

How closed is cosmology?

Sean Gryb*

University College Groningen, Groningen, Netherlands

David Sloan†

Department of Physics, Lancaster University, Lancaster UK

Classical cosmology exhibits a particular kind of scaling symmetry. The dynamics of the invariants of this symmetry forms a system that exhibits many of the features of open systems such as the non-conservation of mechanical energy and the focusing of measures along the dynamical flow. From these properties, we show that important dynamical features emerge that are not present in closed systems. In particular, a large and physically plausible class of cosmological models give rise to a natural arrow of time. We then argue that the appropriate notion of closure in cosmology is dynamical closure - that a system can be integrated without reference to external factors. This is realised in physical systems in terms of the algebraic closure of the equations of motion such that the system is autonomous. Remarkably, in a growing class of models it can be shown that the autonomous system obtained remains regular and can be integrated through the big bang.

I. INTRODUCTION

What is a closed system? In trying to answer this question, we will find standard responses wanting in many respects. Our proposed solution — one that makes sense for self-gravitating systems — will have some seemingly unusual features. We will show that these features have unexpected virtues and only seem surprising because of the inadequacies of naive definitions of closed systems. In fact, our proposal will be seen to lead to some rather remarkable conclusions: the universe behaves like an open system where non-conservation of energy leads to the elimination of the conventional Big Bang and a solution to the problem of the Arrow of Time (AoT).

An important aspect of our proposal is the claim that energy need not be conserved in a closed system. Common intuition suggest that if a system is truly closed then there is nowhere for the energy to escape. But such an argument begs the question. Assuming that the energy, which is just some function on the state space of the system, must ‘escape’ somewhere when its value decreases is precisely the assumption that the energy must be conserved.

A more rigorous argument for conservation of energy in a closed system is to require the autonomy of the Lagrangian used to generate the system’s laws. Closure of the system could be taken to motivate the formal requirement that the Lagrangian be a function only of an algebraically closed set of dependent variables (e.g., field configurations and their derivatives) with no explicit dependence on the independent variables (e.g., spacetime coordinates x or t) of the system. In this case, it’s easy to show¹ that the equations of motion are equivalent to a symplectic Hamiltonian system where the Hamiltonian is a function only of phase space. A straightforward application of the theorems of symplectic geometry then suggests that the

* s.b.gryb@rug.nl

† d.sloan@lancaster.ac.uk

¹ Modulo spatial boundary conditions that could be taken to zero for a closed system.

Hamiltonian function is a constant of motion. Identifying the Hamiltonian with the energy then gives conservation of energy.

This argument, though general and elegant, has several difficulties. First, it's not always appropriate to identify the Hamiltonian function with the energy of the system. Any reparametrization-invariant theory, such as general relativity, will have a vanishing Hamiltonian, and this cannot not lead to an interesting notion of energy. Second, the argument assumes that the laws of the theory are generated by Hamilton's principle, which under the assumption of the autonomy of the Lagrangian leads to a symplectic system. But many interesting physical systems, as we will argue later, are not symplectic. There is thus no guarantee, in general, that the Hamiltonian — let alone the energy — will be conserved in a non-symplectic system.

An alternative motivation for conservation of energy could run through Noether's first theorem and the assumption of time-translation invariance of the Lagrangian. Noether's first theorem, however, also assumes that the equations of motion are generated by Hamilton's principle and therefore also excludes non-symplectic systems. Moreover, it's not clear that time-translation invariance of the Lagrangian should be a necessary condition for a closed system. In order to have a definition of a closed system that is general enough to include non-symplectic systems, one must therefore abandon the notion that energy must be conserved.

There are, however, naive arguments against the existence of non-symplectic closed systems. These arguments run through Liouville's theorem, which gives the unique *Liouville measure* that is conserved in a (finite-dimensional) symplectic system.² The existence of such a measure guarantees that conserved probability distributions can be defined within the theory. Thus, within symplectic systems there is a straightforward way to 'conserve probabilities.' The conditions for Liouville's theorem are not, however, necessary conditions for being able to define conserved probabilities outside of symplectic systems. For example, in contact systems, which are not symplectic and will play an important role in our considerations below, generalisations of Liouville's theorem guarantee a unique dynamically conserved measure. Thus, there exists perfectly well-defined notions of probability for closed non-symplectic systems.

