A Discrete Analog of General Covariance: Could the world be fundamentally set on a lattice?

Daniel Grimmer\textsuperscript{1,2,3,*}

\textsuperscript{1}Reuben College, University of Oxford, Oxford, OX2 6HW United Kingdom
\textsuperscript{2}Faculty of Philosophy, University of Oxford, Oxford, OX2 6GG United Kingdom
\textsuperscript{3}Barrio RQI, Waterloo, Ontario N2L 3G1, Canada

A crucial step in the history of General Relativity was Einstein’s adoption of the principle of general covariance which demands a coordinate independent formulation for our spacetime theories. General covariance helps us to disentangle a theory’s substantive content from its merely representational artifacts. It is an indispensable tool for a modern understanding of spacetime theories, especially regarding their background structures and symmetry. Motivated by quantum gravity, one may wish to extend these notions to quantum spacetime theories (whatever those are). Relatedly, one might want to extend these notions to discrete spacetime theories (i.e., lattice theories). This paper delivers such an extension with surprising consequences.

One’s first intuition regarding discrete spacetime theories may be that they introduce a great deal of fixed background structure (i.e., a lattice) and thereby limit our theory’s possible symmetries down to those which preserve this fixed structure (i.e., only certain discrete symmetries). By so restricting symmetries, lattice structures appear to be both theory-distinguishing and fundamentally “baked-into” our discrete spacetime theories. However, as I will discuss, all of these intuitions are doubly wrong and overhasty. Discrete spacetime theories can and do have continuous translation and rotation symmetries. Moreover, the exact same theory can be given a wide variety of lattice structures and can even be described with no lattice at all. As my discrete analog of general covariance will reveal: lattice structure is rather less like a fixed background structure or part of an underlying manifold and rather more like a coordinate system, i.e., merely a representational artifact. Ultimately, I show that the lattice structure supposedly underlying any discrete “lattice” theory has the same level of physical import as coordinates do, i.e., none at all. Thus, the world cannot be “fundamentally set on a square lattice” (or any other lattice) any more than it could be “fundamentally set in a certain coordinate system”. Like coordinate systems, lattice structures are just not the sort of thing that can be fundamental; they are both thoroughly representational. Spacetime cannot be discrete (even when it might be representable as such).

\section{I. INTRODUCTION}

A crucial step in the history of General Relativity (GR) was Einstein’s adoption of the principle of general covariance\textsuperscript{1} which states that the form of our physical laws should be independent of any choice of coordinate systems. At first, Einstein thought this property was unique to GR and that this is what set his theory apart from all of its predecessors. However, in 1917 Kretschmann pointed out that any physical theory can be written in a generally covariant form (i.e., in a coordinate-independent way). See \textsuperscript{2} for a historical review of this point.

The modern understanding of the principle of general covariance is best summarized by Friedman \textsuperscript{3}:

\ldots the principle of general covariance has no physical content whatever: it specifies no particular physical theory; rather it merely expresses our commitment to a certain style of formulating physical theories.

However, despite this lack of physical content\textsuperscript{2} the conceptual benefits of writing a theory in a coordinate-free way are immense. A generally covariant formulation of a theory has at least two major benefits: 1) it more clearly exposes the theory’s geometric background structure, and 2) it thereby helps clarify our understanding of the theory’s symmetries (i.e., its structure/solution preserving transformations). It does both of these by disentangling the theory’s substantive content from representational artifacts which arise in particular coordinate representations \textsuperscript{1} \textsuperscript{2} \textsuperscript{3}. Thus, general covariance is an indispensable tool for a modern understanding of spacetime theories. This paper seeks to extend this tool to discrete spacetime theories (i.e., lattice theories\textsuperscript{1}). One aim of this extension is to (albeit indirectly) contribute

\begin{itemize}
  \item Some argue that this principle does, in fact, have physical content at least when it is applied to isolated subsystems \textsuperscript{4}. E.g., in Galileo’s thought experiment when only the ship subsystem is boosted relative to the un-boosted shore.
  \item As I will discuss later, calling these “lattice theories” is a bit of a misnomer. This would be analogous to referring to classical spacetime theories as “coordinate theories”. As I will discuss, in both cases the coordinate systems/lattice structure are mere
\end{itemize}
to the debate regarding background structures outlined below with possible applications towards quantum gravity.

It is now widely believed that the key conceptual shift which sets GR apart from its predecessors is not its general covariance as Einstein thought but rather its background independence, i.e., its complete lack of background structure. This consensus, however, is largely only verbal as there is still substantial debate about what exactly should and should not count as background structure \cite{1,2,6-10}. The central goal in this debate is to find a relatively simple notion of background structure such that all intuitively GR-like spacetime theories lack background structure while all intuitively GR-unlike spacetime theories (e.g., special relativity, Newtonian gravity, etc.) have background structure. That is, the task is to populate this landscape with interesting theories and then compare different ways of carving this space up \cite{11-17}.

In addition to its foundational importance, progress in this debate is also expected to help guide us in our future theorizing. Indeed, it is often claimed that any successor theory to GR (e.g., quantum gravity) should follow GR’s lead and be similarly background independent \cite{18-23} (whatever this ends up meaning \cite{8}). One may think that this claim is frustratingly aspirational in two ways: If we cannot agree on what precisely “background structure” means in the context of classical spacetime theories, what hope do we have of extending this notion to quantum spacetime theories (whatever those are exactly)? Isn’t it too soon to try to extend these unclear ideas into new territory?

This sort of thinking goes wrong in two ways: firstly, thinking that conceptual clarity must precede extension, and secondly, thinking that consequently clarity can help with extension but not vice-versa. However, a brief review of the history of science suggests otherwise: the puzzles and uncertainties of old concepts are often solved or clarified by extending them into new domains. While extending the scope of our unsettled debates into new territory might ultimately only add to our confusion, it might also give us exactly what we need: new clarifying examples and intuitions. Extending the landscape gives us both new room to populate and new ways to carve. In his light, our revised task is to extend, populate, and carve the landscape.

An extension of the background structure debate towards various approaches to quantum gravity has been carried out in \cite{8}. From such an analysis one can expect to achieve two things: 1) directly, a better understanding of what kinds of background structures and symmetries different theories of quantum gravity might have, and 2) indirectly, a better understanding of how our different notions of background structure relate to one another.

