On the mathematical foundations of the statistical analysis in physics

Dissertation
an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Paula Reichert-Schürmer

München, den 24.05.2018

On the mathematical foundations of the statistical analysis in physics

Paula Reichert-Schürmer

Erstgutachter: Prof. Dr. Detlef Dürr
Zweitgutachter: Prof. Dr. Stefan Teufel
Termin der Verteidigung: 27.07.2018

Abstract

The work of Ludwig Boltzmann on statistical mechanics has shown that a statistical analysis of subsystems of the universe needs to be embedded within a statistical analysis of the universe which above all is needed to explain the origin of low-entropy initial states. I aim to provide such an analysis. The main advantage of this account as compared to standard explanations is that it does not invoke a Past Hypothesis, i.e., it works without the assumption of a very special (atypical) state at the beginning of the universe. Instead everything is typical.

To obtain this explanation, I relate a proposal of Carroll and Chen [2004], [2005] to the recent work of Barbour, Koslowski and Mercati [2013], [2015]. To draw the connection I introduce a notion of entropy for the Newtonian universe (which is a model of N particles moving through infinite space thereby attracting each other according to the Newtonian gravitational force law). I show that, with respect to this notion of entropy, the Newtonian universe is a Carroll-type universe, featuring a U-shaped entropy curve. This explains the observation of an entropy gradient, but it leaves us with an non-normalizable measure which cannot provide a statistical analysis.

Next I deal with the statistical analysis. I show that the measure suggested by Barbour et al. is indeed the correct measure for the statistical analysis of the Newtonian universe. For that purpose I derive the formula which they use to construct the measure from the geometry of the underlying space, the space of physically distinct mid-point data. This space is obtained by reducing the standard phase space of the system with respect to the symmetries of translation and rotation, introducing an internal time parameter and finally using the dynamical similarity of the internal equations of motion. Once we got rid off all the redundant degrees of freedom, we are able to construct a normalizable volume measure in terms of which a statistical analysis can be made. From this we learn that typically at the point of minimal extension of the particles, the Big Bang of the Newtonian universe, the system is in a homogenous state, a state of low entropy. Hence, we got rid off the Past Hypothesis in the end.

Having determined the reduced internal dynamics of the Newtonian gravitational system, we are able to address another topic: the evolution through the points of total collision. Whereas on absolute phase space the physical vector field turns singular at the points of total collision, this does not happen on shape phase space. Instead, the shape degrees of freedom can be evolved uniquely through the points of total collision, determining a unique way to combine *two* solutions on absolute phase space – one which ends at and one which starts at a total collision – to form *one* trajectory passing the singularity.

Zusammenfassung

Die Dissertation beschäftigt sich mit der Frage nach dem Ursprung des Zweiten Hauptsatzes der Thermodynamik und der Begründung der statistischen Analyse in der Physik. Die Arbeiten von Ludwig Boltzmann haben gezeigt, dass eine statistische Analyse von Subsystemen in eine statistische Analyse des ganzen Universums eingebettet sein muss, welche insbesondere die Existenz von Anfangszuständen niedriger Entropie in Subsystemen erklärt. Ziel dieser Dissertation ist eine eben solche Analyse. Der entscheidende Vorteil dieser Darstellung im Vergleich zu anderen ist, dass sie ohne die so genannte „Past Hypothesis“ auskommt, also ohne die Annahme eines speziellen (untypischen) Zustandes zu Beginn des Universums. Stattdessen ist alles typisch.

Für dieses Erklärungsmodell verbinde ich einen Vorschlag von Carroll und Chen [2004], [2005] mit den Arbeiten von Barbour, Koslowski und Mercati [2013], [2015]. Um die Verbindung herzustellen, führe ich einen Entropiebegriff für das Newton'sche Universum (ein Modell von N Teilchen, die sich gemäß dem Newton'schen Gravitationsgesetz durch den unendlichen Raum bewegen) ein. Damit zeige ich, dass das Newton'sche Universum ein Universum im Sinne von Carroll ist, d.h., dass es einen U-förmigen Entropieverlauf aufweist. Dies begründet den Entropiegradienten, birgt aber das Problem eines nicht-normierbaren Maßes, welches keine statistische Analyse zulässt.

Als nächstes beschäftige ich mich mit der Frage nach der statistischen Analyse. Ich zeige, dass das Maß, das Barbour et al. vorschlagen, tatsächlich das richtige Maß für die statistische Analyse des Newton'schen Universums ist. Dafür leite ich die Formel, die sie zur Konstruktion des Maßes benutzen, aus der Geometrie des zugrundeliegenden Raumes ab, dem Raum der physikalisch unterscheidbaren „Mittelpunkts“-Zustände. Dies ist der Raum auf dem das Maß definiert ist. Man konstruiert ihn, indem man den gewöhnlichen Phasenraum des Systems bzgl. Translations- und Rotationssymmetrie reduziert, einen internen Zeitparameter einführt und zuletzt die so genannte dynamische Ähnlichkeit der internen Bewegungsgleichungen berücksichtigt. Beinhaltet die Beschreibung des Systems keine redundanten Freiheitsgrade mehr, kann man ein normierbares Maß konstruieren, mit dem sich eine statistische Analyse durchühren lässt. Aus ihr lernen wir, dass das System zum Zeitpunkt minimaler Ausdehnung der Teilchen, dem so genannten Big Bang des Newton'schen Universums, typischer Weise in einem homogenen Zustand ist. Dies ist zugleich ein Zustand niedriger Entropie. Damit brauchen wir die „Past Hypothesis“ nicht mehr.

Nachdem wir die reduzierten internen Bewegungsgleichungen des Newton'schen gravitierenden Systems hergeleitet haben, können wir ein weiteres Thema behandeln: die Dynamik durch die Punkte der Totalkollision aller Teilchen. Während das Vektorfeld auf dem absoluten Phasenraum an diesen Punkten singulär ist, ist dies auf dem „Shape“ Phasenraum nicht der Fall. Vielmehr können die konformen („Shape“) Freiheitsgrade eindeutig durch die Punkte der Totalkollision hindurch entwickelt werden, was eine Möglichkeit aufweist, zwei auf dem absoluten Phasenraum definierten Lösungen - eine, die mit einer Totalkollision aufhört, und eine, die mit einer Totalkollision anfängt - eindeutig zu *einer* Lösung zu verbinden, welche die Singularität durchläuft.

Acknowledgements

I thank my family – Tobi, Jakob and Clara – for their inestimable love and invaluable support.

I thank my professor, Detlef Dürr, for having been an exceptional supervisor and encouraging person, as well as for being a dedicated physicist.

I thank Dustin Lazarovici for many clarifying discussions and great encouragement. I thank Julian Barbour for sharing his beautiful ideas and Peter Pickl, Ward Struyve and Sahand Tokasi for many fruitful debates.

I thank my mother Monika and my mother in law Hanne for taking care of the children whenever needed. And I thank all my family, friends and the working group for their encouragement.

Contents

1 Introduction

Why does entropy increase or stay the same, but never decrease? Part of an answer to this question has been given by Boltzmann at the end of the 19th century. Boltzmann showed that if we understand a macroscopic system as being constituted of small particles, atoms or molecules, which move according to Newton's laws and if we look at the micro-evolution of the system starting from an initial non-equilibrium state, a state of low entropy, then we should by all reasonable means expect the system to be carried towards equilibrium, a state of high entropy, very quickly.

The reason for that is essentially that there are by far more microstates that realize an equilibrium macrostate than there are microstates that realize a non-equilibrium macrostate. And this does not refer to a proportion of one to a hundred or a thousand, but to an incredibly high number: a proportion of about $1 : 10^N$ where N is the number of particles involved, i.e. $N \approx 10^{23}$!

Hence, fundamental to the understanding of why entropy increases or stays the same, but never decreases, is the distinction between micro- and macro-description of the system. This involves the fact that different macrostates (which are defined by certain macrovariables like volume V, temperature T, and so on) partition phase space into regions – sets of microstates realizing the given macrostate – that differ not only a bit, but vastly in size with an equilibrium state that fills almost the entire phase space volume. Here the notion of "size" is given by the Liouville measure, the natural measure of phase space volume. Now, starting from a low-entropy state which corresponds to a tiny region in phase space, it follows almost directly from the dominance of the equilibrium state that almost all microstates realizing the low-entropy state evolve towards equilibrium rather quickly. For a thorough presentation of the underlying argument, which is also known as the typicality account, see Lebowitz [1981], [1993], Bricmont [1995], Goldstein [2001] or Lazarovici and Reichert [2015].

But there is a caveat to that argument: why should the system start from a low-entropy state if such a state is highly atypical? Typically (where typically refers to the Liouville measure) the system should be in equilibrium or at least close to equilibrium at any moment in time. But this is not what we observe.

When we consider a particular system like a gas in a box and try to trace back the origin of its low-entropy initial state, then we find another, bigger system of which the former system is merely a part (like, in our case, e.g., the box plus the device preparing the initial state of the gas), which has started from an initial state of even lower entropy further in the past. Only if it has started from a state of even lower entropy, the second law of thermodynamics stating that entropy increases throughout will not have been violated. This argument can be repeated on and on continuously enlarging the system under consideration – until we reach the universe as a whole. The universe as a whole is not a subsystem. It is all there is. Hence, we must conclude that the universe has started from a very special, low-entropy initial state and that we are still somewhere on the way towards equilibrium.

However, this line of reasoning is not strictly compelling. There is again a caveat and that is the following: to be honest, the only thing we can say is that the universe is in a state of low entropy *now*. But according to Boltzmann such a state is highly unlikely. Why then don't we conclude that the universe is at the bottom of a deep fluctuation out of equilibrium at this very moment? We can even sharpen this argument: Boltzmann's statistical reasoning tells us that entropy increases towards the future, but the very same reasoning also holds for the past. If we take it seriously, then we must conclude that we are at the bottom of a deep fluctuation at this very moment.

There is only one argument against this conclusion and this is that we want to believe that the past has actually existed. Of course, everything that exists at this very moment, among this the particular configuration of our brains including our memories and our knowledge about the past, might be a large fluctuation. This scenario is known as the *Boltzmann brain* scenario. But this is not what we like to think about our world. It is a solipsistic conception which, like any solipsistic account denying the existence of anything exterior of us, cannot be excluded by an argument or falsified by an experiment.[1] However, it is not a conception we want to adopt.

Another way to think about the *fluctuation hypothesis* is by presuming that we are indeed in a fluctuation, but, since we are sure that we have had a past, concluding that we are already on the way out of it. This is actually what Boltzmann had in mind (cf. Boltzmann [1896a]). He presumes both that time is eternal (which is what he believed to be true) and that there exists an equilibrium state from which the universe departs, occasionally moving into some small or (seldom) larger fluctuation. Being in one of these fluctuations, we are already on the way out of it at some distance away from the minimum, either on the way downwards or upwards both of which we cannot distinguish because we always conceive the past to be where the lower-entropy states are (for arguments for this, see Albert [2009]). This scenario was Boltzmann's explanation in [1896a] and I will refer to it as the Boltzmann model later on.

However, as Feynman [1967] pointed out, this reasoning is ridiculous. Assume that the universe is on the way out of a fluctuation. Then it follows from Boltzmann's statistical reasoning that the fluctuation is only as big as it has to be in order to account for all we know about the past. However, we still learn about the past, e.g. by finding dinosaur bones, and every time we learn about the past, we need to adjust the size of the fluctuation – the fluctuation must be even larger than we assumed before. This continuous adjustment is what Feynman calls ridiculous. He says there is only one reasonable way out and that is by positing a special state of very low entropy at the beginning of the universe. This is Feynman's proposal which has by now become *the* explanation of the arrow time (the thermodynamic asymmetry in time) and the second law of thermodynamics. There is a very nice drawing of it by Roger Penrose [2004] which shows God marking the initial state of the universe with a pin nail in one of the tiniest regions of phase space. And it was David Albert [2009] who coined the name "Past Hypothesis" under which the assumption of a low-entropy initial state of the universe is most commonly known today.

The question arises whether we can do better. Can we get rid of the Past Hypothesis which

[1] This does not say that there is no argument against solipsism, it just says that there is no argument which renders this conception impossible.

says that, at one moment in time – the Big Bang –, the universe has been in a highly unlikely state? Can we maybe find an explanation in which everything is typical? Is there a way to say that *typically* the second law of thermodynamics holds within subsystems of the universe?

There are mainly two proposals which answer this question affirmatively. One is the multiverse scenario proposed by Sean Carroll, the other the shape dynamical approach of Barbour, Koslowski, and Mercati. What they have in common is that they assume an eternal, overall time-symmetric evolution of the universe. There is neither a beginning nor an end. And there is no time asymmetry imposed from the very beginning. Instead, both proposals feature two arrows of time with one past in the middle and two futures in both directions away from it. However, where both approaches differ distinctly is with regards to the statistical analysis and the notion of entropy of the universe. This is where this thesis wants to go beyond, connecting both proposals, solving the remaining difficulties and obtaining a final explanation of why typically the second law holds and why typically there has been a state of lower entropy in the past. Before being more explicit about the aim and scope of the thesis, let me shortly outline the two proposals.

Carroll's proposal. The first serious attempt aimed at getting rid of the Past Hypothesis has been made by Sean Carroll. I call this the Carroll proposal or Carroll model. In his book [2010] and before that in an article with Jennifer Chen [2004] (cf. also Carroll and Chen [2005]), he claims that the second law of thermodynamics can be explained within a multiverse scenario on the basis of a particular evolution of the overall entropy (that is, the entropy of the multiverse). The particular shape of the entropy curve – a U-shape – provides the core of his argument. He claims that, given such a U-shaped entropy curve, the thermodynamic arrow of time (i.e., the fact that entropy increases, but never decreases) is typical and, even more, we don't need a Past Hypothesis. In what follows, I will call any model of the universe featuring a U-shaped entropy curve a Carroll-type universe.

Carroll's explanation is essentially grounded on the assumption that entropy can grow without bound. More specifically, it is based on a U-shaped overall entropy curve with a point of lowest, though arbitrarily high entropy in the middle and with entropy increasing without bound in both directions of time away from that. The increase of entropy determines an arrow of time. Hence, there are two arrows of time directed in opposite directions, while the overall picture is time-symmetric. That is, there is one past, namely at the point of lowest entropy, and there are two futures at the opposite ends of the entropy curve. Since the direction of lower entropy determines what we call the past, the fact that entropy increases towards the future holds at any point apart from the minimum. Let me say this again. There is an entropy gradient at any point of the entropy curve apart from the midpoint (which has measure zero). From this it follows that the increase of entropy is typical.

For Carroll, the overall entropy curve is connected to a multiverse scenario where baby universes arise from quantum fluctuations in Anti-De Sitter space, expand and after black holes have formed eventually evaporate again into almost empty Anti De Sitter space from which further quantum fluctuations give rise to new baby universes. This unbounded birth and decay process of universes allows Carroll to say that, wherever we are on the overall entropy curve, there is, at that moment, a universe similar to ours in a state similar to the one we experience at this very moment. Hence, we may be at any point on that curve. Moreover, Carroll argues, since the curve is unbounded from above, we will typically be somewhere, but not close to the minimum of the curve. This is why we experience to be far away from the minimum at this very moment.

So far this sounds nice, but Carroll's proposal has to deal with two difficulties. First, he does not propose any particular model, that is, he does not propose any particular dynamics nor does he give any particular definition of the overall entropy that would make it a U-shaped function with respect to time. He just says that there is *some* dynamics which makes the *somehow defined* entropy evolve that way. In fact, we will later show that within the $E = 0$ Newtonian universe the entropy evolves just the way Carroll has in mind. To be precise, there exists a sensible definition of entropy for the Newtonian universe (a Boltzmann entropy of the Newtonian gravitational system) such that, taking into account the dynamics of the Newtonian gravitational system, the entropy curve is a U-shaped function in time.

But even though we find a particular model for a Carroll-type universe, we are not done yet. The problem is that the Carroll model features a second difficulty which cannot be dissolved so easily. Given that the entropy is unbounded, which is a necessary assumption for the entropy curve to be U-shaped, the notion of entropy does not relate to a normalizable typicality measure. Entropy, the way it was understood by Boltzmann, is essentially a measure of phase space volume. Now if the entropy is unbounded, this presupposes that the total measure of phase space is infinite. Respectively, that the volume measure of phase space is non-normalizable.

But how can we then perform a statistical analysis of the system? Any regularization procedure with the purpose of rendering the measure normalizable, be it by imposing a cut-off or by conditioning, will lead to different results. Depending on the specific regularization, we may even come to opposite results what regards one and the same physical question. Here we are particularly interested in the question whether we are close to the minimum of the overall entropy curve or not. If we are typically close to the minimum, this in contradiction with observation and we need a Past Hypothesis to fix it. Given a non-normalizable measure, there is no unambiguous mathematical answer to the question whether we are typically close to the minimum of the overall entropy curve or not. Still, there is a way out following a different kind of (not purely mathematical) reasoning (cf. the proposal of Goldstein et al. [2016]).

At this point, Barbour, Koslowski, and Mercati add an essential ingredient to the discussion, suggesting a normalizable measure for the $E = 0$ Newtonian universe.

The account of Barbour, Koslowski, and Mercati. Why is there any question about the measure of typicality and/or the notion of entropy of the universe? When it comes to a system in which gravity is the dominant force, a so-called model universe, the notion of entropy is unclear. It is not clear which state is a state of high entropy and which is not. To my knowledge, there is only one drawing in Roger Penrose's famous book (cf. Penrose [2004]) proposing that the entropy of a gravitating system increases as the system evolves from a homogeneous state (like the Big Bang) to a dilute state of clusters (a state in which galaxies have formed). The reasoning behind this is simple: given that the entropy is at the same time a measure of the typicality or "likeliness" of a certain state, the observed evolution of the universe should be an evolution from an atypical state, a state of low entropy, towards a typical state, a state of high entropy.

Still, the notion of entropy of a gravitating system has so far not been given and this has to do with the following: the Boltzmann entropy of an isolated system like the universe is defined with respect to the microcanonical measure which, in case gravity is taken into account, is non-normalizable. Now, if the measure is non-normalizable, the entropy is not well-defined. We will show a way out of this dilemma and explain how we can still determine the entropy. However, there is a second problem: given that the measure is non-normalizable, we cannot perform the usual statistical analysis. In general, we cannot say which state is typical and which is not. This is where the proposal of Barbour et al. [2015] comes into play.

Barbour, Koslowski and Mercati [2015] were the first who succeeded in determining a normalizable measure, a measure of typicality, for a realistic model of the universe, the Newtonian universe. The crucial idea that led to their success was to define the measure not on full phase space, but on a lower-dimensional space, the space of physically distinct solutions. From this measure they compute an entropy-type quantity, the entaxy. The idea to define the measure on the space of solutions instead of defining it on standard phase space goes back to Gibbons, Hawking, and Stuart [1987]. It is based on an internal time parametrization. Barbour et al. go several steps further and, in addition to the internal time formulation, reduce phase space by several dimensions taking into account the symmetries of the system. This way they obtain the dynamics on the reduced phase space, eventually obtaining a description on shape phase space T^*S. Last but not least, they use the fact that different solutions can be run through by different speeds (called a mechanical or dynamical similarity of the system). Identifying these solutions, this leads to the space of physically distinct solutions PT^*S, which is a compact space. On that space, a normalizable measure can be defined. Even more, the measure can be obtained in a canonical manner from the original Liouville measure.

Once we have the measure, the reasoning is the following: The $E = 0$ Newtonian universe evolves in a certain manner due to the dynamical law. This is necessity. From the dynamics we already get a lot, namely we know that there is a point of minimal extension of the particles whereas the extension of the particles increases in both time directions away from that. This defines two arrows of time with one past in the middle at the so-called Janus point (the Big Bang) and two futures in both time directions away from that. The dynamics also tells us that as the universe expands galaxies and clusters of galaxies form. That way, effectively isolated subsystems with an asymptotically conserved energy relation come into existence and provide the setting in which standard thermodynamics can take place. All we still need the measure for is to statistically analyze the initial data – which, in this model, are really mid-point data/ data at the Janus point. The normalizable measure on the set of mid-point data allows us to make statistical assertions with regards to the macroscopic properties of the universe at that very moment, the moment of minimal extension of the particles, which we identify with the Big Bang.

Now typically (where typically refers to the normalizable measure over mid-point data) the distribution of particles in the universe is homogeneous at the Janus point! But this is just what observation tells us about the Big Bang! And, by the way, this is opposed to the idea of Penrose according to which a typical state is a clustered state whereas a homogeneous state is atypical. Hence, according to the measure on PT^*S, the universe didn't start from a very special, highly unlikely state, but just the opposite - it started from a typical state.

Still, the proposal of Barbour et al. lacks a notion of entropy of the Newtonian universe. This we need in order to explain the increase of entropy in subsystems of the universe, respectively the thermodynamic asymmetry in time. This is where this thesis will go beyond.

Aim and scope of this thesis. The first part of this thesis (Part I – Part III) explains why *typically* within subsystems of the $E = 0$ Newtonian universe entropy increases or stays the same, but never decreases and why *typically at this moment* we are far away from the Big Bang at which entropy has been far lower than it is now.[2] In order to show that the second law of thermodynamics and the low-entropy past are typical features of the Newtonian universe, we need to combine both the ideas of Carroll and Barbour et al. and introduce a notion of entropy for the Newtonian gravitational system.

The thesis also contains a second part (Part IV), which is more or less independent of the first. At least, it is not concerned with the notion of entropy and the statistical analysis of the universe. Still, it makes use of the formulation of the Newtonian dynamics on shape space (which we derive first in order to perform the statistical analysis). This formulation is convenient as it allows us to discuss the singularity of a total collision of the particles of the Newtonian gravitational system. While on absolute phase space the physical vector field is singular at the respective points, on shape phase space this is not the case. Explicitly, I show that the shape degrees of freedom can be evolved uniquely through the points of total collision. This evolution on shape phase space determines a unique way to combine two trajectories on absolute phase space – one which ends at and one which starts at a total collision – to form *one* trajectory passing the singularity.

The outline if the thesis is the following. In Section 2 we introduce the notions of entropy, (stationary) measures and typicality. We discuss the problems that arise in an "ad hoc" definition of the entropy of the Newtonian gravitational system and propose a more sophisticated definition instead. In Section 3 we show that the $E = 0$ Newtonian universe is a Carroll-type universe. This answers many questions, but we are left with the problem of non-normalizability of the measure, which sheds new light on the Past Hypothesis. In order to obtain a normalizable measure with respect to which an unambiguous statistical analysis can be performed, we develop the mathematical framework of the reduced and internal dynamics. This will constitute Sections 4 and 5. There we also derive the formula used by Barbour et al. to construct the normalizable measure on the space of physically distinct states. Sections 6 and 7 provide the statistical analysis, discuss the notions of entaxy and complexity and connect it to the notion of entropy

[2] The $E = 0$ Newtonian universe is a model of particles moving through infinite three-dimensional Euclidean space with total energy $E = 0$ attracting each other according to the Newtonian gravitational force law.

given in Section 2. As a synthesis an explanation of the second law of thermodynamics and the low-entropy past of our universe is obtained. Section 8 stands for its own as it proves that one can evolve the shape degrees of freedom through the points of total collision of the Newtonian gravitational system.

All sections contain results of my own with the main new results presented in Sections 2, 3, 5, 7 and 8. When I use the results of other people, I indicate this at the beginning of the respective section.

Part I

On the Notion of Entropy of the Newtonian Universe

2 How to define the entropy of the Newtonian universe

In this first part of this section, I will introduce the notion of a typicality measure. For further details, cf. Dürr and Teufel [2009], Dürr, Frömel, and Kolb [2017] and Lazarovici and Reichert [2015]. In the second part, I will discuss the technical problems that arise when you try to define a measure of typicality of the universe.

2.1 Introduction: measures of typicality and the notion of entropy

Let us consider a dynamical system: a measure space $(\Gamma, \mathcal{B}(\Gamma), \mu)$ together with a flow T on Γ. Here Γ is a set, $\mathcal{B}(\Gamma)$ the Borel algebra of measurable subsets of Γ and μ a measure on Γ. Note that this definition of a dynamical system is more general than the one commonly used. First, I consider a measure space, not a probability space, that is, the measure need not be normalizable. Second, the flow T need not be measure-preserving (although it will be measure-preserving in many important cases).

In case the flow is measure-preserving, we say that the measure is *invariant under the dynamics*, or *stationary*. To be able to define stationarity, we need the notion of the time-evolved measure μ_t.

Definition 2.1 (Time-evolved measure). Let $(\Gamma, \mathcal{B}(\Gamma), \mu)$ be a measure space and T^t, $t \in \mathbb{R}$, a one-parameter group of transformations on Γ. Let $A \in \mathcal{B}(\Gamma)$. Then

$$\mu_t(A) := \mu(T^{-t}A) \tag{2.1}$$

is the *time-evolved measure.*

In other words, the time-evolved measure μ_t of a set A is just the original measure μ of the original set $T^{-t}A$ (the pre-image of A under backwards time evolution). This equation, in fact, corresponds to the well-known continuity equation for the measure density.[3] Now we can define stationarity.

Definition 2.2 (Stationary measure). Let again $(\Gamma, \mathcal{B}(\Gamma), \mu, T)$ be a dynamical system. Let $A \in \mathcal{B}(\Gamma)$ and let μ_t be as defined in (2.1). A measure is called *stationary* if and only if, for all $t \in \mathbb{R}$,

$$\mu_t(A) = \mu(A). \tag{2.2}$$

[3]Cf. Dürr and Teufel [2009].

A measure is stationary if and only if the flow T is measure-preserving. With the help of (2.1), (2.2) can be rewritten as follows:

$$\mu(T^{-t}A) = \mu(A). \tag{2.3}$$

While Equation (2.1) corresponds to the continuity equation, Equation (2.3) corresponds to the Liouville equation for the measure density.[4]

We saw that stationary measures are invariant under the phase flow, respectively under time evolution. You may say that they behave nicely under the dynamics. But even more than that. A stationary measure is important as it allows us to compare different states of the system at different times. If the measure would change as time evolves, we would essentially not be able to perform a statistical analysis at all.

The importance of stationary measures has first been noticed by Ludwig Boltzmann.[5] Boltzmann was concerned with grounding the thermodynamic notion of entropy (the Clausius entropy) within a microscopic theory of matter. Assuming that matter is composed of atoms or molecules, he found that entropy S is basically the logarithm of the "number" of microstates X realizing a particular macrostate M. While a microstate X is determined by the positions and momenta of all the particles, a macrostate M is a thermodynamic state defined by certain thermodynamic variables like volume V, temperature T, and so on. Of course, $M = M(X)$. Whereas X is one point in phase space Γ, M defines an entire region $\Gamma_M \subset \Gamma$, the subset of all microstates X realizing the macrostate M. Hence, the "number" of microstates really refers to the phase space volume of the given set Γ_M. Now the Boltzmann entropy can be defined as follows.

Definition 2.3 (Boltzmann entropy)**.** Let $(\Gamma, \mathcal{B}(\Gamma), \mu, T)$ be a dynamical system. Let T be a measure-preserving transformation. For a microstate $X \in \Gamma$ and a macrostate $M(X)$ determining a region $\Gamma_M \subset \Gamma$, the *Boltzmann entropy* is

$$S(X) = k_B \log |\Gamma_M(X)|.$$

Here $|\Gamma_M| := \mu(\Gamma_M)$ and k_B is Boltzmann's constant.

Note that for this to be a sensible definition of entropy, the measure needs to be stationary – otherwise the entropy of a macrostate M would change as time evolves which is in contradiction with Clausius' notion of entropy. Now there exist many different stationary measures. In particular, the $6N$-dimensional Lebesgue measure is stationary. This means that phase space volume is conserved under time evolution (since, for N particles, phase space $\Gamma \cong \mathbb{R}^{6N}$). This measure of phase space volume is called the Liouville measure.

Definition 2.4 (Liouville measure)**.** Let q_i, p_i $(i = 1, ..., N)$ be local coordinates on $\Gamma \cong \mathbb{R}^{6N}$. Let $A \subset \Gamma$ be a measurable subset of Γ. Then

$$\mu(A) = \int_A \mathrm{d}\mu \tag{2.4}$$

[4] Cf. Dürr and Teufel [2009].
[5] Cf. Ehrenfest, P. and T. [1911].

with

$$\mathrm{d}\mu = \prod_{i=1}^{3N} \mathrm{d}q_i \mathrm{d}p_i = \mathrm{d}^{3N}q\mathrm{d}^{3N}p \tag{2.5}$$

is the *Liouville measure* or *volume* of the set $A \subset \Gamma$.

There are other stationary measures frequently used in statistical mechanics, like the microcanonical measure, the canonical measure, the grand canonical measure, and so on. Most important for us is the microcanonical measure which is the correct measure for any isolated system, respectively, for any system in which energy is conserved. Hence, it is also the correct measure for the universe. Since we later have to deal with non-normalizable measures, we give here the definition of the non-normalized microcanonical measure.

Definition 2.5 (Microcanonical measure)**.** Let again q_i, p_i be local coordinates on $\Gamma \cong \mathbb{R}^{6N}$. Let $H(q,p) := H(q_1, ..., p_N)$ a smooth function on Γ, the Hamiltonian of the system. Let $A \subset \Gamma$ be a measurable subset of Γ. Then $\mu_E(A) = \int_A \mathrm{d}\mu_E$ with

$$\mathrm{d}\mu_E = \frac{1}{N!h^{3N}} \prod_{i=1}^{3N} \delta(H - E)\mathrm{d}q_i \mathrm{d}p_i \tag{2.6}$$

is the *microcanonical measure* of the set $A \subset \Gamma$.

Note that due to the delta function this is a volume measure on the constant energy surface $\Gamma_E = \{(q,p) \in \Gamma | H(q,p) = E\}$. However, it is not the natural surface area measure ν_E on Γ_E. In fact, μ_E deviates from ν_E by the gradient of H:

$$\mu_E = \frac{\nu_E}{||\nabla H||}. \tag{2.7}$$

The reason behind this deviation is basically that, for different values of E, the curvature of the constant energy surface is different.[6]

So far we neglected normalization, but this does not mean that normalization is not important. Assume we have a normalizable, stationary measure. Let $\Gamma \cong \mathbb{R}^{6N}$, $A \in \mathcal{B}(\Gamma)$ and $(q,p) := (q_1, ..., q_N, p_1, .., p_N) \in \Gamma$. In what follows χ_A denotes the characteristic function of A, i.e. $\chi_A(q,p) = 1$ if (q,p) in A and 0 otherwise. While

$$\mu(A) = \int_\Gamma \chi_A(q,p)\mathrm{d}^{3N}q\mathrm{d}^{3N}p \tag{2.8}$$

determines the *total measure* or *volume* of the set $A \subset \Gamma$,

$$\sigma(A) = \frac{\mu(A)}{\mu(\Gamma)} = \frac{\int_\Gamma \chi_A(q,p)\mathrm{d}^{3N}q\mathrm{d}^{3N}p}{\int_\Gamma \mathrm{d}^{3N}q\mathrm{d}^{3N}p} \tag{2.9}$$

[6] For details, cf. Dürr and Teufel [2009]. See also the discussion of the Faddeev-Popov determinant Δ_{FP} later in this thesis. Δ_{FP} is introduced for the very same reason as $||\nabla H||$ (cf. the remark on the geometrical nature of Δ_{FP} in Sec. 4.3).

determines the *proportion* of the region A as compared to Γ. We say that A is *typical* if $\sigma(A) \approx 1$ and that A is *atypical* if $\sigma(A) \approx 0$. Here $\sigma(A)$ is a *typicality measure* of the set A.

In principle, σ can take values between 0 and 1. However, for a realistic physical system of about $N \approx 10^{23}$ particles and given that we partition phase space into macro-regions according to some (macroscopic) thermodynamic variables, we find that the typicality measure attains only values close to 0 or close to 1. To be precise, there will be one region, the equilibrium region, consisting of the by far largest part of phase space while all other regions (together) have negligible phase space volume. This dominance of the equilibrium state is essentially due to the large number of particles, $N \approx 10^{23}$ for a realistic physical system. For a thorough discussion of this point, see Boltzmann [1896b] and the references given at the beginning of this section.

The typicality measure tells us which state is (overwhelmingly) likely and which is not, which state is *typical* and which is not. And this statement does not depend on the exact form of the measure: if σ is a typicality measure, any other measure σ' which is absolutely continuous with respect to σ will provide the same notion of typicality. This is basically again due to the vast difference in size between the distinct macro-regions. As such the notion of a typicality measure really defines an equivalence of measures (all those which are absolutely continuous with respect to each other). That way it goes beyond the notion of a probability measure.

Asserting which state is (overwhelmingly) likely and which is not, the typicality measure tells us which state is realized in nature and which is not. This is Cournot's principle[7] which gives meaning to the notion of probabilities in physics: Nature realizes what has (probability or typicality) measure close to 1, while She does not realize what has (probability or typicality) measure close to 0.

In what follows, we want to find a typicality measure and determine the entropy of a gravitational system, the Newtonian model of the universe. Before that, let me draw attention to two issues we will come upon. First, what if phase space Γ is infinite? Can we statistically analyze the system in that case? We will discuss the problems regarding non-normalizable measures in Sec. 3.2.

Second, what if for the model under consideration there does not exist a stationary measure? Is there a possibility to perform a statistical analysis at a particular moment of time? What time would that be? And if there was a preferred moment of time, what would then be a good criterion for the choice of the measure, if not stationarity? Certainly, if the theory treats all states at the same footing, a uniform measure would be the natural choice. This can be grounded on the principle of sufficient reason, respectively on Laplace's principle:[8] there is no reason to prefer one state over the other. We will again deal with this question when we introduce the notion of entaxy in Section 7.2. For the moment, note that also the Liouville measure is a uniform measure on phase space.

---

[7] Cf. Cournot [1843].

[8] Cf. Laplace [1814].

2.2 Difficulties regarding the definition of entropy

Let us try to define the notion of entropy for a particular model of the universe, the Newtonian universe. By "Newtonian universe" I refer to a non-relativistic, Newtonian model of N particles moving through infinite, three-dimensional Euclidean space thereby attracting each other according to the Newtonian gravitational force law. What is stationary measure for this model, a measure of typicality, in terms of which the entropy can be defined?

Definitely, the universe as a whole is an isolated system. As such, energy is conserved. Consequently, the microcanonical measure should be the correct measure in order to statistically analyze the system and determine the entropy. However, when applied to the Newtonian universe, the microcanonical measure diverges (see below).

There are mainly two reasons for the measure's divergence. One has to do with the infinity of space leading to a divergence of the q-integral. The other has to do with the singularity of the Newton potential leading to a divergence of the p-integral. What regards the first source of divergence, this occurs for any spatially open model of the universe. Whenever space is infinite, the q-integral diverges. As a remedy you might propose the following: just pick a spatially closed model, which is a possible model of our universe as well.

So let us consider a closed universe. Let V be some finite volume within which the particles are confined, like, e.g., the unit three-sphere S^3. There is still a second source of divergence – a divergence of the p-integral – which cannot be treated so easily. Let

$$H(\mathbf{q},\mathbf{p}) = \sum_{i=0}^{N} \frac{\mathbf{p}_i^2}{2m} - \sum_{i<j,\ i,j=1}^{N} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|} \tag{2.10}$$

be the Hamiltonian of the Newtonian gravitational system, governing the motion of the particles. Then the following holds.[9]

Lemma 2.1. *Let the Hamiltonian $H = H(\boldsymbol{q},\boldsymbol{p})$ given by (2.10). Let V be some finite volume and $\Gamma_V \subset \Gamma$ the set of all points for which the particles are confined in V. Then the microcanonical measure*

$$\mu_E(\Gamma_V) = \frac{1}{N!h^{3N}} \int_{V^N} \mathrm{d}^{3N}q \int_{\mathbb{R}^{3N}} \mathrm{d}^{3N}p\ \delta(H(\boldsymbol{q},\boldsymbol{p}) - E) \tag{2.11}$$

diverges for $N \geq 3$:

$$\mu_E(\Gamma_V) = \infty \quad \textit{for } N \geq 3. \tag{2.12}$$

Proof. The p-integral can be computed. It is

$$\mu_E(\Gamma_V) = C' \int_{V^N} \mathrm{d}^{3N}q \Big(E + \sum_{i<j,\ i,j=1}^{N} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|} \Big)^{\frac{3N-2}{2}}.$$

[9] Cf. Padmanabhan [1990] for the proof and Kiessling [2001] and Heggie and Hut [2003] for a discussion of the result.

Now consider a change of variables. Instead of $\mathbf{q}_1$ we introduce $\mathbf{l} = \mathbf{q}_1 - \mathbf{q}_2$. Then we get

$$\mu_E(\Gamma_V) = C' \int_{V^{N-1}} \mathrm{d}\mathbf{q}_2...\mathrm{d}\mathbf{q}_N \ \mathcal{A}(\mathbf{q}_2, ..., \mathbf{q}_N)$$

where

$$\mathcal{A}(\mathbf{q}_2, ..., \mathbf{q}_N) = \int_V \mathrm{d}\mathbf{l} \Bigg(E + \frac{Gm^2}{|\mathbf{l}|} + \sum_{i=3}^{N} \frac{Gm^2}{|\mathbf{l} + \mathbf{q}_2 - \mathbf{q}_i|} + \sum_{i<j,\ i,j=2}^{N} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|} \Bigg)^{\frac{3N-2}{2}}.$$

This integral is divergent for all $N \geq 3$.

To see this consider the behavior of the integral $\mathcal{A}$ near zero. Near $\mathbf{l} = 0$, the main contribution comes from the $1/|\mathbf{l}|$-term. That is, the behavior of $\mathcal{A}$ is essentially as follows:

$$\int_0^\varepsilon \mathrm{d}l \ l^2 \left(\frac{Gm^2}{l} \right)^{\frac{3N-2}{2}} = Gm^2 \left(\frac{1}{\varepsilon} \right)^{\frac{3N-8}{2}} \longrightarrow \infty \quad \text{if} \quad \varepsilon \to 0.$$

Since $\mathcal{A}$ is divergent, it follows that $\mu_E(\Gamma_V)$ is divergent, too. □

We see that, even if absolute space is assumed to be finite, $V < \infty$, the singularity of the Newton potential leads to a divergence of the p-integral. It follows that the Boltzmannn entropy of the Newtonian gravitational system determined via the microcanonical measure,

$$S = k_B \ln \mu_E(\Gamma_V), \tag{2.13}$$

is infinite. Hence, entropy is not well-defined.

There is one way out of this dilemma choosing different macrovariables. Equation (2.13) essentially determines the phase space volume of macroscopic states (subsets of Γ) of constant total energy E and volume V. That is, E and V are the macrovariables with respect to which we compute the entropy of the system. We know from classical statistical mechanics that this is the correct choice for any isolated, non-gravitating system. We will, however, now give an argument why for a gravitating system we need to choose different macrovariables.

The next two sections present joint work with Dustin Lazarovici.[10] My contribution is, in particular, the proof of the lemmas and the main theorem.

2.3 New choice of macrovariables

In order to determine the entropy of the Newtonian universe with respect to absolute distances and velocities, the first trial was to start from Boltzmann's famous formula

$$S = k_B \ln \mu_E(\Gamma)$$

[10] The work with Dustin Lazarovici is based on private conversation in Munich in between 2015 and 2017 and on a working paper from Dustin Lazarovici and myself called "Entropy and Gravity" from 2017 (last version).

taking the microcanonical measure to compute the phase space volume (where we want to reconsider the original model of N particles moving through infinite space, that is $V = \mathbb{R}^{3N}$):

$$\mu_E(\Gamma) = \frac{1}{N!h^{3N}} \int_{\mathbb{R}^{6N}} \prod_{i=1}^{3N} \delta(H - E)\mathrm{d}q_i \mathrm{d}p_i.$$

We know that this is the correct measure for an isolated system in which energy is conserved. So it should be the correct measure for the Newtonian universe as well. However, we have already seen that if the Newton potential forms part of the Hamiltonian of the system,

$$H = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} - \sum_{i<j} \frac{Gm_i m_j}{|\mathbf{q}_i - \mathbf{q}_j|},$$

then both the position part (given that $V = \mathbb{R}^{3N}$) and the momentum part of the microcanonical phase space integral diverge. How do we handle these divergencies?

Everything here depends on the choice of macrovariables. While the total volume of the constant energy hypersurface is infinite, this doesn't have to be the case for particular macro-regions. In fact, once we choose the correct macrovariables, we find that phase space is decomposed into (infinitely many) macro-regions of finite measure.

Recall that for an isolated, non-gravitating system like the ideal gas in the box, the correct macrovariables are the volume V of the box and the total (= kinetic) energy of the system: $E = T$. Also within the Newtonian universe total energy E is conserved. Thus, we have to keep it as a macrovariable, respectively, we have to keep the microcanonical measure as the (correct) stationary measure in terms of which we later define the entropy. But what about volume V?

What regards the Newtonian universe, there is no box – instead, space is infinite. Of course, we could still take the volume V as a macrovariable referring to the finite volume (the hypothetical box) within which the N particles are contained at a given moment of time. But there is a problem with that and this is the following: when we integrate over V^N, we sum over *all* possible configurations of N particles distributed within the volume V. This includes configurations that fill this volume more or less homogeneously, but also configurations in which the particles occupy only a small fraction of V. This means, in other words, that what we compute is not the phase space volume corresponding to a macrostate in which the particles actually *occupy* a certain volume V, but rather the phase space volume corresponding to *all possible configurations* of the N particles *within the boundaries of* V.

For the ideal gas, this difference is negligible. The reason is that, in that case, almost the entire phase space is occupied by the gases equilibrium state, respectively, the configurations corresponding to a homogeneous distribution of the gas over the accessible volume. (Just think of the number of microstates which look macroscopically like a gas filling half a volume compared to the number of microstates which look macroscopically like a gas filling the entire volume, which is roughly $1 : 2^N$ with $N \approx 10^{23}$.) For the gravitating system, this is distinctly different because the spatial configurations are correlated with the kinetic energy, respectively, with the possible momentum configurations of the system. The closer the particles, the larger the kinetic energy.

Thus, a macrostate describing a system of small spatial extension is not necessarily one of small (or even negligible) phase space volume.

This implies that a) the total volume V is not a good macrovariable to describe a gravitating system and, more specifically, b) if we want to know whether the entropy of a gravitating system increases as the system "collapses", forming one or several clusters, we have to consider a macroscopic variable that allows us to distinguish between a more "concentrated" and a more "spread out" configuration.

Hence first, instead of taking volume V, we better consider the moment of inertia

$$I = \sum_{i=1}^{N} m\mathbf{q}_i^2 \tag{2.14}$$

as an intrinsic measure of the total spatial extension of all the particles.[11]

However, the moment of inertia I is still to coarse a variable to distinguish between, for example, a homogeneous distribution of particles and a concentrated cluster with a few residual particles far far away – which, in the case of gravitating systems, again amounts to very different phase space volumes due to the respective momentum configurations. This is the case because these configurations depend on the gravitational potential which is different for different spatial configurations, even for one and the same total spatial extension. In particular, two particles that are very close (or even arbitrarily close) to each other imply very high (or even arbitrarily high) momenta since the absolute value of the gravitational potential is then also very (or arbitrarily) high, thereby implying that the respective macrostate corresponds to a very big (or arbitrarily big) region in phase space. In order to be able to distinguish between a homogeneous and a clustered state (or a cold and a hot state, respectively), we will thus have to introduce a further variable thereby obtaining a higher "resolution" of macrostates. To this end, it is convenient to consider the system's potential energy U. This will lead to a partition of phase space into macro-regions of finite volume.

Thus, as an additional macrovariable, we choose the (minus the) potential energy of the gravitating system,

$$U = \sum_{\substack{i<j \\ i,j=1}}^{N} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|}. \tag{2.15}$$

You can think of U as a macrovariable reflecting how much the system is clustered. However, from $E = T - U$ it follows that a description in terms of E, I and U is equivalent to a description in terms of E, I and T where T is the total kinetic energy. But again, this is just analogous to the Boltzmann entropy of an isolated ideal gas: also in case of the ideal gas, the total kinetic energy is fixed simply due to the fact that the total energy $E = T$ is fixed. For a gravitating system, if we want to fix the total kinetic energy, or temperature, T we need a further macrovariable in addition to E. This is just the potential energy U.

[11] In fact, (2.14) is the moment of inertia in the center-of-mass frame. We can determine it in that frame without loss of generality since the system is invariant under spatial translations.

In what follows, we will therefore denote by

$$S = k_B \ln \mu_E(\Gamma_{U,I}) \tag{2.16}$$

with

$$\mu_E(\Gamma_{U,I}) = \frac{1}{N!h^{3N}} \int_{\mathbb{R}^{3N}} d^{3N}p \int_{\mathbb{R}^{3N}} d^{3N}q\ \delta\big(H(\mathbf{q},\mathbf{p}) - E\big)\delta\Big(\sum_{\substack{i<j\\ i,j=1}} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|} - U\Big)\delta\Big(\sum_{i=1}^{N} m\mathbf{q}_i^2 - I\Big). \tag{2.17}$$

the entropy of the Newtonian universe.

2.4 Phase space integral and entropy

Before we actually compute the phase space integral (2.17) and determine the entropy (2.16), the following preliminary considerations are due.

2.4.1 Preliminary considerations

We start with a well-known result about the moment of inertia.[12] Instead of taking the distances between the particles and the center of mass $|\mathbf{q}_i - \sum_i \mathbf{q}_i|$, the moment of inertia can also be expressed in terms of the inter-particle distances $|\mathbf{q}_i - \mathbf{q}_j|$. Let us, for simplicity, consider the equal mass case: $m_i = m$ $(i = 1, ..., N)$.

Lemma 2.2. *Let $\sum_{i=1}^{N} m\boldsymbol{q}_i = 0$ (the origin is fixed to the center of mass). Let $M = Nm$ denote the total mass of the N particles. Then the moment of inertia*

$$I = \sum_{i=1}^{N} m\Big(\boldsymbol{q}_i - \frac{1}{M}\sum_{i=1}^{N} m\boldsymbol{q}_i\Big)^2 \tag{2.18}$$

is given by

$$I = \frac{m}{N} \sum_{\substack{i<j\\ i,j=1}}^{N} |\boldsymbol{q}_i - \boldsymbol{q}_j|^2. \tag{2.19}$$

Proof. From the definition of I given by (2.18) and the fact that the origin is fixed to the center of mass, $\sum_{i=1}^{N} m\mathbf{q}_i = 0$, we get

$$I = \sum_{i=1}^{N} m\mathbf{q}_i^2$$

(cf. 2.14). Using this equation, it is

$$\sum_{\substack{i<j\\ i,j=1}}^{N} m^2|\mathbf{q}_i - \mathbf{q}_j|^2 = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} m^2(\mathbf{q}_i^2 + \mathbf{q}_j^2 - 2\mathbf{q}_i\mathbf{q}_j) = \frac{1}{2}(2MI - 2\sum_{i=1}^{N} m\mathbf{q}_i \sum_{j=1}^{N} m\mathbf{q}_j) = MI$$

[12]Cf. Saari [1971] for the result and proof.

Here we obtain the first equation from adding the terms with equal indices $i = j$. This can be done because they are zero anyway. Then we rewrite everything in terms of M and I. The last step uses again the fact that we are in the center-of-mass frame: $\sum_{i=1}^{N} m\mathbf{q}_i = 0$. With $M = Nm$, the assertion follows. $\square$

Let us now analyze the constraints on $\Gamma \cong \mathbb{R}^{6N}$ as formulated by the delta functions in (2.17). Note that due to the U and I constraints there is an upper and a lower bound on the distances between pairs of particles.

Lemma 2.3. *Let $(\boldsymbol{q}_1, ..., \boldsymbol{p}_N) \in \Gamma_{E,U,I}$ where $\Gamma_{E,U,I} \subset \Gamma$ is the subset of Γ fix E, U and I (determined via delta-functions). Then the inter-particle distances $|\boldsymbol{q}_i - \boldsymbol{q}_j|$ are bounded from above and below. That is, $\forall i \neq j$, $i, j = 1, .., N$,*

$$\frac{Gm^2}{U} \leq |\boldsymbol{q}_i - \boldsymbol{q}_j| \leq \sqrt{\frac{NI}{m}}. \tag{2.20}$$

Proof. From

$$U = \sum_{\substack{i<j \\ i,j=1}}^{N} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|}$$

it follows that

$$\inf_{\substack{(\mathbf{q}_i,\mathbf{q}_j) \\ (\mathbf{q},\mathbf{p})\in\Gamma_U}} |\mathbf{q}_i - \mathbf{q}_j| = \frac{Gm^2}{U}$$

where $\Gamma_U \subset \Gamma$ is the hypersurface of constant U, $\Gamma_U = \{(\mathbf{q}, \mathbf{p}) \in \Gamma | \sum_{i<j} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|} = U\}$. On the other hand, from I given by (2.19), that is,

$$I = \sum_{\substack{i<j \\ i,j=1}}^{N} \frac{m}{N} |\mathbf{q}_i - \mathbf{q}_j|^2,$$

it follows that

$$\sup_{\substack{(\mathbf{q}_i,\mathbf{q}_j) \\ (\mathbf{q},\mathbf{p})\in\Gamma_I}} |\mathbf{q}_i - \mathbf{q}_j| = \sqrt{\frac{NI}{m}}$$

where $\Gamma_I \subset \Gamma$ is the hypersurface of constant I, $\Gamma_I = \{(\mathbf{q}, \mathbf{p}) \in \Gamma | \sum_{i<j} \frac{m}{N} |\mathbf{q}_i - \mathbf{q}_j|^2 = I\}$.

Since the conservation of total energy E does not impose a constraint on the q-variables (only on the p-variables given that the q-variables are fixed) it follows that

$$\inf_{\substack{(\mathbf{q}_i,\mathbf{q}_j) \\ (\mathbf{q},\mathbf{p})\in\Gamma_U}} |\mathbf{q}_i - \mathbf{q}_j| = \inf_{\substack{(\mathbf{q}_i,\mathbf{q}_j) \\ (\mathbf{q},\mathbf{p})\in\Gamma_{E,U}}} |\mathbf{q}_i - \mathbf{q}_j|,$$

where $\Gamma_{E,U} \subset \Gamma$ is the hypersurface of constant E and U. An analogous relation holds for Γ_I and $\Gamma_{E,I}$ and the supremum of the inter-particle distances. Taking both conditions together, we get an upper and lower bound on all the $|\mathbf{q}_i - \mathbf{q}_j|$, $i \neq j$.

Let $(\mathbf{q},\mathbf{p}) = (\mathbf{q}_1, ..., \mathbf{q}_N, \mathbf{p}_1, ..., \mathbf{p}_N) \in \Gamma_{E,U,I}$ where $\Gamma_{E,U,I} \subset \Gamma$ is the hypersurface of fix E, U, and I. Then for all pairs of particles $(\mathbf{q}_i, \mathbf{q}_j)$ with $i \neq j$ it is

$$\frac{Gm^2}{U} \leq |\mathbf{q}_i - \mathbf{q}_j| \leq \sqrt{\frac{NI}{m}}.$$

□

We will later use that, in particular,

$$\frac{Gm^2}{U} \leq |\mathbf{q}_1 - \mathbf{q}_2| \leq \sqrt{\frac{NI}{m}}. \tag{2.21}$$

Of course, this bound is very crude. When we simultaneously fix U and I, there is no way for any of the $|\mathbf{q}_i - \mathbf{q}_j|$ to actually realize (nor, in general, come close to) one of the above bounds. This is due to the fact that the infimum of $|\mathbf{q}_i - \mathbf{q}_j|$ is attained if and only if, at the same time, all other inter-particle distances are infinite – which is excluded by the upper bound on the inter-particle distances. Analogously, the supremum is attained if and only if, at the same time, all other inter-particle distances are zero – which is excluded by the lower bound on these distances.

Let us further analyze the constraint surface. We want to determine the volume of the hypersurface of constant E, U, and I. Since the conservation of total energy E does not impose a constraint on the q-variables, we can start from configuration space $Q = \mathbb{R}^{3N}$ and consider the U and I constraints on that space. Each U and I separately determine a $(3N-1)$-dimensional hypersurface within $3N$-dimensional Q. What about the intersection of these two hypersurfaces?

Fixing U and I imposes two different constraints on the coordinates, so there exist three possibilities for the common constraint surface $\Gamma_{U,I}$. Either the two constraint surfaces (determined by U and I separately) do not intersect. Or they just "touch" each other and do not properly intersect. Or, and this is what we call the generic case, they do intersect and the surface of intersection is a $(3N-2)$-dimensional hypersurface Σ.

In order for the two surfaces to "touch" each other (except in isolated points of measure zero) the gradients ∇U and ∇I have to be parallel. This they are not.

Lemma 2.4. *Let U given by (2.15) and I given by (2.14). Then ∇U and ∇I are not parallel:*

$$\nabla U \nparallel \nabla I. \tag{2.22}$$

Proof. On the one hand, it is

$$\frac{\partial U}{\partial \mathbf{q}_i} = -\sum_{j \neq i,\ j=1}^{N} \frac{Gm^2(\mathbf{q}_i - \mathbf{q}_j)}{|\mathbf{q}_i - \mathbf{q}_j|^3}.$$

On the other hand,

$$\frac{\partial I}{\partial \mathbf{q}_i} = \frac{2m}{N} \sum_{j\neq i,\ j=1}^{N} (\mathbf{q}_i - \mathbf{q}_j).$$

We find that, in general, there does not exist a k such that $\forall i$: $\partial U/\partial \mathbf{q}_i = k \cdot \partial I/\partial \mathbf{q}_i$. Hence ∇U and ∇I are not parallel. $\square$

There remain two possibilities: either the two constraint surfaces intersect properly or they don't intersect at all. Of course, whether the two surfaces intersect at all depends on the values of I and U. For a given I, there exists a minimum value of U – only then there exist common solutions to the constraint equations – and vice versa. (The assertion that U has to be larger than a minimum follows directly from the fact that U is a function of the reciprocal inter-particle distances $|\mathbf{q}_i - \mathbf{q}_j|^{-1}$ which are bounded from below by $\sim 1/\sqrt{I}$).

Let, in what follows, U and I be such that there exist common solutions and the constraint surface is a $(3N-2)$-dimensional hypersurface.

Now what about the volume of that hypersurface? Within $Q = \mathbb{R}^{3N}$, the $(3N-1)$-dimensional hypersurface of fix I is a $(3N-1)$-sphere S_R^{3N-1} of radius $R = \sqrt{I/m}$. If one more dimension is "taken out" by fixing U, we end up with a "curve" on that sphere (a hypersurface of $3N-2$ dimensions). Unfortunately, there is no way to actually compute the volume $|\cdot|$ of that hypersurface. Instead, we want to estimate its volume in powers of the radius $R = \sqrt{I/m}$ of the sphere S_R^{3N-1}. To be precise, what we will find (cf. (2.37)) is that

$$\int_{\mathbb{R}^{3N}} d^{3N}q\ \delta\Big(\sum_{\substack{i<j \\ i,j=1}} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|} - U \Big) \delta\Big(\sum_{i=1}^{N} m\mathbf{q}_i^2 - I \Big) \sim |\Sigma| \tag{2.23}$$

where

$$|\Sigma| = \int \mathrm{d}\mathbf{q}_N ... \mathrm{d}\mathbf{q}_2 \int \mathrm{d}\phi_{\mathbf{q}_{12}} \int \mathrm{d}\theta_{\mathbf{q}_{12}} \delta\Big(\Big[\sum m\mathbf{q}_i^2 \Big]_{|\mathbf{q}_{12}|^*} - I \Big) \tag{2.24}$$

and $|\mathbf{q}_{12}|^* = h(\phi_{\mathbf{q}_{12}}, \theta_{\mathbf{q}_{12}}, \mathbf{q}_2, ..., \mathbf{q}_N)$ is a function of U and all the other q-variables. Now we want to say that $|\Sigma|$ is, for large N, approximately equal to λR^{3N} where $\lambda = \lambda(N)$ is some positive constant:

$$|\Sigma| \approx \lambda R^{3N}. \tag{2.25}$$

Where does this come from? To have a picture in mind imagine a two-sphere S_r^2 of radius r. Imagine we cut the sphere by a plane through the origin such that we end up with a great cycle. This is our constraint hypersurface. The volume (= length) of this cycle is $2\pi r$. Now assume we do not know exactly the form of the curve of intersection γ. Still, any one-dimensional curve which is not too different from a great circle has a volume of the order of the radius: $|\gamma| \sim r$. This is what we want to use (cf. (2.25)). In our example, when we estimate the length of the curve in terms of the radius r, then a) the curve must not bend too much (otherwise it is much larger than the radius: $|\gamma| >> r$) nor b) must it define a very small cycle (then it's much shorter than the radius: $|\gamma| << r$). Since we consider a space of many, many dimensions, $d \approx 3N/2$, this effect becomes even more pronounced and most of all possible curves γ (arbitrarily chosen) will be of a length in between the two extrema. In this sense, we want to argue that $|\gamma| \sim r$ is the generic case. As long as we have no reason to believe otherwise, this is what should be expected.

The following consideration sheds light on this issue from another perspective. For a given I and large N, the number of solutions (points on $Q = \mathbb{R}^{3N}$) for which one of the inter-particle distances $|\mathbf{q}_i - \mathbf{q}_j|$ contributes significantly to I, is negligibly small as compared to all solutions.

Lemma 2.5. *Let $Q \cong \mathbb{R}^{3N}$ and $\boldsymbol{q} := (\boldsymbol{q}_1, ..., \boldsymbol{q}_N)$ a point on Q. Let $\mathbb{P} = \int \mathrm{d}^{3N}q$ the Lebesgue measure (natural volume measure) on Q and $\mathcal{A} := \{\boldsymbol{q} \in Q | \sum_{i=1}^N \boldsymbol{q}_i^2 = I\}$. Let $\varepsilon > 0$. Then*

$$\frac{\mathbb{P}(\{\boldsymbol{q} \in \mathcal{A} | |\sum_{i=1}^N \boldsymbol{q}_i^2 - \sum_{j=3}^N \boldsymbol{q}_j^2| \geq \varepsilon\})}{\mathbb{P}(\mathcal{A})} \to 0 \text{ for } N \to \infty \tag{2.26}$$

Proof. On the one hand:

$$\begin{aligned}
& \mathbb{P}\Big(\Big\{\mathbf{q} \in \mathcal{A} | \Big|\sum_{i=1}^N \mathbf{q}_i^2 - \sum_{j=3}^N \mathbf{q}_j^2\Big| \geq \varepsilon\Big\}\Big) \\
= \ & \mathbb{P}\Big(\Big\{(\mathbf{q}_1, ..., \mathbf{q}_N) \in Q | \sum_{i=1}^N \mathbf{q}_i^2 = I \ \wedge (\mathbf{q}_1^2 + \mathbf{q}_2^2) \geq \varepsilon\Big\}\Big) \\
= \ & \int \mathrm{d}\mathbf{q}_1 \int \mathrm{d}\mathbf{q}_2 \mathbb{I}_{\{\varepsilon \leq \mathbf{q}_1^2 + \mathbf{q}_2^2 \leq I\}} \int \mathrm{d}\mathbf{q}_3 ... \mathrm{d}\mathbf{q}_N \mathbb{I}_{\{\mathbf{q}_3^2 + ... + \mathbf{q}_N^2 = I - \mathbf{q}_1^2 - \mathbf{q}_2^2\}} \\
= \ & \frac{1}{2}\Omega^{3N-7} \int \mathrm{d}\mathbf{q}_1 \int \mathrm{d}\mathbf{q}_2 \mathbb{I}_{\{\varepsilon \leq \mathbf{q}_1^2 + \mathbf{q}_2^2 \leq I\}} (I - \mathbf{q}_1^2 - \mathbf{q}_2^2)^{\frac{3N-8}{2}} \\
= \ & \frac{1}{2}\Omega^{3N-7}\Omega^5 \int_{\sqrt{\varepsilon}}^{\sqrt{I}} \mathrm{d}r r^5 (I - r^2)^{\frac{3N-8}{2}} \\
\leq \ & \frac{1}{2}\Omega^{3N-7}\Omega^5 I^2 \int_{\sqrt{\varepsilon}}^{\sqrt{I}} \mathrm{d}r r (I - r^2)^{\frac{3N-8}{2}} \\
= \ & \frac{1}{2}\Omega^{3N-7}\Omega^5 I^2 \frac{1}{3N-6} (I - \varepsilon)^{\frac{3N-6}{2}}.
\end{aligned} \tag{2.27}$$

On the other hand:

$$\mathbb{P}(\mathcal{A}) = \mathbb{P}(\{\mathbf{q}_1, ..., \mathbf{q}_N\} \in \Gamma | \sum_{i=1}^N \mathbf{q}_i^2 = I) = \int d^{3N}q \mathbb{I}_{\{\mathbf{q}_1^2 + ... + \mathbf{q}_N^2 = I\}} = \frac{1}{2}\Omega^{3N-1} I^{\frac{3N-2}{2}} \tag{2.28}$$

For large N, the fraction of both terms is small. Explicitly, for large N, this fraction is

$$\frac{\Omega^{3N-7}\Omega^5}{\Omega^{3N-1}} \frac{I^2}{3N} \left(1 - \frac{\varepsilon}{I}\right)^{\frac{3N}{2}} \to 0 \text{ for } N \to \infty.$$

This shows the assertion. □

To interpret the lemma, remember the geometric picture from the beginning. There we saw that the $(3N-1)$-dimensional sphere S_R^{3N-1} of radius $R = \sqrt{I/m}$ is intersected and the surface of intersection is a $(3N-2)$-dimensional "curve" on that sphere. What we do in this lemma (cf. (2.26)) is that we take out even more, namely 6 dimensions by taking out $\mathbf{q}_1$ and $\mathbf{q}_2$ and consider a hypersurface of $3N-7$ dimensions. Lemma 2.5 now says that for almost all configurations $\mathbf{q} \in Q$ the contribution of the variables $\mathbf{q}_1$ and $\mathbf{q}_2$ to I is negligible and, hence, the volume of that hypersurface is about the volume of the sphere which is about $\sim I^{3N/2}$. This implies that, for almost all constraints $|\mathbf{q}_1-\mathbf{q}_2| = |\mathbf{q}_1-\mathbf{q}_2|^*$ (with $|\mathbf{q}_1-\mathbf{q}_2|^*$ arbitrary and where $|\mathbf{q}_1-\mathbf{q}_2|^*$ may be a function of all the other coordinates) determining the $(3N-2)$-dimensional "curve" on $S^{3N-1}_{\sqrt{I/m}}$, the volume of the curve is approximately the volume of the sphere. Of course, this is not a proof when it comes to the *particular* curve determined by fixing U. But it gives us another reason to believe that this particular curve, i.e., the intersection surface Σ, can be approximated by the volume of the sphere $S^{3N-1}_{\sqrt{I/m}}$. Hence, it is reasonable to assume that, for large N, $|\Sigma| \approx \lambda R^{3N}$ with $R = \sqrt{I/m}$.