Our proposal for defining a closed system does not require the system to be symplectic. Instead, it makes use of an essential element of the proposals given above: the requirement of the autonomy of the dynamical system in question. While symplectic proposals rely on the autonomy of the Lagrangian and the use of Hamilton's principle to generate autonomous equations of motion, ours will focus more directly on the autonomy of the equations of motion themselves. In particular, we will define a closed dynamical system as a system whose dynamical laws are deterministic and algebraically closed in terms of a particular algebra of observables. We will be silent on the nature of the observable algebra in question so that the algebraic closure of our equations does not in principle prevent us from treating probabilistic or quantum theories (modulo the measurement problem).

To begin, we will give a simple example to illustrate different senses in which a dynamical system can be closed. This will help both to motivate our proposal and to illustrate the consequences for cosmology.

A. Closure in the harmonic oscillator

Let us first consider an example of a system that is generally accepted to be closed, and is so under all the definitions discussed above. This is the example of the harmonic oscillator

² No such finite measure exists in a infinite-dimensional system.

consisting of two constants: the mass m and spring constant k ; and an equation of motion relating acceleration a to position x ,

$$a = -kx. \quad (1.1)$$

In this system, the mechanical energy,

$$E = \frac{mv^2}{2} + \frac{kx^2}{2} \quad (1.2)$$

which is also numerically equal to the Hamiltonian of the system, is constant. If we plot the trajectory of such a particle on phase space (i.e. plotting momentum against position) we find that the orbits form ellipses, which we can directly see is the result of the energy, E , being constant along a classical trajectory. Given an ensemble of such systems, each with some initial position and momentum determined such that the ensemble spreads over some region R in the phase space, the dynamical flow of the ensemble is such that R is rotated about the origin at some fixed angular velocity while the volume it occupies on phase space is constant. This is a direct consequence of Liouville's theorem, which applies to this system.

Let us now imagine generalising the harmonic-oscillator model in the following standard way by adding a damping term proportional to the velocity v :

$$ma = -kx - fv \quad (1.3)$$

where f is a constant. There are now at least two senses in which the system is no longer closed. First, the mechanical energy of the system decreases over time because dE/dt is proportional to $-v^2 < 0$. Second, the orbits on phase space are no longer ellipses but spirals, which eventually come to rest at the origin. The region R then shrinks over time to a point at the origin.

But although our model is open in the sense that neither energy nor Liouville volume are preserved, it is nevertheless autonomous because the equations of motion are entirely self-contained: given an initial x and v and the value of f , we can solve the dynamical equations for all time without reference to any other information. This damped model is then autonomous, and therefore closed, in the sense we have defined above.

At this point, one might contend that damped oscillators should be understood as open systems because they are usually thought to interact dynamically with their environments; for example, through air resistance. But the model itself makes no explicit reference to such effects, and including these explicitly would ultimately involve knowing more information about the environment and its interactions with the system. On the other hand, the damped oscillator can be taken as a complete system making no reference to the outside world, where the value of f is akin to a constant of nature. Absent further modeling assumptions, the system can thus be seen to be closed in the sense that all future behaviour can be inferred from current, self-contained knowledge.

The models treated in this section illustrate the three distinct notions of 'closure' discussed in the introduction. Note that these notions are completely independent: physical systems may exhibit any or none of these. But because the implications of each notion can differ drastically, it is important identify which notions are relevant (or not) a system in question.

II. CONTACT SYSTEMS

One class of systems with autonomous equations of motion that gives rise neither to conserved Liouville forms nor to conserved energies is the class of contact systems. Contact mechanics can be thought of as the odd-dimensional counterpart to symplectic mechanics. Here,

for reasons of length, we will introduce only the necessary parts for our analysis, and work directly in Darboux coordinates. Comprehensive and general introductions can be found in [5, 13].