This paper develops another extension, not towards quantum spacetime theories but towards discrete spacetime theories (i.e., lattice theories). These extensions are not unrelated: quantization literally means discretization. Many approaches to quantum gravity assume some kind of discrete spacetime: causal sets, cellular automata, loop quantum gravity, spin foams, etc. Indeed, we have good reason to believe that a full non-perturbative theory of quantum gravity will have something like a finite density of degrees of freedom \cite{8} to borrow a phrase from Achim Kempf \cite{24,28}. From such an analysis one can expect to achieve two things: 1) directly, a better understanding of what kinds of background structures and symmetries our discrete spacetime theories might have, and 2) indirectly, a better understanding of how our different notions of background structure relate to one another.

Central Claims

Why do we need a discrete analog of general covariance? Intuitively, the world might be fundamentally set on a lattice. Indeed, as discussed above, quantum gravity seems to point towards this possibility.

Let’s consider the following empirical situation and follow our first intuitions interpreting it. Suppose that by empirically investigating the world on the smallest scales

representational artifacts and so do not deserve “first billing” so to speak. Hence, I prefer to call them “discrete spacetime theories”. Although I am not quite satisfied with this term either. Indeed, as I will discuss, taking the spacetime itself to be discrete also causes issues: systematically under-predicting symmetries. Presently, I think “discretely representable spacetime theories” is the most apt term for them.

\footnote{This runs parallel in many ways to recent efforts to extend, populate, and carve the quantum-classical landscape \cite{11,17}.}

\footnote{Firstly, basic thought experiments in quantum gravity suggest the existence of something like a minimum possible length scale at approximately the Planck length. Measurements which resolve things at this scale, or attempts to store information at this scale both seem to lead to the creation of black holes. Relatedly, the Bekenstein bound suggests that only a finite amount of information can be stored in any given volume (with this bound scaling with the region’s area) before a black hole forms.}
we discover microscopic symmetry restriction. Namely, we find that only quarter turns or one-sixth turns (i.e., not continuous rotations) preserve the dynamics.

Intuitively, these are “lattice artifacts” which reflect the symmetries of the underlying lattice structure. For instance, Fig. 1 shows a square lattice and a hexagonal lattice. Intuitively, a theory set on a square lattice can only have the symmetries of that lattice (i.e., discrete shifts, quarter rotations, and certain mirror reflections). Similarly for a hexagonal lattice but with one-sixth rotations. Even with an unstructured lattice we would still be restricted to discrete symmetries (i.e., permutations of lattice sites).

Thus, assuming that the world has an underlying lattice structure would explain our restricted rotation symmetries we found at a microscopic scale. Moreover, under this assumption we could discover which kind of lattice structure the world has (e.g., square vs hexagonal) by investigating the theory’s dynamical symmetries and finding the matching lattice structure.

Buying into this underlying-lattice-assumption, one might investigate further and try to discover what kind of interactions there are on this lattice: nearest neighbor, next-to-nearest neighbor, infinite-range, etc. Intuitively, one could discover this sort of thing through microscopic investigations.

Suppose that after substantial empirical investigation we find such “lattice artifacts” and moreover we have great predictive success when modeling the world as being set on (for instance) a square lattice with next-to-nearest neighbor interactions. Does this really prove that the world is fundamentally set on such a lattice? No, all this would prove is: The dynamics of the world can be faithfully represented on such a lattice with such interactions, at least empirically. Anything can be faithfully represented in any number of ways. Some extra-empirical work must be done to know which of these representations we should take seriously. That is, we must ask which parts of the theory are substantive and which parts are representational? We need a discrete analog of general covariance.

Proceeding without one for the moment, there are some intuitive reasons for taking such lattice structures seriously. One’s might have the following three interconnected first intuitions regarding the substantive role that the lattice and lattice structure play in discrete spacetime theories:

1. They restrict our possible symmetries. Taking the lattice structure to be a part of the theory’s fixed background structure, our possible symmetries are limited to those which preserve this fixed structure. As discussed above, intuitively a theory set on a square lattice can only have the symmetries of that lattice. Similarly for a hexagonal lattice, or even an unstructured lattice.

2. Differing lattice structures distinguishes our theories. Two theories with different lattice structures (e.g., square vs hexagonal) cannot be identical. They have different fixed background structures and as therefore (as suggested above) have different symmetries.

3. The lattice is fundamentally “baked-into” the theory. Firstly, it is what the fundamental fields are defined over: they map lattice sites (and possibly times) into some value space. Secondly, the lattice is what the lattice structure structures. Thirdly, it is what limits us to discrete permutation symmetries in advance of further limitations from the lattice structure.

However, as this paper demonstrates, each of the above intuitions are doubly wrong and overhasty.

What goes wrong with the above intuitions is that we attempted to directly transplant our notions of background structure and symmetry from continuous to discrete spacetime theories. This is an incautious way to proceed and is apt to lead us astray. Recall that, as discussed above, our notions of background structure and symmetry are best understood in light of general covariance. It is only once we understand what is substantial and what is representational in our theories, that we have any hope of understanding them. Therefore, we ought to instead first transplant a notion of general covariance into our discrete spacetime theories and then see what conclusions we are led to regarding the role that the lattice and lattice structure play in our discrete spacetime theories. This paper does just that.

As my discrete analog of general covariance will reveal: lattice structure is rather less like a fixed background structure and rather more like a coordinate system, i.e., merely a representational artifact. Indeed, this paper develops a rich analogy between the lattice structures which appear in our discrete spacetime theories and the coordinate systems which appear in our continuum spacetime theories. Three lessons learned throughout this paper⁶ point us in this direction. Each of these lessons negate one of the above discussed intuitions.

Firstly, as I will show, taking any lattice structure seriously as a fixed background structure systematically underpredicts the symmetries that discrete theories can and do have. Indeed, as I will show neither the lattice itself nor the lattice structure in any way restrict a theory’s possible symmetries. In fact, there is no conceptual barrier to having a theory with continuous translation and rotation symmetries formulated on a discrete lattice. As I will discuss, this is analogous to the familiar fact that there is no conceptual barrier to having a continuum theory with rotational symmetry formulated on a Cartesian coordinate system.