2.4.2 Result

Theorem 2.1 (Phase space integral)**.** *Let $Q = \mathbb{R}^{3N}$ and $\Gamma = T^*Q \cong \mathbb{R}^{6N}$. Let E be the total energy, U minus the potential energy, T the kinetic energy and I the moment of inertia as above. Let U and I be such that they determine a $(3N-2)$-dimensional hypersurface $\Sigma \subset Q$ given by (2.24) of volume*

$$|\Sigma| \approx \lambda\big(\sqrt{I/m}\big)^{3N}. \tag{2.29}$$

Here N is large and λ is some positive constant. Then the microcanonical measure of a hypersurface of fix U and I,

$$\mu_E(\Gamma_{U,I}) = \frac{1}{N!h^{3N}} \int d^{3N}p \int d^{3N}q \; \delta\big(H(\boldsymbol{q},\boldsymbol{p}) - E\big)\delta\Big(\sum_{\substack{i<j\\ i,j=1}} \frac{Gm^2}{|\boldsymbol{q}_i - \boldsymbol{q}_j|} - U\Big)\delta\Big(\sum_{i=1}^{N} m\boldsymbol{q}_i^2 - I\Big), \tag{2.30}$$

is, for large N, bounded by

$$\frac{C}{Gm^2}(E+U)^{\frac{3N-2}{2}}|\Sigma|\left(\frac{Gm^2}{U}\right)^4 \leq \mu_E(\Gamma_{U,I}) \leq \frac{C}{Gm^2}(E+U)^{\frac{3N-2}{2}}|\Sigma|\left(\frac{NI}{m}\right)^2 \tag{2.31}$$

with $C = \frac{1}{N!h^{3N}} m\Omega^{3N-1}(2m)^{3N/2-1}$.

Proof. For later purposes, let us include within the integral a characteristic function $\mathbb{I}$ expressing the bounds on $|\mathbf{q}_1 - \mathbf{q}_2|$ given by (2.21). As we have seen, for points on Γ of fix E, U, and I, this condition must be fulfilled anyway. Thus, by inserting the respective characteristic function there is no change to the integral:

$$\begin{aligned}
\mu_E(\Gamma_{U,I}) &= \frac{1}{N!h^{3N}} \int \mathrm{d}^{3N}p \int \mathrm{d}^{3N}q \; \delta\big(H(\mathbf{q},\mathbf{p}) - E\big)\delta\Big(\sum_{i<j} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|} - U\Big)\delta\Big(\sum_i m\mathbf{q}_i^2 - I\Big)\\
&= \frac{1}{N!h^{3N}} \int \mathrm{d}^{3N}p \int \mathrm{d}^{3N}q \; \delta(H-E)\delta\Big(\sum_{i<j} - U\Big)\delta\Big(\sum_i - I\Big)\mathbb{I}_{\left\{\frac{Gm^2}{U} \leq |\mathbf{q}_1 - \mathbf{q}_2| \leq \sqrt{\frac{NI}{m}}\right\}}.
\end{aligned}$$

Here $\sum_{i<j} := \sum_{i<j} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|}$ and $\sum_i := \sum_i m\mathbf{q}_i^2$. At this point, the p-integral can be computed and we have

$$C \int \mathrm{d}^{3N} q \Big(E + \sum_{i<j} \Big)^{\frac{3N-2}{2}} \delta\Big(\sum_{i<j} - U \Big) \delta\Big(\sum_i - I \Big) \mathbb{I}_{\left\{ \frac{Gm^2}{U} \le |\mathbf{q}_1 - \mathbf{q}_2| \le \sqrt{\frac{NI}{m}} \right\}} \tag{2.32}$$

where $C = 1/2 \frac{1}{N! h^{3N}} \Omega^{3N-1} (2m)^{3N/2-1}$. Here Ω^{3N-1} denotes the $(3N-1)$-dimensional volume element of the sphere.

In what follows, let us rewrite the integral. Let us make a transformation of variables from $\mathbf{q}_1$ to $\mathbf{q}_1 - \mathbf{q}_2$. Notice that the determinant of the Jacobian is equal to 1: $\det\left(\frac{\partial(\mathbf{q}_1 - \mathbf{q}_2)}{\partial \mathbf{q}_1}\right) = 1$. Hence, the integral (2.32) can be rewritten as

$$C \int \mathrm{d}\mathbf{q}_N \ldots \mathrm{d}\mathbf{q}_2 \int \mathrm{d}(\mathbf{q}_1 - \mathbf{q}_2) \Big(E + \sum_{i<j} \Big)^{\frac{3N-2}{2}} \delta\Big(\sum_{i<j} - U \Big) \mathbb{I}_{\left\{ \frac{Gm^2}{U} \le |\mathbf{q}_1 - \mathbf{q}_2| \le \sqrt{\frac{NI}{m}} \right\}} \delta\Big(\sum_i - I \Big). \tag{2.33}$$

Let us compute the $(\mathbf{q}_1 - \mathbf{q}_2)$-integral separately. Using polar coordinates,

$$\begin{aligned} &\int \mathrm{d}(\mathbf{q}_1 - \mathbf{q}_2) \Big(E + \sum_{i<j} \Big)^{\frac{3N-2}{2}} \delta\Big(\sum_{i<j} - U \Big) \mathbb{I}_{\left\{ \frac{Gm^2}{U} \le |\mathbf{q}_1 - \mathbf{q}_2| \le \sqrt{\frac{NI}{m}} \right\}} \delta\Big(\sum_i - I \Big) \\ = &\int \mathrm{d}\phi_{\mathbf{q}_{12}} \int \mathrm{d}\theta_{\mathbf{q}_{12}} \int \mathrm{d}|\mathbf{q}_1 - \mathbf{q}_2| |\mathbf{q}_1 - \mathbf{q}_2|^2 \Big(E + \sum_{i<j} \Big)^{\frac{3N-2}{2}} \delta\Big(\sum_{i<j} - U \Big) \mathbb{I}_{\{.\}} \delta\Big(\sum_i - I \Big) \end{aligned} \tag{2.34}$$

with $\mathbb{I}_{\{.\}} = \mathbb{I}_{\left\{ \frac{Gm^2}{U} \le |\mathbf{q}_1 - \mathbf{q}_2| \le \sqrt{\frac{NI}{m}} \right\}}$.

Now we can evaluate the delta-function fixing U by integration over the variable $|\mathbf{q}_1 - \mathbf{q}_2|$. Notice that due to $\delta(f(x)) = \delta(x)/|f'(x)|$, this leads to an additional factor of $|\mathbf{q}_1 - \mathbf{q}_2|^2$ in the numerator. Explicitly, $\delta(\sum_{i<j} - U)$ can be rewritten as follows:

$$\delta\Bigg(\frac{Gm^2}{|\mathbf{q}_1 - \mathbf{q}_2|} + \sum_{\substack{i<j \\ (i,j) \neq (1,2)}} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|} - U \Bigg) = \frac{|\mathbf{q}_1 - \mathbf{q}_2|^2}{Gm^2} \, \delta\Bigg(|\mathbf{q}_1 - \mathbf{q}_2| - \frac{Gm^2}{U - \sum \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|}} \Bigg). \tag{2.35}$$

Define $|\mathbf{q}_{12}| := |\mathbf{q}_1 - \mathbf{q}_2|$ and let $g(|\mathbf{q}_{12}|, \phi_{\mathbf{q}_{12}}, \theta_{\mathbf{q}_{12}}, \mathbf{q}_2, ..., \mathbf{q}_N) := |\mathbf{q}_{12}| - Gm^2 \big(U - \sum \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|} \big)^{-1}$. Let $|\mathbf{q}_{12}|^* = h(\phi_{\mathbf{q}_{12}}, \theta_{\mathbf{q}_{12}}, \mathbf{q}_2, ..., \mathbf{q}_N)$ be a solution to $g = 0$: $g(|\mathbf{q}_{12}|^*, \phi_{\mathbf{q}_{12}}, \theta_{\mathbf{q}_{12}}, \mathbf{q}_2, ..., \mathbf{q}_N) = 0$.

Now evaluate the delta-function fixing U by integration over $|\mathbf{q}_{12}|$. Then (2.34) turns into

$$\frac{1}{Gm^2} (E+U)^{\frac{3N-2}{2}} \int \mathrm{d}\phi_{\mathbf{q}_{12}} \int \mathrm{d}\theta_{\mathbf{q}_{12}} (|\mathbf{q}_{12}|^*)^4 \, \mathbb{I}_{\left\{ \frac{Gm^2}{U} \le |\mathbf{q}_{12}|^* \le \sqrt{\frac{NI}{m}} \right\}} \delta\Bigg(\Big[\sum_i \Big]_{|\mathbf{q}_{12}| = |\mathbf{q}_{12}|^*} - I \Bigg). \tag{2.36}$$

At this point, let us return to the full integral (2.33) – the phase space integral we started with at the beginning. Inserting (2.36) in (2.33), we get

$$\frac{C}{Gm^2}(E+U)^{\frac{3N-2}{2}}\int d\mathbf{q}_N...d\mathbf{q}_2\int \mathrm{d}\phi_{\mathbf{q}_{12}}\int \mathrm{d}\theta_{\mathbf{q}_{12}}\big(|\mathbf{q}_{12}|^*\big)^4\ \mathbb{I}_{\{|\mathbf{q}_{12}|^*\}}\delta\bigg(\Big[\sum_i\Big]_{|\mathbf{q}_{12}|^*}-I\bigg) \tag{2.37}$$

where $\mathbb{I}_{\{|\mathbf{q}_{12}|^*\}}=\mathbb{I}_{\left\{\frac{Gm^2}{U}\le|\mathbf{q}_{12}|^*\le\sqrt{\frac{NI}{m}}\right\}}$.

We can now determine an upper and lower bound on this integral:

$$\begin{aligned}&\int d\mathbf{q}_N...d\mathbf{q}_2\int \mathrm{d}\phi_{\mathbf{q}_{12}}\int \mathrm{d}\theta_{\mathbf{q}_{12}}\big(|\mathbf{q}_{12}|^*\big)^4\ \mathbb{I}_{\left\{\frac{Gm^2}{U}\le|\mathbf{q}_{12}|^*\le\sqrt{\frac{NI}{m}}\right\}}\ \delta\bigg(\Big[\sum_i\Big]_{|\mathbf{q}_{12}|^*}-I\bigg)\\ \le\ &\Big(\frac{NI}{m}\Big)^2\int d\mathbf{q}_N...d\mathbf{q}_2\int \mathrm{d}\phi_{\mathbf{q}_{12}}\int \mathrm{d}\theta_{\mathbf{q}_{12}}\ \delta\bigg(\Big[\sum_i\Big]_{|\mathbf{q}_{12}|^*}-I\bigg).\end{aligned} \tag{2.38}$$

Analogously,

$$\begin{aligned}&\int d\mathbf{q}_N...d\mathbf{q}_2\int \mathrm{d}\phi_{\mathbf{q}_{12}}\int \mathrm{d}\theta_{\mathbf{q}_{12}}\big(|\mathbf{q}_{12}|^*\big)^4\ \mathbb{I}_{\left\{\frac{Gm^2}{U}\le|\mathbf{q}_{12}|^*\le\sqrt{\frac{NI}{m}}\right\}}\ \delta\bigg(\Big[\sum_i\Big]_{|\mathbf{q}_{12}|^*}-I\bigg)\\ \ge\ &\Big(\frac{Gm^2}{U}\Big)^4\int d\mathbf{q}_N...d\mathbf{q}_2\int \mathrm{d}\phi_{\mathbf{q}_{12}}\int \mathrm{d}\theta_{\mathbf{q}_{12}}\ \delta\bigg(\Big[\sum_i\Big]_{|\mathbf{q}_{12}|^*}-I\bigg).\end{aligned} \tag{2.39}$$

Let us analyze what is left. The remaining integral is a high-dimensional phase space integral which resembles much the surface integral connected to the $(3N-1)$-dimensional hypersurface of fix I. The only difference is that, in the above case, $|\mathbf{q}_1-\mathbf{q}_2|=|\mathbf{q}_{12}|^*$ where $|\mathbf{q}_{12}|^*$ is a given function $h=h(\phi_{\mathbf{q}_{12}},\theta_{\mathbf{q}_{12}},\mathbf{q}_2,...,\mathbf{q}_N)$ of U and all the other coordinates. Geometrically, the remaining integral determines the volume of the intersection surface Σ which is determined by fixing I and U. We can now use the assumption from the beginning according to which Σ is a $(3n-2)$-dimensional hypersurface with volume $|\Sigma|\approx\lambda R^{3N}$ (for large N, $R=\sqrt{I/m}$ and some positive constant $\lambda=\lambda(N)$).

Hence, by assumption (cf. (2.29))

$$\int \mathrm{d}\mathbf{q}_N...\mathrm{d}\mathbf{q}_2\int \mathrm{d}\phi_{\mathbf{q}_{12}}\int \mathrm{d}\theta_{\mathbf{q}_{12}}\delta\bigg(\Big[\sum\Big]_{|\mathbf{q}_{12}|^*}-I\bigg)\approx\lambda R^{3N}$$

where $\lambda=\lambda(N)$ and $R=\sqrt{I/m}$.

Putting everything back together, we get the following bounds on the full phase space integral. For large N,

$$\frac{C}{Gm^2}(E+U)^{\frac{3N-2}{2}}|\Sigma|\Big(\frac{Gm^2}{U}\Big)^4\le\mu_E(\Gamma_{U,I})\le\frac{C}{Gm^2}(E+U)^{\frac{3N-2}{2}}|\Sigma|\Big(\frac{NI}{m}\Big)^2$$

where

$$|\Sigma|\approx\lambda\Big(\frac{I}{m}\Big)^{\frac{3N}{2}}.$$

This shows the assertion.

□

From this result, the following corollary can be obtained.

Corollary 2.1 (Entropy of the Newtonian universe)**.** *Let* $S = k_B \ln \mu_E(\Gamma_{U,I})$ *the Boltzmann entropy of the Newtonian universe. Then, for large* N*,*

$$S \approx \frac{3N}{2} k_B \ln(E+U) + \frac{3N}{2} k_B \ln I + S'(N), \tag{2.40}$$

where $S'(N)$ *depends on* N *and the other constants, but not on* E*,* U*, or* I*.*

Proof. In Theorem 2.1 we found positive constants f_1 and f_2 such that, for large N,

$$f_1 C(E+U)^{\frac{3N}{2}} |\Sigma| \leq \mu_E(\Gamma_{U,I}) \leq f_2 C(E+U)^{\frac{3N}{2}} |\Sigma| \tag{2.41}$$

where $|\Sigma| \approx \lambda(I/m)^{\frac{3N}{2}}$. Here $f_1 = (Gm^2)^3 U^{-4}$ and $f_2 = (Gm^2)^{-1}(NI)^2 m^{-2}$, $\lambda = \lambda(N)$ and $C = \frac{1}{N! h^{3N}} m\Omega^{3N-1}(2m)^{3N/2-1}$. From this we conclude that, for large N,

$$\mu_E(\Gamma_{U,I}) \approx C(E+U)^{\frac{3N}{2}} \lambda(I/m)^{\frac{3N}{2}}. \tag{2.42}$$

Hence,

$$S = k_B \ln \mu_E(\Gamma_{U,I}) \approx \frac{3N}{2} \ln(E+U) + \frac{3N}{2} \ln I + S'(N)$$

where $S'(N)$ depends on N, but not on E, U, and I. □

We see that, once we suitably adapt the classical macrovariables (that is, in particular, volume V) to the case of gravitation, the entropy of the Newtonian universe is well-defined. It is defined with respect to a stationary measure, the microcanonical measure, it is finite and its macrovariables are analogous to the macrovariables of a non-gravitating system. Moreover, it captures well our intuition about the gravitating system. The entropy given by (2.40) is proportional to I and U. Hence, it increases as the system expands (leading to an increase in the configurational part of the phase space integral) and as it forms clusters (leading to an increase in the momentum part of the phase space integral).

In what follows we will show that, due to the dynamics, the entropy of the $E = 0$ Newtonian universe is a U-shaped function of time. As such the $E = 0$ Newtonian universe is a Carroll-type universe, respectively, it is an example for the model Carroll has in mind.

3 The $E = 0$ Newtonian universe as a Carroll-type universe

Let us consider the $E = 0$ Newtonian universe. From (2.42) together with $E = 0$ it follows that the microcanonical measure of states of fix U and I is

$$\mu_E(\Gamma_{U,I}) \approx C\lambda m^{-\frac{3N}{2}} (U \cdot I)^{\frac{3N}{2}}. \tag{3.1}$$

Hence, the entropy of the $E = 0$ Newtonian universe $S = k_B \ln \mu_E(\Gamma_{U,I})$ is

$$S \approx \frac{3N}{2} \ln U + \frac{3N}{2} \ln I + S'(N), \tag{3.2}$$

where $S'(N)$ depends on N and the other constants, but not on U or I.

In what follows, we will show that the $E = 0$ Newtonian universe is precisely a Carroll-type universe. Explicitly, given the formula for the entropy (Eq. (3.2)) together with the dynamics of the $E = 0$ Newtonian universe, it will turn out that the universal entropy curve is a U-shaped function in time.

At this point be aware that U and I are macrovariables which change as the system evolves in time: $U(t) = U(\mathbf{q}_1(t), ..., \mathbf{q}_N(t))$, $I(t) = I(\mathbf{q}_1(t), ..., \mathbf{q}_N(t))$. Hence, the entropy is a function of time, $S = S(t)$, governed by the time evolution of U and I.

3.1 Evolution of the entropy due to the dynamics

To obtain the time evolution of U and I, we have to analyze the dynamics. The dynamics of the Newtonian gravitational N-body system is governed by the Lagrange-Jacobi equation.[13]

Lemma 3.1 (Lagrange-Jacobi equation)**.** *Let $I = \sum m_i \boldsymbol{q}_i^2$ be the center-of-mass moment of inertia, E the total energy and U minus the potential energy as defined above. Then*

$$\ddot{I} = 4E + 2U. \tag{3.3}$$

Proof. The proof of this equation can be found in Appendix C. □

The Lagrange-Jacobi equation provides a first classification of motion of the Newtonian gravitational system. In particular, it states that if $E = 0$, then, since $U > 0$,

$$\ddot{I} > 0. \tag{3.4}$$

This means that $I(t)$ is concave upwards. In addition, there exists a result by Pollard [1967] on the asymptotic behavior of I. It tells us that, for $E = 0$, $I \to \infty$ as $t \to \pm\infty$. Since I is strictly positive, this means that there exists a positive, global minimum: $I = I_{min}$.

Hence, for the $E = 0$ Newtonian universe the following scenario is due. At some moment τ of time the moment of inertia is minimal, $I = I_{min}$, whereas I increases in both time directions away from that.

Due to the results of Saari [1971] and Marchal and Saari [1974] we have an even more precise idea of the asymptotic behavior of the gravitational N-particle system. Saari [1971] studies the inter-particle distances $|\mathbf{q}_i - \mathbf{q}_j|$ ($i \neq j$; $i, j = 1, ..., N$) of the Newtonian gravitational system as $t \to \infty$, independent of the total energy of the system. He shows that, in the absence of oscillatory and pulsating motion,[14] the Newtonian gravitational system is quite well-behaved.

[13] This equation has been found by Lagrange and Jacobi at the end of the 18th century, cf. Moeckel [2007] for a historical introduction.

[14] For the notion of "oscillatory" and "pulsating" and the cited result, cf. Saari [1971].

To be precise, if pulsating and oscillatory motion is excluded, then either

$$|\mathbf{q}_i - \mathbf{q}_j| \sim C_{ij} t \tag{3.5}$$

or

$$|\mathbf{q}_i - \mathbf{q}_j| \approx t^{2/3} \tag{3.6}$$

or

$$|\mathbf{q}_i - \mathbf{q}_j| = \mathcal{O}(1). \tag{3.7}$$

Here C_{ij} is some positive constant. Since the Newtonian dynamics is time-reversal invariant, this result holds for $t \to -\infty$ as well.

Saari interprets this behavior as follows. As $t \to \infty$ the system forms clusters consisting of particles whose inter-particle distances are bounded. The centers of mass of these clusters recede from each other at a rate of about $t^{2/3}$. Moreover, the system forms subsystems (clusters of clusters) whose centers of mass recede from each other at a rate proportional to t. This, according to Saari, reflects well the actual behavior of our universe. It shows that, as time evolves, galaxies form which recede from each other according to the Newtonian version of the Hubble law of expansion: $|\dot{\mathbf{q}}_{ij}|/|\mathbf{q}_{ij}| = 1/t$ with $|\mathbf{q}_{ij}| := |\mathbf{q}_i - \mathbf{q}_j|$.

In addition to Saari's result, there exists a result by Pollard [1967] on the behavior of the maximal and minimal distance between the particles. Let R denote the maximal distance, $R = \max_{i \neq j} |\mathbf{q}_i - \mathbf{q}_j|$ with $i, j = 1, ..., N$, and r the minimal distance: $r = \min_{i \neq j} |\mathbf{q}_i - \mathbf{q}_j|$. Pollard shows that, as $t \to \infty$,

$$r \to 0 \quad \text{iff} \quad R/t \to \infty. \tag{3.8}$$

It follows that in the absence of oscillatory and pulsating motion, it cannot happen that $r \to 0$ as $t \to \infty$ (since, in the absence of oscillatory and pulsating motion, (3.4)–(3.6) hold, that is, $R/t < \infty$ as $t \to \infty$). Hence, $U(t)$ cannot grow without bounds. Now we can determine the asymptotic behavior of U and I.

Let us assume that the Newtonian gravitational system forms at least two subsystems where each subsystem consists of at least two clusters. This we want to call the generic behavior of the N-particle system. Given this assumption, it follows that, as $t \to \pm\infty$, the moment of inertia I given by $I(t) = \frac{m}{N} \sum_{i<j} |\mathbf{q}_i(t) - \mathbf{q}_j(t)|^2$ (cf. (2.19)) increases as

$$I(t) \sim t^2 \tag{3.9}$$

whereas the potential energy $U(t) = \sum_{i<j} \frac{Gm^2}{|\mathbf{q}_i(t) - \mathbf{q}_j(t)|}$ evolves as

$$U(t) \sim 1. \tag{3.10}$$

Here we are just interested in the order of t. It follows that, as $t \to \pm\infty$, $I \cdot U$ increases as

$$I(t) \cdot U(t) \sim t^2. \tag{3.11}$$

Recall that we are interested in the time evolution of the entropy of the $E = 0$ Newtonian universe, where the entropy is given by (3.1),

$$S \approx 3N/2 \ln(I \cdot U) + S'(N).$$

We get the behavior of S for $t \to \pm\infty$ from (3.11). What about the behavior in between? We know that $I(t)$ is a function that is concave upwards with a positive, global minimum at $t = \tau$: $I(\tau) = I_{min}$. Let us assume that

$$I(t) = \alpha(t - \tau)^2 + \beta \tag{3.12}$$

where α and β are positive constants ($\beta = I_{min}$). This gives us the qualitatively correct behavior of I. Now, let us furthermore, assume that U is suitably well-behaved. For that we have to exclude point-particle collisions and "near point-particle collisions" (close encounters of particles). Then U is finite and $\dot{U}$ is bounded (i.e. the graph of U has no narrow peaks). Since U is strictly positive, $U > 0$, we conclude that $I(t) \cdot U(t)$ has a global minimum, more or less at $t = \tau$, and the qualitative behavior of $I \cdot U$ is given by

$$I(t) \cdot U(t) = \gamma(t - \tau)^2 + \delta \tag{3.13}$$

where γ and δ are positive constants. Note that this is really a simplified picture. In reality, $I \cdot U$ will fluctuate both around the minimum and as it increases with $(t - \tau)^2$. However, (3.13) is qualitatively correct. It gives us the correct asymptotic behavior ($I \cdot U \sim t^2$ as $t \to \pm\infty$) and it captures the fact that, apart from fluctuations, $I(t) \cdot U(t)$ is concave upwards with a positive minimum at about $t = \tau$.

In fact, numerical simulations by Barbour et al. [2013] and [2015] for $N = 1000$ particles and typical initial data strongly support the claim that the actual evolution of $I \cdot U$ is well approximated by (3.13).[15]

Recall that we are only interested in the qualitative behavior of the entropy of the $E = 0$ Newtonian universe. Hence, we may assume that (3.13) holds and that

$$\mu_E(\Gamma_{U,I}) = (I \cdot U)^{3N/2}, \tag{3.14}$$

where in (3.1) we set all positive constants equal to 1 and replace "$\approx$" by "$=$" for simplicity. This gives us the correct qualitative behavior. Then the entropy is

$$S = k_B \ln \mu_E(\Gamma_{U,I}) = \frac{3N}{2} k_B \ln(I \cdot U) \tag{3.15}$$

where $S(t)$ is determined by $I(t) \cdot U(t)$. From (3.13) we obtain that $S(t)$ has a global minimum, $S = 3N/2 k_B \ln \beta$, at $t = \tau$ and S increases without bound in both time directions away from it. Hence, $S(t)$ is a U-shaped function in time!

[15] Cf. Barbour, Koslowski and Mercati [2013] and [2015]. Here "typical" initial data presumably refers to arbitrarily chosen "initial" positions $\mathbf{q}_i(\tau)$ and momenta $\mathbf{p}_i(\tau)$.

3.2 Statistical analysis: the problem of non-normalizable measures

We want to find out whether we are typically close to the minimum of the entropy curve or far away from it. If we are typically far away from the minimum, we don't need a Past Hypothesis to explain our low-entropy past. If we are typically close to the minimum, we do need a Past Hypothesis. To decide upon this question, we need to statistically analyze the $E = 0$ Newtonian universe.

Remark (Boltzmann statistics)**.** Recall that according to Boltzmann we are typically close to the minimum of the overall entropy curve. In other words, if we are in a state of low entropy *now*, it is highly unlikely that we had been in a state of even lower entropy *before* (or will be *afterwards*). This is an unambiguous statistical assertion given Boltzmann's model of the universe. Now there is a crucial difference between Boltzmann's model and Carroll's model of the universe (here exemplified by the $E = 0$ Newtonian universe). In Carroll's model the entropy can grow without bounds whereas in Boltzmann's model the entropy is finite with a maximum attained when the universe is in thermal equilibrium. This difference is reflected in the fact that in the Carroll model phase space – or rather the constant energy hypersurface Γ_E – is infinite whereas in the Boltzmann model it is finite. Since the measure of typicality is a measure on phase space, respectively on the constant energy hypersurface Γ_E, this entirely changes the statistical analysis. In the Boltzmann model the measure is normalizable, whereas in the Carroll model it is not. This leads to unambiguous statistical assertions in the first case, and to apparent mathematical contradictions in the other. Let us look at this in more detail.

Let us statistically analyze the $E = 0$ Newtonian universe.[16] Let $y(t) := I(t) \cdot U(t)$ and $\Gamma_y := \Gamma_{U,I}$. From (3.13) we know that $y(t) = \alpha(t - \tau)^2 + \beta$ where α, β are positive constants. Let, for simplicity, $\alpha = 1$. From (3.13) we know that $\mu_E(\Gamma_y) = y^{3N/2}$. Let $x := t - \tau$ denote the difference between time t and the moment τ at which the entropy is minimal. Hence, $y = x^2 + \beta$.

To answer the question whether we are typically close to the minimum (x small) or far away from it (x large), let us consider the projection of the microcanonical measure μ_E onto the $x - y$ plane.

Lemma 3.2. *Let everything be as above. Let $\mu_E(\Gamma_y) = y^{3N/2}$ with $y = x^2 + \beta$. Then the projection of the measure μ_E onto the $x - y$ plane is, for large N,*

$$\rho(x, y) = (y - Ux^2)^{3N/2}. \tag{3.16}$$

Proof. Let

$$\Gamma_{x',y'} = \{(\mathbf{q}, \mathbf{p}) \in \Gamma_E | x(\mathbf{q}, \mathbf{p}) = x', y(\mathbf{q}, \mathbf{p}) = y'\}.$$

[16] The following analysis is similar to the analysis performed in Goldstein et al. [2016]. There the authors analyze what they call the Carroll toy model: free particles in infinite space. As I show in my master thesis [2012], the Carroll toy model does not feature a U-shaped entropy curve (as opposed to what it is meant to do), whereas the $E = 0$ Newtonian universe does. Hence, the following analysis really applies to the $E = 0$ Newtonian universe. Over and above this convenient incidence, it provides a very realistic picture of the actual universe.

What we need to compute is $\mu_E(\Gamma_{x',y'})$.

Note that, for $E = 0$, the momentum part of the microcanonical phase space integral determines a $(3N-1)$-dimensional sphere S_r of radius $r = \sqrt{m^{-1}(\sum_{i<j} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|})}$ (where, without loss of generality, we set $m_i = m \; \forall i = 1, ..., N$). This follows from the fact that, in case $E = 0$, $\Gamma_E = \{(\mathbf{q}, \mathbf{p}) \in \Gamma | \sum_i m\mathbf{p}_i^2 = \sum_{i<j} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|}\}$ with $\Gamma = \mathbb{R}^{6N}$ (cf. also (2.32)).

Now we use that

$$\Gamma_{x',y'} = \bigcup_{\mathbf{p} \in \mathbb{R}^{3N} \cap S_r} Q^{\mathbf{p}}_{x',y'} \times \{\mathbf{p}\}$$

where

$$Q^{\mathbf{p}}_{x',y'} = \{\mathbf{q} \in \mathbb{R}^{3N} | x(\mathbf{q}, \mathbf{p}) = x', y(\mathbf{q}, \mathbf{p}) = y'\}.$$

If the volume of $Q^{\mathbf{p}}_{x',y'}$ is independent of $\mathbf{p}$, then $\mu_E(\Gamma_{x',y'})$ can be determined as follows. We know from (2.32) that in that case the microcanonical measure turns into

$$\mathrm{d}\mu'_E = C\left(\sum_{i<j} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|}\right)^{\frac{3N-2}{2}} \mathrm{d}^{3N}q$$

where C depends on N and the other constants. Now μ'_E is a measure on $Q = \mathbb{R}^{3N}$ and $\mu_E(\Gamma_{x',y'}) = \mu'_E(Q_{x',y'})$ where $Q_{x',y'} = \{\mathbf{q} \in \mathbb{R}^{3N} | x(\mathbf{q}, \mathbf{p}) = x', y(\mathbf{q}, \mathbf{p}) = y'\}$ (where we dropped the index $\mathbf{p}$ to indicate that the volume of $Q_{x',y'}$ does not depend on $\mathbf{p}$).

Recall that, in the microcanonical phase space integral, U and I are fixed separately (cf. (2.30)). Let $U = U'$ with $U = \sum_{i<j} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|}$ and let $I = I'$ with $I = \sum_i m\mathbf{q}_i^2$. Recall, in addition, that the equation $y = x^2 + \beta$ was obtained by assuming that the time-evolution of y (with $y = U \cdot I$) is essentially governed by the time evolution of I. In this case we can use the result of Goldstein et al. [2016]. They consider a system of free particles with kinetic (= total) energy $E > 0$, take I as the only macrovariable and compute the projection of the microcanonical measure onto the $x - y$ plane (where $I = x^2 + \alpha$ due to the dynamics). Let

$$A_{x',I'} = \{\mathbf{q} \in \mathbb{R}^{3N} | x(\mathbf{q}, \mathbf{p}) = x', I(\mathbf{q}, \mathbf{p}) = I'\}.$$

They show that the volume of the set $A_{x',I'}$ is indeed independent of $\mathbf{p}$ and that, for large N,

$$Vol(A_{x',I'}) \sim \big(I' - (x')^2\big)^{3N/2}.$$

We can use this result since we assumed that $y \sim x^2$ because $I \sim x^2$. Hence, the projection onto the $x - I$ plane determines the projection onto the $x - y$ plane, only that now the total energy constraint is replaced by the constraint $U = U'$ with $U = \sum_{i<j} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|}$. This determines the radius of the momentum sphere S_r. It is $r = \sqrt{m^{-1}U'}$.

We obtain that $\mu'_E(Q_{x',y'}) = Vol(A_{x',I'})S_{\sqrt{m^{-1}U'}}$ where $Vol(A_{x',I'}) \sim \big(I' - (x')^2\big)^{3N/2}$ and $S_{\sqrt{m^{-1}U'}} \sim (U')^{3N/2}$ for large N. With $y' = U' \cdot I'$ it follows that, for large N,

$$\mu'_E(Q_{x',y'}) \sim (y' - U'(x')^2)^{3N/2}.$$

Now $\mu'_E(Q_{x',y'}) = \mu_E(\Gamma_{x',y'})$. Setting $\rho(x,y) = \mu_E(\Gamma_{x',y'})$ we get $\rho(x,y) = (y - Ux^2)^{3N/2}$.

□

A measure which is easier to analyze than (3.16), but which is qualitatively the same is

$$\rho(x,y) = e^{y-x^2}. \tag{3.17}$$

Let us in what follows use this measure.

Let now A be a small stripe (say of width 4) around the minimum of the entropy curve. That is, $A = \{(x,y) \in \Sigma | x \in [-2,2]\}$ where the physical region is $\Sigma = \{(x,y) | y \geq x^2\}$. We want to find out whether A is typical – then we are typically close to the minimum of the entropy – or not. Unfortunately, $\rho(x,y)$ does not allow for a statistical assertion about A. To be precise,

$$\rho(A) = \rho(\Sigma \backslash A) = \infty. \tag{3.18}$$

To do the statical analysis, let us regularize the measure by conditioning. To keep it simple, let us consider the $x-y$ half plane $\Gamma = \{(x,y) | x \in \mathbb{R}, y \in \mathbb{R}_0^+\}$ instead of Σ (which is qualitatively the same). We will find that, depending on the way we condition, the measure of A will be different.

Lemma 3.3. *Let $\rho(x,y) = e^{y-x^2}$ on $\Gamma\{(x,y) | x \in \mathbb{R}, y \in \mathbb{R}_0^+\}$ with $y = x^2 + \beta$. Let $x \in [-2,2]$. Let*

$$\sigma_{y'}(x) = \frac{\rho(x,y|y=y')}{\int_{-\infty}^{\infty} \rho(x,y|y=y')\mathrm{d}x}, \tag{3.19}$$

respectively

$$\sigma_{\beta'}(x) = \frac{\rho(x,y|\beta=\beta')}{\int_{-\infty}^{\infty} \rho(x,y|\beta=\beta')\mathrm{d}x} \tag{3.20}$$

the (normalized) conditional measures. It is

$$\sigma_{y'}([-2,2]) > 0.9 \tag{3.21}$$

and

$$\sigma_{\beta'}([-2,2]) << 1. \tag{3.22}$$

Proof. Assume we condition on the macrostate $y = y'$. Then, for all y',

$$\sigma_{y'}([-2,2]) = \frac{\int_{-2}^{2} \rho(x,y')dx}{\int_{-\infty}^{\infty} \rho(x,y')dx} = \frac{1}{\sqrt{\pi}} \int_{-2}^{2} e^{-x^2} dx > 0.9.$$

In contrast, let us condition on the particular curve by setting $\beta = \beta'$. In that case, the conditional measure is uniform in x: $\sigma_{\beta'}(x) \propto 1$. It follows that, for all β',

$$\sigma_{\beta'}([-2,2]) = \frac{\int_{-2}^{2} \rho_{\beta'}(x)dx}{\int_{-\infty}^{\infty} \rho_{\beta'}(x)dx} << 1.$$

□

With this result we seem to arrive at two different conclusions what regards the measure of the set A. To see this first notice that the two ways of conditioning reflect two different partitions of Σ into fibres. If we condition on the macrostates, $y = y'$, we condition on horizontal lines. The set of all horizontal lines spans Σ. On the other hand, if we condition on the curves, $\beta = \beta'$, we condition on parabolas. Again, the set of all parabolas spans Σ.

Now, since $\sigma_{y'}([-2, 2]) > 0.9$ for all y', it seems to follow that

$$\sigma(A) = \rho(A)/\rho(\Gamma) > 0.9. \tag{3.23}$$

On the other hand, since $\sigma_{\beta'}([-2, 2]) << 1$ for all β', it seems to follow that

$$\sigma(A) = \rho(A)/\rho(\Gamma) << 1. \tag{3.24}$$

But this is a mathematical contradiction.

The contradiction is resolved by taking into account the fact that ρ is non-normalizable. The above reasoning is justified if and only if ρ is normalizable. Only then, the measure of a set $A \subset \Gamma$ is given by the average of the conditional measures of A on a partition of Γ into fibres!

Remark (Limits of the statistical analysis)**.** Non-normalizable measures only allow for a very limited statistical analysis. Let again Γ denote the space of all possible states of the system. Let μ be the volume measure on Γ and $\mu(\Gamma) = \infty$. Basically, as long as the total measure of Γ is infinite, statistical assertions can be made only about macro-regions $A \subset \Gamma$ of finite measure – respectively, about their complements $\Gamma \backslash A$. If some macrostate M_A defines a region A of *finite* measure, then we can (unambiguously) say that this macrostate is *atypical* with typicality measure

$$\sigma(A) = \frac{\mu(A)}{\mu(\Gamma)} = 0. \tag{3.25}$$

On the other hand, we can (unambiguously) say that its negation $\neg M_A$ defining the complement of A (i.e. $\Gamma \backslash A$) is *typical* with typicality measure

$$\sigma(\Gamma \backslash A) = 1 - \frac{\mu(A)}{\mu(\Gamma)} = 1. \tag{3.26}$$

For any other macrostate M_B defining a region B of infinite phase space volume and where the complement $\Gamma \backslash B$ has infinite volume as well, $\mu(B) = \mu(\Gamma \backslash B) = \infty$, we cannot make any assertion (like in the case above, cf. (3.18)).[17]

[17] In the physics literature, non-normalizable measures are often turned into normalizable measures by some regularization procedure (like by conditioning or imposing a cut-off). However, depending on the specific regularization we get different statistics which can even lead to opposite results like in the model discussed above or like in the case of Carroll and Tam [2010] and Gibbons and Turok [2008] who both study the probability of inflation at the example of the minisuperspace model by imposing a cut-off on the scale factor. There the authors impose two different cut-offs and arrive at opposite conclusions. For a discussion and resolution of the latter contradiction, cf. Schiffrin and Wald [2012].

3.3 Interpretation of the result

The mathematical analysis of the Newtonian/Carroll-type universe did not provide an answer to the question whether we are typically close to the minimum of the overall entropy curve or not. Still, Carroll claims that we *can* answer this question. Goldstein, Tumulka, and Zanghì [2016] explain how. Explicitly, they say that to understand Carroll one has to replace both the mathematical reasoning (from which we conclude that there is no answer to the question whether we are close to the minimum of the entropy curve or not) as well as what they call the evidential reasoning (which tells us that we should be at the minimum of the entropy curve, see below) by a different type of reasoning, a kind of theoretical reasoning.

Their approach is the following. We have some pre-theoretic knowledge about the world. Among this is our knowledge about the past: we have strong evidence that we have had a past and that this past was ordered, respectively ordered structure existed. For example, we have strong evidence for the case that dinosaurs existed because of the bones we find today. This knowledge about the past has to be taken into account when we build our best physical theories (in that sense, it is pre-theoretic). It also has to be taken into account when we give weights to the microstates compatible with our current macrostate, contrary to what evidential reasoning tells us. According to evidential reasoning we give *equal* weight to all microstates compatible with the current macrostate – from which we conclude that we should currently be at the minimum of the entropy curve.

In the Carroll model the increase of entropy is typical. This, according to Goldstein et al., is essentially the best we can hope for. It is much more than Boltzmann's model can give us. Moreover, it is *consistent* with our knowledge about the past. This is basically all we need. We need no further explanation for why we actually had a past. In particular, we don't need a Past Hypothesis.

Certainly this reasoning is correct, but it does not provide a genuine *explanation* of the fact that we had a low-entropy past. While the Carroll model is consistent with this fact, it does not provide an argument for why we really should have a low-entropy past, rather than not. In that sense it does not provide a genuine explanation. Usually, we build a physical theory in order to explain what we find in the world. The more it can explain, the better the theory. Respectively, the less pre-theoretic knowledge we have to put in, the better the theory. Given the Carroll model, there is no statistical argument which tells us that we should expect to be far away from the minimum of the entropy curve. If we had such an argument, we could say that we have a genuine explanation of the fact that we actually had a past.

This is where the work of Barbour, Koslowski and Mercati[18] adds an essential ingredient. It shows that, for the Newtonian universe, there *exists* a normalizable measure of typicality. With respect to this measure, it is an unambiguous mathematical result that *typically*, at this moment, we are far away from the minimum of the entropy curve. This will be shown in Part III. Part II will provide the mathematical framework to perform the statistical analysis.

[18] Cf. Barbour, Koslowski, and Mercati [2013], [2015].

Part II

Mathematical Framework

This part shall provide the mathematical framework which we need to statistically analyze the Newtonian universe (Part III). It also provides the grounds to later discuss the Newtonian dynamics through the points of total collision (Part IV).

4 Dynamics and measure on (generalized) phase space

Barbour, Koslowski, and Mercati [2015] present a measure of typicality on the space of mid-point data PT^*S of the $E = \mathbf{L} = \mathbf{P} = 0$ Newtonian universe. To understand why PT^*S is indeed the correct space for the statistical analysis of the Newtonian universe, we will introduce the notions of a generalized and an internal phase space. For means of generality, everything shall be formulated in coordinate-free geometric terms.

4.1 Notation and basic lemmas

To introduce the notation, I first present the standard Hamiltonian description. After that, I include time t and as a next step its canonical conjugate p_t among the phase space coordinates – this will enable us to construct the so-called generalized phase space where time has disappeared (Sec. 4.2). On generalized phase space we will reintroduce time internally and we will get back Hamiltonian equations together with a stationary measure, described on the so-called internal phase space (Sec. 4.3).

For the time-independent and parts of the time-dependent case, cf. Abraham and Marsden [1978] and Scheck [2003]. The generalized formalism is sketched in Rovelli [2000]. I give a precise account of the generalized formalism and develop the internal Hamiltonian description.

4.1.1 Hamiltonian dynamics on Γ

Let us consider a system of particles. The positions of the particles are represented by n position variables q^i (in case of N particles in three-dimensional space: $n = 3N$).[19] These variables form a complete set of local coordinates of n-dimensional configuration space Q. Together with the momenta p_i they form a collection of local coordinates of the cotangent bundle of Q, $T^*Q =: \Gamma$. We call Γ the $2n$-dimensional *phase space* of the system, q^i the *canonical coordinates* and p_i the *canonical momenta*. In the Hamiltonian formalism, every point $(q, p) := (q^1, ..., q^n, p_1, .., p_n) \in \Gamma$ represents one possible state of the system, respectively one possible initial condition. For N particles moving through three-dimensional Euclidean space: $Q = \mathbb{R}^{3N}$ and $T^*Q \cong \mathbb{R}^{6N}$.

[19] We will use upper and lower indices as long as we are particularly interested in the geometry while we will later, when we come to the physics, use only lower indices for all the variables.

On the cotangent bundle of configuration space T^*Q there exists a *natural* or *canonical one-form* θ (an element of $T^*(T^*Q)$)[20] which can be expressed in local coordinates as follows.

Definition 4.1 (Natural one-form on Γ)**.** [21] Let $\Gamma = T^*Q$. Let q^i, p_i local coordinates on Γ.

$$\theta = \sum_{i=1}^{n} p_i \mathrm{d}q^i \tag{4.1}$$

is the natural one-form on Γ.

There also exists a *natural two-form* ω on $T^*\mathcal{Q}$ which can be constructed from the natural one-form θ by taking the negative exterior derivative.

Definition 4.2 (Natural two-form on Γ)**.** Let θ the natural one-form on $\Gamma = T^*Q$. Then

$$\omega = -d\theta \tag{4.2}$$

is the natural two-form on Γ.

Here the minus sign is a matter of convention. In local coordinates,

$$\omega = \sum_{i=1}^{n} \mathrm{d}q^i \wedge \mathrm{d}p_i. \tag{4.3}$$

This two-form is symplectic.[22] That is, it has the following properties.

Definition 4.3 (Symplectic form)**.** Let M be a manifold of $2n$ dimensions. Let ω a two-form on M. Then ω is symplectic if and only if it is *closed* ($\mathrm{d}\omega = 0$), *alternating* (i.e., $\omega(X, X) = 0$ for all $X \in \mathcal{V}(\Gamma)$ where $\mathcal{V}(\Gamma)$ denotes the set of smooth vector fields on Γ) and *nondegenerate* (i.e., there exists no non-zero $X \in \mathcal{V}(\Gamma)$ such that $\omega(X, Y) = 0$ for all $Y \in \mathcal{V}(\Gamma)$).

Volume form and Liouville measure. Given a symplectic two-form, there exists a natural, oriented volume form.

Definition 4.4 (Natural volume form)**.** Let ω the symplectic two-form on $\Gamma = T^*Q$. Then

$$\Omega = \frac{(-1)^{[n/2]}}{n!} \omega^n \tag{4.4}$$

is the natural volume form on Γ. Here $[n/2]$ is the biggest integer smaller or equal to $n/2$.

[20] For the naturalness of this definition, cf. Scheck [2003]. There it is also shown that on the cotangent bundle T^*M of any smooth manifold M, there always exists a natural one-form θ and a natural symplectic two-form $\omega := -\mathrm{d}\theta$. From Darboux's theorem it follows that there exist local coordinates (q^i, p_i) such that θ and ω can be written in the given form.

[21] Alternatively, there exists a coordinate-free definition of the natural one-form. What regards the one-form, we don't benefit from giving its coordinate-free definition, so we don't do it here – all we need to know is that there exists a natural one-form which, in local coordinates, is of the given form.

[22] For a proof, cf. Scheck [2003].

Notice that this is really a volume form on Γ. To be a volume form it has to be a nowhere-zero top-dimensional form on Γ. It is top-dimensional because it is a $2n$-form on a $2n$-dimensional space and it is nowhere zero due to the non-degeneracy of ω.

In local coordinates,

$$\Omega = (-1)^{[n/2]} \mathrm{d}q^1 \wedge \mathrm{d}p_1 \wedge ... \wedge \mathrm{d}q^n \wedge \mathrm{d}p_n. \tag{4.5}$$

Connected to this oriented volume form, there exists a so-called *volume element* or *density* $\mu = |\Omega|$ which is non-oriented. It is this volume element which defines a measure on Borel sets. Borel sets are Lebesgue integrable. Hence, the existence of the density $\mu = |\Omega|$ allows for the use of the Lebesgue integral in order to integrate functions on Γ. It is just the measure we need to determine the volume of regions $A \subset \Gamma$. In fact, it is just the Liouville measure (2.5).

Definition 4.5 (Natural volume measure)**.** Let Ω the natural volume element on $\Gamma = T^*Q$. Then

$$\mathrm{d}\mu = |\Omega| \tag{4.6}$$

is the natural volume measure, or Liouville measure, on Γ.

To see that this is the Liouville measure, let us rewrite it in local coordinates. From (4.5) it follows that $\mu = |\Omega|$ is

$$\mathrm{d}\mu = \prod_{i=1}^{n} \mathrm{d}q^i \mathrm{d}p_i. \tag{4.7}$$

This coincides with our definition of the Liouville measure above (cf. (2.5)).

Hamiltonian dynamics. In addition to phase space itself, there exists a smooth function $H = H(q^i, p_i) : \Gamma \to \mathbb{R}$ called the *Hamiltonian* of the system. The Hamiltonian H defines the physical vector field X_H on $\Gamma = T^*Q$.

Definition 4.6 (Physical vector field X_H)**.** Let H be a smooth function on Γ, the Hamiltonian of the system, and ω the symplectic two-form on Γ. Let $X_H \in T\Gamma$ such that

$$\omega(X_H, \cdot) = \mathrm{d}H. \tag{4.8}$$

Then X_H is the physical vector field.

X_H is a Hamiltonian vector field[23] and it defines a Hamiltonian phase flow T on Γ. Since ω determines a measure μ on Γ and H determines a flow T on Γ, the quadruple $(\Gamma, \mathcal{B}(\Gamma), \omega, H)$ form a dynamical system. Since the flow is Hamiltonian, we also call it a *Hamiltonian system.*

In local coordinates, the physical vector field can be written as follows.

[23] We call a vector field X_f *Hamiltonian* if and only if there exists a smooth function f on $M = T^*V$ such that $\omega(X_f, \cdot) = \mathrm{d}f$. Here V is a vector space and ω is the symplectic two-form on $M = T^*V$. In contrast, the *physical* vector field is determined by the *physical Hamiltonian* H of the system. For a given physical model, there is only one physical vector field while there can be many different Hamiltonian vector fields fulfilling the relation $\omega(X_f, \cdot) = \mathrm{d}f$ for some f.

Lemma 4.1 (X_H in local coordinates)**.** *Let everything as above. Let q^i, p_i local coordinates on Γ. It is*

$$X_H = \sum_{i=1}^{n} \frac{\partial H(q^i,p_i)}{\partial p_i}\frac{\partial}{\partial q^i} + \sum_{i=1}^{n}\Big(-\frac{\partial H(q^i,p_i)}{\partial q^i}\Big)\frac{\partial}{\partial p_i}. \tag{4.9}$$

Proof. In local coordinates, a vector field $X \in T(T^*\mathcal{Q})$ is of the general form

$$X = \sum_{i=1}^{n} f^i(q^i,p_i)\frac{\partial}{\partial q^i} + \sum_{i=1}^{n} g_i(q^i,p_i)\frac{\partial}{\partial p_i}$$

with f^i, g_i arbitrary functions.

According to Definition 4.6, X_H is the physical vector field if and only if $\omega(X_H,\cdot) = \mathrm{d}H$ where $\omega = \sum_{i=1}^{n} \mathrm{d}q^i \wedge \mathrm{d}p_i$. Let $Y \in \mathcal{V}(\Gamma)$ an arbitrary, smooth vector field on Γ. Now

$$\begin{aligned}
\omega(X_H,Y) &= \Big(\sum_{i=1}^{n} \mathrm{d}q^i \wedge \mathrm{d}p_i\Big)(X_H,Y) \\
&= \sum_{i=1}^{n}\big[\mathrm{d}q_i(X_H)\mathrm{d}p_i(Y) - [\mathrm{d}p_i(X_H)\mathrm{d}q^i(Y)\big] \\
&= \sum_{i=1}^{n}\big[f^i\mathrm{d}p_i(Y) - g_i\mathrm{d}q^i(Y)\big]
\end{aligned}$$

where

$$\sum_{i=1}^{n}\big[f^i\mathrm{d}p_i(Y) - g_i\mathrm{d}q^i(Y)\big] = \mathrm{d}H(Y)$$

if and only if f^i, g_i fulfill the equations

$$f^i(q^i,p_i) = \frac{\partial H(q^i,p_i)}{\partial p_i}, \quad g_i(q^i,p_i) = -\frac{\partial H(q^i,p_i)}{\partial q^i}.$$

This shows the assertion. □

X_H is the physical vector field. That is, the integral curves $\gamma(t)$ along X_H are possible trajectories of the system:

$$\dot{\gamma}(t) = (X_H)_{\gamma(t)} \tag{4.10}$$

Here $\cdot$ denotes the derivative with respect to time t. In local coordinates, $\gamma(t) = (q^i(t), p_i(t))$ and (4.10) turns into

$$\frac{\mathrm{d}q^i}{\mathrm{d}t} = \frac{\partial H(q^i,p_i)}{\partial p_i}, \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = -\frac{\partial H(q^i,p_i)}{\partial q^i}. \tag{4.11}$$

These are the well-known *Hamiltonian laws of motion.*

All integral curves $\gamma(t)$ together constitute the physical phase flow (which is a Hamiltonian phase flow) T_t.

Definition 4.7 (Physical phase flow)**.** Let X_H defined by (4.8) the physical vector field on γ.

Let $\gamma(t) = (q^i(t), p_i(t))$ an integral curve along X_H, i.e., $\dot{\gamma}(t) = (X_H)_{\gamma(t)}$. Then $T_t : \Gamma \to \Gamma$,

$$T_t(q^i, p_i) = (q^i(t), p_i(t)), \tag{4.12}$$

is the physical phase flow.

Invariants of motion and Liouville's theorem. Let us now study the behavior of the measure under the physical phase flow and derive Liouville's theorem. For that we need to introduce the notion of the Lie derivative.

Definition 4.8 (Lie derivative of functions)**.** Let M be a manifold, $X \in \mathcal{V}(M)$ a smooth vector field, f a smooth function on M. Then

$$L_X f = \mathrm{d}f(X) \tag{4.13}$$

is the Lie derivative of f along X.

That is, the Lie derivative of f along X is just the derivative of f in the direction of X (which is the application of X to f): $L_X f = \mathrm{d}f(X) = X(f)$.

There exists another useful formulation of the Lie derivative. For any smooth curve $\gamma(t)$ on M parametrized by t with $\gamma(0) = p$ and tangent vector $X_p = \dot{\gamma}(0)$, the following relation holds: $\mathrm{d}f_p(X_p) = \frac{\mathrm{d}}{\mathrm{d}t} f(\gamma(t))|_{t=0}$. Now let the integral curve $\gamma(t)$ along the vector field X passing through a point p be given by the respective flow line $T_t p$ with $T_0(p) = p$ and $\frac{\mathrm{d}}{\mathrm{d}t} T_0(p) = X_p$. Then, at the point $p \in M$, the Lie derivative can also be written as

$$\mathrm{d}f_p(X_p) = \frac{\mathrm{d}}{\mathrm{d}t} f(T_t p)\big|_{t=0}. \tag{4.14}$$

Using the definition of the pullback, this can be reformulated as follows:

$$\frac{\mathrm{d}}{\mathrm{d}t} f(T_t p)\big|_{t=0} = \frac{\mathrm{d}}{\mathrm{d}t} T_t^* f(p)\big|_{t=0}. \tag{4.15}$$

Together with (4.13), (4.14) says that the Lie derivative L_{X_H} of f along the Hamiltonian vector field X_H can be identified with the total time derivative of f.

In addition, there exists a definition of the Lie derivative on the space of differential forms.

Definition 4.9 (Lie derivative of forms)**.** Let M be a manifold, $X \in \mathcal{V}(M)$ a smooth vector field, $\alpha \in \Omega^k(M)$ a differential k-form. Then

$$L_X \alpha = (\mathrm{d}\alpha)(X, \cdot\ k\ \mathit{times}\ \cdot) + \mathrm{d}(\alpha(X, \cdot\ k-1\ \mathit{times}\ \cdot)). \tag{4.16}$$

This equation is called *Cartan's formula* or *Cartan's identity.*[24] For a given two-form ω, it turns into

$$L_X \omega = (\mathrm{d}\omega)(X, \cdot, \cdot) + \mathrm{d}(\omega(X, \cdot)). \tag{4.17}$$

[24]Cf. Abraham and Marsdem [1978].

It is an important result that the Lie derivative of the symplectic two-form ω along X_H vanishes.

Lemma 4.2. *Let $X_H \in T\Gamma$ and ω the symplectic two-form on Γ. Then*

$$L_{X_H}\omega = 0. \tag{4.18}$$

Proof. We use (4.16) and (4.8) and the fact that ω is closed: $\mathrm{d}\omega = 0$. From (4.8) we get that $\mathrm{d}\omega(X_H, \cdot) = \mathrm{d} \circ \mathrm{d}H = 0$. Inserting this into (4.16), we find that

$$\begin{aligned} L_{X_H}\omega &= (\mathrm{d}\omega)(X_H, \cdot, \cdot) + \mathrm{d}(\omega(X_H, \cdot)) \\ &= 0 + \mathrm{d} \circ \mathrm{d}H = 0. \end{aligned}$$

□

It follows from this result that the two-form ω and the volume element $\mu = |\Omega| = |\omega|^n/n!$ are invariant under the Hamiltonian phase flow. Hence, volume is conserved under time evolution. This is Liouville's theorem.

Corollary 4.1 (Liouville's theorem)**.** *Let everything be as above. Let T_t the Hamiltonian phase flow and $\mu = |\omega|^n/n!$. Then*

$$T_t^*\omega = \omega \tag{4.19}$$

and

$$T_t^*\mu = \mu. \tag{4.20}$$

Proof. Using (4.13)-(4.15) and (4.18), we find that

$$\frac{\mathrm{d}}{\mathrm{d}t}T_t^*\omega = L_{X_H}\omega = 0.$$

This shows (4.19). In addition, (4.18) together with the fact that $L_X(\alpha \wedge \beta) = (L_X\alpha) \wedge \beta + \alpha \wedge (L_X\beta)$ implies that each form ω^k with $k = 1, ..., n$ is invariant under the Hamiltonian phase flow. Hence, in particular, the $2n$-form ω^n and, consequently, also the oriented volume form $\Omega = (-1)^{[n/2]}\omega^n/n!$ and its density $\mathrm{d}\mu = |\Omega|$ are invariant under the flow. □

In addition, the Hamiltonian $H = H(q^i, p_i)$ is itself invariant under the Hamiltonian phase flow. This means that total energy is conserved under time evolution.

Lemma 4.3 (Energy conservation)**.** *Let H be the Hamiltonian of the system and X_H the Hamiltonian vector field on Γ. Then*

$$L_{X_H}H = 0 \tag{4.21}$$

Proof. Let q^i, p_i a set of local coordinates on Γ. It is

$$L_{X_H}H = \mathrm{d}H(X_H) = X_H(H) = \sum_{i=1}^{n} \frac{\partial H}{\partial p^i}\frac{\partial H}{\partial q^i} + \sum_{i=1}^{n}\Big(- \frac{\partial H}{\partial q^i}\Big)\frac{\partial H}{\partial p_i} = 0.$$

$\square$

You can use (4.14) to rewrite Eq. (4.21) as follows:

$$\frac{\mathrm{d}}{\mathrm{d}t} H(q^i(t), p_i(t)) = 0. \tag{4.22}$$

That is, H is invariant under time evolution.