Herglotz [18] considered a generalization of the usual Lagrangian mechanics to include Lagrangians that are dependent on the action, i.e.

$$S = \int \mathcal{L}(q_i, \dot{q}_i, S) dt \quad (2.1)$$

the minimization of which gives rise to a slightly modified version of the usual Euler-Lagrange equations, henceforth called the Herglotz-Lagrange equations:

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) - \frac{\partial \mathcal{L}}{\partial q_i} - \frac{\partial \mathcal{L}}{\partial S} \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \quad (2.2)$$

Here we see that when the Lagrangian is independent of the action the usual equations are recovered. The Lagrangian is taken to be a function of the tangent bundle over configuration space and the action. This is an odd dimensional space. By construction, our system is autonomous because the Lagrangian is a function only of the coordinates, their derivatives and the action. We can perform a Legendre transform and arrive at a contact Hamiltonian following the usual scheme. Let $p = \frac{\partial \mathcal{L}}{\partial \dot{q}}$. Then, the contact Hamiltonian is

$$\mathcal{H}^c = \sum_{i=1}^n p_i \dot{q}_i - \mathcal{L} \quad (2.3)$$

From this have the contact Hamilton equation:

$$\dot{q}_i = \frac{\partial \mathcal{H}^c}{\partial p_i} \quad \dot{p}_i = -\frac{\partial \mathcal{H}^c}{\partial q_i} - \frac{\partial \mathcal{H}^c}{\partial S} \frac{\partial \mathcal{H}^c}{\partial p_i} \quad \dot{S} = \sum p_i \frac{\partial \mathcal{H}^c}{\partial p_i} - \mathcal{H}^c \quad (2.4)$$

Again, the usual Hamilton equations are recovered when the contact Hamiltonian is independent of the action. It is a straightforward exercise to see that the contact Hamiltonian is not conserved under time evolution with

$$\dot{\mathcal{H}}^c = -\frac{\partial \mathcal{H}^c}{\partial S} \mathcal{H}^c \quad (2.5)$$

The contact form, $\eta = \sum_{i=1}^n p_i dq_i - dS$, is the odd-dimensional counterpart to the symplectic potential. To obtain a volume form on our contact phase space, written in the coordinates q_i, p_i and S , we take

$$\Omega^c = \eta \wedge d\eta^n = -dS \wedge \bigwedge_{i=1}^n (dq_i \wedge dp_i) \quad (2.6)$$

Since the space is odd dimensional, we cannot construct a Liouville-type measure because the symplectic form is a two-form, and hence any product of symplectic forms is an even order form that cannot serve as a volume form for an odd-dimensional space. In contrast to the Liouville form, the privileged volume form on a contact manifold is *not* conserved under time evolution and obeys

$$\dot{\eta} = -\frac{\partial \mathcal{H}^c}{\partial S} \eta \rightarrow \dot{\Omega}^c = -(n+1) \frac{\partial \mathcal{H}^c}{\partial S} \Omega^c \quad (2.7)$$

Let us illustrate our construction by considering the damped harmonic oscillator. If we begin with the Lagrangian

$$\mathcal{L} = \frac{m\dot{x}^2}{2} - \frac{kx^2}{2} - \frac{\mu S}{m} \quad (2.8)$$

then the Herglotz–Lagrange equation for x is precisely that of the damped harmonic oscillator 1.3. Identifying $p = m\dot{x}$, the contact Hamiltonian is thus

$$\mathcal{H}^c = \frac{p^2}{2m} + \frac{kx^2}{2} + \frac{\mu S}{m} \quad (2.9)$$

It is important to note that this does not correspond to what we would think of as the mechanical energy of the oscillator as it contains a term proportional to the action. The evolution of the contact Hamiltonian shows that it has an exponential decay, as does the contact volume-form.

We can interpret the decay of the contact volume-form as a focusing of solutions on the contact space. This focusing of solutions, in turn, introduces a local time asymmetry into the system along particular solutions, and thus can be used to define a local Arrow of Time. More globally, because this focusing is monotonic along *all* solutions and points in the same direction as the decrease in the contact Hamiltonian, the asymptotic future of any damped oscillator is a system at rest, $x = 0 = \dot{x}$. The future of all solutions converges on to this point. Hence, for all solutions we can further identify a global Arrow of Time: the direction pointing towards this final state. Equivalently, this Arrow of Time is the direction in which the energy (either mechanical, or the contact Hamiltonian itself) is decreasing. See [13] for an extensive treatment of how an arrow of time can emerge through such a mechanism. The damped oscillator thus has a qualitative feature distinct from its undamped counterpart.