⁶ These lessons are also visible in some corners of the physics literature, particularly in the work of Achim Kempf [24-33] among others [35-37]. For an overview see [34].
Secondly, as I will show, discrete theories which are presented to us with very different lattice structures (i.e., square vs. hexagonal) may nonetheless turn out to be completely equivalent theories. Moreover, given any discrete theory with some lattice structure we can always re-describe it using a different lattice structure. As I will discuss, this is analogous to the familiar fact that our continuum theories can be described in different coordinates, and moreover we can switch between these coordinate systems freely.

Thirdly, as I will show, in addition to being able to switch between lattice structures, we can also reformulate any discrete theory in such a way that it has no lattice structure whatsoever. Indeed, we can always do away with the lattice altogether. As I will discuss, this is analogous to the familiar fact that any continuum theory can be written in a generally covariant (i.e., coordinate-free) way.

These three lessons combine to give us a rich analogy between lattice structures and coordinate systems. As I will discuss, there are actually two ways of fleshing out this analogy. Thus, in actuality, we have two discrete analogs to general covariance. These two approaches differ in how they treat lattice structure once it has been revealed to be coordinate-like and so merely representational. In light of this difference I shall call them internal and external, see Sec. VIII. I find reason to favor the external approach, but this will be discussed later. In either case, as one would hope, these discrete analogs of general covariance help us disentangle a discrete theory’s substantive content from its merely representational artifacts.

Having exposed lattice structure as a merely representational artifact, it becomes clear that the lattice structure supposedly underlying any discrete “lattice” theory has the same level of physical import as coordinates do, i.e., none at all. Thus, the world cannot be “fundamentally set on a square lattice” (or any other lattice) any more than it could be “fundamentally set in a certain coordinate system”. Lattice structures are just not the sort of thing that can be fundamental; they are thoroughly representational. Spacetime cannot be discrete (even when it might be representable as such).

Outline of the Paper

In Sec. II I will follow [1] in overviewing the differences between the concepts of general covariance, diffeomorphism invariance, and background independence. To demonstrate these ideas and to lay the groundwork needed to extend them to discrete spacetime theories, I will work through several example theories. Namely, I will consider the Klein Gordon equation and the heat equation. This section will also make an (ultimately wrong) attempt at transferring some of these ideas to discrete spacetime theories generally.

To make this more concrete, in Sec. III I will introduce seven discrete heat equations in an interpretation-neutral way and solve their dynamics. Then, in Sec. IV I will make a first attempt at interpreting these theories. I will (ultimately wrongly) identify their underlying manifold, locality properties, and symmetries. Among other issues, a central problem with this first attempt is that it takes the lattice to be a fundamental part of the underlying manifold and thereby unequivocally cannot support continuous translation and rotation symmetries. This systematically under predicts the symmetries that these theories can and do have.

In Sec. V I will provide a second attempt at interpretation which fixes this issue (albeit in a less than satisfying way). In particular, in this second attempt I deny that the lattice is a fundamental part of the underlying manifold. Instead I “internalize” it. That is, in this second attempt the lattice is associated with the theory’s value space and not the underlying manifold. Fruitfully, this second interpretation does allow for continuous translation and rotation symmetries. Indeed, it exposes such hidden symmetries in our seven discrete heat theories. However, the key move here of “internalization” has several unsatisfying consequences. For instance, the continuous translation and rotation symmetries we find are here classified as internal (i.e., associated with the value space as opposed to the manifold) whereas intuitively they ought to be external.

We thus will need a third attempt at interpreting these theories. However, before that in Sec. VI I will provide an informal overview of the primary mathematical tools used in this paper. Namely, I will review the basics of Nyquist-Shannon sampling theory and bandlimited functions.

With this review complete, in Sec. VII I will use these tools to provide a third and final attempt at interpreting these theories. A perspective similar to this third interpretation has been put forward in the physics literature by Achim Kempf [24–34] among others [35–37]. For an overview see [31]. Like my second attempt, this third interpretation can support continuous translation and rotation symmetries. However, unlike the second attempt it realizes them as external symmetries (i.e., associated with the underlying manifold, not the theory’s value space). Roughly, this is accomplished by (in a principled way) inventing a continuous manifold for the fields to live on. The discrete theory is then embedded onto this manifold as a bandlimited function.

In Sec. VIII I will review the lessons learned in these three attempts at interpretation. As I will discuss, the lessons learned combine to give us a rich analogy between lattice structures and coordinate systems. As I will discuss, there are actually two ways of fleshing out this analogy: one internal and one external. This section spells out these analogies in detail, each of which gives

---

7 There is some subtlety here which will be discussed in Sec. VII
us a discrete analog of general covariance. I find reason to prefer the external notion, but either is likely to be fruitful for further investigation/ use.

Finally, Sec. [X] and Sec. [X] will summarize the results of this paper and provide an outlook of future work.

II. BRIEF REVIEW OF GENERAL COVARIANCE, DIFFEOMORPHISM INVARIANCE, AND BACKGROUND INDEPENDENCE

As discussed in the introduction, a crucial step in the history of GR was Einstein’s adoption of the principle of general covariance. While, ultimately, this principle is merely stylistic (with no physical content per se) it nonetheless commits us to a good and useful style of theorizing. As discussed above, a generally covariant formulation of a theory disentangles its substantive content from its merely representational artifacts. In particular, reformulating in this way more clearly exposes the theory’s geometric background structure, and thereby helps clarify our understanding of the theory’s symmetries.

In the physics literature, three closely related concepts are often confused: general covariance, diffeomorphism invariance, and background independence (i.e., a complete lack of background structure). To demonstrate these ideas and to lay the groundwork needed to extend them to discrete spacetime theories, let’s go through some examples.

