alpha$ , are characteristic of the geometry of the spheres. Of the six weight functions, four are scalars

$$\begin{aligned} \omega_3(\mathbf{r}) &= \Theta(R - r), & \omega_2(\mathbf{r}) &= \delta(R - r), \\ \omega_1(\mathbf{r}) &= \frac{\delta(R - r)}{4\pi R}, & \omega_0(\mathbf{r}) &= \frac{\delta(R - r)}{4\pi R^2}, \end{aligned}$$

and two are vectors (indicated by bold indices)

$$\omega_2(\mathbf{r}) = \mathbf{e}_r \delta(R - r), \quad \omega_1(\mathbf{r}) = \mathbf{e}_r \frac{\delta(R - r)}{4\pi R},$$

where  $\mathbf{e}_r = \mathbf{r}/r$  is a unit vector.

Applying the definition (24) for  $c_p^{(1)}$  to the free energy (36) generates the following approximate form for the one-body direct correlation function:

$$c_p^{(1)}(\mathbf{r}_1) = - \sum_\alpha \int d\mathbf{r}_2 \Phi'_\alpha(\mathbf{r}_2) \omega_\alpha(\mathbf{r}_2), \quad (38)$$

where the summation runs over all scalar and vector indices,  $\Phi'_\alpha = \partial \Phi / \partial n_\alpha$ , and  $\mathbf{r}_{21} = \mathbf{r}_2 - \mathbf{r}_1$ . The function  $\Phi'_\alpha$  is a vector quantity when  $\alpha$  takes the value **1** or **2**, in which case a scalar product with the corresponding vectorial weight function is implied in Eq. (38), otherwise it is a scalar function.

Taking two functional derivatives of the free energy (36) generates the following expression for the two-body direct

correlation function:

$$c^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = - \sum_{\alpha\beta} \int d\mathbf{r}_3 \omega_\alpha(\mathbf{r}_{31}) \Phi''_{\alpha\beta}(\mathbf{r}_3) \omega_\beta(\mathbf{r}_{32}), \quad (39)$$

where  $\Phi''_{\alpha\beta} = \partial^2 \Phi / \partial n_\alpha \partial n_\beta$ . For a detailed descriptions how to implement Eq. (39) in planar and spherical geometries we refer the reader to Ref. [34].

The inhomogeneous OZ equation

$$h(\mathbf{r}_1, \mathbf{r}_2) = c^{(2)}(\mathbf{r}_1, \mathbf{r}_2) + \int d\mathbf{r}_3 h(\mathbf{r}_1, \mathbf{r}_3) \rho(\mathbf{r}_3) c^{(2)}(\mathbf{r}_3, \mathbf{r}_2) \quad (40)$$

connects the two-body direct correlation function,  $c^{(2)}$ , with the total correlation function,  $h$ . The latter is related to the two-body density according to

$$\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = \rho(\mathbf{r}_1) \rho(\mathbf{r}_2) (h(\mathbf{r}_1, \mathbf{r}_2) + 1). \quad (41)$$

Substitution of Eq. (39) into the inhomogeneous OZ Eq. (40) yields a linear integral equation which can be solved for  $h$ , given the one-body density as input. The two-body density defined by Eq. (41) is thus a known functional of the one-body density, as required for implementation of the force-DFT.

Numerical evaluation of the right-hand side of Eq. (39) followed by iterative solution of the inhomogeneous OZ Eq. (40) is a demanding, yet well-defined and ultimately manageable, task. For researchers familiar with standard potential-DFT implementations (which operate purely on the one-body level) working with two-body numerics represents a significant step. However, having explicit access to the two-body correlations provides a much deeper insight into the particle microstructure and this benefit thus outweighs the increased computational complexity.

### G. Implementation in planar geometry

Now that we have specified the model of interest (hard-spheres) we choose to henceforth restrict our attention to planar geometry, for which the two-body correlation functions can be expressed using the cylindrical coordinates,  $z_1$ ,  $z_2$ , and  $r_2$  (see Refs. [14,34]). Although neither the potential- nor the force-DFT are limited to any particular geometry, our choice to focus on the planar case enables us to make connection to the contact sum-rules proven analytically in Secs. III D and III E.

To implement force-DFT, as expressed by the central Eq. (20), we begin by integrating Eq. (18) to obtain the function  $c_f^{(1)}$ . This yields

$$c_f^{(1)}(z) - c_f^{(1)}(0) = \int_0^z dz_1 \frac{2\pi}{\rho(z_1)} \int_{z_1-1}^{z_1+1} dz_2 (z_1 - z_2) \rho^{(2)}(z_1, z_2, r_2^*), \quad (42)$$

with  $r_2^* \equiv \sqrt{1 - (z_1 - z_2)^2}$  and where the particle diameter has been set to unity. Both the factor  $(z_1 - z_2)$  and the argument  $r_2^*$  appearing in the two-body density are consequences of the gradient of the hard-sphere potential in Eq. (18); a detailed derivation of Eq. (42) is given in Appendix C. The integration constant  $c_f^{(1)}(0)$  is unknown, but it does not have to be determined to calculate the density profile. Defining

a new parameter  $\alpha \equiv \beta\mu + c_f^{(1)}(0)$ , we obtain the following expression:

$$\rho(z) = e^{\beta(\mu - V_{\text{ext}}(z) + c_f^{(1)}(z))} = e^\alpha e^{-\beta V_{\text{ext}}(z) + \int_0^z dz_1 \frac{2\pi}{\rho(z_1)} \int_{-\infty}^{\infty} dz_2 (z_1 - z_2) \rho^{(2)}(z_1, z_2, r_2^*)}. \quad (43)$$

If we choose the average number of particles  $\langle N \rangle = \int_{-\infty}^{\infty} dz \rho(z)$  to be conserved, then the corresponding value of  $\alpha$  can be determined from

$$e^\alpha = \frac{\langle N \rangle}{\int_{-\infty}^{\infty} dz e^{-\beta V_{\text{ext}}(z) + 2\pi \int_0^z dz_1 \int_{-\infty}^{\infty} dz_2 (z_1 - z_2) \frac{\rho^{(2)}(z_1, z_2, r_2^*)}{\rho(z_1)}}, \quad (44)$$

which circumvents the need to prescribe  $c_f^{(1)}(0)$  and  $\mu$  independently. Given a method to calculate the two-body density from a given one-body profile, Eqs. (43) and (44) provide a closed system for numerical determination of the equilibrium density profile. This is possible since the two-body direct correlation function,  $c^{(2)}$ , is given as a functional of the one-body density in Eq. (39). The connection between  $c^{(2)}$  and  $\rho^{(2)}$  is given by combining the inhomogeneous OZ Eq. (40) with the definition (41). The three-dimensional integral appearing in Eq. (40) can be reduced to a manageable one-dimensional integral using the method of the Hankel transforms, as described in detail in Ref. [34]. The Hankel transform of the OZ Eq. (40) is given by

$$\bar{h}(z_1, z_2, k) = \bar{c}^{(2)}(z_1, z_2, k) + \int_{-\infty}^{\infty} dz_3 \bar{h}(z_1, z_3, k) \rho(z_3) \bar{c}^{(2)}(z_3, z_2, k), \quad (45)$$

where an overbar indicates a Hankel transformed quantity. In Ref. [34] a convenient analytical expression is given for the Rosenfeld form of  $\bar{c}^{(2)}$ .

On the other hand, the implementation of potential-DFT to obtain the density profile in planar geometry is a standard procedure in FMT studies. The EL Eq. (23) in planar geometry reads

$$\rho(z) = e^{\beta(\mu - V_{\text{ext}}(z) + c_p^{(1)}(z))}, \quad (46)$$

where  $c_p^{(1)}(z)$  is given by the planar version of Eq. (38) (see Ref. [34]).

### H. Numerical results for hard-spheres at a hard-wall

To calculate the density profile from force-DFT, we need to choose as input a value for the average number of particles,  $\langle N \rangle$ . In contrast, the potential-DFT takes as input the reduced chemical potential,  $\beta\mu$ . To enable comparison of the results from the two different approaches, we first calculate the potential-DFT density profiles for a given value of the reduced chemical potential and then calculate the average number of particles in the system by spatial integration. This value is then used as input in the force-DFT calculation. The quantities of relevance for testing the wall contact theorem are the bulk density [taken as  $\rho_b \approx \rho(z \rightarrow \infty)$ ] and the density at the wall,  $\rho_w$ . As we employ the Rosenfeld functional, the compressibility and virial pressures are identical to those of

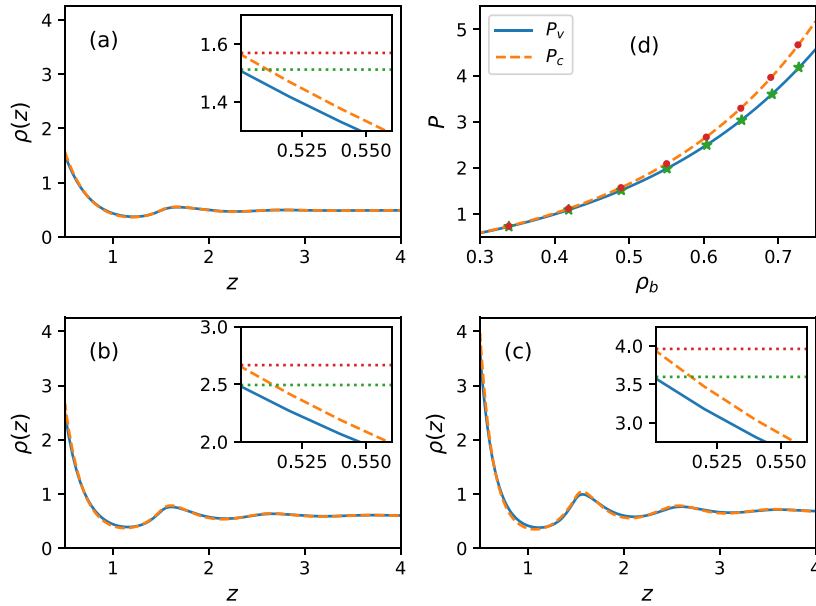


FIG. 2. Density profiles and pressure at a hard-wall. Numerical results for hard-spheres at a hard-wall calculated using potential-DFT (dashed orange lines), force-DFT (full blue lines). The density curves from potential-DFT shown in the panels (a), (b), and (c) were calculated at reduced chemical potentials  $\beta\mu = 3, 5,$  and  $7,$  respectively. The corresponding force-DFT density curves shown in the same panels were obtained by setting the average number of particles to match the values from potential-DFT. The top right panel (d) shows the analytic PY compressibility and virial pressures, together with numerical contact-values from potential-DFT (red filled circles) and force-DFT (green stars). These are also indicated by dotted horizontal lines in the zoomed density profiles shown in the insets in panels (a), (b), and (c), following the same color scheme.

the well-known Percus-Yevick (PY) integral equation theory [1], which are given by

$$\begin{aligned} \frac{\beta P^c}{\rho_b} &= \frac{1 + \eta + \eta^2}{(1 - \eta)^3}, \\ \frac{\beta P^v}{\rho_b} &= \frac{1 + 2\eta + 3\eta^2}{(1 - \eta)^2}, \end{aligned} \quad (47)$$

where  $\eta = \frac{4\pi}{3} \rho_b R^3$  is the packing fraction.

In Fig. 2, we show three sets of representative density profiles and the contact density as a function of  $\rho_b$  compared with the expected pressure. In the panel of Fig. 2(a) the reduced chemical potential is rather low,  $\beta\mu = 3,$  and the resulting potential- and force-DFT density profiles are very similar. Zooming to inspect the contact value highlights the difference between the two profiles and shows that our numerical data are highly consistent with the expected analytical pressures (indicated by the horizontal dotted lines, red for the compressibility route and green for the virial route). The panel of Fig. 2(b) is for  $\beta\mu = 5$  and, although some slight differences begin to emerge in the oscillations of the two density profiles, the contact values remain in excellent agreement with the respective analytical predictions. The panel of Fig. 2(c) is for  $\beta\mu = 7$  and shows more significant deviation of the density oscillations, but the contact densities still remain consistent with Eqs. (47). We find that both the oscillation amplitude

and contact density from force-DFT are lower than those of the potential-DFT.

We have performed similar calculations for a wider set of reduced chemical potentials. The panel of Fig. 2(d) shows the resulting contact densities as a function of  $\rho_b.$  This discrete set of points are shown together with the analytical curves from Eqs. (47), exhibiting an excellent level of agreement for a broad range of bulk densities. From our numerical results it is clear that the force-DFT does correspond to the virial route. This demonstrates that we have constructed a method by which DFT calculations can be reliably performed within the “virial realm” instead of the “compressibility realm,” which seemed to be the only one accessible before.

#### IV. DYNAMICAL THEORY

##### A. Force-DDFT

The tools we have developed can be readily extended to explore the dynamics of the one-body density out-of-equilibrium. In the following we consider systems subject to overdamped Brownian dynamics (BD). These model dynamics are suitable for the present investigation for two primary reasons. First, in overdamped BD the temperature is per construction constant. Hence relating the dynamics to an equilibrium ensemble is more straightforward than it is in molecular dynamics. Second, the absence of inertia in BD leads to simpler dynamical behavior emerging on the

one-body level [33]. An example is acceleration-dependent viscosity, which arises in molecular dynamics, but not in overdamped BD [54].

For overdamped motion the dynamics of the  $N$ -body distribution function is dictated by the Smoluchowski equation [30]. Integration over  $N - 1$  position coordinates generates the exact equation of motion for the one-body density

$$\frac{\partial \rho(\mathbf{r}_1, t)}{\partial t} = -\nabla_{\mathbf{r}_1} \cdot \mathbf{j}(\mathbf{r}_1, t), \quad (48)$$

where the current is given by

$$\begin{aligned} \mathbf{j}(\mathbf{r}_1, t) = & -D_0 \rho(\mathbf{r}_1, t) \left( \nabla_{\mathbf{r}_1} \ln(\rho(\mathbf{r}_1, t)) + \nabla_{\mathbf{r}_1} \beta V_{\text{ext}}(\mathbf{r}_1) \right) \\ & + \int d\mathbf{r}_2 \frac{\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2, t)}{\rho(\mathbf{r}_1, t)} \nabla_{\mathbf{r}_1} \beta \phi_{12}, \end{aligned} \quad (49)$$

where  $D_0$  is the diffusion coefficient. The interparticle force,  $-\nabla_{\mathbf{r}_1} \beta \phi_{12}$ , appears explicitly in the integral term. Equations (48) and (49) form the basis of the force-DDFT. Calculation of the current requires the exact time-dependent two-body density,  $\rho^{(2)}$ , as an input quantity, which is not available for any interacting model of real interest. A workable approximation can be obtained by making the assumption that  $\rho^{(2)}$  is instantaneously equilibrated to the nonequilibrium density. This adiabatic approximation enables one to employ the two-body correlations calculated using the inhomogeneous OZ Eq. (40) (which is an equilibrium relation) to obtain the average interaction force at each time-step. Note that in-equilibrium the current (49) vanishes. Since the density is nonzero, the sum of the three terms in parentheses in Eq. (49) must also vanish and we recover the YBG Eq. (16). The time-dependent density of force-DDFT thus relaxes to the density profile of force-DFT in the long-time limit.

As for the equilibrium case, we only consider hard-spheres subject to external fields of planar geometry. The gradient inside the integral term of Eq. (49) must therefore be treated carefully to correctly capture the discontinuous hard-sphere interaction potential. Fortunately, for planar geometry the integral can be conveniently reduced to one-dimension and Eqs. (48) and (49) can be combined and rewritten as

$$\begin{aligned} \frac{1}{D_0} \frac{\partial \rho(z_1, t)}{\partial t} = & \frac{\partial}{\partial z_1} \left( \frac{\partial \rho(z_1, t)}{\partial z_1} + \rho(z_1, t) \frac{\partial \beta V_{\text{ext}}(z_1)}{\partial z_1} \right. \\ & \left. - 2\pi \int_{-\infty}^{\infty} dz_2 (z_1 - z_2) \rho^{(2)}(z_1, z_2, r_2^*, t) \right), \end{aligned} \quad (50)$$

where  $r_2^* = \sqrt{1 - (z_1 - z_2)^2}$  for the particle diameter set to unity. This corresponds to evaluating the two-body density only on the contact shell where the interparticle forces act. The force-DDFT generates the dynamics of the density profile in the virial realm, which contrasts and complements the standard potential-DDFT, which we recall in the following.

### B. Potential-DDFT

The current for potential-DDFT [30] is given by

$$\mathbf{j}(\mathbf{r}, t) = -D_0 \rho(\mathbf{r}, t) \nabla_{\mathbf{r}} (\ln(\rho(\mathbf{r}, t)) + \beta V_{\text{ext}}(\mathbf{r}) - c_p^{(1)}(\mathbf{r}, t)). \quad (51)$$

In the construction of the force-DDFT, in Sec. IV A, we applied an adiabatic approximation to the two-body density,  $\rho^{(2)}$ , and thus to the entire average interaction force. This approach explicitly implements the idea of instantaneous equilibration of  $\rho^{(2)}$  at each time-step. Here we exploit an equilibrium sum-rule (see Ref. [30]) to approximate the average interaction force using the gradient of the one-body direct correlation function,  $c_p^{(1)}$ , which results in Eq. (51). The consequence of making this approximation is that the potential-DDFT operates within the compressibility realm. The long-time limit of the density time-evolution then reduces to that of the potential-DFT. As already pointed out in the previous subsection, the current must vanish at equilibrium. In the present case this implies that the sum of terms in parentheses in Eq. (51) must vanish, which recovers the (gradient of) the EL Eq. (23).

For the present case of planar geometry, combining Eqs. (48) and (51) yields the following one-dimensional equation of motion

$$\begin{aligned} \frac{1}{D_0} \frac{\partial \rho(z, t)}{\partial t} = & \frac{\partial}{\partial z} \left( \frac{\partial \rho(z, t)}{\partial z} + \rho(z, t) \frac{\partial \beta V_{\text{ext}}(z)}{\partial z} \right. \\ & \left. - \rho(z, t) \frac{\partial c_p^{(1)}(z, t)}{\partial z} \right) \end{aligned} \quad (52)$$

for the density profile. This can be compared with the exact Eq. (50). Note that if we follow the adiabatic approximation scheme on the one-body density functional  $\rho^{(2)}[\rho]$ , then we get back Eq. (52) but with  $c_f^{(1)}$ , defined by Eq. (18), instead of  $c_p^{(1)}$ .

### C. Numerical results for hard-spheres in a harmonic-trap

To compare the predictions of force-DDFT with those of potential-DDFT we consider a simple benchmark test of the relaxational dynamics. The density is first equilibrated to a planar harmonic external potential,  $\beta V_{\text{ext}}(z) = A(z - z_0)^2$ , where we use the values  $A = 0.75$  and  $z_0 = 5$  inside of a computational domain covering the range from  $z = 0$  to  $z = 10$ . At time  $t = 0$  we instantaneously switch the harmonic-trap amplitude to the value  $A = 0.5$  and then use either Eq. (50) or Eq. (52) to calculate the relaxational time-evolution of the density toward the equilibrium state of the new trap. The time integration of Eqs. (50) and (52) is performed using forward Euler integration, which amounts to approximating the partial time derivative according to the following finite difference expression:

$$\frac{\partial \rho(z, t)}{\partial t} \approx \frac{\rho(z, t + \Delta t) - \rho(z, t)}{\Delta t},$$

where  $\Delta t$  is the time-step. In practice, the numerical realization of equilibrium as a long-time limit of the dynamics may be difficult to achieve due to the accumulation of discretization errors over many time-steps.

The left panels (a) and (b) of Fig. 3 show the time-evolution of the density obtained from the potential-DDFT. In the upper panel we show only the left-half of the symmetric density profile and in the lower panel we show a zoom of the density peak. The black dashed curves indicate the equilibrium densities obtained from potential-DFT and can be compared with our grand-canonical Monte Carlo simulation data [55], given by the silver dotted lines. The simulation is equilibrated

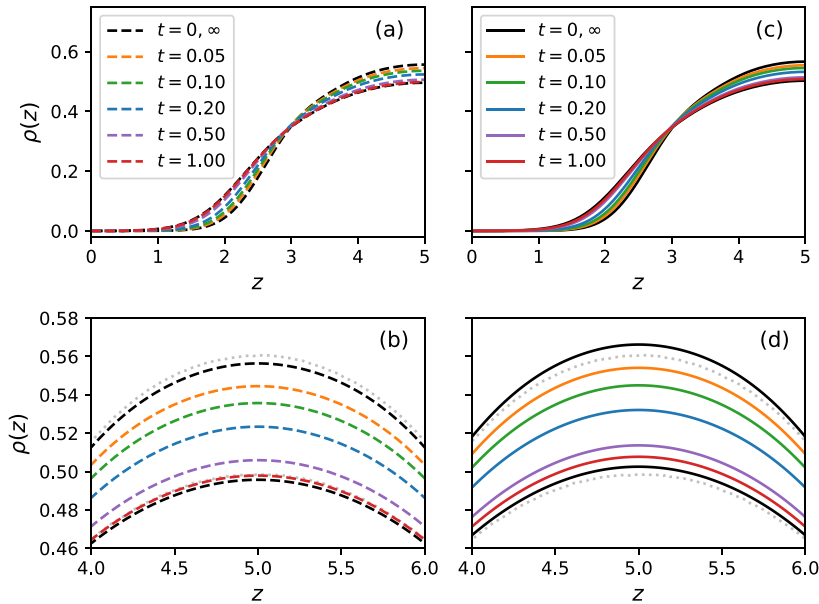


FIG. 3. Transient dynamics in a harmonic-trap. Time-evolution of the density following a discontinuous change in the trap amplitude from  $A = 0.75$  to  $0.5$  at time  $t = 0$ . The left panels (a) and (b) show the density obtained from potential-DDFT and the right panels (c) and (d) show the density from force-DDFT. The black lines (dashed for potential-DFT, solid for force-DFT) give the equilibrium initial and final states. The silver dotted lines in the lower row, in panels (b) and (d), show the initial equilibrium state and the final equilibrium state, as obtained from grand-canonical Monte Carlo (GCMC) simulations.

for  $10^5$  sweeps and sampled for  $10^7$  sweeps, the box size is  $30 \times 30 \times 20$ , where the unit of length is a hard-sphere diameter, and on average there are 2142 particles in the system.

The colored dashed lines in Fig. 3 show density profiles obtained from potential-DDFT for a selection of different times, which we give in units of particle diameter squared over diffusion coefficient,  $D_0$ . These results should be compared with those of the force-DFT and force-DDFT shown in the right panels (c) and (d) of Fig. 3, where we used the same colors as before to identify curves at equal times. We clearly see that, in this case, the force-DDFT relaxes more slowly than potential-DDFT, which implies that the average (repulsive) interaction force is stronger in the latter approximation. A possible explanation for that phenomenon is that the hard-sphere system has a very harshly repulsive interparticle potential, which strongly influences the spatial distribution of the particles and which is captured more effectively in force-DDFT.

The implementation of potential-DDFT is rather quick and simple since it only involves one-body functions and requires only a single Picard update at each time-step. On the other hand the force-DDFT is way more demanding since it involves solving the OZ equation at every time-step and then also requires a Picard update of the density. Not only are the analytical expressions more complicated, but also the numerical computational work. The shown curves therefore took significantly more computational time to be obtained, but they can nevertheless be calculated to high accuracy.

## V. CONCLUSIONS AND OUTLOOK

Starting from fundamental principles of Noether invariance, we have developed a force-based theory for the density profile both in- and out-of-equilibrium. The equilibrium theory shows that density profiles can be calculated via the virial route by following our explicit force-DFT scheme. This situation can be contrasted with the standard potential-DFT that is known to follow the compressibility route. The latter is a well-used result and often a crucial test in a significant number of DFT studies, so we provided a mathematical proof that explicitly shows that the planar hard-wall contact density from potential-DFT is given by the reduced *compressibility* pressure.

Our analytical proofs have been tailored to highlight the different outcomes from the two routes. If we had access to the exact Helmholtz free energy functional, then there would be no route-dependency. A more general proof of the contact theorem (shown in Appendix B or in Refs. [1,36,37]) would then be sufficient. We thus suggest to exploit the differences between the density profiles from the virial and compressibility routes to test, scrutinize, and ultimately attempt to improve, approximate DFT schemes. Working with inhomogeneous two-body correlation functions, as implemented explicitly in the force-DFT, is both analytically and numerically more demanding than using the standard potential-DFT scheme. However, facing the increase in complexity is rewarded by gaining deeper insight into the theoretical structure of DFT.

Moreover working on the two-body level allows to explicitly incorporate the pairwise interparticle interactions and take direct account of their influence on the spatial distribution of the particles. While carrying out force-DFT calculations comes at an increased numerical cost, the additional workload (both in terms of implementation and runtime) is far from prohibitive and practical research can be efficiently performed.

The distinction between the virial and compressibility routes is known to be important in the integral-equation theory of bulk liquids [1]. Here we reveal an analogous scenario for the theory of inhomogeneous fluids, which is an interesting result in its own right. While both approaches, the conventional potential-DFT and the force-DFT, construct the density-functional dependencies in alternative forms, both approaches start from the same approximation for the excess free energy functional. This offers clear pathways toward improved theories that enforce self-consistency in a variety of ways. For example, the virial and compressibility routes could be mixed in the spirit of liquid-state integral-equation theories, using approximations analogous to the Rogers-Young [56] or Carnahan-Starling theories [57]. Another possibility would be to enforce the exact core-condition on the total correlation function  $h(\mathbf{r}_1, \mathbf{r}_2)$  and thus improve the description of the inhomogeneous two-body correlations.

A particularly appealing feature of the force-DFT is that it naturally generalizes to treat nonequilibrium systems. At the most fundamental level, particles are moved by forces, rather than by potentials, and hence forces form a solid basis for developing a dynamical theory [24,25,32,33]. The adiabatic approach that we have employed closes the dynamical description on the level of the one-body density. On this basis we have explored the dynamical behavior of hard-spheres inside of a harmonic-trap under a temporal switching protocol. We found that the density dynamics that follow from potential- and force-DDFT differ significantly from each other. Not only are the equilibrium (long-time) profiles different, but so are the relaxation rates.

The starting equations of force-DDFT, namely, Eqs. (48) and (49), are exact. If we had access to the exact  $\rho^{(2)}$  as a functional of the one-body density, then we could calculate the exact time-evolution of  $\rho$ . As this information is not available, we close the theory by making an adiabatic approximation for  $\rho^{(2)}$ , thus assuming that it equilibrates at each time-step. This thinking is also captured in the adiabatic construction of power functional theory [32,33].

A point of interest is to attempt to close at a higher level of the correlation function hierarchy, with the aim to provide a first-principles superadiabatic dynamical theory. It is hard to conceive that such progress could be made without a force-based approach. The force-DFT that we present here thus represents a first step toward full treatment of nonequilibrium. Furthermore force-DDFT, when compared with potential-DDFT, has the clear benefit that the average interparticle interaction force does not appear automatically as a gradient term in the exact Eq. (49). If this were the case, as it is in potential-DDFT, then it would exclude *de facto* all nonconservative forces. Force-DDFT thus leaves the door open for future studies of driven systems such as systems with shear flows. It is well known that the adiabatic approximation within standard DDFT fails for shear fields [58]. However,

it would be interesting to investigate shear flows with higher order force-DDFT to check on the validity of these considerations and approximations.

Our derivation of the force balance (YBG) relationship from local Noether invariance in Sec. III A is based on considering a local displacement field  $\epsilon(\mathbf{r})$ . This object bears similarities with the vector field that maps between positions in the Lagrangian and Eulerian picture in continuum mechanics. The connections between this thinking and a local density functional treatment were recently explored by Sprik. Specifically he considered the case of dielectric fluids [59]. Our more microscopic formulation could possibly help to shed some light on the relationship of the continuum mechanical force balance and the DFT equilibrium equation. The connections to the crystalline state and crystal deformations are also worth exploring, as addressed by Sprik within continuum mechanics [60], by Fuchs and coworkers from a more microscopic point of view [61,62] and recently by Lin et al. [63] within DFT.

We have shown that the hard-sphere system is described within fundamental measure theory to a good level of self-consistency. Going beyond hard-spheres and exploring the force-DFT for functionals that describe interparticle attraction would be interesting. Such work could be already revealing in the context of the standard mean-field functional [16].

In the context of power functional theory [33] the force-based theories could play a role in the description of the adiabatic state as applied to bulk and interfaces of active Brownian particles [64–68], to flow phenomena in overdamped systems [25,69], to shear [70,71], and the van Hove function [72,73].

We would expect the treatment of long-ranged forces as they occur in charged fluids to require extra care in dealing with divergent integrals. Nevertheless, application of force-DFT to Coulombic systems could be revealing for the behaviour of the electrical double layer [74], the differential capacitance [75], as well as for long-ranged decay of correlations, as recently explored for the restricted primitive model [76].

Burke and his collaborators have recently put forward a new approach to electronic DFT. Their “blue electron approximation” [28] offers a concrete way to work efficiently at finite temperatures within what they call the conditional probability DFT [29]. In the high-temperature limit an analogy to Percus’ classical test particle limit arises [28]. As their method works on the two-body level cross fertilization with our present approach is not inconceivable.

Molecular DFT generalizes classical DFT to systems with orientational degrees of freedom, see, e.g., Refs. [4,5,77,78]. Various ingenious ways of dealing efficiently with the associated numerical burdens of accounting for the molecular Euler angles have been formulated, see, e.g., Ref. [79]. Whether the present approach can help to describe the corresponding forces and torques in such systems is an interesting point for future work. Also going beyond the planar (effective one-dimensional) geometry and addressing fully inhomogeneous three-dimensional situations [80–83] constitutes an exciting, yet formidable, research task.

As the two-body correlation functions upon which the force-DFT is built are directly accessible via many-body simulation (see Ref. [84] aimed at the direct correlation function),

one can wonder whether using simulations data as input would allow to construct force-DFT approximations. This could possibly be aided by machine-learning techniques [85].

The two-body density gives information about the probability to find a particle at position  $\mathbf{r}_2$  given that there is a particle at position  $\mathbf{r}_1$ . This enables the pair interaction forces acting within the fluid to be analysed in detail. Moreover, multiplying the two-body density with the gradient of the pair-potential allows the average pair interaction force to be calculated explicitly and thus, in the case of hard interparticle interactions, incorporates the particle geometry directly.

#### ACKNOWLEDGMENTS

S.M.T. and J.M.B. thank G. T. Hamsler for her critical judgment and for taking the time to go through the whole manuscript several times. M.S. acknowledges useful discussions with Daniel de las Heras. This work is partially supported by the German Research Foundation (DFG) via Project No. 436306241.

#### APPENDIX A: CANONICAL TRANSFORMATION

The transformation given by Eqs. (2) and (3) is canonical and it hence preserves the phase-space volume element. That the transformation is canonical can be demonstrated by considering a generating function  $\mathcal{G}$  [86], which for the present transformation has an explicit form given by

$$\mathcal{G} = \sum_{i=1}^N \mathbf{p}'_i \cdot (\mathbf{r}_i + \boldsymbol{\epsilon}(\mathbf{r}_i)). \quad (\text{A1})$$

As  $\mathcal{G}$  is a function of the original coordinates and of the new momenta, the transformation equations are generated via  $\mathbf{r}'_i = \partial\mathcal{G}/\partial\mathbf{p}'_i$  and  $\mathbf{p}_i = \partial\mathcal{G}/\partial\mathbf{r}_i$ . Using the explicit form (A1) and expanding to lowest order in  $\boldsymbol{\epsilon}(\mathbf{r})$  yields Eqs. (2) and (3) in a straightforward way.