Remark (Time-dependent Hamiltonian)**.** Almost everything we said so far also applies to the time-dependent case. Only then, $H : \mathbb{R} \times \Gamma \to \mathbb{R}, H = H(t, q^i, p_i)$ and the same for the vector field, $X_H : \mathbb{R} \times \Gamma \to T\Gamma, X_H = X_H(t, q^i, p_i)$. But what is a time-dependent vector field? Notice that, for any *fix* moment of time, H is a Hamiltonian function on Γ and X_H is a Hamiltonian vector field on Γ. That is, we can define a time-dependent Hamiltonian H_t and a time-dependent vector field X_{H_t} on Γ as follows.

Definition 4.10 (Time-dependent Hamiltonian H_t and vector field X_{H_t})**.** Let $H : \mathbb{R} \times \Gamma \to \mathbb{R}, H = H(t, q^i, p_i)$ and $X_H : \mathbb{R} \times \Gamma \to T\Gamma, X_H = X_H(t, q^i, p_i)$. Then

$$H_t : \Gamma \to \mathbb{R}, \quad H_t(q^i, p_i) = H(t, q^i, p_i) \tag{4.23}$$

is the time-dependent Hamiltonian H_t and

$$X_{H_t} : \Gamma \to T\Gamma, \quad X_{H_t}(q^i, p_i) = X_H(t, q^i, p_i). \tag{4.24}$$

is the time-dependent physcial vector field X_{H_t} on Γ.

Notice that, also in this case, the time-dependent vector-field X_{H_t} is determined by (4.8), that is, by demanding that $\omega(X_{H_t}, \cdot) = \mathrm{d}H_t$.

Whereas a time-independent Hamiltonian defines a one-parameter phase flow T_t on Γ, a time-dependent Hamiltonian defines a two-parameter flow $T_{t,s}$ on Γ.

Definition 4.11 (Time-dependent phase flow $T_{t,s}$)**.** Let $(q^i, p_i) := (q^i(s), p_i(s))$ an integral curve along the time-dependent vector field X_{H_t} from (4.24). Then

$$T_{t,s}(q^i, p_i) = (q^i(t), p_i(t)) \tag{4.25}$$

is the time-dependent Hamiltonian phase flow on Γ.

Again, possible trajectories of the system are the flow lines of $T_{t,s}$. But what about the conservation of phase space volume? The proof of Liouville's theorem (Lemma 4.2 and Corollary 4.1) applies directly to the time-dependent case replacing H by H_t and X_H by X_{H_t}. That is, also for a time-dependent Hamiltonian system, phase space volume is conserved. Since this is an important result, I have attached another version of the proof using local coordinates in Appendix A.

Remark (Extended phase space)**.** In what follows, let us include time t among the coordinates. As a first step we will consider extended space $\mathbb{R} \times \Gamma$, a space of $2n+1$ dimensions. As a second step we will consider the cotangent bundle of extended configuration space $T^*(\mathbb{R} \times Q)$, a space of $2n+2$ dimensions. Whereas the first approach neglects the canonical conjugate of time, the second approach treats both time t as well as its canonical conjugate p_t as a coordinate. Later we will start from generalized coordinates q^a, p_a and develop an internal Hamiltonian description by identifying some internal time parameter τ (a monotonic function of the generalized coordinates) and its canonical conjugate p_τ.

Having the extended phase space picture in mind, it is clear that trajectories on Γ are really just a projection of the actual trajectories which lie in $\mathbb{R} \times \Gamma$. In case the Hamiltonian H is time-independent, this projection is somewhat trivial. However, the projection also applies to the time-dependent case as soon as we make use of the definition of the time-dependent Hamiltonian H_t and vector field X_{H_t} from above ((4.23) and (4.24)). Having projected the trajectories onto Γ, phase space Γ fulfills a dual role. It is both the space of initial conditions and the space in which the trajectories lie.

4.1.2 Hamiltonian dynamics on $\mathbb{R} \times \Gamma$

By adding the time variable t to the other phase space coordinates, one can construct the extended phase space $\Sigma = \mathbb{R} \times \Gamma = \mathbb{R} \times T^*Q$. Notice that in contrast to Γ, $\Sigma = \mathbb{R} \times \Gamma$ is not symplectic. This follows from the fact that it is a space of odd dimensions (cf. Def. 4.3).

However, Σ is a manifold. In local coordinates, any point $p \in \Sigma$ can be written as $p = (t, q^i, p_i)$. Like before, there exists a natural one-form θ' on this space.

Definition 4.12 (Natural one-form on Σ)**.** Let t, q^i, p_i local coordinates on $\Sigma = \mathbb{R} \times T^*Q$. Let H_t the time-dependent Hamiltonian on $\Gamma = T^*Q$ given by (4.23).

$$\begin{aligned} \theta' &= \sum_{i=1}^{n} p_i \mathrm{d}q^i - H_t(q^i, p_i)\mathrm{d}t \\ &= \theta - H_t \mathrm{d}t \end{aligned} \tag{4.26}$$

is the natural one-form on Σ.

From this one-form, one can construct a two-form $\omega' = -\mathrm{d}\theta'$, analogous to the construction on Γ (cf. Def. 4.2). Of course, since Σ is a space of $2n+1$ dimensions, ω' is not a symplectic form.

Let us now express the dynamics with respect to Σ. On Σ, there again exists a physical vector field X'. In other words, expressing the dynamics with respect to Σ, there exists an X' such that the integral curves along X' are possible trajectories of the system.[25]

Definition 4.13 (Physical vector field on Σ)**.** Let H_t the (time-dependent) Hamiltonian and

[25] For this definition, cf. Abraham and Marsden [1978].

$X_{H_t} : \Gamma \to T\Gamma$ the physical vector field on Γ. Then

$$X' = \frac{\partial}{\partial t} + X_{H_t} \tag{4.27}$$

is the physical vector field on Σ.

Let t, q^i, p_i local coordinates on Σ. In local coordinates, X' is of the following form:

$$X' = \frac{\partial}{\partial t} + \frac{\partial H_t(q^i, p_i)}{\partial p_i} \frac{\partial}{\partial q^i} - \frac{\partial H_t(q^i, p_i)}{\partial q^i} \frac{\partial}{\partial p_i}. \tag{4.28}$$

This follows directly from (4.27) together with the expression for $X_H = X_{H_t}$ on Γ (4.8).

Whereas X_{H_t} defines a two-parameter flow $T_{t,s}$ on Γ, X' defines a one-parameter flow T'_t on $\Sigma = \mathbb{R} \times \Gamma$. Both flows are connected as follows.

Definition 4.14 (Flow on Σ)**.** Let $(s, p) \in \mathbb{R} \times \Gamma$ and let $T_{t,s}$ the flow on Γ. Then

$$T'_t(s, p) = (t + s, T_{t,s}(p)) \tag{4.29}$$

is the flow on Σ.

Again, T'_t is the set of integral curves along X' on Σ and $T_{t,s}$ is the set of integral curves along X_{H_t} on Γ. In each case, the flow lines represent the physical trajectories. The two formulations of the dynamics (on Γ respectively Σ) provide merely two different mathematical representations of the very same actual physical trajectories.

Just like on Γ, on $\Sigma = \mathbb{R} \times \Gamma$ there exists a coordinate-free way of specifying the physical vector field X'. Consider the natural two-form $\omega' = -\mathrm{d}\theta'$. With respect to this two-form, the physical vector field X' is determined as follows.

Lemma 4.4 (X' coordinate-free)**.** *Let θ' the natural one-form on Σ and $\omega' = -\mathrm{d}\theta'$. Let X' be a smooth vector field on Σ such that*

$$\omega'(X', \cdot) = 0. \tag{4.30}$$

Then X' is the physical vector field (4.27) (up to multiplication by a scalar).

Note that Eq. (4.30) is nothing but the assertion that ω' is degenerate. Hence, it is not symplectic (cf. Def. 4.3).

Proof. It follows from the definition of θ' given by (4.26) that $\omega' = -\mathrm{d}\theta'$ is of the form $\omega - \mathrm{d}t \wedge \mathrm{d}H$. Hence,

$$\omega'(X', \cdot) = \omega(X', \cdot) - (\mathrm{d}t \wedge \mathrm{d}H)(X', \cdot).$$

Now any vector field X' on $\Sigma = \mathbb{R} \times \Gamma$ is of the general form $h \cdot \partial/\partial t + X$ where $h = h(t, q^i, p_i)$ is a function on Σ and X is a vector field on Γ. It follows that

$$0 \stackrel{!}{=} \omega(X', \cdot) - (\mathrm{d}t \wedge \mathrm{d}H_t)(X', \cdot) = \omega(X, \cdot) - (\mathrm{d}t \wedge \mathrm{d}H_t)\left(h(t, q^i, p_i)\frac{\partial}{\partial t}, \cdot\right) = \omega(X, \cdot) - h\mathrm{d}H_t$$

if and only if $\omega(X,\cdot) = h\mathrm{d}H_t$. But this is fulfilled if and only if (apart from scalar multiplication) $h(t, q^i, p_i) = 1$ and $X = X_{H_t}$. In that case, we get that

$$\omega(X,\cdot) = \omega(X_{H_t},\cdot) = \mathrm{d}H_t = h\mathrm{d}H_t$$

where the second equation is obtained by (4.8) with $X_H = X_{H_t}$. We get the full set of solutions by multiplication with a scalar s. Then $h = s$ and $X = sX_{H_t}$ and $\omega(X,\cdot) = \omega(sX_{H_t},\cdot) = s\mathrm{d}H_t = h\mathrm{d}H_t$. That is, $X' = X_{H_t} + \partial/\partial t$ is unique up to scalar multiplication. □

Note that multiplication by a scalar is nothing but reparametrization of the trajectories, that is, both time t and the physical vector field X_{H_t} are reparametrized.

On Σ, there also exists an invariant volume form Ω' and, hence, an invariant volume element $\mu' = |\Omega'|$.

Lemma 4.5 (Invariant volume form on Σ). *Let θ' from (4.26) and t, q^i, p_i local coordinates on Σ. Then*

$$\Omega' = -\mathrm{d}t \wedge (\mathrm{d}\theta')^n \tag{4.31}$$

is a volume form on Σ. It is invariant under the flow T'_t on Σ given by (4.29):

$$T'_t\Omega' = \Omega'. \tag{4.32}$$

Proof. The form $-\mathrm{d}t \wedge (\mathrm{d}\theta')^n$ is $(n+1)$-dimensional. Hence, it is a top-dimensional form on Σ. Moreover, it is nowhere zero. This follows from the fact that $-\mathrm{d}t \wedge (\mathrm{d}\theta')^n = \mathrm{d}t \wedge \Omega$ and Ω is nowhere zero. It follows that $-\mathrm{d}t \wedge (\mathrm{d}\theta')^n$ is a volume form on Σ.

The Lie derivative of this volume form vanishes:

$$\begin{aligned}
L_{X'}(-\mathrm{d}t \wedge (\mathrm{d}\theta')^n) &= L_{X'}(-\mathrm{d}t) \wedge (\mathrm{d}\theta')^n - \mathrm{d}t \wedge L_{X'}[(\mathrm{d}\theta')^n] \\
&= [-\mathrm{d} \circ \mathrm{d}t)(X',\cdot) + \mathrm{d}(-\mathrm{d}t(X'))] \wedge (\mathrm{d}\theta')^n - 0 \\
&= \left[0 + \mathrm{d}\Big(- \mathrm{d}t\Big(\frac{\partial}{\partial t}\Big)\Big)\right] \wedge (\mathrm{d}\theta')^n - 0 \\
&= 0 \wedge (\mathrm{d}\theta')^n = 0.
\end{aligned}$$

Here we used in particular that $L_{X'}\omega' = 0$.

Since $\frac{\mathrm{d}}{\mathrm{d}t}(T'_t\Omega') = L_{X'}\Omega'$ and $L_{X'}\Omega' = 0$, it follows that $T'_t\Omega' = \Omega'$. That is, $\Omega' = -\mathrm{d}t \wedge (\mathrm{d}\theta')^n$ is invariant under the flow T'_t.

□

4.1.3 Hamiltonian dynamics on $T^*(\mathbb{R} \times Q)$

By adding the time variable to configuration space, we arrive at the extended configuration space $\tilde{Q} = \mathbb{R} \times Q$. Local coordinates of $\tilde{\mathcal{Q}}$ are (t, q^i). In contrast to before, we now treat the time variable like an ordinary position variable which means that we have to include its canonical conjugate p_t into the description. The phase space $\tilde{\Gamma}$ is now the cotangent bundle of the extended configuration space $\tilde{\Gamma} = T^*\tilde{Q} = T^*(\mathbb{R} \times Q)$, which is a factor space $T^*(\mathbb{R} \times Q) = T^*\mathbb{R} \times T^*Q$. Analogous to Γ, $\tilde{\Gamma}$ is again a symplectic space and local coordinates of $\tilde{\Gamma}$ are (t, q^i, p_t, p_i).

Let us now describe the dynamics on $\tilde{\Gamma} = T^*(\mathbb{R} \times Q)$. On $T^*(\mathbb{R} \times Q)$ there exists a smooth function $\mathcal{H}$ which for all physical motion vanishes identically.

Definition 4.15 (Hamiltonian constraint)**.** Let H, t and p_t be as above. Let $\mathcal{H}$ a smooth function on $\tilde{\Gamma} = T^*(\mathbb{R} \times Q)$ such that

$$\mathcal{H}(t, q^i, p_t, p_i) = p_t + H(t, q^i, p_i) = 0. \tag{4.33}$$

Then $\mathcal{H} = 0$ is the *Hamiltonian constraint.*

Within $(2n + 2)$-dimensional phase space $\tilde{\Gamma}$, the Hamiltonian constraint $\mathcal{H} = 0$ defines a $(2n+1)$-dimensional hypersurface Σ to which any physical motion is restricted. By construction, this is just the surface Σ with local coordinates (t, q^i, p_i) which we know from before (cf. Sec. 4.1.2). (Of course, this is where we got the Hamiltonian constraint from in the first place.) This can be seen as follows.

Analogous to θ on $\Gamma = T^*Q$, there exists a natural one form $\tilde{\theta}$ on $\tilde{\Gamma} = T^*(\mathbb{R} \times Q)$,

$$\tilde{\theta} = \sum_{i=1}^{n} p_i \mathrm{d}q^i + p_t \mathrm{d}t. \tag{4.34}$$

Now it follows from (4.33) that $p_t = -H$ on Σ. In that case, $\tilde{\theta}$ is equal to θ' (cf. (4.26)),

$$\tilde{\theta}\big|_{\mathcal{H}=0} = \theta'. \tag{4.35}$$

Here $\tilde{\theta}\big|_{\mathcal{H}=0}$ denotes the restriction of $\tilde{\theta}$ on $\tilde{\Gamma}$ to Σ. Since $\mathcal{H}$ is a smooth function on $\tilde{\Gamma}$, this is just the pullback, $\tilde{\theta}\big|_{\mathcal{H}=0} = i_{\mathcal{H}}^*\tilde{\theta}$, where $i_{\mathcal{H}} : \Sigma \to \tilde{\Gamma}$ denotes the embedding of Σ in $\tilde{\Gamma}$.

Just like before, there also exists a natural symplectic two-form $\tilde{\omega} = -d\tilde{\theta}$ on $\tilde{\Gamma}$. In addition, there again exists a vector field $\tilde{X}$ on $\tilde{\Gamma} = T^*(\mathbb{R} \times \mathcal{Q})$ which can be determined in a coordinate-free way as follows.

Definition 4.16 (Physical vector field on $\tilde{\Gamma}$)**.** Let $\tilde{X}$ on $\tilde{\Gamma} = T^*(\mathbb{R} \times \mathcal{Q})$ such that

$$-\mathrm{d}\tilde{\theta}\big|_{\mathcal{H}=0}(\tilde{X}, \cdot) = 0. \tag{4.36}$$

Then $\tilde{X}$ is the physical vector field on $\tilde{\Gamma}$.

Again, possible trajectories are the flow lines along this vector field. In fact, these are just the trajectories from the previous section. To see this, note that due the Hamiltonian constraint every trajectory lies in the *constraint surface* Σ. Let again $i_{\mathcal{H}} : \Sigma \to \tilde{\Gamma}$ denote the embedding of Σ in $\tilde{\Gamma}$. Then $i_{\mathcal{H}}^*\mathrm{d}\tilde{\theta} = \mathrm{d}\theta'$ and $i_{\mathcal{H}}^*\tilde{X} = X'$. Thus, we are back to the setting we discussed in Section 3.1.2. Again, any physical trajectory is given by an integral curve along the vector field X' on Σ.

4.2 Generalized coordinates: Hamiltonian dynamics on $\tilde{\Gamma}$

The formalism we developed so far can now be applied to a very general setting: a formulation of the dynamics where time has disappeared. Let us start with generalized phase space $\tilde{\Gamma} = T^*\tilde{Q}$, the cotangent bundle of generalized configuration space $\tilde{Q}$. The configurational variables q^a with $a = 1, ..., n$ form local coordinates of $\tilde{Q}$ and the q^a together with their conjugate momenta p_a form local coordinates of $\tilde{\Gamma}$. Since t is not specified, there is also no special function H to which t is connected via its canonical conjugate ($p_t = -H$) as it had been the case in the previous section. However, we can adapt the formalism also to this very general setting.

Analogous to the definitions of θ and ω on Γ, we can define a natural one-form $\tilde{\theta}$ and a natural symplectic two-form $\tilde{\omega}$ on $\tilde{\Gamma}$. Hence, analogous to Def.'s 4.1 and 4.2,

$$\tilde{\theta} = p_a \mathrm{d}q^a \tag{4.37}$$

and

$$\tilde{\omega} = -\mathrm{d}\tilde{\theta} = \sum_a \mathrm{d}q^a \wedge \mathrm{d}p_a. \tag{4.38}$$

In addition, also in this generalized setting, there is a function $\mathcal{H}$ on $\tilde{\Gamma} = T^*\tilde{\mathcal{Q}}$ in terms of which the Hamiltonian constraint

$$\mathcal{H}(q^a, p_a) = 0 \tag{4.39}$$

is formulated. Now again (cf. Def. 4.16), the physical vector field can be defined by

$$-\mathrm{d}\tilde{\theta}\big|_{\mathcal{H}=0}(\tilde{X}, \cdot) = 0. \tag{4.40}$$

Since q^a and p_a are local coordinates of $\tilde{\Gamma}$, we know that $\tilde{X}$ on $\tilde{\Gamma}$ must be of the form

$$\tilde{X} = \sum_a f^a(q^a, p_a)\frac{\partial}{\partial q^a} + \sum_a g_a(q^a, p_a)\frac{\partial}{\partial p_a} \tag{4.41}$$

with f^a and g_a determined by $\mathcal{H}$.

Now everything is analogous to the case in which $\tilde{\Gamma} = T^*(\mathbb{R} \times Q)$ (see the previous section). Also on $2n$-dimensional generalized phase space $\tilde{\Gamma} = T^*\tilde{Q}$ the Hamiltonian constraint defines a $(2n-1)$-dimensional hypersurface $\Sigma \subset \tilde{\Gamma}$ to which any physical solution, respectively any possible trajectory, is restricted. That is, $\mathcal{H}$ is a conserved quantity of the system: it is constant (constantly zero) along the trajectories. Let again $i_{\mathcal{H}} : \Sigma \to \tilde{\Gamma}$ denote the embedding of Σ in $\tilde{\Gamma}$ where Σ is defined by $\mathcal{H} = 0$ (4.39). Analogous to before, there exists a one-form $\theta' = i^*_{\mathcal{H}}\tilde{\theta}$ on Σ and a degenerate two-form $\omega' = -\mathrm{d}\theta'$ which determines the trajectories of the system via the equation $\omega'(X', \cdot) = 0$. Again $i^*_{\mathcal{H}}\tilde{X} = X'$ and possible trajectories are the integral curves along X' on Σ. Let us, in what follows, look at this in more detail.

4.3 Internal Hamiltonian description

I will show that for the generalized Hamiltonian system there exists an internal Hamiltonian description.

Lemma 4.6 (Internal Hamiltonian description)**.** *Let there be a symplectic space $\tilde{\Gamma}$ with a symplectic two-form $\tilde{\omega}$. Let $\mathcal{H}$ a smooth function on $\tilde{\Gamma}$. Let $\Sigma = \{(q,p) \in \tilde{\Gamma} | \mathcal{H}(q,p) = 0\}$ and $i_{\mathcal{H}} : \Sigma \to \tilde{\Gamma}$ the embedding of Σ in $\tilde{\Gamma}$. Let $\tilde{X}$ such that*

$$\tilde{\omega}\big|_{\mathcal{H}=0}(\tilde{X}, \cdot) = 0. \tag{4.42}$$

Then there exist locally on $\tilde{\Gamma}$ smooth functions τ and F with $F = \mathcal{H} - p_\tau$ such that $i^*_{\mathcal{H}}\tilde{X} = X'$ can be written as follows (up to multiplication by a scalar):

$$X' = \frac{\partial}{\partial \tau} + X_F \tag{4.43}$$

where X_F is determined by

$$\omega(X_F, \cdot) = \mathrm{d}F. \tag{4.44}$$

Here p_τ is the canoncial conjugate of τ, $\omega = i^*_\tau i^*_{\mathcal{H}} \tilde{\omega}$ is the pullback of $\tilde{\omega}$ on $\tilde{\Gamma}$ to $\Gamma = \{(q,p) \in \Sigma | \tau(q,p) = \tau^*\}$, $i_{\mathcal{H}} : \Sigma \to \tilde{\Gamma}$ denotes the embedding of Σ in $\tilde{\Gamma}$ and $i_\tau : \Gamma \to \Sigma$ the embedding of Γ in Σ.

Proof. We apply Darboux's theorem to $\tilde{\omega}$ on $\tilde{\Gamma}$.[26] Darboux's theorem says that given some differentiable, locally non-zero function p_τ on $\tilde{\Gamma}$, there exists locally some τ on $\tilde{\Gamma}$ such that $\tilde{\omega}$ can be written in the form

$$\tilde{\omega} = \mathrm{d}\tau \wedge \mathrm{d}p_\tau + \omega.$$

Choose such a p_τ. For later convenience, let us go further and choose p_τ such that it is some differentiable, locally non-zero function of the coordinates on which $\mathcal{H}$ depends. We can always make such a choice.

Since $\mathcal{H}$ is a smooth function on $\tilde{\Gamma}$ and $\mathcal{H}$ depends on p_τ, it follows that there exists a smooth function F on $\tilde{\Gamma}$ such that the Hamiltonian constraint $\mathcal{H} = 0$ can be written as

$$\mathcal{H} = F + p_\tau = 0. \tag{4.45}$$

Here F is some function of the coordinates, but not of p_τ. This allows us to write $\tilde{\omega}$ in the general form:

$$\tilde{\omega} = \mathrm{d}\tau \wedge \mathrm{d}p_\tau + \omega = (\mathrm{d}F - \mathrm{d}\mathcal{H}) \wedge \mathrm{d}\tau + \omega.$$

Let again $i_{\mathcal{H}} : \Sigma \to \tilde{\Gamma}$ denote the embedding of Σ in $\tilde{\Gamma}$ where Σ is the submanifold determined by the Hamiltonian constraint $\mathcal{H} = 0$. Then the pullback of $\tilde{\omega}$ to Σ is

$$i^*_{\mathcal{H}}\tilde{\omega} = \mathrm{d}F \wedge \mathrm{d}\tau + \omega.$$

[26] For Darboux's theorem, cf. Scheck [2003].

Let $i_\tau : \Gamma \to \Sigma$ denote the embedding of $\Gamma = \{(q,p) \in \Sigma | \tau(q,p) = \tau^*\}$ in Σ. Then ω is the pullback of $i^*_{\mathcal{H}}\tilde{\omega}$ to Γ: $\omega = i^*_\tau i^*_{\mathcal{H}}\tilde{\omega}$.

Now the two-form $i^*_{\mathcal{H}}\tilde{\omega}$ determines the physical vector field $\tilde{X}$ via the equation $\tilde{\omega}\big|_{\mathcal{H}=0}(\tilde{X},\cdot) = i^*_{\mathcal{H}}\tilde{\omega}(\tilde{X},\cdot) = 0$. Let $Y \in \mathcal{V}(\tilde{\Gamma})$ a smooth function on $\tilde{\Gamma}$. Then

$$\tilde{\omega}\big|_{\mathcal{H}=0}(\tilde{X},Y) = i^*_{\mathcal{H}}\tilde{\omega}(\tilde{X},Y) = \mathrm{d}F \wedge \mathrm{d}\tau(\tilde{X},Y) + \omega(\tilde{X},Y).$$

Here the right hand side vanishes for arbitrary Y if and only if the pullback of $\tilde{X}$ to Σ, $i^*_{\mathcal{H}}\tilde{X} = X'$, is (up to scalar multiplication) given by

$$X' = \frac{\partial}{\partial\tau} + X_F$$

and X_F such that

$$\omega(X_F,\cdot) = \mathrm{d}F.$$

In that case, the right hand side turns into

$$\mathrm{d}F \wedge \mathrm{d}\tau(X',Y) + \omega(X',Y) = -\mathrm{d}F + \mathrm{d}F = 0$$

as demanded. Note that $X' = \partial/\partial\tau + X_F$ is unique up to multiplication by a scalar. $\square$

Definition 4.17 (Internal Hamiltonian description)**.** Let everything be as in Lemma 4.6. If τ is strictly monotonic along the trajectories, we call τ the internal time parameter, Γ the internal phase space, F the internal Hamiltonian and X_F the internal Hamiltonian vector field.

Here we demand the function τ to be strictly monotonic along the trajectories. That is, the following definition applies.

Definition 4.18 (Monotonicity)**.** Let f and α smooth functions on $\tilde{\Gamma}$. Let $\tilde{X}$ the physical vector field determined by (4.42). Then $\tilde{X} = X'$ on $\Sigma = \{(q,p) \in \Gamma | \mathcal{H}(q,p) = 0\}$ and f is strictly monotonic along the trajectories if and only if

$$L_{X'}f = \alpha \tag{4.46}$$

with $\alpha = \alpha(q,p) > 0$ or $\alpha = \alpha(q,p) < 0$ for all $(q,p) \in \Sigma$.

According to this definition, α may vary (it may be a function of the coordinates), but it must be either strictly positive or strictly negative along the trajectories.

Does any τ from Lemma 4.6 fulfill the criterium of monotonicity? The answer is no. While Lemma 4.6 guarantees that locally there exists a function τ fulfilling the condition $L_{\tilde{X}}\tau \neq 0$ (otherwise τ would not be conjugate to the Hamiltonian and the construction failed), there need not exist a τ such that globally, for all $(q,p) \in \Sigma$, $L_{\tilde{X}}\tau = \alpha$ with $\alpha > 0$ or $\alpha < 0$. In other words, τ need not be monotonic along the trajectories – it might increase, then decrease, and so on, and Lemma 4.6 would still hold.

Still, in order to interpret τ as an internal time parameter, we need its monotonicity. If it were not monotonic along the trajectories, it would not foliate phase space into constant-time hypersurfaces Γ_τ such that each surface is crossed once and only once by each of the trajectories – which is just what we need in order to perform a statistical analysis of the system (and which, by the way, relates to the only sensible definition of time running in one direction). Hence, let us demand that τ is monotonic.

In what follows, we can complete the internal Hamiltonian description by proving Liouville's theorem on the internal space.

Lemma 4.7 (Invariance of the internal measure)**.** *Let everything be as in Lemma 4.6. Then the two-form ω on Γ is symplectic and conserved under the flow T_τ along X_F on Γ,*

$$T_\tau^* \omega = \omega. \tag{4.47}$$

Let $\Omega = \frac{(-1)^{[\frac{n-1}{2}]}}{(n-1)!}\omega^{n-1}$ the natural volume form on Γ as defined by (4.4) and $\mathrm{d}\mu = |\Omega|$ the natural volume measure on Γ as defined by (4.6). Then Ω and μ are conserved under T_τ as well:

$$T_\tau^* \Omega = \Omega \tag{4.48}$$

and

$$T_\tau^* \mu = \mu. \tag{4.49}$$

We call μ on Γ the internal (volume) measure.

Note that this is Liouville's theorem on the internal space Γ.

Proof. The internal space Γ is an even-dimensional submanifold of $\tilde{\Gamma}$. As such it inherits the symplectic structure from $\tilde{\Gamma}$. That is,

$$\omega = i_\tau^* i_{\mathcal{H}}^* \tilde{\omega}$$

is a symplectic two-form on Γ. Hence, in particular, ω is closed: $\mathrm{d}\omega = 0$. From this together with (4.44), it follows that the Lie derivative of ω along X_F on Γ vanishes:

$$L_{X_F}\omega = (\mathrm{d}\omega)(X_F, \cdot, \cdot) + \mathrm{d}(\omega(X_F, \cdot)) = 0 + \mathrm{d} \circ \mathrm{d}F = 0.$$

We have already shown in Corollary 4.1 that if the Lie derivative of the symplectic form ω along X_F vanishes, then ω is invariant under the flow T_τ along that vector field,

$$\frac{\mathrm{d}}{\mathrm{d}\tau} T_\tau^* \omega = L_{X_H}\omega = 0,$$

which implies that the natural volume form Ω and the natural volume measure μ are invariant under the flow (cf. Cor. 4.1). This proves the assertion. □

Remark (Space of solutions Γ_{sol})**.** Be aware that the internal phase space Γ is both the space of solutions where each point represents an entire solution and the space on which the trajectories lie (on which the internal dynamics is defined). Just like ordinary phase space, it fulfills a dual role.

Let us assume that we have a physical system and we have the standard Hamiltonian formulation of that system, that is, a formulation of the dynamics with respect to external time t. Also in this case, there may exist an internal time parameter and it may be possible to develop an internal Hamiltonian description. This is what we will later use when we look for a relational description of the Newtonian universe, trying to get rid off the notion of an absolute, external time and diminishing the number of variables that enter the statistical analysis, by that means clearing the way towards a normalizable typicality measure (cf. Sec. 6.4).

In case we start with the standard Hamiltonian formulation of a physical system (the formulation with respect to external time t), an internal time parameter is most easily found by help of the Poisson bracket.

Remark (Poisson bracket)**.** The Poisson bracket $\{\cdot,\cdot\}$ is a mathematical structure closely related to the Lie derivative and the symplectic two-form ω. Thus, we can often use one notion instead of the other. In Appendix A.2 I discuss this relation in detail. At this point, let me just remark that

$$\frac{\mathrm{d}f}{\mathrm{d}t} = \{f, H\} = L_{X_H} f = \omega(X_H, X_f)$$

where X_f is defined via the equation $\omega(X_f, \cdot) = \mathrm{d}f$. For further details, see the appendix.

Since $\mathrm{d}f/\mathrm{d}t = \{f, H\}$ (which is proven in the appendix), the monotonicity of a function f can now be determined via the Poisson bracket. Explicitly, f is (strictly) monotonic along the trajectories if and only if

$$\{f, H\} = \alpha \tag{4.50}$$

with $\alpha = \alpha(q,p) < 0$ or $\alpha = \alpha(q,p) > 0$ for all $(q,p) \in \Gamma_E$ (where $\Gamma_E = \{(q,p) \in \Gamma | H(q,p) = E\}$).

Any such f can be taken as an internal time parameter, that is, it serves as a parameter τ with respect to which the internal Hamiltonian F (the canonical conjugate of τ) and the internal Hamiltonian vector field X_F can be constructed (cf. Def. 4.17).

Example (Two free particles)**.** Let us develop the internal Hamiltonian description and construct the invariant measure on the internal space for the simple model of two classical free particles. Let the particles have unit masses $m_1 = m_2 = 1$. Consider the standard Hamiltonian formulation of that model. The particles have positions q^1 and q^2 and (conjugate) momenta p_1 and p_2. Together, q^1, q^2, p_1, p_2 form canonical coordinates of phase space Γ. The motion of the particles is governed by the Hamiltonian

$$H = \frac{1}{2}p_1^2 + \frac{1}{2}p_2^2. \tag{4.51}$$

The equations of motion with respect to external time t are

$$p_1 = p_1(0), p_2 = p_2(0) \quad \text{and} \quad q^1 = p_1 t + q^1(0), q^2 = p_2 t + q^2(0). \tag{4.52}$$

These are the solutions to the standard Hamiltonian laws of motion. Since we consider an isolated system, energy is conserved, $H = E$, and the trajectories are restricted to the constant energy hypersurface Γ_E.

In what follows, we want to read this system as a generalized Hamiltonian system. That is, we read Γ as generalized phase space $\tilde{\Gamma}$ on which there exists a symplectic two-form

$$\tilde{\omega} = \mathrm{d}q^1 \wedge \mathrm{d}p_1 + \mathrm{d}q^2 \wedge \mathrm{d}p_2. \tag{4.53}$$

In addition, we have a Hamiltonian constraint $\mathcal{H} = H - E = 0$. Since $H = H(q,p) = \frac{1}{2}p_1^2 + \frac{1}{2}p_2^2$, the Hamiltonian constraint can be written as

$$\frac{1}{2}p_1^2 + \frac{1}{2}p_2^2 - E = 0. \tag{4.54}$$

Let us rearrange terms to separate one of the variables. We get

$$p_1 - \sqrt{2E - p_2^2} = 0. \tag{4.55}$$

This way of writing the Hamiltonian constraint already suggests a particular internal time parameter, $\tau = q^1$, and a particular internal Hamiltonian

$$F = -\sqrt{2E - p_2^2}. \tag{4.56}$$

This can be seen as follows. The way F is defined, it fulfills the equation $p_1 + F(q^a, p_a) = 0$ with $a = 2, ..., n$ which ensures that F is the canonical conjugate of $\tau = q^1$ governing the evolution of the internal coordinates q^a, p_a $(a = 2, ..., n)$ with respect to internal time $\tau = q^1$. It remains to check that q^1 is monotonic. This can be done by help of the Poisson bracket. It is

$$\frac{\mathrm{d}q^1}{\mathrm{d}t} = \{q^1, H\} = \sum_{i=1}^{n} \left(\frac{\partial q^1}{\partial q^i}\frac{\partial H}{\partial p_i} + \frac{\partial q^1}{\partial p_i}\left(-\frac{\partial H}{\partial q^i}\right)\right) = p_1 \tag{4.57}$$

and

$$\frac{\mathrm{d}p_1}{\mathrm{d}t} = \{p_1, H\} = \sum_{i=1}^{n} \left(\frac{\partial p_1}{\partial q^i}\frac{\partial H}{\partial p_i} + \frac{\partial p_1}{\partial p_i}\left(-\frac{\partial H}{\partial q^i}\right)\right) = 0. \tag{4.58}$$

That is,

$$\frac{\mathrm{d}q^1}{\mathrm{d}t} = p_1(0). \tag{4.59}$$

So the rate of change of q^1 along the trajectories is constant (constantly $p_1(0)$). As such, q^1 is strictly monotonic along the trajectories (except if $p_1(0) = 0$ which we want to exclude for means of generality) and serves as an internal time parameter: $\tau = q^1$.

Let now $i_{\mathcal{H}} : \Sigma \to \tilde{\Gamma}$ be the embedding of Σ in $\tilde{\Gamma}$ where Σ is defined by $\mathcal{H} = H(q,p) - E = 0$. In accordance with the generalized formalism, the physical vector field is defined by (4.42), that is, by the equation

$$\tilde{\omega}\big|_{\mathcal{H}=0}(\tilde{X}, \cdot) = i^*_{\mathcal{H}}\tilde{\omega}(\tilde{X}, \cdot) = 0.$$

Here $\tilde{X} = X'$ on Σ and $i^*_{\mathcal{H}}\tilde{\omega}$ is the pullback of the two-form $\tilde{\omega}$ on $\tilde{\Gamma}$ to Σ. To find the pullback, let us rewrite $\tilde{\omega}$ using the functional form of $\mathcal{H}$ to replace p_1. Then

$$\begin{aligned}\tilde{\omega} &= \mathrm{d}q^1 \wedge \mathrm{d}p_1 + \mathrm{d}q^2 \wedge \mathrm{d}p_2 \\ &= \mathrm{d}q^1 \wedge \left(\frac{1}{\sqrt{2(\mathcal{H}+E) - p_2^2}}\mathrm{d}\mathcal{H} - \frac{p_2}{\sqrt{2(\mathcal{H}+E) - p_2^2}}\mathrm{d}p_2\right) + \mathrm{d}q^2 \wedge \mathrm{d}p_2. \end{aligned} \tag{4.60}$$

The pullback of $\tilde{\omega}$ to Σ can now be obtained directly. On Σ, $\mathcal{H} = 0$ and $\mathrm{d}\mathcal{H} = 0$ and, consequently,

$$i^*_{\mathcal{H}}\tilde{\omega} = -\frac{p_2}{\sqrt{2E - p_2^2}}\mathrm{d}q^1 \wedge \mathrm{d}p_2 + \mathrm{d}q^2 \wedge \mathrm{d}p_2. \tag{4.61}$$

This two-form determines the physical vector field X' on Σ. Remember that the physical vector field is determined by $i^*_H\tilde{\omega}(X', \cdot) = 0$ (cf. Eq. (4.42)). With $i^*_H\tilde{\omega}$ given by (4.61) if follows that (up to scalar multiplication)

$$X' = \frac{\partial}{\partial q^1} + \frac{p_2}{\sqrt{2E - p_2^2}}\frac{\partial}{\partial q^2} + 0\frac{\partial}{\partial p_2}. \tag{4.62}$$

This can be seen as follows. Let $Y \in \mathcal{V}(\Sigma)$. Then

$$\begin{aligned} i^*_H\omega(X', Y) &= \left[-\frac{p_2}{\sqrt{2E - p_2^2}}\mathrm{d}q^1 \wedge \mathrm{d}p_2 + \mathrm{d}q^2 \wedge \mathrm{d}p_2\right]\left(\frac{\partial}{\partial q_1} + \frac{p_2}{\sqrt{2E - p_2^2}}\frac{\partial}{\partial q^2}, Y\right) \\ &= -\frac{p_2}{\sqrt{2E - p_2^2}}\mathrm{d}p_2(Y) + \frac{p_2}{\sqrt{2E - p_2^2}}\mathrm{d}p_2(Y) = 0. \end{aligned}$$

Let us now rewrite everything with respect to internal time $\tau = q^1$ and internal Hamiltonian $F = -\sqrt{2E - p_2^2}$. Then (4.61) becomes

$$i^*_{\mathcal{H}}\tilde{\omega} = \mathrm{d}F \wedge \mathrm{d}\tau + \mathrm{d}q^2 \wedge \mathrm{d}p_2 \tag{4.63}$$

and (4.62) turns into

$$\begin{aligned} X' &= \frac{\partial}{\partial \tau} + \frac{\partial F}{\partial p_2}\frac{\partial}{\partial q^2} + 0\frac{\partial}{\partial p_2} \\ &= \frac{\partial}{\partial \tau} + \frac{\partial F}{\partial p_2}\frac{\partial}{\partial q^2} - \frac{\partial F}{\partial q^2}\frac{\partial}{\partial p_2} \end{aligned} \tag{4.64}$$

where we used that $-\partial F/\partial q^2 = 0$.

Let us now determine the internal Hamiltonian description on the space of constant internal time $\tau = \tau^*$. Let therefore $i_\tau : \Gamma_\tau \to \Sigma$ denote the embedding of Γ_τ in Σ where Γ_τ is the hypersurface of constant $\tau = \tau^*$. Let $\omega := i^*_\tau i^*_{\mathcal{H}}\tilde{\omega}$ denote the pullback of $i^*_{\mathcal{H}}\tilde{\omega}$ on Σ to Γ_τ. On Γ_τ,

$\mathrm{d}\tau = 0$ and with $i_{\mathcal{H}}^*\tilde{\omega}$ given by (4.63), it follows that $\omega = i_\tau^* i_{\mathcal{H}}^* \tilde{\omega}$ is

$$\omega = \mathrm{d}q^2 \wedge \mathrm{d}p_2. \tag{4.65}$$

Now rewrite X' from (4.64) as $X' = \partial/\partial\tau + X_F$. That is,

$$X_F = \frac{\partial F}{\partial p_2}\frac{\partial}{\partial q^2} - \frac{\partial F}{\partial q^2}\frac{\partial}{\partial p_2}. \tag{4.66}$$

It is now easy to see that X_F fulfills the standard equation

$$\omega(X_F, \cdot) = \mathrm{d}F. \tag{4.67}$$

That is, X_F is a Hamiltonian vector field on Γ_τ. It governs the evolution of the internal variables q^2, p_2 with respect to internal time τ.

To be precise, describing the system on the internal space, internal time $\tau = q^1$ has become an external time parameter and F given by (4.56) has become a standard Hamiltonian determining a Hamiltonian vector field X_F via the Hamiltonian equation $\omega(X_F, \cdot) = \mathrm{d}F$. Again, possible trajectories are the integral curves along X_F parametrized by τ: $\gamma'(\tau) = (X_F)_{\gamma(\tau)}$. Here $\gamma'(\tau)$ denotes the derivative of γ with respect to τ and $\gamma(\tau) = (q^i(\tau), p_i(\tau))$. In local coordinates,

$$\frac{\mathrm{d}q^2}{\mathrm{d}\tau} = \frac{\partial F}{\partial p_2} = \frac{p_2}{\sqrt{2E - p_2^2}}, \qquad \frac{\mathrm{d}p_2}{\mathrm{d}\tau} = -\frac{\partial F}{\partial q^2} = 0 \tag{4.68}$$

where, again, we used that $F = -\sqrt{2E - p_2^2}$.

How does one interpret these equations? What we found are the equations of motion of particle 2 with respect to "internal time" $\tau = q^1$. In other words, the position of particle 1 provides a clock with respect to which the motion of particle 2 can be described. This description is again Hamiltonian. We can say that it is the Hamiltonian description of a subsystem (particle 2) relative to its environment (particle 1).

One can, of course, also get the equations of motion of particle 2 with respect to time $\tau = q^1$ directly. Start from full phase space $\tilde{\Gamma}$ and solve the equations of motion on $\tilde{\Gamma}$ for given initial conditions $q^1(0) = q^2(0) = 0$ and $p_1(0) = p_1$, $p_2(0) = p_2$. The solutions are $q^1(t) = p_1 t$ and $q^2(t) = p_2 t$ as well as $p_1(t) = p_1$ and $p_2(t) = p_2$. Rewriting these equations for particle 2 with respect to "time" $q^1 = p_1/t$ leads to

$$q^2(q^1) = \frac{p_2}{p_1}q^1 = \frac{p_2}{\sqrt{2E - p_2^2}}q^1, \qquad p_2(q^1) = p_2.$$

These are (for the given initial data) solutions to the equations of motion (4.68) from above.

Now what about the measure on the internal space Γ_τ? Note that $\omega = \mathrm{d}q^2 \wedge \mathrm{d}p_2$ is already the final volume form Ω on Γ_τ. This is because, in this example, the internal space is two-dimensional.

From Ω we get a volume measure:

$$\mathrm{d}\mu = |\Omega| = \mathrm{d}q^2 \mathrm{d}p_2. \tag{4.69}$$

Again, it follows from the fact that ω is closed and $\omega(X_f, \cdot) = \mathrm{d}F$ that ω is conserved under the Hamiltonian flow T_τ connected to X_F:

$$L_{X_F}\omega = 0 \tag{4.70}$$

From this it follows that the volume measure μ is invariant with respect to internal time evolution:

$$T_\tau^* \mu = \mu. \tag{4.71}$$

This is Liouville's theorem for the subsystem consisting of particle 2 alone. It just says that the uniform measure $\mu = \mathrm{d}q^2 \mathrm{d}p_2$, suitable for the statistical analysis of the subsystem, is conserved under internal time evolution.

In Appendix B you find another example for the internal Hamiltonian description, based on the minisuperspace model. In Section 6.4.1, we will apply the general scheme of an internal Hamiltonian formulation to the physical model we are interested in: the $E = 0$ Newtonian universe.

Remark (GHS measure)**.** The idea to construct a measure on hypersurfaces of constant internal time has first been introduced by Gibbons, Hawking, and Stuart in their [1987] paper "A natural measure on the set of all universes". While they do not develop the full internal Hamiltonian description I give above, they construct the measure by demanding it to fulfill certain criteria, like to be invariant with respect to some monotonic parameter foliating relativistic spacetime. They then construct the measure on one of the hypersurfaces of the given foliation and take it to be a measure on the set of solutions. In fact, the measure they construct is just the measure we obtain from the internal Hamiltonian formulation.

As an example Gibbons, Hawking, and Stuart discuss the minisuperspace model, for which I describe the full internal Hamiltonian description in Appendix B. That particular model (and measure) have later been discussed by Gibbons and Turok [2008], Carroll and Tam [2010], and Carroll [2014]. None of them, however, gives the internal Hamiltonian description.

5 A measure on reduced phase space Γ_{red}

Let us now consider a Hamiltonian system with symmetries. Later we want to discuss the Newtonian universe, that is, the model of N particles moving through three-dimensional Euclidean space according to Newton's law of gravitation. This system is invariant under spatial translations and rotations. While usually the dynamics is formulated on $6N$-dimensional phase space Γ, the symmetries allow for a description of the system in terms of less than $6N$ variables, respectively, the dynamics can be formulated on a lower-dimensional space – so called reduced phase space Γ_{red}.

Just like the internal time parametrization, the reduction by symmetries will help us to diminish the number of variables entering the statistical analysis of the system. This will in the end enable us to obtain a normalizable typicality measure for the Newtonian universe.

Reduced phase space Γ_{red} can be obtained from phase space Γ by a method called symplectic reduction. Sometimes Γ_{red} can be identified with the cotangent bundle of reduced configuration space T^*Q_{red} which would then provide an alternative way of determining reduced phase space (cf. Sec. 5.1). In what follows, I want to present the general method of symplectic reduction before I apply it to the special case of three particles and the symmetry group of translations, rotations, and dilations.

For the mathematical details connected to symplectic reduction, see Marsden and Weinstein [1974], the respective section in Abraham and Marsden [1978], Iwai [1987], Iwai and Yamaoka [2005] and the respective section in Arnol'd [1989]. For the symmetry group of dilations, see also Tokasi [2017|. My notation will follow the one of Arnol'd.

5.1 Introduction and notation

Let there be a Lie group G of symmetries acting on configuration space Q. Clearly, this defines an action on the cotangent bundle of configuration space, phase space $\Gamma = T^*Q$, as well. This allows for the reduction of phase space.

5.1.1 Nöther's theorem

Nöther's theorem tells us that connected to every symmetry (which is usually defined by the Lagrangian or Hamiltonian being invariant under the respective mapping) there exists a conserved quantity (also called a first integral) of motion. If there are several first integrals of motion, the common level manifold $M \subset \Gamma$ of these first integrals is an invariant manifold of the phase flow. That is, trajectories which start on M do not leave M. Now consider the subgroup G_p of the symmetry group G which maps this manifold onto itself. The subgroup G_p is then said to leave M fixed. In that case, we are allowed to factorize M by the action of G_p. The resultant quotient space is the reduced phase space $\Gamma_{red} = M/G_p$. It is again a symplectic space (where the symplectic structure is induced by the symplectic structure on Γ) and the reduced dynamics is again Hamiltonian (induced by the Hamiltonian dynamics on Γ).

Let us look at this more closely. Let G be the Lie group of symmetry transformations acting on Q (with an induced action on $\Gamma = T^*Q$). In the physical examples to come, $Q = \mathbb{R}^{3N}$ and $\Gamma \cong \mathbb{R}^{6N}$ and G will either be the Euclidean group E(3) of translations and rotations or the similarity group Sim(3) of translations, rotations, and dilations. Each one-parameter group of transformations g_t defines a phase flow on Γ. This flow connects to a Hamiltonian function H via Nöther's formula.

Definition 5.1 (Nöther's formula)**.** Let Q be an n-dimensional manifold and $\Gamma = T^*Q$. Let G be a Lie group of symmetry transformations and $\mathbf{g}$ the Lie algebra of G. Let $g_t \in \mathbf{g}$, $x \in Q$ and let q^j, p_j be local coordinates on Γ. Then

$$H(x) = \sum_{j=1}^{n} p_j \mathrm{d}q^j \left(\left. \frac{\mathrm{d}}{\mathrm{d}t} \right|_{t=0} g_t x \right) \tag{5.1}$$

is the Hamiltonian function connected to g_t.

Let us reintroduce physics notation. Let $\mathbf{q} = (\mathbf{q}^1, ..., \mathbf{q}^N) \in Q$ where $Q = \mathbb{R}^{3N}$. Nöther's formula shows that related to translations $\mathbf{q} \to \mathbf{q} + \mathbf{a}$ the conserved quantity is total linear momentum $\mathbf{P} = \sum_{i=1}^{N} \mathbf{p}_i$. Related to rotations $\mathbf{q} \to \mathbf{q} + \boldsymbol{\theta} \times \mathbf{q}$ the conserved quantity is total angular momentum $\mathbf{L} = \sum_{i=1}^{N} \mathbf{p}_i \times \mathbf{q}^i$ and related to dilations (scalings) $\mathbf{q} \to \mathbf{q} + \lambda \mathbf{q}$ the conserved quantity is dilational momentum $D = \sum_{i=1}^{N} \mathbf{p}_i \cdot \mathbf{q}^i$. Here $\mathbf{a}, \boldsymbol{\theta}$, and λ are infinitesimal parameters of translation, rotation, and dilation.

5.1.2 Symplectic reduction of phase space

Marsden and Weinstein [1974] develop the general method of symplectic reduction. Let us, in what follows, outline it. This is really meant only as an outline, so I refer to the literature for more details. I will exemplify this method – and this is all we need – on the example of three particles and the similarity group Sim(3)=$\mathbb{R}^3 \times SO(3) \times \mathbb{R}^+$ in Section 5.2.

Symplectic reduction in general. Let G be a Lie group of symmetry transformations acting on $\Gamma = T^*Q$. Let $\mathbf{g}$ be the Lie algebra of G and $\mathbf{g}^*$ the dual of $\mathbf{g}$.[27] Then we can define the momentum mapping

$$P : \Gamma \to \mathbf{g}^* \tag{5.2}$$

as a mapping from phase space to the dual space of the Lie algebra of the group G. We want an object which, for any point $x \in \Gamma$, takes an element g_i of the Lie algebra $\mathbf{g}$ of G and assigns to it the value of the respective Hamiltonian function H_i according to (5.1):

$$p_x(g_i) = H_i(x). \tag{5.3}$$

This p_x is the element of the dual space of the Lie algebra $\mathbf{g}^*$ which is connected to x:

$$P(x) = p_x. \tag{5.4}$$

For $p \in \mathbf{g}^*$, let

$$M_p = P^{-1}(p) \tag{5.5}$$

denote the level set $M_p \subset \Gamma$.

Let G_p the subgroup of G which leaves M_p fixed (maps M_p onto itself). The reduced phase space Γ_p is obtained from M_p by forming the quotient with respect to the stationary subgroup

[27] The Lie algebra is a vector space $\mathbf{g}$ over a field F together with a binary operation which is bilinear, alternating and fulfills the Jabobi identity: $[\cdot, \cdot] : \mathbf{g} \times \mathbf{g} \to \mathbf{g}$, called the Lie bracket. Every Lie group gives rise to a Lie algebra. The dual space of the Lie algebra $\mathbf{g}^*$ is the set of all linear maps (linear functionals) $\phi : \mathbf{g} \to F$.

G_p,

$$\Gamma_p = \frac{M_p}{G_p}. \tag{5.6}$$

It is the base space of a fibre bundle

$$\pi : M_p \to \Gamma_p \tag{5.7}$$

where the fibres are the orbits of G_p. Here an orbit $O(x)$ is the set of points in M_p to which a given point $x \in M_p$ can be moved by the elements of the group G_p, that is,

$$O(x) = \{gx | g \in G_p\}. \tag{5.8}$$

In other words, the quotient space is the set of equivalence classes of points which are connected by the respective symmetry transformation. Alternatively, the quotient space can be represented by choosing one point of the respective orbit. Thus, every symmetry connects to a "gauge" degree of freedom which we are allowed to fix.

In what follows, it will be important that the reduced phase space Γ_p has a symplectic structure and that there exist reduced Hamiltonian dynamics on Γ_p which are induced by the symplectic structure and the Hamiltonian dynamics on Γ, respectively.

Definition 5.2 (Symplectic form ω_p)**.** Let Γ_p and M_p be as above. Let $i : M_p \to \Gamma$ be the embedding of M_p in Γ and let $\pi : M_p \to \Gamma_p$ be the fibre bundle. Let i^* and π^* denote the pullback. Let ω be the symplectic form on Γ. Then

$$i^*\omega = \pi^*\omega_p \tag{5.9}$$

determines the symplectic form ω_p on Γ_p.

Later we will construct the reduced phase space by choosing a particular hypersurface perpendicular to the orbits (a particular gauge).

From ω_p we can determine a natural volume form Ω_p in accordance with Def. 4.4. Let the reduced phase space Γ_p be a space of $2k$ dimensions. Then

$$\Omega_p = (-1)^{[\frac{k}{2}]}\frac{\omega_p^k}{k!}. \tag{5.10}$$

Here $[k/2]$ is the biggest integer smaller or equal to $k/2$. This volume form Ω_p determines a natural volume measure μ_p on Γ_p in accordance with Def. 4.5. That is,

$$\mathrm{d}\mu_p = |\Omega_p|. \tag{5.11}$$

Moreover, we have reduced Hamiltonian equations on Γ_p. Let H be a smooth function on Γ, the Hamiltonian of the system. Let H be invariant under the action of the group G_p. Then the reduced Hamiltonian H_p on Γ_p is determined as follows.

Definition 5.3 (Reduced Hamiltonian H_p)**.** Let everything be as in the previous definition. Let H be a smooth function on Γ, the Hamiltonian of the system, invariant under the action of G_p.

Then

$$i^*H = \pi^*H_p \tag{5.12}$$

determines the reduced Hamiltonian H_p on Γ_p.

For more details, cf. Marsden and Weinstein [1974] and Arnol'd [1989].

Symplectic reduction for translations. Let us consider a Hamiltonian system of N particles specified by the quadruple $(\Gamma, \mathcal{B}(\Gamma), \omega, H)$ with $\Gamma = T^*Q$ and $Q = \mathbb{R}^{3N}$. Let the Hamiltonian H of the system be invariant under spatial translations $\mathbf{q}^i \to \mathbf{q}^i + \mathbf{a}\ \forall i = 1, ..., N$ with $\mathbf{a} \in \mathbb{R}^3$. Then the total linear momentum $\mathbf{P} = \sum_{i=1}^N \mathbf{p}_i$ is conserved: $\{H, \mathbf{P}\} = 0$. Let M_p denote the level set related to total linear momentum: $M_p = \{(\mathbf{q}, \mathbf{p}) \in \Gamma | \sum_{i=1}^N \mathbf{p}_i = p\}$. The group G_p which, for any value p of $\mathbf{P}$, leaves M_p fixed is the group of translations: $G_p = \mathbb{R}^3$. The reduced phase space is the quotient space

$$\Gamma_p = \frac{M_p}{G_p}. \tag{5.13}$$

As we have seen, factoring by the action of a symmetry group is equivalent to choosing one point of the respective orbit. In this case, one possible choice is made by fixing the center of mass to the origin: $\mathbf{Q}_{cm} = \sum_{i=1}^N m_i \mathbf{q}^i = 0$. Thus, in total, Γ_p can be obtained from Γ by imposing the following six constraints:

$$\mathbf{P} = \sum_{i=1}^N \mathbf{p}_i \quad = \quad p, \tag{5.14}$$

$$\mathbf{Q}_{cm} = \sum_{i=1}^N m_i \mathbf{q}^i \quad = \quad 0. \tag{5.15}$$

These constraints can be imposed by help of a convenient choice of coordinates (e.g., Jacobi coordinates and their canonical conjugates).

What is the dimension of reduced phase space Γ_p? We reduce by three dimensions when going to the level set M_p (which is defined by $\mathbf{P} = p$) and we reduce by another three dimensions when factoring out translations (by setting $\mathbf{Q}_{cm} = 0$). Hence, the resultant quotient space $\Gamma_p = M_p/G_p$ is $(6N-6)$-dimensional.

Symplectic reduction for rotations. Let, in addition, the Hamiltonian H be invariant under rotations $\mathbf{q}_i \to \boldsymbol{\theta} \times \mathbf{q}_i\ \forall i = 1, ..., N$. The symmetry group G related to spatial translations and rotations is the Euclidean group: $G = \mathrm{E}(3) = \mathbb{R}^3 \times SO(3)$. Since the system is invariant under rotations, total angular momentum $\mathbf{L} = \sum_{i=1}^N \mathbf{q}^i \times \mathbf{p}_i$ is conserved: $\{H, \mathbf{L}\} = 0$. The level set M_l related to total angular momentum is $M_l = \{(\mathbf{q}, \mathbf{p}) \in \Gamma_p | \mathbf{L} = \sum_{i=1}^N \mathbf{q}^i \times \mathbf{p}_i = l\}$. Let $G_l \subset G$ denote the subgroup of G which leaves M_l fixed. The reduced phase space Γ_l is now the quotient space

$$\Gamma_l = \frac{M_l}{G_l}. \tag{5.16}$$

Remark (Singular configurations)**.** We want to only consider non-singular configurations. The term "singular configurations" refers to the points of total collision (where all particles are at one point) and collinear configurations – any other configuration is called non-singular. Total collisions, collinear configurations and non-singular configurations define different strata of configuration space and, in principle, reduction has to be done separately for each stratum since each stratum has its own symmetry group. This leads to the method of stratified reduction of configuration space.[28] It has, however, been shown by Iwai [2005] that the Hamiltonian equations of motion for non-singular configurations reduce in the limit of going to the boundary of the stratum to those for collinear configurations. This is all we need, so that, in what follows, we will stick to non-singular configurations and present the method of symplectic reduction for those.

What is the subgroup G_l which maps M_l onto itself? To determine G_l, we first need to take into account that M_l is a subset of Γ_p. If $\mathbf{P} = 0$, then the full group $SO(3)$ acts on Γ_p. This follows from the fact that, for $\mathbf{P} = 0$, the equations (4.15) and (4.16) determining Γ_p are invariant under $SO(3)$. However, this is not the case if $\mathbf{P} = p \neq 0$. Then only a subgroup of $SO(3)$ acts on Γ_p. Assume, in the following, that $\mathbf{P} = 0$.

It turns out that G_l is different for $\mathbf{L} = 0$ and $\mathbf{L} \neq 0$, respectively. Let $M_{l=0} \subset \Gamma_p$ denote the level set for $\mathbf{L} = l = 0$. If $\mathbf{L} = 0$, then the full rotational group $SO(3)$ acts on M_0 leaving it fixed. In contrast, if $\mathbf{L} \neq 0$, this is only $SO(2)$. This is the case because, for $\mathbf{L} = 0$, the system is rotationally invariant around all three axes while, for $\mathbf{L} \neq 0$, the system is rotationally invariant only with respect to the axis which is parallel to $\mathbf{L}$.[29] Hence, in the latter case, M_l is mapped onto itself only by the action of $SO(2)$. Note that $SO(2) = S^1$ is the circle action.

Since $SO(3)$ is a three-dimensional group whereas $SO(2)$ is a one-dimensional group, the respective quotient spaces will be different in dimension by two:

$$dim\left(\frac{M_{l=0}}{SO(3)}\right) = dim\left(\frac{M_{l\neq 0}}{SO(2)}\right) + 2. \tag{5.17}$$

In case $\mathbf{P} = 0$ and $\mathbf{L} = 0$, the reduced phase space $\Gamma_l = M_{l=0}/SO(3)$ is a $(6N - 12)$-dimensional subspace of Γ. However, in contrast to Γ_p, Γ_l is not a submanifold. This means that, having the fibre bundle picture in mind, Γ_l cannot be identified with a smooth hypersurface cutting the fibres transversally. The reason is the following. If we walk along a closed cycle on Γ_l – the base space of the fibre bundle $\pi_l : M_l \to \Gamma_l$ – we will end up at a different point on the fibre

[28] Cf. Iwai [2005].