A. Klein Gordon Equation

Consider a real scalar field \( \phi : M \rightarrow \mathbb{R} \) with mass \( M \), satisfying the Klein Gordon equation,

\[
\partial_t^2 \phi(t, x, y, z) = (\partial_x^2 + \partial_y^2 + \partial_z^2 - M^2) \phi(t, x, y, z). \tag{1}
\]

This formulation is not generally covariant since when it is rewritten in different coordinates it changes form. For instance, in the coordinates \( t' = t, x' = x, y' = y \) and \( z' = z + \frac{1}{2}a t^2 \), we have,

\[
\partial_{t'}^2 \phi(t', x', y', z') = (\partial_{x'}^2 + \partial_{y'}^2 + \partial_{z'}^2 - M^2) \phi(t', x', y', z') - a \partial_{z'} \phi(t', x', y', z'). \tag{2}
\]

An extra term shows up when we move into a non-inertial coordinate system. Let’s fix this. Introducing a fixed Lorentzian metric field, \( \eta^{ab} = \text{diag}(-1,1,1,1) \) we can rewrite Eq. (1) as,

\[
(\eta^{\mu\nu} \partial_\mu \partial_\nu - M^2) \phi = 0, \tag{3}
\]

where \( x^\mu = (t, x, y, z) \) and \( \partial_\mu = (\partial_t, \partial_x, \partial_y, \partial_z) \). Unfortunately this is still not generally covariant. If we rewrite Eq. (3) in arbitrary coordinates, \( x'^\mu \), we find,

\[
(\eta^{\sigma\rho} \frac{\partial x'^\mu}{\partial x^{\sigma}} \frac{\partial x'^\nu}{\partial x^{\rho}} - M^2) \phi + \eta^{\sigma\rho} \frac{\partial^2 x'^\mu}{\partial x^{\sigma} \partial x^{\rho}} \partial_\mu \phi = 0.
\]

This formulation, however, is coordinate-invariant. If we change coordinates again to \( x'^\mu \), the equation keeps the same form except with \( x'^\mu \rightarrow x''^\mu \).

This demonstrates an ambiguity in the usage of the term generally covariant above [38]; coordinate-independent can mean coordinate-invariant (but still written in terms of coordinates) or coordinate-free (written without any reference to coordinates at all). The real benefits of general covariance come from having a coordinate-free representation. This is the notion of general covariance relevant throughout this paper. To achieve general covariance we need to reformulate Eq. (3) in the coordinate-free language of differential geometry.

Before this however, I need to introduce some terminology. Throughout this paper I will associate with any classical spacetime theory with two spaces of models [1,39]: kinematically possible models (KPMs) and dynamically possible models (DPMs). Roughly, these are off-shell and on-shell solutions.

KPMs are all of the mathematical objects which have the right sort of structures to make sense as models of our theory (regardless of whether they satisfy the dynamics). These are represented as an ordered collection of the theory’s manifold together with its geometric fields and matter fields. For our Klein Gordon example, the KPMs are given by \( \mathcal{M} \)

\[
\text{SR1: KPMs:} \quad (\mathcal{M}, \eta_{ab}, \phi) \tag{4}
\]

where \( \mathcal{M} \) is a differentiable (3+1)-manifold, \( \eta_{ab} \) is a fixed Lorentzian metric field\(^8\), and \( \phi : \mathcal{M} \rightarrow \mathbb{R} \) is a dynamical real scalar field.

By contrast, a theory’s dynamically possible models (DPMs), are the subset of the KPMs which obey the theory’s dynamical equations. For SR1 the DPMs are picked out by,

\[
\text{SR1: DPMs:} \quad (\eta_{ab} \nabla_a \nabla_b - M^2) \phi = 0,
\]

where \( \nabla_a \) is the unique covariant derivative operator compatible with the metric, (i.e. with \( \nabla_a \eta_{ab} = 0 \)). This formulation of the Klein Gordon equation is now generally covariant (in the coordinate-free sense).

\[
\text{Klein Gordon - Symmetries}
\]

Let us now use this generally covariant formulation to help us understand this theory’s symmetries. It is important to distinguish two kinds of symmetry [5] (spacetime

---

\(^8\) The name SR1 is picked to follow the notation set in [1].

\(^9\) It’s important to note the difference between \( \eta_{\mu\nu} \) and \( \eta_{ab} \). Any tensor object with greek indices (e.g., \( \eta_{\mu\nu} \)) is to be understood as the components of a certain tensor in a particular coordinate system. By contrast, any tensor object with roman indices (e.g., \( \eta_{ab} \)) is coordinate-free, its “indices” are merely there to remind us of the rank of this tensor and to help us see how it interacts with the other tensor objects.
symmetries and dynamical symmetries) related to two
different kinds of fields [1] (fixed fields and dynamical
fields). The latter distinction is that fixed fields are fixed
by fiat to be the same in every model. By contrast, dy-
namical fields can vary from model to model. In SR1,
$\eta_{ab}$ is fixed whereas $\phi$ is dynamical.

The distinction regarding symmetries is as fol-
lows. Spacetime symmetries are those diffeomorphisms,
$d \in \text{Diff}(\mathcal{M})$, which preserve the theory’s fixed fields (re-
gardless of the dynamical equations). Dynamical sym-
metries are those diffeomorphisms, $d \in \text{Diff}(\mathcal{M})$, which
map solutions to solutions when applied to the dynami-
cal fields of our models (leaving the fixed fields fixed). In
either case, let us call these external symmetries.

For SR1 our only fixed field is $\eta_{ab}$. Thus SR1’s
spacetime symmetries are those diffeomorphisms$^{10}$ with
$d^*\eta_{ab} = \eta_{ab}$. Only a small subset of $\text{Diff}(\mathcal{M})$ have this
property, namely the Poincaré group.

For SR1, given a generic DPM, $(\mathcal{M}, \eta_{ab}, \phi)$ we can ap-
ply a generic diffeomorphism $d \in \text{Diff}(\mathcal{M})$ to its dynamical
fields to get some KPM, $(\mathcal{M}, \eta_{ab}, d^*\phi)$. This dif-
feomorphism $d$ is a dynamical symmetry when for every
input DPM this output KPM is a solution to the dynami-
cal equations (i.e., is also a DPM). It turns out that all and only $d$
in the Poincaré group maps solutions to solutions. Thus
the dynamical symmetry group of SR1 is the Poincaré
group.

In the above example, the theory’s spacetime symme-
tries and its dynamical symmetries match. There are
good reasons$^{11}$ for these to match in general$^{5}$ but they
won’t always.$^{12}$ In any case it is important to keep
them separate conceptually. Unless otherwise specified,
throughout this paper unqualified references to symme-
tries should be understood as meaning dynamical sym-
metries.

For our later reasoning, it is important to stress two
points here. Firstly, it’s important to note that neither of
these notions of symmetry have anything to do with coor-
dinate representations. Recall from Sec.I Kretschmann’s
point that any theory can be represented in terms of any
coordinates (or, indeed, in terms of no coordinates at all).
Another immediate consequence of this is the fa-
miliar fact that coordinate systems do not in any way
restrict a theory’s possible symmetries. It is a familiar
fact that there is no conceptual barrier to having a con-
tinuum theory with rotational symmetry formulated on
a Cartesian coordinate system.