The canonical generator  $\mathcal{G}$  defined in Eq. (3) is a function of the original coordinates  $\mathbf{r}_1, \dots, \mathbf{r}_N$  and of the new momenta  $\mathbf{p}'_1, \dots, \mathbf{p}'_N$ . For the case of such dependence the original Hamiltonian  $H$  and the transformed Hamiltonian  $H'$  are related by the general transformation [86]:

$$H' = H + \frac{\partial\mathcal{G}}{\partial t}. \quad (\text{A2})$$

As the generator (3) carries no explicit time dependence, the last term in Eq. (A2) vanishes, and  $H' = H$ . This invariance of the Hamiltonian under the considered transformation implies the trivial replacement of variables, i.e., that the transformed Hamiltonian depends on the transformed coordinates and momenta, i.e.,  $H'(\mathbf{r}'_1, \dots, \mathbf{r}'_N, \mathbf{p}'_1, \dots, \mathbf{p}'_N)$ . Then by construction, the equations of motion, when expressed in the new phase space variables, are generated from the standard Hamiltonian procedure:  $d\mathbf{p}'_i/dt = -\partial H'/\partial\mathbf{r}'_i$  and  $d\mathbf{r}'_i/dt = \partial H'/\partial\mathbf{p}'_i$ .

#### APPENDIX B: GENERAL DERIVATION OF THE CONTACT THEOREM

The following Appendix shows a derivation of the contact theorem appropriate to situations in which all quantities are known exactly. For this reason we use the generic notation

$c^{(1)}$  and  $P$  for the one-body direct correlation function and the pressure, respectively.

Let us consider a hard-wall such that the distance of closest approach of a particle is located at  $z = 0$ . We assume that the system reaches a bulk-like state at and around a (large) distance  $L$  away from the wall. To have a closed system in the  $z$  direction, we consider a second “ultrasoft” wall that vanishes for  $z < L$ , and then gives a slowly rising energy penalty upon increasing  $z$ , which ultimately diverges  $V_{\text{ext}}(z \rightarrow \infty) = \infty$ .

We recall the global Noether identity of vanishing total interparticle force

$$\int_{-\infty}^{\infty} dz \rho(z) \frac{dc^{(1)}(z)}{dz} = 0, \quad (\text{B1})$$

where  $c^{(1)}(\mathbf{r}) = -\beta\delta F_{\text{exc}}[\rho]/\delta\rho(\mathbf{r})$  is the one-body direct correlation function. The integrand in Eq. (B1) is, up to a factor of thermal energy, the locally resolved interparticle force density,  $k_B T \rho(\mathbf{r}) \nabla_{\mathbf{r}} c^{(1)}(\mathbf{r})$ , acting in the  $z$  direction. One can argue equivalently and independently (see, e.g., Ref. [50]), that Eq. (B1) holds on the basis of Newton’s third law.

Here we rather start from the alternative form

$$\int_{-\infty}^{\infty} dz c^{(1)}(z) \frac{d\rho(z)}{dz} = 0, \quad (\text{B2})$$

which is straightforwardly obtained from the Noether sum-rule (B1) via integration by parts; circumstances must be such that the boundary terms at infinity vanish. More significantly, within a DFT context, it is the form (B1) that is the primary result from applying Noether’s theorem to the invariance of the excess free energy functional  $F_{\text{exc}}[\rho]$  upon spatial shifting of the system [50].

Here we proceed directly with the form (B2), treating three spatial regions separately: the vicinity of the hard-wall,  $-\Delta < z < \Delta$ , where  $\Delta$  is a small parameter (as compared to all other lengthscales in the system); the region from the wall to the bulk-like state, i.e.,  $\Delta < z < L$ ; and the soft wall region,  $z > L$ . In the following, the limit  $\Delta \rightarrow 0$  is implicit.

In the vicinity of the hard-wall we can identify the leading term as

$$\begin{aligned} \int_{-\Delta}^{\Delta} dz c^{(1)}(z) \frac{d\rho(z)}{dz} &= \int_{-\Delta}^{\Delta} dz c^{(1)}(z) \delta(z) \rho(0) \\ &= c^{(1)}(0) \rho_w \\ &= \rho_w \ln(\rho_w) - \beta\mu \rho_w, \end{aligned} \quad (\text{B3})$$

where  $\rho_w = \rho(0)$  and in the last step we have used the EL equation

$$c^{(1)}(z) = \ln[\rho(z)] + \beta V_{\text{ext}}(z) - \beta\mu,$$

to express the one-body direct correlation function at contact,  $c^{(1)}(0)$ ; note that the external potential term gives no contribution as  $V_{\text{ext}}(0^+) = 0$ ; furthermore,  $c^{(1)}(z)$  is continuous at  $z = 0$ .

In the region from between outside the wall and the bulk, i.e., for  $\Delta < z < L$ , the external potential vanishes and we have

$$\begin{aligned} \int_{\Delta}^L dz c^{(1)}(z) \frac{d\rho(z)}{dz} &= \int_{\Delta}^L dz [\ln(\rho(z)) - \beta\mu] \frac{d\rho(z)}{dz} \\ &= [\rho(\ln(\rho) - 1) - \beta\mu\rho]_{\rho_w}^{\rho_b} \end{aligned}$$



$$= \rho_b(\ln(\rho_b) - 1 - \beta\mu) - \rho_w(\ln(\rho_w) - 1 - \beta\mu), \quad (\text{B4})$$

where in the first step we have again used the EL Eq. (B1) and the bulk density is defined as  $\rho_b = \rho(L)$ .

In the soft wall regime, i.e., for  $L < z$ , the density inhomogeneity is so weak that a local density approximation becomes accurate and hence

$$\int_L^\infty dz c^{(1)}(z) \frac{d\rho(z)}{dz} = \int_{\rho_b}^0 d\rho c^{(1)}(\rho) = f_{\text{exc}}(\rho_b) \\ = -P - \rho_b[\ln(\rho_b) - 1] + \mu\rho_b. \quad (\text{B5})$$

The upper limit in the density integral is  $\rho(z \rightarrow \infty) = 0$ , and  $f_{\text{exc}}(\rho_b)$  is the bulk excess free energy density per volume as a function of  $\rho_b$ . The value at the upper boundary of the density integration vanishes, as the system is infinitely dilute. The last step identifies the pressure  $P$ .

Adding up the three contributions (B3), (B4), and (B5) gives according to Noether invariance (B2) a vanishing result. Rewriting yields

$$\rho_w = \beta P, \quad (\text{B6})$$

which is the general form of the hard-wall sum-rule.

### APPENDIX C: PLANAR HARD-SPHERE FORCE INTEGRAL

In the following we derive the one-body direct correlation function  $c_f^{(1)}$ , given by Eq. (42), for hard-spheres in planar geometry. We start with the gradient of  $c_f^{(1)}$ , from Eq. (18), namely,

$$\nabla_{\mathbf{r}_1} c_f^{(1)}(\mathbf{r}_1) = - \int d\mathbf{r}_2 \frac{\rho^{(2)}(\mathbf{r}_1, \mathbf{r}_2)}{\rho(\mathbf{r}_1)} \nabla_{\mathbf{r}_1} \beta\phi_{12}, \quad (\text{C1})$$

and as a first step exploit the planar geometry. The symmetry simplifies the dependence on the position variables such that the one-body distributions only depend on the  $z$  coordinate. The two-body density,  $\rho^{(2)}$ , depends on  $z_1$ ,  $z_2$ , and  $r_2$  (see Ref. [34]). The distance between the two particle positions is then  $r_{12} = \sqrt{r_2^2 + (z_1 - z_2)^2}$ .

As  $c_f^{(1)}$  only depends on  $z$ , the gradient on the left-hand side of Eq. (C1) reduces to  $\mathbf{e}_z d/dz$ , where  $\mathbf{e}_z$  is the unit vector in the  $z$  direction. The interparticle interaction potential,  $\phi$ , depends only on  $r_{12}$ . This allows us to rewrite the gradient of  $\phi$  as a derivative with respect to this distance,  $\mathbf{e}_{r_{12}} d/dr_{12}$ , where  $\mathbf{e}_{r_{12}} = (\mathbf{r}_1 - \mathbf{r}_2)/r_{12}$  indicates the radial unit vector. Equation (C1) thus simplifies to

$$\frac{d c_f^{(1)}(z_1)}{dz_1} \mathbf{e}_z = - \int d\mathbf{r}_2 \frac{\rho^{(2)}(z_1, z_2, r_2)}{\rho(z_1)} \frac{d\beta\phi_{12}}{dr_{12}} \mathbf{e}_{r_{12}}. \quad (\text{C2})$$

We next express the  $\mathbf{r}_2$  integral in cylindrical coordinates, such that the  $z$  component of Eq. (C2) becomes

$$\frac{d c_f^{(1)}(z_1)}{dz_1} = - \frac{2\pi}{\rho(z_1)} \int_{-\infty}^{\infty} dz_2 \int_0^{\infty} dr_2 r_2 \rho^{(2)}(z_1, z_2, r_2) \\ \times \frac{d\beta\phi_{12}}{dr_{12}} \frac{(z_1 - z_2)}{r_{12}}. \quad (\text{C3})$$

To deal with the hard-sphere potential,  $\phi$ , we proceed as previously in Eq. (27). We therefore multiply the integrand in Eq. (C3) by  $1 = e^{\beta\phi_{12}} e^{-\beta\phi_{12}}$ . The second Boltzmann factor can be grouped together with the derivative of the interaction potential as  $e^{-\beta\phi_{12}} \frac{d\beta\phi_{12}}{dr_{12}} = -d e^{-\beta\phi_{12}}/dr_{12}$ . For the hard-sphere interaction potential the Boltzmann factor can be identified as a step function,  $e^{-\beta\phi_{12}} = \Theta(r_{12} - 1)$ , where  $\Theta$  indicates the Heaviside step function. The radial derivative then gives a Dirac  $\delta$  distribution,

$$\frac{d\Theta(r_{12} - 1)}{dr_{12}} = \delta(r_{12} - 1) = \frac{\delta(r_2 - r_2^*)}{|r_2/r_{12}|},$$

where  $r_2^* = \sqrt{1 - (z_1 - z_2)^2}$  is the cylindrical radial distance at contact for given coordinates  $z_1$  and  $z_2$ . This yields

$$\frac{d c_f^{(1)}(z_1)}{dz_1} = - \frac{2\pi}{\rho(z_1)} \int_{z_1-1}^{z_1+1} dz_2 (z_1 - z_2) \rho^{(2)}(z_1, z_2, r_2^*).$$

To obtain the desired Eq. (42), we then integrate with respect to  $z_1$  from 0 to  $z$ .

- 
- [1] J. P. Hansen and I. R. McDonald, *Theory of Simple Liquids*, 4th ed. (Academic Press, London, 2013).
  - [2] R. Evans, D. Frenkel, and M. Dijkstra, From simple liquids to colloids and soft matter, *Phys. Today* **72**, 38 (2019).
  - [3] S. R. Nagel, Experimental soft-matter science, *Rev. Mod. Phys.* **89**, 025002 (2017).
  - [4] M. Levesque, R. Vuilleumier, and D. Borgis, Scalar fundamental measure theory for hard spheres in three dimensions: Application to hydrophobic solvation, *J. Chem. Phys.* **137**, 034115 (2012).
  - [5] G. Jeanmairet, M. Levesque, and D. Borgis, Molecular density functional theory of water describing hydrophobicity at short and long length scales, *J. Chem. Phys.* **139**, 154101 (2013).
  - [6] R. Evans, M. C. Stewart, and N. B. Wilding, A unified description of hydrophilic and superhydrophobic surfaces in terms of the wetting and drying transitions of liquids, *Proc. Natl. Acad. Sci. USA* **116**, 23901 (2019).
  - [7] R. Evans and N. B. Wilding, Quantifying Density Fluctuations in Water at a Hydrophobic Surface: Evidence for Critical Drying, *Phys. Rev. Lett.* **115**, 016103 (2015).
  - [8] T. Eckert, N. C. X. Stuhlmüller, F. Sammüller, and M. Schmidt, Fluctuation Profiles in Inhomogeneous Fluids, *Phys. Rev. Lett.* **125**, 268004 (2020).
  - [9] R. Evans, M. C. Stewart, and N. B. Wilding, Critical Drying of Liquids, *Phys. Rev. Lett.* **117**, 176102 (2016).
  - [10] B. Chacko, R. Evans, and A. J. Archer, Solvent fluctuations around solvophobic, solvophilic, and patchy nanostructures and the accompanying solvent mediated interactions, *J. Chem. Phys.* **146**, 124703 (2017).
  - [11] D. Martin-Jimenez, E. Chacón, P. Tarazona, and R. Garcia, Atomically resolved three-dimensional structures of electrolyte aqueous solutions near a solid surface, *Nat. Commun.* **7**, 12164 (2016).

- [12] J. Hernández-Muñoz, E. Chacón, and P. Tarazona, Density functional analysis of atomic force microscopy in a dense fluid, *J. Chem. Phys.* **151**, 034701 (2019).
- [13] J. Muscatello, E. Chacón, P. Tarazona, and F. Bresme, Deconstructing Temperature Gradients Across Fluid Interfaces: The Structural Origin of the Thermal Resistance of Liquid-Vapor Interfaces, *Phys. Rev. Lett.* **119**, 045901 (2017).
- [14] S. M. Tschopp, H. D. Vuijk, A. Sharma, and J. M. Brader, Mean-field theory of inhomogeneous fluids, *Phys. Rev. E* **102**, 042140 (2020).
- [15] R. Evans, The nature of the liquid-vapour interface and other topics in the statistical mechanics of nonuniform, classical fluids, *Adv. Phys.* **28**, 143 (1979).
- [16] R. Evans, Density functionals in the theory nonuniform fluids, in *Fundamentals of Inhomogeneous Fluids*, edited by D. Henderson (Dekker, New York, NY, 1992).
- [17] For an overview of new developments in classical density functional theory, see R. Evans, M. Oettel, R. Roth, and G. Kahl, New developments in classical density functional theory, *J. Phys.: Condens. Matter* **28**, 240401 (2016).
- [18] J. Yvon, *Actualités Scientifiques et Industrielles* (Hermann & Cie, Hamburg, 1935).
- [19] M. Born and H. S. Green, A general kinetic theory of liquids I. The molecular distribution functions, *Proc. R. Soc. London A* **188**, 10 (1946).
- [20] B. Rotenberg, Use the force! Reduced variance estimators for densities, radial distribution functions, and local mobilities in molecular simulations, *J. Chem. Phys.* **153**, 150902 (2020).
- [21] D. de las Heras and M. Schmidt, Better Than Counting: Density Profiles from Force Sampling, *Phys. Rev. Lett.* **120**, 218001 (2018).
- [22] D. Borgis, R. Assaraf, B. Rotenberg, and R. Vuilleumier, Computation of pair distribution functions and three-dimensional densities with a reduced variance principle, *Mol. Phys.* **111**, 3486 (2013).
- [23] A. Purohit, A. J. Schultz, and D. A. Kofke, Force-sampling methods for density distributions as instances of mapped averaging, *Mol. Phys.* **117**, 2822 (2019).
- [24] D. de las Heras and M. Schmidt, Velocity Gradient Power Functional for Brownian Dynamics, *Phys. Rev. Lett.* **120**, 028001 (2018).
- [25] D. de las Heras and M. Schmidt, Flow and Structure in Nonequilibrium Brownian Many-Body Systems, *Phys. Rev. Lett.* **125**, 018001 (2020).
- [26] W. Tarantino and C. A. Ullrich, A reformulation of time-dependent Kohn–Sham theory in terms of the second time derivative of the density, *J. Chem. Phys.* **154**, 204112 (2021).
- [27] M.-L. M. Tchenkoue, M. Penz, I. Theophilou, M. Ruggenthaler, and A. Rubio, Force balance approach for advanced approximations in density functional theories, *J. Chem. Phys.* **151**, 154107 (2019).
- [28] R. J. McCarty, D. Perchak, R. Pederson, R. Evans, Y. Qiu, S. R. White, and K. Burke, Bypassing the Energy Functional in Density Functional Theory: Direct Calculation of Electronic Energies from Conditional Probability Densities, *Phys. Rev. Lett.* **125**, 266401 (2020).
- [29] R. Pederson, J. Chen, S. R. White, and K. Burke, Conditional probability density functional theory, *Phys. Rev. B* **105**, 245138 (2022).
- [30] A. J. Archer and R. Evans, Dynamical density functional theory and its application to spinodal decomposition, *J. Chem. Phys.* **121**, 4246 (2004).
- [31] U. M. B. Marconi and P. Tarazona, Dynamic density functional theory of fluids, *J. Chem. Phys.* **110**, 8032 (1999).
- [32] M. Schmidt and J. M. Brader, Power functional theory for Brownian dynamics, *J. Chem. Phys.* **138**, 214101 (2013).
- [33] M. Schmidt, Power functional theory for many-body dynamics, *Rev. Mod. Phys.* **94**, 015007 (2022).
- [34] S. M. Tschopp and J. M. Brader, Fundamental measure theory of inhomogeneous two-body correlation functions, *Phys. Rev. E* **103**, 042103 (2021).
- [35] P. Attard, *Thermodynamics and Statistical Mechanics* (Academic Press, San Diego, CA, 2002).
- [36] J. L. Lebowitz, Asymptotic value of the pair distribution near a wall, *Phys. Fluids* **3**, 64 (1960).
- [37] R. Lovett and M. Baus, A family of equivalent expressions for the pressure of a fluid adjacent to a wall, *J. Chem. Phys.* **95**, 1991 (1991).
- [38] F. van Swol and J. R. Henderson, Wetting and drying transitions at a fluid-wall interface: Density-functional theory versus computer simulation, *Phys. Rev. A* **40**, 2567 (1989).
- [39] P. Tarazona and R. Evans, A simple density functional theory for inhomogeneous liquids, *Mol. Phys.* **52**, 847 (1984).
- [40] R. Roth, Fundamental measure theory for hard-sphere mixtures: A review, *J. Phys.: Condens. Matter* **22**, 063102 (2010).
- [41] A. J. F. Siegert and E. Meeron, Generalizations of the virial and wall theorems in classical statistical mechanics, *J. Math. Phys.* **7**, 741 (1966).
- [42] D. Henderson, L. Blum, and J. L. Lebowitz, An exact formula for the contact value of the density profile of a system of charged hard spheres near a charged wall, *J. Electroanal. Chem.* **102**, 315 (1979).
- [43] J. R. Henderson, Statistical mechanics of fluids at spherical structureless walls, *Mol. Phys.* **50**, 741 (1983).
- [44] L. Blum, Contact theorems for rough interfaces, *J. Stat. Phys.* **75**, 971 (1994).
- [45] P. J. Upton, Fluids Against Hard Walls and Surface Critical Behavior, *Phys. Rev. Lett.* **81**, 2300 (1998).
- [46] P. Malfaretti and M. Bier, Local pressure for confined systems, *Phys. Rev. E* **97**, 022102 (2018).
- [47] D. Henderson and L. Blum, Some exact results and the application of the mean spherical approximation to charged hard spheres near a charged hard wall, *J. Chem. Phys.* **69**, 5441 (1978).
- [48] S. L. Carnie and D. Y. C. Chan, The statistical mechanics of the electrical double layer: Stress tensor and contact conditions, *J. Chem. Phys.* **74**, 1293 (1981).
- [49] J. P. Mallarino, G. Téllez, and E. Trizac, The contact theorem for charged fluids: From planar to curved geometries, *Mol. Phys.* **113**, 2409 (2015).
- [50] S. Hermann and M. Schmidt, Noether’s theorem in statistical mechanics, *Commun. Phys.* **4**, 176 (2021).
- [51] J. M. Brader and M. Schmidt, Free power dissipation from functional line integration, *Mol. Phys.* **113**, 2873 (2015).
- [52] M. Kasch, X. S. Chen, and F. Forstmann, The calculation of correlation functions in fluids by a weighted density concept, *Mol. Phys.* **75**, 415 (1992).

OPEN ACCESS

IOP Publishing

Journal of Physics A: Mathematical and Theoretical

J. Phys. A: Math. Theor. 55 (2022) 464003 (18pp)

<https://doi.org/10.1088/1751-8121/aca12d>

# Force balance in thermal quantum many-body systems from Noether's theorem

Sophie Hermann\*  and Matthias Schmidt\* 

Theoretische Physik II, Physikalisches Institut, Universität Bayreuth, D-95447 Bayreuth, Germany

E-mail: [Sophie.Hermann@uni-bayreuth.d](mailto:Sophie.Hermann@uni-bayreuth.d) and [Matthias.Schmidt@uni-bayreuth.de](mailto:Matthias.Schmidt@uni-bayreuth.de)

Received 31 July 2022; revised 2 November 2022

Accepted for publication 8 November 2022

Published 18 November 2022



CrossMark

## Abstract

We address the consequences of invariance properties of the free energy of spatially inhomogeneous quantum many-body systems. We consider a specific position-dependent transformation of the system that consists of a spatial deformation and a corresponding locally resolved change of momenta. This operator transformation is canonical and hence equivalent to a unitary transformation on the underlying Hilbert space of the system. As a consequence, the free energy is an invariant under the transformation. Noether's theorem for invariant variations then allows to derive an exact sum rule, which we show to be the locally resolved equilibrium one-body force balance. For the special case of homogeneous shifting, the sum rule states that the average global external force vanishes in thermal equilibrium.

Keywords: force balance, quantum statistical mechanics, YBG equation, density functional theory, Noether's theorem

## 1. Introduction

When investigating global equilibrium properties such as the equation of state for a given many-body Hamiltonian, the strategies in classical and quantum statistical mechanical treatments differ markedly from each other. Obtaining the partition sum in the classical case requires, in principle, to carry out the high-dimensional phase space integral over the

\* Authors to whom any correspondence should be addressed.



Original Content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Boltzmann factor of the Hamiltonian [1]. The quantum mechanical analog thereof is the trace over the Boltzmann factor of the Hamiltonian, where the latter is viewed as an infinite-dimensional matrix expressed in a suitable basis of e.g. energy eigenfunctions [2]. In both cases, quantum and classical, the leap from the dynamics of the particle-based many-body description to the thermal average is both powerful and abstract. As a result physically meaningful quantities, such as the pressure, chemical and thermal susceptibilities etc become systematically available, at least in principle, through derivatives of the free energy, which are readily available from the partition sum. On a higher level of detail, locally resolved correlation functions are available as statistical averages and they characterize the microscopic structure of the system and allow to obtain global properties via suitable integration.

On the other hand the concept of forces, while being at the very heart of mechanics, often receives less attention in both statistical and quantum contexts. Nevertheless in the realm of quantum many-body systems several recent publications [3–5] addressed in detail the force balance relationship on the one-body level of correlation functions. Here the forces are resolved in position and also in time in the dynamic case. Tarantino and Ullrich [3] reformulated time-dependent Kohn–Sham density functional theory (DFT) in terms of the second time derivative of the density. In their approach forces feature prominently. They argue that the causal structure of their formulation is more transparent than that of the standard Kohn–Sham formalism of DFT. Tchenkoue *et al* [4] have addressed the force balance in several advanced approximations in DFT. They state that their approach avoids differentiability and causality issues and having to carry out the optimized-effective-potential procedure of orbital-dependent functionals.

Earlier than these advancements, Tokatly had already honed in on the force balance relationship in the framework of his time-dependent deformation functional theory [6]. The theory is based on considering a hydrodynamic Lagrangian view of quantum many-body dynamics [7, 8]. The force balance equation plays a role of a gauge condition that fixes the reference frame [7]. The approach yields formally exact equations of motion and conservation laws [7] and it provided the basis for a geometric formulation of time-dependent DFT [8]. Ullrich and Tokatly were then able to address important nonadiabatic effects in the electron dynamics in time-dependent density-functional theory [9].

Locally resolved force fields play a prominent role in the recent power functional framework for many-body dynamics [5]. Besides the time-dependent density profile, this variational approach includes the locally resolved current and acceleration distributions as its fundamental physical variables. The theory has been formulated for classical [10, 11] and quantum [12, 13] systems; a recent review [5] gives much background. The respective variational equation has the clear physical interpretation of a nonequilibrium force balance relationship and it allows to categorize flow and structural forces [5] and acceleration viscous forces [5], all of which go beyond the adiabatic forces that are captured by the dynamical classical DFT [14–18]. Standard DFT is recovered as the equilibrium limit of the power functional theory [5].

Furthermore, in the classical context, forces were recently put to the fore in methods to obtain statistically averaged quantities, such as the density profile of a spatially inhomogeneous system, via computer simulation of the many-body problem. In his recent review [19], Rotenberg gives a clear account of such force-sampling techniques; see e.g. [20–22] for original work. On the theoretical side, classical DFT [14–16] offers access to forces via building the gradient of the Euler–Lagrange minimization equation [5]. An alternative that applies to pairwise interparticle interactions is the force integral over the two-body density correlation function [23, 24]. The two-body density is explicitly available within state-of-the-art classical density functionals, such as fundamental measure theory, see e.g. [25]. Two-body density correlation functions are also central to the recently developed conditional probability DFT for quantum systems [26, 27].

In prior work we have applied Noether's theorem of invariant variations [28, 29] to the classical statistical mechanics of particle-based many-body systems [30–33]. Rather than starting with the invariance properties of an action functional, the approach is based on considering the symmetry properties of appropriate statistical functionals, such as the partition sum, in order to derive exact identities. These 'sum rules' carry clear physical interpretation as interrelations between forces when starting with spatial displacement, and between torques when starting with spatial rotations. Different types of identities result, depending on whether the elementary free energy is displaced (leading to external force sum rules), the excess free energy density functional (internal force identities) or the power functional (memory identities [30] that connect time direct correlation functions [5]).

We emphasize that Noether's original work [28] is not restricted to the action integral of a physical system. She rather deals with functionals of a general nature, formulating carefully necessary (and for our practical purposes very mild) assumptions of analyticity. Background from an entirely mathematical perspective can be found in [34]. Descriptions of the standard application to the action can be found in many sources, including [35] and on a more popular level [36]. For the classical case, the differences between the present use in thermal physics and the standard deterministic form are discussed in [31]. Briefly, within our present setting, we require to identify a functional  $F[\epsilon]$  of a position-dependent vector field  $\epsilon(\mathbf{r})$ . On the one hand  $\epsilon(\mathbf{r})$  parametrizes the functional dependence on further fields (suppressed in the notation). Noether's theorem applies, when despite this apparent dependence, on the other hand the functional is invariant under changes of  $\epsilon(\mathbf{r})$ . Hence trivially  $\delta F[\epsilon]/\delta \epsilon(\mathbf{r}) = 0$ . Quite remarkably this leads to a nontrivial identity, when taken as a concrete recipe for calculation of the left-hand side.

The invariant variational techniques have aided the development of a force-based approach to classical DFT [33]. Here the fundamental starting equation is the locally resolved equilibrium force balance relationship, which (for pairwise interparticle forces) is a classical result [1] that dates back to Yvon [23], and to Born and Green [24]. The derivation of this fundamental equation is performed by considering an inhomogeneous spatial displacement of the entire system, as described by a vector field  $\epsilon(\mathbf{r})$  in three-dimensional space. Together with a corresponding change of momenta (described in detail below) the change of variables constitutes a canonical transformation on classical phase space, and hence it preserves the phase space volume element [37]. The specific form of the transformation (in particular it being independent of time) also preserves the Hamiltonian. Hence the partition sum itself is unchanged under the transformation and so is the free energy. (The Hamiltonian, via its associated Boltzmann factor, and the phase space integral are the only nontrivial ingredients in the partition sum.) One is hence faced with an invariant variational problem, as addressed succinctly by Emmy Noether in her classical work [28]; see [29] for a historical account.

In the present contribution, we demonstrate that Noether's theorem is applicable to the equilibrium statistical properties of quantum many-body systems. We present a quantal shifting transformation of the position and momentum operators that reduces to the transformation of [33] in the classical case. Quantum mechanically, the transformation is canonical [38], i.e. it preserves the fundamental commutator relation between position and momentum. Such transformations represent unitary transformations on the Hilbert space of the considered system. The partition sum is hence invariant under the transformation, as it is given as the trace of the Hamiltonian's Boltzmann factor, with both the trace and the Hamiltonian being invariants, as is the case classically. The result, to first order in the displacement field, is the locally resolved equilibrium force balance relationship [3–5]. While one could expect on general grounds that

the Noether line of thought would indeed apply to quantum systems, the details of the derivation differ markedly from the classical case, and we spell out the details in the following.

## 2. Thermal invariance theory

### 2.1. Quantum canonical transformation

We consider Hamiltonians of the form

$$H = \sum_i \frac{\mathbf{p}_i^2}{2m} + u(\mathbf{r}^N) + \sum_i V_{\text{ext}}(\mathbf{r}_i), \quad (1)$$

where the sums run over all particles  $i = 1, \dots, N$ , with the total number of particles  $N$ . All particles possess identical mass  $m$  and each particle  $i$  is characterized by its position ( $\mathbf{r}_i$ ) and momentum ( $\mathbf{p}_i$ ) operator. The interparticle interaction potential  $u(\mathbf{r}^N)$  depends on all particle positions and we use the compact notation  $\mathbf{r}^N \equiv \mathbf{r}_1, \dots, \mathbf{r}_N$ . The system is under the influence of an external one-body potential  $V_{\text{ext}}(\mathbf{r})$ , where  $\mathbf{r}$  is a generic position variable.

Position and momentum satisfy the fundamental commutator relations

$$[\mathbf{r}_i, \mathbf{p}_j] = i\hbar\delta_{ij}\mathbb{1}, \quad (2)$$

where  $i$  is the imaginary unit,  $\hbar$  denotes the reduced Planck constant,  $\delta_{ij}$  is the Kronecker symbol and  $\mathbb{1}$  indicates the  $3 \times 3$ -unit matrix. The commutator of two vectors involves transposition according to  $[\mathbf{r}_i, \mathbf{p}_j] = \mathbf{r}_i\mathbf{p}_j - \mathbf{p}_j\mathbf{r}_i^T$ , where the multiplication of two vectors is dyadic and the superscript  $T$  denotes the transpose of a  $3 \times 3$ -matrix. Hence in component notation  $[r_i^\alpha, p_j^\gamma] = r_i^\alpha p_j^\gamma - p_j^\gamma r_i^\alpha$ , where Greek superscripts  $\alpha, \gamma$  denote Cartesian components of position and momentum. Equation (2) then reads as  $[r_i^\alpha, p_j^\gamma] = i\hbar\delta_{ij}\delta_{\alpha\gamma}$ . We work in position representation, such that the momentum operator of particle  $i$  is given by  $\mathbf{p}_i = -i\hbar\nabla_i$ , where  $\nabla_i$  indicates the derivative with respect to  $\mathbf{r}_i$ .