A given contact system can be symplectified. Doing so will produce a symplectic system in which the Hamiltonian is conserved under evolution. This is achieved at the cost of embedding the contact phase space into a phase space with one extra dimension. This process will imply that a single solution to the contact system is mapped onto multiple solutions of the symplectic system, which will appear as a symmetry in the symplectic representation. A simple way to achieve this is to expand our configuration space by introducing a new coordinate Q orthogonal to the contact space. We can then define the symplectic Hamiltonian and symplectic potential by

$$\mathcal{H}^s = Q\mathcal{H}^c \quad \theta = Q\eta \quad (2.10)$$

It is a simple exercise to show that both the symplectic Hamiltonian and potential are conserved as a consequence of Hamilton's equations, and that the equation of motion for each of the contact variables remains unchanged by this process. In our damped harmonic oscillator example, our symplectic system is then

$$\mathcal{H}^s = \frac{\pi^2}{2mQ} + \frac{kQx^2}{2} + \frac{\mu QS}{m} \quad \omega = dS \wedge dQ + d\pi \wedge dx \quad (2.11)$$

Note that the observable algebra of this system is not generated by the original x and $p = v$ variables but x and $\pi = Qv$. This system, as any system made from such symplectifications, exhibits dynamical similarity of order 1 under the similarity vector $D = Q\frac{\partial}{\partial Q} + \pi\frac{\partial}{\partial \pi}$, i.e

$$\mathfrak{L}_D \mathcal{H}^s = \mathcal{H}^s \quad \mathfrak{L}_D \omega = \omega. \quad (2.12)$$

This reflects the redundancy brought to the system through the introduction of the coordinate Q .

At this point it is natural to ask why, if an open (contact) system can always be embedded within a closed (symplectic) system of this type, do we not simply declare the closed system to give a complete description? Our answer comes from demanding autonomy only on the smallest algebraic set required to retain empirical adequacy [14]. Let us motivate this demand now.

In our symplectification process we have, by necessity, introduced an auxiliary structure Q into the system. In the case of the damped harmonic oscillator, the momenta depend upon this auxiliary variable Q . Similarly, both the conserved volume-form and the conserved energy (ie, the symplectic Hamiltonian) are also dependent upon Q , which can take arbitrary values. All these quantities will therefore take arbitrary values themselves, and thus cannot be used to represent physical structures. The value of Q can therefore have no empirical significance. We see that autonomy in the symplectic system is only achieved in terms of an algebra that is strictly larger than the minimum required for empirical adequacy. In terms of the measure, this means that states are counted as distinct even if they differ only by the (arbitrary) value of Q . This is clearly not a desirable state of affairs as it would introduce a distinctions without a difference.

In the case of the damped oscillator, we thus find that the non-conservative, contact formulation should be treated as a complete description of the system as it involves an autonomous system that is independent of auxiliary structures. We conclude that certain apparently ‘open’ systems can reasonably be considered to be ‘closed’ in the sense that they provide complete descriptions.

III. COSMOLOGY

Let us now apply our framework to cosmology. In our analysis we shall restrict to the simplest case: flat ($k = 0$) Friedmann-Lemaître-Robertson-Walker (FLRW) models. The results can be generalized to include open ($k = -1$) and closed ($k = 1$) spatial slices, and include anisotropies, together with general matter sources [19, 23, 24]. Our considerations here thus illustrate a much more general procedure that can be applied to a variety of different situations. For reasons of space, we refer the interested reader to the cited papers for details.

There is an unfortunate clash of nomenclature here wherein cosmologically the terms ‘open’ and ‘closed’ (and separately ‘flat’) are used to refer to the geometry of a spatial manifold, which can either be a three dimensional hyperboloid, a three-sphere or a compactified section of \mathbb{R}^3 . The issue of whether or not a spatial slice is compact is not at issue here since the stringent demands of homogeneity and isotropy render such considerations orthogonal to the issue of ‘open’ versus ‘closed’ in terms of conservation laws or autonomy.

Cosmology provides an interesting test case for distinguishing between the various demands of being ‘closed’. Since the object under consideration is taken to be the entire universe, it may be argued that if some quantity is to be conserved in a complete accounting of physics, then there should be nowhere to hide any losses. However, one may counter that cosmology is simply the consideration of long wavelength modes of a full theory of general relativity, and that a complete model should account for the transfer of, e.g., energy between modes of differing wavelengths. Thus, a failing of strict conservation laws within the cosmological context does not rule out such conservation for the full theory.