Secondly, it is important to note that both types of ex-
ternal symmetry discussed above (dynamical and spacetime)
are restricted to be within $\text{Diff}(\mathcal{M})$. Why the $\mathcal{M}$ in $\text{Diff}(\mathcal{M})$? Because this is the place where the the-
ory’s fundamental fields map from. But why the Diff in
$\text{Diff}(\mathcal{M})$? Because this is the relevant class of automor-
phisms for differentiable manifolds. Of course, more can
be said about each of these points, however here it suf-
fices to note that if our fields are mapped out of some
manifold-like space $Q$ then we would expect to find our
symmetries only within $\text{Auto}(Q)$ for some relevant notion
of automorphism.

For completeness I ought to mention another type of
symmetry that a theory might have: internal symme-
tries and relatedly gauge symmetries. In any given theory,
the dynamical fields will map$^{13}$ from the manifold $\mathcal{M}$
into some value space $\mathcal{V}$ as $\phi : \mathcal{M} \rightarrow \mathcal{V}$. We may find
additional internal symmetries of our theory within the
automorphisms of this value space, $\text{Auto}(\mathcal{V})$. We might
also find our theory has gauge symmetries by allowing
these internal automorphisms to vary smoothly across
the manifold.

For SR1, the value space of our dynamical field $\phi$ is
the real numbers $\mathcal{V} = \mathbb{R}$. Taking $\mathbb{R}$ to be a vector space
here, the relevant automorphisms are $\text{Aff}(\mathbb{R})$ consisting of
linear-affine transformations. Among these, SR1’s inter-
ernal symmetries are only global re-scalings of $\phi$ as $\phi \mapsto A \phi$
for some $A \in \mathbb{R}$. We find no gauge symmetry trying to
localize this.

**Klein Gordon - Background Structure**

Let us now use this generally covariant formulation to
help us understand this theory’s background structure.
As mentioned in the introduction, there is ongoing de-
bate$^{14} 2 6 10$ about what exactly should and should
not count as background structure. However, it is widely
agreed that any fixed field ought to count as background
structure. Thus, for SR1 the fixed Lorentzian metric $\eta_{ab}$
counts as background structure.

From this one may be tempted to reason (poorly) as
follows. A theory’s spacetime symmetries are just those
transformations which preserve its fixed fields. Any fixed

$^{10}$ In this paper for simplicity I will not differentiate between the
pull back and push forward of $d$. Both of these will be referred
to as $d^*$ which will be called the drag along map. $d^*$ is whatever
modification of $d$ is demanded by the context.

$^{11}$ If there are more dynamical symmetries than spacetime symme-
tries, then some of the theory’s fixed fields are not being used
by any of the dynamics, in which case why are they there? Con-
versely, if there are more spacetime symmetries than dynamical
symmetries then it appears some necessary piece of spacetime
structure is missing. E.g., consider a case where the dynamics
are not boost invariant and so implicitly pick out a rest frame,
but somehow the spacetime comes equipped with no rest frame.

$^{12}$ They would not match for instance if, as a piece of fixed field,
we had included a time orientation field, $\chi$ which distinguishes
the future light cone from the past light cone at each event.

In this case, the dynamical symmetries would still be the Poincaré
group, but the spacetime symmetries would only be the time-
orientation preserving subset of these. In general, the spacetime
symmetries can be smaller than the dynamical symmetries if
there is some piece of spacetime structure which is unused by
the dynamics.

$^{13}$ More generally, the fields might be defined as sections of a fiber
bundle of $\mathcal{V}$ over $\mathcal{M}$, but let’s neglect that complication here.
fields count as background structures. As such, minimizing background structure is the same as maximizing spacetime symmetry. Therefore, background independence is the same concept as diffeomorphism invariance.

However, this reasoning and its conclusion are in error. There can be other kinds of background structure than fixed fields. Indeed, following [1], we can reformulate SR1 as,

\[
\begin{align*}
\text{SR2:} & \quad \text{KPMs:} \quad \langle \mathcal{M}, g_{ab}, \phi \rangle, \\
& \quad \text{DPMs:} \quad (g^{ab} \nabla_a \nabla_b - M^2) \phi = 0 \\
& \quad R^a_{bcd} = 0.
\end{align*}
\]

Here the fixed Lorentzian metric field, \(\eta_{ab}\), has been replaced with a dynamical metric field, \(g_{ab}\), with signature \((-1,1,1,1)\). The dynamical metric field varies from model to model and obeys a new dynamical equation, \(R^a_{bcd} = 0\), where \(R^a_{bcd}\) is the Riemann tensor associated with \(\nabla_a\). Note that SR2 has the same KPMs as the Klein Gordon equation in GR would:

\[
\begin{align*}
\text{GR:} & \quad \text{KPMs:} \quad \langle \mathcal{M}, g_{ab}, \phi \rangle, \\
& \quad \text{DPMs:} \quad (g^{ab} \nabla_a \nabla_b - M^2) \phi = 0 \\
& \quad G_{ab} = 8\pi T_{ab}.
\end{align*}
\]

Indeed, SR2 and GR only differ in their dynamical equations. One may favor SR2 over SR1 on these grounds.

What are SR2’s symmetries? SR2 has no fixed fields. As such, its spacetime symmetries are the full diffeomorphism group, \(\text{Diff}(\mathcal{M})\). This theory’s dynamical symmetries are also the full diffeomorphism group, \(\text{Diff}(\mathcal{M})\). Thus, SR2 is diffeomorphism invariant.

Ought we to conclude (using the above erroneous logic) that because SR2 is diffeomorphism invariant that it is also background independent? Clearly not. Intuitively SR1 and SR2 should have the same background structures. The difference is that in SR2 this background structure is hidden whereas SR1 is in a sense more honest, declaring its background structure up front as a fixed field. For this reason one may prefer SR1 to SR2.

However, as SR2 clearly demonstrates, we cannot in general expect our theories to be so honest about their background structures. There can be other sorts of background structure in our theories than those which are declared upfront as fixed fields. There is a wide literature attempting to find a way of systematically identifying these hidden background structures [1] [2] [5] [10].

To demonstrate these ideas further and lay the groundwork for what is to come, let’s consider one more example.