We consider the following transformation of position and momenta

$$\mathbf{r}_i \rightarrow \mathbf{r}_i + \boldsymbol{\epsilon}(\mathbf{r}_i), \quad (3)$$

$$\mathbf{p}_i \rightarrow \{(\mathbb{1} + (\nabla_i\boldsymbol{\epsilon}_i))^{-1} \cdot \mathbf{p}_i + \mathbf{p}_i \cdot (\mathbb{1} + (\nabla_i\boldsymbol{\epsilon}_i)^T)^{-1}\}/2, \quad (4)$$

where  $\boldsymbol{\epsilon}(\mathbf{r})$  is a given real-valued three-dimensional vector field with  $\mathbf{r}$  indicating position,  $\boldsymbol{\epsilon}_i = \boldsymbol{\epsilon}(\mathbf{r}_i)$  is a shorthand notation, and the superscript  $-1$  indicates matrix inversion. In equation (4) the gradient operator  $\nabla_i$  acts only on  $\boldsymbol{\epsilon}_i$ , as is indicated by the surrounding parentheses; hence in position representation each entry of the  $3 \times 3$ -matrix  $(\nabla_i\boldsymbol{\epsilon}_i)$ , which is obtained as a dyadic product of the vectors  $\nabla_i$  and  $\boldsymbol{\epsilon}_i$ , acts only as a multiplication operator on the wave function. In our notation the dot product of a matrix  $\mathbf{A}$  and a vector  $\mathbf{x}$  is understood in the standard way as  $(\mathbf{A} \cdot \mathbf{x})_\alpha = A_{\alpha\gamma}x_\gamma$ , with summation being implied over the repeated index  $\gamma$ . For convenience we also define this product with the reversed order of factors as  $(\mathbf{x} \cdot \mathbf{A})_\alpha = x_\gamma A_{\gamma\alpha}$ . This appears in the second term in the sum in equation (4).

We assume throughout that the vector field  $\boldsymbol{\epsilon}(\mathbf{r})$  is such that a bijection is established between old and new coordinates. Hence the transformations (3) and (4) need to be invertible. (A poignant counterexample is  $\boldsymbol{\epsilon}(\mathbf{r}) = -\mathbf{r}$ , which renders equation (3) to be  $\mathbf{r}_i \rightarrow 0$  and the matrix inversion in the momentum transformation (4) becoming ill-defined.) The momentum transformation is the self-adjoint version of the classical phase space transformation considered in [33], which is in linear order, as given in [33], simply  $\mathbf{p}_i \rightarrow \mathbf{p}_i - (\nabla_i\boldsymbol{\epsilon}_i) \cdot \mathbf{p}_i$ . Here there is no need to pay attention to the ordering of terms, as the classical phase space variables commute with each other. The finite version thereof is  $\mathbf{p}_i \rightarrow (\mathbb{1} + \nabla_i\boldsymbol{\epsilon}_i)^{-1} \cdot \mathbf{p}_i$ , as can be

shown via a generating function that via differentiation yields the transformation equations. Equation (4) is obtained as the arithmetic mean of this expression and its adjoint. Very briefly, the classical generator (see the appendix of [33]) is  $\mathcal{G} = \sum_{i=1}^N \tilde{\mathbf{p}}_i \cdot (\mathbf{r}_i + \boldsymbol{\epsilon}(\mathbf{r}_i))$ , and the transformation equations are obtained via the identities  $\tilde{\mathbf{r}}_i = \partial\mathcal{G}/\partial\tilde{\mathbf{p}}_i$  and  $\mathbf{p}_i = \partial\mathcal{G}/\partial\mathbf{r}_i$ , where the tilde indicates the transformed variables.

We expand the inverse matrix in equation (4) to linear order in the gradient of the displacement field according to:  $(\mathbb{1} + (\nabla_i\boldsymbol{\epsilon}_i))^{-1} = \mathbb{1} - (\nabla_i\boldsymbol{\epsilon}_i)$ , where terms of the order  $(\nabla_i\boldsymbol{\epsilon}_i)^2$  and higher have been omitted. Using this expansion, equation (4) in component notation is:  $p_i^\alpha \rightarrow p_i^\alpha - \sum_\gamma \{(\nabla_i^\alpha\epsilon_i^\gamma)p_i^\gamma + p_i^\gamma(\nabla_i^\alpha\epsilon_i^\gamma)\}/2$ , which when resorting back to vector notation is:

$$\mathbf{p}_i \rightarrow \mathbf{p}_i - \{(\nabla_i\boldsymbol{\epsilon}_i) \cdot \mathbf{p}_i + \mathbf{p}_i \cdot (\nabla_i\boldsymbol{\epsilon}_i)^\top\}/2. \quad (5)$$

We first ascertain that the new coordinates  $\tilde{\mathbf{r}}_i$  and new momenta  $\tilde{\mathbf{p}}_i$ , as defined by the right-hand sides of the transformation (3) and (4), also satisfy canonical commutation relations. We start with the prominent case of position and momentum:

$$[\tilde{\mathbf{r}}_i, \tilde{\mathbf{p}}_j] = [\mathbf{r}_i + \boldsymbol{\epsilon}_i, \mathbf{p}_j - \{(\nabla_j\boldsymbol{\epsilon}_j) \cdot \mathbf{p}_j + \mathbf{p}_j \cdot (\nabla_j\boldsymbol{\epsilon}_j)^\top\}/2] \quad (6)$$

$$= [\mathbf{r}_i, \mathbf{p}_j] + [\boldsymbol{\epsilon}_i, \mathbf{p}_j] - [\mathbf{r}_i, (\nabla_j\boldsymbol{\epsilon}_j) \cdot \mathbf{p}_j]/2 - [\mathbf{r}_i, \mathbf{p}_j \cdot (\nabla_j\boldsymbol{\epsilon}_j)^\top]/2 \quad (7)$$

$$= i\hbar\delta_{ij}\mathbb{1}, \quad (8)$$

where we have truncated in (7) at linear order in the displacement field and its gradient. As the left-hand side of (6) involves no coupling between different particles, it is straightforward to see that for distinct particles,  $i \neq j$ , the result vanishes, as is indeed the case in equation (8). For  $i = j$  we use the explicit form of the momentum operator  $\mathbf{p}_i = -i\hbar\nabla_i$  to find that the second term in (7) is  $[\boldsymbol{\epsilon}_i, \mathbf{p}_i] = i\hbar(\nabla_i\boldsymbol{\epsilon}_i)$ . This contribution is precisely cancelled by the sum of the third and the fourth term in (7), which can be shown to have the form  $-[\mathbf{r}_i, (\nabla_i\boldsymbol{\epsilon}_i) \cdot \mathbf{p}_i] = -(\nabla_i\boldsymbol{\epsilon}_i) \cdot (i\hbar\mathbb{1}) = -i\hbar(\nabla_i\boldsymbol{\epsilon}_i)$ . Hence the first term in (7) alone gives the result (8) upon using the fundamental commutator (2).

For completeness, the new variables also satisfy  $[\tilde{\mathbf{r}}_i, \tilde{\mathbf{r}}_j] = 0$  and  $[\tilde{\mathbf{p}}_i, \tilde{\mathbf{p}}_j] = 0$ . The former relationship is trivial, as in position representation only coordinates are involved according to the transformation (3). The momentum identity can be worked out straightforwardly, as we show in appendix A.1. Furthermore the new degrees of freedom are self-adjoint operators. For the positions this is trivial, as we have  $\tilde{\mathbf{r}}_i^\dagger = \mathbf{r}_i^\dagger + \boldsymbol{\epsilon}(\mathbf{r}_i)^\dagger = \mathbf{r}_i + \boldsymbol{\epsilon}(\mathbf{r}_i) \equiv \tilde{\mathbf{r}}_i$ , because  $\boldsymbol{\epsilon}(\mathbf{r})$  is real-valued. For the momenta:  $\tilde{\mathbf{p}}_i^\dagger = \mathbf{p}_i^\dagger - \{(\nabla_i\boldsymbol{\epsilon}_i) \cdot \mathbf{p}_i + \mathbf{p}_i \cdot (\nabla_i\boldsymbol{\epsilon}_i)^\top\}^\dagger/2 = \mathbf{p}_i - \{\mathbf{p}_i \cdot (\nabla_i\boldsymbol{\epsilon}_i)^\top + (\nabla_i\boldsymbol{\epsilon}_i) \cdot \mathbf{p}_i\}/2 \equiv \tilde{\mathbf{p}}_i$ . For completeness we demonstrate that the transformation is quantum canonical beyond linear order in appendix A.2.

## 2.2. Functional derivatives by local shift

Having ascertained that the new variables form a sound basis for the description of the quantum mechanics, we wish to illustrate the effect of the transformation on the system. The following considerations will be an essential ingredient in the thermal physics addressed further below. We wish to investigate the effect on the Hamiltonian  $H[\boldsymbol{\epsilon}]$ , which is obtained by applying the operator replacements (3) and (4) in the form (1) of the original Hamiltonian. We consider the functional derivative of the transformed Hamiltonian with respect to the displacement field:

$$\frac{\delta H[\boldsymbol{\epsilon}]}{\delta \boldsymbol{\epsilon}(\mathbf{r})} = \frac{\delta}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \left( \sum_i \frac{\tilde{\mathbf{p}}_i^2}{2m} + u(\tilde{\mathbf{r}}^N) + \sum_i V_{\text{ext}}(\tilde{\mathbf{r}}_i) \right). \quad (9)$$

To make progress, we first address the fundamental derivatives of the new position and new momentum with respect to the displacement field. These are easily obtained as follows:

$$\left. \frac{\delta \tilde{\mathbf{r}}_i}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \right|_{\boldsymbol{\epsilon}=0} = \delta(\mathbf{r} - \mathbf{r}_i) \mathbb{1}, \quad (10)$$

$$\left. \frac{\delta \tilde{\mathbf{p}}_i}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \right|_{\boldsymbol{\epsilon}=0} = \nabla \{ \delta(\mathbf{r} - \mathbf{r}_i) \mathbf{p}_i + \mathbf{p}_i \delta(\mathbf{r} - \mathbf{r}_i) \} / 2, \quad (11)$$

where  $\delta(\cdot)$  denotes the (three-dimensional) Dirac distribution and the derivatives are taken at vanishing displacement field,  $\boldsymbol{\epsilon}(\mathbf{r}) = 0$ , as is indicated in the notation on both left-hand sides. The right-hand side of equation (10) constitutes the density operator of particle  $i$  times the unit matrix. The right-hand side of equation (11) is the spatial gradient of the momentum density operator of particle  $i$ . That both correctly localized operators appear naturally as functional derivatives is an initial indication that the considered transformation indeed can be used as a successful probe for the spatially resolved behaviour of the system.

For completeness, in index notation equation (11) reads as

$$\delta \tilde{p}_i^\alpha / \delta \epsilon^\gamma = \nabla^\alpha (\delta_i p_i^\gamma + p_i^\gamma \delta_i) / 2, \quad (12)$$

where we have introduced the shorthand notations  $\delta_i = \delta(\mathbf{r} - \mathbf{r}_i)$  and  $\epsilon^\gamma = \epsilon^\gamma(\mathbf{r})$  and the derivative is again evaluated at vanishing displacement field (such that higher than linear powers in the displacement gradient, as they occur in the *finite* momentum transformation (4), vanish).

In order to obtain the functional derivative (9) of the Hamiltonian we proceed by first differentiating the kinetic energy. We defer the detailed calculations to appendix B, which contains both the simpler one-dimensional case, where matrix-vector complexities are absent (appendix B.1), as well as the present three-dimensional case (appendix B.2). The latter calculation, carried out in index notation in appendix B.2, gives the following result:

$$\left. \frac{\delta}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \sum_i \frac{\tilde{\mathbf{p}}_i^2}{2m} \right|_{\boldsymbol{\epsilon}=0} = \nabla \cdot \sum_i \frac{\mathbf{p}_i \delta_i \mathbf{p}_i + \mathbf{p}_i \delta_i \mathbf{p}_i^\top}{2m} - \frac{\hbar^2}{4m} \nabla \nabla^2 \sum_i \delta(\mathbf{r} - \mathbf{r}_i). \quad (13)$$

We recall that the transpose (superscript  $\top$ ) acts on the entire  $3 \times 3$ -matrix  $\mathbf{p}_i \delta_i \mathbf{p}_i$  and that the multiplication of vector and matrix, as is relevant for the divergence, contracts the vector index with the first matrix index; we recall our description thereof after equation (4). Equation (13) is also given in index notation in appendix B.2.

The first term on the right-hand side of equation (13) is directly analogous to the classical case [33], upon viewing the momentum operators as phase space variables. The second term on the right-hand side of equation (13) is genuinely quantum mechanical, as it is quadratic in  $\hbar$  and hence vanishes in the classical limit  $\hbar \rightarrow 0$ . This contribution can be rewritten upon expressing the gradient of the Laplace operator as  $\nabla \nabla^2 = \nabla^2 \nabla = \nabla \cdot \nabla \nabla$ . Then one can express the second term in equation (13) as  $-\nabla \cdot \nabla \nabla \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \hbar^2 / (4m)$ . Here the Hessian of the density operator,  $\nabla \nabla \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$ , together with the factor  $\hbar^2 / (4m)$  forms the quantal kinetic stress contribution  $\nabla \nabla \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \hbar^2 / (4m)$ .

Together with the first term in equation (13), which already is of divergence form, we can define the position-resolved kinetic stress operator (see e.g. [5]) as

$$\hat{\boldsymbol{\tau}}(\mathbf{r}) = - \sum_i \frac{\mathbf{p}_i \delta_i \mathbf{p}_i + \mathbf{p}_i \delta_i \mathbf{p}_i^\top}{2m} + \frac{\hbar^2}{4m} \nabla \nabla \sum_i \delta_i. \quad (14)$$

We have hence adopted the convention to include the wave-like contribution  $\hbar^2 \nabla \nabla \hat{\rho}(\mathbf{r}) / (4m)$  into the kinematic stress, where  $\hat{\rho}(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$  is the standard form of the one-body density operator. The classical kinetic stress is recovered by letting  $\hbar \rightarrow 0$ , such that  $\hat{\boldsymbol{\tau}}(\mathbf{r})$  reduces to  $-\sum_i \delta(\mathbf{r} - \mathbf{r}_i) \mathbf{p}_i \mathbf{p}_i / m$ , where here  $\mathbf{p}_i$  denotes the classical phase



space variable, which trivially commutes with the spatial delta distribution, and the transpose becomes irrelevant as for the phase space variable  $\mathbf{p}_i \mathbf{p}_i = \mathbf{p}_i \mathbf{p}_i^\top$ .

We have so far shown that the considered quantum canonical transformation the functional derivative of kinetic energy with respect to the displacement field creates the following fundamental result:

$$\left. \frac{\delta H_{\text{kin}}[\boldsymbol{\epsilon}]}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \right|_{\boldsymbol{\epsilon}=0} = -\nabla \cdot \hat{\boldsymbol{\tau}}(\mathbf{r}). \quad (15)$$

Here we have split the Hamiltonian (1) according to  $H = H_{\text{kin}} + H_{\text{pot}}$ , where the potential energy contains the interparticle and external contributions,  $H_{\text{pot}} = u(\mathbf{r}^N) + \sum_i V_{\text{ext}}(\mathbf{r}_i)$ . As already laid out above, the functional dependence on  $\boldsymbol{\epsilon}(\mathbf{r})$  that is indicated on the left-hand side of (15) arises from expressing the original positions and momenta in the Hamiltonian (1) via the transformation (3) and (4). One could view the result (15) as being unexpectedly simple, despite the technical complexity of the kinematic stress operator  $\hat{\boldsymbol{\tau}}(\mathbf{r})$ . Recall that the kinematic stress occurs in the Heisenberg equation of motion for the one-body current density [5, 7, 8] and that it hence constitutes a meaningful physical object in its own right. That it is created here from the functional derivative of kinetic energy with respect to the shift field is a strong indicator that the thermal Noether invariance against the local shifting transformation given by equations (3) and (4) carries actual physical significance.

Treating the effects of the local displacement transformation on the potential energy is comparatively easier than the above kinetic energy consideration, as here only position coordinates are involved and hence the commutator structure is trivial. The calculation is very closely analogous to the classical case [31]. We obtain

$$\left. \frac{\delta H_{\text{pot}}[\boldsymbol{\epsilon}]}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \right|_{\boldsymbol{\epsilon}=0} = \left. \frac{\delta u(\mathbf{r}^N)}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \right|_{\boldsymbol{\epsilon}=0} + \sum_i \left. \frac{\delta V_{\text{ext}}(\mathbf{r}_i)}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \right|_{\boldsymbol{\epsilon}=0} \quad (16)$$

$$= \sum_i (\nabla_i u(\mathbf{r}^N)) \delta_i + \sum_i (\nabla_i V_{\text{ext}}(\mathbf{r}_i)) \delta_i \quad (17)$$

$$= -\hat{\mathbf{F}}_{\text{int}}(\mathbf{r}) + \hat{\rho}(\mathbf{r}) \nabla V_{\text{ext}}(\mathbf{r}), \quad (18)$$

where we have defined the one-body interparticle force density operator  $\hat{\mathbf{F}}_{\text{int}}(\mathbf{r}) = -\sum_i (\nabla_i u(\mathbf{r}^N)) \delta_i$ . The rewriting that involves the external force field  $-\nabla V_{\text{ext}}(\mathbf{r})$  in (18) is possible as the derivatives  $\nabla_i$  and  $\nabla$ , as well as the positions  $\mathbf{r}$  and  $\mathbf{r}_i$ , can be identified with each other due to the presence of the delta function. The negative external force field  $\nabla V_{\text{ext}}(\mathbf{r})$  can then be taken as a common factor outside of the second sum in equation (17) and the density operator remains.

Summing up the kinetic energy derivative (15) and the potential energy identity (18) we obtain

$$-\left. \frac{\delta H[\boldsymbol{\epsilon}]}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \right|_{\boldsymbol{\epsilon}=0} = \nabla \cdot \hat{\boldsymbol{\tau}}(\mathbf{r}) + \hat{\mathbf{F}}_{\text{int}}(\mathbf{r}) - \hat{\rho}(\mathbf{r}) \nabla V_{\text{ext}}(\mathbf{r}), \quad (19)$$

which makes explicit that the Hamiltonian generates, via its negative functional derivative with respect to the displacement field, the sum of all one-body force density distributions that act in the system.

That the transformation (3) and (4) has an effect on the Hamiltonian could have been expected from the outset, as the transformation has a nontrivial spatial structure via its dependence on the vector field  $\boldsymbol{\epsilon}(\mathbf{r})$ . Hence Noether's theorem seemingly does not apply, due to the absence of a direct corresponding invariance. In contrast to this standard application, here we proceed differently and search for an invariance that applies in thermal equilibrium. This requires an

average to be an invariant rather than the corresponding operator itself being an invariant, as we lay out in the following.

### 2.3. Force balance from thermal Noether invariance

We hence turn to a statistical mechanical description which we base on the free energy in the canonical ensemble, expressed as

$$F = -k_B T \ln Z, \quad (20)$$

$$Z = \text{Tr} e^{-\beta H} \quad (21)$$

$$= \sum_n \langle n | e^{-\beta H} | n \rangle, \quad (22)$$

where  $k_B$  denotes the Boltzmann constant and  $T$  is absolute temperature. The trace in Hilbert space is denoted by  $\text{Tr}$  and it is made explicit in (22) with  $|n\rangle$  denoting the complete set of orthonormal eigenstates of  $H$  labelled by index  $n$ . (Possibly degenerate energy eigenstates occur multiple times in the sum.) In more explicit notation, using position representation,  $|n\rangle = \phi_n(\mathbf{r}^N)$  such that  $\langle n | \cdot | n \rangle = \int d\mathbf{r}^N \phi_n^*(\mathbf{r}^N) \cdot \phi_n(\mathbf{r}^N)$ , where the integral is over all position coordinates,  $\int d\mathbf{r}^N = \int d\mathbf{r}_1 \int d\mathbf{r}_2 \dots \int d\mathbf{r}_N$  and the asterisk denotes complex conjugation. Here and throughout, we assume that the partition sum (22) and hence the free energy (20) exists, see Giesbertz and Ruggenthaler's [39] account of the divergences that occur in even simple unbounded systems. In our case, we assume (as we do classically [31]) that the system is bounded via the influence of appropriate container walls, as modelled by a corresponding form of the external potential  $V_{\text{ext}}(\mathbf{r})$ . We hence adopt a pragmatic stance to the existence of the free energy [40].

We expand the free energy (20) in the transformation parameter according to:

$$\left. \frac{\delta F[\boldsymbol{\epsilon}]}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \right|_{\boldsymbol{\epsilon}=0} = - \frac{k_B T}{Z} \left. \frac{\delta Z[\boldsymbol{\epsilon}]}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \right|_{\boldsymbol{\epsilon}=0} \quad (23)$$

$$= - \frac{k_B T}{Z} \text{Tr} \left. \frac{\delta e^{-\beta H[\boldsymbol{\epsilon}]}}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \right|_{\boldsymbol{\epsilon}=0} \quad (24)$$

$$= \sum_n \frac{e^{-\beta E_n}}{Z} \langle n | \left. \frac{\delta H[\boldsymbol{\epsilon}]}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \right|_{\boldsymbol{\epsilon}=0} | n \rangle, \quad (25)$$

where in equation (23) we have used the definition (20) of the free energy via the partition sum  $Z[\boldsymbol{\epsilon}]$  of the transformed system. In equation (24) we have used the form (21) of the partition sum and have exchanged the order of the functional derivative and building the trace. Equation (25) constitutes a thermal equilibrium average, where  $E_n$  denotes the energy eigenvalue corresponding to the energy eigenstate  $|n\rangle$ . We have hence obtained

$$\left. \frac{\delta F[\boldsymbol{\epsilon}]}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \right|_{\boldsymbol{\epsilon}=0} = \left\langle \left. \frac{\delta H[\boldsymbol{\epsilon}]}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \right|_{\boldsymbol{\epsilon}=0} \right\rangle_{\text{eq}}, \quad (26)$$

where on the right-hand side we have used the notation  $\langle \cdot \rangle_{\text{eq}}$  to indicate the average over the canonical ensemble as it occurs in equation (25); explicitly this is  $\langle \cdot \rangle_{\text{eq}} = \sum_n Z^{-1} e^{-\beta E_n} \langle n | \cdot | n \rangle$ . The identity (26) is remarkable as it indicates that the local transformation (3) and (4) to lowest order in the displacement field generates a well-defined and physically meaningful thermal average, that of the functional derivative of the Hamiltonian. This mathematical structure mirrors closely that of standard partial derivatives of the free energy with respect to thermodynamic variables, such as e.g. obtaining the entropy via  $S = -\partial F / \partial T$ .

Before exploiting the specific form of the right-hand side of equation (26) further, we first proceed with the general invariance argument. We expand the free energy of the transformed system to linear order in the displacement field according to:

$$F[\epsilon] = F + \int d\mathbf{r} \left. \frac{\delta F[\epsilon]}{\delta \epsilon(\mathbf{r})} \right|_{\epsilon=0} \cdot \epsilon(\mathbf{r}) \quad (27)$$

$$= F + \int d\mathbf{r} \left\langle \left. \frac{\delta H[\epsilon]}{\delta \epsilon(\mathbf{r})} \right|_{\epsilon=0} \right\rangle_{\text{eq}} \cdot \epsilon(\mathbf{r}), \quad (28)$$

where equation (27) is the functional Taylor expansion to linear order and the form (28) follows from using equation (26).

On the other hand, the free energy is an invariant under the quantum canonical transformation, and hence:

$$F[\epsilon] = F, \quad (29)$$

where  $F$  is the free energy (20) of the original representation of the system. Equation (29) holds due to the fact that canonical transformations are analogous to unitary transformations on the underlying Hilbert space of the considered system; see e.g. the account given by Anderson [38]. We will return to this point below.

From comparison of the Taylor expansion (28) with the free energy invariance (29) we can conclude that the linear term in the expansions vanishes identically and it has to do so irrespective of the form of  $\epsilon(\mathbf{r})$ . This can only hold provided that the prefactor vanishes:

$$\left\langle \left. \frac{\delta H[\epsilon]}{\delta \epsilon(\mathbf{r})} \right|_{\epsilon=0} \right\rangle_{\text{eq}} = 0. \quad (30)$$

Equation (30) is a bare consequence of the invariance of the free energy under the displacement operation, and it is obtained immediately from  $\delta F[\epsilon]/\delta \epsilon(\mathbf{r})|_{\epsilon=0} = 0$ , as mentioned in the introduction, upon skipping the Taylor expansion argument expressed in equations (27) and (28). The identity (30) holds irrespective of the precise form of the interparticle interaction potential  $u(\mathbf{r}^N)$  and of the external potential  $V_{\text{ext}}(\mathbf{r})$  as they appear in the Hamiltonian (1).

In order to reveal the physical significance of the Noether sum rule (30) we proceed by inserting the explicit force form of the functional derivative of the Hamiltonian given by equation (19), which yields

$$\nabla \cdot \boldsymbol{\tau}(\mathbf{r}) + \mathbf{F}_{\text{int}}(\mathbf{r}) - \rho(\mathbf{r}) \nabla V_{\text{ext}}(\mathbf{r}) = 0. \quad (31)$$

Here we have introduced the equilibrium averages for the locally resolved kinetic stress:  $\boldsymbol{\tau}(\mathbf{r}) = \langle \hat{\boldsymbol{\tau}}(\mathbf{r}) \rangle_{\text{eq}}$ , for the interparticle force density:  $\mathbf{F}_{\text{int}}(\mathbf{r}) = \langle \hat{\mathbf{F}}_{\text{int}}(\mathbf{r}) \rangle_{\text{eq}}$ , and for the one-body density distribution:  $\rho(\mathbf{r}) = \langle \hat{\rho}(\mathbf{r}) \rangle_{\text{eq}}$ . The force density balance relationship (31) is a known exact equilibrium sum rule, see e.g. [3–5]. Our derivation demonstrates its origin in the invariance of the free energy under the quantum canonical transformation (3) and (4).

As a special case we consider a uniform displacement such that  $\epsilon(\mathbf{r}) = \epsilon_0 = \text{const.}$  For classical systems the invariance of the free energy under such homogeneous displacement leads to the sum rule of vanishing global external force in thermal equilibrium [30, 31]. This result readily translates to the quantum case as follows.

First we obtain the global identity by starting with the locally resolved force balance relationship (31) and integrating over all positions. Two of the resulting integrals vanish,  $\int d\mathbf{r} \nabla \cdot \boldsymbol{\tau}(\mathbf{r}) = 0$  and  $\int d\mathbf{r} \mathbf{F}_{\text{int}}(\mathbf{r}) = 0$ , where the former identity can be shown via integration by parts and the latter identity is a consequence of the translational invariance of the interparticle interaction potential:  $\int d\mathbf{r} \mathbf{F}_{\text{int}}(\mathbf{r}) = \int d\mathbf{r} \langle \hat{\mathbf{F}}_{\text{int}}(\mathbf{r}) \rangle_{\text{eq}} = \langle \int d\mathbf{r} \hat{\mathbf{F}}_{\text{int}}(\mathbf{r}) \rangle_{\text{eq}} = \langle \hat{\mathbf{F}}_{\text{int}}^0 \rangle_{\text{eq}} = 0$ . This holds due to the global force operator vanishing identically:  $\hat{\mathbf{F}}_{\text{int}}^0 = -\sum_i (\nabla_i u(\mathbf{r}^N)) \equiv 0$ , which can be

seen straightforwardly by displacing all positions arguments in  $u(\mathbf{r}^N)$  and observing that this leaves its value invariant. Explicitly the invariance is  $u(\mathbf{r}_1, \dots, \mathbf{r}_N) = u(\mathbf{r}_1 + \epsilon_0, \dots, \mathbf{r}_N + \epsilon_0)$ , as the global shift leaves all distance vectors  $\mathbf{r}_i - \mathbf{r}_j$  unchanged. The Taylor expansion of the right-hand side is to first order  $u(\mathbf{r}_1, \dots, \mathbf{r}_N) + \epsilon_0 \cdot \partial u(\mathbf{r}_1 + \epsilon_0, \dots, \mathbf{r}_N + \epsilon_0) / \partial \epsilon_0 |_{\epsilon_0=0}$ . The linear term can be rewritten as  $\epsilon_0 \cdot \sum_i \nabla_i u(\mathbf{r}^N)$ . As this vanishes for any  $\epsilon_0$ , the prefactor vanishes identically which provides the anticipated vanishing of the global interparticle force.

This reasoning is analogous to Newton's third law, actio equals reactio, which holds due to the interparticle forces being conservative. The standard derivation does not require (nor identify) the translational invariance. Typically one addresses the special but important case of pairwise interparticle interactions, with given pair potential  $\phi(r)$  as a function of interparticle distance  $r$ . Then the global interparticle potential energy is  $u(\mathbf{r}^N) = \sum'_{k,l} \phi(|\mathbf{r}_k - \mathbf{r}_l|) / 2$ , where the primed sum indicates that the case  $k=l$  has been omitted and the factor 1/2 corrects for double counting. The global interparticle force is then the (negative) sum of all gradients,  $\hat{\mathbf{F}}_{\text{int}}^0 = -\sum_i \sum'_{k,l} \nabla_i \phi(|\mathbf{r}_k - \mathbf{r}_l|) / 2$ . Via re-organizing the nested sums one obtains  $\hat{\mathbf{F}}_{\text{int}}^0 = \sum'_{i,k} [\nabla_i \phi(|\mathbf{r}_i - \mathbf{r}_k|) - \nabla_i \phi(|\mathbf{r}_i - \mathbf{r}_k|)] / 2 = 0$ , identical to the above result based on invariance.

The only term that remains of equation (31) after carrying out the position integral is the external force contribution, which reads as:

$$-\int d\mathbf{r} \rho(\mathbf{r}) \nabla V_{\text{ext}}(\mathbf{r}) = 0. \quad (32)$$

Equation (32) expresses the vanishing of the average global external force in thermal equilibrium.

Briefly, in our second route to equation (32) we start from the free energy (20) and directly perform the transformation for the special case of a homogeneous displacement  $\epsilon_0$ . In this case the momenta are unchanged, as the gradient of the (constant) displacement field vanishes identically. Hence kinetic energy is trivially invariant. As laid out above, the coordinate change does not affect the interparticle potential energy, as the difference vectors are unaffected. Hence the only change in the Hamiltonian occurs in the external contribution and we obtain, following the same argumentation as in the case of position-dependent shifting, the result

$$-\left\langle \sum_i \nabla_i V_{\text{ext}}(\mathbf{r}_i) \right\rangle_{\text{eq}} = 0, \quad (33)$$

which is analogous to the previous form (32) upon prepending  $1 = \int d\mathbf{r} \delta(\mathbf{r} - \mathbf{r}_i)$  to equation (33), moving the delta function into the thermal average, and identifying the density profile  $\rho(\mathbf{r}) = \langle \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \rangle_{\text{eq}}$ .