Such an arguments does not appear to hold in the vicinity of space-time singularities where, per the Belinski-Khalatnikov-Lifschitz [3, 4, 15, 20, 26] conjecture, the dynamics of the full

theory of general relativity asymptotes to those of a homogeneous solution at each spatial point. Further, since cosmological solutions are exact solutions of the full theory, as opposed to truncations in which some terms have been disregarded, any features of these solutions are necessarily features of at least a subset of solutions to the full theory. Thus we have a counter-example to the claim that conservation should hold in a full theory.

For our purposes it will suffice to demonstrate our results in the flat FLRW model with simple scalar fields as matter.³ The behaviour of such a cosmology can be found from the Einstein–Hilbert action following a symmetry reduction to homogeneous and isotropic subsystems. After some technical considerations, the foremost of which is checking compatibility with the principle of symmetric criticality, we can render the action as

$$S_{\text{EH}} = \int v \left(-\frac{\dot{v}^2}{3v^2} + \frac{\dot{\phi}_i^2}{2} - V \right) dt, \quad (3.1)$$

where $V = V(\vec{\phi})$ is a potential function and, to streamline our algebra, we adopt the convention $8\pi G = 1$. We here note that a cosmological constant can be included trivially in this system by having a non-zero minimum of V . The geometric sector is described by v , which represents the volume of a fiducial cell in comoving coordinates. Note that although the scale factor is often referred to as being the ‘size of the universe’, this is undefined in the case of an infinite spatial slice. From this action we can derive (following a Legendre transformation) the Hamiltonian description of our theory. We write this in terms of a Hamiltonian function and symplectic structure on phase space

$$\begin{aligned} 0 := \mathcal{H} &= v \left(-\frac{3P_v^2}{8} + \frac{P_i^2}{2v^2} + V \right) \\ \omega &= dP_v \wedge dv + dP_i \wedge d\phi_i. \end{aligned} \quad (3.2)$$

Here, $P_v = -\frac{4\dot{v}}{3v} = 3H$, where H is the Hubble parameter. Note that the Hamiltonian is a constraint and must be zero — a fact that follows from the time reparametrization invariance of general relativity. From the existence of these structures it is easy to verify that the ‘energy’ (the Hamiltonian) and the volume-form $\Omega = \omega^{1+n}$ where n is the number of scalar fields, are conserved over time.⁴ Furthermore, from Hamilton’s equations we can find an equations of motion of each of the phase-space variables written in terms of itself and the other phase-space variables. Our system is therefore autonomous. It would hence seem apparent that cosmology is a closed system following all three notions that we have discussed.

There is, however, a catch: The algebra generated by the set of variables just described is strictly *larger* than what is needed for empirical adequacy. In particular, the volume v can be set to any value by re-scaling its initial condition without affecting any empirical predictions of the theory. This is because such predictions only ever rely on the value of v relative to some reference value (usually taken to be an initial or final condition).

The equations of motion of our system are the Friedmann equations and the Klein-Gordon

³ We will deal exclusively with classical models in this paper, sidestepping the important question of how to build a full-fledged quantum model of the early universe. Such models, for instance, may involve modified dynamics on a coarse-grained level.

⁴ Note that while the numerical value of the Hamiltonian, treated as a phase space function, is preserved, the functional form of its Legendre transform will depend, in general, on the time parameter chosen for the evolution.

equation:

$$H^2 = \frac{1}{3} \left(\frac{\dot{\phi}_i^2}{2} + V \right) \quad (3.3)$$

$$\dot{H} + H^2 = -\frac{1}{3} \left(\dot{\phi}_i^2 + V \right) \quad (3.4)$$

$$\ddot{\phi}_i + 3H\dot{\phi}_i + \frac{\partial V}{\partial \phi_i} = 0 \quad (3.5)$$