**B. Heat Equation**

Let us next consider a real scalar field \(\psi : \mathcal{M} \to \mathbb{R}\) satisfying the following one and two-dimensional heat equations:

**Heat Equation 00 (H0):**

\[
\partial_t \psi(t, x) = \alpha_0 \nabla^2 \psi(t, x) = \alpha_0 \frac{\partial^2}{\partial x^2} \psi(t, x)
\]

**Heat Equation 0 (H0):**

\[
\partial_t \psi(t, x, y) = \frac{\alpha_0}{2} (\partial_t^2 + \partial_x^2 + \partial_y^2) \psi(t, x, y)
\]

with some diffusion rate \(\alpha_0 \geq 0\). Focusing on the two-dimensional case, after substantial work, one can rewrite this equation in the coordinate-free language of differential geometry as follows:

\[
\begin{align*}
\text{H0:} & \quad \text{KPMs:} \quad \langle \mathcal{M}, t_{ab}, h^{ab}, \nabla_a, T^a, \psi \rangle \\
& \quad \text{DPMs:} \quad T^a \nabla_a \psi = \frac{\alpha_0}{2} h^{bc} \nabla_b \nabla_c \psi.
\end{align*}
\]

The geometric objects used in this formulation are as follows. \(\mathcal{M}\) is a 2+1 dimensional differentiable manifold. \(h^{ab}\) and \(t_{ab}\) are space and time metric fields. They are each symmetric with signatures \((0,1,1)\) and \((1,0,0)\) respectively. Moreover, these metrics are orthogonal to each other (i.e., with \(h^{ab} t_{bc} = 0\)). These metrics allow us to compute lengths and durations. \(\nabla_a\) is a covariant derivative operator which is compatible with these metrics (i.e., \(\nabla_a h^{bc} = 0\) and \(\nabla_a t_{bc} = 0\)). Note that \(\nabla_a\) is not uniquely determined by the metric in this case because neither \(h^{ab}\) nor \(t_{ab}\) are invertible. We take \(\nabla_a\) to be flat, \(R^a_{bcd} = 0\). The covariant derivative operator \(\nabla_a\) allows us to do parallel transport and compute derivatives of non-scalar fields.

The quadruple, \(\langle \mathcal{M}, t_{ab}, h^{ab}, \nabla_a \rangle\), picked out by the above discussed objects is a Galilean spacetime [5]. However, in addition to these structures, our spacetime also has a constant unit time-like vector field \(T^a\) with \(\nabla_a T^b = 0\) and \(t_{ab} T^a T^b = 1\). This vector field picks out a standardized way of moving forward in time (i.e, translation generated by \(T^a \nabla_a\)). That is, \(T^a\) provides a rest frame.

Each of the above discussed objects is considered fixed in this formulation: they do not vary from model to model and do not evolve dynamically. The only dynamical field in this formulation is the real scalar field \(\psi : \mathcal{M} \to \mathbb{R}\) representing the temperature field. Mathematical structures satisfying the above conditions (independent of whether it satisfies the dynamics) are the theory’s KPMs. This theory’s DPMs are the subset of the KPMs which additionally satisfy the theory’s dynamics. The dynamical equation in Eq. (10) says that the derivative of the temperature field in the \(T^a\) direction is proportional to its second derivative in space.

What are the heat equation’s symmetries? In contrast with SR1, the H0 has many more fixed fields: \(t_{ab}, h^{ab}, \nabla_a,\) and \(T^a\) are all fixed. Ultimately, this restricts the theory’s spacetime symmetries down to the two-dimensional Euclidean group (spacial translations, rotations and reflections) plus constant time shifts. Note that time inversions and Galilean boosts are not symmetries because each of these fail to preserve our rest frame \(T^a\).
This theory’s dynamical symmetries match its spacetime symmetries.

And what about internal symmetries? Taking $V = \mathbb{R}$ to be a vector space here, the relevant automorphisms are $\text{Aff}(\mathbb{R})$ consisting of linear-affine transformations. Within $\text{Aff}(\mathbb{R})$ any transformation preserves solutionhood, $\psi \rightarrow A\psi + b$ for $A, b \in \mathbb{R}$. Localizing this internal symmetry does not lead to any gauge symmetries.

What are the heat equation’s background structures? The above discussed fixed fields will surely count as background structures. While conceivably there could be more background structures than just these, given the simplicity of this theory that seems unlikely.

C. First (Naive) Intuitions about Discrete Spacetime Theories

Let us now attempt (naively) to apply some of the above discussion to discrete theories. Ultimately, these first intuitions are naive because we have not yet developed a discrete analog of general covariance. Thus, we have little clarity as to what parts of these theories are substantive as opposed to merely representational. Proceeding anyway despite this, we are apt to get things wrong.

This subsection has two purposes. The first is to informally introduce a few discrete spacetime theories (i.e., lattice theories) before giving a more general characterization. Secondly, I would like to give some (faulty) support for the (ultimately wrong) first intuitions given in the introduction preceding the central claims. The following two sections will then repeat these objectives in much more detail applied to several discrete heat equations.

Let’s begin with three such discrete heat equations: one set on a uniform 1D lattice, one set on square lattice and one set on a hexagonal lattice. Each of the following have only nearest-neighbor (N.N.) interactions.

1D N.N. Heat Equation (H1):

$$\frac{d}{dt} \phi_n(t) = \alpha [\phi_{n-1}(t) - 2\phi_n(t) + \phi_{n+1}(t)]$$

Square N.N. Heat Equation (H4):

$$\frac{d}{dt} \phi_{n,m}(t) = \frac{\alpha}{2} [\phi_{n-1,m}(t) - 2\phi_{n,m}(t) + \phi_{n+1,m}(t) + \phi_{n,m-1}(t) - 2\phi_{n,m}(t) + \phi_{n,m+1}(t)]$$

Hexagonal N.N. Heat Equation (H5):

$$\frac{d}{dt} \phi_{n,m}(t) = \frac{\alpha}{3} [\phi_{n-1,m}(t) - 2\phi_{n,m}(t) + \phi_{n+1,m}(t) + \phi_{n,m-1}(t) - 2\phi_{n,m}(t) + \phi_{n,m+1}(t) + \phi_{n+1,m-1}(t) - 2\phi_{n,m}(t) + \phi_{n-1,m+1}(t)]$$

These theories are named H1, H4, and H5 in anticipation of their treatment in Sec. VII! In the latter two cases the lattice sites are indexed $(n, m)$ as shown in Fig. 1. In each case, the right-hand side is the best approximation of the second derivative possible using only nearest neighbor interactions.