[29] In practice, the dimension of the symmetry subgroup which leaves the common level set fixed can always be determined by computing the Poisson brackets of the first integrals of motion. Consider, for example, the three components of angular momentum $\mathbf{L}$. While Nöther's theorem tells us that a rotation around the $\hat{x}$-axis conserves the angular momentum L_x in that direction and analogously for a rotation around the $\hat{y}$-axis, it is not the case that a rotation around the $\hat{x}$-axis leaves L_y invariant and vice versa. This is captured by the fact that $\{L_x, L_y\} \neq 0$. Here L_x can be read as the generator of rotations around the $\hat{x}$-axis and the given Poisson bracket determines the change of L_y with respect to that generator. The number of first integrals whose mutual Poisson brackets vanish gives us the dimension of the symmetry subgroup.

(corresponding to a performed rotation/a change in orientation). But this cannot happen if we walk along a smooth hypersurface transversally intersecting the fibres. The physical meaning of this peculiar global property of Γ_l is the following. While total angular momentum is constantly zero, $\mathbf{L} = 0$, a change in the internal variables (local coordinates of Γ_l) can induce an overall rotation. In other words, a deformation of the shape of a body can induce a rotation. In the literature, this is known as the phenomenon of the "falling cat".[30] A cat can induce a rotation just by deforming its body. This allows the cat to come down on her feet even if she starts falling with the feet pointing to the sky.

Littlejohn and Reinsch [1997], however, show that we can still loosely identify Γ_l with a section Σ cutting the fibres transversely.[31] This identification is "loose" in the sense that it allows us to determine the correct local structure, but not the global one. The section Σ is determined by fixing constraints just the way we did it above. One way to do this is by setting the off-diagonal terms of the center-of-mass inertia tensor $\mathbf{I}_{cm} = \sum_{i=1}^{3} m_i(\mathbf{q}^i_{cm} \cdot \mathbf{q}^i_{cm}\, \mathbb{I} - \mathbf{q}^i_{cm} \otimes \mathbf{q}^i_{cm})$ to zero. Since the off-diagonal terms of $\mathbf{I}_{cm}$ define the so-called principal axes of inertia, this is known as the principal-axis-gauge. If we set them to zero, we fix the axes of our coordinate system to the principal axes of inertia. That way, the orientation is fixed by definition whereas, if we walk along a closed cycle on Γ_l, we can change the orientation. Still, we obtain the correct local structure of Γ_l from Γ_p by imposing the following six constraints:

$$\mathbf{L} = \sum \mathbf{q}_i \times \mathbf{p}_i = 0, \tag{5.18}$$

$$\mathbf{I}^{12}_{cm} = 0, \qquad \mathbf{I}^{13}_{cm} = 0, \qquad \mathbf{I}^{23}_{cm} = 0. \tag{5.19}$$

While the constraints (5.18) determine the level set $M_l \subset \Gamma_p$, the constraints (5.19) determine the section Σ which we loosely identify with Γ_l.

With respect to the reduced Hamiltonian equations of motion the following occurs. For total angular momentum $\mathbf{L} \neq 0$, when we project the motion on the internal space (reduced phase space), then there appear some additional terms which we identify as "Coriolis forces". For $\mathbf{L} = 0$, these terms vanish.

Symplectic reduction for dilations. Let, in addition, the Hamiltonian H be invariant under dilations (scalings) $\mathbf{q}^i \to \lambda \mathbf{q}^i\ \forall i = 1, ..., N$ with $\lambda \in \mathbb{R}^+$. Then also dilational momentum $D = \sum \mathbf{q}^i \cdot \mathbf{p}_i$ is conserved: $\{H, D\} = 0$. In that case, $D = \sum \mathbf{q}^i \cdot \mathbf{p}_i = d$. Reduced phase space Γ_d is now the quotient space

$$\Gamma_d = \frac{M_d}{G_d} \tag{5.20}$$

where $M_d \subset \Gamma_l$ is the set of all points $x \in \Gamma_l$ for which $\mathbf{D} = d$ and $G_d \subset G$ is the isotropy group of dilations leaving M_d fixed.

Now what is the correct subgroup G_d leaving M_d fixed? Again, G_d is different for $D = 0$ and $D \neq 0$. If $D = 0$, then the full dilational group $\mathbb{R}^+$ acts on $M_{d=0}$ leaving it fixed. However, if

[30] For a thorough discussion of the falling cat problem, see Littlejohn and Reinsch [1997].
[31] For this, cf. Littlejohn and Reinsch [1997].

$D = d \neq 0$, then G_d is the identity e.

Since $\mathbb{R}^+$ is a one-dimensional group whereas the identity e is zero-dimensional, the respective quotient spaces will be different in dimension by one:

$$dim\left(\frac{M_{d=0}}{\mathbb{R}^+}\right) = dim\left(\frac{M_{d\neq 0}}{e}\right) + 1. \tag{5.21}$$

Given that $\mathbf{P} = 0$, $\mathbf{L} = 0$, and $D = 0$, the reduced phase space $\Gamma_d = M_{d=0}/\mathbb{R}^+$ is a space of $6N - 14$ dimensions. It can be obtained from translationally and rotationally reduced phase space Γ_l by imposing the following two constraints:

$$D = \sum \mathbf{q}^i \cdot \mathbf{p}_i = 0 \tag{5.22}$$

and

$$I = \sum m_i |\mathbf{q}^i|^2 = 1. \tag{5.23}$$

Here I is the moment of inertia in the center-of-mass frame (where, due to the translational symmetry, $\mathbf{Q}_{cm} = 0$). While (5.22) fixes the constant of motion, (5.23) fixes the scale of the system.

5.1.3 Reduced configuration space Q_{red} and its relation to Γ_{red}

In case all the conserved quantities (like $\mathbf{P}$, $\mathbf{L}$ or D) are identically zero, there is an alternative way to construct the reduced phase space. In that case, Γ_{red} can be identified with the cotangent bundle of reduced configuration space T^*Q_{red} where reduced configuration space Q_{red} is a quotient space obtained from Q by factoring out the symmetries of the system. To be precise, if the conserved quantities are identically zero, then Γ_{red} and T^*Q_{red} are diffeomorphic.[32]

How do we obtain Q_{red}? Let G be a Lie group of symmetry transformations acting on $3N$-dimensional configuration space Q. Then $Q_{red} = Q/G$.

Definition 5.4 (Center-of-mass configuration space Q_{cm})**.** Let $Q = \mathbb{R}^{3N}$. Let $G_p = \mathbb{R}^3$ the group of translations. Then

$$Q_{cm} = \frac{Q}{G_p} \tag{5.24}$$

is the reduced configuration space with respect to the group of translations. We also call it the *center-of-mass configuration space.*

Definition 5.5 (Shape space with scale S_R)**.** Let $Q = \mathbb{R}^{3N}$. Let E(3)= $\mathbb{R}^3 \times SO(3)$ the Euclidean group of translations and rotations. Then

$$S_R = \frac{Q}{\mathrm{E}(3)} \tag{5.25}$$

is the reduced configuration space with respect to the Euclidean group. We call S_R *shape space with scale.*

[32] Cf. Arnol'd [1989].

While Q_{cm} is a $(3N-3)$-dimensional space, S_R is a $(3N-6)$-dimensional space. Why do we call it shape space with scale? In the standard Hamiltonian formulation, the system's configuration is described by $3N$ position variables. If we factor out translations and rotations, that is, if we identify all those configurations which are connected by an overall translation and/or rotation, then all that remains in the description is the shape determined by the angles between the particles and the scale. If we finally take out scale by quotienting with respect to the group of dilations, we end up with *shape space* S.

Definition 5.6 (Shape space S)**.** Let $Q = \mathbb{R}^{3N}$. Let Sim(3)$= \mathbb{R}^3 \times SO(3) \times \mathbb{R}^+$ the similarity group of translations, rotations and (orientation-preserving) dilations. Then

$$S = \frac{Q}{\mathrm{Sim}(3)} \tag{5.26}$$

is the reduced configuration space with respect to the similarity group. We call S *shape space.*

On shape space, all that remains of the description of the system's configuration are angles, respectively relative distances between particles. There is neither an absolute position nor an absolute orientation nor an absolute scale left. Since we quotient by the similarity group, which is 7-dimensional, shape space S is a $(3N-7)$-dimensional space. In other words, the shape of the system is determined by $3N-7$ shape degrees of freedom (angles or relative distances).

Definition 5.7 (Shape phase space with scale)**.** Let S_R be shape space with scale. Its cotangent bundle T^*S_R is called *shape phase space with scale.*

Definition 5.8 (Shape phase space)**.** Let S be shape space. Its cotangent bundle T^*S is called *shape phase space.*

It can now be shown that, in case the Hamiltonian is invariant under translations and rotations and $\mathbf{P} = \mathbf{L} = 0$, then the reduced phase space Γ_l is symplectic and diffeomorphic to the cotangent bundle of the reduced configuration space S_R:

$$\Gamma_l \cong T^*S_R. \tag{5.27}$$

Analogously, in case the Hamiltonian is invariant under translations, rotations, and dilations and $\mathbf{P} = \mathbf{L} = D = 0$, then the reduced phase space Γ_d is symplectic and diffeomorphic to the cotangent bundle of shape space S:

$$\Gamma_d \cong T^*S. \tag{5.28}$$

5.2 Reduction of phase space for 3 particles and $\mathbf{P} = \mathbf{L} = D = 0$

As an example consider a system of $N = 3$ particles moving through three-dimensional Euclidean space. On ordinary phase space $\Gamma \cong \mathbb{R}^{6N}$, the system is described by totally $6N = 18$ coordinates, nine position coordinates $\mathbf{q}^1, \mathbf{q}^2, \mathbf{q}^3$ and nine (conjugate) momenta $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$. Assume the Hamiltonian is invariant under translations, rotations, and dilations. Hence, $\mathbf{P} = \sum \mathbf{p}_i$, $\mathbf{L} = \sum \mathbf{q}^i \times \mathbf{p}_i$, and $D = \sum \mathbf{q}^i \cdot \mathbf{p}_i$ are conserved. Let $\mathbf{P} = \mathbf{L} = D = 0$. Remember that, in this particular case, the reduced phase space is equal to the cotangent bundle of the reduced configuration space. In what follows, let me present the reduction of phase space Γ with respect to the similarity group Sim(3) and $\mathbf{P} = \mathbf{L} = D = 0$.

For three particles, the reduction of configuration space Q with respect to translations and rotations can be found in Montgomery [2002]. Barbour, Koslowski, and Mercati [2013] and [2015] use the results of Montgomery and extend it to phase space Γ given that $\mathbf{P} = \mathbf{L} = D = 0$ – at least as far as they need in order to compute the measure from the Faddeev-Popov formula (cf. (5.63) and (6.31)). I will present the full symplectic reduction focusing on the symplectic two-form from which I will directly construct the volume measure.

Remark (Reduction for the Newtonian universe)**.** Later we will be interested in the Newtonian universe – particles moving through three-dimensional Euclidean space and attracting each other according to Newton's law of gravitation. The Hamiltonian of that system is invariant only with respect to translations and rotations, but not with respect to dilations. This can be checked by computing the Poisson brackets of H and $\mathbf{P}$, $\mathbf{L}$, and D, respectively. It is $\{H, \mathbf{P}\} = \{H, \mathbf{L}\} = 0$, but $\{H, D\} \neq 0$. This means that $\mathbf{P}$ and $\mathbf{L}$ are conserved quantities of motion – they can be set equal to zero –, but D is not. Hence, the reduced dynamics will have to be formulated on translationally and rotationally reduced phase space $\Gamma_l \cong T^*S_R$ and not on T^*S (which one obtains when reducing with respect to dilations as well).

However, we will find that, also for the Newtonian gravitational system, there exists an internal time parameter and an internal Hamiltonian description the way it was described in Section 4.3 and that internal description will be formulated on T^*S. Hence, we later need the geometry of both spaces. For that reason we will construct both T^*S_R and T^*S not assuming any particular Hamiltonian, and we will say more about the actual dynamics of the Newtonian gravitational system later (see Section 6.3).

5.2.1 Translational invariance

Before we start with reduction, note that matters can be facilitated essentially taking into account the fact that, for the special case of three particles and zero angular momentum $\mathbf{L} = 0$, the motion of the particles is restricted to a plane.[33] This means that the $\mathbf{q}^i, \mathbf{p}_i$ are really two-component vectors $\mathbf{q}^i \in \mathbb{R}^2, \mathbf{p}_i \in \mathbb{R}^2$ and phase space Γ is really $4N$-dimensional: $\Gamma \cong \mathbb{R}^{4N}$, i.e., $\Gamma \cong \mathbb{R}^{12}$. In that case, the symmetry group of translations is $\mathbb{R}^2$, the group of rotations is $SO(2)$ and the group of dilations is (like in three dimensions) $\mathbb{R}^+$.

Let us start with translational symmetry. We know that we can construct reduced phase space Γ_p from Γ via the constraints (5.14) and (5.14): $\sum \mathbf{p}_i = 0$ and $\sum m_i \mathbf{q}^i = 0$. This can be done by help of a convenient choice of coordinates. One such choice are the Jacobi coordinates $\boldsymbol{\rho}^i \in \mathbb{R}^2$ and their canonical conjugates $\boldsymbol{\kappa}_i \in \mathbb{R}^2$. These coordinates are convenient because they split into translationally-invariant coordinates $\boldsymbol{\rho}^1$, $\boldsymbol{\rho}^2$ and center-of-mass coordinates $\boldsymbol{\rho}^3 \sim \sum m_i \mathbf{q}^i$ which are fixed by the center-of-mass constraint (5.15). Similarly, the conjugate momenta split

[33] For this assertion, see any standard textbook on mechanics.

into relative momenta $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$ and total linear momentum $\boldsymbol{\kappa}_3 \sim \sum \mathbf{p}_i$ which is fixed by the momentum constraint (5.14).

Explicitly, the Jacobi coordinates are:

$$\begin{aligned}
\boldsymbol{\rho}^1 &= \sqrt{\frac{m_1 m_2}{m_1+m_2}}(\mathbf{q}^1-\mathbf{q}^2),\\
\boldsymbol{\rho}^2 &= \sqrt{\frac{m_3(m_1+m_2)}{m_1+m_2+m_3}}\left(\mathbf{q}^3-\frac{m_1\mathbf{q}^1+m_2\mathbf{q}^2}{m_1+m_2}\right),\\
\boldsymbol{\rho}^3 &= \frac{m_1\mathbf{q}^1+m_2\mathbf{q}^2+m_3\mathbf{q}^3}{\sqrt{m_1+m_2+m_3}}.
\end{aligned} \tag{5.29}$$

Connected to the Jacobi coordinates, the conjugate momenta are:

$$\begin{aligned}
\boldsymbol{\kappa}_1 &= \frac{m_1\mathbf{p}_2-m_2\mathbf{p}_1}{\sqrt{m_1m_2(m_1+m_2)}},\\
\boldsymbol{\kappa}_2 &= \sqrt{\frac{(m_1+m_2)}{(m_1+m_2+m_3)m_3}}\left(\mathbf{p}_3-\frac{m_3\mathbf{p}_1+m_3\mathbf{p}_2}{m_1+m_2}\right),\\
\boldsymbol{\kappa}_3 &= \frac{\mathbf{p}_1+\mathbf{p}_2+\mathbf{p}_3}{\sqrt{m_1+m_2+m_3}}.
\end{aligned} \tag{5.30}$$

Together the $\boldsymbol{\rho}^i$ and $\boldsymbol{\kappa}_i$ form a canonical set of local coordinates on Γ. To be precise, the change of coordinates from the canonical coordinates $\mathbf{q}^i, \mathbf{p}_i$ to the Jacobi coordinates and their canonical conjugates, $\boldsymbol{\rho}^i, \boldsymbol{\kappa}_i$ is canonical. This means that it preserves both the volume,

$$\mathrm{d}^3\mathbf{q}\mathrm{d}^3\mathbf{p} = \mathrm{d}^3\boldsymbol{\rho}\mathrm{d}^3\boldsymbol{\kappa},$$

and the symplectic structure. The symplectic structure is conserved if and only if the new coordinates again fulfill the canonical Poisson bracket relations, respectively if and only if the symplectic two-form is again of the standard form:[34]

$$\omega = \mathrm{d}\boldsymbol{\rho}^1 \wedge \mathrm{d}\boldsymbol{\kappa}_1 + \mathrm{d}\boldsymbol{\rho}^2 \wedge \mathrm{d}\boldsymbol{\kappa}_2 + \mathrm{d}\boldsymbol{\rho}^3 \wedge \mathrm{d}\boldsymbol{\kappa}_3.$$

Lemma 5.1. *Let $\Gamma = T^*Q$ with $Q = \mathbb{R}^6$. Let the Hamiltonian H on Γ be invariant under translations. Let $M_p = \{(\boldsymbol{q}, \boldsymbol{p}) \in \Gamma | \sum_{i=1}^3 \boldsymbol{p}_i = 0\}$ and $G_p = \mathbb{R}^2$, i.e.*

$$\Gamma_p = \frac{M_p}{G_p} \tag{5.31}$$

is the (translationally) reduced phase space. Let $\boldsymbol{\rho}^i \in \mathbb{R}^2$, $\boldsymbol{\kappa}_i \in \mathbb{R}^2$, $i = 1, 2$ be the Jacobi coordinates (5.29) and their canonical conjugates (5.30). Then the symplectic form ω_p on Γ_p can be written as

$$\omega_p = \mathrm{d}\boldsymbol{\rho}^1 \wedge \mathrm{d}\boldsymbol{\kappa}_1 + \mathrm{d}\boldsymbol{\rho}^2 \wedge \mathrm{d}\boldsymbol{\kappa}_2. \tag{5.32}$$

[34] It is proven in Appendix A.2 that the coordinates fulfill the canonical Poisson bracket relations if and only if there exists a symplectic two-form of the standard form.

Proof. We know from Section 5.1.2 that in order to construct translationally reduced phase space Γ_p, the constraints (5.14) and (5.15) need to be fulfilled. These are, for the case at hand,

$$\sum_{i=1}^{3} \mathbf{p}_i = 0$$

and

$$\sum_{i=1}^{3} m_i \mathbf{q}^i = 0 \quad .$$

To be able to impose these constraints we change from the canonical coordinates $\mathbf{q}^i, \mathbf{p}_i$ with $i = 1, 2, 3$ to the Jacobi coordinates $\boldsymbol{\rho} = (\boldsymbol{\rho}^1, \boldsymbol{\rho}^2, \boldsymbol{\rho}^3)$ and their canonical conjugates $\boldsymbol{\kappa} = (\boldsymbol{\kappa}_1, \boldsymbol{\kappa}_2, \boldsymbol{\kappa}_3)$. Since this change of coordinates is canonical, there again exists a symplectic two-form on Γ of the standard form:

$$\omega = \mathrm{d}\boldsymbol{\rho}^1 \wedge \mathrm{d}\boldsymbol{\kappa}_1 + \mathrm{d}\boldsymbol{\rho}^2 \wedge \mathrm{d}\boldsymbol{\kappa}_2 + \mathrm{d}\boldsymbol{\rho}^3 \wedge \mathrm{d}\boldsymbol{\kappa}_3.$$

From the definition of the Jacobi coordinates (5.29) it follows that the center-of-mass constraint (5.15) is fulfilled if and only if

$$\boldsymbol{\rho}^3 = 0. \tag{5.33}$$

This equation fixes the center of mass to the origin. The remaining translationally-invariant coordinates $\boldsymbol{\rho}^1$ and $\boldsymbol{\rho}^2$ are local coordinates of Q_{cm}.

Similarly, given the definition of the $\boldsymbol{\kappa}_i$ (5.30), it follows that the linear momentum constraint (5.14) is fulfilled if and only if

$$\boldsymbol{\kappa}_3 = 0. \tag{5.34}$$

Together with $\boldsymbol{\rho}^1$ and $\boldsymbol{\rho}^2$, $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$ form local coordinates of 8-dimensional reduced phase space $\Gamma_p = T^*Q_{cm}$ and the symplectic form on Γ_p can be written as

$$\omega_p = \mathrm{d}\boldsymbol{\rho}^1 \wedge \mathrm{d}\boldsymbol{\kappa}_1 + \mathrm{d}\boldsymbol{\rho}^2 \wedge \mathrm{d}\boldsymbol{\kappa}_2.$$

□

Remark ($\Gamma_p \cong T^*Q_{cm}$)**.** The way we found the internal variables of Γ_p sheds light on the diffeomorphism between the reduced phase space Γ_p and the cotangent bundle of the reduced configuration space T^*Q_{cm}. One way to construct the reduced phase space Γ_p is to fix the level set $\mathbf{P} = 0$ (5.14) and to fix the gauge (4.15) which leaves us with the variables $\boldsymbol{\rho}^1, \boldsymbol{\rho}^2, \boldsymbol{\kappa}_1, \boldsymbol{\kappa}_2$ as local coordinates of Γ_0. Another way is to first construct the quotient of configuration space Q_{cm} by introducing the Jacobi coordinates and setting $\boldsymbol{\rho}^3 = 0$ (fixing the gauge). From $\boldsymbol{\rho}^1, \boldsymbol{\rho}^2$ we then construct the canonical conjugates $\boldsymbol{\kappa}_1, \boldsymbol{\kappa}_2$ which together with $\boldsymbol{\rho}^1, \boldsymbol{\rho}^2$ form a local basis of the cotangent bundle $T^*(Q_{cm}) \cong \Gamma_p$.

5.2.2 Rotational invariance

Let us now deal with rotational symmetry. Convenient coordinates are such that they split into rotationally-invariant coordinates and coordinates which fix the orientation of the system. A particular good choice for the description of the system on (translationally and rotationally) reduced phase space $\Gamma_l = \Gamma_p/SO(2) \cong T^*S_R$ are the rotationally-invariant Hopf coordinates $\mathbf{w} = (w^1, w^2, w^3)$ and their canonical conjugates $\mathbf{z} = (z_1, z_2, z_3)$. The $\mathbf{w}$ coordinates we introduce have originally been proposed by Hopf in his discovery of the Hopf fibration.[35] They are:

$$w^1 = \frac{||\boldsymbol{\rho}^1||^2 - ||\boldsymbol{\rho}^2||^2}{2}, \qquad w^2 = \boldsymbol{\rho}^1 \cdot \boldsymbol{\rho}^2, \qquad w^3 = \boldsymbol{\rho}^1 \times \boldsymbol{\rho}^2. \tag{5.35}$$

The canonical momenta are:

$$\begin{aligned}
z_1 &= \frac{\boldsymbol{\rho}^1 \cdot \boldsymbol{\kappa}_1 - \boldsymbol{\rho}^2 \cdot \boldsymbol{\kappa}_2}{|\boldsymbol{\rho}^1|^2 + |\boldsymbol{\rho}^2|^2}, \\
z_2 &= \frac{\boldsymbol{\rho}^1 \cdot \boldsymbol{\kappa}_2 + \boldsymbol{\rho}^2 \cdot \boldsymbol{\kappa}_1}{||\boldsymbol{\rho}^1||^2 + ||\boldsymbol{\rho}^2||^2} - \frac{1}{2}\frac{\boldsymbol{\rho}^1 \times \boldsymbol{\rho}^2(||\boldsymbol{\rho}^1||^2 - ||\boldsymbol{\rho}^2||^2)}{||\boldsymbol{\rho}^1|| \cdot ||\boldsymbol{\rho}^2||}\frac{\boldsymbol{\kappa}_1 \times \boldsymbol{\rho}^1 + \boldsymbol{\kappa}_2 \times \boldsymbol{\rho}^2}{||\boldsymbol{\rho}^1||^2 + ||\boldsymbol{\rho}^2||^2}, \\
z_3 &= \frac{||\boldsymbol{\rho}^2||^2 \boldsymbol{\rho}^1 \times \boldsymbol{\kappa}_2 - ||\boldsymbol{\rho}^1||^2 \boldsymbol{\rho}^2 \times \boldsymbol{\kappa}_1}{||\boldsymbol{\rho}^1||^2 + ||\boldsymbol{\rho}^2||^2} - \frac{1}{2}\frac{\boldsymbol{\rho}^1 \times \boldsymbol{\rho}^2(||\boldsymbol{\rho}^1||^2 - ||\boldsymbol{\rho}^2||^2)}{||\boldsymbol{\rho}^1|| \cdot ||\boldsymbol{\rho}^2||}\frac{\boldsymbol{\kappa}_1 \cdot \boldsymbol{\rho}^1 - \boldsymbol{\kappa}_2 \cdot \boldsymbol{\rho}^2}{||\boldsymbol{\rho}^1||^2 + ||\boldsymbol{\rho}^2||^2}.
\end{aligned} \tag{5.36}$$

The $\mathbf{w}$ and $\mathbf{z}$ coordinates together with the off-diagonal component of the center-of-mass inertia tensor I_L and the (normalized) angular momentum $L' = L/(|\boldsymbol{\rho}^1|^2 + |\boldsymbol{\rho}^2|^2)$ form a set of local coordinates of Γ_p. Let $\mathbf{I}_{cm}$ denote the center-of-mass inertia tensor. In terms of the $\boldsymbol{\rho}$ and $\boldsymbol{\kappa}$ coordinates, it can be written as

$$\mathbf{I}_{cm} = \sum_{i=1}^{2} m_i(\boldsymbol{\rho}^i \cdot \boldsymbol{\rho}^i\, \mathbb{I} - \boldsymbol{\rho}^i \otimes \boldsymbol{\rho}^i). \tag{5.37}$$

Note that since we are on a plane, $\mathbf{I}_{cm}$ is 2×2 matrix, so there is only one off-diagonal component. In terms of $\boldsymbol{\rho}$ and $\boldsymbol{\kappa}$, the angular momentum is $L = \sum_{i=1}^{2} \boldsymbol{\rho}^i \times \boldsymbol{\kappa}_i$. Hence, the constraints (5.18) and (5.19) which determine rotationally reduced phase space Γ_l can be written as

$$L = \sum_{i=1}^{2} \boldsymbol{\rho}^i \times \boldsymbol{\kappa}_i \quad \text{and} \quad I_L = -m_1\rho^{12}\rho^{11} - m_2\rho^{22}\rho^{21}. \tag{5.38}$$

While L' and I_L will be fixed by these constraints, the rotationally-invariant $\mathbf{w}$ and $\mathbf{z}$ coordinates will form a canonical set of local coordinates of Γ_l. To be precise, the following can be shown.

Lemma 5.2. *Let $\Gamma = T^*Q$ with $Q = \mathbb{R}^6$. Let $\boldsymbol{\rho}^i$, $\boldsymbol{\kappa}_i$ with $i = 1, 2$ given by (5.29), (5.30) and Γ_p given by (5.31). Let the Hamiltonian H on Γ be invariant under translations and rotations. Let*

[35] The Hopf fibration is a non-trivial fibre bundle found by Hopf in 1931. Hopf discovered that S^3 is locally a product space $S^2 \times S^1$ (cf. Montgomery [2002]).

$M_l = \{(\boldsymbol{\rho}, \boldsymbol{\kappa}) \in \Gamma_p | \sum_{i=1}^2 \boldsymbol{\rho}^i \times \boldsymbol{\kappa}_i = 0\}$ and $G_l = SO(2)$, i.e.

$$\Gamma_l = \frac{M_l}{G_l} \tag{5.39}$$

is the (translationally and rotationally) reduced phase space. Let $\boldsymbol{w} = (w^1, w^2, w^3)$ the Hopf coordinates (5.35) and $\boldsymbol{z} = (z_1, z_2, z_3)$ their canonical conjugates (5.36). Then the symplectic form ω_l on Γ_l can be written as

$$\omega_l = \mathrm{d}w^1 \wedge \mathrm{d}z_1 + \mathrm{d}w^2 \wedge \mathrm{d}z_2 + \mathrm{d}w^3 \wedge \mathrm{d}z_3. \tag{5.40}$$

Proof. We know from Section 5.1.2 that we can obtain the local structure of Γ_l from Γ_p by imposing the angular momentum constraint $L = \sum \mathbf{q}^i \times \mathbf{p}_i = 0$ (5.18) and the principal-axis constraint $I_L = 0$ (5.19) where I_L denotes the off-diagonal component of the center-of-mass inertia tensor $\mathbf{I}_{cm}$. While (5.18) fixes the level set where angular momentum is zero, (5.19) fixes the orientation of the system (it fixes the direction of the axis of our local coordinate system to the direction of the principal axis of inertia).

To find the local structure of translationally and rotationally reduced phase space Γ_l, it suffices to evaluate the constraints (5.18) and (5.19) on translationally reduced phase space Γ_p. In terms of the internal coordinates $\boldsymbol{\rho}^i$, $\boldsymbol{\kappa}_i$ of Γ_p, these constraints are

$$L = \sum_{i=1}^{2} \boldsymbol{\rho}^i \times \boldsymbol{\kappa}_i = 0 \quad \text{and} \quad I_L = -m_1 \rho^{12} \rho^{11} - m_2 \rho^{22} \rho^{21} = 0.$$

This form we have already derived above (cf. (5.38)). There we changed from the Jacobi coordinates and their canonical conjugates $\boldsymbol{\rho}^i, \boldsymbol{\kappa}_i$ with $i = 1, 2$ to the Hopf coordinates $\mathbf{w} = (w^1, w^2, w^3)$ and their canonical conjugates $\mathbf{z} = (z_1, z_2, z_3)$ which together with $L' = L/(|\boldsymbol{\rho}^1|^2 + |\boldsymbol{\rho}^2|^2)$ and I_L form a set of local coordinates of Γ_p. From $L = 0$ it follows that $L' = 0$. Hence, we obtain that the rotationally invariant coordinates $\mathbf{w} = (w^1, w^2, w^3)$ and their canonical conjugates $\mathbf{z} = (z_1, z_2, z_3)$ form a canonical set of coordinates of 6-dimensional reduced phase space $\Gamma_l = T^*S_R$ and the symplectic form on Γ_l can then be written as

$$\omega_l = \mathrm{d}w^1 \wedge \mathrm{d}z_1 + \mathrm{d}w^2 \wedge \mathrm{d}z_2 + \mathrm{d}w^3 \wedge \mathrm{d}z_3.$$

□

Remark (Special Euclidean group)**.** Let me be more precise about the reduction performed here. It is actually the reduction with respect to $SE(3)$, the special Euclidean group. $SE(3)$ is the group of translations and rotations, but not of reflections. In the literature, $SE(3)$ is also called the symmetry group of rigid motions (cf. Montgomery [2002]). This makes sense because by rotating and translating a rigid body within Euclidean space $\mathbf{E}^3$ it cannot be reflected. The fact that we do not quotient by reflections implies that shape space S is the whole two-sphere S^2, not only the upper half of it. Defined this way, shape space S is actually the space of all oriented shapes.

5.2.3 Scale invariance

Let us now deal with dilations. Convenient coordinates are coordinates which separate scale from shape degrees of freedom (which are either angles or relative distances). For this purpose, let us introduce spherical coordinates R, ψ, ϕ via the relation

$$w^1 = R \sin\psi \cos\phi, \qquad w^2 = R \sin\psi \sin\phi, \qquad w^3 = R \cos\psi. \tag{5.41}$$

The canonical conjugates p_R, p_ψ, p_ϕ satisfy:

$$\begin{aligned} z_1 &= \frac{1}{R}(\cos\phi(Rp_R \sin\psi - p_\psi \cos\psi) - p_\phi \sin^{-1}\psi \sin\phi) \\ z_2 &= \frac{1}{R}(\sin\phi(Rp_R \sin\psi - p_\psi \cos\psi) + p_\phi \sin^{-1}\psi \cos\phi) \\ z_3 &= -p_R \cos\psi - \frac{1}{R} p_\psi \sin\psi. \end{aligned} \tag{5.42}$$

You find these equations by demanding that the change of coordinates from the $\mathbf{w}$ and $\mathbf{z}$ coordinates to the spherical coordinates R, ψ, ϕ and their canonical conjugates is canonical. You can check that the change of coordinates is volume-preserving:

$$\det \frac{\mathrm{d}(w^1, w^2, w^3)}{\mathrm{d}(R, \psi, \phi)} \det \frac{\mathrm{d}(z^1, z^2, z^3)}{\mathrm{d}(p_R, p_\psi, p_\phi)} = (R^2 \sin\psi)(R^{-2} \sin^{-1}\psi) = 1. \tag{5.43}$$

It follows that

$$\mathrm{d}^3 w \mathrm{d}^3 z = \mathrm{d}R \mathrm{d}\psi \mathrm{d}\phi \mathrm{d}p_R \mathrm{d}p_\psi \mathrm{d}p_\phi. \tag{5.44}$$

In addition, it preserves the symplectic structure, i.e., also in terms of the new coordinates the symplectic two-form ω_l on Γ_l is of the standard form:

$$\omega_l = \mathrm{d}R \wedge \mathrm{d}p_R + \mathrm{d}\psi \wedge \mathrm{d}p_\psi + \mathrm{d}\phi \wedge \mathrm{d}p_\phi. \tag{5.45}$$

Hence, R, θ, ϕ and p_R, p_ψ, p_ϕ form a canonical set of local coordinates of Γ_l. When constructing the dilationally reduced phase space Γ_d, R and p_R will be fixed by the constraints while the angular degrees of freedom ψ, ϕ and p_ψ, p_ϕ will become canonical coordinates of translationally, rotationally and dilationally reduced phase space $\Gamma_d \cong T^*S$.

Lemma 5.3. *Let $\Gamma = T^*Q^+$ with $Q^+ = Q \backslash \{\boldsymbol{q}^1 = \boldsymbol{q}^2 = \boldsymbol{q}^3\}$ and $Q = \mathbb{R}^6$. Let $\Gamma_l \cong T^*S_R$ be as above but where now $S_R = Q^+/(\mathbb{R}^2 \times SO(2))$. Let the Hamiltonian H on Γ be invariant under translations, rotations, and dilations. Let $\boldsymbol{w}, \boldsymbol{z}$ be defined by (5.35), (5.36), $M_d = \{(\boldsymbol{w}, \boldsymbol{z}) \in \Gamma_l | \sum_{i=1}^3 w^i \cdot z_i = 0\}$ and $G_d = \mathbb{R}^+$, i.e.*

$$\Gamma_d = \frac{M_d}{G_d} \tag{5.46}$$

is the (translationally, rotationally and dilationally) reduced phase space. Let R, θ, ϕ be the spherical coordinates (5.41) and p_R, p_θ, p_ϕ their canonical conjugates (5.42). Then the symplectic form

ω_d on Γ_d can be written as

$$\omega_d = \mathrm{d}\psi \wedge \mathrm{d}p_\psi + \mathrm{d}\phi \wedge \mathrm{d}p_\phi. \tag{5.47}$$

Proof. We need to impose the constraints (5.22) and (5.23):

$$D = \sum_{i=1}^{3} \mathbf{q}^i \cdot \mathbf{p}_i = 0$$

and

$$I = \sum_{i=1}^{3} m_i |\mathbf{q}^i|^2 = 1.$$

We do this by help of a canonical change of coordinates from the $\mathbf{w}$ and $\mathbf{z}$ coordinates to the spherical coordinates and their canonical conjugates defined in (5.41) and (5.42). Let us now formulate the constraints (5.22) and (5.23) in terms of the new coordinates. The dilational momentum $D = \sum \mathbf{q}^i \cdot \mathbf{p}_i$ can be written as

$$D = 2\ \mathbf{w} \cdot \mathbf{z} + \boldsymbol{\kappa}_3 \cdot \boldsymbol{\rho}^3 = 2\ R \cdot p_R + \boldsymbol{\kappa}_3 \cdot \boldsymbol{\rho}^3. \tag{5.48}$$

Due to the translational invariance, it is $\boldsymbol{\kappa}_3 = \boldsymbol{\rho}^3 = 0$ on Γ_l. It follows that, on Γ_l,

$$D = 2\ R \cdot p_R. \tag{5.49}$$

In addition, the moment of inertia can be written as

$$I = 2(||\boldsymbol{\rho}_1||^2 + ||\boldsymbol{\rho}_2||^2) = 2||\mathbf{w}|| = 2R. \tag{5.50}$$

It follows that on Γ_l the constraints (5.22) and (5.23) can be written as

$$D = R \cdot p_R = 0 \quad \text{and} \quad I = 2R = 1.$$

These constraints are fulfilled if and only if $R = 1/2$ and $p_R = 0$.

Note, at this point, that the reduction with respect to dilations can only be done on $Q^+ = Q \backslash \{\mathbf{q}^1 = \mathbf{q}^2 = \mathbf{q}^3\}$, respectively on $T^*Q^+ = T^*(Q \backslash \{\mathbf{q}^1 = \mathbf{q}^2 = \mathbf{q}^3\})$. This follows from the fact that only on Q^+ the constraint $R = 1/2$ can be imposed. This is why we changed the definition of Q in the lemma. Having the dynamical system in mind, this means that three-particle collisions are excluded. Now $Q^+/\mathbb{R}^2 = Q_{cm} \backslash \{0\}$ and $Q^+/(\mathbb{R}^2 \times SO(2)) = S_R \backslash \{0\}$.

Again, $\Gamma_d \cong T^*S$ is determined by the constraints $R = 1/2$ and $p_R = 0$. It follows that ψ and ϕ are local coordinates of two-dimensional shape space S and ψ and ϕ together with p_ψ and p_ϕ are local coordinates of four-dimensional shape phase space $\Gamma_d \cong T^*S$. From (5.45) it follows that the coordinates ψ, ϕ and p_ψ, p_ϕ are canonical and the symplectic form ω_d on $\Gamma_d \cong T^*S$ is of the standard form

$$\omega_d = \mathrm{d}\psi \wedge \mathrm{d}p_\psi + \mathrm{d}\phi \wedge \mathrm{d}p_\phi.$$

□

Remark (Three-particle shape space)**.** Let us, for simplicity, consider the equal-mass case ($m_i = m$ with $i = 1, 2, 3$). Topologically, the three-particle shape space S is a two-sphere S^2 with local coordinates ψ and ϕ. Every point on that sphere represents a different triangle shape (specified by the two angles ψ and ϕ). Let ψ be the polar angle and ϕ the azimuthal angle of the shape sphere. Then we find the collinear configurations (where all particles are in a line) on the equator of the sphere, $\psi = \pi/2$, and the two equilateral triangles (the equilateral triangle and its reflected version) at the top and bottom of the sphere, $\psi = 0$ and $\psi = \pi$. The upper half-sphere is a reflection of the lower half-sphere (if we had factored out reflections, we would have ended up with one half-sphere only). There are three points of binary collision (for the three possibilities to have two particles at one point and one particle spatially separated) and three Euler configurations where all particles are on a line with one particle centered between the other two (there are three possibilities to have one particle in the middle). Both the binary collision points and the Euler configurations lie on the equator of the shape sphere at equal distance from each other (with one Euler configuration between two binary collision points).

5.2.4 Volume measures on T^*S_R and T^*S

From the symplectic two-forms ω_l, respectively ω_d, we can obtain a natural volume measure on $\Gamma_l \cong T^*S_R$, respectively $\Gamma_d \cong T^*S$.

Lemma 5.4. *Let ω_l given by (5.40) the symplectic form on T^*S_R and ω_d given by (5.47) the symplectic form on T^*S. Then*

$$\mathrm{d}\mu_R = \mathrm{d}R\mathrm{d}p_R\mathrm{d}\psi\mathrm{d}p_\psi\mathrm{d}\phi\mathrm{d}p_\phi \tag{5.51}$$

is the natural volume measure on T^*S_R and

$$\mathrm{d}\mu = \mathrm{d}\psi\mathrm{d}p_\psi\mathrm{d}\phi\mathrm{d}p_\phi \tag{5.52}$$

is the natural volume measure on T^*S.

Proof. Recall the natural volume form Ω on $2n$-dimensional $\Gamma = T^*Q$ as arising from the natural, symplectic two-form ω (cf. Eq. (4.4)): $\Omega = \frac{(-1)^{[n/2]}}{n!}\omega^n$. From Ω we get the natural volume measure μ (cf. Eq. (4.6)): $\mathrm{d}\mu = |\Omega|$.

Now consider ω_l on T^*S_R given by (5.40): $\omega_l = \mathrm{d}R \wedge \mathrm{d}p_R + \mathrm{d}\psi \wedge \mathrm{d}p_\psi + \mathrm{d}\phi \wedge \mathrm{d}p_\phi$. It follows that, in the given spherical coordinates and with $n = 3$, the volume measure on T^*S_R is

$$\mathrm{d}\mu_R = \frac{|\omega_l^3|}{3!} = \mathrm{d}R\mathrm{d}p_R\mathrm{d}\psi\mathrm{d}p_\psi\mathrm{d}\phi\mathrm{d}p_\phi.$$

Here we used that $\mathrm{d}a \wedge \mathrm{d}a = 0$ for any a and that $\mathrm{d}a \wedge \mathrm{d}b = -\mathrm{d}b \wedge \mathrm{d}a$ for any a, b.

Analogously, we get the volume measure on T^*S. Consider ω_d on T^*S given by (5.47): $\omega_d = \mathrm{d}\psi \wedge \mathrm{d}p_\psi + \mathrm{d}\phi \wedge \mathrm{d}p_\phi$. It follows that, in the given coordinates and with $n = 2$, the volume measure on T^*S is

$$\mathrm{d}\mu = \frac{|\omega_l^2|}{2!} = \mathrm{d}\psi\mathrm{d}p_\psi\mathrm{d}\phi\mathrm{d}p_\phi.$$

□

5.3 Faddeev-Popov construction of the measure

In their [2015] paper, Barbour et al. construct a volume measure on the reduced internal phase space of the Newtonian universe via a formula of Faddeev and Popov, which they introduce without further justification. To understand this formula, I will derive it and compare it to the volume measure obtained from the reduced symplectic two-form.

To provide some intuition, let me introduce the method of Faddeev and Popov on the original example of the Yang-Mills field, based on their [1967] paper. Then I apply their method to a general Hamiltonian system with constraints. This has originally been done by Faddeev [1969]. I will give a proof similar to his.

Faddeev-Popov measure for a gauge field. Faddeev and Popov [1967] studied the Feynman path integral for the Yang-Mills field when they realized that the measure used in the definition of the integral had to be changed in order to appropriately treat the gauge degrees of freedom. To be precise, they realized that the gauge degrees of freedom should be factored out before starting the perturbative analysis of the integral. This led them to the introduction of the so-called Faddeev-Popov determinant.[36]

Since I only want to convey some intuition, let me use a somewhat informal notation. Let G be a group of gauge transformations acting on Γ. Let $x \in \Gamma$ and $A(x)$ be a gauge field on Γ, that is, A transforms under gauge transformations as

$$A_\mu \to A_\mu^\Omega \tag{5.53}$$

where Ω takes values in the gauge group G. The Lagrangian $\mathcal{L}$ is invariant under this transformation (this is why we call it a gauge transformation).

Originally, the Feynman integral that is used to compute the S-matrix element between incoming and outgoing states has been formulated as

$$< in|out > \sim \int e^{iS[A]} \prod_x \mathrm{d}A(x). \tag{5.54}$$

Here $e^{iS[A]}$ and $\int \prod_x \mathrm{d}A(x)$ are invariant under the gauge transformations – the first because the Lagrangian is invariant, the latter since we integrate over all fields (including all fields which can be obtained from another by a gauge transformation). Now the idea is that the integrand in (5.54) should be constant on the gauge orbits. In other words, the integral should be proportional

[36] The Faddeev-Popov determinant, when rewritten such that it constitutes part of the action S, gives rise to a Gaussian representing fermionic particles – the so-called Faddeev-Popov ghost. This is well-known in quantum field theory where the method of Faddeev and Popov is nowadays standard when calculating the Feynman path integral for arbitrary gauge fields.

to the volume of the orbits, which can then be factored out. The volume of the orbits is

$$\int \prod_x \mathrm{d}\Omega(x). \tag{5.55}$$

One way to separate this term within the integral is by integrating along the orbits and over some transversal surface. A surface transversal to the orbits is determined by choosing a particular gauge, which can be done by demanding that $\partial_\mu A_\mu$ is equal to zero. Now the "trick" is to rewrite the Feynman integral inserting an extra term which is equal to one (we insert a one because we don't want to change the integral as we factor out the volume of the orbits). At this point, the Faddeev-Popov determinant Δ_{FP} is introduced. Let Δ_{FP} be such that

$$\int \Delta_{FP} \prod_x \delta(\partial_\mu A_\mu^\Omega(x))\mathrm{d}\Omega(x) = 1 \tag{5.56}$$

with δ being the Dirac delta function. Then the Feynman integral (the right hand side of (5.54)) can be rewritten as

$$\begin{aligned} &\int e^{iS[A]} \int \Delta_{FP} \prod_x \delta(\partial_\mu A_\mu^\Omega(x))\mathrm{d}\Omega(x) \prod_x \mathrm{d}A(x) \\ = &\int e^{iS[A]} \Delta_{FP} \prod_x \delta(\partial_\mu A_\mu(x))\mathrm{d}A(x) \int \prod_x \mathrm{d}\Omega(x). \end{aligned} \tag{5.57}$$

In the last step, we used that we can replace A_μ^Ω by A_μ because we integrate over all $A(x)$. This means that all dependence on the particular choice of gauge has vanished and we can factor out the volume of the gauge orbits.

Remark (Geometrical nature of Δ_{FP})**.** The geometrical interpretation of the Faddeev-Popov determinant Δ_{FP} is essentially analogous to the meaning of $||\nabla H||$ in the definition of the microcanonical measure (cf. (2.6)). Both Δ_{FP} and $||\nabla H||$ take care of the angle between the orbit and the hypersurface transversally intersecting the orbit (as determined by the delta-function). That way both the Faddeev-Popov determinant and the gradient of H take care of the difference between the natural volume measure, respectively the natural measure of surface area, of the underlying constraint space and the volume measure, respectively measure of surface area, as obtained by the delta-function.

Faddeev-Popov measure for Hamiltonian systems with constraints. Let us now derive the Faddeev-Popov measure for a Hamiltonian system with constraints. In Sections 5.1 and 5.2 we found that a system with symmetries can be treated as a Hamiltonian system with constraints. Let us therefore consider a Hamiltonian system with m first class constraints $\phi_j = \phi_j(q,p)$ in the sense of Dirac [1964].

Definition 5.9 (First class constraints)**.** Let $(\Gamma, \mathcal{B}(\Gamma), \mu, H)$ be a Hamiltonian system with m

linearly independent constraints $\phi_j = 0$ $(j = 1, ..., m)$. If

$$\{\phi_i, \phi_j\} = 0 \tag{5.58}$$

for all $j \neq i$ $(i, j = 1, ..., m)$, then the ϕ_j are *first class constraints.*

That is, the constraints are first class if their mutual Poisson brackets vanish. In accordance with Dirac, we say that they are "weakly zero" and write

$$\phi_i \approx 0 \tag{5.59}$$

in order to indicate that, if we want to determine the dynamics by help of the Poisson bracket[37], all Poisson brackets need to be worked out before the ϕ_j's are actually taken to be zero.

Since they are the conserved quantities of motion, together the ϕ_j's determine the constraint surface $M \subset \Gamma$ in which the trajectories lie. If there are m first class constraints, the constraint surface is a hypersurface of $2n - m$ dimensions.

Analyzing the equations of motion, one finds that all the ϕ_j's can be interpreted as generators of gauge transformations. Let $\varepsilon_j \in \mathbb{R}^+$, $j = 1, ..., m$. For each ϕ_j and all phase space variables q_i, p_i[38] with $i = 1, ..., n$, the respective transformation is given by

$$\begin{aligned} q_i \to q_i + \delta_{\phi_j} q_i &= q_i + \varepsilon_j \{q_i, \phi_j\}, \\ p_i \to p_i + \delta_{\phi_j} p_i &= p_i + \varepsilon_j \{p_i, \phi_j\}. \end{aligned} \tag{5.60}$$

Here $\delta_{\phi_j} q_i = \varepsilon_j \{q_i, \phi_j\}$ describes a deviation of q_i of the amount ε_j along the gauge orbit connected to the generator ϕ_j (and the same for p_i). As Dirac [1964] shows, the transformations (5.60) do not affect the dynamics.[39] This is why they are called gauge transformations.

In order to "fix the gauge", we can choose one representative of the equivalence class of states connected by the given gauge transformation. In case there are m first class constraints $\phi_j \approx 0$, the gauge fixing leads to m additional constraints $\chi_l \approx 0$ which together with the first class constraints form m pairs of second class constraints. Also these additional constraints are weakly zero meaning that, in order to determine the dynamics by help of the Poisson bracket, all Poisson brackets need to be worked out before the χ_l are actually taken to be zero.

Definition 5.10 (Second class constraint)**.** Let $(\Gamma, \mathcal{B}(\Gamma), \mu, H)$ be a Hamiltonian system. Let $\phi_j \approx 0$ and $\chi_l \approx 0$ with

$$\{\chi_l, \phi_j\} \neq 0. \tag{5.61}$$

Then ϕ_j, χ_l form a pair of *second class constraints.*

The condition $\{\chi_l, \phi_j\} \neq 0$ ensures that $\chi_l \approx 0$ is really a gauge fixing of the gauge transformation connected to ϕ_j – for if it were zero, χ_l would not change along the gauge orbit which

[37] Cf. Appendix A.2 for the dynamical law in terms of the Poisson bracket.

[38] We no longer distinguish between upper and lower indices since we now deal with dynamical systems and not with geometry. From now on all coordinates come with lower indices.

[39] Cf. Dirac [1964].

is just what is demanded. The constraint $\chi_l \approx 0$ fixes a surface transversally intersecting the gauge orbits. Together, the $2m$ constraints χ_l, ϕ_j define the $(2n-2m)$-dimensional constraint space $\Gamma^* \subset M$, the space of initial conditions of the reduced equations of motion. It is again a symplectic space inheriting the symplectic structure from Γ.

Note that Γ^* is exactly the reduced phase space we constructed in the preceding section. The ϕ_j are the conserved quantities of motion/the momentum functions determining the common level set $M \subset \Gamma$ and the χ_l are the connected gauge functions determining the base space of the fibre bundle $\pi : M \to \Gamma^*$ related to the respective symmetry transformation. Since the $\phi_j' s$ are equal to zero by definition of being a first class constraint, the reduced phase space is really a space of $2n-2m$ (and not less) dimensions.

Dirac captures the symplectic structure of Γ^* via the so-called Dirac bracket. While we can obtain the volume measure on Γ^* directly from the symplectic two-form connected to the Dirac bracket, there is another way, namely by means of the Faddeev-Popov construction.[40] This will be presented in what follows.

Theorem 5.1 (Faddeev-Popov measure). *Let $(\Gamma, \mathcal{B}(\Gamma), \mu, H)$ be a Hamiltonian system with m first class constraints $\phi_j \approx 0$. Then there exist m "gauge-fixing" functions $\chi_l \approx 0$ with $\{\chi_l, \phi_j\} \neq 0$ such that the Faddeev-Popov measure $\mathrm{d}\mu_{red}$ on Γ^* defined by*

$$\mathrm{d}\mu\big|_M = \mathrm{d}\mu_{red} \cdot \prod_{j=1}^{m} \mathrm{d}\varepsilon_j \tag{5.62}$$

can be written as

$$\mathrm{d}\mu_{red} = \det|\{\chi_l, \phi_j\}| \prod_{i=1}^{n} \prod_{j,l=1}^{m} \delta(\chi_l)\delta(\phi_j)\mathrm{d}q_i\mathrm{d}p_i. \tag{5.63}$$

Here M and Γ^* are as defined above, $\mathrm{d}\mu\big|_M$ is the Liouville measure on Γ restricted to M by delta functions and $\int \prod_j \mathrm{d}\varepsilon_j$ with ε_j from (5.60) is the volume of the gauge orbits.

Proof. To construct the volume measure on Γ^* as defined by (5.62), we start from the Liouville measure μ on Γ restricted to the constraint surface $M \subset \Gamma$ where M is determined by the m first class constraints. The restricted (or projected) measure is obtained by imposing delta functions $\prod_j \delta(\phi_j)$. We now find that we have a setting similar to above when we discussed the Feynman part integral. Again, on this constraint surface M, we have gauge transformations and gauge orbits and in order to factor out that part of the Liouville measure which measures the gauge volume,

$$\int \prod_{j=1}^{m} \mathrm{d}\varepsilon_j,$$

[40] Remember that every Poisson bracket relates to a symplectic two-form (cf. Appendix A). Both capture the symplectic structure of the manifold. In the same way as the usual Poisson bracket relates to the symplectic two-form on Γ, the Dirac bracket (defined with respect to the $2n-2m$ second class constraints ϕ_j, χ_l) relates to the symplectic form on the reduced space Γ^* (cf. Dirac [1964]).

we can use the Faddeev-Popov "trick". In that case,

$$\begin{aligned}\int \mathrm{d}\mu\big|_M &= \int \prod_{i=1}^{n}\prod_{j=1}^{m}\delta(\phi_j)\mathrm{d}q_i\mathrm{d}p_i \\ &= \int \prod_{i=1}^{n}\prod_{j=1}^{m}\delta(\phi_j)\int \Delta_{FP}\prod_{l=1}^{m}\delta(\chi_l^{\varepsilon_j})\mathrm{d}\varepsilon_j\mathrm{d}q_i\mathrm{d}p_i\end{aligned}$$

where Δ_{FP} must be such that

$$\int \Delta_{FP}\prod_{j=1}^{m}\prod_{l=1}^{m}\delta(\chi_l^{\varepsilon_j})\mathrm{d}\varepsilon_j = 1. \tag{5.64}$$

Here $\chi_l^{\varepsilon_j}$ is some particular gauge fixing. The subscript ε_j indicates that we consider some point which deviates from (q,p) by a small transformation $(\delta_{\phi_j}q, \delta_{\phi_j}p)$ induced by the generator ϕ_j. Here it is understood that (q,p) is a short writing for $(q_1,..,q_n,p_1,...,p_n)$ and $(\delta_{\phi_j}q,\delta_{\phi_j}p)$ for $(\delta_{\phi_j}q_1,...,\delta_{\phi_j}q_n,\delta_{\phi_j}p_1,...,\delta_{\phi_j}p_n)$. Hence, just like in case of the Feynman path integral, we integrate over the orbits $\int\prod_j \mathrm{d}\varepsilon_j$ and over some transversal surface determined by the particular gauge fixings $\delta(\chi_l^{\varepsilon_j})$. Now

$$\begin{aligned}\delta(\chi_l^{\varepsilon_j}(q,p)) &= \delta\big(\chi_l(q+\varepsilon_j\{q,\phi_j\}, p+\varepsilon_j\{p,\phi_j\})\big) \\ &= \frac{1}{|\partial\chi_l/\partial\varepsilon_j|}\delta(\varepsilon_j-\varepsilon_j^*) \\ &= \left[\sum_i \frac{\partial\chi_l}{\partial q_i}\{q_i,\phi_j\} + \frac{\partial\chi_l}{\partial p_i}\{p_i,\phi_j\}\right]^{-1}\delta(\varepsilon_j-\varepsilon_j^*) \\ &= \left[\sum_i \frac{\partial\chi_l}{\partial q_i}\frac{\partial\phi_j}{\partial p_i} + \frac{\partial\chi_l}{\partial p_i}\Big(-\frac{\partial\phi_j}{\partial q_i}\Big)\right]^{-1}\delta(\varepsilon_j-\varepsilon_j^*) \\ &= \frac{1}{\{\chi_l,\phi_j\}}\delta(\varepsilon_j-\varepsilon_j^*).\end{aligned}$$

Here the third equation is obtained by help of (5.60) and the forth and fifth equation by help of the definition of the Poisson bracket.

Inserting this result into (5.64), we can finally determine the Faddeev-Popov determinant. It is

$$\begin{aligned}1 &= \int \Delta_{FP}\prod_{j=1}^{m}\prod_{l=1}^{m}\delta(\chi_l^{\varepsilon_j})\mathrm{d}\varepsilon_j \\ &= \int \Delta_{FP}\prod_{j=1}^{m}\prod_{l=1}^{m}\frac{1}{\{\chi_l,\phi_j\}}\delta(\varepsilon_j-\varepsilon_j^*)\mathrm{d}\varepsilon_j.\end{aligned}$$

With

$$\int \delta(\varepsilon_j-\varepsilon_j^*)\mathrm{d}\varepsilon_j = 1$$

the following must hold:

$$\Delta_{FP} = \prod_{j=1}^{m}\prod_{l=1}^{m}\{\chi_l, \phi_j\} = \det|\{\chi_l, \phi_j\}|. \tag{5.65}$$

This allows us to rewrite the volume integral $\int \mathrm{d}\mu\big|_M$ as follows:

$$\begin{aligned}
\int \mathrm{d}\mu\big|_M &= \int \prod_{i=1}^{n}\prod_{j=1}^{m}\delta(\phi_j)\det|\{\chi_l, \phi_j\}| \int \prod_{l=1}^{m}\delta(\chi_l^{\varepsilon_j})\mathrm{d}\varepsilon_j \mathrm{d}q_i \mathrm{d}p_i \\
&= \int \det|\{\chi_l, \phi_j\}| \prod_{i=1}^{n}\prod_{j,l=1}^{m}\delta(\chi_l)\delta(\phi_j)\mathrm{d}q_i\mathrm{d}p_i \int \prod_{j=1}^{m}\mathrm{d}\varepsilon_j.
\end{aligned}$$

In the last step we were allowed to drop the subscript ε_j of χ_l because we integrate over all q_i, p_i. After that, the volume integral splits naturally into the volume of the reduced space and the volume of the gauge orbits. Let us finally factor out the gauge volume, $\int \prod_j \mathrm{d}\varepsilon_j$. This gives us the volume measure $\mathrm{d}\mu_{red}$ on the reduced space. From (5.64), that is, from the equation

$$\mathrm{d}\mu\big|_M = \mathrm{d}\mu_{red} \cdot \prod_{j=1}^{m}\mathrm{d}\varepsilon_j$$

we get that

$$\mathrm{d}\mu_{red} = \det|\{\chi_l, \phi_j\}| \prod_{i=1}^{n}\prod_{j,l=1}^{m}\delta(\chi_l)\delta(\phi_j)\mathrm{d}q_i\mathrm{d}p_i.$$

□

Corollary 5.1. *Let everything be as in the above theorem. In particular, let μ_{red} be the Faddeev-Popov measure on Γ^* with*

$$\mathrm{d}\mu_{red} = \det|\{\chi_l, \phi_j\}| \prod_{i=1}^{n}\prod_{j,l=1}^{m}\delta(\chi_l)\delta(\phi_j)\mathrm{d}q_i\mathrm{d}p_i.$$

This measure is invariant under the choice of gauge, that is, it is invariant under the choice of the functions χ_l.

Proof. The Faddeev-Popov measure μ_{red} is defined via equation (5.62), that is, via

$$\mathrm{d}\mu\big|_M = \mathrm{d}\mu_{red} \cdot \prod_j \mathrm{d}\varepsilon_j.$$

Now both $\int \mathrm{d}\mu\big|_M$ and $\int \prod_j \mathrm{d}\varepsilon_j$ are invariant under the choice of the functions χ_l fixing the gauge. Hence, also the total volume $\int \mathrm{d}\mu_{red}$ of the reduced space and, consequently, also the volume measure $\mathrm{d}\mu_{red}$ are invariant under the choice of the χ_l. □

Note the significance of this corollary. Let $\chi_l \approx 0$ with $\{\phi_j, \chi_l\} \neq 0$ a gauge fixing. The corollary tells us that instead of χ_l we can choose any other smooth function $\psi_l \approx 0$ with

$\{\phi_j, \psi_l\} \neq 0$ to do the gauge fixing and we still obtain the same measure. At the same time, we do not have to fix the gauge to zero. Let $\chi_l \approx 0$ a gauge fixing. Then also $\chi_l \approx c$ with $c \in \mathbb{R}$ is a gauge fixing because we can define a new function $\tilde{\chi}_l = \chi_l - c$ such that $\tilde{\chi}_l \approx 0$ and $\{\phi_j, \tilde{\chi}_l\} \neq 0$.

5.4 Comparison of the two approaches

Let us compare the measure obtained from the reduced symplectic two-form (cf. Sec. 5.2.4) to the measure obtained by the Faddeev-Popov construction (cf. Sec. 5.3). We have shown that both are volume measures on reduced phase space and both are obtained in a canonical, "gauge-invariant" manner. In fact, both measures coincide.