It is clear from these that the dynamical algebra closes in terms of the ϕ and H without ever needing to refer to v . We should therefore seek a description of our system that never makes reference to v in the first place. To do so, we note that the freedom to choose the value of v at any time represents a scaling symmetry, and thus we can follow the analysis of [6, 13, 21] and reduce our symplectic system to a contact system with one fewer dimension. In fact, it was shown in [23, 24] that the dynamics of this system can be derived from an action principle that never refers to scale in the first place by extremizing the action

$$S_h = \int_0^t \left(\frac{3S_h^2(t')}{2} + \frac{\dot{\phi}_i^2}{2} - V \right) dt'. \quad (3.6)$$

This is an action of the Herglotz type as discussed above. Together with the identification $S_h = -H$ this produces the original equations of motion. The action is equivalent to the second Friedmann equation and the Herglotz–Lagrange equation for ϕ is the Klein–Gordon equation. The contact Hamiltonian is then equivalent to the first Friedmann equation:

$$0 := \mathcal{H}^c = -\frac{3S^2}{2} + \frac{p_i^2}{2} + V \quad \eta = -dS + dp_i \wedge d\phi_i \quad (3.7)$$

with $p_i = \dot{\phi}_i$.

There are several notable features of this description of our system. The contact Hamiltonian and contact structure are expressed directly in terms of the minimum autonomous system $(\phi, \dot{\phi}, S = -H)$ required for empirical adequacy rather than the usual symplectic variables used in cosmology. The identification of the action S with the Hubble parameter makes the nature of ‘Hubble friction’ more apparent. From equation 2.6, we see that the natural volume form over physical quantities will shrink in time. This solution focusing is a hallmark of ‘open’ systems.

The action is monotonically increasing, and can therefore act as a physically observable proxy for time. It is common when addressing issues of measure in cosmology to construct a space of solutions by restricting to a surface whereon the Hubble parameter is constant [1, 11]. This is the restriction to a constant extrinsic curvature slicing in the cosmological context. Doing so demonstrates clearly the distinction between the symplectic and contact cases: If we restrict the Liouville form in each case to this slice, we find that $\Omega^s = \dot{\phi} dv \wedge d\phi$ whereas $\Omega^c = \dot{\phi} d\phi$ since the contact phase space is one dimension smaller. Thus in the symplectic case, when evaluating measures on phase space, one must choose an interval in v over which to integrate. This can lead to problems because the volume is not empirically significant. However, since the amount by which volume changes between two times is, the freedom to chose different initial conditions for v leads to ambiguous statements [2, 7, 8, 25]. In the contact case, the measure is unambiguous but not preserved under time evolution. The time evolution of the measure is proportional to the action, and hence solutions focuses on those which maximize the integral of the Hubble parameter between two time slices.

These considerations show that in cosmology, as in the damped oscillator, ‘open’ dynamical systems can reasonably be understood to give complete, though idealised, descriptions of the universe.

A further enticing feature of the open description of cosmology is that the system of differential equations it produces remain deterministic at, and beyond, the initial big bang singularity. The status of singularities within general relativity as a whole is a complex topic with the Hawking–Penrose theorems proving that, under fairly generic circumstances, there will exist points in a space-time beyond which the paths of physical objects within the system become non-deterministic as the geodesics that objects follow are incomplete. Thus, Einstein’s equations do not admit a unique continuation beyond such points. In the cosmological setup, our considerations are somewhat simpler: we need only concern ourselves with the determinism of the system of equations derived from the symplectic Hamiltonian description, equations III. The relevant point in evolution is when the Hubble parameter becomes infinite as does (necessarily) at least one contribution to the matter energy density. A theorem due to Foster [10] shows that under a broad range of conditions this corresponds to the kinetic energy of the scalar field becoming infinite.

Infinities, even of physically observable quantities, do not necessary signal the breakdown of the evolution equations. If we consider the harmonic oscillator described in terms of $y = 1/x$, we see that its evolution results regularly in infinities of y in finite time. The physical system, with x interpreted as the distance to equilibrium, sees such points as nothing special and continues beyond them. The infinities may appear, on first glance, to inhibit the integration of the equations of motion. However, when rendered in the usual manner through a change of variables, we see that they are in fact an artefact of the description.