Note that in each of the above theories time is still treated as continuous. This doesn’t have to be the case as we could also consider the following example:

Square N.N. Klein Gordon Equation:

$$\phi_{j-1,n} - 2\phi_{j,n} + \phi_{j+1,n} = \phi_{j-1,n} - 2\phi_{j,n} + \phi_{j,n+1} - \mu^2 \phi_{j,n}$$

with $j$ indexing time and $n$ indexing space and $\mu$ playing the role of the field’s mass.

Let’s start (naively) interpreting these theories by taking their above formulations seriously, i.e. Eq. (11), Eq. (12) and Eq. (13). Taken literally what are these theories about?

Beginning with H1, H4 and H5, these theories are intuitively about a field $\phi(t)$ which maps lattice sites ($\ell \in \mathbb{L} \cong \mathbb{Z}^1$) and times ($t \in \mathbb{R}$) into temperatures ($\phi(t) \in \mathbb{R}$). That is a field $\phi : \mathbb{Q} \rightarrow \mathbb{V}$ with manifold $\mathbb{Q} = \mathbb{L} \times \mathbb{R}$ and value space $\mathbb{V} = \mathbb{R}$.

Similarly, for the discrete Klein Gordon example the fundamental field seems to be $\phi : \mathbb{Q} \rightarrow \mathbb{V}$ except that it has a manifold $\mathbb{Q} = \mathbb{L}$ with no times. While this paper will only explicitly deal with theories with $\mathbb{Q} = \mathbb{L} \times \mathbb{R}$, I expect that the central claims hold true in either case.

In either case, taking $\phi : \mathbb{Q} \rightarrow \mathbb{V}$ seriously as a fundamental field leads us to thinking of $\mathbb{Q}$ as the theory’s underlying manifold and $\mathbb{V} = \mathbb{R}$ as the theory’s value space. It is important to note that in this interpretation, $\mathbb{Q}$ is the entire manifold, it is not being thought of as embedded in some larger manifold. (However, a view like this will be considered in our third interpretation in Sec. VII)

With this manifold and value space picked out, what can we expect of these theories’ symmetries? As discussed earlier in this section, if $\mathbb{Q}$ is a theory’s underlying manifold then its external symmetries (either spacetime or dynamical) are restricted $\text{Auto}(\mathbb{Q})$ for some relevant notion of automorphism. Similarly, the theory’s internal symmetries are restricted to $\text{Auto}(\mathbb{V})$. We might also have gauge symmetries which mix these two. Returning to $\text{Auto}(\mathbb{Q})$ however what is the relevant notion of automorphism?

Answering this question will require us to distinguish what structures are “built into” $\mathbb{Q}$ and what are “built on top of” $\mathbb{Q}$. The analogous distinction in the continuum case is that we generally take the manifold’s differentiable structure to be built into it while the Minkowski metric, for instance, is something additional built on top of the manifold. In this paper, I am officially agnostic on where we draw this line in the discrete case. However, for didactic purposes I will here be as conservative as possible giving $\mathbb{Q}$ as little structure as is sensible. Note that the more structure we associate with $\mathbb{Q}$ the smaller the class of relevant automorphisms $\text{Auto}(\mathbb{Q})$. Thus, I am taking $\text{Auto}(\mathbb{Q})$ to be as large as it can reasonably be.

For the second kind of discrete theory (i.e., those with both space and time discrete) the minimal structure we
can reasonably associate with $Q = L$ is that of a set. As such the largest $\text{Auto}(Q)$ could reasonably be is permutations of the lattice sites, $\text{Auto}(Q) = \text{Perm}(L)$.

For the first kind of discrete theory (i.e., those with continuous time) the minimal structure we can reasonably associate with $Q = L \times \mathbb{R}$ is that of a set times a differentiable manifold. In this case, the largest $\text{Auto}(Q)$ could reasonably be is permutations of the lattice sites together with time reparametrizations, $\text{Auto}(Q) = \text{Perm}(L) \times \text{Diff}(\mathbb{R})$.

Recall that in addition to $\text{Auto}(Q)$ we might also have internal or gauge symmetries. While in general there may be abundant internal or gauge symmetries, for the present cases there are not many. In particular, for all of the above-mentioned theories we only have $\mathcal{V} = \mathbb{R}$. Moreover, these theories are all linear (solutions sum to solutions) such that it makes sense to structure $\mathbb{R}$ as an affine vector space. Thus, $\text{Auto}(\mathcal{V}) = \text{Aff}(\mathbb{R})$ such that our internal symmetries are linear-affine rescalings of $\phi$. Moreover, localizing these internal symmetries reveal no gauge symmetries in these examples.

Thus, in total, for H1 H4 and H5 the widest scope of symmetry transformations available to us is:

$$s : \phi(t) \mapsto c_1 \phi_{P(t)}(\tau(t)) + c_2$$

for some $c_1, c_2 \in \mathbb{R}$, some smooth $\tau(t)$, and some permutation $P : L \to L$. Similarly, for the discrete Klein Gordon example we have

$$s : \phi(t) \mapsto c_1 \phi_{P(t)} + c_2$$

dropping the time label.

In either case, however, it should be clear that our theory cannot have continuous spacial translations and rotations (at least not while interpreting $\phi : Q \to \mathcal{V}$ and consequently $Q$ as fundamental as we are here). I have thus spelled out some (ultimately faulty) support for the (ultimately wrong) first intuitions put forward in Sec. II. To summarize, taking the initial presentation of these theories literally, we are led to think of $\phi : Q \to \mathcal{V}$ to be fundamental with manifold $Q$ and value space $\mathcal{V}$. Reasoning from here we found that the lattice itself (in addition to further lattice structure) restricts our theories’ symmetries to be discrete: it appears we cannot have continuous spacial and rotational symmetries.

As I will show in Sec. III this systematically under predicts the symmetries that discrete spacetime theories can and do have. Fixing this issue will lead us to develop two discrete analogs of general covariance. However, before this let me introduce several discrete heat equations around which the rest of the paper will be framed.