Lemma 5.5 (Equality of measures)**.** *Let there be a Hamiltonian system with constraints as described in Theorem 5.1. Let* $\mathrm{d}\mu$ *be the natural volume measure on* Γ^* *(obtained from* ω^**) and* $\mathrm{d}\mu_{red}$ *the Faddeev-Popov measure given by (5.63). Then*

$$\mathrm{d}\mu = \mathrm{d}\mu_{red}. \tag{5.66}$$

Proof. The proof is essentially based on Darboux's theorem which guarantees the existence of canonical local coordinates on any symplectic space.

Let us start with the method of symplectic reduction of Γ (as described in Sec.'s 5.1 and 5.2). Let us choose canonical coordinates which split naturally into internal coordinates q_i^*, p_i^* $(i = 1, ..., k)$ and external coordinates $\tilde{q}_j, \tilde{p}_j$ $(j = k+1, ..., n)$ with

$$\tilde{p}_j = 0 \qquad \text{and} \qquad \tilde{q}_j = c \tag{5.67}$$

with $c \in \mathbb{R}$. While the external coordinates are fixed by the constraints, the internal coordinates form a canonical set of local coordinates of reduced phase space. Here Darboux's theorem guarantees the existence of the respective coordinates and, thus, we can write the symplectic form ω on Γ as

$$\omega = \sum_i \mathrm{d}q_i^* \wedge \mathrm{d}p_i^* + \sum_j \mathrm{d}\tilde{q}_j \wedge \mathrm{d}\tilde{p}_j.$$

With respect to these coordinates, the symplectic form ω^* on reduced phase space Γ^* becomes

$$\omega^* = \sum_i \mathrm{d}q_i^* \wedge \mathrm{d}p_i^*$$

and the volume measure on Γ^* is

$$\mathrm{d}\mu = \frac{|(\omega^*)^k|}{k!} = \frac{|(\sum_i \mathrm{d}q_i^* \wedge \mathrm{d}p_i^*)^k|}{k!} = \prod_{i=1}^{k} \mathrm{d}q_i^* \mathrm{d}p_i^*. \tag{5.68}$$

On the other hand, Eq.'s (5.67) define $2n - 2k$ second class constraints which one can take to construct the Faddeev-Popov measure on Γ^*. Since the external coordinates $\tilde{q}_j, \tilde{p}_j$ with $j = k+1, ..., n$ are canonical by construction, it follows that the Faddeev-Popov determinant (5.65) is equal to one:

$$\Delta_{FP} = \det|\{\chi_l, \phi_j\}| = \det|\{\tilde{q}_l, \tilde{p}_j\}| = 1$$

(where now $\chi_l = \tilde{q}_l$ and $\phi_j = \tilde{p}_j$). Using that the change of coordinates from the q_i, p_i to the $q_i^*, p_i^*, \tilde{q}_j, \tilde{p}_j$ is canonical and, hence, volume-preserving, i.e.

$$\prod_{i=1}^{n} \mathrm{d}q_i \mathrm{d}p_i = \prod_{i=1}^{k} \prod_{j=k+1}^{n} \mathrm{d}q_i^* \mathrm{d}p_i^* \mathrm{d}\tilde{q}_j \mathrm{d}\tilde{p}_j,$$

the Faddeev-Popov measure (4.64) becomes

$$\begin{aligned} \mathrm{d}\mu_{red} &= \det|\{\chi_l, \phi_j\}| \prod_{i=1}^{n} \prod_{j,l=1}^{m} \delta(\chi_l)\delta(\phi_j)\mathrm{d}q_i \mathrm{d}p_i \\ &= \prod_{i=1}^{k} \prod_{j=k+1}^{n} \delta(\tilde{q}_j)\delta(\tilde{p}_j)\mathrm{d}q_i^* \mathrm{d}p_i^* \mathrm{d}\tilde{q}_j \mathrm{d}\tilde{p}_j = \prod_{i=1}^{k} \mathrm{d}q_i^* \mathrm{d}p_i^*. \end{aligned} \tag{5.69}$$

Comparing (5.68) and (5.69), we find that $\mathrm{d}\mu = \mathrm{d}\mu_{red}$. □

One can also prove the equality of both measures by direct computation. I do this for two constraints to show how it works.

Direct computation. Let us consider one pair of second class constraints. That is, let $\tilde{\Gamma}$ be a symplectic manifold with a symplectic two-form $\tilde{\omega}$ and H and χ two smooth functions on $\tilde{\Gamma}$ such that $H \approx 0$ and $\chi \approx 0$ and

$$\{H, \chi\} \neq 0.$$

Since the Poisson bracket is unequal to zero, it is clear that $H \approx 0$ and $\chi \approx 0$ determine a $(2n-2)$-dimensional constraint surface $\Gamma \subset \tilde{\Gamma}$.

Let $i_H : \Sigma \to \tilde{\Gamma}$ be the embedding of Σ in $\tilde{\Gamma}$ where $\Sigma \subset \tilde{\Gamma}$ is determined by $H \approx 0$ and let $i_\chi : \Gamma \to \Sigma$ be the embedding of Γ in Σ where $\Gamma \subset \Sigma$ is determined by $\chi \approx 0$. Then $\omega = i_\chi^* i_H^* \tilde{\omega}$ is a symplectic two-form on Γ. From this you can construct the natural volume form

$$\Omega = \frac{(-1)^{[\frac{n-1}{2}]}}{(n-1)!}\omega^{n-1}$$

and natural volume measure $\mathrm{d}\mu = |\Omega|$ on Γ. Let us, in what follows, determine ω in local coordinates q_i, p_i and from that $|\Omega|$. Explicitly, the following can be shown.

Lemma 5.6. *Let $\tilde{\Gamma}$, $\tilde{\omega}$, Γ, ω, H, χ and Ω be as above. Let q_i, p_i be local coordinates on $\tilde{\Gamma}$. On Γ, the natural volume measure $\mathrm{d}\mu = |\Omega|$ is*

$$|\Omega| = \left|1 + \sum_{i=2}^{n} \left(\frac{\partial \tilde{f}_0}{\partial q_i}\frac{\partial g_0}{\partial p_i} - \frac{\partial \tilde{f}_0}{\partial p_i}\frac{\partial g_0}{\partial q_i} \right)\right| \prod_{i=2}^{n} \mathrm{d}q_i \mathrm{d}p_i. \tag{5.70}$$

Here $\tilde{f}_0$ and g_0 solve the constraints $H \approx 0$ and $\chi \approx 0$ for q_1 and p_1. To be precise, it is $\tilde{f}_0 =$

$\tilde{f}_{H=0,\chi=0}$ with $q_1 = \tilde{f}_{H,\chi}(\hat{q}_1, ..., q_n, \hat{p}_1, ..., p_n)$ and $g_0 = g_{H=0}$ with $p_1 = g_H(q_1, ..., q_n, \hat{p}_1, ..., p_n)$. Here $\hat{q}_1$ denotes the omission of q_1.

Proof. Since H and χ are conjugates, i.e. $\{H, \chi\} \neq 0$, we can use the functional form of H and χ to express one of the (canonical) pairs of phase space variables q_i, p_i in terms of H and χ and all the other phase space variables. Here we use the implicit function theorem.

Let us, without loss of generality, consider the pair q_1, p_1 and let, again without loss of generality, H depend on p_1 and χ on q_1. Then there exist two smooth functions f and g depending on χ, respectively H such that

$$q_1 = f_\chi(\hat{q}_1, q_2, ..., q_n, p_1, ..., p_n), \qquad p_1 = g_H(q_1, ..., q_n, \hat{p}_1, p_2, ..., p_n).$$

Using the functional form of p_1, we can also express q_1 as a function of H and χ and the remaining phase space coordinates: $q_1 = \tilde{f}_{H,\chi}(\hat{q}_1, q_2, ..., q_n, \hat{p}_1, ..., p_n)$.

Expressing p_1 in terms of H and the other phase space variables, the symplectic form becomes

$$\begin{aligned} \tilde{\omega} &= \sum_{i=1}^{n} \mathrm{d}q_i \wedge \mathrm{d}p_i \\ &= \mathrm{d}q_1 \wedge \left(\frac{\partial g_H}{\partial H} \mathrm{d}H + \sum_{i=2}^{n} \frac{\partial g_H}{\partial q_i} \mathrm{d}q_i + \sum_{j=2}^{n} \frac{\partial g_H}{\partial p_j} \mathrm{d}p_j \right) + \sum_{k=2}^{n} \mathrm{d}q_k \wedge \mathrm{d}p_k \end{aligned}$$

where the sum over $\mathrm{d}q_i$ starts from $i = 2$ because $\mathrm{d}q_1 \wedge \mathrm{d}q_1 = 0$.

It follows that the pullback of the two-form $\tilde{\omega}$ on $\tilde{\Gamma}$ to Σ is

$$i_H^* \tilde{\omega} = \mathrm{d}q_1 \wedge \left(\sum_{i=2}^{n} \frac{\partial g_0}{\partial q_i} \mathrm{d}q_i + \sum_{j=2}^{n} \frac{\partial g_0}{\partial p_j} \mathrm{d}p_j \right) + \sum_{k=2}^{n} \mathrm{d}q_k \wedge \mathrm{d}p_k$$

with $g_0 = g_{H=0}$.

Let us now use $H \approx 0$ to express q_1 in terms of χ and the remaining phase space variables. That is, $q_1 = \tilde{f}_{0,\chi}(\hat{q}_1, q_2, ..., q_n, \hat{p}_1, ..., p_n)$, where $\tilde{f}_{0,\chi}$ is obtained by solving $f_\chi(\hat{q}_1, ..., q_n, g_0, p_2, ..., p_n)$ with $g_0 = g_0(q_1, ..., q_n, \hat{p}_1, ..., p_n)$ for q_1. The two-form $i_H^* \tilde{\omega}$ then becomes

$$\begin{aligned} i_H^* \tilde{\omega} &= \left(\frac{\partial \tilde{f}_{0,\chi}}{\partial \chi} \mathrm{d}\chi + \sum_{i=2}^{n} \frac{\partial \tilde{f}_{0,\chi}}{\partial q_i} \mathrm{d}q_i + \sum_{j=2}^{n} \frac{\partial \tilde{f}_{0,\chi}}{\partial p_j} \mathrm{d}p_j \right) \\ &\quad \wedge \left(\sum_{i=2}^{n} \frac{\partial g_0}{\partial q_i} \mathrm{d}q_i + \sum_{j=2}^{n} \frac{\partial g_0}{\partial p_j} \mathrm{d}p_j \right) \\ &\quad + \sum_{k=2}^{n} \mathrm{d}q_k \wedge \mathrm{d}p_k . \end{aligned}$$

It follows that the pullback of $i_H^*\tilde{\omega}$ on Σ to Γ is

$$\begin{aligned}\omega = i_\chi^* i_H^* \tilde{\omega} \quad &= \quad \left(\sum_{i=2}^n \frac{\partial \tilde{f}_0}{\partial q_i}\mathrm{d}q_i + \sum_{j=2}^n \frac{\partial \tilde{f}_0}{\partial p_j}\mathrm{d}p_j\right) \wedge \left(\sum_{i=2}^n \frac{\partial g_0}{\partial q_i}\mathrm{d}q_i + \sum_{j=2}^n \frac{\partial g_0}{\partial p_j}\mathrm{d}p_j\right) \\ &\quad + \sum_{k=2}^n \mathrm{d}q_k \wedge \mathrm{d}p_k.\end{aligned}$$

This two-form is again symplectic (since Γ is even-dimensional) and, hence, can be taken to construct a volume form Ω via the relation

$$\Omega = \frac{(-1)^{[\frac{n-1}{2}]}}{(n-1)!}\omega^{n-1}$$

where $n-1$ is half the dimension of Γ (which ensures that Ω is a top-dimensional form on Γ). From this we get a volume measure $\mu = |\Omega|$. Inserting the above expression for ω and observing that $\mathrm{d}q_i \wedge \mathrm{d}q_i = \mathrm{d}p_i \wedge \mathrm{d}p_i = 0$, we get

$$\frac{\omega^{n-1}}{(n-1)!} = \sum_{i=2}^n \left(\frac{\partial \tilde{f}_0}{\partial q_i}\frac{\partial g_0}{\partial p_i} - \frac{\partial \tilde{f}_0}{\partial p_i}\frac{\partial g_0}{\partial q_i}\right)\mathrm{d}q_2 \wedge ... \wedge \mathrm{d}p_n + \mathrm{d}q_2 \wedge ... \wedge \mathrm{d}p_n.$$

We also used that each of the contributing terms will turn up exactly $(n-1)!$ times. This is due to the $(n-1)!$ different possible orderings of pairs which are the result of the $(n-1)$-fold multiplication. This factor will then just cancel the $1/(n-1)!$ factor from the definition of the volume form.

Hence,

$$|\Omega| = \left|1 + \sum_{i=2}^n \left(\frac{\partial \tilde{f}_0}{\partial q_i}\frac{\partial g_0}{\partial p_i} - \frac{\partial \tilde{f}_0}{\partial p_i}\frac{\partial g_0}{\partial q_i}\right)\right| \prod_{i=2}^n \mathrm{d}q_i \mathrm{d}p_i.$$

□

Let us compare this to the Faddeev-Popov construction.

Lemma 5.7. *Let everything be as in the previous lemma. In particular, let there be given a pair of second class constraints H, χ. In that case, the Faddeev-Popov formula (5.63) for the measure becomes*

$$\mathrm{d}\mu_{red} = |\{H, \chi\}|\delta(H)\delta(\chi)\prod_{i=1}^n \mathrm{d}q_i \mathrm{d}p_i \tag{5.71}$$

and

$$\mathrm{d}\mu_{red} = |\Omega| \tag{5.72}$$

where $|\Omega|$ is the natural volume measure on Γ given by (5.70).

Proof. For one pair of second class constraints, the Faddeev-Popov formula gives

$$\begin{aligned}
\mathrm{d}\mu_{red} &= |\{H,\chi\}|\delta(H)\delta(\chi)\prod_{i=1}^{n}\mathrm{d}q_i\mathrm{d}p_i \\
&= |\{H,\chi\}|\left|\frac{\partial H}{\partial q_1}\frac{\partial \chi}{\partial p_1}-\frac{\partial H}{\partial p_1}\frac{\partial \chi}{\partial q_1}\right|^{-1}\delta(q_1-q_1^*)\delta(p_1-p_1^*)\prod_{i=1}^{n}\mathrm{d}q_i\mathrm{d}p_i.
\end{aligned}$$

Here the last equation makes use of the generalized formula for multivariable delta functions

$$\delta^{(n)}(\mathbf{f}(\mathbf{x}))=\frac{1}{|\det\partial\mathbf{f}_i/\partial\mathbf{x}_j|}\delta^{(n)}(\mathbf{x}-\mathbf{x}_0)$$

where the $\mathbf{x}_0$'s are the zeros of $\mathbf{f}(\mathbf{x})$. In this case, $\mathbf{f}=(H,\chi)$, $\mathbf{x}=(q_1,p_1)$ and $\mathbf{x}_0=(q_1^*,p_1^*)$ is the solution to the constraint equations $H\approx 0$ and $\chi\approx 0$. The q_1^* and p_1^* are functions of the remaining phase space coordinates q_i,p_i with $i=2,...,n$. In particular, they fulfill the equations $H=q_1-q_1^*\approx 0$ and $\chi=p_1-p_1^*\approx 0$. We can use the latter to rewrite H and χ. Then we get

$$\begin{aligned}
|\{H,\chi\}| &= \left|\sum_{i=1}^{n}\frac{\partial(q_1-q_1^*)}{\partial q_i}\frac{\partial(p_1-p_1^*)}{\partial p_i}-\frac{\partial(q_1-q_1^*)}{\partial p_i}\frac{\partial(p_1-p_1^*)}{\partial q_i}\right| \\
&= \left|1+\sum_{i=2}^{n}\frac{\partial q_1^*}{\partial q_i}\frac{\partial p_1^*}{\partial p_i}-\frac{\partial q_1^*}{\partial p_i}\frac{\partial p_1^*}{\partial q_i}\right|.
\end{aligned}$$

In addition,

$$\left|\frac{\partial H}{\partial q^1}\frac{\partial \chi}{\partial p_1}-\frac{\partial H}{\partial p_1}\frac{\partial \chi}{\partial q_1}\right|^{-1}=\left|\frac{\partial(q_1-q_1^*)}{\partial q_1}\frac{\partial(p_1-p_1^*)}{\partial p_1}-\frac{\partial(q_1-q_1^*)}{\partial p_1}\frac{\partial(p_1-p_1^*)}{\partial q_1}\right|^{-1}=1.$$

Both results follow from the fact that q_1^* and p_1^* do not depend on q_1 and p_1, but only on the other variables. Let us reinsert the two expressions in the above equation for μ_{red}. Since

$$\delta(q_1-q_1^*)\delta(p_1-p_1^*)\prod_{i=1}^{n}\mathrm{d}q_i\mathrm{d}p_i=\prod_{i=2}^{n}\mathrm{d}q_i\mathrm{d}p_i,$$

it follows that the reduced measure can be written as

$$\mu_{red}=\left|1+\sum_{i=2}^{n}\left(\frac{\partial q_1^*}{\partial q_i}\frac{\partial p_1^*}{\partial p_i}-\frac{\partial q_1^*}{\partial p_i}\frac{\partial p_1^*}{\partial q^i}\right)\right|\prod_{i=2}^{n}\mathrm{d}q_i\mathrm{d}p_i.$$

Identifying $\tilde{f}_0$ with q_1^* and g_0 with p_1^* we find that this is the canonical volume measure $|\Omega|$ given by (5.70). □

Clearly, the last two lemmas can be generalized to multiple pairs of second class constraints. Only then many more terms will arise, in particular many mixed terms displaying the inter-dependence of the constraint functions, so that the computation will not be as short and nice.

Part III

A Statistical Analysis of the Newtonian Universe

6 A normalizable measure of typicality for the Newtonian universe

In this section, we construct and discuss a (normalizable) measure with respect to which a statistical analysis of the Newtonian universe can be performed. This measure has been obtained by Barbour, Koslowski, and Mercati [2015] using the formula of Faddeev and Popov. We show how to construct the measure from the underlying geometric structure. We analyze its behavior under the dynamics and elaborate on its explanatory value and its connection to the notion of entropy in the end.

6.1 Introduction

In their 2015 paper, Barbour, Koslowski, and Mercati introduce a normalizable measure on the space of (physically distinct) mid-point data PT^*S of the $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian universe. From this measure they obtain an entropy-type quantity: the entaxy. The entaxy is essentially the volume of (macro)states of the universe of constant complexity. Here "volume" refers to the normalizable measure over mid-point data (solution-determining data) on PT^*S and "complexity" is a scale-invariant macro-variable specifying the shape of the system (approximately given by the largest distance divided by the smallest distance between the particles). Since the entaxy is defined with respect to a measure over mid-point data, it determines the volume of macrostates of the universe at that particular point (the so-called Janus point or Big Bang) of the Newtonian universe.

The crucial point of this analysis is that the volume measure on the space of mid-point data PT^*S – which, as I explain, is the correct space for the statistical analysis of the system – is normalizable. That way the usual statistical analysis can be performed.

Barbour, Koslowski, and Mercati [2015] obtain the measure on PT^*S by elimination of all the redundant, non-physical degrees of freedom of the system. To get rid off these extra degrees of freedom, Barbour et al. in a first step reduce the phase space of the $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian universe making use of the symmetries of the system. Adopting a Machian perspective, they claim that all states are equal which can be arrived at by an overall translation and/or rotation. This explains why they consider the $\mathbf{P} = \mathbf{L} = 0$ Newtonian universe – only then the system is fully translationally and rotationally invariant – and take the reduced phase space to be the space of physically distinct states.

Further they go to the constant energy hypersurface and by help of a monotonic parameter, an internal time τ, they determine the dynamics on a hypersurface of constant internal time cutting the trajectories transversally. This hypersurface is a representative of the space of solutions Γ_{sol}

where each point represents an entire solution and on which the internal (reduced) dynamics can be defined (cf. Lemma 4.6). The space of solutions turns out to be isomorphic to the cotangent bundle of shape space: $\Gamma_{sol} \cong T^*S$.

Finally, Barbour, Koslowski, and Mercati use the remaining dynamical similarity (the invariance of the reduced internal equations of motion under a simultaneous scaling of internal time τ and the shape momenta) to reduce the space of solutions by one further dimension. That way, they end up on a compact space: the projective cotangent bundle of shape space PT^*S which is the space of physically distinct solutions, respectively, the space of physically distinct mid-point data. On that space, a normalizable volume measure – the measure of typicality of the $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian universe – can be defined.

Remark (Machian universe)**.** Given a description of the Newtonian universe with respect to absolute space and time, Barbour, Koslowski, and Mercati aim to derive a description free of what they consider to be the unphysical degrees of freedom. Starting from first principles in the spirit of Mach,[41] they claim that all configurations of the system are physically equivalent which can be transformed into one another by an overall spatial translation, rotation or dilation. Unfortunately, however, the Newtonian dynamics is only invariant with respect to translations and rotations (and with respect to the full rotational group only if the total angular momentum vanishes – this is why we set $\mathbf{L} = 0$), but not with respect to dilations.

In some of his earlier work,[42] Barbour constructed a theory of gravitation that is invariant under the full similarity group (including dilations). Clearly, this theory of gravitation is different from the Newtonian, but it recovers the Newtonian force law to some extent. Later Barbour gave up on this project and now claims that from that model nothing interesting – in particular, no structure formation like the one we observe in our universe – can be obtained.[43] Everything of interest (like structure formation, arrows of time, etc.) appears naturally as soon as we include scale in the description. This is why in their 2013 and 2015 papers, Barbour, Koslowski and Mercati start from the Newtonian theory of gravitation, set $\mathbf{P} = \mathbf{L} = 0$ (and also $E = 0$) for Machian reasons and reduce the dynamics with respect to translations and rotations, but not with respect to scalings. What regards the measure of typicality of that system they are particularly lucky because due the existence of a monotonic variable, an internal time parameter τ, together with the dynamical similarity of the internal equations of motion they get rid of scales in the end.

6.2 Dynamics of the $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian universe

Let us consider the $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian universe. In the standard Hamiltonian formulation, there is a Hamiltonian $H = T + V_N$ on Γ where T is the kinetic energy and V_N the

[41] Cf. Barbour and Pfister (Eds.) [1995] for a discussion of Mach's principles.
[42] Cf. Barbour [2003]
[43] From a talk by J. Barbour at the summer school on "Philosophy of Physics" in Saig in 2017. Cf. also Barbour et al. [2013].

Newtonian gravitational potential V_N, that is,

$$H = T + V_N = \sum_{i=1}^{N} \frac{\mathbf{p}_i^2}{2m_i} - \sum_{\substack{i<j\\ i,j=1}}^{N} \frac{Gm_i m_j}{|\mathbf{q}_i - \mathbf{q}_j|}. \tag{6.1}$$

We presented the most important results concerning the dynamics of the $E = 0$ Newtonian universe in Section 3.1. The dynamics is governed by the Lagrange-Jacobi equation,

$$\ddot{I} = 4E - 2V_N > 0 \quad \text{if} \quad E = 0 \tag{6.2}$$

and by Pollard's result stating that, for $E = 0$, $I \to \infty$ as $t \to \pm\infty$.[44]

Hence, for the $E = 0$ Newtonian universe the following scenario is due. At some moment of time the moment of inertia is minimal, $I = I_{min}$, whereas I increases in both time directions away from that. Now clearly, I is a measure of the total extension of the particles. In other words, at some moment in time the particles are closest while they spread starting from that moment in both time directions.

The point at which the particles are closest represents the Big Bang within the $E = 0$ Newtonian universe. In addition, the increase of I determines a gravitational arrow of time. Hence, in this scenario, there are two arrows of time pointing in opposite directions away from the point of minimal extension. Let me say this again. There is one common past at the moment at which $I = I_{min}$ (the Big Bang) and there are two futures in both directions away from that (for $I \to \infty$, respectively $t \to \pm\infty$).

Due to the results of Saari [1971], which we described in Section 3.1, we have an even more precise idea of the behavior of the N-particle system as $t \to \infty$. According to Saari's results the generic behavior is the following. As $t \to \infty$ the system forms clusters consisting of particles whose inter-partice distances are bounded. The centers of mass of these clusters recede from each other at a rate of about $t^{2/3}$. Moreover, the system forms subsystems whose centers of mass recede from each other at a rate proportional to t.

Later we will, for means of simplicity, discuss the three-particle universe. For three particles, the generic behavior has been shown to be the following.[45] At some moment in time the particles are closest while in both time directions away from that point a binary forms[46] (given some mild assumptions on the initial conditions). As time evolves, the binary becomes a more and more perfect Kepler pair with the third particles receding from the center of mass of the binary asymptotically proportional to time t.

From now on we want to restrict the discussion to the three-body system as the simplest

[44] Cf. Pollard [1967] and the Appendix C.

[45] Cf. Barbour et al. [2013] and Moeckel [1981], [2007] for a study of the generic behavior of the three-particle system.

[46] A binary is a two-particle system. We say "a binary forms" when two particles stick together from some moment of time onwards. We say that the two particles become "a more and more perfect Kepler pair" to express that the gravitational interaction between the third particle and the binary becomes more and more negligible as the third particle moves away from the binary.

non-trivial model of the Newtonian universe. For the N-body system, some of the results follow immediately while others will be hard to show. Though the N-body system is much more complicated, we expect that analogous results can be obtained.

6.3 Dynamics on T^*S_R and T^*S

Let us study the dynamics of the three-particle Newtonian universe on reduced and internal phase space. Recall the symplectic reduction of phase space for the three-particle system (Sec. 5.2). The Newtonian theory of gravity is invariant under spatial translations and rotations, but not under scalings. This follows from the transformation properties of the Hamiltonian.

Lemma 6.1 (Invariance of H)**.** *Let $H = T + V_N$ defined by (6.1) the Hamiltonian of the system. H is invariant under translations and rotations, but not under scalings.*

Proof. The invariance properties of the Hamiltonian follow directly from the invariance properties of the Newton potential. The Newton potential V_N is invariant under translations and rotations, but not under scalings because it depends on (nothing but) the inter-particle distances, $V_N = V_N(|\mathbf{q}_i - \mathbf{q}_j|)$, which are invariant under translations and rotations of the whole system, but not under scalings. □

Equivalently, we can determine the symmetries of the system via Nöther's theorem by computing the conserved quantities of the system. Next to total energy $H = E$ these are the total linear momentum $\mathbf{P}$ and the total angular momentum $\mathbf{L}$, but not the dilational momentum $D = \sum_i \mathbf{q}_i \mathbf{p}_i$ since $\{H, D\} \neq 0$.

It follows that the reduced dynamics of the $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian universe are dynamics on shape phase space with scale T^*S_R.

6.3.1 Reduced Hamiltonian dynamics on T^*S_R

Recall the symplectic reduction of the three-particle $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian universe with respect to translations and rotations described in Sections 5.2.1 and 5.2.2. Reduced phase space is shape phase space with scale T^*S_R. Local coordinates of $T^*\mathrm{S}_R$ are the three translationally and rotationally invariant Hopf variables $\mathbf{w} = (w_1, w_2, w_3)$ and their three conjugate momenta $\mathbf{z} = (z_1, z_2, z_3)$ defined in (5.35) and (5.36). They form a set of canonical coordinates. On T^*S_R, there exists a reduced Hamiltonian $H = H(\mathbf{w}, \mathbf{z})$ which determines the reduced Hamiltonian equations of motion. For three particles, the reduced Hamiltonian has been obtained by Montgomery [2002]. Before we derive the reduced Hamiltonian, let me add one definition.

Definition 6.1 (Shape potential V_S)**.** Let V_N the Newton potential given in (6.1) and $\mathbf{w} = (w_1, w_2, w_3)$ the Hopf coordinates (5.35). We call

$$V_S = V_N \cdot \sqrt{||\mathbf{w}||} \tag{6.3}$$

the *shape potential*.

It follows from a simple dimensional analysis that V_S is homogeneous of degree zero and as such invariant under scalings. Now the reduced Hamiltonian can be obtained.[47]

Lemma 6.2 (Reduced Hamiltonian on T^*S_R). *Let $\boldsymbol{w}$ and $\boldsymbol{z}$ the Hopf coordinates (5.35) and their canonical conjugates (5.36). For the $\boldsymbol{P} = \boldsymbol{L} = 0$ Newtonian universe with Hamiltonian H on $\Gamma = T^*Q$ given by (6.1) there exists a reduced Hamiltonian $H = H(\boldsymbol{w}, \boldsymbol{z})$ on T^*S_R of the form*

$$H = ||\boldsymbol{w}|| \cdot ||\boldsymbol{z}||^2 + \frac{V_S}{\sqrt{||\boldsymbol{w}||}} \tag{6.4}$$

with

$$V_S = -\sqrt{2} \sum_{\substack{i<j \\ i,j=1}}^{3} \frac{G(m_i m_j)^{\frac{3}{2}} (m_i + m_j)^{-\frac{1}{2}}}{\sqrt{1 - \boldsymbol{w} \cdot \boldsymbol{b}_{ij}/||\boldsymbol{w}||}}. \tag{6.5}$$

Here the $\mathbf{b}_{ij}$ are three unit vectors representing the three binary collisions.

Proof. The reduced Hamiltonian can be obtained from the original Hamiltonian

$$\begin{aligned} H(\mathbf{q}, \mathbf{p}) &= T + V_N \\ &= \sum_{i=1}^{3} \frac{\mathbf{p}_i^2}{2m_i} - \sum_{\substack{i<j \\ i,j=1}}^{3} \frac{G m_i m_j}{|\mathbf{q}_i - \mathbf{q}_j|} \end{aligned}$$

by imposing the constraints $\mathbf{P} = \mathbf{L} = 0$. To do that, we rewrite H in terms of the $\mathbf{w}$ and $\mathbf{z}$ coordinates and in terms of L given by (5.38) and $\boldsymbol{\kappa}_3$ given by (5.30).

The kinetic term can be rewritten as follows:

$$T = \sum_{i=1}^{3} \frac{\mathbf{p}_i^2}{2m_i} = \sum_{i=1}^{3} \frac{\boldsymbol{\kappa}_i^2}{2} = ||\mathbf{w}|| ||\mathbf{z}||^2 + \frac{L^2}{2||\mathbf{w}||} + \frac{|\boldsymbol{\kappa}_3|^2}{2}.$$

Setting $\mathbf{P} = 0$ means setting $\boldsymbol{\kappa}_3 = \sum_{i=1}^{3} \mathbf{p}_i = 0$. Moreover, from $\mathbf{L} = 0$ we get that $L = |\mathbf{L}| = 0$. Hence, the reduced kinetic term is

$$T = ||\mathbf{w}|| ||\mathbf{z}||^2.$$

With respect to the $\mathbf{w}$ coordinates, the Newton potential can be rewritten as

$$\begin{aligned} V_N &= -\sum_{i<j} \frac{G m_i m_j}{|\mathbf{q}_i - \mathbf{q}_j|} \\ &= -\sum_{i<j} \frac{\sqrt{2} G (m_i m_j)^{\frac{3}{2}} (m_i + m_j)^{-\frac{1}{2}}}{\sqrt{||\mathbf{w}|| - \mathbf{w} \cdot \mathbf{b}_{ij}}}, \end{aligned}$$

where the $\mathbf{b}_{ij} \in \mathbb{R}^3$ are unit vectors representing the three binary collisions. Recall that, for three particles, there are three different binary collisions, the collision between particles 1 and 2,

[47] Cf. also Montgomery [2002] for a derivation of the Newton potential in terms of the Hopf coordinates.

2 and 3, and 1 and 3, respectively. These are represented by three points on shape space with scale S_R, where S_R can be depicted as a two-sphere of radius $R = ||\mathbf{w}||$, cf. (5.50). Explicitly, the three vectors representing the three binary collisions are $||\mathbf{w}|| \cdot \mathbf{b}_{12}$ for the collision between particles 1 and 2 (where $|\mathbf{q}_1 - \mathbf{q}_2| = 0$), $||\mathbf{w}|| \cdot \mathbf{b}_{23}$ for the collision between particles 2 and 3, and so on. Now the distance $|\mathbf{q}_i - \mathbf{q}_j|$ between any two particles scales with $\sqrt{||\mathbf{w}||}$ (since $\sqrt{||\mathbf{w}||} = 1/2\sqrt{I}$ (cf. (5.50)) and I can be expressed in terms of the inter-particles distances according to (2.19)). It follows from (2.19) that $|\mathbf{q}_i - \mathbf{q}_j| = \sqrt{||\mathbf{w}||}$ (times the correct mass factor) if there is a collision between one of the particles i, j and the third particle k (and the same for the equilateral triangle, with a different factor). And, of course, $|\mathbf{q}_i - \mathbf{q}_j| = 0$ in case there is a collision between particles i and j. Every other distance $|\mathbf{q}_i - \mathbf{q}_j|$ lies in between 0 and $\sqrt{||\mathbf{w}||}$ depending only on the "distance" between the actual configuration $\mathbf{w}$ and the binary collision points $\mathbf{b}_{ij}$ (as expressed in $\mathbf{w} \cdot \mathbf{b}_{ij}/||\mathbf{w}||$). This can be obtained explicitly from (2.19) for $N = 3$ and different masses, from where we also get the correct mass factor. Factoring out scale $\sqrt{||\mathbf{w}||}$ and with V_S defined by (6.3) the above assertion follows. □

From this, the following corollary can be obtained.

Corollary 6.1. *Let (R, ψ, ϕ) and (p_R, p_ψ, p_ϕ) be the spherical coordinates defined in (5.41) and (5.42). In terms of these coordinates, the reduced Hamiltonian is*

$$H = \frac{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + R^2 p_R^2}{R} + \frac{V_S(\psi, \phi)}{\sqrt{R}} \tag{6.6}$$

with

$$V_S = -\sum_{i<j} \frac{G(m_i m_j)^{\frac{3}{2}}(m_i + m_j)^{-\frac{1}{2}}}{\sqrt{1 - \sin\psi\cos(\phi - \phi_{ij})}}. \tag{6.7}$$

Proof. The form of the kinetic part of the Hamiltonian follows directly from expressing the $\mathbf{w}$ and $\mathbf{z}$ in the spherical coordinates (R, ψ, ϕ) and (p_R, p_ψ, p_ϕ).

To determine V_S in the new coordinates, note that the binary collision vectors $\mathbf{b}_{ij}$ always lie in the $w_3 = 0$ plane. This follows from the fact that $w_3 = \boldsymbol{\rho}_1 \times \boldsymbol{\rho}_2$ (cf. (5.35)) where $\boldsymbol{\rho}_1$ is the vector between particles 1 and 2 and $\boldsymbol{\rho}_2$ the vector between particle 3 and the center of mass between particles 1 and 2. Accordingly, $w_3 = 0$ if and only if the three particles are collinear. For the given choice of spherical coordinates, the collinear configurations lie on the equator of the shape sphere where the equator is specified by the angle $\psi = \pi/2$. Since $\mathbf{b}_{ij}$ is a unit vector pointing in the direction of the binary collision points, it can be written as $\mathbf{b}_{ij} = (\sin\psi_{ij}\cos\phi_{ij}, \sin\psi_{ij}\sin\phi_{ij}, \cos\psi_{ij})$ with $\psi_{ij} = \pi/2$ and where the three angles ϕ_{ij} represent the position of the three binary collision points on the equator of the shape sphere. It follows that

$$\mathbf{w} \cdot \mathbf{b}_{ij} = ||\mathbf{w}|| \sin\psi\cos(\phi - \phi_{ij}).$$

Here the $(\phi - \phi_{ij})$ are the angles between the $\mathbf{b}_{ij}$ and the projection of $\mathbf{w}$ onto the $w_3 = 0$ plane and the term $||\mathbf{w}|| \sin\psi$ is the component of $\mathbf{w}$ which is parallel to the $w_3 = 0$ plane. □

Equations of motion on T^*S_R. The reduced Hamiltonian determines reduced Hamiltonian equations of motion. Consider H from (6.6). The respective equations of motion are

$$\begin{aligned}
\frac{\mathrm{d}\psi}{\mathrm{d}t} &= \frac{2p_\psi}{R}, & \frac{\mathrm{d}p_\psi}{\mathrm{d}t} &= \frac{2\sin^{-3}\psi\cos\psi p_\phi^2}{R} - \frac{\partial V_S/\partial\psi}{\sqrt{R}}, \\
\frac{\mathrm{d}\phi}{\mathrm{d}t} &= \frac{2\sin^{-2}\psi p_\phi}{R}, & \frac{\mathrm{d}p_\phi}{\mathrm{d}t} &= -\frac{\partial V_S/\partial\phi}{\sqrt{R}}, \\
\frac{\mathrm{d}R}{\mathrm{d}t} &= 2R\cdot p_R, & \frac{\mathrm{d}p_R}{\mathrm{d}t} &= \frac{p_\psi^2+\sin^{-2}\psi p_\phi^2 - R^2p_R^2}{R^2} + \frac{1}{2}\frac{V_S(\psi,\phi)}{R^{3/2}}
\end{aligned} \tag{6.8}$$

where $\mathrm{d}\psi/\mathrm{d}t = \partial H/\partial p_\psi$, $\mathrm{d}p_\psi/\mathrm{d}t = -\partial H/\partial\psi$ and so on. These equations govern the motion on T^*S_R.

6.3.2 Internal Hamiltonian dynamics on T^*S

At this point the setting is the following. The dynamics of the $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian universe is described by the equations of motion on the translationally and rotationally reduced phase space $T^*\mathrm{S}_R$. This space is symplectic with the symplectic two-form given by

$$\omega = \mathrm{d}R\wedge\mathrm{d}p_R + \mathrm{d}\psi\wedge\mathrm{d}p_\psi + \mathrm{d}\phi\wedge\mathrm{d}p_\phi.$$

Moreover, since energy is conserved ($E = 0$), there exists a Hamiltonian constraint on T^*S_R given by

$$H = 0. \tag{6.9}$$

This constraint determines a five-dimensional submanifold of T^*S_R. Possible trajectories of the system are restricted to that submanifold.

If we could find an internal time parameter, we could reduce by one further dimension and describe the dynamics on the internal space, a hypersurface of constant internal time (cf. Sec. 3.3). Luckily, there exists a monotonic dynamical variable for all solutions on T^*S_R with $E \geq 0$: the dilational momentum $D = \sum \mathbf{q}_i\cdot\mathbf{p}_i$.

Lemma 6.3. *Let $D = \sum_{i=1}^N \boldsymbol{q}_i\cdot\boldsymbol{p}_i$ and $H = E$ with $H = \sum_{i=1}^N \boldsymbol{p}_i^2/2m_i + V_N$ given by (6.1). Then*

$$\frac{\mathrm{d}D}{\mathrm{d}t} = \{D,H\} = 2E - V_N. \tag{6.10}$$

For $E = 0$, we have

$$\{D,H\} = -V_N > 0. \tag{6.11}$$

Proof. Computing the Poisson bracket, we have

$$\begin{aligned}\frac{\mathrm{d}D}{\mathrm{d}t} = \{D, H\} &= \sum_{i=1}^{3} \left(\frac{\partial D}{\partial \mathbf{q}_i} \frac{\partial H}{\partial \mathbf{p}_i} - \frac{\partial H}{\partial \mathbf{q}_i} \frac{\partial D}{\partial \mathbf{p}_i} \right) \\ &= \sum_{i=1}^{3} \left(\mathbf{p}_i \cdot \mathbf{p}_i / m_i - \mathbf{q}_i \cdot \frac{\partial V_N}{\partial \mathbf{q}_i} \right) \\ &= 2H - 2V_N - \mathbf{q}_i \cdot \frac{\partial V_N}{\partial \mathbf{q}_i}.\end{aligned}$$

Let us now recall Euler's homogeneous function theorem which says that, for any function f which is homogeneous of a certain degree k, that is, for a function such that $f(\alpha q_1, .., \alpha q^n) = \alpha^k f(q_1, ..., q^n)$ for some α, it holds that $kf(q_1, ...q_n) = \sum_{i=1}^n q_i \partial f / \partial q_i$.

For the Newton potential, $k = -1$ and

$$-V_N = \sum_{i=1}^{3} \mathbf{q}_i \cdot \frac{\partial V_N}{\partial \mathbf{q}_i}.$$

With this and $H = T + V_N = E$, it follows that

$$\{D, H\} = 2H - 2V_N - \mathbf{q}_i \cdot \frac{\partial V_N}{\partial \mathbf{q}_i} = 2H - V_N = 2E - V_N.$$

For $E = 0$, we have $\{D, H\} = -V_N$ with $V_N < 0$. Hence, Eq. (6.11) follows. $\square$

Since H and D are translationally and rotationally invariant, the Poisson bracket $\{D, H\}$ can equally well be evaluated on reduced phase space $T^* S_R$. Let us check this. Let $\{\cdot, \cdot\}^*$ denote the (reduced) Poisson bracket on $T^* S_R$ and consider H given by (6.6). In addition, $D = 2R \cdot p_R$ (cf. (5.49)). Hence, in the Poisson bracket $\{\cdot, \cdot\}^*$ all derivatives other than those with respect to R and p_R vanish and you get

$$\begin{aligned}\{D, H\}^* &= \frac{\partial D}{\partial R} \frac{\partial H}{\partial p_R} - \frac{\partial H}{\partial R} \frac{\partial D}{\partial p_R} \\ &= \big(2p_R\big)\big(2Rp_R\big) - \left(- \frac{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + R^2 p_R^2}{R^2} - \frac{1}{2} \frac{V_S(\psi, \phi)}{R^{3/2}} + 2p_R^2 \right) 2R \\ &= 2H - V_S / \sqrt{R} = 2H - V_N = 2E - V_N \qquad (6.12)\end{aligned}$$

on $H = E$ in agreement with Lemma 6.2.

The monotonicity of D also follows directly from the Lagrange-Jacobi inequality (6.2) noting that, since $D = \sum \mathbf{q}_i \mathbf{p}_i$ and $I = \sum m_i \mathbf{q}_i^2$,

$$D = \frac{1}{2} \dot{I}. \qquad (6.13)$$

That is, D is half the first time derivative of the moment of inertia. From the Lagrange-Jacobi inequality we know that, if $E = 0$, I is concave upwards, $\ddot{I} > 0$. It follows that D is monotonically increasing, $\dot{D} > 0$, with

$$D = 0 \qquad \text{iff} \qquad I = I_{min}. \tag{6.14}$$

Of course, this argument is essentially equivalent to our computation above (Lemma 6.3). Note aside that D is not increasing at a constant rate, since $\mathrm{d}D/\mathrm{d}t = -V_N \neq const.$

With respect to D, the internal Hamiltonian F governing the motion is given as follows.

Lemma 6.4 (Internal Hamiltonian on T^*S)**.** *Let H be the reduced Hamiltonian on T^*S_R given by (6.6) and $D = 2R \cdot p_R$ (cf. (5.49)). The internal Hamiltonian F governing the motion on T^*S with respect to internal time D is*

$$F = \log\left[\frac{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + \frac{1}{4}D^2}{-V_S(\psi,\phi)}\right]. \tag{6.15}$$

Proof. Remember that the internal Hamiltonian F is the canonical conjugate of D restricted to the constant energy hypersurface (as determined by the Hamiltonian constraint $H = 0$). It is a function of the internal variables and possibly of D. Let again $\{\cdot,\cdot\}^*$ denote the Poisson bracket on T^*S_R. First observe that $1/2 \log R$ is canonical conjugate to $D = 2R \cdot p_R$, since

$$\{1/2\log R, D\}^* = \frac{\partial(1/2\log R)}{\partial R}\frac{\partial D}{\partial p_R} - \frac{\partial(1/2\log R)}{\partial p_R}\frac{\partial D}{\partial R} = \frac{1}{2R}\cdot 2R - 0 = 1.$$

We now express $1/2\log R = \log\sqrt{R}$ in terms of the internal variables ψ, ϕ, p_ψ, and p_ϕ and possibly of D by help of the Hamiltonian constraint $H = 0$. Inserting $D = 2R \cdot p_R$ in (6.6), the Hamiltonian constraint reads

$$H = \frac{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + \frac{1}{4}D^2}{R} + \frac{V_S(\psi,\phi)}{\sqrt{R}} = 0,$$

hence

$$\sqrt{R}\big|_{H=0} = \frac{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + \frac{1}{4}D^2}{-V_S(\psi,\phi)}.$$

Hence, (6.15) follows.

From (4.44) we know that F gives us the equations of motion on a hypersurface of constant internal time $D = D^*$ transversally intersecting the trajectories. What is the geometry of that hypersurface? From Lemma 4.7 we know that it is a symplectic space. For this particular model, we know even more than that.

Let us without loss of generality consider the hypersurface determined by $D = 0$. Remember that, apart from being the internal time parameter, D is also the generator of dilations. Now $H = 0$ can be interpreted as its gauge fixing. This is possible since $\{D, H\} \neq 0$. Consequently, we can read the pair of constraints $D = 0$ and $H = 0$ as specifying the dilationally reduced phase space $\Gamma_d \cong T^*S$. (Whereas in Section 5.2.3 we fixed the gauge related to dilations by fixing the moment of inertia to one, $I = 1$, here we fix the gauge by setting $H = 0$. Since the construction of reduced phase space is gauge invariant, we end up on the same space: T^*S). Hence, by imposing the constraints $D = H = 0$ we end up on T^*S and F determines internal

Hamiltonian equations of motion on T^*S. $\square$

Be aware that, although we end up on T^*S, we did not reduce the dynamics with respect to dilations. This is impossible from the very beginning because the Newtonian dynamics is not scale-invariant. Instead, we used the fact that D is a monotonic parameter with respect to which we can formulate internal Hamiltonian dynamics as described in Section 4.3. By chance, this internal phase space is isomorphic to dilationally reduced phase space T^*S.

Note that $F = \log \sqrt{R}|_{H=0}$ is a function depending on D. That is, we have here a "time"-dependent Hamiltonian $F = F_D(\psi, \phi, p_\psi, p_\phi)$. This time-dependent Hamiltonian F_D generates a time-dependent vector field X_D.

Equations of motion on T^*S. The internal Hamiltonian equations of motion with respect to D are:

$$\begin{aligned} \frac{\mathrm{d}\psi}{\mathrm{d}D} &= \frac{2p_\psi}{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + \frac{1}{4}D^2}, & \frac{\mathrm{d}p_\psi}{\mathrm{d}D} &= \frac{2\sin^{-3}\psi\cos\psi p_\phi^2}{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + \frac{1}{4}D^2} + \frac{\partial \log(-V_s)}{\partial \psi}, \\ \frac{\mathrm{d}\phi}{\mathrm{d}D} &= \frac{2\sin^{-2}\psi p_\phi}{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + \frac{1}{4}D^2}, & \frac{\mathrm{d}p_\phi}{\mathrm{d}D} &= \frac{\partial \log(-V_s)}{\partial \phi}, \end{aligned} \tag{6.16}$$

where $\mathrm{d}\psi/\mathrm{d}D = \mathrm{d}F_D/\mathrm{d}p_\psi$, $\mathrm{d}p_\psi/\mathrm{d}D = -\mathrm{d}F_D/\mathrm{d}\psi$, and analogously for ϕ, p_ϕ.

Remark (Dual role of T^*S)**.** Note that T^*S plays a dual role. Just like ordinary phase space Γ it is both the space in which the trajectories lie and the space of initial conditions where each point represents an entire solution. The latter is the case because T^*S is obtained from the (reduced) constant energy surface $T^*S_R\big|_{H=0}$ in which the trajectories lie by fixing internal time $D = 0$. Since D is monotonic, the $D = 0$ hypersurface is cut once and only once by each of the trajectories. As such each point on T^*S represents one solution. This is why we also call it the space of solutions $\Gamma_{sol} \cong T^*S$. At the same time, formulating the dynamics with respect to internal time D allows us to project the trajectories onto T^*S, that is, the trajectories parametrized by D lie in T^*S.

6.4 A volume measure on $\Gamma_{sol} \cong T^*S$

Both the reduction with respect to the symmetries together with the choice of an internal time parameter have helped us to reduce the number of dimensions of phase space from 18 to 4. In the end we obtained a Hamiltonian formulation of the dynamics on four-dimensional T^*S. Let us now determine the invariant volume measure on T^*S.

6.4.1 Internal Hamiltonian description and measure

We can derive the measure directly from the underlying symplectic structure of T^*S. We determined the natural volume measure on dilationally reduced phase space $\Gamma_d \cong T^*S$ in (5.52):

$$\mathrm{d}\mu = \mathrm{d}\psi \mathrm{d}\phi \mathrm{d}p_\psi \mathrm{d}p_\phi.$$

This is also the natural volume measure when we interpret T^*S as the internal phase space. In what follows, we will present the entire internal Hamiltonian description to show the way in which the measure arises in that description and to be able to prove that the measure is invariant under internal time evolution.

We have seen that, since the $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian gravitational system is invariant under translations and rotations, the dynamics can be formulated on reduced phase space $\Gamma_l = T^*S_R$. By construction, reduced phase space is a symplectic space and it is equipped with a unique symplectic two-form ω_l. With respect to the spherical coordinates defined in (5.41) and (5.42), the two-form ω_l is

$$\omega_l = \mathrm{d}R \wedge \mathrm{d}p_R + \mathrm{d}\psi \wedge \mathrm{d}p_\psi + \mathrm{d}\phi \wedge \mathrm{d}p_\phi.$$

This has already been shown before (cf. (5.45)).

Let now $i_H : \Sigma \to T^*S_R$ the embedding of Σ in T^*S_R where Σ is defined by the constraint $H = 0$ and let $i_D : \Gamma_D \to \Sigma$ the embedding of Γ_D in Σ where Γ_D is the hypersurface of constant $D = 0$. In order to determine the pullback of ω_l to the constraint surface Γ_D, let us rewrite ω_l with respect to D. Since $D = 2R \cdot p_R$, we have $\mathrm{d}p_R = \frac{1}{2R}\mathrm{d}D - \frac{D}{2R^2}\mathrm{d}R$ and

$$\begin{aligned} \omega_l &= \mathrm{d}R \wedge \left(\frac{1}{2R}\mathrm{d}D - \frac{D}{2R^2}\mathrm{d}R\right) + \mathrm{d}\psi \wedge \mathrm{d}p_\psi + \mathrm{d}\phi \wedge \mathrm{d}p_\phi \\ &= \frac{1}{2R}\mathrm{d}R \wedge \mathrm{d}D + \mathrm{d}\psi \wedge \mathrm{d}p_\psi + \mathrm{d}\phi \wedge \mathrm{d}p_\phi \\ &= \mathrm{d}\left(\frac{1}{2}\log R\right) \wedge \mathrm{d}D + \mathrm{d}\psi \wedge \mathrm{d}p_\psi + \mathrm{d}\phi \wedge \mathrm{d}p_\phi. \end{aligned} \tag{6.17}$$

Let us first determine the pullback $i_H^*\omega_l$ of ω_l onto the hypersurface of constant $H = 0$. With $F = 1/2 \log R|_{H=0}$ we can bring it into the well-know form:

$$i_H^*\omega_l = \mathrm{d}F \wedge \mathrm{d}D + \mathrm{d}\psi \wedge \mathrm{d}p_\psi + \mathrm{d}\phi \wedge \mathrm{d}p_\phi. \tag{6.18}$$

We already know that D is a suitable time parameter and F is its canonical conjugate. In accordance with (4.42), the physical vector field is determined by

$$i_H^*\omega_l(X_H, \cdot) = 0. \tag{6.19}$$

Explicitly,

$$X_H = \frac{\partial}{\partial D} + \frac{\partial F}{\partial p_\psi}\frac{\partial}{\partial \psi} - \frac{\partial F}{\partial \psi}\frac{\partial}{\partial p_\psi} + \frac{\partial F}{\partial p_\phi}\frac{\partial}{\partial \phi} - \frac{\partial F}{\partial \phi}\frac{\partial}{\partial p_\phi} \tag{6.20}$$

with $F = F_D(\psi, \phi, p_\psi, p_\phi)$ from (6.15).

The pullback of $i_H^*\omega_l$ onto the internal space Γ_D, that is the form $\omega := i_D^* i_H^*\omega_l$, is

$$\omega = \mathrm{d}\psi \wedge \mathrm{d}p_\psi + \mathrm{d}\phi \wedge \mathrm{d}p_\phi. \tag{6.21}$$

This two-form coincides with the symplectic form ω_d on dilationally reduced phase space T^*S

determined in (5.47). Hence, we see explicitly that $\Gamma_D \cong T^*S$ and Γ_D is a symplectic space. On that space, there exists an internal Hamiltonian vector field determined by

$$\omega(X_F, \cdot) = \mathrm{d}F \tag{6.22}$$

in accordance with (4.44). Explicitly,

$$X_F = \frac{\partial F}{\partial p_\psi}\frac{\partial}{\partial \psi} - \frac{\partial F}{\partial \psi}\frac{\partial}{\partial p_\psi} + \frac{\partial F}{\partial p_\phi}\frac{\partial}{\partial \phi} - \frac{\partial F}{\partial \phi}\frac{\partial}{\partial p_\phi}. \tag{6.23}$$

The physical trajectories on $\Gamma_D \cong T^*S$ are the integral curves along X_F parametrized by D. Using $\gamma'(D) = (X_F)_{\gamma(D)}$ you get back the Hamiltonian equations (6.16) from above.

Connected to the symplectic two-form ω on T^*S, there exists a volume measure $\mu = |\Omega|$ with $|\Omega| = |\omega|^2/2!$. Hence,

$$\mathrm{d}\mu = \mathrm{d}\psi \mathrm{d}\phi \mathrm{d}p_\psi \mathrm{d}p_\phi. \tag{6.24}$$

This is just the natural volume measure on T^*S (see (5.52)).

6.4.2 Invariance of the measure under time-evolution

The time-dependent Hamiltonian equations (6.16) determine a two-parameter flow $T_{\tau,\sigma}$ to describe the motion of the system on T^*S. For any point $p \in T^*S$ with $p = p(\sigma)$: $T_{\tau,\sigma}(p) = p(\tau)$. The volume measure μ derived from the volume form ω is invariant under the flow $T_{\tau,\sigma}$ along X_F.

Lemma 6.5 (Invariance under time evolution)**.** *Let $\mu = |\Omega|$ with $|\Omega| = |\omega|^2/2!$ and ω from (5.20). Let X_F given by (5.22) and $T_{\tau,\sigma}$ defined by $\frac{\mathrm{d}}{\mathrm{d}\tau}T_{\sigma,\sigma}(p) = X_F(p)$. Then*

$$L_{X_F}\omega = 0 \tag{6.25}$$

and

$$T^*_{\tau,\sigma}\mu = \mu. \tag{6.26}$$

Proof. The proof of this lemma is analogous to the proof of Lemma 4.7, respectively Corollary 4.1. That is, the Lie derivative vanishes because ω on Γ is closed ($\mathrm{d}\omega = 0$) and X_F is a Hamiltonian vector field, that is, $\omega(X_F, \cdot) = \mathrm{d}F$ (with $\mathrm{d} \circ \mathrm{d}F = 0$). From this we get

$$L_{X_F}\omega = (\mathrm{d}\omega)(X_F, \cdot, \cdot) + \mathrm{d}(\omega(X_F, \cdot)) = 0.$$

It follows that $L_{X_F}\Omega = 0$ and, in addition, $L_{X_F}\mu = 0$. From this we get that μ is transported invariantly by the flow: $\frac{\mathrm{d}}{\mathrm{d}t}(T^*_{\tau,\sigma}\mu) = L_{X_F}\mu = 0$ and thus $T^*_{\tau,\sigma}\mu = \mu$. □

Hence, $\mathrm{d}\mu = \mathrm{d}\psi \mathrm{d}p_\psi \mathrm{d}\phi \mathrm{d}p_\phi$ is conserved under internal time translation. To be precise, let $\mu(\tau)$ denote the volume of a region $A(\tau) \subset T^*\mathrm{S}$ at time τ and let $A(t) = T_{\tau,\sigma}A(\sigma)$. This last equation just asserts that every point $p(\sigma) \in A(\sigma)$ is transported by the Hamiltonian phase flow

$T_{\tau,\sigma}$ to another point $p(\tau) \in A(\tau)$. That is,

$$\mu(\tau) = \int_{A(\tau)} \mathrm{d}\psi \mathrm{d}p_\psi \mathrm{d}\phi \mathrm{d}p_\phi \tag{6.27}$$

and analogously $\mu(\sigma) = \int_{A(\sigma)} \mathrm{d}\psi \mathrm{d}p_\psi \mathrm{d}\phi \mathrm{d}p_\phi$ and we have shown that

$$\mu(\sigma) = \mu(\tau) \tag{6.28}$$

for all τ, σ. This is Liouville's theorem on T^*S.

6.4.3 The measure obtained from the Faddeev-Popov formula

Barbour, Koslowski, and Mercati [2015] do not refer to the symplectic structure of the internal space, but instead use the formula of Faddeev and Popov (5.63) for the computation of the measure on T^*S. They introduce this formula starting from the consideration that the space of solutions $\Gamma_{sol} \cong T^*S$ is determined by a set of second class constraints H_a, χ_b with

$$\{H_a, \chi_b\} \neq 0. \tag{6.29}$$

For these constraints, they compute the Faddeev-Popov measure

$$\mu_{red} = |\det\{H_a, \chi_b\}| \prod_{a,b} \prod_i \delta(H_a)\delta(\chi_b)\mathrm{d}q_i \mathrm{d}p_i. \tag{6.30}$$

In the paper of Barbour et al. there is no further justification for why this is the correct volume measure on T^*S nor do they show that it is conserved under internal time evolution. This is basically why I went through all of symplectic reduction and developed the internal Hamiltonian formulation on the symplectic internal phase space.

Lemma 6.6 (Faddeev-Popov measure on T^*S)**.** *Let $\boldsymbol{P}$, $\boldsymbol{Q}_{cm}$, $\boldsymbol{L}, \boldsymbol{I}_L$, H and D smooth functions on Γ defined in Section 4.2. Let $\boldsymbol{P} = \boldsymbol{Q}_{cm} = \boldsymbol{L} = \boldsymbol{I}_L = H = D = 0$. Then the Faddeev-Popov measure (4.64) on T^*S becomes*

$$\mathrm{d}\mu_{red} = |\det\{(\boldsymbol{P}, \boldsymbol{L}, H), (\boldsymbol{Q}_{cm}, \boldsymbol{I}_L, D)\}| \prod_i \delta(H)\delta(\boldsymbol{P})\delta(\boldsymbol{L})\delta(D)\delta(\boldsymbol{Q}_{cm})\delta(\boldsymbol{I}_L)\mathrm{d}q_i \mathrm{d}p_i. \tag{6.31}$$

Let $\psi, \phi, p_\psi, p_\phi$ local coordinates on T^*S defined in (4.42) and (4.43). Then

$$\mathrm{d}\mu_{red} = \mathrm{d}\psi \mathrm{d}\phi \mathrm{d}p_\psi \mathrm{d}p_\phi. \tag{6.32}$$

Proof. Since $\{\mathbf{P}, \mathbf{Q}_{cm}\} \neq 0$, $\{\mathbf{L}, \mathbf{I}_L\} \neq 0$, $\{H, D\} \neq 0$ and $\mathbf{P} = \mathbf{L} = H = 0$ we have a set of second class constraints in the sense of Dirac (cf. Section 4.3). Hence, the Faddeev-Popov formula for the measure (4.64) can be applied and we obtain (5.30).