In our cosmological model it is the scalar fields that become infinite. We can change coordinates through a gnomonic projection. For two scalar fields, this is achieved by projecting a point on a plane onto one half of a sphere. The unit radius sphere is tangential to the plane, touching at the origin. Each point P on the plane is then mapped uniquely to a point Q on the lower half of the sphere by considering a line from the center of the sphere to the point P , and taking Q to be the point of intersection of this line with the sphere. Thus two scalar field values, ϕ_1 and ϕ_2 are represented as a pair of longitude and latitude on the sphere. The singularity in the original coordinates (where at least one of the fields has become infinite) is mapped to the equator of the sphere. It is then a matter of technical calculations [17, 22] to show that the equations of motion of the system remain well defined and predictive when expressed in terms of these sphere based coordinates.

The basic mechanism that allows for evolution through the big bang in this formalism is clear: when the volume v is recognised as an empirically irrelevant quantity, then the geodesics of the space-time metric lose some of their operational significance. Integrability of the contact equations then replaces geodesic completeness as a criteria for singularity resolution. While some progress has been made in understanding this new criterion more physically (see, for example, [16]), ongoing efforts of underway to provide further clarifications. See, however, [17, 22] for the full details of the technical construction and motivations.

IV. CONCLUSIONS

Throughout this work, we have investigated different inequivalent ways in which dynamical systems can be considered closed. These different notions include closure in the sense of conserved energy, conserved measure density and dynamical autonomy. While we found merits for

each, no single notion of closure was completely adequate for all purposes. In particular, we found conventional notions of conserved energy and measure to be wanting when applied to the cosmological setting. In that context, we compared a closed (in the sense of preserved measure density and Hamiltonian) symplectic description of cosmology to an open contact formulation. We found that the latter should be taken to give a better complete description of the system.

Central to our argument was a requirement to treat as physical only those structures that are strictly necessary for maintaining empirical adequacy. The symplectic ‘closed’ descriptions violate this principle because they treat as distinct states that differ only by the overall scale of the universe (as determined by v), which is not empirically relevant. In contrast, the contact description does satisfy this principle as it excludes v . In other words, while the symplectic and contact systems are both empirically adequate, they are not physically equivalent in the sense that the symplectic system makes definite claims about the value of v — claims which are not necessary for describing cosmological phenomena. While the contact system does not conserve energy or measure density, it is nevertheless autonomous in the empirically sufficient variables, and is thus ‘closed’ in the third sense defined in the introduction. This suggests that the appropriate notion of closure in the cosmological setting is that of dynamical autonomy in terms of the contact system.

In addition to this epistemic argument favouring the contact description, we also showed that the contact description of cosmology has several other explanatory advantages over the standard symplectic description. Firstly, the contact description provides a natural dynamical explanation for the Arrow of Time.⁵ This arrow comes from the focusing of solutions along the dynamical evolution — particularly when this focusing is monotonic across a dense set of solutions that are consistent with observation. Indeed, we saw that this is precisely what occurs in the cosmological models we considered since the solution focusing is proportional to the Hubble parameter, and Hubble expansion is monotonic under a dense set of solutions in the presence of a positive cosmological constant.⁶ In symplectic description, the natural measure is the time-independent Liouville measure, which exists because the symplectic theory counts states related by different values of v as distinct. It is thus extremely unnatural to use time-dependent measures in the symplectic description. As a result, the Arrow of Time is usually explained in terms of some version of a past hypothesis. The limitations of a past hypothesis have been discussed by many authors (e.g., [9, 12]). In the description we are advocating, these difficulties are avoided because a past hypothesis is unnecessary.

A second advantage of the ‘open’ model⁷ is that, while it agrees with the ‘closed’ description wherever that description is valid, it continues to be regular when the ‘closed’ system breaks down. The ‘open’ model can be integrated through the initial singularity unambiguously. The contact system is thus sufficient to save the phenomena while avoiding the problematic aspects of a big bang singularity.

[1] Ashtekar, A., & Sloan, D. (2010). Loop quantum cosmology and slow roll inflation. *Phys.Lett., B694*, 108–112.

[2] Ashtekar, A., & Sloan, D. (2011). Probability of inflation in loop quantum cosmology. *General Relativity and Gravitation*, 43(12).

⁵ See [13] for more details about which aspects of the problem of the Arrow of Time can be addressed with this formalism.

⁶ Particularly when the matter content is required to obey any of the energy conditions of general relativity.

⁷ Here, we mean ‘open’ in the sense of having no conserved measure density and Hamiltonian but closed in the sense of dynamical autonomy.