For completeness and as a final step, we make explicit that the quantum canonical transformation corresponds indeed to a unitary transformation on Hilbert space, as is relevant for the invariance (29) of the free energy under the transformation. For the present transformation, to linear order in  $\epsilon(\mathbf{r})$ , the transformed Hamiltonian is obtained via functional Taylor expansion in the following form:

$$H[\epsilon] = H + \int d\mathbf{r} \left. \frac{\delta H[\epsilon]}{\delta \epsilon(\mathbf{r})} \right|_{\epsilon=0} \cdot \epsilon(\mathbf{r}), \quad (34)$$

and we recall the explicit one-body force density form (19) of the functional derivative of the Hamiltonian. We treat the second term in equation (34) as a perturbation to the original Hamiltonian  $H$ . (We recall that the thermal average over the functional derivative  $\delta H[\epsilon] / \delta \epsilon(\mathbf{r})$

directly leads to the static force density balance relationship (31).) Then the transformed (perturbed) energy eigenstates  $|\tilde{n}\rangle$  are given by

$$|\tilde{n}\rangle = \sum_k U_{nk}|k\rangle, \quad (35)$$

$$U_{nk} = \delta_{nk} + \frac{1 - \delta_{nk}}{E_k - E_n} \int d\mathbf{r} \langle n | \left. \frac{\delta H[\boldsymbol{\epsilon}]}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \right|_{\boldsymbol{\epsilon}=0} |k\rangle \cdot \boldsymbol{\epsilon}(\mathbf{r}). \quad (36)$$

The form (36) of the matrix that performs the change of basis follows from applying time-independent first order perturbation theory, as is appropriate to capture the effects to linear order in  $\boldsymbol{\epsilon}(\mathbf{r})$  that we consider. (We imply that the prefactor of the integral in equation (36) vanishes for  $k = n$ .) The matrix elements of the Hermitian conjugate to equation (36) can be obtained via exchanging indices  $n$  and  $k$  as

$$U_{nk}^\dagger = \delta_{kn} + \frac{1 - \delta_{kn}}{E_n - E_k} \int d\mathbf{r} \langle k | \left. \frac{\delta H[\boldsymbol{\epsilon}]}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \right|_{\boldsymbol{\epsilon}=0}^\dagger |n\rangle \cdot \boldsymbol{\epsilon}(\mathbf{r}) \quad (37)$$

$$= \delta_{nk} - \frac{1 - \delta_{nk}}{E_k - E_n} \int d\mathbf{r} \langle n | \left. \frac{\delta H[\boldsymbol{\epsilon}]}{\delta \boldsymbol{\epsilon}(\mathbf{r})} \right|_{\boldsymbol{\epsilon}=0} |k\rangle \cdot \boldsymbol{\epsilon}(\mathbf{r}), \quad (38)$$

where to obtain the matrix elements (38) we have exploited that the functional derivative of the Hamiltonian is self-adjoint. We observe that the sole difference between equations (36) and (38) is the minus sign. Hence we can see explicitly that unitarity holds,  $\sum_k U_{nk}^\dagger U_{km} = \delta_{nm}$  to linear order in  $\boldsymbol{\epsilon}(\mathbf{r})$ , as was expected on general grounds [38].

### 3. Outlook and conclusions

In conclusion we have investigated the consequences of a specific local displacement operation for the free energy of a quantum mechanical many-body system. The transformation consists of position-dependent shifting, as parameterized by a real-valued displacement (or ‘shift’) field, and a corresponding transformation of the quantum mechanical momentum operator of each particle. The entirety of the transformation can be viewed as the self-adjoint version of the corresponding local shifting transformation of the classical phase space variables [33]. We have explicitly shown that the new position and momentum operators are self-adjoint and that they satisfy the fundamental commutator relations and hence form a valid and complete set of degrees of freedom of the considered system. The transformation can be viewed as a basis change of the underlying Hilbert space of the quantal system and we have spelled out explicitly the corresponding unitary transformation between the original and the new basis.

The resulting invariance of the free energy under changes in the displacement field then leads, following Noether's theorem for invariant variations, to an exact local identity (‘sum rule’) which we have shown to be the thermal equilibrium force balance. The present derivation of this known and fundamental result from Noether's theorem sheds new light on the very nature of the identity. Existing derivations are based e.g. on the second time derivative of the one-body density profile [3] or, equivalently, on the first time derivative of the one-body current distribution [5] and then taking the equilibrium limit.

Our results hold for the ground state of the quantum system, as it is obtained in the limit  $T \rightarrow 0$  of the free energy of the thermal system. We have used the canonical ensemble throughout as it captures the essence of the required thermal physics. We expect the reasoning to carry over straightforwardly to the grand ensemble with fluctuating particle number, as the classical canonical [31] and grand canonical cases lead to analogous results upon identifying the respective statistical averages.

Future work could be addressed at investigating how functional differentiation can be used to obtain quantum sum rules for higher-body correlation functions, as previously shown for classical systems [30]. It would be interesting to address the effects beyond linear order in the displacement field; classically the variance of the global external force was shown to be constrained by the external potential energy curvature [32]. Last but not least it would be worthwhile to find possible relationships of our displacement field and the strain field that is central to elasticity theory, see e.g. [41, 42] for recent work again in classical systems.

Identifying connections with Tokatly's work [6–9] would be highly interesting. His approach is more general than what we cover here, as it allows for the treatment of the dynamical and nonlinear cases. Clearly, attempting to generalize our approach to the dynamics of statistical quantum systems is an exciting and demanding research task. (We re-iterate that we have here only considered systems in static thermal equilibrium.)

The Noether argument itself is not restricted to linear transformations. The second order was shown, for the case of a global invariance, to relate the variance of fluctuations with the mean potential energy curvature [32]. Carrying through this concept for the quantum case is a further very worthwhile research task.

### Data availability statement

No new data were created or analysed in this study.

### Acknowledgment

This paper is dedicated to Sir Michael Berry on the occasion of his 80th birthday. M S is grateful for the manifold inspirations and reliable judgements, in scientific and other matters, that Michael provided over many years. S H and M S acknowledge useful discussions with Bob Evans and Daniel de las Heras on the topic of the present paper. We also thank the Referees for constructive comments and one of them for pointing out the striking counterexample  $\epsilon(\mathbf{r}) = -\mathbf{r}$ . This work is supported by the German Research Foundation (DFG) via Project No. 436306241.

## Appendix A. Momentum and position commutators

### A.1. Momentum-momentum commutator

To derive the commutator of the new momenta,  $[\tilde{\mathbf{p}}_i, \tilde{\mathbf{p}}_j]$ , we insert the definition of the transformation (5) and consider terms up to linear order in the displacement gradient. In index notation this reads as follows:

$$2[\tilde{p}_i^\alpha, \tilde{p}_j^\gamma] = 2[p_i^\alpha, p_j^\gamma] - [p_i^\alpha, (\nabla_j^\gamma \epsilon_j^\delta) p_j^\delta] - [p_i^\alpha, p_j^\delta (\nabla_j^\gamma \epsilon_j^\delta)] - [(\nabla_i^\alpha \epsilon_i^\delta) p_i^\delta, p_j^\gamma] - [p_i^\delta (\nabla_i^\alpha \epsilon_i^\delta), p_j^\gamma]. \quad (\text{A1})$$

The correlator of the original momenta, as it appears in the first term on the right-hand side, vanishes trivially,  $[p_i^\alpha, p_j^\gamma] = 0$ . This identity also allows to take the operators  $p_i^\delta$  and  $p_j^\delta$  out of the commutator in the remaining terms on the right-hand side of equation (A1). We obtain

$$\begin{aligned}
 2[\tilde{p}_i^\alpha, \tilde{p}_j^\gamma] &= -[p_i^\alpha, (\nabla_j^\gamma \epsilon_j^\delta)] p_j^\delta - p_j^\delta [p_i^\alpha, (\nabla_j^\gamma \epsilon_j^\delta)] + [p_j^\gamma, (\nabla_i^\alpha \epsilon_i^\delta)] p_i^\delta \\
 &\quad + p_i^\delta [p_j^\gamma, (\nabla_i^\alpha \epsilon_i^\delta)],
 \end{aligned} \tag{A2}$$

where we have exploited the anti-symmetry of the commutator,  $[A, B] = -[B, A]$ , for rewriting the third and the fourth term on the right-hand side of equation (A2).

Writing out explicitly the commutator in the first contribution in equation (A2) yields  $[p_i^\alpha, (\nabla_j^\gamma \epsilon_j^\delta)] = p_i^\alpha (\nabla_j^\gamma \epsilon_j^\delta) - (\nabla_j^\gamma \epsilon_j^\delta) p_i^\alpha$ . Hence the momentum operator only acts on the gradient of the displacement field,  $(p_i^\alpha \nabla_j^\gamma \epsilon_j^\delta) = -i\hbar (\nabla_i^\alpha \nabla_j^\gamma \epsilon_j^\delta)$ , where as before the parentheses indicate that the derivative(s) only act on the displacement field and we have expressed the momentum operator in position representation. Analog manipulation of all remaining commutators in equation (A2) then yields

$$\begin{aligned}
 \frac{2i}{\hbar} [\tilde{p}_i^\alpha, \tilde{p}_j^\gamma] &= -(\nabla_i^\alpha \nabla_j^\gamma \epsilon_j^\delta) p_j^\delta - p_j^\delta (\nabla_i^\alpha \nabla_j^\gamma \epsilon_j^\delta) + (\nabla_j^\gamma \nabla_i^\alpha \epsilon_i^\delta) p_i^\delta \\
 &\quad + p_i^\delta (\nabla_j^\gamma \nabla_i^\alpha \epsilon_i^\delta).
 \end{aligned} \tag{A3}$$

For  $i \neq j$  it is now straightforward to see that each term on the right-hand side of equation (A3) vanishes individually: As the displacement field  $\epsilon_i$  only depends on positions  $\mathbf{r}_i$ , derivatives with respect to  $\mathbf{r}_j$  vanish for  $i \neq j$ . For  $i = j$  the first and the third term, as well as the second and the fourth term, on the right-hand side of equation (A3) cancel each other pairwise, as the derivatives  $\nabla_i^\alpha$  and  $\nabla_j^\gamma$  commute. Collecting the cases  $i = j$  and  $i \neq j$  the commutator of the new momentum operators hence vanishes,

$$[\tilde{p}_i^\alpha, \tilde{p}_j^\gamma] = 0, \tag{A4}$$

which ascertains that the position-dependent momentum transformation does not generate any spurious terms.

#### A.2. Momentum-position commutator for finite transformations

The transformations (3) and (4) are canonical not only in linear order of  $\nabla \epsilon(\mathbf{r})$  and  $\epsilon(\mathbf{r})$ , but also for finite values thereof. To demonstrate this property we show that the canonical commutation relations are satisfied given the finite transformations (3) and (4). The position-position commutator  $[\tilde{\mathbf{r}}_i, \tilde{\mathbf{r}}_j]$  is unchanged compared to the derivation in linear order. This is due to the transformed position operator  $\tilde{\mathbf{r}}_i$  (3) containing no higher than linear terms in  $\epsilon_i$ .

In contrast, the transformed momenta  $\tilde{\mathbf{p}}_i$  do contain higher contributions. We express the transformed momentum  $\tilde{\mathbf{p}}_i$  given by equation (4) as an infinite Taylor series in matrix powers of  $(\nabla_i \epsilon_i)$  as

$$\tilde{\mathbf{p}}_i = \mathbf{p}_i + \frac{1}{2} \sum_{n=1}^{\infty} \left( (-\nabla_i \epsilon_i)^n \cdot \mathbf{p}_i + \mathbf{p}_i \cdot (-\nabla_i \epsilon_i)^{\text{T}n} \right). \tag{A5}$$

Note that here the order of transposing and raising the power can be interchanged, i.e.  $(\nabla_i \epsilon_i)^{\text{T}n} = (\nabla_i \epsilon_i)^n \text{T}$ .

We consider the commutator of position and momentum  $[\tilde{\mathbf{r}}_i, \tilde{\mathbf{p}}_j]$ . Only the case  $i = j$  needs to be considered, since otherwise the commutator vanishes trivially. Insertion of the transformations (3) and (A5) and exploiting the linearity of the commutator gives

$$2[\tilde{\mathbf{r}}_i, \tilde{\mathbf{p}}_i] = 2[\mathbf{r}_i, \mathbf{p}_i] + \sum_{n=1}^{\infty} \left( [\mathbf{r}_i, (-\nabla_i \boldsymbol{\epsilon}_i)^n \cdot \mathbf{p}_i] + [\mathbf{r}_i, \mathbf{p}_i \cdot (-\nabla_i \boldsymbol{\epsilon}_i)^{\top n}] \right) + \sum_{n=0}^{\infty} \left( [\boldsymbol{\epsilon}_i, (-\nabla_i \boldsymbol{\epsilon}_i)^n \cdot \mathbf{p}_i] + [\boldsymbol{\epsilon}_i, \mathbf{p}_i \cdot (-\nabla_i \boldsymbol{\epsilon}_i)^{\top n}] \right), \quad (\text{A6})$$

where the contribution  $2[\boldsymbol{\epsilon}_i, \mathbf{p}_i]$  is included as the  $n=0$  term of the second series in equation (A6).

We rewrite the first term in the first series of equation (A6) using index notation:

$$[r_i^\alpha, (-\nabla_i \epsilon_i)_{\beta\gamma}^n p_i^\gamma] = (-\nabla_i \epsilon_i)_{\beta\gamma}^n [r_i^\alpha, p_i^\gamma] \quad (\text{A7})$$

$$= (-\nabla_i \epsilon_i)_{\beta\gamma}^n i\hbar \delta_{\alpha\gamma} \quad (\text{A8})$$

$$= i\hbar (-\nabla_i \epsilon_i)_{\beta\alpha}^n. \quad (\text{A9})$$

Here for readability the indices indicating the Cartesian components of the matrix  $(-\nabla_i \boldsymbol{\epsilon}_i)^n$  are written as subscripts. The factor  $(-\nabla_i \boldsymbol{\epsilon}_i)_{\beta\gamma}^n$  is local and hence commutes with position,  $[r_i^\alpha, (-\nabla_i \boldsymbol{\epsilon}_i)_{\beta\gamma}^n] = 0$ . Therefore this term can be taken outside of the commutator in equation (A7). We have inserted the usual position-momentum commutator in equation (A8) and evaluated the Kronecker delta in equation (A9).

Similarly we express the first term of the second series in equation (A6) as

$$[\epsilon_i^\alpha, (-\nabla_i \epsilon_i)_{\beta\gamma}^n p_i^\gamma] = (-\nabla_i \epsilon_i)_{\beta\gamma}^n [\epsilon_i^\alpha, p_i^\gamma] \quad (\text{A10})$$

$$= (-\nabla_i \epsilon_i)_{\beta\gamma}^n i\hbar (\nabla_i \epsilon_i)_{\gamma\alpha} \quad (\text{A11})$$

$$= -i\hbar (-\nabla_i \epsilon_i)_{\beta\alpha}^{n+1}, \quad (\text{A12})$$

where again the fact that  $\boldsymbol{\epsilon}_i$  and  $\nabla_i \boldsymbol{\epsilon}_i$  commute allows to take  $(-\nabla_i \boldsymbol{\epsilon}_i)_{\beta\gamma}^n$  out of the commutator in equation (A10). In equation (A11) we have inserted the commutator  $[\epsilon_i, \mathbf{p}_i] = i\hbar (\nabla_i \boldsymbol{\epsilon}_i)$ .

Recall that both expressions (A9) and (A12) are part of a sum in equation (A6). It becomes apparent that the  $(n+1)$ th term of the first sum cancels with the  $n$ th contribution of the second sum. (This amounts to renaming the summation index in the first sum of equation (A6) as  $n \rightarrow n+1$ . Then the first part of both occurring sums become identical up to a minus sign.) The only remaining terms are both second contributions to the sums of equation (A6). These corresponding transposed terms also cancel each other following an analogous argumentation. Thus no contribution to the sums in equation (A6) remains and we determine the canonical commutator as

$$[\tilde{\mathbf{r}}_i, \tilde{\mathbf{p}}_j] = [\mathbf{r}_i, \mathbf{p}_j] = i\hbar \delta_{ij} \mathbb{1}. \quad (\text{A13})$$

The above considerations generalize equations (6)–(8) from linear order to the general case.

Showing explicitly that the momentum self commutator vanishes,  $[\tilde{\mathbf{p}}_i, \tilde{\mathbf{p}}_j] = 0$  can be done similarly to the treatment of the position-momentum commutator by explicitly using the transformation (A5). The corresponding calculation is straightforward though tedious, and we omit it here.



## Appendix B. Local shift derivative of kinetic energy

### B.1. One dimension

As a preparation for the general three-dimensional case shown below, we first consider the simpler case of systems in one spatial dimension, with position  $x_i$  and momentum  $p_i = -i\hbar\partial/\partial x_i$  of particle  $i = 1, \dots, N$ . We consider a one-dimensional displacement  $\epsilon(x)$  of the position coordinate  $x$ , such that  $x_i \rightarrow x_i + \epsilon(x_i)$ , in analog to the three-dimensional case of equation (3). The one-dimensional momentum transformation [corresponding to equation (5)] is  $p_i \rightarrow p_i - \{\epsilon'(x_i)p_i + p_i\epsilon'(x_i)\}/2 \equiv \tilde{p}_i$ , where the prime denotes the derivative by the argument.

We wish to derive the one-dimensional analogue of equation (15), which reads as

$$\frac{\delta}{\delta\epsilon(x)} \sum_i \frac{\tilde{p}_i^2}{2m} = \frac{\partial}{\partial x} \sum_i \frac{p_i \delta_i p_i}{m} - \frac{\hbar^2}{4m} \frac{\partial^3}{\partial x^3} \sum_i \delta_i. \quad (\text{B1})$$

Here the density operator of particle  $i$  is defined as  $\delta_i = \delta(x_i - x)$  and the identity holds at  $\epsilon(x) = 0$ . The functional derivative on the left-hand side of equation (B1) can be moved inside of the sum over all particles and we hence need to consider

$$\frac{\delta \tilde{p}_i^2}{\delta\epsilon(x)} = p_i \frac{\delta \tilde{p}_i}{\delta\epsilon(x)} + \frac{\delta \tilde{p}_i}{\delta\epsilon(x)} p_i. \quad (\text{B2})$$

This equality holds to first order in  $\epsilon(x)$ , as we have replaced  $\tilde{p}_i$  by  $p_i$  on the right-hand side.

The first term on the right-hand side of equation (B2) becomes

$$p_i \frac{\delta \tilde{p}_i}{\delta\epsilon(x)} = \partial_x (p_i \delta_i p_i + p_i p_i \delta_i) / 2 \quad (\text{B3})$$

$$= \partial_x \{p_i \delta_i p_i + p_i (p_i \delta_i) + p_i \delta_i p_i\} / 2, \quad (\text{B4})$$

where  $\partial_x = \partial/\partial x$  is a shortcut notation. We have used  $\delta p_i / \delta\epsilon(x) = \partial_x (\delta_i p_i + p_i \delta_i) / 2$ , i.e. the one-dimensional analogue of equation (12), in the first equality and the product rule of differentiation for  $p_i$  in the second equality. The remaining second term in equation (B2) is

$$\frac{\delta \tilde{p}_i}{\delta\epsilon(x)} p_i = \partial_x (\delta_i p_i p_i + p_i \delta_i p_i) / 2 \quad (\text{B5})$$

$$= \partial_x \{p_i \delta_i p_i - (p_i \delta_i) p_i + p_i \delta_i p_i\} / 2. \quad (\text{B6})$$

The minus sign in equation (B6) allows to simplify the sum of the respective second terms:

$$p_i (p_i \delta_i) - (p_i \delta_i) p_i = (p_i p_i \delta_i) \quad (\text{B7})$$

$$= -\hbar^2 \left( \frac{\partial^2}{\partial x_i^2} \delta_i \right) \quad (\text{B8})$$

$$= -\hbar^2 \partial_x^2 \delta_i, \quad (\text{B9})$$

where in the second equality we have expressed the effect of the momentum operator on the delta function by the (negative) position gradient, i.e.

$$(p_i \delta_i) = -i\hbar \frac{\partial \delta(x - x_i)}{\partial x_i} = i\hbar \frac{\partial \delta(x - x_i)}{\partial x}. \quad (\text{B10})$$

Collecting all terms yields

$$\frac{\delta \tilde{p}_i^2}{\delta\epsilon(x)} = \frac{\partial}{\partial x} 2p_i \delta_i p_i - \frac{\hbar^2}{2} \frac{\partial^3}{\partial x^3} \delta_i, \quad (\text{B11})$$

and summation over  $i$  and division by  $2m$  then yields equation (B1), as desired.

The three-dimensional case covered below is closely related, with the additional complexity of the matrix and tensor indices interfering very little with the operator structure.

### B.2. Three dimensions

We wish to derive equation (13), which we reproduce for convenience:

$$\frac{\delta}{\delta\epsilon^\gamma} \sum_i \frac{\tilde{\mathbf{p}}_i^2}{2m} \Big|_{\epsilon=0} = \nabla \cdot \sum_i \frac{\mathbf{p}_i \delta_i \mathbf{p}_i + \mathbf{p}_i \delta_i \mathbf{p}_i^\top}{2m} - \frac{\hbar^2}{4m} \nabla \nabla^2 \sum_i \delta(\mathbf{r} - \mathbf{r}_i). \quad (\text{B12})$$

We use Einstein summation convention over pairs of Greek indices and after taking the functional derivative set  $\epsilon(\mathbf{r}) = 0$  throughout. We consider the  $\gamma$ th component of the left-hand side of equation (B12) for particle  $i$  only, which yields

$$\frac{\delta}{\delta\epsilon^\gamma} \tilde{p}_i^\alpha \tilde{p}_i^\alpha = p_i^\alpha \frac{\delta \tilde{p}_i^\alpha}{\delta\epsilon^\gamma} + \frac{\delta \tilde{p}_i^\alpha}{\delta\epsilon^\gamma} p_i^\alpha, \quad (\text{B13})$$

where the sum over  $\alpha$  (repeated index) generates the square of momentum, as it occurs in the kinetic energy. The first term on the right-hand side, using the explicit form of the transformed momentum (5), becomes

$$p_i^\alpha \frac{\delta \tilde{p}_i^\alpha}{\delta\epsilon^\gamma} = \nabla^\alpha (p_i^\alpha \delta_i p_i^\gamma + p_i^\alpha p_i^\gamma \delta_i) / 2 \quad (\text{B14})$$

$$= \nabla^\alpha \{ p_i^\alpha \delta_i p_i^\gamma + p_i^\alpha (p_i^\gamma \delta_i) + p_i^\alpha \delta_i p_i^\gamma \} / 2, \quad (\text{B15})$$

where we have used equation (12) in the first equality and the product rule of differentiation for the application of  $p_i^\gamma$  in the second equality. The remaining second term in equation (B13) is

$$\frac{\delta \tilde{p}_i^\alpha}{\delta\epsilon^\gamma} p_i^\alpha = \nabla^\alpha (\delta_i p_i^\gamma p_i^\alpha + p_i^\gamma \delta_i p_i^\alpha) / 2 \quad (\text{B16})$$

$$= \nabla^\alpha \{ p_i^\gamma \delta_i p_i^\alpha - (p_i^\gamma \delta_i) p_i^\alpha + p_i^\gamma \delta_i p_i^\alpha \} / 2. \quad (\text{B17})$$

The appearance of the minus sign in equation (B17) allows to carry out the following cancellation of the respective ‘middle’ terms:

$$p_i^\alpha (p_i^\gamma \delta_i) - (p_i^\gamma \delta_i) p_i^\alpha = (p_i^\alpha p_i^\gamma \delta_i) \quad (\text{B18})$$

$$= -\hbar^2 (\nabla_i^\alpha \nabla_i^\gamma \delta_i) \quad (\text{B19})$$

$$= -\hbar^2 \nabla^\alpha \nabla^\gamma \delta_i, \quad (\text{B20})$$

where in the second step we have rewritten the effect of the momentum operator on the delta function by the (negative) position gradient, i.e.

$$(p_i^\gamma \delta_i) = -i\hbar (\nabla_i^\gamma \delta_i) = i\hbar (\nabla^\gamma \delta_i). \quad (\text{B21})$$

Collecting all terms we obtain the overall result for the shift derivative of kinetic energy,

$$\frac{\delta}{\delta\epsilon^\gamma} \sum_i \frac{\tilde{p}_i^\alpha \tilde{p}_i^\alpha}{2m} \Big|_{\epsilon=0} = \nabla^\alpha \left\{ \sum_i \frac{p_i^\alpha \delta_i p_i^\gamma + p_i^\gamma \delta_i p_i^\alpha}{2m} - \frac{\hbar^2}{4m} \nabla^\gamma \nabla^\alpha \sum_i \delta_i \right\} \quad (\text{B22})$$

$$= -\nabla^\alpha \hat{\tau}^{\alpha\gamma}. \quad (\text{B23})$$

As desired, equation (B22) is the  $\gamma$ th Cartesian component of equation (13) [reproduced above as equation (B12)] and equation (B23) is analogous to equation (15), with tensor contractions

and matrix transpositions expressed in index notation, and the definition of the kinetic stress operator as given by equation (14).

### ORCID iDs

Sophie Hermann  <https://orcid.org/0000-0002-4012-9170>

Matthias Schmidt  <https://orcid.org/0000-0002-5015-2972>

### References

- [1] Hansen J P and McDonald I R 2013 *Theory of Simple Liquids* 4th edn (London: Academic)
- [2] Reif F 2007 *Statistical and Thermal Physics* 18th edn (New York: McGraw-Hill)
- [3] Tarantino W and Ullrich C A 2021 A reformulation of time-dependent Kohn-Sham theory in terms of the second time derivative of the density *J. Chem. Phys.* **154** 204112
- [4] Tchenkoue M-L M, Penz M, Theophilou I, Ruggenthaler M and Rubio A 2019 Force balance approach for advanced approximations in density functional theories *J. Chem. Phys.* **151** 154107
- [5] Schmidt M 2022 Power functional theory for many-body dynamics *Rev. Mod. Phys.* **94** 015007
- [6] Tokatly I V 2007 Time-dependent deformation functional theory *Phys. Rev. B* **75** 125105
- [7] Tokatly I V 2005 Quantum many-body dynamics in a Lagrangian frame: I. Equations of motion and conservation laws *Phys. Rev. B* **71** 165104
- [8] Tokatly I V 2005 Quantum many-body dynamics in a Lagrangian frame: II. Geometric formulation of time-dependent density functional theory *Phys. Rev. B* **71** 165105
- [9] Ullrich C A and Tokatly I V 2006 Nonadiabatic electron dynamics in time-dependent density-functional theory *Phys. Rev. B* **73** 235102
- [10] Schmidt M and Brader J M 2013 Power functional theory for Brownian dynamics *J. Chem. Phys.* **138** 214101
- [11] Schmidt M 2018 Power functional theory for Newtonian many-body dynamics *J. Chem. Phys.* **148** 044502
- [12] Schmidt M 2015 Quantum power functional theory for many-body dynamics *J. Chem. Phys.* **143** 174108
- [13] Brütting M, Trepl M T, de las Heras D and Schmidt M 2019 Superadiabatic forces via the acceleration gradient in quantum many-body dynamics *Molecules* **24** 3660
- [14] Evans R 1979 The nature of the liquid-vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids *Adv. Phys.* **28** 143
- [15] Evans R 1992 Density functionals in the theory nonuniform fluids *Fundamentals of Inhomogeneous Fluids* ed D Henderson (New York: Dekker)
- [16] For an overview of new developments in classical density functional theory, see: Evans R, Oettel M, Roth R and Kahl G 2016 New developments in classical density functional theory *J. Phys.: Condens. Matter.* **28** 240401
- [17] Marconi U M B and Tarazona P 1999 Dynamic density functional theory of fluids *J. Chem. Phys.* **110** 8032
- [18] Archer A J and Evans R 2004 Dynamical density functional theory and its application to spinodal decomposition *J. Chem. Phys.* **121** 4246
- [19] Rotenberg B 2020 Use the force! Reduced variance estimators for densities, radial distribution functions and local mobilities in molecular simulations *J. Chem. Phys.* **153** 150902
- [20] de las Heras D and Schmidt M 2018 Better than counting: density profiles from force sampling *Phys. Rev. Lett.* **120** 218001
- [21] Borgis D, Assaraf R, Rotenberg B and Vuilleumier R 2013 Computation of pair distribution functions and three-dimensional densities with a reduced variance principle *Mol. Phys.* **111** 3486
- [22] Purohit A, Schultz A J and Kofke D A 2019 Force-sampling methods for density distributions as instances of mapped averaging *Mol. Phys.* **117** 2822
- [23] Yvon J 1935 *Actualités Scientifiques et Industrielles* (Paris: Hermann & Cie.)
- [24] Born M and Green H S 1946 A general kinetic theory of liquids I. The molecular distribution functions *Proc. R. Soc. A* **188** 10
- [25] Tschopp S M and Brader J M 2021 Fundamental measure theory of inhomogeneous two-body correlation functions *Phys. Rev. E* **103** 042103

- [26] McCarty R J, Perchak D, Pederson R, Evans R, Qiu Y, White S R and Burke K 2020 Bypassing the energy functional in density functional theory: direct calculation of electronic energies from conditional probability densities *Phys. Rev. Lett.* **125** 266401
- [27] Pederson R, Chen J, White S R and Burke K 2022 Conditional probability density functional theory *Phys. Rev. B* **105** 245138
- [28] Noether E 1918 Invariante Variationsprobleme *Nachr. König. Ges. Wiss. Gött. Math.-Phys. Klasse* **235** 183  
Tavel M A 1971 Invariant variation problems *Transp. Theory Stat. Phys.* **1** 186 (Engl. transl.) for a version in modern typesetting see: Wang F Y 2018 arXiv:[physics/0503066v3](https://arxiv.org/abs/physics/0503066v3)
- [29] Byers N 1998 E. Noether's discovery of the deep connection between symmetries and conservation laws (arXiv:[physics/9807044](https://arxiv.org/abs/physics/9807044))
- [30] Hermann S and Schmidt M 2021 Noether's theorem in statistical mechanics *Commun. Phys.* **4** 176
- [31] Hermann S and Schmidt M 2022 Why Noether's theorem applies to statistical mechanics *J. Phys.: Condens. Matter* **34** 213001
- [32] Hermann S and Schmidt M 2022 Variance of fluctuations from Noether invariance *Commun. Phys.* **5** 276
- [33] Tschopp S M, Sammüller F, Hermann S, Schmidt M and Brader J M 2022 Force density functional theory in- and out-of-equilibrium *Phys. Rev. E* **106** 014115
- [34] Sardanashvily G 2016 *Noether's Theorems, Applications in Mechanics and Field Theory (Atlantis Studies in Variational Geometry)* ed D Krupka and H Sun (New York: Atlantis Press)
- [35] Kosmann-Schwarzbach Y 2018 *The Noether Theorems, Invariance and Conservation Laws in the Twentieth Century* (New York: Springer)
- [36] Neuenschwander D E 2011 *Emmy Noether's Wonderful Theorem* (Baltimore: Johns Hopkins University Press)
- [37] Goldstein H, Poole C and Safko J 2002 *Classical Mechanics* (New York: Addison-Wesley)
- [38] Anderson A 1994 Canonical transformations in quantum mechanics *Ann. Phys.* **232** 292
- [39] Giesbertz K J H and Ruggenthaler M 2019 One-body reduced density-matrix functional theory in finite basis sets at elevated temperatures *Phys. Rep.* **806** 1–47
- [40] In his recollections Mermin N D 2003 My life with Fisher *J. Stat. Phys.* **110** 467, David Mermin describes his encounter with Bob Griffiths who “let it be known that what he was up to was proving that the free energy of a spin system exists. ‘That it *what?*’ I said.” (Mermin's emphasis)
- [41] Sprik M 2021 Continuum model of the simple dielectric fluid: consistency between density based and continuum mechanics methods *Mol. Phys.* **119** e1887950
- [42] Sprik M 2021 Chemomechanical equilibrium at the interface between a simple elastic solid and its liquid phase *J. Chem. Phys.* **155** 244701

# communications physics

ARTICLE

<https://doi.org/10.1038/s42005-022-01046-3>

OPEN

## Variance of fluctuations from Noether invariance

Sophie Hermann <sup>1</sup>✉ & Matthias Schmidt <sup>1</sup>✉

The strength of fluctuations, as measured by their variance, is paramount in the quantitative description of a large class of physical systems, ranging from simple and complex liquids to active fluids and solids. Fluctuations originate from the irregular motion of thermal degrees of freedom and statistical mechanics facilitates their description. Here we demonstrate that fluctuations are constrained by the inherent symmetries of the given system. For particle-based classical many-body systems, Noether invariance at second order in the symmetry parameter leads to exact sum rules. These identities interrelate the global force variance with the mean potential energy curvature. Noether invariance is restored by an exact balance between these distinct mechanisms. The sum rules provide a practical guide for assessing and constructing theories, for ensuring self-consistency in simulation work, and for providing a systematic pathway to the theoretical quantification of fluctuations.