Let us now compute the measure. $A := \{(\mathbf{P}, \mathbf{L}, H), (\mathbf{Q}_{cm}, \mathbf{I}_L, D)\}$ is as matrix

$$A = \begin{pmatrix} \{\mathbf{P}, \mathbf{Q}_{cm}\} & \{\mathbf{L}, \mathbf{Q}_{cm}\} & \{H, \mathbf{Q}_{cm}\} \\ \{\mathbf{P}, \mathbf{I}_L\} & \{\mathbf{L}, \mathbf{I}_L\} & \{H, \mathbf{I}_L\} \\ \{\mathbf{P}, D\} & \{\mathbf{L}, D\} & \{H, D\} \end{pmatrix}.$$

For the given system, this matrix can be simplified. On the constraint surface where $\mathbf{P} = \mathbf{L} = H = 0$, we have

$$\{\mathbf{P}, D\} = \{\mathbf{L}, D\} = 0.$$

Hence, we get

$$A = \begin{pmatrix} \{\mathbf{P}, \mathbf{Q}_{cm}\} & \{\mathbf{L}, \mathbf{Q}_{cm}\} & \{H, \mathbf{Q}_{cm}\} \\ \{\mathbf{P}, \mathbf{I}_L\} & \{\mathbf{L}, \mathbf{I}_L\} & \{H, \mathbf{I}_L\} \\ 0 & 0 & \{H, D\} \end{pmatrix}$$

and

$$|\det A| = |\det\{(\mathbf{P}, \mathbf{L}, H), (\mathbf{Q}_{cm}, \mathbf{I}_L, D)\}| = |\{\mathbf{P}, \mathbf{Q}_{cm}\}\{\mathbf{L}, \mathbf{I}_L\}\{H, D\}| = |\{H, D\}|. \tag{6.33}$$

Here the last equation holds since

$$\{\mathbf{P}, \mathbf{Q}_{cm}\} = \{\mathbf{L}, \mathbf{I}_L\} = 1.$$

Now Faddeev-Popov formula becomes very simple. We just need to do a coordinate transformation from the $\mathbf{q}_i, \mathbf{p}_i$ to the Jacobi and Hopf coordinates and their canonical conjugates introduced in (5.29), (5.30), (5.35) and (5.36). Hence, with (6.33) we have

$$\begin{aligned} \int \mathrm{d}\mu_{red} &= |\det\{(\mathbf{P}, \mathbf{L}, H), (\mathbf{Q}_{cm}, \mathbf{I}_L, D)\}| \int \prod_{i=1}^{3} \delta(H)\delta(D)\delta(\mathbf{P})\delta(\mathbf{Q}_{cm})\delta(\mathbf{L})\delta(\mathbf{I}_L)\mathrm{d}\mathbf{q}_i\mathrm{d}\mathbf{p}_i \\ &= |\{H, D\}| \int \delta(H)\delta(D)\delta(\boldsymbol{\kappa}_3)\delta(\boldsymbol{\rho}_3)\delta(L)\delta(I_L)\mathrm{d}\boldsymbol{\rho}_3\mathrm{d}\boldsymbol{\kappa}_3\mathrm{d}L\mathrm{d}I_L \cdot \mu_R \\ &= |\{H, D\}^*| \int \delta(H)\delta(D)\mathrm{d}\mu_R. \end{aligned} \tag{6.34}$$

Here $\mathrm{d}\mu_R = \mathrm{d}R\mathrm{d}\psi\mathrm{d}\phi\mathrm{d}p_R\mathrm{d}p_\psi\mathrm{d}p_\phi$ as defined in (4.52) and $\{H, D\}^*$ denotes the (reduced) Poisson bracket on T^*S_R.

Recalling (6.12), we have

$$\left|\{H, D\}_{\mathbf{P}=L=0}\right| = |\{H, D\}^*| = \left|2H - \frac{V_S}{\sqrt{R}}\right|$$

and (6.34) becomes

$$\int \mathrm{d}\mu_{red} = \int \left|2H - \frac{V_S}{\sqrt{R}}\right| \delta(H)\delta(D)\mathrm{d}\mu_R.$$

Using

$$\delta^{(n)}(\mathbf{f}(\mathbf{x})) = \frac{1}{|\det \partial \mathbf{f}_i/\partial \mathbf{x}_j|}\delta^{(n)}(\mathbf{x}-\mathbf{x}_0),$$

where the $\mathbf{x}_0$'s are the zeros of $\mathbf{f}(\mathbf{x})$. With $\mathbf{f} = (H, D)^T$, $\mathbf{x} = (R, p_R)^T$ and $\mathbf{x}_0 = (R^*, 0)^T$ solving $H = 0$, $D = 0$, it follows that

$$\delta(H)\delta(D) = \frac{1}{|\{H, D\}^*|}\delta(R-R^*)\delta(p_R - 0) = \left|2H - \frac{V_S}{\sqrt{R}}\right|^{-1}\delta(R-R^*)\delta(p_R).$$

Thus,

$$\int \mathrm{d}\mu_{red} = \int \left|2H - \frac{V_S}{\sqrt{R}}\right|\left|2H - \frac{V_S}{\sqrt{R}}\right|^{-1}\delta(R-R^*)\delta(p_R)\mathrm{d}R\mathrm{d}\psi\mathrm{d}\phi\mathrm{d}p_R\mathrm{d}p_\psi\mathrm{d}p_\phi$$

from which it follows that

$$\mathrm{d}\mu_{red} = \mathrm{d}\psi\mathrm{d}\phi\mathrm{d}p_\psi\mathrm{d}p_\phi.$$

□

In what follows, we will show that we obtain the same measure if in the Faddeev-Popov formula for the measure (6.31) we interchange H and D.

Corollary 6.2. *Let everything as in the preceding lemma, but let the volume measure on T^*S now be given by*

$$\mathrm{d}\mu'_{red} = |\det\{(\boldsymbol{P}, \boldsymbol{L}, D), (\boldsymbol{Q}_{cm}, \boldsymbol{I}_L, H)\}| \prod_i \delta(D)\delta(\boldsymbol{P})\delta(\boldsymbol{L})\delta(H)\delta(\boldsymbol{Q}_{cm})\delta(\boldsymbol{I}_L)\mathrm{d}q_i\mathrm{d}p_i \tag{6.35}$$

(this is (6.31) with H and D interchanged). It follows that

$$\mathrm{d}\mu'_{red} = \mathrm{d}\psi\mathrm{d}\phi\mathrm{d}p_\psi\mathrm{d}p_\phi. \tag{6.36}$$

Hence, $\mathrm{d}\mu'_{red} = \mathrm{d}\mu_{red}$ where $\mathrm{d}\mu_{red}$ is given by (5.31).

Proof. The only difference in the definitions of $\mathrm{d}\mu'_{red}$ (6.35) and $\mathrm{d}\mu_{red}$ (6.31) is in the Faddeev-Popov determinant. Hence, the result follows as soon as we have shown that on the constraint surface

$$|\det\{(\mathbf{P}, \mathbf{L}, D), (\mathbf{Q}_{cm}, \mathbf{I}_L, H)\}| = |\det\{(\mathbf{P}, \mathbf{L}, H), (\mathbf{Q}_{cm}, \mathbf{I}_L, D)\}|. \tag{6.37}$$

This we find by inspection of the matrix $B := \{(\mathbf{P}, \mathbf{L}, D), (\mathbf{Q}_{cm}, \mathbf{I}_L, H)\}$. Since $\mathbf{P}$ and $\mathbf{L}$ are conserved by the dynamics, we have

$$\{\mathbf{P}, H\} = \{\mathbf{L}, H\} = 0.$$

Hence, the determinant of B is again the product of the diagonal matrix elements,

$$|\det\{(\mathbf{P}, \mathbf{L}, D), (\mathbf{Q}_{cm}, \mathbf{I}_L, H)\}| = |\{H, D\}\{\mathbf{P}, \mathbf{Q}_{cm}\}\{\mathbf{L}, \mathbf{I}_L\}|$$

and as such equals the determinant of A (cf. Eq. (6.33)). This shows the assertion. $\square$

How do we interpret the fact that H and D can be interchanged without changing the measure ($\mathrm{d}\mu_{red} = \mathrm{d}\mu'_{red}$)? On the one hand, we group together $H, \mathbf{P}$, and $\mathbf{L}$ which are the conserved quantities of motion. They determine the hypersurface on which the trajectories lie. In this reading, while $\mathbf{Q}_{cm} = 0$ and $\mathbf{I}_L = 0$ are the gauge fixings, $D = 0$ is not a gauge-fixing, but it determines a hypersurface which is cut by the trajectories once and only once. This is the space of solutions Γ_{sol}. On the other hand, if we group together $\mathbf{P}, \mathbf{L}$ and D, we group together the generators of translations, rotations, and dilations. In that reading, $H, \mathbf{Q}_{cm}$ and $\mathbf{I}_L$ can be interpreted as the respective gauges. Hence, the space which is constructed this way is just reduced phase space T^*S (where reduction has been done with respect to the similarity group). It follows that we can identify the space of solutions with T^*S,

$$\Gamma_{sol} \cong T^*S. \tag{6.38}$$

6.5 Dynamical similarity on T^*S and a normalizable measure on PT^*S

In this section we identify one further redundant degree of freedom in the description of the Newtonian universe due to the dynamical similarity of the system. This will allow us to reduce the space of solutions T^*S by one further dimension and finally construct the *space of physically distinct solutions PT^*S*.

Dynamical similarity. Let us recall the form of the Hamiltonian equations of motion on T^*S (cf. (6.16)):

$$\begin{aligned}
\frac{\mathrm{d}\psi}{\mathrm{d}D} &= \frac{2p_\psi}{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + \frac{1}{4}D^2}, \qquad & \frac{\mathrm{d}p_\psi}{\mathrm{d}D} &= \frac{2\sin^{-3}\psi\cos\psi p_\phi^2}{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + \frac{1}{4}D^2} + \frac{\partial \log(-V_s)}{\partial\psi}, \\
\frac{\mathrm{d}\phi}{\mathrm{d}D} &= \frac{2\sin^{-2}\psi p_\phi}{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + \frac{1}{4}D^2}, \qquad & \frac{\mathrm{d}p_\phi}{\mathrm{d}D} &= \frac{\partial \log(-V_s)}{\partial\phi}
\end{aligned}$$

where $\mathrm{d}\psi/\mathrm{d}D = \mathrm{d}F/\mathrm{d}p_\psi$, $\mathrm{d}p_\psi/\mathrm{d}D = -\mathrm{d}F/\mathrm{d}\psi$ and analogously for ϕ, p_ϕ. These equations are invariant under the following transformation:

$$D \to kD, \quad \psi \to \psi, \quad \phi \to \phi, \quad p_\psi \to kp_\psi, \quad p_\phi \to kp_\phi \tag{6.39}$$

with $k \in \mathbb{R}\backslash\{0\}$ arbitrary. This property is called a mechanical or *dynamical similarity* of the system.[48] It defines an equivalence relation on the set of solutions.

Definition 6.2. Let the equations of motion on T^*S be given by (6.16). Then

$$(\psi, \phi, p_\psi, p_\phi)(D) \sim (\psi, \phi, kp_\psi, kp_\phi)(kD) \tag{6.40}$$

with $k \in \mathbb{R}\backslash\{0\}$ defines an equivalence relation $\sim$.

[48] Cf. Landau and Lifshitz [1967].

Hence, two solutions are equivalent if their solution-determining data are related according to (6.40). In what follows we want to identify all solutions that belong to the same equivalence class of solutions. This is reasonable because two solutions which are equivalent according to (6.40) determine two geometrically similar curves on internal phase space T^*S. In fact, they determine *one and the same geometrical curve* on S which is *run through at different speeds* and *with time adapted appropriately.* Since we deal with the universe as a whole and not with some subsystem for which there exists an external frame of reference, there is no way to distinguish these curves from within the universe/by observation. Adopting a relational conception not only of space, but also of time, we take them to actually refer to *one and the same physical solution* of the Newtonian model of the universe.

Note that, in contrast to before, the transformation (6.39) does not reflect a symmetry of the system. It is not a gauge transformation nor is there any corresponding conserved quantity of motion. Instead, it specifies the inter-dependence of time, position and momentum. More precisely, it determines how one can scale time, positions and momenta without changing the equations of motion. Such a dynamical similarity exists for every system with a potential which is homogeneous of a certain degree.[49]

Remark (Kepler orbits)**.** Of course, the invariance under the transformation (6.39) also applies to subsystems of the universe, like to the planets orbiting around the sun. Only then we do not identify dynamically similar curves because we have an external frame of reference (the frame of the fix stars) to hold the curves apart. In case of the planets of our solar system, dynamical similarity tells us that there exist different orbits, different in size and with different periods and velocities of revolution, but where the proportions of the radii and periods and velocities are the same: the proportions are exactly what is specified by the respective scaling behavior of the equations of motion. This way one may obtain the Kepler laws.[50]

Let us now finally get rid off all redundant degrees of freedom and specify the space of physically distinct solutions PT^*S by identifying dynamically similar curves. Note that, in general, the above equivalence relation identifies different points of T^*S at different times D and kD, respectively (cf. (6.40)). Thus, $D = 0$ plays a special role. For $D = 0$, the transformation property of the equations of motion allows us to identify different points of T^*S at one and the same moment of time:

$$(\psi, \phi, p_\psi, p_\phi)(0) \sim (\psi, \phi, kp_\psi, kp_\phi)(0). \tag{6.41}$$

In other words, at $D = 0$, we may identify all points which are connected by the transformation

$$\psi \to \psi, \qquad \phi \to \phi, \qquad p_\psi \to kp_\psi, \qquad p_\phi \to kp_\phi. \tag{6.42}$$

This transformation relates dynamically similar solutions on the surface of mid-point data.

The $D = 0$ hypersurface – the surface of "initial" or better mid-point conditions – has been

[49]Cf. Landau and Lifshitz [1967]. See also the Appendix C.

[50]Cf. Landau and Lifshitz [1967] for more details.

called the *Janus surface* by Barbour et al. [2013]. It is a representative of the space of solutions Γ_{sol}. That is, each point corresponds to one physical trajectory, respectively one solution. Dynamical similarity gives us one remaining redundancy, one further dimension we can get rid off by identifying all those points which are related by (6.41). That way we end up with the space of *physically distinct solutions*, respectively *physically distinct mid-point data*: PT^*S.

A measure on PT^*S. To construct the space of physically distinct solutions PT^*S we express the shape momenta p_ψ, p_ϕ in terms of polar coordinates $R_p \sim \sqrt{p_\psi^2 + p_\phi^2}, \chi_p \sim p_\psi/p_\phi$ and identify all points with the same χ_p in accordance with (6.41). Let $\mathbf{p}_S = (p_\psi, p_\phi)^T$ denote the shape momentum vector. From (6.42) we know that two curves are dynamically similar if and only if, at $D = 0$, their shape coordinates ψ and ϕ and the orientation of the shape momentum vector $\chi_p \sim p_\psi/p_\phi$ coincide. The three "initial" data (which are rather mid-point data specified at $D = 0$) $\psi_0 = \psi(0)$, $\phi_0 = \phi(0)$, and $[p_\psi/p_\phi]_0 = [p_\psi/p_\phi](0)$ completely determine the solutions. In other words, the dynamics is invariant under scaling of the radial component of the shape momentum vector, $R_p \to kR_p$ with $k \in \mathbb{R}^+$, at $D = 0$. Hence, we arrive at the space of physically distinct solutions PT^*S if and only if we fix the radial component R_p to some constant value: $R_p = c$ with $c \in \mathbb{R}^+$. This way we see that the space of physically distinct solutions PT^*S is actually a projective vector bundle (a bundle of the projective cotangent spaces where a projective space is just the set of lines through the origin).

To construct PT^*S, we make use of a metric. If we start with a metric $\mathrm{d}s^2$ on configuration space Q which is invariant under translations, rotations and dilations, this induces a quotient metric $\mathrm{d}s_s^2$ on shape space S.[51] This is called the Riemannian quotient. It is the quotient of the metric $\mathrm{d}s^2$ by the similarity group $\mathrm{Sim}(3)$ which, in this case, is the metric's isometry group. Hence, to begin with we need a metric on Q which is invariant under translations, rotations, and dilations. One simple choice is the mass metric $\sum_i m_i \mathrm{d}\mathbf{q}_i \cdot \mathrm{d}\mathbf{q}_i$ divided by the center-of-mass moment of inertia $I = \sum_i m_i \mathbf{q}_i^2$. This is the choice of Barbour et al. [2015]. Of course, any choice is possible which renders the metric invariant under $\mathrm{Sim}(3)$.

Lemma 6.7. *Consider shape space $S = Q/\mathrm{Sim}(3)$. Let $\boldsymbol{q}_i$ with $i = 1, 2, 3$ be local coordinates on Q and ψ, ϕ be local coordinates on S given by (5.41). Let $I = \sum_i m_i \boldsymbol{q}_i^2$ and*

$$\mathrm{d}s^2 = \sum_i \frac{m_i \mathrm{d}\boldsymbol{q}_i \cdot \mathrm{d}\boldsymbol{q}_i}{I} \tag{6.43}$$

the (conformal) metric on Q. Then

$$\mathrm{d}s_S^2 = \mathrm{d}\psi^2 + \sin^2\psi \mathrm{d}\phi^2 \tag{6.44}$$

is the Riemannian quotient of $\mathrm{d}s^2$ *with respect to* $\mathrm{Sim}(3)$*. It is a metric on S.*

Proof. This result can be obtained directly via the coordinate transformations from Section 5.2 where ψ and ϕ are introduced (cf. (5.41)). Just express the $\mathrm{d}\mathbf{q}_i$ and I in terms of the new

[51] Cf. Montgomery [2002].

coordinates and apply the constraints, then you find the quotient metric $\mathrm{d}s_S^2$. □

From the norm of the shape vectors $\mathbf{s}_S = (\psi, \phi)^T$ expressed in terms of the metric we get the norm of the cotangent vectors (shape momenta) $\mathbf{p}_S = (p_\psi, p_\phi)^T$ via the inverse metric.

Lemma 6.8. *Let ψ, ϕ be local coordinates on S and $\psi, \phi, p_\psi, p_\phi$ be local coordinates on T^*S given by (5.41) and (5.42). Let $\mathrm{d}s_S^2$ given by (6.42) the metric on S. Then*

$$\mathrm{d}p_S^2 = \mathrm{d}p_\psi^2 + \sin^{-2}\psi \mathrm{d}p_\phi^2 \tag{6.45}$$

is the metric on the cotangent space.

Proof. This form of $\mathrm{d}p_S^2$ follows directly from the fact that it need to be the inverse of the metric $\mathrm{d}s_S^2$ given by (6.44). □

By help of the metric (6.45) on the cotangent space, we can identify the shape momenta which differ in their length, but not in their orientation. For that purpose, let me introduce polar coordinates. For a given ψ, we have

$$R_p = \sqrt{p_\psi^2 + \sin^{-2}\psi p_\phi^2}, \qquad \chi_p = \arctan\frac{\sin^{-1}\psi p_\phi}{p_\psi}. \tag{6.46}$$

We can now identify the shape momenta which differ in their length R_p, but not in their orientation χ_p. This we do by help of a delta function, gauge fixing R_P. That way we can construct a measure ε on PT^*S which is the restriction of μ on T^*S to PT^*S.

Lemma 6.9. *Let μ be the volume measure on T^*S with $\mathrm{d}\mu$ given by (6.24) and let R_p given by (6.46) the length of the shape momentum vector. Let ε be a measure with*

$$\mathrm{d}\varepsilon = \delta(R_p - 1)\cdot \mathrm{d}\mu. \tag{6.47}$$

Then ε is a volume measure on PT^*S and

$$\mathrm{d}\varepsilon = \sin\psi \mathrm{d}\psi \mathrm{d}\phi \mathrm{d}\chi. \tag{6.48}$$

Proof. With R_p given by (6.46) and $\mathrm{d}\mu = \mathrm{d}\psi\mathrm{d}\phi\mathrm{d}p_\psi\mathrm{d}p_\phi$, it is

$$\begin{aligned}\int \mathrm{d}\varepsilon &= \int \delta(R_p - 1)\mathrm{d}\mu = \int \delta(R_p - 1)\sin\psi\mathrm{d}\psi\mathrm{d}\phi\mathrm{d}p_\psi\mathrm{d}p_\phi' \\ &= \int \delta(R_p - 1) R_p \sin\psi \mathrm{d}\psi\mathrm{d}\phi\mathrm{d}R_p\mathrm{d}\chi_p \\ &= \int \sin\psi\mathrm{d}\psi\mathrm{d}\phi\mathrm{d}\chi_p\end{aligned}$$

where we substituted $p_\phi' = \sin^{-1}\psi p_\phi$ and used that $\mathrm{d}p_\psi\mathrm{d}p_\phi' = R_p\mathrm{d}R_p\mathrm{d}\chi_p$. The condition $R_p = 1$ ensures the this is a measure on PT^*S. □

Corollary 6.3. *Let ε be the volume measure on PT^*S (with $\mathrm{d}\varepsilon$ from (6.48)). Then*

$$\sigma_\varepsilon = \frac{\varepsilon}{\varepsilon(PT^*S)} \tag{6.49}$$

is the normalized volume measure on PT^*S with

$$\mathrm{d}\sigma_\varepsilon = \frac{\sin\psi \mathrm{d}\psi \mathrm{d}\phi \mathrm{d}\chi_p}{8\pi^2}. \tag{6.50}$$

Proof. The normalization is obtained by integrating $\mathrm{d}\varepsilon$ over $\psi \in [0,\pi], \phi \in [0, 2\pi[, \chi_p \in [0, 2\pi[$. □

In what follows σ_ε will be the measure with respect to which we statistically analyze the $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian universe.

Why is ε (respectively σ_ε) the correct measure? The reason is that it is *uniform*. It is the uniform volume measure on the set of physically distinct solutions, respectively (physically distinct) mid-point data PT^*S. With respect to this measure, each (physically distinct) solution has the same weight. This is *Laplace's principle*. To adopt Laplace's principle is the most reasonable thing we can do as long as we have nothing like stationarity to prefer one measure over the other. And we don't have stationarity here because we don't have a time-evolved measure in the first place. All we have is a measure over mid-point data, that is, a measure defined at one particular moment in time. And this measure is uniform. Moreover, it is normalizable. As such it allows us to (unambiguously) statistically analyze the Newtonian universe at $D = 0$!

Remark (Lack of stationarity)**.** When we reduce the description from T^*S to PT^*S we loose the invariance of the measure under time evolution. Of course, trajectories can still be projected onto PT^*S, but we do no longer have Hamiltonian equations guiding the motion and phase space volume is lost as the system evolves away from the Janus point.

Intuitively, this can be seen as follows. For three particles, the generic evolution is such that as $t \to \pm\infty$ a more and more perfect Kepler pair forms with the third particle receding from it. At the same time, the two angles ψ and ϕ approach certain fix values ψ_{ij} and ϕ_{ij} specifying the three binary collision points $\mathbf{b}_{ij}$ (with $i < j; i, j = 1, 2, 3$) in S. These three points $\mathbf{b}_{ij}$ represent the three possibilities to form a Kepler pair with a single particle away from it. Now consider some arbitrary region A in shape space S. Almost any trajectory which is in A at time $D = 0$ will at some later time $D = D^*$ be close to one of the binary collision points (approaching $\mathbf{b}_{ij}$ as $D \to \infty$).

On T^*S, this loss of phase space volume in the configurational part of the measure is equilibrated by a gain of volume in the momentum part (remember that, on T^*S, Liouville's theorem holds). On PT^*S, this is no longer possible since we fixed the absolute value of the shape momentum vector to one. Hence, on PT^*S, we face a total loss of phase space volume as the system evolves away from the Janus point. The non-stationarity of the volume measure is proven in Section 7.3.

7 Complexity, entaxy and entropy

7.1 Complexity C_S as a macrovariable

In oder to perform the statistical analysis of the Newtonian universe on the space of mid-point data PT^*S we need to distinguish different macrostates. We want to find out whether a certain macroscopic property is typical or not. That is, we need a macrovariable (or set of macrovariables) which allows us to distinguish between different macrostates[52] of the universe. These macrostates define macro-regions – the sets of all points realizing the respective macroscopic property (or macroscopic properties) – on PT^*S. In order to partition PT^*S, we need a macrovariable which distinguishes between different shapes of the system.

Barbour, Koslowski, and Mercati [2015] suggest the complexity C_S as a meaningful macrovariable for the Newtonian model of the universe.[53]

Definition 7.1 (Complexity)**.** Let V_N be the Newton potential given by (6.1). Let $I = \sum_i m_i \mathbf{q}_i^2$ be the center-of-mass moment of inertia. Then

$$C_S = -V_N \cdot \sqrt{I} \tag{7.1}$$

is the *complexity* of the system.

It follows from this definition and the definition of the shape potential $V_S = V_N \sqrt{||\mathbf{w}||}$ with $||\mathbf{w}|| = I$ (cf. (6.3) and (5.50)) that C_S on PT^*S is just minus the scale potential:

$$C_S = -V_S. \tag{7.2}$$

Let us write down the explicit form of C_S for a system of three particles. Expressed in terms of the internal coordinates $\psi, \phi, p_\psi, p_\phi$ on T^*S introduced in (5.41) and (5.42), the complexity C_S can be written as

$$C_S = \sum_{i<j} \frac{(m_i m_j)^{\frac{3}{2}} (m_i + m_j)^{-\frac{1}{2}}}{\sqrt{1 - \sin\psi \cos(\phi - \phi_{ij})}}. \tag{7.3}$$

This follows from (6.7) with $C_S = -V_S$.

The quantity C_S is an interesting macrovariable for the Newtonian N-particle system because it distinguishes between homogenous and inhomogenous states, respectively, between states in which the particles are at approximately equal distance from each and states of clusters. This can be seen as follows. Let $i, j = 1, ..., N$. Let

$$R = \max_{i \neq j} |\mathbf{q}_i - \mathbf{q}_j| \tag{7.4}$$

[52] For the notion of a macrovariable, cf. Dürr and Teufel [2009]. In thermodynamics, macrovariables are volume V, temperature T, pressure p, and so on.

[53] The quantity C_S has to my knowledge first been introduced by Saari in the context of central configurations where it has been called the *configurational measure*. This is because it “measures” the actual “configuration” – what we call shape – of the multi-particle system. Cf. the paper of Saari on central configurations.

denote the largest distance between the particles and

$$r = \min_{i \neq j} |\mathbf{q}_i - \mathbf{q}_j| \tag{7.5}$$

the smallest distance. Then C_S is approximately equal to the largest distance R divided by the smallest distance r of the system.

Lemma 7.1. *Let C_S from (7.1) be the complexity. Let R and r be as defined above. Then there exist positive constants α, β such that*

$$\alpha \frac{R}{r} \leq C_S \leq \beta \frac{R}{r}. \tag{7.6}$$

In that case, we write

$$C_S \approx \frac{R}{r}. \tag{7.7}$$

Proof. The square-root of the moment of inertia of the system $\sqrt{I}$ serves as a measure of the maximal distance R between any two particles. Explicitly, there exist positive constants c_1, c_2 such that

$$c_1 R \leq \sqrt{I} \leq c_2 R.$$

On the other hand, the reciprocal value of minus the Newton potential $1/(-V_N)$ serves as a measure of the minimum distance r. There exist positive constants c_1', c_2' such that

$$c_1' r \leq \frac{1}{-V_N} \leq c_2' r.$$

Combining both inequalities and setting $\alpha = c_1/c_2'$ and $\beta = c_2/c_1'$, Eq. (7.5) follows. From (7.5) we conclude that C_S is approximately equal to R/r and write $C_S \approx R/r$. □

It follows from the definition of the complexity that there exists no upper bound on C_S. On the other hand, for every given particle number N, there exists a minimal value $C_{min} = C_{min}^N$, where the complexity C_S is minimal if and only if the distances between the particles are equal (or "as equal as possible" for there need not exist a configuration of equal distances). In that case, $R \approx r$ and $C_S \approx 1$. In contrast, C_S grows without bound as the proportion of the maximal distance R to the minimal distance r increases.

For a system of three particles, the minimum $C_S = C_{min}$ is attained if and only if the particles form an equilateral triangle, while C_S grows without bound as two particles form a more and more perfect Kepler pair with one particle receding from the other two.

For a system of N particles, C_S tells us something about the homogeneity of the configuration. If the particles are more or less homogeneously distributed, then C_S is close to the minimum: $C_S \approx C_{min}$. In contrast, C_S grows without bound (with small fluctuations) as the particles cluster (galaxies form) and the clusters recede from each other.

The results of Saari [1971] on the final evolution of the $E = 0$ Newtonian universe give us an estimate on the behavior of C_S for $t \to \pm\infty$.

Lemma 7.2. *Consider a generic evolution of the N-particle universe. That is, as $t \to \pm\infty$, $I(t) \sim t^2$ and $-V_N(t) \sim 1$ (cf. (3.9) and (3.10)). Then, as $t \to \pm\infty$,*

$$C_S(t) \sim |t|. \tag{7.8}$$

Proof. This follows directly from the assumptions and the definition of C_S (cf. (7.1)). $\square$

Moreover, let us again assume that $I(t)$ is well approximated by $I(t) = \alpha(t-\tau)^2 + \beta$ and that the Newton potential V_N is suitably well-behaved (like we did in Sec. 3.1). Hence, we exclude point-particle collisions and "near point-particle collisions" (close encounters of particles). Then we find that the behavior of C_S is governed by the behavior of I. As such it is well approximated by

$$C_S(t) = \gamma|t - \tau| + \delta \tag{7.9}$$

in analogy to (3.12). Here γ and δ are positive constants and τ is the moment at which the moment of inertia is minimal: $I(\tau) = I_{min}$.

From an analysis of the dynamics we obtain that, at the Janus point ($t = \tau$), C_S is minimal while it grows without bound in both directions away from it. Still, C_S can be arbitrarily high at the Janus point. In the next section, we will show that, at $t = \tau$, C_S is typically close to its absolute minimum C_{min}.

Again, the numerical results of Barbour et al. (2013) and (2015) for $N = 1000$ particles and random initial data show that the C_S-curve is well approximated by (7.9). That is, C_S features a distinct minimum at $t = \tau$ ($D = 0$) and it increases with t with small fluctuations in both time directions away from the minimum.[54]

7.2 Solution entaxy ε_{sol}

We can now define the solution entaxy ε_{sol} which determines the volume of sets of constant complexity C_S on PT^*S (sets of constant complexity at the Janus point $D = 0$).

Definition 7.2 (Solution entaxy). Let $C_S = C_S(\psi, \phi)$ from (7.7) be the complexity. Let ψ, ϕ, and χ_p be local coordinates on PT^*S defined in (5.41) and (6.44). Then

$$\varepsilon_{sol}(C^*) = \int_{PT^*S} \delta(C_S(\psi, \phi) - C^*) \sin\psi \mathrm{d}\psi \mathrm{d}\phi \mathrm{d}\chi_p \tag{7.10}$$

is the *solution entaxy.*

Consider the volume measure ε on PT^*S with $\mathrm{d}\varepsilon = \sin\psi \mathrm{d}\psi \mathrm{d}\phi \mathrm{d}\chi$ from (6.48). Let $\Gamma_{C^*} = \{(\phi, \psi, \chi_p) \in PT^*S | C_S(\psi, \phi) = C^*\}$. It follows that the solution entaxy $\varepsilon_{sol}(C^*)$ is simply the volume of the region of constant complexity,

$$\varepsilon_{sol}(C^*) = \varepsilon(\Gamma_{C^*}). \tag{7.11}$$

[54] Cf. the discussion in Sec. 3.1 and Barbour et al. [2013] and [2015].

While the macro-variable C_S defines a macro-partition of PT^*S (a partition of PT^*S into regions/sets of constant complexity C_S), the solution entaxy ε_{sol} determines the volume of the respective macro-regions (the sets of all points on PT^*S realizing a certain macrostate $C_S = C^*$). In that sense, it is an entropy-type quantity. However, it is *not* the relational (scale-invariant) analogue of the Boltzmann entropy. This will be explained in Section 7.3. Still, the solution entaxy determines the typical (respectively, atypical) values of the complexity C_S of the $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian universe at $D = 0$.[55]

When we compare different macrostates of the $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian universe, we are not so much interested in particular values of C_S, but rather in a range of values $C_S \in [C_1, C_2]$ and the measure of the respective region. Let $\Gamma_{[C_1,C_2]} = \{(\psi, \phi, \chi_p) \in PT^*S | C_S(\psi, \phi) \in [C_1, C_2]\}$. While a fix value $C_S = C^*$ determines a two-dimensional subset of PT^*S, a range of values $C_S \in [C_1, C_2]$ determines a three-dimensional subset/region of PT^*S. In analogy to (7.10), the solution entaxy of a range of values $C_S \in [C_1, C_2]$ is

$$\varepsilon_{sol}([C_1, C_2]) = \varepsilon(\Gamma_{[C_1,C_2]}) = \int_{PT^*S} \mathbb{I}_{\{C_S \in [C_1,C_2]\}} \sin\psi \mathrm{d}\psi \mathrm{d}\phi \mathrm{d}\chi_p. \tag{7.12}$$

In what follows, we are interested in those configurations which have a low complexity – corresponding to a more or less homogenous state of the universe – as compared to those configurations which have a high complexity – corresponding to a dilute state of clusters.

Definition 7.3. Let $1 << \alpha < \infty$. Let

$$I_1 = [C_{min}, \alpha \cdot C_{min}] \tag{7.13}$$

and

$$I_\infty =]\alpha \cdot C_{min}, \infty[\tag{7.14}$$

Then I_1 refers to the *homogeneous states* of the universe and I_∞ to the *inhomogenous states.*

The value of α determines which sets of configurations we consider as macroscopically distinct. Note that the exact value of α is not important, important is just the fact that, for a given α, the intervals I_1 and I_∞ represent macroscopically distinct states. As it turns out I_1 and I_∞ determine regions in PT^*S that differ vastly in size – just like it is the case when we choose a particular macro-partition in classical Boltzmannian statistical mechanics.

For a system of three particles, I_1 represents configurations similar to an equilateral triangle while I_∞ represents configurations similar to a Kepler pair with one particle far away from it. For N particles, low complexity $C_S \in I_1$ refers to a more or less homogeneous distribution while high complexity $C_S \in I_\infty$ refers to a dilute states of clusters where some particles are close to each other forming clusters while the distances between clusters are large.

[55] By the way note that the solution entaxy is essentially itself a measure. Still we want to distinguish between the measure and the entaxy where the latter is a quantity which is defined with respect to a macrovariable and a measure. If we would not make this distinction, it is as if we would identify the entropy and the microcanonical measure.

Let us determine the solution entaxy of I_1 and I_∞, respectively. Let us restrict the attention to a system of three particles of unit masses ($m_i = 1$ with $i = 1, 2, 3$). Let $\varepsilon(I_1) := \varepsilon(\Gamma_{I_1})$ and $\varepsilon(I_\infty) := \varepsilon(\Gamma_{I_\infty})$ and analogously for σ_ε.

Lemma 7.3. *Let everything be as in the preceding paragraph. Let I_1 as defined in (7.13) and I_∞ as defined in (7.14). Then*

$$\sigma_\varepsilon(I_\infty) = \frac{\varepsilon(I_\infty)}{\varepsilon(PT^*S)} \to 0 \quad \text{as} \quad \alpha \to \infty, \tag{7.15}$$

respectively,

$$\sigma_\varepsilon(I_1) = \frac{\varepsilon(I_1)}{\varepsilon(PT^*S)} \to 1 \quad \text{as} \quad \alpha \to \infty. \tag{7.16}$$

Proof. We can find the solution entaxy of a certain interval $C_S \in [C_1, C_2]$ by help of the contour lines of C_S (lines of constant C_S). Let ψ, ϕ defined in (5.41) be local coordinates of the shape sphere S^2. In terms of these coordinates, C_S is given by (7.7). With $m_i = 1\ \forall i = 1, 2, 3$ we have

$$C_S = \frac{1}{\sqrt{2}} \sum_{i<j} \frac{1}{\sqrt{1 - \sin\psi \cos(\phi - \phi_{ij})}}$$

where the ϕ_{ij} with $i < j$ specify the three binary collision points. Now the contour lines of C_S can be determined by symmetry considerations. At the top and bottom of the shape sphere (representing the two equilateral triangles) C_S is minimal ($C_S = C_{min}$); there $\psi = 0$, respectively $\psi = \pi$, hence $C_S = 3/\sqrt{2}$.

In contrast, C_S is infinite at the three binary collision points which lie at equal distance from each other on the equator of the shape sphere; there $\psi = \pi/2$ and $\phi = \phi_{ij}$ for one of the ϕ_{ij}, hence $\sin\psi\cos(\phi - \phi_{ij}) = 1$ for one of the ϕ_{ij} and, consequently, $C_S = \infty$.

In addition, there exist three saddle points of C_S at the three Euler configurations which lie on the equator with one Euler configuration centered between two binary collision points; there $\psi = \pi/2$ and $\phi - \phi_{ij} = \{-\pi/6, \pi/6, \pi\}$, hence

$$C_S = \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{1 - \cos(-\pi/6)}} + \frac{1}{\sqrt{1 - \cos(\pi/6)}} + \frac{1}{\sqrt{1 - \cos(\pi)}} \right) \approx 6/\sqrt{2}.$$

Now everything is symmetric. Hence, the contour lines of C_S are circles around the binary collision points and circles around the top and bottom of the shape sphere. In between, the contour lines circling the top and bottom get more and more deformed, bending towards the Euler configurations the closer they get to the equator, finally following the shape of the three outmost circles around the binary collision points.

At the Euler configurations we have $C_S \approx 2\ C_{min}$. Let, therefore, $2 < \alpha < \infty$. Then $C_S \in I_\infty$ determines a region on S^2 which consists of three spherical caps around the binary collision points. Let K_α denote the surface area of one such cap. The larger α, the smaller the caps and vice versa. Hence, the solution entaxy is

$$\begin{aligned}\varepsilon(I_\infty) &= \int_{[0,2\pi[\times I_\infty} \sin\psi \mathrm{d}\psi \mathrm{d}\phi \mathrm{d}\chi_p \\ &= 2\pi \cdot \int_{I_\infty} \sin\psi \mathrm{d}\psi \mathrm{d}\phi = 2\pi \cdot 3K_\alpha.\end{aligned}$$

In contrast,

$$\begin{aligned}\varepsilon(I_1) &= \int_{[0,2\pi[\times I_1} \sin\psi \mathrm{d}\psi \mathrm{d}\phi \mathrm{d}\chi_p \\ &= \int_{PT^*S} \sin\psi \mathrm{d}\psi \mathrm{d}\phi \mathrm{d}\chi_p - \int_{[0,2\pi[\times I_\infty} \sin\psi \mathrm{d}\psi \mathrm{d}\phi \mathrm{d}\chi_p \\ &= 8\pi^2 - 2\pi \cdot 3K_\alpha.\end{aligned}$$

As α increases the caps become smaller approaching the binary collision points $\mathbf{b}_{ij}$ $(i < j; i, j = 1, 2, 3)$ as $\alpha \to \infty$. Since these are merely three points in S^2, $K_\alpha \to 0$ as $\alpha \to \infty$. It follows that the normalized solution entaxy σ_ε of regions I_∞ and I_1 is

$$\sigma_\varepsilon(I_\infty) = \frac{1}{\varepsilon(PT^*S)} \int_{[0,2\pi[\times I_\infty} \sin\psi \mathrm{d}\psi \mathrm{d}\phi \mathrm{d}\chi_p = \frac{3K_\alpha}{4\pi} \to 0 \quad \text{as} \quad \alpha \to \infty,$$

respectively,

$$\sigma_\varepsilon(I_1) = \frac{1}{\varepsilon(PT^*S)} \int_{[0,2\pi[\times I_1} \sin\psi \mathrm{d}\psi \mathrm{d}\phi \mathrm{d}\chi_p = 1 - \frac{3K_\alpha}{4\pi} \to 1 \quad \text{as} \quad \alpha \to \infty.$$

□

It can be computed that already for $\alpha = 5$ the proportion of I_∞ to PT^*S is small, $\sigma_\varepsilon(I_\infty) \approx 0.1$, while the proportion of I_1 to PT^*S is large: $\sigma_\varepsilon(I_1) = 1 - \sigma_\varepsilon(I_\infty) \approx 0.9$. Now $\alpha = 5$ still refers to a quite homogeneous distribution. There the maximal distance between the particles is about five times the minimal distance: $R \approx 5r$. Hence, we conclude that at the Janus point the three-particle universe is most likely to be in a state of low complexity, corresponding to what we call a homogeneous state!

The fact that in the given example $(\alpha = 5, N = 3)$ we don't find values of $\sigma_\varepsilon(I_\infty)$ closer to 0 and $\sigma_\varepsilon(I_1)$ closer to 1 has to do with the small number of particles, $N = 3$. The difference in volume becomes more pronounced as soon as we consider a greater number of particles. This follows both from numerical computations[56] as well as from pure reasoning: for N particles, shape space is of much higher dimension $(3N - 7)$. This dimension enters exponentially in the computation of the volume. It follows that, for N particles, I_1 and I_∞ specify regions that differ vastly in size. In that case, we can make a proper typicality statement concluding that at the Janus point/Big Bang homogenous distributions are typical while inhomogeneous, dilute states of clusters are atypical!

[56] Cf. Barbour et al. [2015] for numerical computations on the proportion of sets of constant complexity for more than three particles.

Last but not least, be aware that this result does not depend on the particular choice of the normalized volume measure σ_ε on PT^*S. In that sense, σ_ε is really a typicality measure on PT^*S (cf. Sec. 2.1).

Corollary 7.1. *Let everything be as in Lemma 7.3. Let σ_ε from (6.49) be the normalized volume measure on PT^*S and let ν_ε be another volume measure on PT^*S absolutely continuous with respect to σ_ε. Then*

$$\nu_\varepsilon(I_\infty) \to 0 \quad \text{and} \quad \nu_\varepsilon(I_1) \to 1 \quad \text{as } \alpha \to \infty. \tag{7.17}$$

Proof. Since the measure ν_ε is absolutely continuous with respect to σ_ε, it follows that there exists an $f \in L^1(PT^*S, \sigma_\varepsilon)$ such that, for every measurable set $A \subset PT^*S$, ν_ε can be written as

$$\nu_\varepsilon(A) = \int_A f \cdot \mathrm{d}\sigma_\varepsilon. \tag{7.18}$$

It follows that

$$\nu_\varepsilon(I_\infty) = \int_{I_\infty} f \cdot \mathrm{d}\sigma_\varepsilon \leq ||f||_\infty \cdot \sigma_\varepsilon(I_\infty) \to 0 \quad \text{as } \alpha \to \infty.$$

Since ν_ε is normalized, i.e. $\nu_\varepsilon(PT^*S) = 1$, and with $I_1 = PT^*S \backslash I_\infty$ it follows that

$$\nu_\varepsilon(I_1) = 1 - \nu_\varepsilon(I_\infty) \to 1 \quad \text{as } \alpha \to \infty.$$

□

7.3 Entaxy, complexity and entropy

This section is meant to clarify the relation between the notions of entaxy, complexity and entropy. We will show that there does not exist a shape analogue of the Boltzmann entropy. Instead, we need absolute quantities (in particular, absolute size) to define the entropy of the Newtonian universe – bringing us back to the notion of entropy presented in Section 2.

Entaxy and entropy. Barbour, Koslowski, and Mercati [2015] introduce the notion of an entaxy at time D. This implicitly suggests to read the entaxy as the shape analogue of the Boltzmann entropy. However, this analogy is not correct. The main reason is that the entaxy is not defined with respect to a stationary measure. It only serves for the statistical analysis of the universe at one moment in time.

When Barbour et al. [2015] introduce the notion of an entaxy at time D, they simply take the definition of the solution entaxy ε_{sol} (cf. (7.10)+(7.11)) and insert in that definition the shape of the system at time D. Let $\psi(D)$ and $\phi(D)$ be solutions to the Hamiltonian equations of motion on T^*S (cf. (6.16)). Then they define the entaxy ε^D at time D as

$$\varepsilon^D(\Gamma_{C^*}) = \int_{PT^*S} \delta(C_S(\psi_D, \phi_D) - C^*) \sin\psi \mathrm{d}\psi \mathrm{d}\phi \mathrm{d}\chi_p \tag{7.19}$$

with $\psi_D = \psi(D)$ and $\phi_D = \phi(D)$.

This formula does not involve a stationary measure. Instead, defined this way, the entaxy decreases (respectively, increases) because phase space volume is lost (respectively, gained) during time evolution.

Let me be more explicit. Let ε be the volume measure on PT^*S with $\mathrm{d}\varepsilon$ given in (6.48). Contrary to what the index D suggests, the measure ε^D defined in (7.19) – determining the entaxy at time D – is not the time-evolved measure since it is not the (original) measure of the pre-image of the given set (cf. Def. 2.1). Explicitly,

$$\varepsilon^D(\Gamma_{C^*}) \neq \varepsilon(T^{-D}\Gamma_{C^*}) \tag{7.20}$$

where T^D is the projection of the Hamiltonian flow on T^*S onto PT^*S.

Let in what follows ε_D denote the time-evolved volume measure. Let us define ε_D not with respect to the pre-image $T^{-D}A$, but (which is equivalent) with respect to the time-evolved set T^DA. That is, for $A \subset PT^*S$ and $D \in \mathbb{R}$, we define

$$\varepsilon_D(T^DA) := \varepsilon(A). \tag{7.21}$$

Clearly, the flow T^D is not Hamiltonian and we expect that, as the system evolves away from the Janus point and continuously approaches the binary collision points in shape space, phase space volume is lost (cf. the remark in Section 6.5). Hence, we expect that

$$\varepsilon(T^DA) < \varepsilon(A), \tag{7.22}$$

respectively $\varepsilon(T^DA) < \varepsilon_D(T^DA)$. In other words, we expect that ε is not stationary (for ε is stationary if and only if $\varepsilon(B) = \varepsilon_D(B)$ for every $B \subset PT^*S$). The following lemma will make this precise.

When we say that the following lemma shows non-stationarity, this has to be taken with a grain of salt. We will show that the measure ε is not stationary assuming that the time-evolution of points on PT^*S is generic. In reality, there will be some points which evolve differently. If they form a set of measure zero, there is no problem at all. If they form a small set, we should in order to rigorously prove non-stationarity give precise bounds on the size of this set. To keep things simple, we don't do that here, but simply assume that the set of points that evolve non-generically is small enough in order to not disturb the result (which is a very reasonable assumption).

Thus let us consider a generic evolution of the $E = 0$ universe. That is, as $t \to \pm\infty$, $I(t) \sim t^2$ as in (3.9) and $C_S(t) \sim |t|$ as in (7.8). Moreover, I is concave upwards, $\ddot{I} > 0$, and I has a minimum at $t = \tau$: $I(\tau) = I_{min}$. Let us therefore, like in (3.12), assume that

$$I(t) = \alpha(t-\tau)^2 + \beta$$

where α and β are positive constants. This gives us the qualitatively correct behavior of I.

Let, in addition, the Newton potential V_N be suitably well-behaved, that is we exclude point-particle and near point-particle collisions (cf. Sec. 3.1). Consequently, the qualitative behavior of $C_S = -V_N \cdot \sqrt{I}$ is, like in (7.9), given by

$$C_S(t) = \gamma|t-\tau| + \delta$$

where γ and δ are positive constants.

Lemma 7.4 (Non-stationarity)**.** *Let (3.12) and (7.9) hold. Let ε be the volume measure on PT^*S as given in (6.45) and let ε_D be the time-evolved measure as defined in (7.21). Let C_S be the complexity as given in (7.3). Then, for $\Gamma_{C^*} \subset PT^*S$,*

$$\varepsilon(T^D\Gamma_{C^*}) < \varepsilon_D(T^D\Gamma_{C^*}). \tag{7.23}$$

Hence, ε is not stationary.

Proof. From $I = \sum_i m\mathbf{q}_i^2$ and $D = \sum_i \mathbf{q}_i\mathbf{p}_i$ it follows that $D = 1/2\dot{I}$ (cf. (6.13)). Given that $I(t) \sim (t-\tau)^2$, it follows that $D \sim |t-\tau|$. Reparametrizing $C_S(t) \sim |t-\tau|$ with respect to D, we find that, $C_S(D) \sim |D|$. Let, without loss of generality, $D > 0$ and

$$C_S(D) = D + C_S(0).$$

Let $p := (\psi(0), \phi(0), \chi_p(0))$ and $p(D) := (\psi(D), \phi(D), \chi_p(D))$. Consider the set of points for which, at $D = 0$, $C_S(0) = C^*$. This is the set $\Gamma_{C^*} = \{p \in PT^*S | C_S(p) = C^*\}$.

Now let $\varepsilon_D(\Gamma_{C^*})$ be the time-evolved measure of the set Γ_{C^*} (cf. (7.21)), i.e.,

$$\varepsilon_D(T^D\Gamma_{C^*}) = \varepsilon(\Gamma_{C^*}). \tag{7.24}$$

Since $C_S(D) = D + C^*$ where $C_S(D) = C_S(p(D))$, we have

$$\begin{aligned} T^D\Gamma_{C^*} &= \{p(D) \in PT^*S | T^{-D}C_S(p(D)) = C^*\} \\ &= \{p(D) \in PT^*S | C_S(p(D)) = D + C^*\} = \Gamma_{D+C^*}. \end{aligned} \tag{7.25}$$

Consequently,

$$\varepsilon_D(T^D\Gamma_{C^*}) = \varepsilon_D(\Gamma_{D+C^*}).$$

Inserting this into (7.24), we get

$$\varepsilon_D(\Gamma_{D+C^*}) = \varepsilon(\Gamma_{C^*}). \tag{7.26}$$

Now recall the shape of the contour lines of C_S (cf. Lemma 7.3). For every $C' \gtrsim 2C_{min}$, the contour line $C_S(\psi, \phi) = C'$ (the set of points of constant C') consists of three circles around the binary collision points. The larger C', the smaller the circles. Hence, for all D and all $C^* \gtrsim 2C_{min}$,

$$\varepsilon(\Gamma_{D+C^*}) < \varepsilon(\Gamma_{C^*}).$$

Using (7.26) this turns into

$$\varepsilon(\Gamma_{D+C^*}) < \varepsilon_D(\Gamma_{D+C^*}).$$

With (7.25) Equation (7.23) follows. That is, we found a set $B \subset PT^*S$, namely $B := \Gamma_{D+C^*}$ $(= T^D\Gamma_{C^*})$ such that $\varepsilon(B) \neq \varepsilon_D(B)$. Hence, ε is not stationary.

$\square$

Since the entaxy is not defined with respect to a stationary measure, it is not a shape analogue of the Boltzmann entropy.

In what follows, we will show that, also on absolute phase space, there is nothing like a shape entropy where by "shape entropy" I refer to a notion of entropy on absolute phase space which is defined with respect to shape macrovariables (like the complexity) alone.

Shape entropy. Let us return to a description of the $E = 0$ Newtonian universe on absolute phase space $\Gamma = \mathbb{R}^{6N}$ and let us determine the absolute phase space volume of regions of constant complexity. Of course, complexity is a macrovariable with respect to absolute phase space as well. Starting from the notion of entaxy and the idea that shape is all there is, it seems natural, as a first step, to define the entropy with respect to some shape macrovariable. For example, one could have the idea to define the entropy with respect to the absolute phase space volume of regions of constant complexity. However, as we will show in the following, the entropy cannot be defined via some shape macrovariable (or set of shape macrovariables) alone.[57]

Definition 7.4 (Shape macrovariable)**.** Let $Q = \mathbb{R}^{3N}$ and let M be a smooth function on Q, invariant under translations, rotations, and scalings. Then we call M a *shape macrovariable.*

Note that according to this definition the complexity C_S is really a shape macrovariable. Now recall that $\mathrm{d}\mu_E = \delta(H(\mathbf{q},\mathbf{p}) - E)\mathrm{d}^{3N}q\mathrm{d}^{3N}p$ (up to some multiplicative constant) with

$$H(\mathbf{q},\mathbf{p}) = \sum_i \frac{\mathbf{p}_i^2}{2m} - \sum_{i<j} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|}.$$

Let again Γ_E denote the constant energy hypersurface: $\Gamma_E = \{(\mathbf{q},\mathbf{p}) \in \Gamma | H(\mathbf{q},\mathbf{p}) = E\}$. We are interested in $\mu_E(\Gamma_{C^*})$ where now Γ_{C^*} is a subset of Γ_E, i.e.,

$$\Gamma_{C^*} = \{(\mathbf{q},\mathbf{p}) \in \Gamma_E | C_S(\mathbf{q}) = C^*\}.$$

Here $C_S = -V_N \cdot \sqrt{I}$ where V_N is the Newton potential, $V_N = -\sum_{i<j} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|}$, and I the center-of-mass moment of inertia: $I = \frac{m}{N}\sum_{i<j} |\mathbf{q}_i - \mathbf{q}_j|^2$.

In what follows we show that every shape macrovariable M (like the complexity C_S) determines a region $\Gamma_M \subset \Gamma_E$ of measure zero or infinity.

Definition 7.5 (Shape macrovariable)**.** Let $Q = \mathbb{R}^{3N}$ and let M be a smooth function on Q, invariant under translations, rotations, and scalings. Then we call M a *shape macrovariable.*

[57] The idea that there cannot be a "shape entropy" due to the scaling properties of the measure is due to Dustin Lazarovici.

Note that according to this definition the complexity C_S is really a shape macrovariable.

Lemma 7.5. *Let everything be as above. Let, in particular, H be the Hamiltonian of the system and μ_E the microcanonical measure as given above. Let M be a shape macrovariable and $\Gamma_{M^*} = \{(\boldsymbol{q}, \boldsymbol{p}) \in \Gamma_E | M(\boldsymbol{q}) = M^*\}$. Let $E = 0$. Then*

$$\mu_E(\Gamma_{M^*}) = \infty \qquad \text{or} \qquad \mu_E(\Gamma_{M^*}) = 0. \tag{7.27}$$

Proof. Since M is a shape macrovariable, it is a function of the positions only: $M = M(\mathbf{q}_1, ...\mathbf{q}_N)$. Now the momentum part of the microcanonical phase space integral determines a $(3N-1)$-dimensional sphere S_r of radius $r = \sqrt{m^{-1}(E + \sum_{i<j} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|})}$ (where, for simplicity, we set $m_i = m \; \forall i = 1, ..., N$). This follows from the fact that $\Gamma_E = \{(\mathbf{q}, \mathbf{p}) \in \Gamma | \sum_i m\mathbf{p}_i^2 - \sum_{i<j} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|} = E\}$ with $\Gamma = \mathbb{R}^{6N}$. We know from (2.32) that, in that case, the microcanonical measure turns into

$$\mathrm{d}\mu_E' = C\biggl(E + \sum_{i<j} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|}\biggr)^{\frac{3N-2}{2}} \mathrm{d}^{3N} q$$

where C depends on N and the other constants, but not on E. Now μ_E' is a measure on $Q = \mathbb{R}^{3N}$ and $\mu_E(\Gamma_{M^*}) = \mu_E'(Q_{M^*})$ where $Q_{M^*} = \{\mathbf{q} \in \mathbb{R}^{3N} | M(\mathbf{q}) = M^*\}$.

For $E = 0$ and large N, the measure turns into

$$\mathrm{d}\mu_E' = C\biggl(\sum_{i<j} \frac{Gm^2}{|\mathbf{q}_i - \mathbf{q}_j|}\biggr)^{\frac{3N}{2}} \mathrm{d}^{3N} q.$$

This is a measure homogeneous of degree $3N/2$. Hence, for $\lambda \in \mathbb{R}^+$, we have

$$\mu_E'(\{\lambda\mathbf{q} \in \mathbb{R}^3 | M(\mathbf{q}) = M^*\}) = \lambda^{3N/2} \cdot \mu_E'(\{\mathbf{q} \in \mathbb{R}^{3N} | M(\mathbf{q}) = M^*\}).$$

Now

$$\begin{aligned} \mu_E'(\{\lambda\mathbf{q} \in \mathbb{R}^{3N} | M(\mathbf{q}) = M^*\}) &= \mu_E'(\{\mathbf{q}' \in \mathbb{R}^{3N} | M(\lambda^{-1}\mathbf{q}') = M^*\}) \\ &= \mu_E'(\{\mathbf{q}' \in \mathbb{R}^{3N} | M(\mathbf{q}') = M^*\}) \end{aligned}$$

where the first equation follows from simple substitution $\mathbf{q}' = \lambda\mathbf{q}$ and the second equation from the fact that M is a shape macrovariable, hence, in particular, it is scale-invariant, $M(\lambda\mathbf{q}) = M(\mathbf{q})$.

Putting everything back together, we get

$$\mu_E'(\{\mathbf{q} \in \mathbb{R}^{3N} | M(\mathbf{q}) = M^*\}) = \lambda^{3N/2} \cdot \mu_E'(\{\mathbf{q} \in \mathbb{R}^{3N} | M(\mathbf{q}) = M^*\}).$$

Hence,

$$\mu_E(\Gamma_{M^*}) = \lambda^{3N/2} \mu_E(\Gamma_{M^*}).$$

This equation can only be fulfilled if Γ_{M^*} has measure zero or infinity. $\square$

This result implies that we cannot define the entropy solely via a shape macrovariable (or set of shape macrovariables). In that sense, there is nothing like a "shape entropy".

Corollary 7.2 ("Shape entropy"). *Let M be a shape macrovariable. Let Γ_{M^*} and μ_E be as above. Then the "shape entropy"*

$$S = k_B \ln \mu_E(\Gamma_{M^*})$$

is not well-defined.

Proof. It follows directly from $\mu_E(\Gamma_{M^*}) = \infty$ or $\mu_E(\Gamma_{M^*}) = 0$ (cf. (7.27)) that $S = \pm\infty$. Hence, S is not well-defined. □

The proof of Lemma 7.5 has shown that, in order for the entropy to be well-defined, we need a macrovariable which keeps track of scales. Accordingly, we can keep the complexity as an interesting macrovariable as long as we introduce a second macrovariable keeping track of absolute scales. Here the moment of inertia I (or its square-root $\sqrt{I}$) is the simplest choice. But this exactly matches the notion of entropy we introduced in Section 2!

Entropy and complexity. In Section 2 we gave a definition of the entropy in terms of (minus) the Newton potential U and the center-of-mass moment of inertia I,

$$S = k_B \ln \mu_E(\Gamma_{U,I}).$$

For the $E = 0$ Newtonian universe, we can express the entropy in terms of C_S and $\sqrt{I}$ as well.

Lemma 7.6. *Let $\mu_E(\Gamma_{U,I})$ be the microcanonical measure (cf. (2.17)). Let $S = k_B \ln \mu_E(\Gamma_{U,I})$ be the entropy (cf. (2.16)) and C_S the complexity (cf. (7.1)). Let $E = 0$. Then*

$$\mu_E(\Gamma_{U,I}) \approx C' \sqrt{I}^{\frac{3N}{2}} C_S^{\frac{3N}{2}} \tag{7.28}$$

where $C' = C\lambda m^{-\frac{3N}{2}}$ with $C = m\Omega^{3N-1}(2m)^{3N/2-1}$ and

$$S \approx \frac{3N}{2} k_B \ln C_S + \frac{3N}{2} k_B \ln \sqrt{I} + S''(N) \tag{7.29}$$

where S'' depends on N and the other constants, but not on C_S or I.

Proof. We know that, for $E = 0$, $\mu_E(\Gamma_{U,I})$ is given by (3.1). From (3.1) with $U = -V_N$ it follows that

$$\mu_E(\Gamma_{U,I}) \approx C' \cdot (-V_N)^{\frac{3N}{2}} I^{\frac{3N}{2}}$$

where $C' = C\lambda m^{-\frac{3N}{2}}$ and $C = m\Omega^{3N-1}(2m)^{3N/2-1}$. Given that $C_S = -V_N \cdot \sqrt{I}$, we have

$$\mu_E(\Gamma_{U,I}) \approx C' \sqrt{I}^{\frac{3N}{2}} C_S^{\frac{3N}{2}}.$$

With $S = k_B \ln \mu_E(\Gamma_{U,I})$ Equation (7.29) follows. □

Lemma 7.6 tells us that a state of high complexity and high moment of inertia is a state of large entropy whereas a state of small complexity and small moment of inertia is a state of low entropy. This follows from the fact that the absolute phase space volume of a state of high complexity and high moment of inertia is by far larger than the absolute phase space volume of a state of small complexity and small moment of inertia (cf. (7.28)) and this does not refer to a proportion of a hundred or a thousand, but to a proportion of about 2^N where N is the number of particles involved (in this case, N is the number of particles in the universe!).