- [3] Belinsky, V. a., Khalatnikov, I. m., & Lifshitz, E. m. (1982). A General Solution of the Einstein Equations with a Time Singularity. *Adv. Phys.*, *31*, 639–667.
- [4] Berger, B. K., Garfinkle, D., Isenberg, J., Moncrief, V., & Weaver, M. (1998). The Singularity in generic gravitational collapse is space - like, local, and oscillatory. *Mod. Phys. Lett.*, *A13*, 1565–1574.
- [5] Bravetti, A., Cruz, H., & Tapias, D. (2017). Contact Hamiltonian mechanics. *Annals of Physics*, *376*, 17–39.
- [6] Bravetti, A., Jackman, C., & Sloan, D. (2023). Scaling symmetries, contact reduction and poincaré’s dream. *Journal of Physics A: Mathematical and Theoretical*, *56*(43), 435203.
URL <https://dx.doi.org/10.1088/1751-8121/acfddd>
- [7] Corichi, A., & Karami, A. (2011). On the measure problem in slow roll inflation and loop quantum cosmology. *Phys.Rev.*, *D83*, 104006.
- [8] Corichi, A., & Sloan, D. (2014). Inflationary Attractors and their Measures. *Class.Quant.Grav.*, *31*, 062001.
- [9] Earman, J. (2006). The “past hypothesis”: Not even false. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, *37*(3), 399–430.
- [10] Foster, S. (1998). Scalar field cosmologies and the initial space-time singularity. *Class. Quant. Grav.*, *15*, 3485–3504.
- [11] Gibbons, G., Hawking, S., & Stewart, J. (1987). A Natural Measure on the Set of All Universes. *Nucl.Phys.*, *B281*, 736.
- [12] Gryb, S. (2021). New difficulties for the past hypothesis. *Philosophy of Science*, *88*(3), 511–532.
- [13] Gryb, S. (2024). Phd thesis. In preparation.
- [14] Gryb, S., & Sloan, D. (2021). When scale is surplus. *Synthese*, *199*(5-6), 14769–14820.
- [15] Heinzle, J. M., & Uggla, C. (2009). Mixmaster: Fact and Belief. *Class. Quant. Grav.*, *26*, 075016.
- [16] Hoffmann, J., & Sloan, D. (2024). Continuation of bianchi spacetimes through the big bang. *Phys. Rev. D*, *110*, 063539.
URL <https://link.aps.org/doi/10.1103/PhysRevD.110.063539>
- [17] Koslowski, T. A., Mercati, F., & Sloan, D. (2018). Through the big bang: Continuing Einstein’s equations beyond a cosmological singularity. *Phys. Lett.*, *B778*, 339–343.
- [18] Lazo, M. J., Paiva, J., Amaral, J. T. S., & Frederico, G. S. F. (2018). An action principle for action-dependent lagrangians: Toward an action principle to non-conservative systems. *Journal of Mathematical Physics*, *59*(3), 032902.
URL <http://dx.doi.org/10.1063/1.5019936>
- [19] Mercati, F. (2019). Through the Big Bang in inflationary cosmology. *JCAP*, *10*, 025.
- [20] Ringstrom, H. (2001). The Bianchi IX attractor. *Annales Henri Poincare*, *2*, 405–500.
- [21] Sloan, D. (2018). Dynamical Similarity. *Phys. Rev.*, *D97*(12), 123541.
- [22] Sloan, D. (2019). Scalar Fields and the FLRW Singularity. *Class. Quant. Grav.*, *36*(23), 235004.
- [23] Sloan, D. (2021). New action for cosmology. *Phys. Rev. D*, *103*(4), 043524.
- [24] Sloan, D. (2023). Herglotz action for homogeneous cosmologies. *Class. Quant. Grav.*, *40*(11), 115008.
- [25] Sloan, D., & Silk, J. (2016). Volume Weighting the Measure of the Universe from Classical Slow-Roll Expansion. *Phys. Rev. D*, *93*(10), 104030.
- [26] Uggla, C. (2007). The Nature of generic cosmological singularities. In *Recent developments in theoretical and experimental general relativity, gravitation and relativistic field theories. Proceedings, 11th Marcel Grossmann Meeting, MG11, Berlin, Germany, July 23-29, 2006. Pt. A-C*, (pp. 73–89).