<sup>1</sup>Theoretische Physik II, Physikalisches Institut, Universität Bayreuth, D-95447 Bayreuth, Germany. ✉email: [Sophie.Hermann@uni-bayreuth.de](mailto:Sophie.Hermann@uni-bayreuth.de); [Matthias.Schmidt@uni-bayreuth.de](mailto:Matthias.Schmidt@uni-bayreuth.de)

Applying Noether's theorem<sup>1</sup> to a physical problem requires identifying and hence exploiting the fundamental symmetries of the system under consideration. Independent of whether such work is performed in a Hamiltonian setting or on the basis of an action functional, typically it is a conservation law that results from each inherent symmetry of the system. The merits of the Noetherian strategy have been demonstrated in a variety of contexts from classical mechanics to field theory<sup>2</sup>. However, much of modern condensed matter physics is focused on seemingly entirely different physical behavior, namely that of fluctuating, disordered, spatially random, yet strongly interacting systems that possess a large number of degrees of freedom. Recent examples include active particles that display freezing<sup>3</sup> and wetting<sup>4</sup>, hydrophobicity rationalized as critical drying<sup>5</sup>, the structure of two-dimensional colloidal liquids<sup>6</sup> and that of fluid interfaces<sup>7,8</sup>.

Relating the fluctuations that occur in complex systems to the underlying symmetries has been investigated in a variety of contexts. Such work addressed the symmetries in fluctuations far from equilibrium<sup>9</sup>, isometric fluctuation relations<sup>10</sup>, fluctuation relations for equilibrium states with broken symmetry<sup>11</sup>, and fluctuation-response out of equilibrium<sup>12</sup>. The fluctuation theorems of stochastic thermodynamics provide a systematic setup to address such questions<sup>13</sup>. Beyond its widespread use in deterministic settings, Noether's theorem was formulated and used in a stochastic context<sup>14</sup>, for Markov processes<sup>15</sup>, for the quantification of the asymmetry of quantum states<sup>16</sup>, for formulating entropy as a Noether invariant<sup>17,18</sup>, and for studying the thermodynamical path integral and emergent symmetry<sup>19</sup>. Early work was carried out by Revzen<sup>20</sup> in the context of functional integrals in statistical physics and a recent perspective from an algebraic point of view was given by Baez<sup>21</sup>.

Noether's theorem has recently been suggested to be applicable in a genuine statistical mechanical fashion<sup>22–24</sup>. Based on translational and rotational symmetries the theorem allows to derive exact identities ("sum rules") with relative ease for relevant many-body systems both in and out of equilibrium. The sum rules set constraints on the global forces and torques in the system, such as the vanishing of the global external force in equilibrium<sup>22,25</sup> and of the global internal force also in nonequilibrium<sup>22</sup>.

Here we demonstrate that Noether's theorem allows to go beyond mere averages and systematically address the strength of fluctuations, as measured by the variance (auto-correlation). We demonstrate that this variance is balanced by the mean potential curvature, which hence restores the Noether invariance. The structure emerges when going beyond the usual linear expansion in the symmetry parameter. The relevant objects to be transformed are cornerstones of Statistical Mechanics, such as the grand potential in its elementary form and the free energy density functional. The invariances constrain both density fluctuations and direct correlations, where the latter are generated from functional differentiation of the excess (over ideal gas) density functional.

## Results and discussion

**External force variance.** We work in the grand ensemble and express the associated grand potential in its elementary form<sup>26</sup> as

$$\Omega[V_{\text{ext}}] = -k_B T \ln \text{Tr} \exp \left( -\beta \left( H_{\text{int}} + \sum_i V_{\text{ext}}(\mathbf{r}_i) - \mu N \right) \right), \quad (1)$$

where  $k_B$  indicates the Boltzmann constant,  $T$  is absolute temperature, and  $\beta = 1/(k_B T)$  is inverse temperature. The grand ensemble "trace" is denoted by  $\text{Tr} = \sum_{N=0}^{\infty} 1/(N! h^{3N}) \int d\mathbf{r}_1 \dots d\mathbf{r}_N d\mathbf{p}_1 \dots d\mathbf{p}_N$ , where  $\mathbf{r}_i$  is the position and  $\mathbf{p}_i$  is the momentum of

particle  $i = 1, \dots, N$ , with  $N$  being the total number of particles and  $h$  the Planck constant. The internal part of the Hamiltonian is  $H_{\text{int}} = \sum_i \mathbf{p}_i^2 / (2m) + u(\mathbf{r}_1, \dots, \mathbf{r}_N)$ , where  $m$  indicates the particle mass,  $u(\mathbf{r}_1, \dots, \mathbf{r}_N)$  is the interparticle interaction potential, and  $V_{\text{ext}}(\mathbf{r})$  is the external one-body potential as a function of position  $\mathbf{r}$ . The thermodynamic parameters are the chemical potential  $\mu$  and temperature  $T$ .

Clearly, the value of the grand potential  $\Omega[V_{\text{ext}}]$  depends on the function  $V_{\text{ext}}(\mathbf{r})$  and we have indicated this functional dependence by the brackets. We consider a spatial displacement by a constant vector  $\boldsymbol{\epsilon}$ , applied to the entire system. The external potential is hence modified according to  $V_{\text{ext}}(\mathbf{r}) \rightarrow V_{\text{ext}}(\mathbf{r} + \boldsymbol{\epsilon})$ . This displacement leaves the kinetic energy invariant (the momenta are unaffected) and it does not change the interparticle potential  $u(\mathbf{r}_1, \dots, \mathbf{r}_N)$ , as its dependence is only on difference vectors  $\mathbf{r}_i - \mathbf{r}_j$ , which are unaffected by the global displacement. Throughout we do not consider the dynamics of the shifting and rather only compare statically the original with the displaced system, with both being in equilibrium. (Hermann and Schmidt<sup>22</sup> present dynamical Noether sum rules that arise from invariance of the power functional<sup>27</sup> at first order in a time-dependent shifting protocol  $\boldsymbol{\epsilon}(t)$ .) The invariance with respect to the displacement can be explicitly seen by transforming each position integral in the trace over phase space as  $\int d\mathbf{r}_i = \int d(\mathbf{r}_i - \boldsymbol{\epsilon})$ . No boundary terms occur as the integral is over  $\mathbb{R}^3$ ; the effect of system walls is explicitly contained in the form of  $V_{\text{ext}}(\mathbf{r})$ . This coordinate shift formally "undoes" the spatial system displacement and it renders the form of the partition sum identical to that of the original system. (See the work of Tschopp et al.<sup>24</sup> for the generalization from homogeneous shifting to a position-dependent operation.)

The Taylor expansion of the grand potential of the displaced system around the original system is

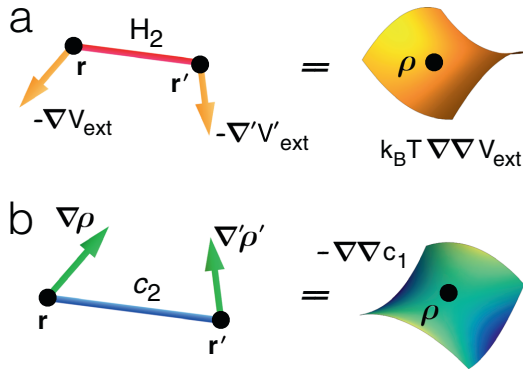
$$\begin{aligned} \Omega[V_{\text{ext}}^\epsilon] &= \Omega[V_{\text{ext}}] + \int d\mathbf{r} \rho(\mathbf{r}) \nabla V_{\text{ext}}(\mathbf{r}) \cdot \boldsymbol{\epsilon} \\ &+ \frac{1}{2} \int d\mathbf{r} \rho(\mathbf{r}) \nabla \nabla V_{\text{ext}}(\mathbf{r}) : \boldsymbol{\epsilon} \boldsymbol{\epsilon} \\ &- \frac{\beta}{2} \int d\mathbf{r} d\mathbf{r}' H_2(\mathbf{r}, \mathbf{r}') \nabla V_{\text{ext}}(\mathbf{r}) \nabla' V_{\text{ext}}(\mathbf{r}') : \boldsymbol{\epsilon} \boldsymbol{\epsilon}, \end{aligned} \quad (2)$$

where we have truncated at second order in  $\boldsymbol{\epsilon}$  and have used the shortcut notation  $V_{\text{ext}}^\epsilon(\mathbf{r}) = V_{\text{ext}}(\mathbf{r} + \boldsymbol{\epsilon})$  for the functional argument on the left hand side of Eq. (2). The colon indicates a double tensor contraction and  $\nabla V_{\text{ext}}(\mathbf{r}) \nabla' V_{\text{ext}}(\mathbf{r}')$  is the dyadic product of the external force field with itself. ( $\nabla'$  denotes the derivative with respect to  $\mathbf{r}'$ .) The occurrence of the one-body density profile  $\rho(\mathbf{r})$  and of the correlation function of density fluctuations  $H_2(\mathbf{r}, \mathbf{r}')$  is due to the functional identities  $\rho(\mathbf{r}) = \delta \Omega[V_{\text{ext}}] / \delta V_{\text{ext}}(\mathbf{r})$  and  $H_2(\mathbf{r}, \mathbf{r}') = -k_B T \delta^2 \Omega[V_{\text{ext}}] / \delta V_{\text{ext}}(\mathbf{r}) \delta V_{\text{ext}}(\mathbf{r}')$ <sup>26–29</sup>.

The Noetherian invariance against the displacement implies that the value of the grand potential remains unchanged upon shifting, and hence  $\Omega[V_{\text{ext}}^\epsilon] = \Omega[V_{\text{ext}}]$ <sup>23</sup>. As a consequence, both the first and the second-order terms in the Taylor expansion (2) need to vanish identically, and this holds irrespectively of the value of  $\boldsymbol{\epsilon}$ ; i.e. both the orientation and the magnitude of  $\boldsymbol{\epsilon}$  can be arbitrary. This yields, respectively, the first<sup>22,25</sup> and second-order<sup>30,31</sup> identities

$$- \int d\mathbf{r} \rho(\mathbf{r}) \nabla V_{\text{ext}}(\mathbf{r}) = 0, \quad (3)$$

$$\int d\mathbf{r} d\mathbf{r}' H_2(\mathbf{r}, \mathbf{r}') \nabla V_{\text{ext}}(\mathbf{r}) \nabla' V_{\text{ext}}(\mathbf{r}') = k_B T \int d\mathbf{r} \rho(\mathbf{r}) \nabla \nabla V_{\text{ext}}(\mathbf{r}). \quad (4)$$



**Fig. 1 Illustrations of the sum rules for the variance of fluctuations.** The sum rules arise from Noether invariance against spatial displacement. Shown are the different types of identical integrals. Thick dots indicate position variables that are integrated over. **a** External sum rule, Eq. (4), which relates the correlation function of density fluctuations  $H_2(\mathbf{r}, \mathbf{r}')$  and the external force field  $-\nabla V_{\text{ext}}(\mathbf{r})$  with the product of the density profile  $\rho(\mathbf{r})$  and the Hessian of the external potential  $k_B T \nabla \nabla V_{\text{ext}}(\mathbf{r})$ . This curvature is indicated by a schematic heat map. **b** Internal sum rule, Eq. (8), where the density gradient at two different positions is bonded by the direct correlation function  $c_2(\mathbf{r}, \mathbf{r}')$ . This integral is identical to the integrated Hessian  $-\nabla \nabla c_1(\mathbf{r})$  (indicated by a schematic heat map) weighted by the local density  $\rho(\mathbf{r})$ .

We can rewrite the sum rule (3) in the compact form  $\langle \hat{\mathbf{F}}_{\text{ext}}^0 \rangle = 0$ , where we have introduced the global external force operator  $\hat{\mathbf{F}}_{\text{ext}}^0 \equiv -\sum_i \nabla_i V_{\text{ext}}(\mathbf{r}_i) = -\int d\mathbf{r} \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \nabla_i V_{\text{ext}}(\mathbf{r}_i)$ . The angular brackets denote the equilibrium average  $\langle \cdot \rangle = \text{Tr} \Psi$ , where the grand ensemble distribution function is  $\Psi = e^{-\beta(H - \mu N)} / \Xi$ , with  $H = H_{\text{int}} + \sum_i V_{\text{ext}}(\mathbf{r}_i)$  and the grand partition sum is  $\Xi = \text{Tr} e^{-\beta(H - \mu N)}$ . Using these averages, and defining the density operator  $\hat{\rho}(\mathbf{r}) = \sum_i \delta(\mathbf{r} - \mathbf{r}_i)$ , where  $\delta(\cdot)$  denotes the Dirac distribution, allows us to express the density profile as  $\rho(\mathbf{r}) = \langle \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \rangle$ . The covariance of the density operator is  $H_2(\mathbf{r}, \mathbf{r}') = \langle \hat{\rho}(\mathbf{r}) \hat{\rho}(\mathbf{r}') \rangle - \rho(\mathbf{r}) \rho(\mathbf{r}')$ , which complements the above definition of  $H_2(\mathbf{r}, \mathbf{r}')$  via the second functional derivative of the grand potential.

The second-order sum rule (4) constrains the variance of the external force operator on its left hand side:  $\langle \hat{\mathbf{F}}_{\text{ext}}^0 \hat{\mathbf{F}}_{\text{ext}}^0 \rangle - \langle \hat{\mathbf{F}}_{\text{ext}}^0 \rangle \langle \hat{\mathbf{F}}_{\text{ext}}^0 \rangle = \langle \hat{\mathbf{F}}_{\text{ext}}^0 \hat{\mathbf{F}}_{\text{ext}}^0 \rangle$ ; recall that the average (first moment) of the external force vanishes, see Eq. (3). The right-hand side of Eq. (4) balances the strength of these force fluctuations by the mean curvature of the external potential (multiplied by thermal energy  $k_B T$ ), see Fig. 1(a) for an illustration of the structure of the integrals.

The curvature term can be re-written, upon integration by parts, as  $\int d\mathbf{r} (-k_B T \nabla \rho(\mathbf{r})) \nabla V_{\text{ext}}(\mathbf{r})$ , which is the integral of the local correlation of the ideal force density,  $-k_B T \nabla \rho(\mathbf{r})$ , and the negative external force field  $\nabla V_{\text{ext}}(\mathbf{r})$ . (We assume setups with closed walls, where boundary terms vanish.) The sum rule (4) remains valid if one replaces  $H_2(\mathbf{r}, \mathbf{r}')$  by the two-body density  $\rho_2(\mathbf{r}, \mathbf{r}') = \langle \hat{\rho}(\mathbf{r}) \hat{\rho}(\mathbf{r}') \rangle$ , due to the vanishing of the external force (3). Explicitly, the alternative form of Eq. (4) that one obtains via this replacement is:  $\int d\mathbf{r} d\mathbf{r}' \rho_2(\mathbf{r}, \mathbf{r}') \nabla V_{\text{ext}}(\mathbf{r}) \nabla V_{\text{ext}}(\mathbf{r}') = k_B T \int d\mathbf{r} \rho(\mathbf{r}) \nabla \nabla V_{\text{ext}}(\mathbf{r})$ .

It is standard practice<sup>26–29</sup> to split off the trivial density covariance of the ideal gas and define the total correlation function  $h(\mathbf{r}, \mathbf{r}')$  via the identity  $H_2(\mathbf{r}, \mathbf{r}') = \rho(\mathbf{r}) \rho(\mathbf{r}') h(\mathbf{r}, \mathbf{r}') +$

$\rho(\mathbf{r}) \delta(\mathbf{r} - \mathbf{r}')$ . Insertion of this relation into Eq. (4) and then moving the term with the delta function to the right-hand side yields the following alternative form of the second-order Noether sum rule:

$$\int d\mathbf{r} d\mathbf{r}' \rho(\mathbf{r}) \rho(\mathbf{r}') h(\mathbf{r}, \mathbf{r}') \nabla V_{\text{ext}}(\mathbf{r}) \nabla V_{\text{ext}}(\mathbf{r}') = \int d\mathbf{r} (k_B T \nabla \nabla V_{\text{ext}}(\mathbf{r}) - (\nabla V_{\text{ext}}(\mathbf{r})) \nabla V_{\text{ext}}(\mathbf{r})) \rho(\mathbf{r}). \quad (5)$$

For the ideal gas  $h(\mathbf{r}, \mathbf{r}') = 0$  and hence the left hand side of (5) vanishes. That the right-hand side then also vanishes can be seen explicitly by inserting the generalized barometric law<sup>26</sup>  $\rho(\mathbf{r}) \propto \exp(-\beta(V_{\text{ext}}(\mathbf{r}) - \mu))$  and either integrating by parts, or by alternatively observing that  $-(k_B T)^2 \int d\mathbf{r} \nabla \nabla \rho(\mathbf{r}) = 0$  and inserting the barometric law therein.

The right-hand side of (5) makes explicit the balancing of the external force variance with the mean potential curvature, as given by its averaged Hessian. For an interacting (non-ideal) system,  $h(\mathbf{r}, \mathbf{r}')$  is nonzero in general and the associated external force correlation contributions are accumulated by the expression on the left hand side of Eq. (5). For the special case of a harmonic trap, as represented by the external potential  $V_{\text{ext}}(\mathbf{r}) = \kappa r^2/2$ , with spring constant  $\kappa$  and Hessian  $\nabla \nabla V_{\text{ext}}(\mathbf{r}) = \kappa \mathbb{1}$ , where  $\mathbb{1}$  denotes the unit matrix, the mean curvature can be obtained explicitly. The first term on the right-hand side of the sum rule (5) then simply becomes  $k_B T \langle N \rangle \kappa \mathbb{1}$  upon integration. Notably, this result holds independently of the type of interparticle interactions, although the latter affect  $h(\mathbf{r}, \mathbf{r}')$  as is present on the left hand side of Eq. (5). The remaining (second) term on the right-hand side of Eq. (5) turns into  $-\kappa^2 \int d\mathbf{r} \rho(\mathbf{r}) \mathbf{r} \mathbf{r}$ , where the integral is the matrix of second spatial moments of the density profile. The alternative form  $-\kappa^2 \langle \sum_i \mathbf{r}_i \mathbf{r}_i \rangle$  is obtained upon expressing the density profile as the average of  $\hat{\rho}(\mathbf{r})$  and carrying out the integral over  $\mathbf{r}$ . Collecting all terms and dividing by  $\kappa^2$  we obtain the sum rule (5) for the case of an interacting system inside of a harmonic trap as:  $\int d\mathbf{r} d\mathbf{r}' \rho(\mathbf{r}) \rho(\mathbf{r}') h(\mathbf{r}, \mathbf{r}') \mathbf{r} \mathbf{r}' = \int d\mathbf{r} \rho(\mathbf{r}) (k_B T \kappa^{-1} \mathbb{1} - \mathbf{r} \mathbf{r})$ .

**Internal force variance.** In light of the external force fluctuations, one might wonder whether the global interparticle force also fluctuates. The corresponding operator is the sum of all interparticle forces:  $\hat{\mathbf{F}}_{\text{int}}^0 \equiv -\sum_i \nabla_i u(\mathbf{r}_1, \dots, \mathbf{r}_N) = -\int d\mathbf{r} \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \nabla_i u(\mathbf{r}_1, \dots, \mathbf{r}_N)$ , where the integrand in the later expression (including the minus sign) is the position-resolved force density operator<sup>27</sup>. However, for each microstate  $\hat{\mathbf{F}}_{\text{int}}^0 = 0$ , as can be seen e.g. via the translation invariance of the interparticle potential<sup>22</sup>, which ultimately expresses Newton's third law *actio est reactio*. Hence trivially the average vanishes,  $\langle \hat{\mathbf{F}}_{\text{int}}^0 \rangle = 0$ , as do all higher moments,  $\langle \hat{\mathbf{F}}_{\text{int}}^0 \hat{\mathbf{F}}_{\text{int}}^0 \rangle = 0$ , as well as cross correlations,  $\langle \hat{\mathbf{F}}_{\text{int}}^0 \hat{\mathbf{F}}_{\text{ext}}^0 \rangle = 0$ , etc. Thus the total internal force does not fluctuate. This holds beyond equilibrium, as the properties of the thermal average are not required in the argument. Identical reasoning can be applied to a nonequilibrium ensemble, where these identities hence continue to hold.

While these *probabilistic* correlators vanish, deeper inherent structure can be revealed by addressing direct correlations, as introduced by Ornstein and Zernike in 1914 in their treatment of critical opalescence and to great benefit exploited in modern liquid state theory<sup>26</sup>. We use the framework of classical density functional theory<sup>26,28,29</sup>, where the effect of the interparticle interactions is encapsulated in the intrinsic Helmholtz excess free energy  $F_{\text{exc}}[\rho]$  as a functional of the one-body density distribution  $\rho(\mathbf{r})$ . As the excess free energy functional solely depends on the interparticle interactions, it necessarily is invariant against spatial displacements. In technical analogy to the previous case of the

external force, we consider a displaced density profile  $\rho(\mathbf{r} + \boldsymbol{\epsilon})$  and Taylor expand the excess free energy functional up to second order in  $\boldsymbol{\epsilon}$  as follows:

$$\begin{aligned} \beta F_{\text{exc}}[\rho^\epsilon] &= \beta F_{\text{exc}}[\rho] - \int d\mathbf{r} c_1(\mathbf{r}) \nabla \rho(\mathbf{r}) \cdot \boldsymbol{\epsilon} \\ &\quad - \frac{1}{2} \int d\mathbf{r} c_1(\mathbf{r}) \nabla \nabla \rho(\mathbf{r}) : \boldsymbol{\epsilon} \boldsymbol{\epsilon} \\ &\quad - \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' c_2(\mathbf{r}, \mathbf{r}') \nabla \rho(\mathbf{r}) \nabla' \rho(\mathbf{r}') : \boldsymbol{\epsilon} \boldsymbol{\epsilon}, \end{aligned} \quad (6)$$

where  $\rho^\epsilon(\mathbf{r}) = \rho(\mathbf{r} + \boldsymbol{\epsilon})$  is again a shorthand. The one- and two-body direct correlation functions are given, respectively, via the functional derivatives  $c_1(\mathbf{r}) = -\beta \delta F_{\text{exc}}[\rho] / \delta \rho(\mathbf{r})$  and  $c_2(\mathbf{r}, \mathbf{r}') = -\beta \delta^2 F_{\text{exc}}[\rho] / \delta \rho(\mathbf{r}) \delta \rho(\mathbf{r}')$ . Noether invariance demands that  $F_{\text{exc}}[\rho^\epsilon] = F_{\text{exc}}[\rho]$  and hence both the linear and the quadratic contributions in the Taylor expansion (6) need to vanish, irrespective of the value of  $\boldsymbol{\epsilon}$ . This yields, respectively:

$$\int d\mathbf{r} c_1(\mathbf{r}) \nabla \rho(\mathbf{r}) = 0, \quad (7)$$

$$\int d\mathbf{r} d\mathbf{r}' c_2(\mathbf{r}, \mathbf{r}') \nabla \rho(\mathbf{r}) \nabla' \rho(\mathbf{r}') = - \int d\mathbf{r} \rho(\mathbf{r}) \nabla \nabla c_1(\mathbf{r}), \quad (8)$$

where we have integrated by parts on the right-hand side of (8). The first-order sum rule (7) expresses the vanishing of the global internal force  $\langle \hat{\mathbf{F}}_{\text{int}}^0 \rangle = 0^{22}$ . This can be seen by integrating by parts, which yields the integrand in the form  $-\rho(\mathbf{r}) \nabla c_1(\mathbf{r})$ , which is the internal force density scaled by  $-k_B T$ . In formal analogy to the probabilistic variance in Eq. (4), the second-order sum rule (8) could be viewed as relating the “direct variance” of the density gradient (left hand side) to the mean gradient of the internal one-body force field in units of  $k_B T$  (right-hand side), which, equivalently, is the Hessian of the local intrinsic chemical potential  $-k_B T c_1(\mathbf{r})$ , see Fig. 1(b).

As a conceptual point concerning the derivations of Eqs. (7) and (8), we point out that the excess free energy density functional  $F_{\text{exc}}[\rho]$  is an intrinsic quantity, which does not explicitly depend on the external potential  $V_{\text{ext}}(\mathbf{r})$ . Hence there is no need to explicitly take into account a corresponding shift of  $V_{\text{ext}}(\mathbf{r})$ . This is true despite the fact that in an equilibrium situation one would consider the external potential (and the correspondingly generated external force field) as the physical reason for the (inhomogeneous) density profile to be stable. Both one-body fields are connected via the (Euler-Lagrange) minimization equation of density functional theory<sup>26,28,29</sup>:  $k_B T \ln \rho(\mathbf{r}) = k_B T c_1(\mathbf{r}) - V_{\text{ext}}(\mathbf{r}) + \mu$ , where we have set the thermal de Broglie wavelength to unity. For given density profile, we can hence trivially obtain the corresponding external potential as  $V_{\text{ext}}(\mathbf{r}) = -k_B T \ln \rho(\mathbf{r}) + k_B T c_1(\mathbf{r}) + \mu$ , which makes the fundamental Mermin-Evans<sup>26–29</sup> map  $\rho(\mathbf{r}) \rightarrow V_{\text{ext}}(\mathbf{r})$  explicit.

As a consistency check, the second-order sum rules (4) and (8) can alternatively be derived from the hyper virial theorem<sup>30,31</sup> or from spatially resolved correlation identities<sup>22,25</sup>. Following the latter route, one starts with  $\int d\mathbf{r}' H_2(\mathbf{r}, \mathbf{r}') \nabla' V_{\text{ext}}(\mathbf{r}') = -k_B T \nabla \rho(\mathbf{r})$  and  $\int d\mathbf{r}' c_2(\mathbf{r}, \mathbf{r}') \nabla' \rho(\mathbf{r}') = \nabla c_1(\mathbf{r})$ , respectively. The derivation then requires the choice of a suitable field as a multiplier ( $\nabla V_{\text{ext}}(\mathbf{r})$  and  $\nabla \rho(\mathbf{r})$ , respectively), spatial integration over the free position variable, and subsequent integration by parts. However, this strategy i) requires the correct choice for multiplication to be made, and ii) it does not allow to identify the Noether invariance as the underlying reason for the validity. In contrast, the Noether route is constructive and it allows to trace spatial invariance as the fundamental physical reason for the respective identity to hold.

**Thermal diffusion force variance.** Similar to the treatment of the excess free energy functional, one can shift and expand the ideal free energy functional  $F_{\text{id}}[\rho] = k_B T \int d\mathbf{r} \rho(\mathbf{r}) (\ln \rho(\mathbf{r}) - 1)$ . Exploiting the translational invariance at first order leads to vanishing of the total diffusive force:  $-k_B T \int d\mathbf{r} \nabla \rho(\mathbf{r}) = 0$ , and at second order:  $\int d\mathbf{r} \rho(\mathbf{r})^{-1} (\nabla \rho(\mathbf{r})) \nabla \rho(\mathbf{r}) = - \int d\mathbf{r} \rho(\mathbf{r}) \nabla \nabla \ln \rho(\mathbf{r})$ . These ideal identities can be straightforwardly verified via integration by parts (boundary contributions vanish) and they complement the excess results (7) and (8).

**Outlook.** While we have restricted ourselves throughout to translations in equilibrium, the variance considerations apply analogously for rotational invariance<sup>22</sup> and to the dynamics, where invariance of the power functional forms the basis<sup>22,27</sup>. In future work it would be highly interesting to explore connections of our results to statistical thermodynamics<sup>13</sup>, to the study of liquids under shear<sup>32</sup>, to the large fluctuation functional<sup>33</sup>, as well as to recent progress in systematically incorporating two-body correlations into classical density functional theory<sup>34,35</sup>. Investigating the implications of our variance results for Levy-noise<sup>36</sup> is interesting. As the displacement vector  $\boldsymbol{\epsilon}$  is arbitrary both in its orientation and its magnitude our reasoning does not stop at second order in the Taylor expansion, see Eqs. (2) and (6). Assuming that the power series exists, the invariance against the displacement rather implies that each order vanishes individually, which gives rise to a hierarchy of correlation identities of third, fourth, etc. moments that are interrelated with third, fourth, etc. derivatives of the external potential (when starting from  $\Omega[V_{\text{ext}}]$ ) or the one-body direct correlation function (when starting from the excess free energy density functional  $F_{\text{exc}}[\rho]$ ).

Future use of the sum rules can be manifold, ranging from the construction and testing of new theories, such as approximate free energy functionals within the classical density functional framework, to validation of simulation data (to ascertain both correct implementation and sufficient equilibration and sampling) and numerical theoretical results. To give a concrete example, in systems like the confined hard sphere liquid considered by Tschopp et al.<sup>24</sup> on the basis of fundamental measure theory, one could apply and test the sum rule (5) explicitly, as the inhomogeneous total pair correlation function  $h(\mathbf{r}, \mathbf{r}')$  is directly accessible in the therein proposed force-DFT approach.

#### Data availability

Data sharing is not applicable to this study as no datasets were generated or analyzed during the current study.

Received: 22 April 2022; Accepted: 17 October 2022;

Published online: 07 November 2022

#### References

1. E. Noether, *Invariante Variationsprobleme*, [https://gdz.sub.uni-goettingen.de/download/pdf/PPN252457811\\_1918/LOG\\_0022.pdf](https://gdz.sub.uni-goettingen.de/download/pdf/PPN252457811_1918/LOG_0022.pdf) Nachr. d. Königl. Gesellsch. d. Wiss. zu Göttingen, Math.-Phys. Klasse, 235 (1918). English translation by M. A. Tavel: *Invariant variation problems*, Transp. Theo. Stat. Phys. 1, 186 (1971); for a version in modern typesetting see: Frank Y. Wang, <http://arxiv.org/abs/physics/0503066v3> (2018).
2. Byers, N. E. *Noether's Discovery of the Deep Connection Between Symmetries and Conservation Laws*, <https://arxiv.org/abs/physics/9807044> (1998).
3. Turci, F. & Wilding, N. B. Phase separation and many-body effects in three-dimensional active Brownian particles. *Phys. Rev. Lett.* 126, 038002 (2021).
4. Turci, F. & Wilding, N. B. Wetting transition of active Brownian particles on a thin membrane. *Phys. Rev. Lett.* 127, 238002 (2021).