Let us return to the statistical analysis (which is an analysis on PT^*S). From the statistical analysis we know that typically, at $D = 0$, the system is in a state of low complexity. Now Lemma 7.6 tells us that a state of low complexity is a state of low entropy. Hence, typically (where typically refers to the uniform volume measure on PT^*S), at the moment of minimal extension of the particles – the Big Bang of the Newtonian universe –, the universe is in a homogenous, low-entropy state!

Remark (Entropy and entaxy)**.** According to (7.29) the entropy is a function of $\sqrt{I}$ and C_S. Starting from a shape macrovariable like the complexity and adding the moment of inertia as a second macrovariable, this is the simplest possible definition of the entropy (as being defined on absolute phase space). The other way round, projecting the entropy back onto shape phase space, the macrovariables C_S and $\sqrt{I}$ reduce to C_S,

$$C_S\sqrt{I} \quad \text{on}\, \Gamma \ \rightarrow C_S \quad \text{on}\ PT^*S\ ,$$

since we obtain shape phase space (T^*S, respectively PT^*S) by setting $I = 1$ (cf. (5.23)). In that sense, the entropy projects onto the entaxy.

7.4 Discussion

The statistical analysis of the $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian universe on PT^*S provides an unambiguous mathematical result regarding the complexity of the system at time $D = 0$. To be precise, the solution entaxy tells us that typically, at that moment (the moment of minimal extension of the particles, respectively the Big Bang), the system is in a more or less homogenous state!

This is all we need in order to obtain a final explanation of the second law of thermodynamics and the low-entropy past. First, note that the result is in good agreement with observation. Homogeneity is exactly what we find when observing the cosmic microwave background. At the same time, it is counterintuitive. From the behavior of the overall entropy we at first glance conclude that a typical state is one in which the particles form a dilute state of clusters while a homogenous state is atypical (with the entropy increasing as the universe evolves from an atypical, homogeneous state to a typical, non-homogeneous one). Now the opposite holds true, at least for universal macrostates at $D = 0$. At that moment – the Big Bang of the Newtonian universe – the universe is typically in a homogenous state! And all the rest – galaxy formation, expansion, growth of complexity, etc. – is due to the dynamics. There is no evolution form the atypical to the typical here. Everything is typical from the very beginning and what we observe, like galaxies forming and so on, is due to the dynamical law of the Newtonian gravitational system.

In addition, entropy behaves just the way it is supposed to do. It is lowest at the Janus point (where complexity and moment of inertia are lowest) and it increases in both time directions away from that point. As such the Newtonian universe is a Carroll-type universe. This explains the existence of an entropy gradient and, together with the normalizable measure over mid-point data, the fact that we have had a low-entropy past. As such it goes beyond the Carroll proposal. We no longer face the problem of non-normalizability because we *have* a normalizable measure on the space of mid-point data. And this measure tells us that, at that mid-point, the universe is typically in a state of small complexity, which is a low-entropy state! This is all we need the measure for. We just need the assertion that *typically at the moment of minimal extension* the universe is in a homogenous, low-entropy state. All the rest, like galaxy formation and so on *and also the increase of entropy*, is due to the dynamics.

What regards the asymmetry of time, the overall picture is the following. We know from the dynamics that there is a moment in time at which the spatial extension of the system of particles is minimal (the Janus point which we identify with the Big Bang) while the system expands and the particles form clusters in both time directions away from that point. At the same time, the complexity is lowest (more or less) at the Janus point while it increases (with small fluctuations) without bound in both time directions away from it. The increase of the moment of inertia accompanied by the increase of complexity defines a *gravitational arrow of time*. Hence, there are two gravitational arrows of time with one common past at the Janus point and two futures in both directions away from it. At the same time, entropy is lowest at the Janus point and increases in both time directions away from that. The entropy gradient defines a *thermodynamic arrow of time*. Thus again, there are two thermodynamic arrows of time with one common past at the Janus point and two futures in both directions away from it. In this scenario, the gravitational and thermodynamic arrows of time coincide. Moreover, the account is overall time-symmetric exhibiting an asymmetry of time at every moment (apart from the minimum).

Last but not least, how do we connect to the notion of entropy of subsystems? We found that answer already. We showed that there is no shape analogue of the entropy. Instead, we have to define the entropy in terms of absolute quantities (in particular, in terms of a quantity measuring absolute size). But once we do this, the notion of entropy of the universe directly relates to the notion of entropy of subsystems. In case we consider a non-gravitating system, the formula for the entropy of the universe reduces to the formula of Boltzmann. Even more, from the fact that entropy increases as galaxies form we conclude that entropy is lowest at the "birth" of new galaxies. Hence, the galaxies start out from a low-entropy state. And the results of Saari [1971] and Marchal and Saari [1974] tell us that, for each galaxy separately, an asymptotic energy relation holds, that is, the galaxies become more and more isolated, that way forming physical systems in which the usual thermodynamic processes can take place.

Part IV

Dynamics through the Big Bang of the Newtonian universe

This section stands for its own as it is not part of the explanation of the second law of thermodynamics and the low-entropy past of our universe.

During the last sections, the moment of minimal extension of the gravitational N-particle system has represented the Big Bang within the Newtonian universe. At this point we will change the nomenclature. From now on "Big Bang" shall really refer to the moment of zero spatial extension, when all particles collide at one point.[58]

8 Dynamics through the points of total collision

Let us further discuss the $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian universe on shape phase space T^*S and, in particular, the points of total collision of all the particles. We can show that, while the Newtonian trajectory ends at that point, the respective trajectory on shape space can be continued uniquely through the point of total collision. In other words, the shape degrees of freedom can be evolved uniquely through that point. They can be evolved through the singularity! In fact, such a curve on shape space will represent two solutions on absolute phase space – one with a collision in its future, one in its past – glued together at the point of total collision. (Here the collision is taken to happen at $t = 0$ where one solution starts at $t = -\infty$ and ends at $t = 0$, the other starts at $t = 0$ and ends at $t = +\infty$). We show that this "gluing together" is unique. Moreover, we show that the total collision is passed in finite time.

For means of simplicity, I will again discuss the three-particle system. In principle, however, it is conceivable that everything should work for the N-particle model as well.

Koslowski, Mercati, and Sloan [2016] have shown a similar behavior for the Bianchi IX model of general relativity. In that case, they found that the dynamics can be continued through the point of zero spatial volume representing the Big Bang. However, that model is essentially different from the Newtonian. In order to show that the dynamics can be continued through the singularity of the $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian universe, a different strategy of proof needs to be adopted.

8.1 Total collisions in absolute space

In this section, I will discuss the Newtonian universe of N particles while starting from the next section I restrict the discussion to the three-particle system.

We saw in Section 5.2 that the long-time behavior of the Newtonian gravitational system is governed by the Lagrange-Jacobi equation and by Pollard's result. That is, for $E = 0$, there

[58] The idea for discussing total collisions on shape space has been proposed to me by Julian Barbour in a private conversation in Munich in 2017.

exists one point at which the total extension of the system is minimal, $I = I_{min}$, while it grows in both time directions away from it. Now what about the behavior near collisions?

Let me first discuss the notion of a total collision. A total collision represents the Big Bang within the Newtonian universe - it is the moment at which all the particles are at the same point and the total extension of the system of particles is zero. That is, a total collision occurs if and only if, at some moment in time, the moment of inertia I is zero.

Definition 8.1 (Total collision)**.** Let, at some moment t of time,

$$I(t) = \sum_{i=1}^{N} m_i(\mathbf{q}_i(t) - \sum m_i \mathbf{q}_i(t))^2 = 0. \tag{8.1}$$

Then we say that, at that moment, there is a total collision of all the particles.

What about the existence of total collisions? The existence of solutions which end (or start) at a total collision has already been shown by Lagrange and Euler in the 18th century.[59] Lagrange showed that if three particles of equal masses form an equilateral triangle and are released with zero initial velocity, they will collide. This particular configuration of the particles plus its reflected version (the reflected equilateral triangle) are called the two Lagrange configurations. Euler, in turn, showed that if three particles of equal masses are aligned with one particle centered between the other two and they are released with zero initial velocity, they will also collide. These three configurations (one for each possibility to put one particle at the center) are called the Euler configurations.

Sundman [1909] has shown that a total collision can occur only if the total angular momentum vanishes, $\mathbf{L} = 0$. If $\mathbf{L} \neq 0$, I is bounded away from zero by some positive constant: $I \geq I_0$ with $I_0 > 0$. That is, for the $E = \mathbf{L} = 0$ Newtonian universe, total collisions occur.

Saari [1984], [2005] has shown that as the particles approach a total collision, say at time $t = 0$, they form a central configuration and their position vectors $\mathbf{q}_i$ behave as $t^{2/3}$. This we will use in order to show that the solutions can be continued through the total collision and that this happens in finite time.

Let me show how this behavior is attained. For that let me define the notion of a central configuration. Consider the Newtonian gravitational potential:

$$V_N = -\frac{1}{2} \sum_{i \neq j} \frac{G m_i m_j}{|\mathbf{q}_i - \mathbf{q}_j|}. \tag{8.2}$$

This potential forms part of the Newtonian law of gravitation which determines the acceleration $\ddot{\mathbf{q}}_i$ of the i'th particle (with mass m_i) as follows:

$$m_i \ddot{\mathbf{q}}_i = \frac{\partial V_N}{\partial \mathbf{q}_i} = -\sum_{i \neq j} \frac{G m_i m_j (\mathbf{q}_i - \mathbf{q}_j)}{|\mathbf{q}_i - \mathbf{q}_j|^3}. \tag{8.3}$$

[59]Cf. Moeckel [1981] and [2007] for a historical introduction. Cf. also Saari [1971]. Saari shows that among all solutions those which feature a total collision form a set of measure zero. This does not worry us here because in this section we are only interested in whether or not total collisions can be passed (independent of their likeliness to happen).

This is a complicated equation, but sometimes it attains a simpler form. This is the case for central configurations.

Definition 8.2 (Central configuration)**.** Let, at some moment t in time, the acceleration vector $\ddot{\mathbf{q}}_i$ of each particle be in line with its center-of-mass position vector $\mathbf{q}_i^{cm} = \mathbf{q}_i(t) - \sum_i m_i \mathbf{q}_i(t)$, i.e. $\forall i = 1, ..., N$,

$$\ddot{\mathbf{q}}_i(t) = \lambda \mathbf{q}_i^{cm}(t), \tag{8.4}$$

where $\lambda = \lambda(t)$ is some common scalar factor of proportionality. Such a configuration is called a central configuration.

In case the particles form a central configuration, the system mimics a central force problem. For a system of three particles, there exist five central configurations: the two Lagrange and the three Euler configurations.

Saari [1984] showed that, as the particles approach a total collision at $t = 0$, their position vectors behave as

$$\mathbf{q}_i(t) = \mathbf{a}_i t^{\alpha} \tag{8.5}$$

where α is a scalar and $\mathbf{a}_i \neq 0$ is a vector constant. From this the following result is obtained.

Lemma 8.1 (Approach of singularity)**.** *Let $\boldsymbol{q}_i(t) = \boldsymbol{a}_i t^{\alpha}$ $\forall i = 1, ..., N$. Here α is a scalar and $\boldsymbol{a}_i \neq 0$ a vector constant. Then the particles form a central configuration,*

$$\ddot{\boldsymbol{q}}_i(t) = \lambda(t) \boldsymbol{q}_i(t) \tag{8.6}$$

with $\lambda(t) = \alpha(\alpha - 1)t^{-2}$, and the position vector is

$$\boldsymbol{q}_i(t) = \boldsymbol{a}_i t^{2/3}. \tag{8.7}$$

Proof. Let us, without loss of generality, work in the center of mass frame: $\mathbf{q}^{cm} = \sum_{i=1}^{N} m_i \mathbf{q}_i = 0$. In that frame, the center of mass position vector $\mathbf{q}_i^{cm}$ of the i'th particle is $\mathbf{q}_i^{cm} = \mathbf{q}_i - \mathbf{q}^{cm} = \mathbf{q}_i$.

Consider the Newtonian law of motion of the i'th particle:

$$m_i \ddot{\mathbf{q}}_i = \frac{\partial V_{New}}{\partial \mathbf{q}_i} = -\sum_{i \neq j} \frac{G m_i m_j (\mathbf{q}_i - \mathbf{q}_j)}{|\mathbf{q}_i - \mathbf{q}_j|^3}.$$

Given that $\mathbf{q}_i(t) = \mathbf{a}_i t^{\alpha}$ for some α, the Newtonian law turns into

$$\mathbf{a}_i \alpha(\alpha - 1) t^{\alpha - 2} = -\sum_{i \neq j} \frac{G m_j (\mathbf{a}_i - \mathbf{a}_j)}{|\mathbf{a}_i - \mathbf{a}_j|^3} t^{-2\alpha}.$$

This equation can be fulfilled if and only if $\alpha = 2/3$. Using (8.5), this shows (8.7), i.e.,

$$\mathbf{q}_i(t) = \mathbf{a}_i t^{2/3}.$$

In addition, the system forms a central configuration. It is

$$\ddot{\mathbf{q}}_i = \mathbf{a}_i \frac{2}{3}\Big(-\frac{1}{3}\Big) t^{-4/3} = \lambda(t)\mathbf{a}_i t^{2/3} = \lambda(t)\mathbf{q}_i$$

where $\lambda(t) = -4/9t^{-2}$. This shows (8.6). □

This result gives us the behavior of the system near total collisions. In addition, we know from the long-time behavior of the $E = 0$ universe that a total collision can occur only at the "central" time (Janus point) where the extension of the particles is minimal, $I = I_{min}$.

8.2 Passing the singularities on shape space

Let us now develop a description of the total collisions on shape space.

8.2.1 Total collisions on T^*S_R and T^*S

Let us consider the gravitational system of three particles and let us start again from the translationally and rotationally invariant Hopf coordinates w_1, w_2, w_3 and their canonical momenta z_1, z_2, z_3 introduced in (5.35) and (5.36). The Hopf coordinates were

$$w_1 = \frac{|\boldsymbol{\rho}_1|^2 - |\boldsymbol{\rho}_2|^2}{2}, \qquad w_2 = \boldsymbol{\rho}_1 \cdot \boldsymbol{\rho}_2, \qquad w_3 = \boldsymbol{\rho}_1 \times \boldsymbol{\rho}_2 \tag{8.8}$$

and their conjugates are

$$z_1 = \frac{\boldsymbol{\rho}_1 \cdot \boldsymbol{\kappa}_1 - \boldsymbol{\rho}_2 \cdot \boldsymbol{\kappa}_2}{|\boldsymbol{\rho}_1|^2 + |\boldsymbol{\rho}_2|^2}, \quad z_2 = \frac{\boldsymbol{\rho}_1 \cdot \boldsymbol{\kappa}_2 + \boldsymbol{\rho}_2 \cdot \boldsymbol{\kappa}_1}{|\boldsymbol{\rho}_1|^2 + |\boldsymbol{\rho}_2|^2}, \quad z_3 = \frac{\boldsymbol{\rho}_1 \times \boldsymbol{\kappa}_2 - \boldsymbol{\rho}_2 \times \boldsymbol{\kappa}_1}{|\boldsymbol{\rho}_1|^2 + |\boldsymbol{\rho}_2|^2}. \tag{8.9}$$

Here the $\boldsymbol{\rho}_i, \boldsymbol{\kappa}_j$ are the Jacobi coordinates and their canonical conjugates defined in (5.29) and (5.30). We have seen that the w_1, w_2, w_3 and z_1, z_2, z_3 form a complete set of canonical coordinates on the translationally and rotationally reduced phase space T^*S_R (shape phase space with scale) on which the reduced Hamiltonian dynamics of the $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian universe is formulated.

With respect to the Hopf coordinates, the two Lagrange configurations (the equilateral triangle and its reflected version) can be specified as follows. Let us, for simplicity, consider the equal mass case $m_1 = m_2 = m_3 = m$.

Lemma 8.2 (Lagrange configurations). *Let w_1, w_2, w_3 be given by (8.8). They are local coordinates of S_R. Let $m_1 = m_2 = m_3 = m$. Then the two Lagrange configurations are specified by*

$$w_1 = w_2 = 0, \ \ w_3 = \pm||\boldsymbol{w}||. \tag{8.10}$$

Proof. This follows directly from the definition of the Jacobi and Hopf coordinates (cf. (5.29) and (5.35)). □

Analogously, the three Euler configurations (the collinear configurations where one particle is centered between the other two) are determined as follows:

Lemma 8.3 (Euler configurations)**.** *Let everything be as in the previous lemma. The three Euler configurations are determined by*

$$w_3 = 0 \tag{8.11}$$

and

$$(w_1, w_2) \in \left\{ \Big(||\boldsymbol{w}||, 0 \Big), \Big(\frac{1}{2}||\boldsymbol{w}||, \frac{\sqrt{3}}{2}||\boldsymbol{w}|| \Big), \Big(-\frac{1}{2}||\boldsymbol{w}||, -\frac{\sqrt{3}}{2}||\boldsymbol{w}|| \Big) \right\}. \tag{8.12}$$

Proof. Again, this follows directly from the definition of the Jacobi and Hopf coordinates (cf. (5.29) and (5.35)).

□

Hamiltonian on T^*S_R. Let us now consider the dynamics of the $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian universe on reduced phase space T^*S_R.

We know from (6.5) that, in terms of the $\mathbf{w}$ and $\mathbf{z}$ coordinates, the reduced Hamiltonian H on T^*S_R is

$$H = T + V_N = ||\mathbf{w}|| \cdot ||\mathbf{z}||^2 + \frac{V_S}{\sqrt{||\mathbf{w}||}} \tag{8.13}$$

where $V_S = V_S(\mathbf{w})$. Explicitly,

$$V_S = -\sqrt{2} \sum_{i<j} \frac{G(m_i m_j)^{\frac{3}{2}} (m_i + m_j)^{-\frac{1}{2}}}{\sqrt{1 - \mathbf{w} \cdot \mathbf{b}_{ij}/||\mathbf{w}||}}. \tag{8.14}$$

We know that shape space with scale S_R can be depicted as a two-sphere of radius $R = ||\mathbf{w}||$. Introducing spherical coordinates R, ψ, and ϕ, we can write the three unit vectors representing the three binary collisions in the general form $\mathbf{b}_{ij} = (\sin\psi_{ij}\cos\phi_{ij}, \sin\psi_{ij}\sin\phi_{ij}, \cos\psi_{ij})^T$. Here $\mathbf{b}_{12}$ represents the collision between particles 1 and 2 (where $|\mathbf{q}_1 - \mathbf{q}_2| = 0$) and so on. Since the $\mathbf{b}_{ij}$ are unit vectors, they are specified by two angles: ψ_{ij} and ϕ_{ij}.

While the kinetic term T is symmetric with respect to w_1, w_2, w_3 and z_1, z_2, z_3, the shape potential V_S is not. This will be important later when we introduce two different choices of spherical coordinates in order to discuss the Euler, respectively the Lagrange configurations.

The physical vector field on T^*S_R as determined by the Hamiltonian H from (8.13) turns out to be singular at $R = ||\mathbf{w}|| = 0$. This reflects the singularity of the Newton potential at $R = 0$. It is the singularity at the points of total collision which we know from absolute phase space $\Gamma = T^*Q$. We get rid off this singularity when we go to shape phase space with scale T^*S by help of an internal time parameter. On T^*S, scale has vanished and we are left with the evolution equations of the shape degrees of freedom.

Internal time D. Choosing the dilational momentum

$$D = \sum \mathbf{q}_i \cdot \mathbf{p}_i \tag{8.15}$$

as an internal time parameter, we can write down the internal Hamiltonian equations of motion on T^*S. In terms of the Hopf coordinates on T^*S,

$$D = 2\mathbf{w} \cdot \mathbf{z} \tag{8.16}$$

(cf. (5.48)). For a justification of the choice of D as a time variable and a derivation of the internal Hamiltonian equations, cf. Section 6.3.2. These internal Hamiltonian equations are the evolution equations of the shape degrees of freedom (cf. (6.16)).

In what follows, it will turn out that the internal Hamiltonian vector field on T^*S is non-singular at the central configurations (Euler and Lagrange points) which are the only points on T^*S at which a total collision may occur. If these points are passed at $D = 0$ and $R = ||\mathbf{w}|| = 0$, the internal Hamiltonian equations determine precisely the evolution of the shape degrees of freedom through the points of total collision.

In order to formulate the internal Hamiltonian dynamics on T^*S, we need to separate the shape and scale degrees of freedom. This we do by help of spherical coordinates. To discuss the internal Hamiltonian vector field at the central configurations (Euler and Lagrange points) we choose spherical coordinates in such a way that the central configurations lie on the equator of the shape sphere. This is a convenient choice of coordinates insofar as the internal vector field which describes the motion on T^*S is not singular at the equator whereas it is singular at the top and bottom of the shape sphere.

Be aware that this is not a physical singularity, but a coordinate singularity due to the transformation from the $\mathbf{w}$ and $\mathbf{z}$ coordinates to the spherical coordinates R, ψ, ϕ and p_R, p_ψ, p_ϕ as defined in (5.41) and (5.42). There is a coordinate singularity at the top and bottom of the shape sphere specified by $\psi = 0$ and $\psi = \pi$.

Of course, there does not exist one choice of spherical coordinates such that all of the central configurations simultaneously lie on the equator of the shape sphere. But there are two different choices, one which is appropriate for the Euler configurations, the other for the Lagrange configurations. We might, of course, also find coordinates which treat all the central configurations at once (where the central configurations are somewhere on the shape sphere, neither on the equator nor at the top or bottom), but then the vector field will attain a much more difficult form, which is why we don't do that.

Choice of coordinates. The choice of coordinates which allows us to discuss the Euler configurations is the one used by Montgomery [2002] and Barbour, Koslowski, and Mercati [2013], [2015]. It is such that the collinear configurations, i.e., in particular, the Euler configurations lie on the equator of the shape sphere. In that case, the Lagrange configurations are at the top and bottom of the shape sphere. This is the way in which the shape sphere is usually depicted.

The second choice of coordinates can be interpreted as a rotation of the "Hopfian" coordinate system w_1, w_2, w_3 such that w_1 becomes w_2, w_2 becomes w_3, and w_3 becomes w_1. That way, the two Lagrange configurations are "brought onto" the equator of the shape sphere while the Euler configurations are brought onto a meridian (the intersection of the shape sphere and the "new" $w_3 = 0$ plane). Here we use that we can simply rotate the "Hopfian" coordinate system the way we like - this merely reflects different embeddings of the shape sphere within $\mathbb{R}^3$.

Let me emphasize again that the two different choices of coordinates are convenient because each choice will connect to a non-singular physical vector field at the respective central configurations. What we use, at this point, is that the physical vector field is always non-singular on the equator of the shape sphere, independent of the orientation of the Hopfian coordinate system, whereas it is always singular at the top and bottom of the shape sphere. This way we treat the singularity of the spherical coordinates, which is not a singularity of the physical vector field.

- **Euler configurations.** In order to discuss the Euler configurations, let us choose spherical coordinates R, ψ, ϕ such that

$$w_1 = R\sin\psi\cos\phi, \qquad w_2 = R\sin\psi\sin\phi, \qquad w_3 = R\cos\psi. \tag{8.17}$$

Let $\psi = \pi/2$ specify the equator of the shape sphere. It follows that, with respect to these coordinates, the Euler configurations lie on the equator of the shape sphere (since $w_3 = 0$ from (8.11) holds if and only if $\psi = \pi/2$), whereas the Lagrange configurations lie at the top and bottom of the sphere (since $w_1 = w_2 = 0$ from (8.10) holds if and only if $\psi = 0$ or $\psi = \pi$).

Connected to this choice of spherical coordinates R, ψ, ϕ, there exist canonical conjugates p_R, p_ψ, p_ϕ defined as follows:

$$\begin{aligned}
z_1 &= \frac{1}{R}(\cos\phi(Rp_R\sin\psi - p_\psi\cos\psi) - p_\phi\sin^{-1}\psi\sin\phi)\\
z_2 &= \frac{1}{R}(\sin\phi(Rp_R\sin\psi - p_\psi\cos\psi) + p_\phi\sin^{-1}\psi\cos\phi)\\
z_3 &= -p_R\cos\psi - \frac{1}{R}p_\psi\sin\psi.
\end{aligned} \tag{8.18}$$

We know from (4.46) that R, ψ, ϕ and p_R, p_ψ, p_ϕ are canonical coordinates on T^*S_R.

- **Lagrange configurations.** In order to discuss the Lagrange configurations, we choose spherical coordinates R', ψ', ϕ' with $R' = R$ such that

$$w_1 = R'\cos\psi', \qquad w_2 = R'\sin\psi'\cos\phi', \qquad w_3 = R'\sin\psi'\sin\phi'. \tag{8.19}$$

Let now $\psi' = \pi/2$ denote the equator of the shape sphere. It follows that now the two Lagrange configurations lie on the equator of the shape sphere (since $w_1 = 0$ from (8.10) holds if and only if $\psi' = \pi/2$) and the Euler configurations lie on a meridian, the intersection of the shape sphere and the $w_3 = 0$ plane.

Again, their exist canonical conjugates p'_R, p'_ψ, p'_ϕ which are specified by:

$$
\begin{aligned}
z_1 &= -p'_R \cos\psi' - \frac{1}{R'} p'_\psi \sin\psi' \\
z_2 &= \frac{1}{R'}(\cos\phi'(R' p'_R \sin\psi' - p'_\psi \cos\psi') - p'_\phi \sin^{-1}\psi' \sin\phi') \\
z_3 &= \frac{1}{R'}(\sin\phi'(R' p'_R \sin\psi' - p'_\psi \cos\psi') + p'_\phi \sin^{-1}\psi' \cos\phi'). \qquad (8.20)
\end{aligned}
$$

Given that both the $\mathbf{w}$ and $\mathbf{z}$ coordinates and the unprimed spherical coordinates defined in (8.17) and (8.18) are canonical, also the primed spherical coordinates R', ψ', ϕ' and p'_R, p'_ψ, p'_ϕ are canonical. This follows directly from the fact that the primed coordinates are defined by analogy with the unprimed coordinates, the only difference being an overall permutation which does not affect the canonical structure.

Unfortunately, we cannot use one of the two coordinates choices to treat all of the central configurations at once. This is the case because the vector field turns out to be singular at the top and bottom of the shape sphere and once we put either the Euler or the Lagrange configurations on the equator of the shape sphere (where the vector field is non-singular), it follows from Eq.'s (8.10)-(8.12) that at least one of the other configurations is placed at the top or bottom of the sphere.[60]

Dynamics on T^*S_R and T^*S. Since the kinetic term T of the Hamiltonian is symmetric in the w_1, w_2, w_3 and z_1, z_2, z_3 coordinates, we can write down the physical vector field on T^*S_R (respectively, the equations of motion) in such a way that it does not distinguish between the two different choices of coordinates. This is possible as long as we do not write down the explicit form of the shape potential V_S.

From (6.7) we know that, with respect to the unprimed spherical coordinates defined in (8.17) and (8.18), the Hamiltonian H can be written as

$$
H = \frac{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + R^2 p_R^2}{R} + \frac{V_S(\psi, \phi)}{\sqrt{R}}. \qquad (8.21)
$$

Analogously, with respect to the primed spherical coordinates defined in (8.19) and (8.20), H can be written as

$$
H = \frac{p_{\psi'}^2 + \sin^{-2}\psi' p_{\phi'}^2 + (R')^2 (p'_R)^2}{R'} + \frac{V'_S(\psi', \phi')}{\sqrt{R'}}
$$

with $V'_S \neq V_S$. This follows by directly inserting the primed coordinates in (8.13) or by noting that T is invariant under a permutation of the w_i, z_i, but V_S is not.

[60] Of course, we might come up with another, more complicated definition of spherical coordinates which allows us to treat all the central configurations at once, but then the computations will become more complicated, which is why we don't do it here.

The respective Hamiltonian equations of motion on T^*S_R are

$$\begin{aligned}
\frac{\mathrm{d}\psi}{\mathrm{d}t} &= \frac{2p_\psi}{R}, & \frac{\mathrm{d}p_\psi}{\mathrm{d}t} &= \frac{2\sin^{-3}\psi\cos\psi p_\phi^2}{R} - \frac{\partial V_S/\partial\psi}{\sqrt{R}}, \\
\frac{\mathrm{d}\phi}{\mathrm{d}t} &= \frac{2\sin^{-2}\psi p_\phi}{R}, & \frac{\mathrm{d}p_\phi}{\mathrm{d}t} &= -\frac{\partial V_S/\partial\phi}{\sqrt{R}}, \\
\frac{\mathrm{d}R}{\mathrm{d}t} &= 2R\cdot p_R, & \frac{\mathrm{d}p_R}{\mathrm{d}t} &= \frac{p_\psi^2 + \sin^{-2}\psi p_\phi^2 - R^2 p_R^2}{R^2} + \frac{1}{2}\frac{V_S(\psi,\phi)}{R^{3/2}}.
\end{aligned} \tag{8.22}$$

and analogously for the primed coordinates.

Note that the equations of motion diverge in the limit $R \to 0$. That is, the physical vector field is singular at $R = 0$ and the solutions cannot be continued through $R = 0$ (which specifies the total collision of all the particles).

If we now introduce the internal time parameter $\tau = D$, we can write down the physical vector field, respectively, the equations of motion on T^*S (cf. Sec. 6.3.2). These equations of motion are generated by the internal Hamiltonian F_D, the canonical conjugate of $\tau = D$. Again, the equations for the primed variables are analogous (replacing ψ by ψ' etc. and V_S by V_S'). In accordance with (6.16) we can write:

$$\begin{aligned}
\frac{\mathrm{d}\psi}{\mathrm{d}\tau} &= \frac{2p_\psi}{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + \frac{1}{4}\tau^2}, & \frac{\mathrm{d}p_\psi}{\mathrm{d}\tau} &= \frac{2\sin^{-3}\psi\cos\psi p_\phi^2}{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + \frac{1}{4}\tau^2} + \frac{\partial\log(-V_S)}{\partial\psi}, \\
\frac{\mathrm{d}\phi}{\mathrm{d}\tau} &= \frac{2\sin^{-2}\psi p_\phi}{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + \frac{1}{4}\tau^2}, & \frac{\mathrm{d}p_\phi}{\mathrm{d}\tau} &= \frac{\partial\log(-V_S)}{\partial\phi}.
\end{aligned} \tag{8.23}$$

It can be seen by direct computation that this vector field is singular at $\psi = 0$ and $\psi = \pi$ (top and bottom of the shape sphere).

Lemma 8.4 (Singularity of vector field at $\psi = 0$ and $\psi = \pi$)**.** *Let the equations of motion on T^*S be given in (8.23). The corresponding physical vector field X_{F_τ} is singular at $\psi = 0$ and $\psi = \pi$, that is, at that point, at least one of the equations diverges.*

Proof. Consider the equation for p_ψ,

$$\frac{\mathrm{d}p_\psi}{\mathrm{d}\tau} = \frac{2\sin^{-3}\psi\cos\psi p_\phi^2}{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + \frac{1}{4}\tau^2} + \frac{\partial\log(-V_S)}{\partial\psi}.$$

In the limit $\psi \to 0$ and $\psi \to \pi$, it is

$$\begin{aligned}
\lim_{\psi\to 0/\pi}\left[\frac{2\sin^{-3}\psi\cos\psi p_\phi^2}{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + \frac{1}{4}\tau^2}\right] &= \lim_{\psi\to 0/\pi}\left[\frac{-6\sin^{-4}\psi\cos\psi p_\phi^2 - 2\sin^{-2}\psi p_\phi^2}{-2\sin^{-3}\psi p_\phi^2}\right] \\
&= \lim_{\psi\to 0/\pi}\left[\frac{3\cos\psi}{\sin\psi} + \sin\psi\right] = \infty.
\end{aligned}$$

This shows the assertion. □

An analogous result holds for the primed coordinates. In total, the vector field is singular at

$\psi = 0$ and $\psi = \pi$ (if we consider the unprimed variables), respectively at $\psi' = 0$ and $\psi' = \pi$ (if we consider the primed variables). This is why we cannot use one of the two coordinate choices from above to treat all of the central configurations at once.

8.2.2 Result

We can now prove that the physical vector field on T^*S is non-singular at the central configurations at $\tau = 0$ (except if $p_\psi = p_\phi = 0$, which defines a set of measure zero, cf. Lemma 8.5).

Theorem 8.1 (Passing the singularity)**.** *Let everything be as above. Let, in particular, the equations of motion on T^*S be given in (8.23). Let $\tau = 0$. Then there exist local coordinates (different for the Euler and Lagrange configurations) such that the internal Hamiltonian vector field at the central configurations is given by*

$$\begin{aligned} \frac{\mathrm{d}\psi}{\mathrm{d}\tau} &= \frac{2p_\psi}{p_\psi^2 + p_\phi^2}, \qquad & \frac{\mathrm{d}p_\psi}{\mathrm{d}\tau} &= 0, \\ \frac{\mathrm{d}\phi}{\mathrm{d}\tau} &= \frac{2p_\phi}{p_\psi^2 + p_\phi^2}, \qquad & \frac{\mathrm{d}p_\phi}{\mathrm{d}\tau} &= 0. \end{aligned} \tag{8.24}$$

This vector field is non-singular except if $p_\psi = p_\phi = 0$.

From the non-singularity of the vector field (except if $p_\psi = p_\phi = 0$) it follows that the dynamics can be continued uniquely through the Euler and Lagrange configurations (except if $p_\psi = p_\phi = 0$).

Proof. Remember that, for both choices of coordinates, the physical vector field on T^*S can be written as follows:

$$\begin{aligned} \frac{\mathrm{d}\psi}{\mathrm{d}\tau} &= \frac{2p_\psi}{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + \frac{1}{4}\tau^2}, \qquad & \frac{\mathrm{d}p_\psi}{\mathrm{d}\tau} &= \frac{2\sin^{-3}\psi\cos\psi p_\phi^2}{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + \frac{1}{4}\tau^2} + \frac{\partial\log(-V_S)}{\partial\psi}, \\ \frac{\mathrm{d}\phi}{\mathrm{d}\tau} &= \frac{2\sin^{-2}\psi p_\phi}{p_\psi^2 + \sin^{-2}\psi p_\phi^2 + \frac{1}{4}\tau^2}, \qquad & \frac{\mathrm{d}p_\phi}{\mathrm{d}\tau} &= \frac{\partial\log(-V_S)}{\partial\phi}. \end{aligned}$$

These equations are the same for the primed and unprimed coordinates. Remember that we have chosen the spherical coordinates such that the central configurations under consideration lie on the equator of the shape sphere. That is, it suffices to analyze the vector field on the equator which is specified by $\psi = \psi' = \pi/2$. In that case, $\sin\psi = \sin\psi' = 1$ and $\cos\psi = \cos\psi' = 0$. Hence, the equations of motion become (both for the primed and unprimed coordinates)

$$\begin{aligned} \frac{\mathrm{d}\psi}{\mathrm{d}\tau} &= \frac{2p_\psi}{p_\psi^2 + p_\phi^2 + \frac{1}{4}\tau^2}, \qquad & \frac{\mathrm{d}p_\psi}{\mathrm{d}\tau} &= \left[\frac{\partial\log(-V_S)}{\partial\psi}\right]_{\psi=\frac{\pi}{2}}, \\ \frac{\mathrm{d}\phi}{\mathrm{d}\tau} &= \frac{2p_\phi}{p_\psi^2 + p_\phi^2 + \frac{1}{4}\tau^2}, \qquad & \frac{\mathrm{d}p_\phi}{\mathrm{d}\tau} &= \left[\frac{\partial\log(-V_S)}{\partial\phi}\right]_{\psi=\frac{\pi}{2}}. \end{aligned} \tag{8.25}$$

In remains to determine the partial derivatives of $\log(-V_S)$ at the central configurations. Since the shape potential is different for the two different choices of spherical coordinates (the unprimed coordinates (8.17)+(8.18) and the primed coordinates (8.19)+(8.20)), it follows that we have to discuss the Euler and Lagrange configurations separately.

From (8.14) we know that, with respect to the Hopf coordinates, the shape potential $V_S = V_S(\mathbf{w})$ can be written as

$$V_S = -\sqrt{2}\sum_{i<j} \frac{G(m_i m_j)^{\frac{3}{2}}(m_i + m_j)^{-\frac{1}{2}}}{\sqrt{1 - \mathbf{w}\cdot\mathbf{b}_{ij}/||\mathbf{w}||}}.$$

Here the $\mathbf{b}_{ij}$ are the unit vectors which represent the three binary collisions (specified by two angles ψ_{ij} and ϕ_{ij} and where the i, j refer to the collision particles). Let us now determine the form of V_S for the two different choices of spherical coordinates.

- **Euler configurations.** The binary collision vectors always lie on the $w_3 = 0$ plane. For the first choice of coordinates, (8.17)+(8.18), this means that they lie on the equator of the shape sphere. The equator is specified by the angle $\psi = \pi/2$, hence, $\forall i, j$: $\psi_{ij} = \pi/2$. If the particles have equal masses, the three binary collision points are further specified by $\phi_{ij} = \frac{1}{3}\pi$, $\phi_{ij} = \pi$, and $\phi_{ij} = \frac{5}{3}\pi$ (where, again, the i and j refer to the collision particles). Hence, for the given coordinates,

  $$\mathbf{w}\cdot\mathbf{b}_{ij} = ||\mathbf{w}|| \sin\psi \cos(\phi - \phi_{ij}).$$

  Here the $(\phi - \phi_{ij})$ are the angles between the $\mathbf{b}_{ij}$ and the projection of $\mathbf{w}$ onto the $w_3 = 0$ plane and the term $||\mathbf{w}|| \sin\psi$ is the component of $\mathbf{w}$ which is parallel to the $w_3 = 0$ plane.

  The shape potential can then be written as

  $$V_S = -\sum_{i<j} \frac{(m_i m_j)^{\frac{3}{2}}(m_i + m_j)^{-\frac{1}{2}}}{\sqrt{1 - \sin\psi\cos(\phi - \phi_{ij})}}. \tag{8.26}$$

  This is the choice of coordinates which is appropriate to discuss the Euler configurations.

- **Lagrange configurations.** For the second choice of coordinates, (8.19) and (8.20), the $w_3 = 0$ plane is specified by $\phi'_{ij} = 0$, respectively $\phi'_{ij} = \pi$. (In this case, the intersection of the shape sphere and the $w_3 = 0$ plane is a meridian, not the equator.) The three binary collision vectors are further specified by the angles ψ'_{ij} (with $i < j$ and $i, j = 1, 2, 3$). In the equal-mass case, the collision vectors are $\mathbf{b}_{ij} = (\phi'_{ij}, \psi'_{ij}) = (0, \frac{1}{3}\pi)$, $\mathbf{b}_{ij} = (0, \pi)$, and $\mathbf{b}_{ij} = (\pi, \frac{1}{3}\pi)$. We now have

  $$\mathbf{w}\cdot\mathbf{b}_{ij} = ||\mathbf{w}|| \cos\phi' \cos(\psi' - \psi'_{ij}).$$

  Again, $(\psi' - \psi'_{ij})$ is the angle between the projection of $\mathbf{w}$ onto the $w_3 = 0$ plane and the binary collision vector and $||\mathbf{w}|| \cos\phi'$ is the component of $\mathbf{w}$ parallel to the $w_3 = 0$ plane.

The shape potential can now be written as

$$V'_S = -\sum_{i<j} \frac{(m_i m_j)^{\frac{3}{2}}(m_i+m_j)^{-\frac{1}{2}}}{\sqrt{1-\cos\phi'\cos(\psi'-\psi'_{ij})}}. \tag{8.27}$$

This is the choice of coordinates which is appropriate to discuss the Lagrange configurations.

Let us now determine the partial derivatives of $\log(-V_S)$ and $\log(-V'_S)$ at the central configurations (i.e., the right hand side of (8.25)) for each of the two cases separately.

- **Euler configurations.** For the first choice of coordinates, the three Euler configurations are specified by $\psi_E = \pi/2$ and $\phi_E \in \{0, \frac{2}{3}\pi, \frac{4}{3}\pi\}$. This is again the equal mass case. In addition, remember that the binary collision vectors were specified by $\phi_{ij} = \{\frac{1}{3}\pi, \pi, \frac{5}{3}\pi\}$. Then, with V_S given by (8.25), for each choice of ψ_E, ϕ_E:

$$\left[\frac{\partial \log(-V_S)}{\partial \psi}\right]_{\psi_E,\phi_E} = \left[\frac{1}{\sqrt{2}}\sum_{i<j}\frac{\cos\psi\cos(\phi-\phi_{ij})}{\sqrt{1-\sin\psi\cos(\phi-\phi_{ij})}^3}\right]_{\psi_E,\phi_E} = 0$$

and

$$\begin{aligned}\left[\frac{\partial \log(-V_S)}{\partial \phi}\right]_{\psi_E,\phi_E} &= \left[-\frac{1}{\sqrt{2}}\sum_{i<j}\frac{\sin\psi\sin(\phi-\phi_{ij})}{\sqrt{1-\sin\psi\cos(\phi-\phi_{ij})}^3}\right]_{\psi_E,\phi_E} \\ &= -\frac{1}{\sqrt{2}}\left[\sum_{i<j}\frac{\sin(\phi-\phi_{ij})}{\sqrt{1-\cos(\phi-\phi_{ij})}^3}\right]_{\phi_E} = 0.\end{aligned}$$

- **Lagrange configurations.** For the second choice of coordinates, the two Lagrange configurations are specified by $\psi'_L = \pi/2$ and $\phi'_L \in \{\pi/2, 3\pi/2\}$. In this case, remember that the binary collision vectors were specified by $(\phi'_{ij}, \psi'_{ij}) = \{(0, \frac{1}{3}\pi), (0, \pi), (\pi, \frac{1}{3}\pi)\}$. Then, with V_S given by (8.26), we have:

$$\left[\frac{\partial \log(-V'_S)}{\partial \psi'}\right]_{\psi'_L,\phi'_L} = \left[-\frac{1}{\sqrt{2}}\sum_{i<j}\frac{\cos\phi'\sin(\psi'-\psi'_{ij})}{\sqrt{1-\cos\phi'\cos(\psi'-\psi'_{ij})}^3}\right]_{\psi'_L,\phi'_L} = 0$$

and

$$\begin{aligned}\left[\frac{\partial \log(-V'_S)}{\partial \phi'}\right]_{\psi'_L,\phi'_L} &= \left[-\frac{1}{\sqrt{2}}\sum_{i<j}\frac{\sin\phi'\cos(\psi-\psi_{ij})}{\sqrt{1-\cos\phi'\cos(\psi'-\psi'_{ij})}^3}\right]_{\psi'_L,\phi'_L} \\ &= -\frac{1}{\sqrt{2}}\left[\sum_{i<j}\pm\cos(\psi'-\psi'_{ij})\right]_{\psi'_L} = 0.\end{aligned}$$

Hence, both for the Lagrange and the Euler configurations (both for the primed and unprimed coordinates), the equations of motion on T^*S turn into

$$\frac{\mathrm{d}\psi}{\mathrm{d}\tau} = \frac{2p_\psi}{p_\psi^2 + p_\phi^2 + \frac{1}{4}\tau^2}, \qquad \frac{\mathrm{d}p_\psi}{\mathrm{d}\tau} = 0,$$
$$\frac{\mathrm{d}\phi}{\mathrm{d}\tau} = \frac{2p_\phi}{p_\psi^2 + p_\phi^2 + \frac{1}{4}\tau^2}, \qquad \frac{\mathrm{d}p_\phi}{\mathrm{d}\tau} = 0.$$

Now remember that, for the $E = 0$ Newtonian universe, a total collision may occur only at $\tau = 0$. At that moment, the vector field turns into:

$$\frac{\mathrm{d}\psi}{\mathrm{d}\tau} = \frac{2p_\psi}{p_\psi^2 + p_\phi^2}, \qquad \frac{\mathrm{d}p_\psi}{\mathrm{d}\tau} = 0,$$
$$\frac{\mathrm{d}\phi}{\mathrm{d}\tau} = \frac{2p_\phi}{p_\psi^2 + p_\phi^2}, \qquad \frac{\mathrm{d}p_\phi}{\mathrm{d}\tau} = 0.$$

We see that the vector-field is non-singular at $\tau = 0$ except if $p_\psi = p_\phi = 0$. □

Lemma 8.5. *Let $\mathrm{d}\mu = \mathrm{d}\psi\mathrm{d}\phi\mathrm{d}p_\psi\mathrm{d}p_\phi$ be the natural, invariant volume measure on T^*S (cf. 5.52). Then $p_\psi = p_\phi = 0$ defines a set of measure zero on $\Gamma_{E,L} \subset T^*S$ with $\Gamma_{E,L} = \{(\psi, \phi, p_\psi, p_\phi) \in T^*S | (\psi, \phi) \in \{(\psi_E, \phi_E), (\psi_L, \phi_L)\}\}$.*

Proof. We know from (5.52) that, for the given spherical coordinates, $\mathrm{d}\mu = \mathrm{d}\psi\mathrm{d}\phi\mathrm{d}p_\psi\mathrm{d}p_\phi$.

Let now $\Gamma_{E,L} = \{(\psi, \phi, p_\psi, p_\phi) \in T^*S | (\psi, \phi) \in \{(\psi_E, \phi_E), (\psi_L, \phi_L)\}\}$ be the subset of T^*S corresponding to the central configurations. The projection of the measure μ on T^*S onto $\Gamma_{E,L}$ is

$$\mathrm{d}\nu = \mathrm{d}\mu(\cdot|\Gamma_{E,L}) = \delta(\psi - \psi_{E/L})\delta(\phi - \phi_{E,L})\mathrm{d}\psi\mathrm{d}\phi\mathrm{d}p_\psi\mathrm{d}p_\phi = \mathrm{d}p_\psi\mathrm{d}p_\phi.$$

Let $B = \{(p_\psi, p_\phi) \in \Gamma_{E,L} | p_\psi = p_\phi = 0\}$ be the set of points in $\Gamma_{E,L}$ for which the vector field is singular. Note that B consists of one element only, namely the point $(0, 0) \in \Gamma_{E,L}$. Since ν is a continuous measure on $\Gamma_{E,L}$, it follows that B constitutes a set of measure zero:

$$\frac{\nu(B)}{\nu(\Gamma_{E,L})} = \frac{\int \delta(p_\psi - 0)\delta(p_\phi - 0)\mathrm{d}p_\psi\mathrm{d}p_\phi}{\int \mathrm{d}p_\psi\mathrm{d}p_\phi} = \frac{1}{\int \mathrm{d}p_\psi\mathrm{d}p_\phi} = 0.$$

□

Hence, the vector field is non-singular almost everywhere. Whenever the vector field is non-singular, there exists a unique solution. Consequently, we found that the trajectories of the three-particle system on shape space can be continued (uniquely) through the central configurations for almost all "initial" conditions/mid-point data $\psi(\tau = 0) = \psi_{E,L}$, $\phi(0) = \phi_{E,L}$, $p_\psi(0), p_\phi(0)$.

8.3 Passage time

How much external time T does it take for the particles to evolve through the Newtonian singularity of a total collision, represented by the Euler and Lagrange points on T^*S? And how much internal time $\tau = D$ does it take?

Lemma 8.6 (Passage time)**.** *Let R defined by (8.17) and D defined by (8.15). Let there be a Newtonian gravitational system of N particles with infinitesimal "size" R approaching a total collision. The external time T (internal time D) it takes to reach the point of collision is*

$$T = \lambda R^{3/4} \quad (D = \lambda^{2/3} R^{1/4}) \tag{8.28}$$

where λ is some positive constant.

Proof. We know that, in terms of the local coordinates (8.17) and (8.18) on T^*S_R, we have $D = 2R \cdot p_R$ (cf. (5.49)) and $I = 2R$ (cf. (5.50)). Here D is the dilational momentum and I is the moment of inertia, both in the center-of-mass frame.

In addition, we know from the Lagrange-Jacobi equation (6.2) that a total collision can only occur at $\tau = D = 0$. Let this, without loss of generality, coincide with external time $t = 0$ (this can be done because the Newtonian equations are invariant under time translation). Since the equations of motion are time-reversal invariant, we can further assume that the collision is reached from the positive (i.e. for $t \to 0$ where t is positive).

We know that, as the system approaches a total collision at $t = 0$, the particles' position vectors behave as $\mathbf{q}_i(t) = \mathbf{a}_i t^{2/3}$. Consequently, for $R(t) = 1/2\ I(t) = 1/2 \sum m_i \mathbf{q}_i^2(t)$, the following holds: $R(t) = \lambda t^{4/3}$ where $\lambda = 1/2 \sum m_i \mathbf{a}_i^2$. Now, for a system starting with fix moment of inertia I respectively fix "size" R, this means that the (external) time T it takes to reach a total collision is

$$T = \lambda^{-1} R^{3/4}.$$

Moreover, we know from the equations of motion on shape space with scale, T^*S_R, that the scale R of the system, respectively the moment of inertia I (where, in the given coordinates, $2R = I$), changes with respect to external time t as follows:

$$\frac{\mathrm{d}R}{\mathrm{d}t} = 2R \cdot p_R = D.$$

Inserting $R(t) = \lambda t^{4/3}$, we get $\mathrm{d}R/\mathrm{d}t = \lambda \mathrm{d}(t^{4/3})/\mathrm{d}t = \lambda t^{1/3}$. This behavior we also get directly from the equation for $\mathbf{q}_i(t)$ using the definition of D. In that case, $D(t) = \sum \mathbf{q}_i(t)\mathbf{p}_i(t) \propto t^{2/3} t^{-1/3} \propto t^{1/3}$. Hence, given that a system starts with a "size" R, the internal time D it takes to reach a total collision is

$$D = \lambda T^{1/3} = \lambda^{2/3} R^{1/4}.$$

□

That is, in the Newtonian N-particle problem total collisions are reached within a finite time, both in absolute space where time t is taken to be absolute and on shape space where time $\tau = D$ is internal. Since the Newtonian equations are time-reversal invariant, this is also the time it takes to leave the collision and reattain a size R. Hence, we may conclude that the particles pass the singularity in finite (external and internal) time.

8.4 Discussion

We found that on shape phase space T^*S the Newtonian trajectories can be continued uniquely through the Euler and Lagrange points. Going back to a description in absolute space, each of these points determines an equivalence class of central configurations, given by the set of all configurations which are connected by an overall translation, rotation, or scaling. This means that, in particular, the size of the system as it passes the Euler or Lagrange points on T^*S is arbitrary (size is not an artifact of shape space, but of absolute space alone). It may be arbitrarily small. We may even take it to be zero. This is not a problem on shape space and it need not be a problem on absolute space as long as we understand what it means. What we need to understand is how a trajectory with zero size, $R = 0$, passing a central configuration in T^*S at $D = 0$ relates back to a trajectory in absolute space and time.

On absolute phase space Γ, trajectories cannot be continued through the Newtonian singularity at $R = 0$ due to the singularity of the vector field. Consequently, trajectories either end or begin at that point. Still, in shape space there is no way to "feel" this singularity. The shape degrees of freedom simply evolve through that point at $D = 0$. And they do that in a unique way.

Going back to shape phase space with scale T^*S_R, every shape space solution passing a central configuration at $D = 0$ corresponds to a pair of solutions starting at, respectively ending at a central configuration at $D = 0$ and $R = 0$. While $\phi(0), \psi(0), p_\phi(0), p_\psi(0)$ are "mid-point" data (i.e. data specified at $D = 0$) uniquely determining the trajectory on T^*S, these data together with $R = 0$ and $D = 0$ uniquely specify the pair of trajectories on T^*S_R which corresponds to the two halves of the shape space trajectory (cut in two halves at $D = 0$). Of course, since the vector field on T^*S_R is singular at $R = 0$, these provide asymptotic data. They specify the end, respectively the starting point of the two trajectories on T^*S_R which, when "glued together" at $D = 0$, $R = 0$ form one trajectory passing the singularity.

Going back to absolute phase space Γ, the pair of trajectories on T^*S_R corresponds to an equivalence class of pairs of trajectories (all those which can be reached by an overall translation and/or rotation) on Γ. Back to the overall picture this means that when we say we continue the dynamics through the singularity, we "glue together" *two* trajectories on Γ at the point of total collision. This "gluing together" is unique – the shape degrees of freedom can be evolved uniquely through that point of total collision – and the two trajectories on Γ after having been "glued together" form *one* trajectory passing the singularity.

9 Conclusion and outlook

Let us sum up the results on the $E = \mathbf{P} = \mathbf{L} = 0$ Newtonian universe thereby providing an explanation of the second law of thermodynamics and the low-entropy past of our universe.

We get a lot already from the dynamics. To be precise, within the Newtonian $E = 0$ model of the universe there exists one moment in time at which the extension of the particles is minimal, the so-called Janus point. This we identify with the Big Bang. In both time directions away from that point, the system expands – thereby defining two gravitational arrows of time, one in each direction away from the Janus point. Hence, within this model, there is one common past at the point of minimal extension of the particles and there are two futures in both time directions away from it. As the system expands, galaxies form and complexity grows (with small fluctuations) due to the gravitational dynamics.

As the system expands and gets more and more complex, the absolute phase space volume of the respective macro-regions and, as a consequence, the entropy of the Newtonian universe as defined in Section 2 increases. That is, the gravitational arrows of time are directly correlated with thermodynamic arrows of time – arrows of time given by the entropy gradient. Hence, the universal entropy curve is U-shaped just like it has been proposed by Sean Carroll. This explains the observation of an entropy gradient, but it also features a non-normalizable measure which cannot explain the fact that we are, at this moment, far from the minimum of the entropy curve.

Now in addition to the dynamics, we have the statistical analysis. We have a uniform volume measure on the space of physically distinct solutions, respectively physically distinct mid-point data. This is the correct measure for the statistical analysis of the Newtonian universe. It tells us that typically, at the moment of minimal extension of the particles (the Big Bang of the Newtonian universe), the system is in a state of low complexity, that is, in a more or less homogenous state.

Be aware that this is the only statistical assertion which is made in the entire account. This is all we need the measure for: to state that typically, at the Big Bang, the universe was in a homogeneous state. This agrees exactly with what we know about the Big Bang. But even more, it corresponds to a low-entropy state given the notion of entropy introduced in Section 2. So this is basically what we found: that a typical state of the universe at the Big Bang – typical with respect to the uniform measure on the space of physically distinct mid-point data – is a state of low entropy. And entropy increases due to the dynamics as the system expands and galaxies form in both directions away from the minimum. Now everything agrees with the drawing of Roger Penrose: the typical universe evolves from a homogenous state towards a dilute state of clusters with the entropy of the universe increasing all along.[61]

The second part of the thesis concerns the singularity of total collisions. We found that the shape degrees of freedom can be evolved uniquely through the points of total collision, providing a unique way to combine two trajectories on absolute phase space – one ending at and one starting from a total collision – forming one trajectory which passes the singularity!

[61] Cf. Penrose [2004].

Finally let me say something about future research. So far everything has been done explicitly only for the three-particle model. There is no doubt that everything should work for the N-particle model as well. With respect to some parts of the analysis this is clear by inspection, other parts would have to be worked out more carefully.

Another way of research would be to check how far one could get within a purely shape dynamical theory factoring out scales from the very beginning. Starting from first principles and not from the Newtonian theory, this is what we should aim at. For a priori there is no reason to believe that scale should be treated differently from position and orientation. There has been an attempt of Barbour to do this, but later he dismissed his work and now it is unclear whether there is a chance for it to work.[62]

Over and above the issue of entropy and the arrow of time, there are many ways in which a relational theory of space and time can be further pursued. Thus, it would be interesting to consider the quantum case. Quantum theory is fundamental and if we want a quantum theory of the universe we should better consider a relational theory, that is, we should factor out translations, rotations (and scalings) as well.

In addition, we can consider the relativistic case. Shape dynamics provides an alternative formulation of general relativity (at least, of part of the solutions of GR) describing the evolution of a conformal three-geometry (a spatial three-geometry containing only shape degrees of freedom) with respect to a distinguished time, the so-called York time. This might provide the setting for a relativistic quantum theory leading towards a quantum theory of gravity. All in all I believe that shape dynamics still offers a great number of projects, worth to work on, tackling the fundamental questions of nature.

[62] Cf. Barbour [2003] and Barbour et al. [2013].

Appendix A: Additional mathematics

A.1 Liouville's theorem for a time-dependent Hamiltonian

Liouville's theorem is the statement made by the Liouville equation. The Liouville equation states that the volume of a region in phase space that is transported by a Hamiltonian phase flow $T_{t,s}$ is conserved under the flow. This holds both for a time-independent Hamiltonian (in which case $T_{t,s}$ reduces to a one parameter flow T_t) as well as for a time-dependent Hamiltonian $H_t = H(\mathbf{q}, \mathbf{p}, t) : \Gamma \times \mathbb{R} \to \mathbb{R}$. Here $(\mathbf{q}, \mathbf{p}, t) := (\mathbf{q}_1, ..., \mathbf{q}_n, \mathbf{p}_1, ..., \mathbf{p}_n, t) \in \Gamma \times \mathbb{R}$. Since the Liouville equation plays such a central role in the discussion of the measure, I want to proof it here again without reference to differential geometry for the general case of a time-dependent Hamiltonian H_t.[63]

Let $T_{t,s}$ be the Hamiltonian phase flow on Γ, i.e. the flow lines are the integral curves along the Hamiltonian vector field, here denoted by $\mathbf{v}_H(\mathbf{x}, t)$. The Hamiltonian vector field $\mathbf{v}_H(\mathbf{x}, t)$ is given by the vector $(\partial H(\mathbf{q}, \mathbf{p}, t)/\partial \mathbf{p}, -\partial H(\mathbf{q}, \mathbf{p}, t)/\partial \mathbf{q})^T$.

Let $\mu(t) = \mu\big(A(t)\big)$ denote the volume of a region $A(t) \subset \Gamma$ at time t and let $A(t) = T^{t,s}A(s)$. This last equation just asserts that every point $(\mathbf{q}(s), \mathbf{p}(s)) \in A(s)$ is transported by the Hamiltonian phase flow $T_{t,s}$ to another point $(\mathbf{q}(t), \mathbf{p}(t)) \in A(t)$ (in differential form this equation is called the continuity equation). Now the Liouville theorem states that the volume of the respective regions $A(s)$ and $A(t)$ is constant under the flow.

Theorem 9.1 (Liouville)**.** *The Hamiltonian phase flow $T_{t,s}$ leaves the volume unaltered, that is, for all t and s,*

$$\mu(t) = \mu(s). \tag{9.1}$$

Proof. Let us, to shorten the notation, define $\mathbf{x} := (\mathbf{q}(s), \mathbf{p}(s))$. Let $\mathbf{x}$ fulfill a set of differential equations, $\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x}, t)$, whose solutions exist for all times t. In particular, let $\mathbf{x}$ fulfill the Hamiltonian equations of motion: $\dot{\mathbf{x}} = \mathbf{v}_H(\mathbf{x}, t)$. This means that $\mathbf{x}$ is transported by the Hamiltonian phase flow, $\mathbf{x}(t) = T_{t,s}\mathbf{x}$. For small times $(t - s) \to 0$, the phase flow $T_{t,s}$ is given by the group of transformations

$$T_{t,s}(\mathbf{x}) = \mathbf{x} + \mathbf{v}_H(\mathbf{x}, t)(t - s) + \mathcal{O}\big((t - s)^2\big). \tag{9.2}$$

Let us now consider a region $A(s) \subset \Gamma$. The volume $\mu(s)$ of a region $A(s)$ is given by

$$\mu(s) = \int_{A(s)} \mathrm{d}\mathbf{x}. \tag{9.3}$$

Now a region $A(s)$ is transported by the flow $T_{t,s}$ to another region $A(t) = T_{t,s}A(s)$. Setting $\mathbf{y} = T_{t,s}\mathbf{x}$ and using the identity $\int \mathrm{d}\mathbf{y} = \int \det \frac{\partial \mathbf{y}}{\partial \mathbf{x}} \mathrm{d}\mathbf{x}$, we can determine the volume of the region $A(t)$ at time t as follows:

$$\mu(t) = \int_{A(t)} \mathrm{d}\mathbf{y} = \int_{A(s)} \det \frac{\partial T_{t,s}\mathbf{x}}{\partial \mathbf{x}} \, \mathrm{d}\mathbf{x}. \tag{9.4}$$

[63] For this proof of Liouville's theorem, cf. Arnol'd [1989] or Scheck [2003].

For small times $(t-s) \to 0$ we can use the above formula for the flow (Eq. (8.2)) to compute $\partial T_{t,s}\mathbf{x}/\partial \mathbf{x}$ (which is the Jacobi matrix of the transformation). It follows that

$$\frac{\partial T_{t,s}\mathbf{x}}{\partial \mathbf{x}} = \mathbb{I} + \frac{\partial \mathbf{v}_H(\mathbf{x},t)}{\partial \mathbf{x}}(t-s) + \mathcal{O}\big((t-s)^2\big). \tag{9.5}$$

Now for any matrix $B = (b_{ij})$ and for small times $\tau \to 0$ the following relation holds true:

$$\det(\mathbb{I} + B\tau) = 1 + \tau \ \mathrm{Tr}\ B + \mathcal{O}\big(\tau^2\big), \tag{9.6}$$

where $\mathrm{Tr}\ B = \sum_{i=1}^{2n} b_{ii}$ denotes the trace of B. Using this identity, we have

$$\det \frac{\partial T_{t,s}\mathbf{x}}{\partial \mathbf{x}} = 1 + (t-s)\ \mathrm{Tr}\frac{\partial \mathbf{v}_H(\mathbf{x},t)}{\partial \mathbf{x}} + \mathcal{O}\big((t-s)^2\big). \tag{9.7}$$

But $\mathrm{Tr}\frac{\partial \mathbf{v}}{\partial \mathbf{x}} = \sum_{i=1}^{2n} \partial \mathbf{v}_i/\partial \mathbf{x}_i$ is nothing else than the divergence of $\mathbf{v}$. Hence, the volume $\mathrm{Vol}(t)$ is given by

$$\mu(t) = \int_{A(s)} [1 + (t-s)\ \mathrm{div}\ \mathbf{v}_H + \mathcal{O}\big((t-s)^2\big)]\mathrm{d}\mathbf{x}. \tag{9.8}$$

At this point we can use that, for a Hamiltonian system, the divergence of $\mathbf{v}$ vanishes:

$$\mathrm{div}\ \mathbf{v}_H = \frac{\partial}{\partial \mathbf{q}}\left(\frac{\partial H}{\partial \mathbf{p}}\right) + \frac{\partial}{\partial \mathbf{p}}\left(-\frac{\partial H}{\partial \mathbf{q}}\right) = 0. \tag{9.9}$$

From this it follows that the Hamiltonian phase flow is volume-conserving: $\mu(t) = \mu(s)$. □

A.2 Poisson bracket formalism

Let me introduce the notion of the *Poisson bracket* $\{\cdot,\cdot\}$. It is a mathematical structure which is closely connected to the symplectic two-form ω. Often we can use one notion instead of the other. Let me in what follows show how they are related.

Let f and g be two smooth functions on phase space Γ and q^i, p_i a set of canonical coordinates on Γ. Then the Poisson bracket of f and g is defined to be

$$\{f,g\} = \sum_{i=1}^{n} \frac{\partial f}{\partial q^i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q^i}. \tag{9.10}$$

As such the Poisson bracket is skew-symmetric, bilinear, and it fulfills the Jacobi identity.[64]

The Poisson bracket is a convenient tool for the study of Hamiltonian systems because of the following properties. Consider a Hamiltonian system with Hamiltonian H on Γ. This Hamiltonian determines a Hamiltonian vector field as follows: $\mathrm{d}q^i/\mathrm{d}t = \partial H/\partial p_i$, $\mathrm{d}p_i/\mathrm{d}t = -\partial H/\partial q^i$. Now the time derivative of a (possibly time-dependent) smooth function g can be expressed in

[64] For this definition as well as for the following statements and results, cf. Scheck [2003].

terms of the Poisson bracket as follows:

$$
\begin{aligned}
\frac{\mathrm{d}g}{\mathrm{d}t} &= \frac{\partial g}{\partial t} + \sum_{i=1}^{n} \left(\frac{\partial g}{\partial q^i} \frac{\mathrm{d}q^i}{\mathrm{d}t} + \frac{\partial g}{\partial p_i} \frac{\mathrm{d}p_i}{\mathrm{d}t} \right) \\
&= \frac{\partial g}{\partial t} + \sum_{i=1}^{n} \left(\frac{\partial g}{\partial q^i} \frac{\partial H}{\partial p_i} + \frac{\partial g}{\partial p_i} \left(- \frac{\partial H}{\partial q^i} \right) \right) = \frac{\partial g}{\partial t} + \{g, H\}.
\end{aligned}
\tag{9.11}
$$

Assume g is not explicitly time-dependent: $\partial g/\partial t = 0$. Then g is invariant under time evolution if and only if

$$\{g, H\} = 0. \tag{9.12}$$

If this is the case, then g is called a *first integral* of motion. It is a conserved quantity of the dynamics.

We can now state the Hamiltonian equations of motion in terms of the Poisson brackets. They read

$$\frac{\mathrm{d}q^i}{\mathrm{d}t} = \{q^i, H\}, \qquad \frac{\mathrm{d}p_i}{\mathrm{d}t} = \{p_i, H\}. \tag{9.13}$$

The Poisson bracket also serves to define the canonical conjugate. Let again f and g be two smooth functions on Γ. A function f is called the canonical conjugate of g if and only if

$$\{f, g\} = 1. \tag{9.14}$$

Now a set of local coordinates is called canonical if and only if the following canonical Poisson bracket relations hold. For all $i, j = 1, ..., n$:

$$\{q^i, q^j\} = \{p_i, p_j\} = 0, \quad \{p_j, q^i\} = \delta^i_j. \tag{9.15}$$

All these properties of the Poisson bracket indicate that there must be a close connection to the symplectic two-form ω. In what follows, let us make this explicit.

Let again f and g be two differentiable functions on phase space Γ. Connected to any differentiable function f, you can uniquely define a (Hamiltonian) vector field X_f via the relation

$$\omega(X_f, \cdot) = \mathrm{d}f. \tag{9.16}$$

Here the notion of "Hamiltonian" does not refer to the physical Hamiltonian H from above. Instead it is the given mathematical relation which makes the vector field a Hamiltonian vector field (and the physical vector field X_H is a special case of this). In fact, the notion of a (Hamiltonian) vector field X_f in a symplectic space is analogous to the notion of the gradient ∇f in Euclidean space. Both are connected via the symplectic matrix $\mathcal{I}$. With respect to this matrix, X_f can be written as follows:

$$X_f = \mathcal{I} \nabla f \tag{9.17}$$

where

$$\mathcal{I} = \begin{pmatrix} 0 & \mathbb{I} \\ -\mathbb{I} & 0 \end{pmatrix} \tag{9.18}$$

and $\mathcal{I}$ is the identity matrix. Hence,

$$X_f = \sum_{i=1}^{n} \frac{\partial f}{\partial p_i} \frac{\partial}{\partial q^i} + \sum_{i=1}^{n} \Big(- \frac{\partial f}{\partial q^i} \Big) \frac{\partial}{\partial p_i}.$$

Let now X_f be the (Hamiltonian) vector field corresponding to f and X_g the (Hamiltonian) vector field corresponding to g. Then

$$\{f, g\} = \omega(X_g, X_f). \tag{9.19}$$

This relation follows directly from the above definitions.

The symplectic two-form also relates to the Lie derivative as follows. For two smooth functions f and g in Γ,

$$\omega(X_g, X_f) = L_{X_f} g \tag{9.20}$$

where $L_{X_f} g = \mathrm{d}g(X_f)$. This follows again directly from the definitions. Hence, the Poisson bracket relates to the Lie derivative as follows:

$$\{f, g\} = L_{X_f} g. \tag{9.21}$$

This means that a smooth function g on phase space is a first integral of motion if and only if its Lie derivative along the Hamiltonian vector field vanishes:

$$\{g, H\} = L_{X_H} g = 0. \tag{9.22}$$

When we introduced the notion of the Lie derivative above, we already saw that this is the correct condition showing invariance under time-evolution (respectively under the Hamiltonian phase flow).

Note also that the existence of a set of canonical variables (q^i, p_j) fulfilling the canonical Poisson bracket relations is equivalent to the existence of a canonical symplectic two-form $\omega = \sum \mathrm{d}q^i \wedge \mathrm{d}p_i$. This can be checked by direct computation. From (8.19) and (8.16), it follows that

$$\begin{aligned} \{q^i, q^j\} = \omega(X_{q^j}, X_{q^i}) &= \mathrm{d}q^j(X_{q^i}) = \mathrm{d}q^j \Big(- \frac{\partial}{\partial p_i} \Big) = 0, \\ \{p_i, p_j\} = \omega(X_{p_j}, X_{p_i}) = \mathrm{d}p_j(X_{p_i}) &= \mathrm{d}p_j \Big(\frac{\partial}{\partial q^i} \Big) = 0, \end{aligned} \tag{9.23}$$

and

$$\{p_j, q^i\} = \omega(X_{q^i}, X_{p_j}) = \mathrm{d}q^i(X_{p_j}) = \mathrm{d}q^i \Big(\frac{\partial}{\partial q^j} \Big) = \delta^i_j. \tag{9.24}$$

Appendix B: Internal Hamiltonian description and invariant measure for the minisuperspace model

In this appendix we discuss a simple relativistic model of the universe which features a measure which is invariant under internal time evolution: the minisuperspace model. This is another example for the internal Hamiltonian formulation and the construction of the measure presented in Section 3.5.

This model and the respective measure has been discussed for different proposes by Gibbons, Hawking, and Stuart [1987], Gibbons and Turok [2008], Carroll and Tam [2010], and Carroll [2014]. In contrast to these contributions, I will show the way in which the measure arises within an internal Hamiltonian description and I will say again why it is indeed the correct measure for the statistical analysis of the system. Moreover, I will present two different ways to obtain the measure: on the one hand by constructing it from the symplectic form (the way it is done by the above authors), on the other hand using the Faddeev-Popov construction (cf. Section 4.3).

Introduction to the model. The minisuperspace model is the simplest possible relativistic cosmological model. It describes a universe that is homogeneous and isotropic. As such, it features only two physical variables, the scale factor a representing the total extension of the universe and the scalar field Φ representing the homogeneous and isotropic mass distribution. To formulate the dynamics, we need two more variables. These are, in the Hamiltonian formulation of the theory, the canonical momenta p_a and p_Φ.[65]

While the configurational variables a and Φ provide the coordinates of two-dimensional configuration space $\tilde{\mathcal{Q}}$, a and Φ together with their canonical conjugates p_a and p_Φ form a set of local coordinates of four-dimensional phase space $\tilde{\Gamma} = T^*\tilde{\mathcal{Q}}$.

On $T^*\tilde{\mathcal{Q}}$, there exists a natural one-form

$$\tilde{\theta} = p_a \mathrm{d}a + p_\Phi \mathrm{d}\Phi \tag{9.25}$$

and a symplectic two-form $\tilde{\omega} = -\mathrm{d}\tilde{\theta}$:

$$\tilde{\omega} = \mathrm{d}a \wedge \mathrm{d}p_a + \mathrm{d}\Phi \wedge \mathrm{d}p_\Phi. \tag{9.26}$$

In addition, there exists a Hamiltonian constraint

$$\mathcal{H} = -\frac{p_a^2}{12a} + \frac{p_\Phi^2}{2a^3} + a^3 V(\Phi) - 3a\kappa = 0. \tag{9.27}$$

Here $V = V(\Phi)$ is some potential and κ is the spatial curvature of the particular model. The Hamiltonian constraint defines a three-dimensional surface $\Sigma \subset \tilde{\Gamma}$ to which any physical trajectory is restricted. Together with the symplectic form, $\mathcal{H}$ determines the physical vector field $\tilde{X}$

[65] For the description of this model, including the Hamiltonian constraint (8.27) and the expression of the Hubble constant in terms of the canonical variables (8.29), see, e.g., Carroll and Tam [2010].

via

$$\tilde{\omega}\big|_{\mathcal{H}=0}(\tilde{X},\cdot) = \mathrm{d}\mathcal{H}. \tag{9.28}$$

To obtain the internal Hamiltonian description, we still need some function $f(q^a, p_a)$ that serves as an internal time variable τ.

Hubble parameter H as internal time τ. Let us consider the Hubble constant H and let us check whether it can be taken as an internal time parameter. In terms of the Hamiltonian coordinates, H can be written as

$$H = -\frac{p_a}{6a^2}. \tag{9.29}$$

Is this a monotonic parameter? To see this, let us compute the Poisson bracket of H and $\mathcal{H}$ and evaluate it at the constraint surface $\mathcal{H} = 0$.

Lemma 9.1. *Let $\mathcal{H}$ and H on $\tilde{\Gamma}$ as defined by (8.27) and (8.29). Then*

$$\{H, \mathcal{H}\}\big|_{\mathcal{H}=0} = -\frac{p_\Phi^2}{2a^6} + \frac{\kappa}{a^2} < 0 \quad \text{iff} \quad \kappa < 0. \tag{9.30}$$

Proof. Let us compute the Poisson bracket of H and $\mathcal{H}$. There all derivatives with respect to Φ and p_Φ vanish because $H = H(a, p_a)$. Hence,

$$\begin{aligned}
\{H, \mathcal{H}\} &= \frac{\partial H}{\partial a}\frac{\partial \mathcal{H}}{\partial p_a} - \frac{\partial H}{\partial p_a}\frac{\partial \mathcal{H}}{\partial a} \\
&= \frac{p_a}{3a^3}\Big(-\frac{p_a}{6a}\Big) - \Big(-\frac{1}{6a^2}\Big)\Big(\frac{p_a^2}{12a^2} - \frac{3p_\Phi^2}{2a^4} + 3a^2V(\Phi) - 3\kappa\Big) \\
&= -\frac{p_a^2}{24a^4} - \frac{p_\Phi^2}{4a^6} + \frac{V(\Phi)}{2} - \frac{\kappa}{2a^2} = \frac{\mathcal{H}}{2a^3} - \frac{p_\Phi^2}{2a^6} + \frac{\kappa}{a^2}.
\end{aligned} \tag{9.31}$$

It follows that

$$\{H, \mathcal{H}\}\big|_{\mathcal{H}=0} = -\frac{p_\Phi^2}{2a^6} + \frac{\kappa}{a^2}.$$

Since the first term is less or equal to zero, the whole expression is less than zero if and only if $\kappa < 0$. This shows the assertion. □

We have shown that for models with negative spatial curvature $\kappa < 0$, the Hubble constant is strictly monotonic and as such provides a parametrization of the trajectories: $\tau = H$. That is, there exists a family of hypersurfaces Γ_τ of constant $H = H^*$ (respectively, $\tau = \tau^*$) foliating the space $\tilde{\Gamma}\big|_{\mathcal{H}=0}$ on which the trajectories lie such that each Γ_τ is cut once and only once by each of the trajectories. Each of the Γ_τ serves as a space of solutions on which the internal measure can be constructed.

What is the canonical conjugate F of the Hubble parameter, generating the motion with respect to internal time $\tau = H$?

Lemma 9.2. *Let everything as above, in particular $\tau = H$. Then*

$$F = -2a^3 \tag{9.32}$$

is the internal Hamiltonian. Here $a = a(q^*, p^*, \tau)$ is a function of the internal variables $q^*, p^* \in \{a, \Phi, p_a, p_\Phi\}$ and τ by help of the Hamiltonian constraint $\mathcal{H} = 0$.

Proof. We find F by demanding that $\{F, H\} = 1$. Clearly,

$$\left\{ -2a^3, -\frac{p_a}{6a^2} \right\} = 1.$$

In order for $-2a^3$ to be the internal Hamiltonian F, we need to express it in terms of the internal coordinates. This depends on our choice of coordinates $q^*, p^* \in \{a, \Phi, p_a, p_\Phi\}$. $\square$

Let us, in what follows, discuss two different choices of coordinates.

Volume measure for internal variables Φ and p_Φ. We want to find the measure on the space of solutions expressed in terms of local coordinates Φ and p_Φ. With respect to these coordinates, the internal Hamiltonian is $F = -2a^3$ where a is a function of the internal variables, $a = a(\phi, p_\phi)$, by help of the Hamiltonian constraint $\mathcal{H} = 0$. Let us try to write down F explicitly. Therefore let us solve the Hamiltonian constraint for a with $p_a = -6a^2 H$. It is

$$\mathcal{H}\big|_{p_a=-6a^2H} = -\frac{36a^4H^2}{12a} + \frac{p_\Phi^2}{2a^3} + a^3V(\Phi) - 3a\kappa = 0.$$

This can be rewritten as

$$[2V(\Phi) - 6H^2]a^6 - 6\kappa a^4 + p_\Phi^2 = 0.$$

Unfortunately, solving this equation amounts to solving a polynomial of the third order which means that the first solution has to be guessed. But this is a hopeless task. In what follows, let us assume that there exists some internal Hamiltonian $F(\Phi, p_\Phi) = -2a^3(\Phi, p_\phi)$.

Let us now determine the invariant measure on the internal space (the space of solutions) in terms of Φ and p_Φ.

Lemma 9.3. *Let everything as above. Let $\Gamma = \{(q,p) \in \Sigma | \tau(q,p) = \tau^*\}$ and $\Sigma = \{(q,p) \in \tilde{\Gamma} | \mathcal{H}(q,p) = 0\}$. Let Φ, p_Φ local coordinates on Γ. The natural volume measure μ on Γ (as defined by (3.6)) is given by*

$$\mathrm{d}\mu = \mathrm{d}\Phi \mathrm{d}p_\Phi. \tag{9.33}$$

This measure is invariant under internal time evolution.

Proof. Let $i_\mathcal{H} : \Sigma \to \Gamma$ denote the embedding of Σ in Γ where Σ is defined by constant $\mathcal{H} = 0$ and $i_\tau : \Gamma \to \Sigma$ the embedding of Γ_τ in Σ where Γ is defined by constant $H = H_*$. From (8.29) we get that

$$\mathrm{d}p_a = -6a^2\mathrm{d}H - 12aH\mathrm{d}a.$$

This we can use to rewrite the two-form $\tilde{\omega}$ on $\tilde{\Gamma}$ as follows:

$$\begin{aligned}\tilde{\omega} &= \mathrm{d}a \wedge \mathrm{d}p_a + \mathrm{d}\Phi \wedge \mathrm{d}p_\Phi \\ &= -6a^2 \mathrm{d}a \wedge \mathrm{d}H + \mathrm{d}\Phi \wedge \mathrm{d}p_\Phi\end{aligned}$$

where we used that $\mathrm{d}a \wedge \mathrm{d}a = 0$. Now the pullback of $\tilde{\omega}$ on $\tilde{\Gamma}$ to Σ is

$$i_{\mathcal{H}}^* \tilde{\omega} = \mathrm{d}F \wedge \mathrm{d}H + \mathrm{d}\Phi \wedge \mathrm{d}p_\Phi$$

where $F = F(\Phi, p_\Phi)$ is the internal Hamiltonian. Finally, the pullback of $i_{\mathcal{H}}^* \tilde{\omega}$ on Σ to Γ_τ, that is, the form $\omega = i_\tau^* i_{\mathcal{H}}^* \tilde{\omega}$, is

$$\omega = \mathrm{d}\Phi \wedge \mathrm{d}p_\Phi$$

This two-form is symplectic and, since it is a top-dimensional form on Γ_τ, it is already the final volume form Ω on Γ_τ. Hence, we can directly write down the volume measure $\mu = |\Omega|$. It is

$$\mathrm{d}\mu = \mathrm{d}\Phi \mathrm{d}p_\Phi.$$

This shows (8.33).

This measure is invariant under internal time evolution by construction. Remember that it is the natural volume measure on the internal space, constructed from the underlying symplectic two-form. Since the dynamics is Hamiltonian (with internal Hamiltonian F), the measure is invariant under internal time evolution (with internal time $\tau = H$). The invariance has been proven in Lemma 3.7. □

Note that the measure we obtained is just the uniform measure in Φ and p_Φ. It is invariant under internal time evolution where the internal dynamics on Γ is governed by the internal Hamiltonian $F = F(\Phi, p_\Phi)$.

Volume measure for internal variables a and Φ. Let us now determine the measure in terms of a and Φ.

Lemma 9.4. *Let everything as in the preceding lemma, only now a and Φ are local coordinates on Γ. Then the natural volume measure μ on Γ is given by*

$$\mathrm{d}\mu = \frac{|18a^5 H_*^2 - 6a^5 V(\Phi) + 12a^3\kappa|}{\sqrt{6a^6 H_*^2 - 2a^6 V(\Phi) + 6a^4\kappa}} \, \mathrm{d}a\mathrm{d}\Phi. \tag{9.34}$$

Again, this measure is invariant under internal time evolution.

Proof. Let $i_{\mathcal{H}} : \Sigma \to \Gamma$ denote the embedding of Σ in Γ where Σ is defined by constant $\mathcal{H} = 0$ and $i_\tau : \Gamma_\tau \to \Sigma$ the embedding of Γ_τ in Σ where Γ_τ is defined by constant $H = H_*$. Let us use the constraint $\mathcal{H} = 0$ to replace p_Φ. In that case, the pullback $i_{\mathcal{H}} \tilde{\theta} = \tilde{\theta}\big|_{\mathcal{H}=0}$ can be written as

$$i_{\mathcal{H}} \tilde{\theta} = p_a \mathrm{d}a + \sqrt{\frac{p_a^2 a^2}{6} - 2a^6 V(\Phi) + 6a^4\kappa} \; \mathrm{d}\Phi$$

where we get the expression for p_Φ from solving the Hamiltonian constraint $\mathcal{H} = 0$:

$$p_\Phi\big|_{\mathcal{H}=0} = \sqrt{\frac{p_a^2 a^2}{6} - 2a^6 V(\Phi) + 6a^4\kappa}.$$

(Here we are not interested in the sign in front of the square root because we will later on consider the absolute value anyway.)

Connected to this one-form, there exists a degenerate two-form

$$\begin{aligned} i_{\mathcal{H}}^*\tilde{\omega} = -\mathrm{d}(i_{\mathcal{H}}^*\tilde{\theta}) &= \mathrm{d}a \wedge \mathrm{d}p_a + \frac{1/3 p_a a^2}{2\sqrt{1/6 p_a^2 a^2 - 2a^6 V(\Phi) + 6a^4\kappa}} \mathrm{d}\Phi \wedge \mathrm{d}p_a \\ &+ \frac{1/3 p_a^2 a - 12a^5 V + 24a\kappa}{2\sqrt{1/6 p_a^2 a^2 - 2a^6 V(\Phi) + 6a^4\kappa}} \mathrm{d}\Phi \wedge \mathrm{d}a, \end{aligned}$$

where we used that $\mathrm{d}a \wedge \mathrm{d}a = 0$.

Let us now determine the pullback of $i_{\mathcal{H}}^*\tilde{\omega}$ to Γ. Let us use the constraint $H = H_*$ to replace p_a. From $H = -p_a/6a^2$, we get $\mathrm{d}p_a = -6a^2\mathrm{d}H - 12aH\mathrm{d}a$. On Γ, this turns into $\mathrm{d}p_a = -12aH^*\mathrm{d}a$. Moreover, $p_a|_{H=H^*} = -6a^2H^*$. That is, the two-form $\omega = i_\tau^* i_{\mathcal{H}}^*\tilde{\omega}$ on Γ can be written as

$$\omega = \frac{18a^5 H_*^2 - 6a^5 V(\Phi) + 12a^3\kappa}{\sqrt{6a^6 H_*^2 - 2a^6 V(\Phi) + 6a^4\kappa}}\, \mathrm{d}a \wedge \mathrm{d}\Phi.$$

This two-form is symplectic (it inherits the symplectic structure of $\tilde{\Gamma}$) and since it is a top-dimensional form on Γ it is already the final volume form Ω on Γ. Hence, we can directly write down the volume measure $\mu = |\Omega|$. It is

$$\mathrm{d}\mu = \frac{|18a^5 H_*^2 - 6a^5 V(\Phi) + 12a^3\kappa|}{\sqrt{6a^6 H_*^2 - 2a^6 V(\Phi) + 6a^4\kappa}}\, \mathrm{d}a\mathrm{d}\Phi.$$

For the same reason as in the preceding lemma, this volume form is invariant under internal time evolution (cf. the proof given there). □

Remark (Non-conjugate internal variables)**.** How do we interpret this measure which is formulated with respect to a pair of non-conjugate internal variables a and Φ? For sure, the internal Hamiltonian description will be of a peculiar form. Let us look at this in more detail.

With respect to the internal variables a and Φ, the internal Hamiltonian is

$$F = -2a^3. \tag{9.35}$$

We know that F is a conserved quantity of the internal equations of motion. This follows from the fact that F is the internal Hamiltonian, the canonical conjugate of internal time $H = \tau$ (cf. Lemma 8.2). From the fact that F is conserved, it follows that a is conserved. In other words, a is constant on any hypersurface of constant $H = H^*$. Mathematically, this is reflected in the special form of the internal physical vector field. From $\omega(X_F, \cdot) = \mathrm{d}F$ with $F = -2a^3$ we get the internal Hamiltonian vector field

$$X_F = 0\frac{\partial}{\partial a} + \frac{\sqrt{6a^6H_*^2 - 2a^6V(\Phi) + 6a^4\kappa}}{|18a^5H_*^2 - 6a^5V(\Phi) + 12a^3\kappa|} \cdot 6a^2\frac{\partial}{\partial \Phi}. \tag{9.36}$$

We see that as internal time evolves there is no change in a, but a considerable change in Φ.

However, a is not constant along the actual trajectories in $\tilde{\Gamma}$. This can be seen most easily when we reintroduce external time: a is not constant with respect to external time evolution. Let us consider external time t and let $\mathcal{H}$ (apart from formulating the Hamiltonian constraint) determine Hamiltonian equations with respect to external time t. This is the textbook presentation of the minisuperspace model which we have not presented so far. In that formulation the Hamiltonian law tells us that

$$\mathrm{d}a/\mathrm{d}t = \partial\mathcal{H}/\partial p_a \neq 0. \tag{9.37}$$

That is, a changes from one surface of constant $H = H^*$ to another. If a is not conserved, then also $F = -2a^3$ is not conserved with respect to external time evolution.

How do the external and internal Hamiltonian description fit together? In particular, what about the measures which are preserved under external and internal time evolution, respectively? Or, to put it differently, what does this mean for the statistical analysis of the system when we consider a macro-partition involving the scale factor a?

By construction, the internal measure (8.34) is invariant as it is transported along the vector field (8.36). Nevertheless, the internal Hamiltonian equations no longer reflect the actual evolution of the scale factor a along the actual trajectories. This just means that we have to be careful about the dynamical interpretation. The internal measure (8.34) is then simply a measure on possible initial conditions expressed in terms of a and Φ at time $H = H^*$. This measure is a measure on initial data compatible with the Hamiltonian constraint $\mathcal{H} = 0$ at time $H = H^*$. At any other moment, $H = H^{**}$, the measure is again the correct measure on initial data at that very moment.

The measure obtained from the Faddeev construction. In the preceding paragraphs, we constructed the natural volume measure μ on the internal space Γ from the underlying symplectic two-form. There exists an alternative way to construct that measure, namely by the method of Faddeev and Popov described in Section 4.3. Let us do this here for the choice of internal variables a and Φ.

Since $\mathcal{H}$ and H form a pair of second class constraints, $\{\mathcal{H}, H\} \neq 0$, the formula of Faddeev (4.66) can be used. Explicitly, the following can be shown.

Lemma 9.5. *Let everything as above, let a and Φ local coordinates on Γ. Let μ given by (8.34). The measure σ on Γ obtained from the Faddeev formula (4.66) is*

$$\sigma = \int \frac{|18a^5H_*^2 - 6a^5V(\Phi) + 12a^3\kappa|}{\sqrt{6H_*^2a^6 - 2a^6V(\Phi) + 6a^4\kappa}}\mathrm{d}a\mathrm{d}\Phi. \tag{9.38}$$

That is, we obtain that $\sigma = \mu$.

Proof. According to the formula of Faddeev,

$$\sigma = \int |\{\mathcal{H}, H\}|\delta(\mathcal{H}-0)\delta(H-H^*)\mathrm{d}a\mathrm{d}p_a\mathrm{d}\Phi\mathrm{d}p_\Phi.$$

From (8.31) we know that the Poisson bracket of $\mathcal{H}$ and H is

$$\{\mathcal{H}, H\} = \frac{p_a^2}{24a^4} + \frac{p_\Phi^2}{4a^6} - \frac{V(\Phi)}{2} + \frac{\kappa}{2a^2}.$$

Let us evaluate this on the constraint surface where the constraints are given by $p_\Phi = p_\Phi^*$ with $p_\Phi^* = p_\Phi\big|_{\mathcal{H}=0}$ and $p_a = p_a^*$ with $p_a^* = p_a\big|_{H=H_*}$. Explicitly,

$$p_\Phi^* = \sqrt{\frac{(p_a^*)^2a^2}{6} - 2a^6V(\Phi) + 6a^4\kappa}, \qquad p_a^* = -6a^2H_*.$$

It follows that

$$\{\mathcal{H}, H\}\big|_{p_\Phi^*, p_a^*} = 3H_*^2 - V(\Phi) + \frac{2\kappa}{a^2}.$$

Using this equation, the measure becomes

$$\begin{aligned}\sigma &= \int |\{\mathcal{H}, H\}|\delta(\mathcal{H}-0)\delta(H-H^*)\mathrm{d}a\mathrm{d}p_a\mathrm{d}\Phi\mathrm{d}p_\Phi \\ &= \int \frac{|\{\mathcal{H}, H\}|\delta(p_\Phi - p_\Phi^*)\delta(p_a - p_a^*)}{|\partial\mathcal{H}/\partial p_\Phi||\partial H/\partial p_a|}\mathrm{d}a\mathrm{d}p_a\mathrm{d}\Phi\mathrm{d}p_\Phi \\ &= \int \frac{|18a^5H_*^2 - 6a^5V(\Phi) + 12a^3\kappa|}{\sqrt{6H_*^2a^6 - 2a^6V(\Phi) + 6a^4\kappa}}\mathrm{d}a\mathrm{d}\Phi.\end{aligned}$$

This shows (8.38). Since (8.38) coincides with (8.34), it follows that $\sigma = \mu$. □

Discussion. Of course, a measure on the set of solutions of the minisuperspace model can tell us little about the real world. This is simply due to the fact that a homogeneous and isotropic model of the universe is a very simplified picture of our world. This model can not even describe one of the most prominent features of our universe, namely a clumping of matter resembling the galaxies we observe. Since the model includes only four variables, the scale factor a and the scalar field Φ together with their canonical conjugates, there are very few ways to specify a macrostate – a state of the system determining a region in phase space we can then analyze with the canonical volume measure. The only two questions which have been addressed so far – and which are maybe the only questions the measure can address at all – are the questions of whether or not inflation is likely to occur (cf. Gibbons and Turok [2008] and Carroll and Tam [2010]) and whether or not the universe is likely to be flat (cf. Carroll [2014]).

Over and above the problem that the minisuperspace model provides a too simplified picture of our world, the measure in itself allows only for a very restricted statistical analysis due to the fact that it is non-normalizable. While we can find an upper bound on Φ based on certain considerations which I do not want to discuss here (see, for instance, Schiffrin and Wald [2012]), the scale factor a is necessarily unbounded. This means that the total measure of internal phase space is infinite.

What about the two explicit questions, the problem of inflation and the flatness problem, which have been addressed by this measure in the past? It turns out that the measure is ambiguous about the likeliness of inflation due to the fact that inflation depends on the value of Φ. In that case a is not restricted and the measure diverges in a. It has been shown by Schiffrin and Wald [2012] that the procedure of regularization by a cut-off can lead to opposite results (like in the case of Gibbons and Turok [2008] and Carroll and Tam [2010]). On the other hand, the measure is definite about the likeliness of flatness because flatness depends on the value of a. The set of non-flat universes, where a is small as compared to κ, has finite measure (cf. Carroll [2014]). Although the total measure of internal phase space (the space of solutions) is infinite, one can correctly conclude that non-flat universes are atypical whereas flat universes are typical. Hence, Carroll is correct when he says that the flatness problem is solved. It is a typical feature of a universe described within the minisuperspace model.

Appendix C: Classification of motion and final evolution of the Newtonian gravitational system

In this appendix, we present the most important known results on the dynamics of the Newtonian gravitational system. It is based on Landau and Lifshitz [1967]. In the end, we cite recent results of Marchal and Saari about the final evolution of the Newtonian N-body system.

Mechanical similarity. Consider a dynamical system (Lagrangian or Hamiltonian) of N particles, each with coordinate $\mathbf{q}_i$ and velocity $\mathbf{v}_i$. Let the system have total energy $E = T + U$, where $T = \sum_i \frac{1}{2} m_i \mathbf{v}^2$ is the kinetic energy and where the potential energy U is a function homogeneous of degree k in the coordinates, i.e., for some constant α,

$$U(\alpha\mathbf{q}_1, \alpha\mathbf{q}_2, ..., \alpha\mathbf{q}_N) = \alpha^k U(\mathbf{q}_1, \mathbf{q}_2, ...\mathbf{q}_N). \tag{9.39}$$

For this system, there exist trajectories (solutions to the equations of motion) that are geometrically similar, but different in size. They can be turned into one another by simultaneously changing all the positions $\mathbf{q}_i$, velocities $\mathbf{v}_i = \dot{\mathbf{q}}_i$ and time t in an appropriate manner. Let the $\mathbf{q}_i$ and t transform as follows: $\mathbf{q}_i \to \alpha\mathbf{q}_i, t \to \beta t$ for some positive α, β. Then $\mathbf{v}_i \to \frac{\alpha}{\beta}\mathbf{v}_i$ and $T \to \frac{\alpha^2}{\beta^2} T$. By assumption the potential energy transforms as $U \to \alpha^k U$. If now $\frac{\alpha^2}{\beta^2} = \alpha^k$, i.e. if $\beta = \alpha^{1-\frac{1}{2}k}$, then the Lagrangian, respectively the Hamiltonian, changes by a constant factor $\frac{\alpha^2}{\beta^2}$. This means that the equations of motion are left unchanged. Hence, the above transformation describes trajectories that are geometrically the same, though different in size and run through by a different speed. This feature of a dynamical system is called *mechanical similarity.*

You can read the above transformation also as a simultaneous transformation of time t and initial conditions $(\mathbf{q}_i, \mathbf{p}_i)$ of a Hamiltonian system. Mechanically similar trajectories are attained by the following global transformation of these $6N+1$ variables:

$$\mathbf{q}_i \to \alpha \mathbf{q}_i, \qquad t \to \beta t, \qquad \mathbf{p}_i \to \frac{\alpha}{\beta}\mathbf{p}_i, \tag{9.40}$$

for all $i = 1, ..., N$ and for some positive α and $\beta = \alpha^{1-\frac{1}{2}k}$.

Remark (Application to Newton)**.** Let us consider an N-particle system governed by the Newtonian gravitational force law. Each particle has a position $\mathbf{q}_i$ and a velocity $\mathbf{v}_i$ and the potential energy U has the form

$$U = -\sum_{i<j;i,j=1}^{N} \frac{Gm_i m_j}{|\mathbf{q}_i - \mathbf{q}_j|}. \tag{9.41}$$

Thus, U is a function homogeneous of degree $k=-1$ in which case mechanical similarity is given if and only if $\beta = \alpha^{\frac{3}{2}}$. Hence, for the Newton potential, mechanically similar trajectories are related to each other by the following global transformation. For all $i = 1, ..., N$ and for some positive, real-valued l:

$$\mathbf{q}_i \to \frac{1}{l^2}\mathbf{q}_i, \qquad t \to \frac{1}{l^3}t \quad , \mathbf{v}_i \to l\mathbf{v}_i. \tag{9.42}$$

Virial theorem. In order to prove the virial theorem, we need to show Euler's homogeneous function theorem first. It provides a useful mathematical identity for functions $f = f(x_1, .., x_n)$ that are homogeneous of degree k, i.e. for functions of the above form,

$$f(\alpha x_1, ..., \alpha x_n) = \alpha^k f(x_1, ..., x_n)$$

for some constant α. Euler's homogeneous function theorem connects the gradient ∇f to the function f as follows:

Theorem 9.2 (Euler's homogeneous function theorem)**.** *Let $f(\alpha x_1, ..., \alpha x_n) = \alpha^k f(x_1, ...x_n)$, i.e., f is a homogeneous function of degree k. Then*

$$\sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i} = kf. \tag{9.43}$$

Proof. Let $x_i' := \alpha x_i \; \forall i = 1, ..., n$. Then

$$k\alpha^{k-1} f(x_1, ...x_n) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i'} \frac{\partial x_i'}{\partial \alpha} = \sum_{i=1}^{n} x_i \frac{\partial f}{\partial x_i'} = \sum_{i=1}^{n} x_i \frac{\partial f}{\partial (\alpha x_i)}.$$

For $\alpha = 1$, the assertion follows. □

Let us consider a dynamical system of N particles with masses $m_1, ..., m_N$, respectively. Let $\mathbf{q}_i$ denote the coordinate of the i-th particle, $\mathbf{v}_i = \dot{\mathbf{q}}_i$ its velocity, and $\mathbf{p}_i = m_i \dot{\mathbf{q}}_i$ its momentum. Let the motion of the system be governed by a potential $U = U(\mathbf{q}_1, .., \mathbf{q}_N)$, where according to the Newtonian force law

$$\dot{\mathbf{p}}_i = -\frac{\partial U}{\partial \mathbf{q}_i}. \tag{9.44}$$

The virial theorem holds whenever the potential U is a function homogeneous of degree k of the coordinates,

$$U(\alpha \mathbf{q}_1, ..., \alpha \mathbf{q}_n) = \alpha^k U(\mathbf{q}_1, ...\mathbf{q}_n),$$

and when, in addition, the function

$$D := \sum_i \mathbf{p}_i \mathbf{q}_i, \tag{9.45}$$

corresponding to what has historically been called the *virial*, is bounded. This is, for example, the case when the system is confined to a box of finite volume and when the momenta cannot become arbitrarily large (e.g. by coupling the system to a heat bath).

These conditions are true for several subsystems of the universe, but typically fail to hold for the universe itself. In particular, D is not bounded for any open model of the universe (where the total volume of space is infinite) nor for a self-gravitating system of point particles (whether confined to a box or not) without a short distance cut-off (where the kinetic energy can increase without bound by the formation of tight binaries or a core-halo structure).

Explicitly, the virial theorem relates the mean kinetic energy $\bar{T}$ to the mean potential energy $\bar{U}$ as follows:

Theorem 9.3 (Virial theorem)**.** *Let $\boldsymbol{q}_i, \boldsymbol{v}_i, \boldsymbol{p}_i$ as defined above. Let again T the kinetic energy, U the potential energy, and $E = T + U$ the total energy of the system. Let $U(\alpha \boldsymbol{q}_1, ..., \alpha \boldsymbol{q}_N) = \alpha^k U(\boldsymbol{q}_1, ...\boldsymbol{q}_n)$, that is, the potential U is a function homogeneous of degree k. In addition, let $D = \sum_i \boldsymbol{p}_i \boldsymbol{q}_i$ be bounded. Then, for the mean kinetic and potential energy,*

$$2\bar{T} = k\bar{U}, \tag{9.46}$$

where the average is taken with respect to time, $\bar{f} = \lim\limits_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau f(t)dt$.

Proof. According to Euler's homogeneous function theorem,

$$\sum_{i=1}^N \mathbf{v}_i \frac{\partial T}{\partial \mathbf{v}_i} = 2T.$$

Now $\mathbf{p}_i = \partial T / \partial \mathbf{v}_i$ due to the form of T. Hence,

$$2T = \sum_{i=1}^N \mathbf{p}_i \mathbf{v}_i = \frac{d}{dt}\left(\sum_{i=1}^N \mathbf{p}_i \mathbf{q}_i\right) - \sum_{i=1}^N \mathbf{q}_i \dot{\mathbf{p}}_i.$$

Now take the time average of both sides of the equation, where the time average of a function f is defined by $\bar{f} = \lim\limits_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau f(t)dt$, and make use of the fact that, whenever the integrand f is the time derivative of a bounded function F, i.e. $f = dF/dt$, the time average vanishes:

$$\bar{f} = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \frac{dF}{dt} dt = \lim_{\tau \to \infty} \frac{F(\tau) - F(0)}{\tau} = 0.$$

According to our assumptions, $D = \sum_{i=1}^n \mathbf{p}_i \mathbf{q}_i$ is bounded and $\dot{\mathbf{p}}_i = -\partial U / \partial \mathbf{q}_i$ according to the Newtonian force law. Hence,

$$2\bar{T} = \overline{\sum_{i=1}^N \mathbf{q}_i \frac{\partial U}{\partial \mathbf{q}_i}} = k\bar{U},$$

where the last step again follows from Euler's homogeneous function theorem. □

Remark (Application to Newton). The Newtonian gravitational potential $U = -\sum_{i<j} \frac{G m_i m_j}{|\mathbf{q}_i - \mathbf{q}_j|}$ is a function homogenous of degree $k = -1$. In that case, the virial theorem states that the time average of the kinetic energy is twice the absolute value of the time average of the potential energy: $2\bar{T} = -\bar{U}$.

Remark (Virial equilibrium). Note that the virial theorem directly connects to systems at thermodynamic equilibrium. Since equilibrium refers to the macrostate corresponding to the macro-region of the by far largest phase space measure – that is, almost all possible states of the system are equilibrium states with the proportion of equilibrium states to all possible states being $|\Gamma_{eq}|/|\Gamma| \approx 1$ –, it must hold that the equilibrium kinetic energy T_{eq} and the equilibrium potential energy U_{eq} approximately correspond to the mean values:

$$T_{eq} \approx \bar{T}, \qquad U_{eq} \approx \bar{U}. \tag{9.47}$$

This regime is called the *virial equilibrium* of the system. It connects to a negative total energy $E = T_{eq} + U_{eq} \approx \bar{T} - 2\bar{T} = -\bar{T}$.

However, keep in mind that in order for the above theorem to hold D has to be bounded. This means that the total volume of phase space Γ has to be finite and, hence, every macroregion $\Gamma_M \subset \Gamma$ has a finite volume. This is a necessary condition in order to determine the mean value of T and U and relate it to the equilibrium value.

Lagrange-Jacobi equation. In what follows I will proof a relation which is fundamental to the dynamics of the Newtonian gravitational N-body system. It connects the moment of inertia I to the total energy E and the kinetic energy T of the system. The result is due to Lagrange and has been further developed by Jacobi.

Let us consider the same system as above, i.e. a Newtonian gravitational N-particle system with total energy $E = T + U$ and gravitational potential

$$U = - \sum_{i<j, i,j=1}^{N} \frac{G m_i m_j}{|\mathbf{q}_i - \mathbf{q}_j|}.$$

Let $M = \sum_i m_i$ denote the total mass of the system. The moment of inertia I is defined by

$$I = \sum_{i=1}^{N} m_i(\mathbf{q}_i - \mathbf{q}_{cm})^2 \tag{9.48}$$

where $\mathbf{q}_{cm} = \frac{1}{M}\sum_i m_i \mathbf{q}_i$ denotes the position of the center of mass of the system. Let us assume that the center of mass is located at the origin (which we may, without loss of generality, since the equations of motion of the Newtonian N-particle system are invariant under spatial translations). Thus, let in the following

$$\mathbf{q}_{cm} = \sum_{i=1}^{N} \mathbf{q}_i = 0. \tag{9.49}$$

We can now prove the Lagrange-Jacobi equation which relates the second derivative of I with the potential energy U and the kinetic energy T, respectively with the potential or kinetic energy and the total energy E of the system:

Theorem 9.4 (Lagrange-Jacobi equation)**.** *Let the total energy E, the potential energy U, and the moment of inertia I be as above. Then*

$$\ddot{I} = 4E - 2U. \tag{9.50}$$

Proof. Let the center of mass be located at the origin, $\mathbf{q}_{cm} = 0$ (which we may due to invariance of the equations of motion of the N-particle system under spatial translations). Then the moment of inertia can be written as

$$I = \sum_{i=1}^{N} m_i \mathbf{q}_i^2.$$

Take the first derivative, which is

$$\dot{I} = 2\sum_{i=1}^{N} m_i \mathbf{q}_i \dot{\mathbf{q}}_i = 2\sum_{i=1}^{N} \mathbf{q}_i \mathbf{p}_i = 2D, \tag{9.51}$$

and the second derivative, which is

$$\ddot{I} = 2\dot{D} = 2\sum_{i=1}^{N} m_i \dot{\mathbf{q}}_i^2 + 2\sum_{i=1}^{N} m_i \mathbf{q}_i \ddot{\mathbf{q}}_i = 4T + 2U. \tag{9.52}$$

Here the last equation follows from the Newtonian force law ($m_i \ddot{\mathbf{q}}_i = -\frac{\partial U}{\partial \mathbf{q}_i}$) and Euler's homogeneous function theorem (applied to the Newton potential U).

Rewriting this result with respect to the total energy E shows the assertion. □

The Lagrange-Jacobi equation determines the long-time evolution of the N-body system and draws a first picture of an evolving Newtonian universe. We will make this precise in the following.

Classification of motion. The above results help us to classify the motion of the Newtonian gravitational system. When we look for such a classification, the first idea is to try to classify the motion according to the sign of the total energy. Let us see, how far we can get by that.

Let again

$$R = \max_{i \neq j} |\mathbf{q}_i - \mathbf{q}_j|$$

denote the largest distance between two particles and

$$r = \min_{i \neq j} |\mathbf{q}_i - \mathbf{q}_j|$$

the smallest distance.

- $E > 0$: Let us first consider the case in which the total energy E is positive. Since the potential energy U is strictly negative, the Lagrange-Jacobi equation implies that, in this case, the second derivative of the moment of inertia I is strictly positive: $\ddot{I} > 0$. Explicitly, for $U < 0$: $\ddot{I} = 4E - 2U \geq 4E$ and with $E > 0$ the assertion follows. In that case, I is a function that is concave upwards. In other words, there exists a global minimum of I, $I = I_{min}$, and a point t_{min} such that for all times $t < t_{min}$, I is decreasing and for all times $t > t_{min}$, I is increasing. Explicitly, since $\ddot{I} > 4E$, it follows that $I \geq 4Et^2 + \mathcal{O}(t)$. In particular, from this together with the definition of R (the a largest distance between the particles) we get that $R \geq Ct + o(t)$ for some positive constant C.

  This gives us a first idea of an evolving Newtonian universe of positive total energy. Consider N particles distributed within infinite space. The evolution of the N-particle system must be like this: The particles come in from infinity, approach each other until, at some moment of time, they are closest to each other and then fly apart and off to infinity again.

- $E = 0$: For the case of zero total energy, there exists a result by Pollard [1967]. Pollard shows that $I \to \infty$ as $t \to \pm\infty$. To be precise it says that either there exist positive constants C_1, C_2 such that $C_1 t^{2/3} + o(t^{2/3}) \leq r \leq R \leq C_1 t^{2/3} + o(t^{2/3})$ or $R/t^{2/3} \to \infty$. Moreover, from the Lagrange-Jacobi-equation we know that I is concave upwards ($\ddot{I} = -2U > 0$). Hence, there again exists a global minimum of I, $I = I_{min}$, and I increases without bound in both directions away from that.

  This means that qualitatively we have the same scenario as above where $E > 0$. Particles come in from infinity, are closest to each other at some moment in time and spread again towards infinity.

- $E < 0$: Let us now consider the case of negative total energy. In contrast to before, this time the sign of the total energy does not determine the curvature of I – the second time derivative of I may be positive or negative or zero depending on the specific relation of T and U. Hence, no global behavior can be inferred. Explicitly, there exists a critical value $U_C = -2T$ such that for all $U > U_C$ (and $U < -T$ because otherwise the total energy is no longer negative), the second time derivative of I is positive, $\ddot{I} > 0$, and for all $U < U_C$, the second time derivative of I is negative, $\ddot{I} < 0$. If $U = U_C$, it holds that $\ddot{I} = 0$. Note that the critical value of U corresponds to virial equilibrium.

Final evolution of the N-body system. Due to the results of Saari [1971] we have an even more precise idea of the asymptotic behavior of the gravitational N-particle system. Saari [1971] studies the inter-particle distances $|\mathbf{q}_i - \mathbf{q}_j|$ ($i \neq j$; $i, j = 1, ..., N$) of the Newtonian gravitational system as $t \to \infty$, independent of the total energy of the system. He shows that, in the absence of oscillatory and pulsating motion,[66] the Newtonian gravitational system is quite well-behaved. To be precise, if pulsating and oscillatory motion is excluded, then either

$$|\mathbf{q}_i - \mathbf{q}_j| \sim C_{ij} t \tag{9.53}$$

or

$$|\mathbf{q}_i - \mathbf{q}_j| \approx t^{2/3} \tag{9.54}$$

or

$$|\mathbf{q}_i - \mathbf{q}_j| = \mathcal{O}(1). \tag{9.55}$$

Here C_{ij} is some positive constant. Since the Newtonian dynamics is time-reversal invariant, this result holds for $t \to -\infty$ as well.

Saari interprets this behavior as follows. As $t \to \infty$ the system forms clusters consisting of particles whose inter-particle distances are bounded. The centers of mass of these clusters recede from each other at a rate of about $t^{2/3}$. Moreover, the system forms subsystems (clusters of clusters) whose centers of mass recede from each other at a rate proportional to t. This, according to Saari, reflects well the actual behavior of our universe. It shows that, as time evolves, galaxies form which recede from each other according to the Newtonian version of the Hubble law of expansion: $|\dot{\mathbf{q}}_{ij}|/|\mathbf{q}_{ij}| = 1/t$ with $|\mathbf{q}_{ij}| := |\mathbf{q}_i - \mathbf{q}_j|$.

[66] For the notion of "oscillatory" and "pulsating" and the cited result, cf. Saari [1971].

References

[1] Abraham, R., and Marsden, J. E. (1978): *Foundations of Mechanics.* Second Edition. Boston: Addison-Wesley Publishing Company.

[2] Albert, D. (2009): *Time and Chance.* Cambridge, Mass.: Harvard University Press.

[3] Antonov, V. A. (1962): "Most Probable Phase Distribution in Spherical Star Systems and Conditions for its Existence." In J. Goodman and P. Hut (Eds.): *Dynamics of Star Clusters.* Dordrecht: Springer. 525-540.

[4] Arnol'd, V. I. (1989): *Mathematical Methods of Classical Mechanics.* New York: Springer.

[5] Barbour, J. B., and Pfister, H. (Eds.) (1995). *Mach's Principle. From Newton's Bucket to Quantum Gravity.* Einstein Studies. Basel: Birkhäuser.

[6] Barbour, J. (2003). "Scale-invariant Gravity: Particle Dynamics." *Classical and Quantum Gravity* 20. 1543.

[7] Barbour, J., Koslowski, T. and Mercati, F. (2013): "A Gravitational Origin of the Arrows of Time." ArXiv: 1310.5167 [gr-qc].

[8] Barbour, J., Koslowski, T. and Mercati, F. (2014): "Identification of a Gravitational Arrow of Time." *Physical Review Letters* 113:181101.

[9] Barbour, J., Koslowski, T. and Mercati, F. (2015): "Entropy and the Typicality of Universes." ArXiv: 1507.06498 [gr-qc].

[10] Boltzmann, L. (1896a): *Vorlesungen über Gastheorie.* Leipzig: Verlag v. J. A. Barth (Nabu Public Domain Reprints).

[11] Boltzmann, L. (1896b): "Entgegnung auf die wärmetheoretischen Betrachtungen des Hrn. E. Zermelo." *Wiedemann's Annalen* 57. 773-784.

[12] Boltzmann, L. (1897): "Zu Hrn. Zermelos Abhandlung 'Über die mechanische Erklärung irreversibler Vorgänge'." *Wiedemann's Annalen* 60. 392-398.

[13] Bricmont, J. (1995): "Science of Chaos or Chaos in Science." *Physicalia Magazine* 17. 159-208.

[14] Carroll, S., and Chen, J. (2004): "Spontaneous Inflation and the Origin of the Arrow of Time." ArXiv: 0410270 [hep-th].

[15] Carroll, S., and Chen, J. (2005): "Does Inflation Provide Natural Initial Conditions for the Universe?" ArXiv: 0505037 [gr-qc].

[16] Carroll, S. (2010): *From Eternity to Here: the Quest for the Ultimate Theory of Time.* Dutton. Penguin USA.

[17] Carroll, S., and Tam, H. (2010): "Unitary Evolution and Cosmological Fine-tuning." ArXiv: 1007.1417 [hep-th].

[18] Carroll, S. (2014): "In What Sense is the Early Universe Fine-tuned?" ArXiv: 1406.3057 [astro-ph].

[19] Cournot, M. A. A. (1843). "Exposition de la théorie des chances et des probabilités." Paris: Librairie de L. Hachette (Gallica digital library).

[20] Dirac, P. (1964): *Lectures on Quantum Mechanics.* New York: Dover Publications.

[21] Dürr, D., and Teufel, S. (2009): *Bohmian Mechanics.* Berlin: Springer.

[22] Dürr, D, Frömel, A., Kolb, M. (2017): *Einführung in die Wahrscheinlichkeitstheorie als Theorie der Typizität. Mit einer Analyse des Zufalls in Thermodynamik und Quantenmechanik.* Berlin: Springer.

[23] Ehrenfest, P. and T. (1911): "Begriffliche Grundlagen der Statistischen Auffassung in der Mechanik." In: *Enzyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen*, Band 4.

[24] Faddeev, L. D. (1969): "The Feynman Integral for Singular Lagrangians." *Theoretical and Mathematical Physics* 1:1. 1-13

[25] Faddeev, L. D., and Popov, V. N. (1967): "Feynman Diagrams for the Yang-Mills Field." *Physics Letters* 25B:1. 29-30.

[26] Feynman, R. (1967): *The Character of Physical Law.* Cambridge, Mass.: The MIT Press.

[27] Gibbons, G., Hawking, S., and Stewart, J. (1987): "A Natural Measure on the Set of all Universes." *Nuclear Physics B* 281:3. 736-751.

[28] Gibbons, G., and Turok, N. (2008): "Measure Problem in Cosmology." *Physical Review D* 77. 063516.

[29] Goldstein, S. (2001): "Boltzmann's Approach to Statistical Mechanics." In J. Bricmont, D. Dürr, et al. (Eds.): *Chance in Physics. Foundations and Perspectives.* Berlin: Springer. 39-54.

[30] Goldstein, S., Tumulka, R., and Zanghì, N. (2016): "Is the Hypothesis about a Low Entropy Initial State of the Universe Necessary for Explaining the Arrow of Time?". ArXiv:1602.05601v3 [astro-ph.CO].

[31] Heggie, D. and Hut, P. (2003): *The Gravitational Million-Body Problem.* Cambridge, UK: Cambridge University Press.

[32] Henneaux, M., and Teitelboim, C. (1992): *Quantization of Gauge Systems.* Princeton: Princeton University Press.

[33] Iwai, T. (1987): “A Geometric Setting for Classical Molecular Dynamics.” *Annales de l‘I.H.P. Physique théorique* 47:2. 199-219.

[34] Iwai, T., and Yamaoka, H. (2005): “Stratified Reduction of Classical Many-body Systems with Symmetry.” *Journal of Physics A: Mathematical and General* 38:11. 2415-2439.

[35] Kiessling, M. (2001): “How to Implement Boltzmann’s Probabilistic Ideas in a Relativistic World?” In J. Bricmont, D. Dürr, et al. (Eds.): *Chance in Physics. Foundations and Perspectives.* Berlin: Springer. 39-54.

[36] Koslowski, T., Mercati, F., and Sloan, D. (2016): “Relationalism Evolves the Universe through the Big Bang.” Preprint forwarded by J. Barbour.

[37] Littlejohn, R.G., and Reinsch, M. (1997) “Gauge Fields in the Separation of Rotations and Internal Motions in the N-body Problem.” *Review of Modern Physics* 69:1. 213-275.

[38] Lagrange, L. (1772): “Essai sur le problème de trois corps.” In: M. J.-A. Ferret and M. G. Darboux (Eds.): Oeuvres complètes de Lagrange. Paris: Gauthiers-Villars (Gallica digital library). Vol. 6. 229-331.

[39] Landau, L. D., and Lifshitz, E. M. (1967): *Mechanics. Volume 1 (Course on Theoretical Physics).* Oxford: Butterworth-Heinemann.

[40] Laplace, P.-S. (1814): “Essai philosophique sur les probabilités.” Paris: Bachelier (Gallica digital library).

[41] Lazarovici, D., and Reichert, P. (2015): “Typicality, Irreversibility and the Status of Macroscopic Laws.” *Erkenntnis* 80:4. 689-716.

[42] Lebowitz, J. (1981). “Microscopic Dynamics and Macroscopic Laws.” *Annals New York Academy of Sciences.* 220-233.

[43] Lebowitz, J. (1993). “Macroscopic Laws, Microscopic Dynamics, Time’s Arrow and Boltzmann’s Entropy.” *Physica A* 194. 1-27.

[44] Marchal, C., and Saari, D. (1974): “On the Final Evolution of the N-body Problem”. *Journal of differential equations* 20. 150-186.

[45] Marsden, J., and Weinstein A. (1974): “Reduction of Symplectic Manifolds with Symmetries.” *Reports on Mathematical Physics* 5. 121-130.

[46] Moeckel, R. (2007): “Symbolic Dynamics in the Planar Three-Body Problem.” *Regular and Chaotic Dynamics* 12:5. 449-475.

[47] Moeckel, R. (1981): “Orbits of the Three-body Problem which Pass Infinitely Close to Triple Collision.” *American Journal of Mathematics* 103:6. 1323-1341.

[48] Montgomery, R. (2002): “Infinitely many Syzygies.” *Archive for Rational Mechanics and Analysis* 164. 311-340.

[49] Padmanabhan, T. (1990): “Statistical Mechanics of Gravitating Systems.” *Physics Reports* 188:5. 285-362.

[50] Penrose, R. (2004): *The Road to Reality.* London: Vintage.

[51] Penrose, R. (2011): *Cycles of Time.* London: Vintage.

[52] Pollard, H. (1967). “Gravitational systems.” *Journal of Mathematics and Mechanics* 17. 601-612.

[53] Reichert, P. (2012): “Can a Parabolic-like Evolution of the Entropy of the Universe Provide the Foundation for the Second Law of Thermodynamics?” Master thesis (Math. Institut LMU München).

[54] Rovelli, C. (2000): *Quantum Gravity.* Cambridge: Cambridge University Press.

[55] Saari, D. (1971): “Expanding Gravitational Systems.” *Transactions of the American Mathematical Society* 156. 219-240.

[56] Saari, D. (1971): “Improbability of Collisions in Newtonian Gravitational Systems.” *Transactions of the American Mathematical Society* 162. 267-271.

[57] Saari, D. (1984): “Manifold Structure for Collisions and for Hyperbolic-Parabolic Orbits in the N-body Problem.” *Journal of Differential Equations* 55.300-329.

[58] Saari, D.: “Central Configurations – A Problem for the Twenty-first Century.” http://www.math.uci.edu/ dsaari/BAMA-pap.pdf

[59] Saari, D. (2005): *Collisions, Rings and other N-body Problems.* Providence, RI: American Mathematical Society.

[60] Scheck, F. (2003): *Theoretische Physik 1. Mechanik. Von den Newton'schen Gesetzen zum deterministischen Chaos.* Springer.

[61] Schiffrin, J. S. , and Wald, R. M. (2012): “Measure and Probability in Cosmology.” *Physical Review D* 86. 023521.

[62] Smale, S. (1970): “Topology and Dynamics.” *Inventiones Mathematicae* 10:4. 305-331.

[63] Tokasi, S. (2017): *Towards Relational Physics.* Master thesis. Math. Institut LMU München.

[64] Wallace, D. (2010): “Gravity, Entropy, and Cosmology: in Search of Clarity.” *British Journal for Philosophy of Science* 61. 513-540.

[65] Xia, Z. (1992): “The existence of non-collision singularities in Newtonian systems.” *Annals of Mathematics* 135. 411-468.

Eidesstaatliche Versicherung

(siehe Promotionsordnung vom 12.07.11, §8, Abs. 2, Pkt. 5)

Hiermit erkläre ich an Eides statt, dass die Dissertation von mir selbständig ohne unerlaubte Hilfe angefertigt ist.

München, den 24.05.2018 Paula Reichert-Schürmer