5. Coe, M. K., Evans, R. & Wilding, N. B. Density depletion and enhanced fluctuations in water near hydrophobic solutes: identifying the underlying physics. *Phys. Rev. Lett.* **128**, 045501 (2022).
6. Thorneywork, A. L. et al. Structure factors in a two-dimensional binary colloidal hard sphere system. *Mol. Phys.* **116**, 3245 (2018).
7. Höfling, F. & Dietrich, S. Enhanced wavelength-dependent surface tension of liquid-vapour interfaces. *Europhys. Lett.* **109**, 46002 (2015).
8. Parry, A. O., Rascón, C. & Evans, R. The local structure factor near an interface; beyond extended capillary-wave models. *J. Phys.: Condens. Matter* **28**, 244013 (2016).
9. Hurtado, P. I., Pérez-Espigares, C., del Pozo, J. J. & Garrido, P. L. Symmetries in fluctuations far from equilibrium. *Proc. Natl Acad. Sci.* **108**, 7704 (2011).
10. Lacoste, D. & Gaspard, P. Isometric fluctuation relations for equilibrium states with broken symmetry. *Phys. Rev. Lett.* **113**, 240602 (2014).
11. Lacoste, D. & Gaspard, P. Fluctuation relations for equilibrium states with broken discrete or continuous symmetries. *J. Stat. Mech.* **2015**, P11018 (2015).
12. Dechant, A. & Sasa, S. Fluctuation-response inequality out of equilibrium. *Proc. Natl Acad. Sci.* **117**, 6430 (2020).
13. Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **75**, 126001 (2012).
14. Lezcano, A. G. & de Oca, A. C. M. A stochastic version of the Noether theorem. *Found. Phys.* **48**, 726 (2018).
15. Baez, J. C. & Fong, B. A Noether theorem for Markov processes. *J. Math. Phys.* **54**, 013301 (2013).
16. Marvian, I. & Spekkens, R. W. Extending Noether's theorem by quantifying the asymmetry of quantum states. *Nat. Commun.* **5**, 3821 (2014).
17. Sasa, S. & Yokokura, Y. Thermodynamic entropy as a Noether invariant. *Phys. Rev. Lett.* **116**, 140601 (2016).
18. Minami, Y. & Sasa, S. Thermodynamic entropy as a Noether invariant in a Langevin equation. *J. Stat. Mech.* **2020**, 013213 (2020).
19. Sasa, S., Sugiura, S. & Yokokura, Y. Thermodynamical path integral and emergent symmetry. *Phys. Rev. E* **99**, 022109 (2019).
20. Revzen, M. Functional integrals in statistical physics. *Am. J. Phys.* **38**, 611 (1970).
21. Baez, J. C. *Getting to the Bottom of Noether's Theorem*, <https://arxiv.org/abs/2006.14741> (2022).
22. Hermann, S. & Schmidt, M. Noether's theorem in statistical mechanics. *Commun. Phys.* **4**, 176 (2021).
23. Hermann, S. & Schmidt, M. Why Noether's theorem applies to statistical mechanics. *J. Phys.: Condens. Matter* **34**, 213001 (2022). (invited Topical Review).
24. Tschopp, S. M., Sammüller, F., Hermann, S., Schmidt, M. & Brader, J. M. Force density functional theory for fluids in- and out-of-equilibrium. *Phys. Rev. E* **106**, 014115 (2022).
25. Baus, M. Broken symmetry and invariance properties of classical fluids. *Mol. Phys.* **51**, 211 (1984).
26. Hansen, J.-P. & McDonald, I. R. *Theory of Simple Liquids* 4th edn (Academic Press, 2013).
27. Schmidt, M. Power functional theory for many-body dynamics. *Rev. Mod. Phys.* **94**, 015007 (2022).
28. Evans, R. The nature of the liquid-vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids. *Adv. Phys.* **28**, 143 (1979).
29. Evans, R. In *Fundamentals of Inhomogeneous Fluids* (ed Henderson, D.) (Dekker, 1992).
30. Hirschfelder, J. O. Classical and quantum mechanical hypervirial theorems. *J. Chem. Phys.* **33**, 1462 (1960).
31. Haile, J. M. *Molecular Dynamics Simulation: Elementary Methods* (Wiley, 1992).
32. Asheichyk, K., Fuchs, M. & Krüger, M. Brownian systems perturbed by mild shear: comparing response relations. *J. Phys.: Condens. Matter* **33**, 405101 (2021).
33. Jack, R. L. & Sollich, P. Effective interactions and large deviations in stochastic processes. *Eur. Phys. J. Spec. Top.* **224**, 2351 (2015).
34. Tschopp, S. M., Vuijk, H. D., Sharma, A. & Brader, J. M. Mean-field theory of inhomogeneous fluids. *Phys. Rev. E* **102**, 042140 (2020).
35. Tschopp, S. M. & Brader, J. M. Fundamental measure theory of inhomogeneous two-body correlation functions. *Phys. Rev. E* **103**, 042103 (2021).
36. Yuvan, S. & Bier, M. Accumulation of particles and formation of a dissipative structure in a nonequilibrium bath. *Entropy* **24**, 189 (2022).

#### Acknowledgements

This work is supported by the German Research Foundation (DFG) via project number 436306241. We thank Daniel de las Heras, Thomas Fischer, and Gerhard Jung for useful discussions.

#### Author contributions

S.H. and M.S. have jointly carried out the work and written the paper.

#### Funding

Open Access funding enabled and organized by Projekt DEAL.

#### Competing interests

The authors declare no competing interests.

#### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s42005-022-01046-3>.

**Correspondence** and requests for materials should be addressed to Sophie Hermann or Matthias Schmidt.

**Peer review information** *Communications Physics* thanks Matthias Krueger and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022



## Active interface polarization as a state function

Sophie Hermann<sup>✉</sup> and Matthias Schmidt<sup>✉\*</sup>*Theoretische Physik II, Physikalisches Institut, Universität Bayreuth, D-95447 Bayreuth, Germany*

(Received 25 January 2020; accepted 10 March 2020; published 3 April 2020)

We prove three exact sum rules that relate the polarization of active Brownian particles to their one-body current: (i) The total polarization vanishes, provided that there is no net flux through the boundaries, (ii) at any planar wall the polarization is determined by the magnitude of the bulk current, and (iii) the total interface polarization between phase-separated fluid states is rigorously determined by the gas-liquid current difference. This result precludes the influence of the total interface polarization on active bulk coexistence and questions the proposed coupling of interface to bulk.

DOI: [10.1103/PhysRevResearch.2.022003](https://doi.org/10.1103/PhysRevResearch.2.022003)

Systems of active Brownian particles (ABPs) consist of thermally diffusing spheres that self-propel along an intrinsic direction, which itself undergoes free rotational diffusion [1–3]. ABPs form the prototypical statistical model for active matter [1–3]. In order to characterize the local orientational order, the polarization  $\mathbf{M}$  is a measure of the strength and direction of the local preferred alignment of the particle orientations. A multitude of relevant situations have been reported in the literature where ABPs display spontaneous polarization effects [4–13]. In many of these cases, the spontaneous polarization occurs in the absence of any explicit torques that act on the particles: No external torques occur when all external fields depend and act on position only, and no internal torques arise when the particles are spheres. In equilibrium systems of spheres, the absence of torques implies local isotropy, and hence the emergence of nonzero local polarization is a genuine effect of nonequilibrium, as characterized by a nonzero spatially and orientationally resolved local one-body current  $\mathbf{J}$ .

Important examples of these nonequilibrium situations include the spontaneous orientational ordering of ABPs against gravity in the sedimentation profile at large altitudes [4–7], the ordering upon adsorption against a (hard) wall [10–13], and the spontaneous polarization of the free interface between phase-separated active gas and liquid phases [13–17]. There,  $\mathbf{M}$  points toward the active liquid in the case of purely repulsive particles [13–15], but toward the gas in the case of active Lennard-Jones particles [16,17]. A range of different mechanisms and descriptions for the occurrence of the bulk phase separation has been put forward, such as, e.g., kinetic blocking as a feedback mechanism [18,19], the existence of a nonequilibrium chemical potential [13,14], and effective interparticle attraction [20].

The status of the nonequilibrium interface, however, has been claimed to be very different from what is known in equilibrium. Tailleur and coworkers [15,21] find in their approach interface-to-bulk coupling, i.e., the properties of the free interface affect the gas and liquid bulk states, which are in stable nonequilibrium coexistence. Further, one can argue that due to the swim force, any nonvanishing polarization is necessarily associated with a one-body force distribution  $\gamma s \mathbf{M}$ , where  $\gamma$  is the translational friction constant and  $s$  is the speed of free swimming. It is not inconceivable (and consistent with simple interface versus bulk dimensional analysis) that this force density compresses the phase toward which  $\mathbf{M}$  points at the expense of the other phase, and hence that it changes the properties of the coexisting phases.

Here we prove rigorously from first principles that the total interfacial polarization is a straightforward quantitative consequence of differing bulk currents in the coexisting phases and the rotational diffusion current  $D_{\text{rot}}$ . This rules out the total polarization as an underlying physical mechanism for the interface-to-bulk coupling [15,21]. Similarly, the total polarization of particles adsorbed at a wall is solely determined by  $D_{\text{rot}}$  and the current in the corresponding bulk fluid, and thus constitutes a state function. Furthermore, we show that in a system without explicit torques and with no total flux through the boundaries, the global orientational distribution function follows a free diffusion equation, so the global polarization vanishes in steady state; we also address the time-dependent case below. Figure 1 illustrates the three types of orientational ordering phenomena that we address in the following. Our derivation of the corresponding sum rules is based on the exact rotational equation of motion and on the continuity equation.

We describe ABPs on the level of their position- and orientation-resolved microscopic one-body density distribution  $\rho(\mathbf{r}, \boldsymbol{\omega}, t)$ , where  $\mathbf{r}$  indicates position,  $\boldsymbol{\omega}$  (unit vector) orientation, and  $t$  time. Then the local polarization  $\mathbf{M}(\mathbf{r}, t)$  is a vector field defined as the first orientational moment of the density profile,

$$\mathbf{M}(\mathbf{r}, t) = \int d\boldsymbol{\omega} \boldsymbol{\omega} \rho(\mathbf{r}, \boldsymbol{\omega}, t), \quad (1)$$

\*Matthias.Schmidt@uni-bayreuth.de

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

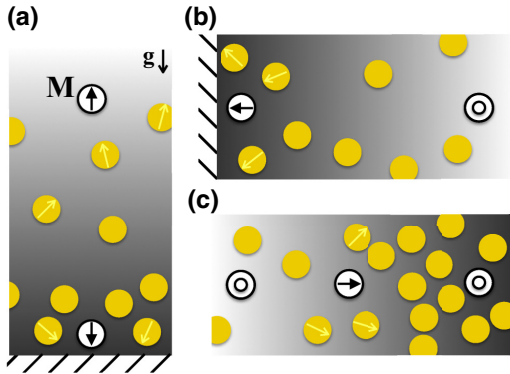


FIG. 1. Schematic of systems in which the sum rules (10), (17), and (18) apply. (a) Sedimentation of ABPs under gravity  $\mathbf{g}$ , which acts toward the lower confining wall (hatched area). (b) Adsorption of active particles at a semi-infinite wall (hatched area). (c) Motility-induced phase separation of purely repulsive interacting ABPs. The yellow circles indicate active Brownian particles (ABPs) and the corresponding yellow arrows show exemplary particle orientations. The local particle polarization  $\mathbf{M}$  (white disks) can either vanish, indicated by a small circle, or point in a specific direction, indicated by a black arrow. The color gradient displays the density modulation from high values (dark) to low values (bright).

where the integral is over all orientations  $\omega$ . The translational one-body current  $\mathbf{J}(\mathbf{r}, \omega, t)$  is the (microscopically resolved) measure of the direction and magnitude of the local flow of particles. As there are no explicit torques, the rotational motion is purely diffusive. Thus, the (in general) inhomogeneous density distribution  $\rho$  generates a nonzero rotational current

$$\mathbf{J}^\omega(\mathbf{r}, \omega, t) = -D_{\text{rot}} \nabla^\omega \rho(\mathbf{r}, \omega, t), \quad (2)$$

where  $D_{\text{rot}}$  is the rotational diffusion constant and  $\nabla^\omega$  indicates the derivative with respect to orientation  $\omega$ . As the dynamics evolve the microstates continuously in time and the total particle number  $N$  remains constant, the one-body distributions satisfy the continuity equation,

$$\dot{\rho}(\mathbf{r}, \omega, t) = -\nabla \cdot \mathbf{J}(\mathbf{r}, \omega, t) - \nabla^\omega \cdot \mathbf{J}^\omega(\mathbf{r}, \omega, t), \quad (3)$$

where  $\dot{\rho} = \partial\rho/\partial t$  with  $\dot{\rho} = 0$  in steady state and  $\nabla$  indicates the derivative with respect to position  $\mathbf{r}$ . Note that the continuity equation (3) holds rigorously, independent of the presence and the type of interparticle interactions, particle-wall interactions, and external forces. The forces influence the translational and rotational motion, but not the form of (3). All occurring terms in (3) can be sampled in computer simulations; see, e.g., Ref. [22].

We first consider the total polarization for systems with vanishing total flux through the boundaries of volume  $V$  at all times  $t$ , i.e.,  $\int_{\partial V} d\mathbf{s} \cdot \mathbf{J}(\mathbf{r}, \omega, t) = 0$ , where  $d\mathbf{s}$  denotes the vectorial surface element and  $\partial V$  indicates the surface of volume  $V$ . Here  $V$  is arbitrary and can be chosen to be either the system volume, an enclosing larger volume that contains the system, or a subvolume of the system. The number of particles inside  $V$  is  $N = \int_V d\mathbf{r} \int d\omega \rho(\mathbf{r}, \omega, t)$ .

We rewrite the spatially integrated density distribution as  $\int_V d\mathbf{r} \rho(\mathbf{r}, \omega, t) = Nf(\omega, t)$ ; this defines the global orientational distribution function  $f(\omega, t)$ , which is normalized at all times  $t$ ,  $\int d\omega f(\omega, t) = 1$ . Building the time derivative of the spatially integrated density distribution  $\rho$  leads to

$$\begin{aligned} N\dot{f}(\omega, t) &= \int_V d\mathbf{r} \dot{\rho}(\mathbf{r}, \omega, t) \\ &= - \int_V d\mathbf{r} (\nabla \cdot \mathbf{J}(\mathbf{r}, \omega, t) + \nabla^\omega \cdot \mathbf{J}^\omega(\mathbf{r}, \omega, t)), \end{aligned} \quad (4)$$

where we used the continuity equation (3) to obtain (5). Assuming the absence of explicit torques and hence a free rotational diffusion current (2), applying the divergence theorem to the translational current contribution in (5) yields

$$\begin{aligned} N\dot{f}(\omega, t) &= - \int_{\partial V} d\mathbf{s} \cdot \mathbf{J}(\mathbf{r}, \omega, t) + \int_V d\mathbf{r} D_{\text{rot}} \Delta^\omega \rho(\mathbf{r}, \omega, t) \\ &= D_{\text{rot}} N \Delta^\omega f(\omega, t), \end{aligned} \quad (6)$$

where the orientational Laplace operator is  $\Delta^\omega = \nabla^\omega \cdot \nabla^\omega$ . The first term on the right-hand side of Eq. (6) vanishes due to the vanishing flux boundary condition and the second term can be rewritten as (7) using the definition of  $f$ . Dividing Eq. (7) by the particle number  $N$  yields a free diffusion equation for the orientational distribution function

$$\dot{f}(\omega, t) = D_{\text{rot}} \Delta^\omega f(\omega, t). \quad (8)$$

Note that Milster *et al.* [23] derived an equation similar to Eq. (8) for the orientational distribution function of two-dimensional ABPs with negligible translational diffusion. We consider the system to be in steady state,  $\dot{\rho}(\mathbf{r}, \omega, t) = 0$ , and thus also  $\dot{f}(\omega, t) = 0$ , which simplifies the diffusion equation (8) to

$$\Delta^\omega f(\omega) = 0. \quad (9)$$

The only solutions in two and three dimensions (2D and 3D) of Eq. (9) are constants,  $f = (2\pi)^{-1}$  for two-dimensional systems and  $f = (4\pi)^{-1}$  in three dimensions. Hence, we conclude [24] that the global orientational distribution function  $f$  is independent of  $\omega$  and the total polarization  $\mathbf{M}_{\text{tot}}$  vanishes,

$$\mathbf{M}_{\text{tot}} = \int_V d\mathbf{r} \mathbf{M}(\mathbf{r}) = 0. \quad (10)$$

In practice, the result (10) can be used as a consistency check in computer simulations and in theoretical descriptions. It is trivially satisfied in equilibrium systems without explicit torques, as such systems imply local isotropy and hence local and total polarization both vanish.

We emphasize that Eq. (10) holds in all steady states, independent of the existence of external potentials or the present type of interparticle or particle-wall interactions, in each (sub)volume  $V$  with zero net flux through its boundaries. The local and hence also the total fluxes through the surface of the considered volume are zero if the orientational distribution function is homogeneous at the surface and the current can be expressed as  $\mathbf{J}(\mathbf{r}, \omega) = J_b \omega$ . The magnitude  $J_b$  is equal to the first Fourier coefficient of the current. The condition for the current is satisfied, e.g., in isotropic bulk states or in regions of vanishing current.

As a first application of Eq. (10), we consider sedimentation of ABPs [5,25]. Figure 1(a) illustrates two prominent effects that occur: (i) The particle orientation points toward the lower confining wall [4–7]. This leads to a particle accumulation at the wall on top of the effect of gravity and may be interpreted as a self-trapping mechanism. (ii) The sedimentation length increases as compared to passive particles due to the alignment of the swimmers against gravity  $\mathbf{g}$  at large distances from the wall [4–7]. Both effects can be interpreted as originating from a dynamical balance between the spatial self-sorting of the active particles and the counteracting mechanism of rotational diffusion. An upward oriented particle, for example, swims on average toward higher altitudes until its orientation changes via particle rotation.

Each of the above phenomena (i) and (ii) generates nonvanishing local polarization [4–7] [cf. Fig. 1(a)] and both have at first sight no relationship with each other. But as the flux through the boundaries is zero [26], the total polarization has to vanish in steady state; cf. (10). Thus, if the volume  $V$  is divided into bottom and top subvolumes, both partial polarizations have to cancel each other, independent of the division itself. This effect is nonlocal as the accumulation and polarization at the bottom determines the overall particle orientation in the remaining volume.

We next consider nonvanishing total flux through the boundaries. We derive a spatially resolved (“local”) version of the sum rule for the ubiquitous two-dimensional system in steady state. In two dimensions, the orientation vector can be written as  $\boldsymbol{\omega} = (\cos \varphi, \sin \varphi)$ , where  $\varphi$  is the angle measured against the positive  $x$  axis, and the orientational derivative  $\nabla^\omega$  reduces to  $\partial/\partial\varphi$ . We assume, as a relevant case, translational invariance along the  $y$  axis. (Note, however, that this restriction is not necessary [27]). Because of the assumption of translational invariance, the density  $\rho(x, \varphi)$  and  $x$  component of the current  $J^x(x, \varphi)$  are even in the angle  $\varphi$  as both are invariant under reflection at  $x$  axis,  $y \rightarrow -y$  and  $\varphi \rightarrow -\varphi$  [14]. So the angular Fourier decomposition of both quantities consists only of cosines. The density thus may be expressed as

$$\rho(x, \varphi) = \sum_{n=0}^{\infty} \rho_n(x) \cos(n\varphi), \quad (11)$$

where  $\rho_n(x)$  indicates the  $n$ th Fourier coefficient of the density profile. Using (11) in the expression for the polarization (1) yields  $\mathbf{M} = (\pi \rho_1(x), 0)$ . The  $y$ -component  $M_y$  vanishes due to the symmetry of the density distribution (11), so the magnitude of the polarization is equal its  $x$  component,  $M = M_x$ . The  $x$  component of the current can be Fourier decomposed similarly as

$$J^x(x, \varphi) = \sum_{n=0}^{\infty} J_n^x(x) \cos(n\varphi), \quad (12)$$

where  $J_n^x(x)$  denotes the  $n$ th Fourier coefficient and thus the  $n$ th orientational moment of the current. As the rotational current consists only of the thermal free diffusion contribution (2), the continuity equation (3) simplifies for steady states to

$$\frac{\partial J^x(x, \varphi)}{\partial x} = D_{\text{rot}} \frac{\partial^2 \rho(x, \varphi)}{\partial \varphi^2}. \quad (13)$$

Equation (13) is satisfied, e.g., for the case of motility-induced phase separation [14]. Insertion of the Fourier decomposition (12) in Eq. (13) and integrating twice in the angle  $\varphi$  allows us to solve the equation for density  $\rho$ . Evaluation of both indefinite integrals leads to

$$\rho(x, \varphi) = -\frac{1}{D_{\text{rot}}} \sum_{n=1}^{\infty} \frac{\partial J_n^x(x)}{\partial x} \frac{\cos(n\varphi)}{n^2} + \rho_0, \quad (14)$$

where we have used the Fourier expansion of the current (12) and the integration constant  $\rho_0$  indicates the average density, i.e., the total number of particles per system volume and per radians. The integration constant of the first integral vanishes, since a linear  $\varphi$  term does not satisfy the  $2\pi$  periodicity in angle  $\varphi$ .

The polarization profile  $M(x)$  (1) can be simplified as  $M(x) = \int_0^{2\pi} d\varphi \rho(x, \varphi) \cos \varphi$  using the present symmetries. Inserting the expansion of  $\rho$  (14) and evaluating the integral over all orientations yields

$$M(x) = -\frac{\pi}{D_{\text{rot}}} \frac{\partial J_1^x(x)}{\partial x}. \quad (15)$$

That is, for each position  $x$  the local polarization is proportional to the spatial change in the first moment of the current. The spatially resolved relation (15) constitutes a local sum rule, similarly determined by Refs. [13,14] in the special case of ABPs. The derivation here is more general and based only on the continuity equation with freely diffusive rotational motion.

In order to derive a global sum rule, we spatially integrate the exact local sum rule (15),

$$M_{\text{tot}} = \int_{x_1}^{x_2} dx \int_{y_1}^{y_2} dy M(x) = \frac{\pi L_y}{D_{\text{rot}}} [J_1^x(x_1) - J_1^x(x_2)], \quad (16)$$

which determines the total polarization  $M_{\text{tot}}$  in the integration volume  $V$ . For simplicity, we restrict ourself to rectangular areas  $V$  aligned with the coordinate axes. The integration limits are set to the arbitrary positions  $x_1$  and  $x_2$  for the  $x$  coordinate and  $y_1$  and  $y_2$  for the  $y$  coordinate. Because of the translational invariance, the  $y$  integral can be explicitly evaluated and gives the length of  $y$  integration,  $L_y = y_2 - y_1$ . In the following, we thus consider the total polarization per unit length in the  $y$  direction,  $M_{\text{tot}}/L_y$ .

Equation (16) holds for ABPs in a large variety of situations. We address two general relevant cases in the following. First, we consider ABPs absorbed at a (hard or soft) planar wall [see Fig. 1(b)]. We set a wall parallel to the  $y$  axis at  $x = 0$ . As the density vanishes inside the wall, the one-body current  $\mathbf{J}(x) = 0$  for  $x \rightarrow -\infty$ . For  $x \rightarrow \infty$ , the semi-infinite system approaches an isotropic bulk fluid, so the current is  $\mathbf{J}(x) = J_b \boldsymbol{\omega}$ , due to symmetry. The (constant) magnitude of the bulk current,  $J_b$ , equals the first Fourier component,  $J_b = J_1^x$ . Setting the limits of integration in Eq. (16) to  $x_1 \rightarrow -\infty$  and  $x_2 \rightarrow \infty$  and using the known expressions for the currents simplifies the total polarization at the wall per unit  $y$  length to

$$\frac{M_{\text{tot}}}{L_y} = -\frac{\pi}{D_{\text{rot}}} J_b. \quad (17)$$

Hence, the absolute value of  $\mathbf{M}_{\text{tot}}$  is solely determined by the bulk current and the rotational diffusion constant. Recall that  $\mathbf{M}_{\text{tot}}$  is oriented along the  $x$  axis due to the translational

symmetry, so the sign of the right-hand side of Eq. (17) determines whether the total polarization points toward or against the wall. As the free swim speed  $s \geq 0$ ,  $J_b$  is greater or equal to zero, and hence the total polarization points towards the wall. (Note that interparticle interactions only tend to reduce the absolute value of the bulk current due to drag effects.) Because of the global sum rule (17), the sign of the total  $x$  polarization per unit length is negative. A vanishing bulk current  $J_b = 0$  constitutes a special case, which leads to a vanishing total polarization as one would expect to occur for a system of passive spheres. We conclude that the total swim force density  $\int d\mathbf{r} d\omega s \gamma \rho \omega$  always points toward the wall, so that the total polarization also points to the wall (except if the total polarization vanishes). The direction of the total polarization is hence independent of both the particle-wall and the interparticle interaction. Furthermore, due to locality of both interactions the bulk itself, in particular the bulk current  $J_b$ , is independent of the wall. Thus, the total polarization  $\mathbf{M}_{\text{tot}}$  only depends on bulk quantities via Eq. (17) and constitutes a state function. This extends the work of Tailleur and coworkers [28,29], who investigated whether pressure is a state function in active fluids. Note that the magnitude and structure of the local polarization profile  $M(x)$  may depend on both the wall-particle and the interparticle interaction potentials.

As a second relevant example, we consider the phase separation of ABPs [schematic sketch in Fig. 1(c)]. The particles phase separate in a dense (liquid) and a dilute (gas) bulk fluid. Since both coexisting bulk states are isotropic, the coexisting bulk currents are proportional to the orientation  $\omega$  and the corresponding magnitudes are  $J_g$  in the gas and  $J_l$  in the liquid bulk phase. Using those relations for the bulk current and setting the limits of integration inside an isotropic bulk phase, i.e.,  $x_1 \rightarrow -\infty$  and  $x_2 \rightarrow \infty$ , simplifies Eq. (16) to

$$\frac{M_{\text{tot}}}{L_y} = \frac{\pi}{D_{\text{rot}}}(J_g - J_l). \quad (18)$$

Hence the difference between both local bulk currents, scaled with the rotational diffusion constant, determines the total polarization per transversal length. Equation (18) constitutes an exact global sum rule. For particles interacting via the Weeks-Chandler-Anderson potential, which is a Lennard-Jones potential cut and shifted at its minimum to be purely repulsive, the swimmers align toward the denser phase in the interfacial region [13,14]; cf. Fig. 1(c). Hence, the total polarization is also directed toward the dense phase and it is positive. According to Eq. (18), one expects a higher current in the dilute phase in comparison to the dense phase, which is in qualitative and quantitative agreement with simulation data [14,30]. In contrast, for active Lennard-Jones particles, the total polarization was found to point toward the dilute phase [17]. A sketch of the system would be similar to Fig. 1(c), but with an inverted polarization arrow. Using the total polarization to calculate the difference between both bulk currents from the global sum rule (18), we predict a higher current in the liquid than in the gas. Note that the particle polarization is primarily located at the interface, since the polarization in bulk vanishes due to isotropy; cf. Eq. (16).

A physical interpretation of the global sum rule (18) is that the interfacial quantity  $M_{\text{tot}}$  is solely determined by the bulk values  $J_g$  and  $J_l$ . In other words, the interfacial polarization is

a mere consequence of the properties of the bulk states. This interpretation follows from the locality of the short-ranged interparticle interactions, which is a similar reasoning as in the case of particles in front of a semi-infinite wall [cf. Fig. 1(b)]. The combination of the expression (18) and the locality of interparticle interactions lets the nonlocal influence of  $\mathbf{M}_{\text{tot}}$  on the entire bulk seem implausible. This questions the conclusion of Solon *et al.* [15, p. 16] that “the phase coexistence densities [...] is controlled by the polar ordering of particles at the gas-liquid interface.” It seems more reasonable that the interface is a consequence of the bulk and not vice versa, especially since no mechanism has been identified which would generate these nonlocal effects. Furthermore, our interpretation is in agreement with the theory of Ref. [14] where no interfacial contributions are required to describe the bulk and the gas-liquid coexistence, as is the case in equilibrium.

We next generalize the steady-state relationship (10) and consider the time dependence of  $\mathbf{M}_{\text{tot}}$ . We restrict ourselves to cases of vanishing total flux through the boundaries of the considered volume  $V$  at all times. The time-dependent total polarization  $\mathbf{M}_{\text{tot}}(t)$  is then given as

$$\mathbf{M}_{\text{tot}}(t) = \int_V d\mathbf{r} \mathbf{M}(\mathbf{r}, t) = N \int d\omega \omega f(\omega, t), \quad (19)$$

and thus can be determined via the global orientational distribution function  $f(\omega, t)$ . We first consider two-dimensional systems. Hence, as above, the orientation vector is  $\omega = (\cos \varphi, \sin \varphi)$ , where  $\varphi$  is an angular coordinate and  $\Delta^\omega$  simplifies to  $\partial^2 / \partial \varphi^2$ . So,  $f(\omega, t)$  is given as the solution of Eq. (8),

$$f(\varphi, t) = \sum_{n=0}^{\infty} [a_n \cos(n\varphi) + b_n \sin(n\varphi)] e^{-n^2 D_{\text{rot}} t}, \quad (20)$$

where the constants  $a_n, b_n$  are determined by the initial conditions. Inserting the global orientational distribution function (20) into Eq. (19) and carrying out the angular integral yields the temporal behavior of the total polarization as an exponential decay,

$$\mathbf{M}_{\text{tot}}(t) = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} e^{-D_{\text{rot}} t}, \quad (21)$$

where  $1/D_{\text{rot}}$  is the time constant and the vector  $(a_1, b_1)$  is the initial polarization at time  $t = 0$ .

In three spatial dimensions, we parametrize  $\omega = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$ , where  $\theta$  and  $\varphi$  indicate polar and azimuthal angles. Then Eq. (8) is solved by

$$f(\theta, \varphi, t) = \sum_{l=0}^{\infty} \sum_{m=-l}^l a_{lm} Y_{lm}(\theta, \varphi) e^{-\frac{D_{\text{rot}}}{l(l+1)} t}, \quad (22)$$

where the constants  $a_{lm}$  are again set by initial conditions and  $Y_{lm}(\theta, \varphi)$  indicate the spherical harmonics. Insertion of (22) into Eq. (19) gives

$$\mathbf{M}_{\text{tot}}(t) = \sum_{l=0}^{\infty} \mathbf{M}_l e^{-\frac{D_{\text{rot}}}{l(l+1)} t}, \quad (23)$$

where we have defined the constants  $\mathbf{M}_l = \int d\omega \sum_m a_{lm} Y_{lm}(\theta, \varphi) \omega$ . Hence, in both the 2D and 3D cases,  $\mathbf{M}_{\text{tot}}$  decays exponentially in time. The dynamics depend only on the rotational diffusion constant and on the initial conditions. Clearly

in the limit  $t \rightarrow \infty$ , the results for the time dependence of  $\mathbf{M}_{\text{tot}}$ , (21) and (23), reduce to the steady-state sum rule (10) of vanishing total polarization.

Furthermore, one can extend the obtained sum rules to higher (e.g., nematic) order moments, e.g.,  $M_n(\mathbf{r}) = \int_0^{2\pi} d\varphi \cos(n\varphi) \rho(\mathbf{r}, \varphi)$  in two-dimensional systems. For a vanishing flux through the surface of the volume, those higher moments in the considered volume are all equal to zero in steady state, as is the polarization [cf. Eq. (10)], and their time evolution can be derived from Eqs. (20) and (22). In translationally invariant two-dimensional systems the sum rules are similar to Eqs. (15) and (16), where the  $n$ th moment  $M_n$  corresponds to the spatial derivative of the  $n$ th moment of the current  $J_n^x$ . Since higher moments of the bulk current  $J_{n>1}^x$  vanish in bulk due to symmetry, the total higher order moments  $M_{n>1}^{\text{tot}}$  are also zero for particle adsorption at a wall or motility induced phase separation. Hence, these total moments cannot contribute to determine the bulk densities. Note, however, that the local structure of these higher order moments is nontrivial in general.

To conclude, we have demonstrated that polarization and current distribution of ABPs are intimately connected. Using the continuity equation, together with the properties of free

rotational diffusion, we have derived three exact global sum rules (10), (17), and (18). These imply, respectively, (i) that the total system polarization vanishes, (ii) that the polarization at a wall is determined by the bulk current and hence represents a state function, and (iii) that for phase-separated fluid states the polarization of the free interface is given by the difference of bulk current in the coexisting active bulk phases. Note that Eq. (18) is indeed satisfied qualitatively and quantitatively in the theory of Refs. [14,31]. These global sum rules, as well as the local sum rule (15), can be useful as consistency checks for simulations and theories and can also be used as an input for theoretical descriptions. One could apply the derived local and global sum rules to further interesting systems: The relations hold in case of spatial inhomogeneous activity  $s(\mathbf{r})$  as considered by Sharma *et al.* [32] and Hasnain *et al.* [33] or for spatially varying translational diffusion [33], as long as the rotational diffusion coefficient is kept constant. It would be interesting to explore in future work the connections of our treatment to the results presented in Refs. [34,35].

We thank D. de las Heras for stimulating discussions and critical reading of the manuscript.

- 
- [1] M. C. Marchetti, J. F. Joanny, S. Ramaswamy, T. B. Liverpool, J. Prost, M. Rao, and R. A. Simha, *Rev. Mod. Phys.* **85**, 1143 (2013).
- [2] P. Romanczuk, M. Bär, W. Ebeling, B. Lindner, and L. Schimansky-Geier, *Eur. Phys. J. Spec. Top.* **202**, 1 (2012).
- [3] C. Bechinger, R. Di Leonardo, H. Löwen, C. Reichhardt, G. Volpe, and G. Volpe, *Rev. Mod. Phys.* **88**, 045006 (2016).
- [4] S. Hermann and M. Schmidt, *Soft Matter* **14**, 1614 (2018).
- [5] M. Enculescu and H. Stark, *Phys. Rev. Lett.* **107**, 058301 (2011).
- [6] J. Vachier and M. G. Mazza, *Eur. Phys. J. E* **42**, 11 (2019).
- [7] F. Ginot, A. Solon, Y. Kafri, C. Ybert, J. Tailleur, and C. Cottin-Bizonne, *New J. Phys.* **20**, 115001 (2018).
- [8] M. Hennes, K. Wolff, and H. Stark, *Phys. Rev. Lett.* **112**, 238104 (2014).
- [9] A. Martín-Gómez, D. Levis, A. Díaz-Guilera, and I. Pagonabarraga, *Soft Matter* **14**, 2610 (2018).
- [10] T. Speck and R. L. Jack, *Phys. Rev. E* **93**, 062605 (2016).
- [11] C. G. Wagner, M. F. Hagan, and A. Baskaran, *J. Stat. Mech.* (2017) 043203.
- [12] J. Elgeti and G. Gompper, *Europhys. Lett.* **101**, 48003 (2013).
- [13] S. Paliwal, J. Rodenburg, R. van Roij, and M. Dijkstra, *New J. Phys.* **20**, 015003 (2018).
- [14] S. Hermann, P. Krinninger, D. de las Heras, and M. Schmidt, *Phys. Rev. E* **100**, 052604 (2019).
- [15] A. P. Solon, J. Stenhammar, M. E. Cates, Y. Kafri, and J. Tailleur, *New J. Phys.* **20**, 075001 (2018).
- [16] V. Prymidis, S. Paliwal, M. Dijkstra, and L. Filion, *J. Chem. Phys.* **145**, 124904 (2016).
- [17] S. Paliwal, V. Prymidis, L. Filion, and M. Dijkstra, *J. Chem. Phys.* **147**, 084902 (2017).
- [18] J. Tailleur and M. E. Cates, *Phys. Rev. Lett.* **100**, 218103 (2008).
- [19] D. Richard, H. Löwen, and T. Speck, *Soft Matter* **12**, 5257 (2016).
- [20] T. F. F. Farage, P. Krinninger, and J. M. Brader, *Phys. Rev. E* **91**, 042310 (2015).
- [21] A. P. Solon, J. Stenhammar, M. E. Cates, Y. Kafri, and J. Tailleur, *Phys. Rev. E* **97**, 020602(R) (2018).
- [22] D. de las Heras, J. Renner, and M. Schmidt, *Phys. Rev. E* **99**, 023306 (2019).
- [23] See Appendix A of S. Milster, J. Nötel, I. M. Sokolov, and L. Schimansky-Geier, *Eur. Phys. J. Spec. Top.* **226**, 2039 (2017).
- [24] In order to show that Eq. (10) holds, we calculate the total polarization, which is defined as the spatial integral of the local polarization (1),  $\mathbf{M}_{\text{tot}} = \int_V d\mathbf{r} \mathbf{M}(\mathbf{r}) = \int d\omega \omega \int_V d\mathbf{r} \rho(\mathbf{r}, \omega)$ , where the spatial and orientational integral were interchanged. The spatially integrated density can be expressed as  $Nf$ , which is constant for two- and three-dimensional systems in case of steady state and a vanishing total flux through the boundaries, cf. Eq. (9). Thus, the total polarization simplifies to  $\mathbf{M}_{\text{tot}} = Nf \int d\omega \omega = 0$ .
- [25] A. P. Solon and M. E. Cates, *Eur. Phys. J. Spec. Top.* **224**, 1231 (2015).
- [26] The density and hence the current through the top and the bottom of the system vanish due to gravity and in the lower confining wall, respectively. The flux through the left system sides cancels with those through the right due to translational invariance. Hence, the total flux through the surface of the system is zero.
- [27] If one discards the translational symmetry along the  $y$  axis, odd and even functions contribute to the density and the current. The Fourier decompositions of the density

- is then  $\rho(x, y, \varphi) = \sum_{n=0}^{\infty} \rho_n^c(x, y) \cos(n\varphi) + \rho_n^s(x, y) \sin(n\varphi)$ , where  $\rho_n^c$  indicates the  $n$ th Fourier coefficient of the cosine contributions and  $\rho_n^s$  denote the coefficients of the sine contributions. Similarly the Fourier expansion of the current is  $\mathbf{J}(x, y, \varphi) = \sum_{n=0}^{\infty} \mathbf{J}_n^c(x, y) \cos(n\varphi) + \mathbf{J}_n^s(x, y) \sin(n\varphi)$ , where  $\mathbf{J}_n^c$  corresponds to the  $n$ th cosine and  $\mathbf{J}_n^s$  to the  $n$ th sine Fourier coefficient of the current. The generalized expression of the local polarization profile (15) is then  $M_x = \pi \rho_1^c = \pi(\partial J_1^{x,c}/\partial x + \partial J_1^{y,c}/\partial y)/D_{\text{rot}}$  for the  $x$  component and  $M_y = \pi \rho_1^s = \pi(\partial J_1^{x,s}/\partial x + \partial J_1^{y,s}/\partial y)/D_{\text{rot}}$  for the  $y$  component.
- [28] A. P. Solon, J. Stenhammar, R. Wittkowski, M. Kardar, Y. Kafri, M. E. Cates, and J. Tailleur, *Phys. Rev. Lett.* **114**, 198301 (2015).
- [29] A. P. Solon, Y. Fily, A. Baskaran, M. E. Cates, Y. Kafri, M. Kardar, and J. Tailleur, *Nat. Phys.* **11**, 673 (2015).
- [30] As an example, based on the simulation data of Ref. [14] for motility-induced phase separation, we obtain the total polarization, i.e., the left-hand side of Eq. (18), as  $3.77/\sigma$ . This agrees well with the value  $3.73/\sigma$ , which we obtain for the right-hand side of Eq. (18). Here  $\sigma$  indicates the particle size; for further simulation details, see Ref. [14].
- [31] S. Hermann, D. de las Heras, and M. Schmidt, *Phys. Rev. Lett.* **123**, 268002 (2019).
- [32] A. Sharma and J. M. Brader, *Phys. Rev. E* **96**, 032604 (2017).
- [33] J. Hasnain, G. Menzl, S. Jungblut, and C. Dellago, *Soft Matter* **13**, 930 (2017).
- [34] W. Yan and J. F. Brady, *J. Fluid Mech.* **785**, R1 (2015).
- [35] A. K. Omar, Z.-G. Wang, and J. F. Brady, *Phys. Rev. E* **101**, 012604 (2020).



**Non-negative Interfacial Tension in Phase-Separated Active Brownian Particles**Sophie Hermann<sup>✉</sup>, Daniel de las Heras<sup>✉</sup>, and Matthias Schmidt<sup>✉\*</sup>*Theoretische Physik II, Physikalisches Institut, Universität Bayreuth, D-95447 Bayreuth, Germany*

(Received 1 August 2019; published 23 December 2019)

We present a microscopic theory for the nonequilibrium interfacial tension  $\gamma_{gl}$  of the free interface between gas and liquid phases of active Brownian particles. The underlying square gradient treatment and the splitting of the force balance in flow and structural contributions is general and applies to inhomogeneous nonequilibrium steady states. We find  $\gamma_{gl} \geq 0$ , which opposes claims by Bialké *et al.* [Phys. Rev. Lett. **115**, 098301 (2015)] and delivers the theoretical justification for the widely observed interfacial stability in active Brownian dynamics many-body simulations.

DOI: 10.1103/PhysRevLett.123.268002

The interfacial tension (or “surface tension”) of the free interface between two coexisting bulk phases is one of the most important quantities in the description of a wide range of interfacial phenomena. The tension  $\gamma_{gl}$  between coexisting gas and liquid bulk phases plays a particularly central role due to the high symmetry of the coexisting fluid phases. It is a key quantity in the Kelvin equation for capillary condensation, for the strength of the thermal capillary wave spectrum, and for the Laplace pressure in droplets.

The typical values of the interfacial tension vary over many orders of magnitude, when going from molecular to colloidal systems. Using the particle size  $\sigma$  and the thermal energy  $k_B T$  as the natural scales, the scaled interfacial tension  $\gamma_{gl} \sigma^2 / k_B T$  is typically of the order of unity. The dependence on  $\sigma$  is particularly dramatic when going from atoms to colloids. An associated factor of  $10^3$  of increase in length scale translates into a decrease of  $\gamma_{gl}$  by a factor of  $10^{-6}$ , as e.g., theoretically [1–3] and experimentally [4–6] demonstrated in phase separated colloid-polymer mixtures, where confocal microscopy can be used to great effect in studying e.g., droplet coalescence [7] and viscous fingering [8].

Very notably, the existence of the interfacial tension is the mechanism by which macroscopic fluid interfaces, such as in droplets and soap bubbles, attain a minimal geometric shape. The phase separated system minimizes the product of  $\gamma_{gl}$  and the interfacial area of the interface. As  $\gamma_{gl}$  is independent of curvature in a first approximation, this amounts to minimizing the interfacial area alone. This effect is e.g., commonly exploited in microscopic computer simulation work, where the use of periodic boundary conditions and suitable elongated box geometries offers the system a preferred (short) direction for the choice of interface orientation, and hence a stabilizing mechanism that truncates large scale fluctuations. This also applies to active Brownian particles, i.e., colloids where the diffusive motion is supplemented by directed self-propulsion and

which phase separate at large enough swimming strength [9–12]. Typical experiments rely on catalyzing a chemical reaction to induce such “swimming” [13].

There is much current progress in the description of free equilibrium interfaces, such as e.g., geometry-induced capillary emptying [14], the local structure factor near an interface [15], and Goldstone modes and resonances in the fluid interfacial region [16]. A variety of related deep theoretical topics have been addressed recently, including the curvature dependence of the surface free energy of liquid drops and bubbles [17], the adsorption of nanoparticles at fluid interfaces [18], the free energy of complex-shaped objects [19], the characterization of the “intrinsic” density profile for liquid surfaces [20–22], and the interface tension of curved interfaces [23].

All of the above physical understanding is necessarily based on the fundamental property  $\gamma_{gl} \geq 0$ . This seemingly indisputable fact was recently challenged based on computer simulation work by Bialké *et al.* [24] in active Brownian particles. The authors of Ref. [24] used the pressure tensor route and found their results for the interfacial tension to be negative. They argue that this “is a genuine nonequilibrium effect that is rationalized in terms of a positive stiffness.” Patch *et al.* [25] reproduce the negative result using an expression for  $\gamma_{gl}$  similar to that of Ref. [24], but with a different method for calculating the active contribution. From analysis of the interfacial (capillary wave) fluctuations, both groups find a positive value for the interfacial stiffness [24,25]. Lee constructs a coarse-grained model with an effective surface tension that is positive, and he is able to describe his simulation data [26]. Solon *et al.* [27] in their numerical analysis find a negative value for the tension, but they also state that their framework supports both positive and negative values. Marconi and Maggi [28] state that the tension would turn out to be negative in their theory. Subsequently, Marconi *et al.* [29] through analytical work have reconsidered the

problem of the mechanical derivation of  $\gamma_{\text{gl}}$ , but these authors do not report numerical results from their theory and they do not comment on the sign of  $\gamma_{\text{gl}}$  in Ref. [29]. Das *et al.* [30] investigated different expressions for the microscopic stress. The authors state that in their treatment the surface tension of active systems can be determined, but they have not done so in Ref. [30]. Considering the influence of activity on the gas-liquid interface of the Lennard-Jones system, Paliwal *et al.* [31] use the pressure tensor route and find a negative contribution from their swim term, but overall positive values for  $\gamma_{\text{gl}}$  across a wide parameter range.

Here we demonstrate, based on a nonequilibrium generalization of the microscopic treatment of the interface pioneered by van der Waals [32], that indeed the tension  $\gamma_{\text{gl}} \geq 0$  for phase-separated active Brownian particles. Its scaled value in natural units is of order unity, and vanishes with a 3/2 (mean-field) exponent near the critical point. This proves, on a sound theoretical footing, the hitherto unexplained stability of the planar active gas-liquid interface and demonstrates the route ahead to the quantitative description of nonequilibrium interfacial properties and phenomena. Our treatment is based on discriminating between structural forces that generate the tension and the flow force balance which does not.

Our mechanism for bulk phase separation is based on the exact translational one-body force balance equation [33,34],

$$\gamma \mathbf{v} = \mathbf{f}_{\text{id}} + \mathbf{f}_{\text{int}} + \gamma s \boldsymbol{\omega}, \quad (1)$$

where the friction force on the left-hand side is balanced by the ideal diffusive force  $\mathbf{f}_{\text{id}}$ , the internal force  $\mathbf{f}_{\text{int}}$ , and the free swim force  $\gamma s \boldsymbol{\omega}$  on the right-hand side. The friction constant is indicated by  $\gamma$  and  $s$  denotes the constant free swim speed. The velocity  $\mathbf{v}$ , the density  $\rho$ ,  $\mathbf{f}_{\text{id}}$ , and  $\mathbf{f}_{\text{int}}$  all depend on position  $\mathbf{r}$  and orientation  $\boldsymbol{\omega}$ , but not on time as we are considering steady states. Furthermore, we assume (i) the interface between the dense (liquid) and the dilute (gas) phases to be perpendicular to the  $x$  axis and (ii) translational invariance with respect to other spatial coordinates. Hence the density varies along the  $x$  axis of the system. The ideal diffusive force field is given exactly as  $\mathbf{f}_{\text{id}} = -k_B T \nabla \ln \rho$ . The internal force field consists of adiabatic and superadiabatic contributions and is defined as

$$\mathbf{f}_{\text{int}} = \mathbf{f}_{\text{ad}} + \mathbf{f}_{\text{sup}} = -\frac{1}{\rho} \left\langle \sum_i \delta_i \nabla_i u(\mathbf{r}_1, \dots, \mathbf{r}_N) \right\rangle, \quad (2)$$

where  $\delta_i = \delta(\mathbf{r} - \mathbf{r}_i) \delta(\boldsymbol{\omega} - \boldsymbol{\omega}_i)$  is used as a shorthand notation with  $\delta$  the Dirac delta function,  $u$  indicates the interparticle interaction potential,  $\nabla_i$  is the derivative with respect to position  $\mathbf{r}_i$  of the  $i = 1, \dots, N$  particle, and  $\langle \cdot \rangle$  is an average in steady state. The adiabatic force field  $\mathbf{f}_{\text{ad}}$  is defined by the right-hand side of Eq. (2) but taken in an

equilibrium system under the influence of an ‘‘adiabatic’’ external potential that generates the true density profile  $\rho$  [35–37]. Here the corresponding equilibrium system has no flow ( $s = 0$ ). Because of the rotational symmetry of the spherical particles considered here  $\mathbf{f}_{\text{ad}}$  is independent of the particle orientation  $\boldsymbol{\omega}$ . From classical density functional theory [38], applied to the adiabatic system, it is known that  $\mathbf{f}_{\text{ad}}$  is a gradient field obtained as  $\mathbf{f}_{\text{ad}} = -\nabla \mu_{\text{ad}}$  [39].

The superadiabatic force field is defined as the difference  $\mathbf{f}_{\text{sup}} = \mathbf{f}_{\text{int}} - \mathbf{f}_{\text{ad}}$ , cf. Eq. (2). From power functional theory [35] follows that  $\mathbf{f}_{\text{sup}}$  is a functional of the density profile, but also of the velocity profile.

We split Eq. (1) into a flow equation and a structural equation, given, respectively, by

$$\gamma \mathbf{v} = \mathbf{f}_{\text{flow}} + \gamma s \boldsymbol{\omega}, \quad (3)$$

$$0 = \mathbf{f}_{\text{id}} + \mathbf{f}_{\text{ad}} + \mathbf{f}_{\text{struc}}, \quad (4)$$

where the superadiabatic force field is the sum of a flow and a structural contribution,  $\mathbf{f}_{\text{sup}} = \mathbf{f}_{\text{flow}} + \mathbf{f}_{\text{struc}}$ . The splitting is unique. The superadiabatic flow force field  $\mathbf{f}_{\text{flow}}$  describes the influence of the internal interactions on the flow. The structural force field  $\mathbf{f}_{\text{struc}}$  is that part of the total internal force field that influences the spatial structure, together with the adiabatic force field  $\mathbf{f}_{\text{ad}}$  and the ideal term (which is small in the present situation). Note that it is the functional dependence of  $\mathbf{f}_{\text{sup}}$  and hence of  $\mathbf{f}_{\text{struc}}$  on velocity which renders Eq. (4) (highly) nontrivial. Since  $\mathbf{f}_{\text{id}}$  and  $\mathbf{f}_{\text{ad}}$  are gradient contributions,  $\mathbf{f}_{\text{struc}}$  necessarily needs to be a gradient field,  $\mathbf{f}_{\text{struc}} = -\nabla \mu_{\text{struc}}$ , which defines  $\mu_{\text{struc}}$  as the negative integral of  $\mathbf{f}_{\text{struc}}$ . Integrating Eq. (4) in space thus leads to

$$\mu_{\text{id}} + \mu_{\text{ad}} + \mu_{\text{struc}} = \mu_b = \text{const}, \quad (5)$$

where  $\mu_b$  is the constant value in the bulk fluid and the sum determines the total chemical potential. The difference to the equilibrium situation is the dependence of  $\mu_{\text{struc}}$  on the (nonvanishing) flow profile. Conceptually, the three chemical potential contributions play the same role as in equilibrium in that their respective gradient is a force field.

The ideal chemical potential  $\mu_{\text{id}} = k_B T \ln \rho$  is for simplicity reduced to the orientation-independent expression

$$\mu_{\text{id}} = k_B T \ln \rho_0, \quad (6)$$

with the rotational averaged density  $\rho_0 = \int d\boldsymbol{\omega} \rho / 2\pi$ . The approximation is reasonable, since the ideal chemical potential is numerically small in the present situation, as is the corresponding ideal diffusive force (see e.g., Ref. [24]). Furthermore,  $\rho_0$  is a main contribution of the Fourier decomposed density  $\rho$  and both densities  $\rho$  and  $\rho_0$  coincide in bulk. Since within the used approximations  $\mu_{\text{id}}$

and  $\mu_{\text{ad}}$  are rotationally invariant, Eq. (5) implies that  $\mu_{\text{struc}}$  and, hence,  $\mathbf{f}_{\text{struc}}$  are also independent of orientation.

We further discriminate between local and nonlocal contributions in Eq. (5). The ideal chemical potential  $\mu_{\text{id}}$  is a purely local expression and  $\mu_{\text{ad}}$  is also a local term since we base it on a local density approximation. Further nonlocal contributions to  $\mu_{\text{ad}}$  were found to be negligible in the present case. Hence the only considerable nonlocal contribution is contained in  $\mu_{\text{struc}}$ , which we split into a sum of local and nonlocal terms,  $\mu_{\text{struc}} = \mu_{\text{struc}}^{\text{loc}} + \mu_{\text{nlloc}}$ . The nonlocal superadiabatic chemical potential is approximated as the lowest order gradient contribution,

$$\mu_{\text{nlloc}} = -\nabla \cdot (m\nabla\rho_0) + \frac{1}{2}(\nabla m) \cdot (\nabla\rho_0), \quad (7)$$

where the coefficient  $m$  can depend on density  $\rho_0$  and on velocity  $\mathbf{v}$ . Note that  $\mu_{\text{nlloc}}$  vanishes in both bulk phases due to the constant density  $\rho_b = \rho_g, \rho_l$ , where  $\rho_g$  and  $\rho_l$  are the constant densities in the gas and liquid phase. Thus, bulk chemical potential and local chemical potential coincide in bulk,  $\mu_b = \mu_{\text{loc}}(\rho_b)$ .

For the local chemical potential,  $\mu_{\text{loc}} = \mu_{\text{id}} + \mu_{\text{ad}} + \mu_{\text{struc}}^{\text{loc}}$ , the corresponding nonequilibrium local pressure,  $P_{\text{loc}} = P_{\text{id}} + P_{\text{ad}} + P_{\text{struc}}^{\text{loc}}$ , can be obtained from the Gibbs-Duhem relation [40]:

$$\frac{\partial P_{\text{loc}}}{\partial \rho_0} = \rho_0 \frac{\partial \mu_{\text{loc}}}{\partial \rho_0}. \quad (8)$$

From Eqs. (5) and (7) follows directly that  $\mu_{\text{loc}}(\rho_l) = \mu_{\text{loc}}(\rho_g) = \mu_b$ , and using the Gibbs-Duhem relation (8) leads to  $P_{\text{loc}}(\rho_l) = P_{\text{loc}}(\rho_g) = P_b$ . The combination of both relations allows us to determine both coexistence densities  $\rho_g$  and  $\rho_l$  and hence the phase diagram of the system, cf. Ref. [41].

As we have identified the structural gradient force contributions, we can proceed in a purely mechanical way. Hence the gas-liquid interfacial tension is given by [32,38]

$$\gamma_{\text{gl}} = \int dx \left( \frac{m}{2} (\nabla\rho_0)^2 - W \right). \quad (9)$$

Equation (9) consists of a nonlocal and a local part. The first, nonlocal contribution results from an (interfacial) square gradient expansion with coefficient  $m$ . The second, local term is given as

$$-W = \psi - \psi_b = (\mu_{\text{loc}} - \mu_b)\rho_0 - (P_{\text{loc}} - P_b), \quad (10)$$

where  $\psi = \mu_{\text{loc}}\rho_0 - P_{\text{loc}}$  and  $\psi_b = \mu_b\rho_0 - P_b$  contain the above introduced nonequilibrium (local) chemical potential and pressure. Note that  $\psi_b$  is not a constant bulk contribution, since  $\rho_0$  still depends on  $x$ . In equilibrium  $\psi$  can

be identified as the local Helmholtz free-energy density and  $\psi_b$  is the corresponding double tangent line.

The chemical potential balance Eq. (5) can then be rewritten as

$$\frac{\partial W}{\partial \rho_0} + \nabla \cdot (m\nabla\rho_0) - \frac{1}{2}(\nabla m) \cdot (\nabla\rho_0) = 0, \quad (11)$$

where we used Eq. (7) to express the nonequilibrium chemical potential and the derivative of Eq. (10) with respect to density,  $-\partial W/\partial \rho_0 = \mu_{\text{loc}} - \mu_b$ . The first integral with respect to  $x$  of Eq. (11) is

$$W + \frac{1}{2}m \left( \frac{\partial \rho_0}{\partial x} \right)^2 = 0, \quad (12)$$

where we used the planar symmetry of the density  $\rho_0$  to simplify the spatial derivative  $\nabla$  to  $\hat{\mathbf{e}}_x \partial/\partial x$ . Rewriting the interfacial tension (9) with relation (12) leads to three alternative forms:

$$\gamma_{\text{gl}} = \int_{-\infty}^{\infty} m \left( \frac{\partial \rho_0}{\partial x} \right)^2 dx \quad (13)$$

$$= -2 \int_{-\infty}^{\infty} W dx \quad (14)$$

$$= \int_{\rho_g}^{\rho_l} \sqrt{-2mW} d\rho_0. \quad (15)$$

The numerical values of Eqs. (13)–(15) only coincide if the functions  $\rho_0$ ,  $m$ , and  $W$  are chosen reasonably and satisfy Eq. (12). Thus, whether a choice of these three functions is appropriate can be gauged by the agreement of the value for  $\gamma_{\text{gl}}$  obtained from either of Eqs. (13)–(15). This provides a check for the approximations for  $m$  and  $W$  as introduced below.

Equation (13) does not depend on the local contribution  $W$  and is thus referred to as the nonlocal route. The relation (14) is independent of the coefficient  $m$  of the nonlocal term. It is denoted as the local route, as the integrand is the local quantity  $W$ . Expression (15) is called the no-profile route, as it is independent of the density distribution  $\rho_0$ . In practice it can be useful to calculate  $\gamma_{\text{gl}}$  without knowledge of  $\rho_0$ . In the equilibrium limit of passive particles ( $s = 0$ ) and vanishing particle velocity  $\mathbf{v} = 0$ , our expressions for the interfacial tension coincide with the known equilibrium relations, cf. e.g., Ref. [32].

We apply our general theory for the nonequilibrium interfacial tension to a system of two-dimensional active particles which interact via a Weeks-Chandler-Anderson potential. This is a Lennard-Jones potential cut at its minimum and shifted to be continuous. The corresponding energy scale is  $\epsilon$  and the characteristic length scale  $\sigma$  is also referred to as the diameter of the spherical particles. The

orientational motion is freely diffusive with rotational diffusion constant  $k_B T / \gamma^\omega$ , where  $\gamma^\omega$  denotes the rotational friction constant. The rotational averaged density can be approximated with high accuracy as a hyperbolic tangent profile [12],

$$\rho_0(x) = \frac{\rho_g + \rho_l}{2} + \frac{\rho_l - \rho_g}{2} \tanh\left(\frac{x}{\lambda}\right), \quad (16)$$

where  $\lambda$  indicates the interfacial width. The coexistence densities  $\rho_g$  and  $\rho_l$  were determined from the pressure and chemical potential balance at theoretical coexistence and coincide with results from simulations very well [41,42].

The chemical potential contributions are chosen in accordance with Ref. [41]. The ideal term is given by relation (6). For the adiabatic chemical potential we use the local density approximation on a scaled particle theory for two-dimensional hard disks [43]. This yields

$$\mu_{\text{ad}} = k_B T [-\ln(1 - \eta') + \eta'(3 - 2\eta')/(1 - \eta')^2], \quad (17)$$

where the rescaled packing fraction  $\eta' = 0.8\eta$  models the soft Weeks-Chandler-Anderson potential. The packing fraction  $\eta = \rho_0/\rho_j$  and  $\rho_j = \text{const}$  indicates the jamming density, where the motion comes to arrest. The remaining  $\mu_{\text{struc}}$  corresponds to the quite life chemical potential [41], which in homogeneous bulk is given as

$$\mu_{\text{struc}}^b = \frac{e_1 \gamma \gamma^\omega}{2 k_B T} v_b^2 \frac{\rho_b}{\rho_j}, \quad (18)$$

where the strength is determined by the dimensionless constant  $e_1$ . The expression (18) is linear in bulk density  $\rho_b$ , quadratic in the bulk speed  $v_b$ , and the resulting force acts toward the liquid phase. Note that due to its velocity dependence  $\mu_{\text{struc}}^b$  is a genuine nonequilibrium expression. To obtain the local structural chemical potential, we expand Eq. (18) across the interface using the orientational averaged density  $\rho_0$  instead of  $\rho_b$  and the known linear decrease  $v_{\text{loc}} = s(1 - \rho_0/\rho_j)$  [11] for the speed  $v_b$ . This yields

$$\mu_{\text{struc}}^{\text{loc}} = \frac{e_1}{6} \text{Pe}^2 k_B T \left(1 - \frac{\rho_0}{\rho_j}\right)^2 \frac{\rho_0}{\rho_j}, \quad (19)$$

where the introduced Péclet number is  $\text{Pe} = s\sigma\gamma/k_B T = 3s\gamma^\omega/k_B T\sigma$ . This dimensionless constant relates active swimming to rotational diffusion.

The nonlocal chemical potential is approximated in the simplest way, with a constant coefficient  $m = e_2 \text{Pe}^2 k_B T / 6\rho_j^2$ , such that Eq. (7) simplifies to  $\mu_{\text{noc}} = -m\nabla^2 \rho_0$  and one obtains

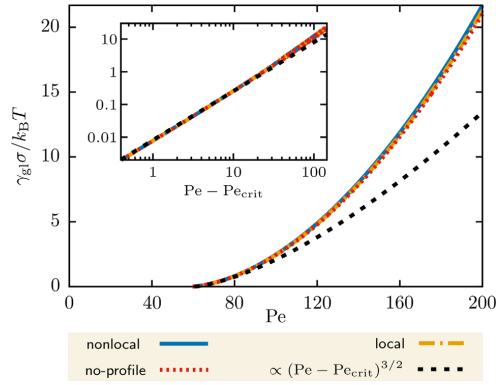


FIG. 1. Interfacial tension  $\gamma_{\text{gl}}$  determined from the nonlocal route Eq. (13) (full blue line), from the local route Eq. (14) (dash-dotted yellow line), and from no-profile route Eq. (15) (dotted red line) in dependence of the Péclet number  $\text{Pe}$ . Close to the critical point the tension increases with a critical exponent of  $3/2$ , as indicated by the dashed black line. The inset also shows  $\gamma_{\text{gl}}$  but in a double-logarithmic plot and the  $x$  axis is shifted by the critical Péclet number and, hence, is  $\text{Pe} - \text{Pe}_{\text{crit}}$ .

$$\mu_{\text{noc}} = -\frac{e_2 \text{Pe}^2 k_B T}{6} \frac{\nabla^2 \rho_0}{\rho_j}, \quad (20)$$

where the amplitude is determined by the dimensionless constant  $e_2$ . One can show within the power functional framework [33,35] that  $\mu_{\text{struc}}$  is an intrinsic quantity and can be written as a kinematic functional, hence only dependent on density  $\rho$  and velocity  $\mathbf{v}$ . Therefore,  $\mu_{\text{struc}}^{\text{loc}}$  and  $\mu_{\text{noc}}$  are “naturally” independent of the swim speed and Eq. (19) can be expressed without  $s$  as an intrinsic expression [41,44]. The local pressure can be determined straightforwardly from the Gibbs-Duhem relation Eq. (8).

The parameters of the system are chosen as follows. The system is at temperature  $k_B T / \epsilon = 0.5$ , has a rotational friction coefficient  $\gamma^\omega / \gamma \sigma^2 = 1/3$ , a jamming density of  $\rho_j 2\pi\sigma^2 = 1.4$ , and the dimensionless prefactors  $e_1 = 0.0865$  and  $e_2 = 0.0385$ . Requiring  $e_2$  to be constant and the chemical potential balance (5) to be satisfied, the interfacial width  $\lambda$  is determined. The swim speed  $s$  changes with Péclet number,  $\text{Pe} = s\sigma\gamma/k_B T$ , while the other parameters are kept constant. We use the approximations for the orientational averaged density profile  $\rho_0$ , the chemical potential contributions Eqs. (6), (17), (19), and (20), and the corresponding pressures to determine the interfacial tension by evaluating the expressions (13)–(15).

The results from the three methods are displayed in Fig. 1. We find the behavior of the function  $W$  and of the interfacial tension to be qualitatively similar to what is found in equilibrium [32]. Figure 1 shows  $\gamma_{\text{gl}}$  as a function of  $\text{Pe}$ . The interfacial tension is only different from zero for

Péclet numbers larger than  $Pe_{\text{crit}} = 59.3$  [41], when the system phase separates. Here the critical value of the Péclet number [41] is determined by the magnitude of  $e_1$ . The tension increases with rising particle activity and hence with the Péclet number (cf. Fig. 1). Close to the critical point  $\gamma_{\text{gl}}$  increases with a critical exponent of  $3/2$ , as indicated by the black dashed line. This corresponds to the theoretical mean-field coefficient of the van der Waals theory, which might be expected since there are many similarities between both descriptions. In order to emphasize the agreement of the interfacial tension with a function proportional to  $(Pe - Pe_{\text{crit}})^{3/2}$ , both quantities are displayed in a double logarithmic plot (cf. the inset of Fig. 1). For Péclet numbers close to the critical point, the functions nearly have the same slope. Far from the critical point the interfacial tension increases faster than with the critical exponent. For a detailed simulation study of the bulk critical behavior of active Brownian particles, see Ref. [45].

The values of the tension are positive,  $\gamma_{\text{gl}} > 0$ , which directly explains the stability of the interface. This is in contrast to Bialké *et al.* [24], who calculated a negative interfacial tension using the pressure tensor. The results for three different methods, the nonlocal route Eq. (13), the local route Eq. (14), and the no-profile route Eq. (15), agree to a very satisfying degree (cf. Fig. 1). Even far from equilibrium, for example, at  $Pe = 200$ , the respective results deviate by only about 3%. This indicates that the chemical potential balance (5) and hence the structural force balance (4) are both satisfied with very good accuracy. Finally, the splitting (3) and (4) [together with (9) within a square gradient approximation] forms a general route toward the interfacial tension of out-of-equilibrium interfaces. We have also ascertained that the “flow” equation of motion (3) creates a vanishing contribution to the interfacial tension in the present system, since after orientational integration the associated pressure tensor contributions either vanish or are isotropic. Hence the splitting (3) and (4) does not imply omission of any relevant terms.

Because of the square gradient character of our treatment, we do not find layering effects at the interface, which would require us to take account of nonlocal interfacial packing effects [20]. Furthermore, our treatment yields the “intrinsic density profile” [21,22], as large scale capillary wave fluctuations are neglected. Thus, interesting future work could be devoted to studying capillary wave fluctuations and the wave vector dependence of the interfacial tension [21,22]. Furthermore, it would be interesting to relate our treatment to that presented in Ref. [27] and to consider fluctuations beyond mean field that could alter the value of the critical scaling exponent.

---

\*Matthias.Schmidt@uni-bayreuth.de

[1] A. Vrij, *Physica (Amsterdam)* **235A**, 120 (1997).

[2] J. M. Brader and R. Evans, *Europhys. Lett.* **49**, 678 (2000).

- [3] J. M. Brader, R. Evans, M. Schmidt, and H. Löwen, *J. Phys. Condens. Matter* **14**, L1 (2002).
- [4] E. H. A. de Hoog and H. N. W. Lekkerkerker, *J. Phys. Chem. B* **103**, 5274 (1999).
- [5] E. H. A. de Hoog and H. N. W. Lekkerkerker, *J. Phys. Chem. B* **105**, 11636 (2001).
- [6] D. G. A. L. Aarts, M. Schmidt, and H. N. W. Lekkerkerker, *Science* **304**, 847 (2004).
- [7] D. G. A. L. Aarts and H. N. W. Lekkerkerker, *J. Fluid Mech.* **606**, 275 (2008).
- [8] S. A. Setu, I. Zacharoudiou, G. J. Davies, D. Bartolo, S. Moulinet, A. A. Louis, J. M. Yeomans, and D. G. A. L. Aarts, *Soft Matter* **9**, 10599 (2013).
- [9] T. F. F. Farage, P. Krinninger, and J. M. Brader, *Phys. Rev. E* **91**, 042310 (2015).
- [10] R. Wittmann, U. M. B. Marconi, C. Maggi, and J. M. Brader, *J. Stat. Mech.* (2017) 113208.
- [11] T. Speck, A. M. Menzel, J. Bialké, and H. Löwen, *J. Chem. Phys.* **142**, 224109 (2015).
- [12] S. Paliwal, J. Rodenburg, R. van Roij, and M. Dijkstra, *New J. Phys.* **20**, 015003 (2018).
- [13] I. Buttinoni, J. Bialké, F. Kümmel, H. Löwen, C. Bechinger, and T. Speck, *Phys. Rev. Lett.* **110**, 238301 (2013).
- [14] C. Rascon, A. O. Parry, and D. G. A. L. Aarts, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 12633 (2016).
- [15] A. O. Parry, C. Rascon, and R. Evans, *J. Phys. Condens. Matter* **28**, 244013 (2016).
- [16] A. O. Parry and C. Rascon, *Nat. Phys.* **15**, 287 (2019).
- [17] B. J. Block, S. K. Das, M. Oettel, P. Virnau, and K. Binder, *J. Chem. Phys.* **133**, 154702 (2010).
- [18] F. Bresme and M. Oettel, *J. Phys. Condens. Matter* **19**, 413101 (2007).
- [19] P. M. König, R. Roth, and K. R. Mecke, *Phys. Rev. Lett.* **93**, 160601 (2004).
- [20] E. Chacón, M. Reinaldo-Falagán, E. Velasco, and P. Tarazona, *Phys. Rev. Lett.* **87**, 166101 (2001).
- [21] P. Tarazona and E. Chacón, *Phys. Rev. B* **70**, 235407 (2004).
- [22] E. Chacón and P. Tarazona, *J. Phys. Condens. Matter* **17**, S3493 (2005).
- [23] A. Tröster, M. Oettel, B. Block, P. Virnau, and K. Binder, *J. Chem. Phys.* **136**, 064709 (2012).
- [24] J. Bialké, J. T. Siebert, H. Löwen, and T. Speck, *Phys. Rev. Lett.* **115**, 098301 (2015).
- [25] A. Patch, D. M. Sussman, D. Yllanes, and M. C. Marchetti, *Soft Matter* **14**, 7435 (2018).
- [26] C. F. Lee, *Soft Matter* **13**, 376 (2017).
- [27] A. P. Solon, J. Stenhammar, M. E. Cates, Y. Kafri, and J. Tailleur, *New J. Phys.* **20**, 075001 (2018).
- [28] U. M. B. Marconi and C. Maggi, *Soft Matter* **11**, 8768 (2015).
- [29] U. M. B. Marconi, C. Maggi, and S. Melchionna, *Soft Matter* **12**, 5727 (2016).
- [30] S. Das, G. Gompper, and R. G. Winkler, *Sci. Rep.* **9**, 6608 (2019).
- [31] S. Paliwal, V. Prymidis, L. Filion, and M. Dijkstra, *J. Chem. Phys.* **147**, 084902 (2017).
- [32] J. S. Rowlinson and B. Widom, *Molecular Theory of Capillarity* (Dover, New York, 2002).
- [33] P. Krinninger and M. Schmidt, *J. Chem. Phys.* **150**, 074112 (2019).

- [34] P. Krinninger, M. Schmidt, and J. M. Brader, *Phys. Rev. Lett.* **117**, 208003 (2016).
- [35] M. Schmidt and J. M. Brader, *J. Chem. Phys.* **138**, 214101 (2013).
- [36] A. Fortini, D. de las Heras, J. M. Brader, and M. Schmidt, *Phys. Rev. Lett.* **113**, 167801 (2014).
- [37] D. de las Heras, J. Renner, and M. Schmidt, *Phys. Rev. E* **99**, 023306 (2019).
- [38] R. Evans, *Adv. Phys.* **28**, 143 (1979).
- [39] Within density functional theory, using the Helmholtz excess free-energy functional  $F_{\text{exc}}[\rho]$ , the internal chemical potential  $\mu_{\text{ad}} = \delta F_{\text{exc}}[\rho]/\delta\rho$  can be written as a functional of (only) the density profile and independent of external forces.
- [40] The Gibbs-Duhem equation results from identifying the negative gradient of a chemical potential  $\mu$  as a force and the negative gradient of a pressure  $P$  as a force density. The combination of both relations leads to  $-\rho\nabla\mu = -\nabla P$ , where the spatial derivative can be rewritten as  $\rho\partial\mu/\partial\rho_0\nabla\rho_0 = \partial P/\partial\rho_0\nabla\rho_0$ . Simplification and averaging over orientation gives Eq. (8) in case of local and rotational independent  $\mu$  and  $P$ .
- [41] S. Hermann, P. Krinninger, D. de las Heras, and M. Schmidt, *Phys. Rev. E* **100**, 052604 (2019).
- [42] The simulation results presented in Ref. [41] are based on  $N = 2000$  particles in rectangular boxes of varying aspect ratio 2.5, 5, and 10. The systems were initialized in configurations with the interface running along the short direction of the simulation box.
- [43] J. A. Barker and D. Henderson, *Rev. Mod. Phys.* **48**, 587 (1976).
- [44] In order to express the swim speed with internal quantities, one can use the forward speed  $v_f = \int d\omega \mathbf{J} \cdot \omega / 2\pi\rho_0$ , the orientational integrated projection of the translational current  $\mathbf{J}$  on the particle orientation  $\omega$ . It is approximately given as  $v_f = s(1 - \rho_0/\rho_j)/[1 + \xi(\nabla\rho_0)^2\rho_0/\rho_j]$  [41], where the constant  $\xi > 0$  sets the amplitude of a square gradient expansion term. Hence the linear decrease in speed  $v_{\text{loc}} = s(1 - \rho_0/\rho_j)$  in Eq. (18) can be replaced by the intrinsic expression  $v_f[1 + \xi(\nabla\rho_0)^2\rho_0/\rho_j]$ .
- [45] J. T. Siebert, F. Dittrich, F. Schmid, K. Binder, T. Speck, and P. Virnau, *Phys. Rev. E* **98**, 030601(R) (2018).

# A Bulk and interfacial theory of motility-induced phase separation

In the following we give an overview of background and hence summarize the main ideas of the theory to describe the phase separation of active Brownian particles. For more theoretical details, especially for the treatment of the flow contributions, please consider the corresponding references [46,94]. Reference [93] contains an earlier version. In reference [46] we construct approximative theoretical expressions by comparing to the results from Brownian dynamic simulation. The determined profiles of the density and the current and the resulting phase diagram agree favourably with simulation results. In reference [94] we lay out all theoretical steps and show the calculations in more detail than given in [46].

Our one-body theory is based on two main equations. The first one is the exact angle-resolved continuity equation, i.e. equation (4.1), which we reproduce for convenience:

$$\dot{\rho}(\mathbf{r}, \boldsymbol{\omega}, t) = -\nabla \cdot \mathbf{J}(\mathbf{r}, \boldsymbol{\omega}, t) - \nabla_{\boldsymbol{\omega}} \cdot \mathbf{J}^{\boldsymbol{\omega}}(\mathbf{r}, \boldsymbol{\omega}, t). \quad (\text{A.1})$$

The um of the negative spatial divergence  $\nabla$  of the translational current  $\mathbf{J}(\mathbf{r}, \boldsymbol{\omega}, t)$  and the negative angular divergence  $\nabla_{\boldsymbol{\omega}}$  of the rotational current  $\mathbf{J}^{\boldsymbol{\omega}}(\mathbf{r}, \boldsymbol{\omega}, t)$  give the temporal change in the density  $\dot{\rho}(\mathbf{r}, \boldsymbol{\omega}, t) = \partial\rho(\mathbf{r}, \boldsymbol{\omega}, t)/\partial t$ . Here  $\dot{\rho}(\mathbf{r}, \boldsymbol{\omega}, t) = 0$  as we only consider steady states. Note that the one-body quantities such as the density  $\rho(\mathbf{r}, \boldsymbol{\omega})$  or the currents  $\mathbf{J}(\mathbf{r}, \boldsymbol{\omega})$ ,  $\mathbf{J}^{\boldsymbol{\omega}}(\mathbf{r}, \boldsymbol{\omega})$  are fully resolved in position and orientation and we have dropped the irrelevant time argument. The active Brownian particle orientations diffuse freely (5.2) as there are neither external nor internal torques present. Hence the rotational current only includes a diffusive contribution which is given by  $\mathbf{J}^{\boldsymbol{\omega}}(\mathbf{r}, \boldsymbol{\omega}) = -D_{\text{rot}}\nabla_{\boldsymbol{\omega}}\rho(\mathbf{r}, \boldsymbol{\omega})$ , where  $D_{\text{rot}}$  denotes the rotational diffusion constant.

The second fundamental equation of the theory is the force balance,

$$\gamma\mathbf{v}(\mathbf{r}, \boldsymbol{\omega}) = \mathbf{f}_{\text{id}}(\mathbf{r}, \boldsymbol{\omega}) + \mathbf{f}_{\text{swim}}(\mathbf{r}, \boldsymbol{\omega}) + \mathbf{f}_{\text{int}}(\mathbf{r}, \boldsymbol{\omega}), \quad (\text{A.2})$$

which is equivalent to the previously considered force density balance (4.18) up to division by  $\rho(\mathbf{r}, \boldsymbol{\omega})$  using the relation between the one-body current and velocity,  $\mathbf{J}(\mathbf{r}, \boldsymbol{\omega}) = \mathbf{v}(\mathbf{r}, \boldsymbol{\omega})\rho(\mathbf{r}, \boldsymbol{\omega})$ . The frictional force on the left-hand side of equation (A.2) is canceled with the ideal term  $\mathbf{f}_{\text{id}}(\mathbf{r}, \boldsymbol{\omega}) = -k_{\text{B}}T\nabla \ln \rho(\mathbf{r}, \boldsymbol{\omega})$ , the swim force  $\mathbf{f}_{\text{swim}}(\mathbf{r}, \boldsymbol{\omega}) = \gamma s\boldsymbol{\omega}$  and the internal contribution  $\mathbf{f}_{\text{int}}(\mathbf{r}, \boldsymbol{\omega})$  resulting from the interparticle interactions. Here  $k_{\text{B}}$  is the Boltzmann,  $T$  denotes the temperature,  $\gamma$  indicates the friction constant and  $s$  is the free swim speed. For each force contribution there are exact and explicit expressions, except for the interaction force for which one has to make approximations in order to make progress. Hence this term is the most difficult one to treat, as it contains the entire complexity of the problem.

It turns out to be useful to split the internal force into a sum of three distinct contributions, which are a flow force, an adiabatic force and a structural force,

$\mathbf{f}_{\text{int}}(\mathbf{r}, \boldsymbol{\omega}) = \mathbf{f}_{\text{flow}}(\mathbf{r}, \boldsymbol{\omega}) + \mathbf{f}_{\text{ad}}(\mathbf{r}, \boldsymbol{\omega}) + \mathbf{f}_{\text{struc}}(\mathbf{r}, \boldsymbol{\omega})$ . The adiabatic force  $\mathbf{f}_{\text{ad}}(\mathbf{r}, \boldsymbol{\omega})$  describes the interparticle interaction force that would act if the system were in equilibrium and have the same one-body density profile as the nonequilibrium system [79]. Therefore this force can be seen as the equilibrium contribution to  $\mathbf{f}_{\text{int}}(\mathbf{r}, \boldsymbol{\omega})$ . In contrast the flow ( $\mathbf{f}_{\text{flow}}(\mathbf{r}, \boldsymbol{\omega})$ ) and the structural ( $\mathbf{f}_{\text{struc}}(\mathbf{r}, \boldsymbol{\omega})$ ) force form the nonequilibrium contribution which together form the superadiabatic (above adiabatic) force. This splitting allows to separate the force balance equation (A.2) into a flow part, which is given by

$$\gamma \mathbf{v}(\mathbf{r}, \boldsymbol{\omega}) = \mathbf{f}_{\text{swim}}(\mathbf{r}, \boldsymbol{\omega}) + \mathbf{f}_{\text{flow}}(\mathbf{r}, \boldsymbol{\omega}). \quad (\text{A.3})$$

Equation (A.3) directly describes the coupled motion and currents in the system. It is coupled to structural relationship and a structural part,

$$0 = \mathbf{f}_{\text{id}}(\mathbf{r}, \boldsymbol{\omega}) + \mathbf{f}_{\text{ad}}(\mathbf{r}, \boldsymbol{\omega}) + \mathbf{f}_{\text{struc}}(\mathbf{r}, \boldsymbol{\omega}). \quad (\text{A.4})$$

which primarily induces the formation of spatial structure and is hence relevant for the phase coexistence. The equations (A.3) and (A.4) are taken to define  $\mathbf{f}_{\text{flow}}(\mathbf{r}, \boldsymbol{\omega})$  and  $\mathbf{f}_{\text{struc}}(\mathbf{r}, \boldsymbol{\omega})$ , so the formulation is without approximations up to this point. One can ascertain a posteriori that the chosen formal splitting is also physically sensible. Further both equations (A.3) and (A.4) are fully resolved with respect to the orientation. This level of description contrasts many other descriptions of motility-induced phase separation which only include the orientational average and sometimes the first and the second orientational moments of force and density distributions [44, 45].

Let us focus on the flow force balance (A.3) first. We use this equation in combination with the continuity equation (A.1) to determine the angle- and position-resolved density and current profiles [94] within assuming the hyperbolic tangent for the orientation-averaged density profile. One then only needs the values for the coexistence densities and a suitable approximation for  $\mathbf{f}_{\text{flow}}(\mathbf{r}, \boldsymbol{\omega})$ . The flow force itself consists again of several superadiabatic contributions,  $\mathbf{f}_{\text{flow}}(\mathbf{r}, \boldsymbol{\omega}) = \mathbf{f}_{\text{sup},0}(\mathbf{r}, \boldsymbol{\omega}) + \mathbf{f}_{\text{sup},1}(\mathbf{r}, \boldsymbol{\omega}) + \mathbf{f}_{\text{sup},2}(\mathbf{r}, \boldsymbol{\omega})$ . The first term  $\mathbf{f}_{\text{sup},0}(\mathbf{r}, \boldsymbol{\omega})$  is a drag contribution. Its origin lies in the motion of a particle that moves through the fluid of surrounding others particles and this force therefore opposes the current  $\mathbf{J}(\mathbf{r}, \boldsymbol{\omega})$ . As a simple approximation we assume that this contribution is spherical and proportional to  $-\mathbf{J}(\mathbf{r}, \boldsymbol{\omega})$ . The contribution  $\mathbf{f}_{\text{sup},1}(\mathbf{r}, \boldsymbol{\omega})$  is a more complex interfacial correction to the spherical drag force, which we determined from comparison with simulation data [46]. The last contribution,  $\mathbf{f}_{\text{sup},2}(\mathbf{r}, \boldsymbol{\omega})$ , is generated from a superadiabatic pressure and tends to expand the active liquid phase. This underlying pressure exactly cancels the swim pressure which originates from the interfacial polarization and the swim force  $\mathbf{f}_{\text{swim}}(\boldsymbol{\omega}) = \gamma s \boldsymbol{\omega}$ .

The flow balance equation (A.3) alone cannot explain the coexistence densities of the phase separated state. Therefore, we consider the structural force balance (A.4), with the primary contributions being the (mutually opposing) structural and adiabatic forces. The ideal gas force is typically negligibly small in the considered systems when considering a high Péclet number. The Péclet-number relates the free swim speed with the rotational diffusion constant,  $\text{Pe} = 3s/\sigma D_{\text{rot}}$  and it can be seen as a measure of the strength of the particle activity. While the adiabatic force tries to expand the denser phase to homogenize the system, the structural term compresses the liquid phase and ultimately stabilizes phase coexistence.

Realizing that all components of the structural force balance are gradient terms allows to rewrite the right-hand side of equation (A.4) as  $-\nabla\mu(\mathbf{r})$ , where we can now



identify the total chemical potential  $\mu(\mathbf{r})$  using the fact that the negative gradient of a chemical potential yields a force. Spatial integration of equation (A.4) thus gives

$$\mu(\mathbf{r}) = \mu_{\text{id}}(\mathbf{r}) + \mu_{\text{ad}}(\mathbf{r}) + \mu_{\text{struc}}(\mathbf{r}) = \text{const}, \quad (\text{A.5})$$

where the occurring chemical potentials  $\mu_\alpha(\mathbf{r})$  are related to the corresponding forces  $\mathbf{f}_\alpha(\mathbf{r}) = -\nabla\mu_\alpha(\mathbf{r})$  for the three contributions  $\alpha \in \{\text{id}, \text{ad}, \text{struc}\}$ . The first term is the well-known ideal gas chemical potential,  $\mu_{\text{id}}(\mathbf{r}, \boldsymbol{\omega}) = k_{\text{B}}T \ln \rho(\mathbf{r}, \boldsymbol{\omega})$ . As this term is negligibly small in the considered systems we simplify it to  $\mu_{\text{id}}(\mathbf{r}) \approx k_{\text{B}}T \ln \rho_0(\mathbf{r})$ , where  $\rho_0(\mathbf{r})$  denotes the orientationally averaged density distribution or respectively the zeroth Fourier coefficient of the density. The adiabatic term  $\mu_{\text{ad}}(\mathbf{r})$  is given via a corresponding equilibrium relation. For simplicity we approximate this chemical potential contribution with the expression from the scaled particle theory [76]. The last contribution is the structural chemical potential for which we constructed an approximation [46]. Due to symmetry we expect this term to have an even dependence on the particle velocity, i.e. the velocity factors only occurs via even powers. We assume for simplicity that the structural contribution is quadratic in the forward velocity and linear in density,  $\mu_{\text{struc}}^{\text{loc}}(\mathbf{r}) = e_1 \gamma v_{\text{f}}^2(\mathbf{r}) \rho_0(\mathbf{r}) / 2D_{\text{rot}} \rho_{\text{jam}}$ , where  $e_1$  is a fit parameter to adjust the magnitude and  $\rho_{\text{jam}}$  denotes the jamming density which is a constant that indicates the density at which the system would come to arrest because of sterical hindrance. The forward speed  $v_{\text{f}}$  describes the velocity that corresponds the current projected on the particle orientation. This velocity is high in the gaseous-like state and low in the denser liquid-like state. Note that we here for simplicity only specify the local contribution of the structural chemical potential. While the ideal and the adiabatic contributions (up to the considered order) just contain a local term, the structural chemical potential also includes a nonlocal interfacial term. For the phase diagram determined as by the Maxwell construction the bulk values and hence the local expressions are sufficient. For the interfacial tension also the interfacial contributions are important, as described in section 5.2.

At this point we have an explicit approximation for the total chemical potential. From this expression one can determine the total pressure  $p(\mathbf{r})$  using a Gibbs-Duhem like equation,

$$-\nabla p(\mathbf{r}) = -\rho(\mathbf{r}) \nabla \mu(\mathbf{r}). \quad (\text{A.6})$$

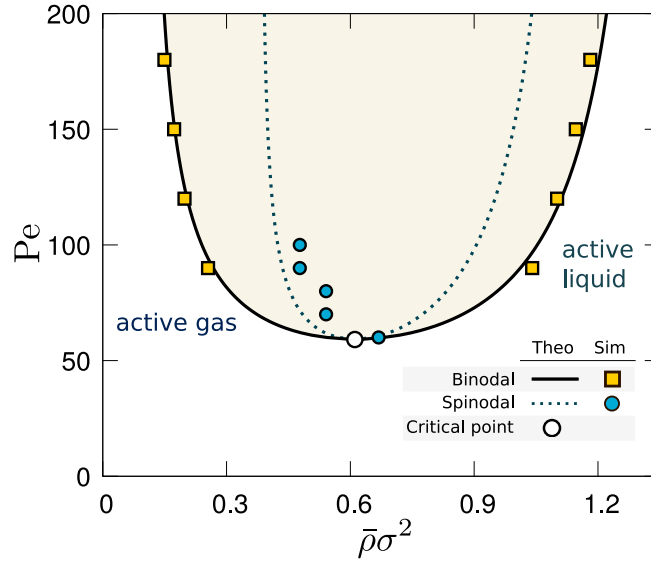
This equation results from the fact that the force and the force density are equal up to a factor  $\rho(\mathbf{r})$ . Thereby the force is expressed as the negative gradient of a chemical potential,  $-\nabla\mu(\mathbf{r})$  and the force density is considered as the negative gradient of a pressure,  $-\nabla P(\mathbf{r})$ . We know that the pressure balance (4.19) exists and that it is valid for MIPS as it is a consequence of the Noether and polarization sum rules, see section 4.3.

With the explicit expressions of the total chemical potential and the total pressure one can use the usual coexistence conditions

$$\mu(\rho^{\text{g}}) = \mu(\rho^{\text{l}}), \quad (\text{A.7})$$

$$p(\rho^{\text{g}}) = p(\rho^{\text{l}}). \quad (\text{A.8})$$

These conditions state that the chemical potential has the same value in the gaseous and in the liquid bulk phase and the same applies to the total pressure. If one would



**Figure A.1:** Phase diagram of two dimensional active Brownian particles in dependence of the Péclet-number  $Pe$  which measures the particle activity. The orange shaded area marks the region of phase coexistence in an active gas and an active liquid and is bounded by the theoretical results for the binodal. The corresponding simulation data from [44] is indicated with orange squares. The spinodal is given by the dashed line (theory) and the blue circles (simulations, [169]). The orientational averaged density  $\bar{\rho} = 2\pi\rho_0$  and  $\sigma$  denoted a length scale resulting from the Weeks-Chandler-Anderson interaction potential. It is used to make expressions dimensionless. The figure was used within [93] and was taken from [46], ©American Physical Society (2011). All rights reserved.

consider the parametric plot of  $\mu(p)$  in dependence of the density, then phase separation occurs when the graph intersects itself and the intersection determines the coexistence densities  $\rho^g$  and  $\rho^l$ . This mechanism is the Maxwell construction as is well-known from equilibrium. However, it turns out that in other approaches the transfer is not straightforward as there the application to simulation data yields unsatisfying results, consider e.g. reference [44, 45]. The deviation was corrected by Solon *et al.* [45] by introducing an additional term that measures the violation of the Maxwell equal area construction in their treatment.

In our approach solving equations (A.7), (A.8) for  $\rho_g$  and  $\rho_l$  determines the coexistence densities, which are shown as a function of the Péclet-number  $Pe$  and the orientational averaged density in the phase diagram shown in figure A.1. For high enough Péclet numbers the system phase separates into an active gas and an active liquid. The theoretical results (black line) are in good quantitative agreement with simulation data (orange squares) [44]. Note that the theory captures the feature of quite high active gas densities as compared to the gas densities in equilibrium liquid-vapour phase separations.

The dashed line in figure A.1 represents the theoretical prediction for the spinodal, which is defined by densities that satisfy the condition  $\partial\mu(\rho)/\partial\rho = 0$ . The data for the spinodal is compared to simulation results from Stenhammar *et al.* [169]. The white circle in Fig. A.1 indicates the critical point and thus the onset of phase coexistence. This successful bulk theory forms the basis of the interfacial treatment described in section 5.2 and published in [SH7].

# Acknowledgments

First of all, a huge thank you goes to my supervisor Matthias Schmidt. During my PhD you went with me through all the good (successful calculations) and bad times (writing thesis, Corona loneliness). I am still impressed by the enormous amount of enthusiasm, patience and trust you seem to have. For me it felt so reassuring to always be aware of your overall support and ability to motivate. Thank you for proofreading texts even after busy days, for the opportunity to visit so many conferences and for giving me the feeling that I can rely on your (actual and mental) help before deadlines. Also, thanks for your open ear for whatever bothered me and for all our (sometimes quite lengthy) private and scientific conversations. Walking in the ÖBG or elsewhere and discussing physics occasionally almost felt like walking in the footsteps of Emmy Noether. :) As you probably already know, you left quite an impression on me during my undergraduate studies. We first met in the classical mechanics lecture in my second semester. I really enjoyed your energetic and sometimes chaotic way to teach, and this course was my favorite Bachelor lecture. (Maybe it partly also became my favorite because of the Lagrange formalism, which I love.) At that time I got interested in theoretical physics, but back then I would have never imagined to actually work together with you. Thank you for becoming my research partner and such a good friend since then.

At this point I would also like to acknowledge the other two referees, Roland Roth and Klaus Kroy. Thank you for indeed reading the document I wrote and for your kind reports. It felt wonderful and a bit overwhelming that you, Klaus, appreciated my results enough to invite me for a presentation of them in a student seminar, which was a great pleasure to me.

I want to thank Joe Brader for his invitation to visit him and his group in Fribourg for six weeks. Thank you, Joe and Salomé, for the warm welcome in your group and for organizing so many shared experiences as sightseeing trips or baking cookies together. I really enjoyed our shared passion for spinners and pumpkins and it was great to see that those two things get even better when combined to a pumpkin-spinner. ;)

In general, I would like to thank all people that hang out in the NW2 building (and of course also in the physics part of the NW1 building) for making the university feel like home to me. It was always fun to stop for a short chat in the foyer and exchange some gossip. Many of you became good friends to me during the last years.

Special thanks go to our group TP2, including all current members and former colleagues of mine. Thank you, Joe, for being such a great office mate all the years. You have been a dear help and support, be it when organizing joined home office in Covid times, just listening to me and trying to answer all the stupid questions that frequently come up my mind. I deeply enjoy our cycling trips, gaming nights and our almost daily joined lunches. Thank you, Dani, for all the discussions and for helping me make

## *Acknowledgments*

up my mind about my future plans. A big thanks also goes to our secretary Sabrina Süß. I am very grateful for all your help, be it with complex organizational stuff as the celebration after my defense or simply telling me how to use the phone correctly.

For keeping me in touch with the real world and regularly showing and explaining me her experimental set up, I want to thank Anna. As you know, I love our shared passion in learning physics as well as all the lectures and exercise sheets we attempted together and sometimes failed (just recall the Casimir forces exercise from QM2). Thanks for all these years of our deep friendship.

My friend Emmi which I first met as she was a participant in one of the tutorials I taught. Back then I would never expected that a friendship could grow out of that. Thank you for always taking the time to listen to me, my problems or everything that burdens my mind and for your understanding and sympathizing character. It is great to have a person that has so many similar interests as crocheting, reading books and watching musicals.

A special thank you goes to my parents and my whole family. You made me the person that I am today. I cannot express how much I appreciate your unconditional love and your overall support. Thanks mum, for phoning with me so many times when I suffered through hard times. Thanks dad, for encouraging me to do what I love, although I ended up as a physicist and not as an electronics technician as you hoped for (sorry). :)

Thanks, Leni and Krissi, for exploring and sharing the passion of our joint love for puzzles. It was and is so much fun to solving riddles, crosswords and exit games together or competitive. And in the end all that training might have actually been useful, as problem solving was part of my job for the last three years. Last but not least thanks to my grandparents for just being their adorable selves. I love you all.

# Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet habe.

Weiterhin erkläre ich, dass ich die Hilfe von gewerblichen Promotionsberatern bzw. Promotionsvermittlern oder ähnlichen Dienstleistern weder bisher in Anspruch genommen habe, noch künftig in Anspruch nehmen werde.

Zusätzlich erkläre ich hiermit, dass ich keinerlei frühere Promotionsversuche unternommen habe.

.....  
Ort, Datum

.....  
Sophie Hermann