III. SEVEN DISCRETE HEAT EQUATIONS

In this section I will introduce seven discrete heat equations (H1-H7) in an interpretation-neutral way and solve their dynamics. In the previous section H1, H4 and H5 were already introduced and analyzed somewhat. Here we are starting again clean.

In particular, in the previous section casual comparison was made between parts of these theories’ dynamical equations and various approximations of the second derivative. While, as I will discuss, such comparisons can be made, to do so immediately is unearned. It comes dangerously close to imagining the lattices shown in Fig. 1 as being embedded in a continuous manifold. This may be something we want to do later, but it is a non-trivial interpretational move which ought not be done so casually.

Crucially, in this section I will be analyzing these theories as discrete-native theories. As such, it’s important to think of the following discrete spacetime theories as self-sufficient theories in their own right. We must not begin by thinking of them as various discretizations or bandlimitations of the continuum theories. While, as I will discuss, these discrete theories have some notable relationships to various continuum theories it is important to resist any temptation to see these continuum theories as “where they came from”.

Moreover, the previous section casually associated these theories with the lattice structures shown in Fig. 1. Namely, H4 was associated with a square lattice and H5 with a hexagonal lattice. Making such associates ab initio is unwarranted. While we may eventually associate these theories with those lattice structures we cannot do so immediately. Such an association would need to be made following careful consideration of the dynamics. Beginning here in an interpretation neutral way these theories ought to be seen as being defined over a completely unstructured lattice. That is, at this point the set of labels for the lattice sites, $L$, is just that, an unstructured set.

With these words of caution in mind, let’s introduce some dynamics. All seven theories consider a field $\phi : Q \to \mathbb{R}$ with $Q = L \times \mathbb{R}$. That is, all seven theories consider a real scalar field $\phi(t)$.

As a first example consider a field $\phi(t)$ which under some convenient relabeling of the lattice sites, $\ell \mapsto n$, satisfies Eq. (11), namely,

$$H1 : \frac{d}{dt} \phi_n(t) = \alpha [\phi_{n-1}(t) - 2\phi_n(t) + \phi_{n+1}(t)].$$

At the risk of repeating myself, these $n \in \mathbb{Z}$ are just labels. The fact that our labels $n \in \mathbb{Z}$ are in a sense equidistant from each other does not force us to think of the lattice sites as being equidistant from each other or on a uniform grid. Nor are we forced to think that “the distance between lattice sites” is to be meaningful at all. Dynamical considerations may later push us in this direction, but the mere convenience of this labeling should not.

For practical applications it is convenient to rewrite the dynamics by collecting these field values $\phi_n(t)$ into an infinite dimensional vector as

$$\phi(t) := (\ldots, \phi_{-1}(t), \phi_0(t), \phi_1(t), \ldots) \in \mathbb{R}^L.$$
In these terms the dynamics of H1 is given by,

**Heat Equation 1 (H1):**

\[
\frac{d}{dt} \phi(t) = \alpha \Delta^2_{(1)} \phi(t)
\]

where \(\Delta^2_{(1)}\) is the following bi-infinite Toeplitz matrix:

\[
\Delta^2_{(1)} = \frac{1}{2} \{ \Delta^+_{(1)}, \Delta^-_{(1)} \} = \text{Toeplitz}(1, -2, 1)
\]

\[
\Delta^+_{(1)} = \text{Toeplitz}(0, -1, 1)
\]

\[
\Delta^-_{(1)} = \text{Toeplitz}(-1, 1, 0)
\]

Recall that Toeplitz matrices are so called diagonal-constant matrices with \([A]_{i,j} = [A]_{i+1,j+1}\). Thus, the values in the above expression give the matrix’s values on either side of the main diagonal.

In addition to H1, I will also consider two more theories with the following dynamics:

**Heat Equation 2 (H2):**

\[
\frac{d}{dt} \phi(t) = \alpha \Delta^2_{(2)} \phi(t)
\]

**Heat Equation 3 (H3):**

\[
\frac{d}{dt} \phi(t) = \alpha D^2 \phi(t).
\]

where

\[
\Delta^2_{(2)} = \text{Toeplitz}(-1, 4, 5, 4, -1)
\]

\[
D = \text{Toeplitz}(\ldots, -1, 1, -1, 1, 1, 2, 1, 0, 1, 1, 2, 3, 4, 5, \ldots)
\]

\[
D^2 = \text{Toeplitz}(\ldots, -2, 2, -2, 2, -2, 2, -2, 2, -2, 2, -2, 2, -2, 2, -2, 2, \ldots).
\]

Although above I warned about thinking in terms of derivative approximations prematurely, a few comments are here warranted. Suppose we were to somehow imagine these \(\phi_n(t)\) values as coming from some continuous function \(\phi(t, x)\) by either sampling it or taking local averages on/around some uniform lattice \(x_n = na\). Vectorizing these values as in Eq. [15] and applying any of the above defined Toeplitz matrices would give us approximations to the derivative of \(\phi(t, x)\). Namely, \(\Delta^+_n / a\) would be associated with the forward derivative approximation, \(\Delta^-_n / a\) would be associated with the backwards derivative approximation, and \(\Delta^2_{(1)} / a^2\) would be associated with the nearest neighbor second derivative approximation,

\[
\Delta^2_{(1)} / a^2 : \quad \partial^2 f(x) \approx \frac{f(x-a) - 2f(x) + f(x+a)}{a^2}.
\]

Similarly \(\Delta^2_{(2)} / a^2\) is related to the next-to-nearest-neighbor approximation to the second derivative.

Notice that the longer range we make our derivative approximations the more accurate they can be. The operator \(D\) (and its square \(D^2\)) in some sense are the best discrete approximations to the derivative (and second derivative) possible. The defining property of \(D\) is that it is diagonal in the (discrete) Fourier basis with spectrum,

\[
\lambda_D(k) = ik
\]

for \(k \in [-\pi, \pi]\). This is in tight connection with the continuum derivative operator \(\partial_x\) which is diagonal in the (continuum) Fourier basis with spectrum \(\lambda_{\partial_x}(k) = ik\) for \(k \in [-\infty, \infty]\).

Alternatively, one can construct \(D^2\) in the following way: generalize \(\Delta^2_{(1)} / a^2\) and \(\Delta^2_{(1)} / a^2\) to \(\Delta^2_{(n)} / a^2\) namely the best second derivative approximation considering which considers \(n\)th neighbors to either side. Taking the limit \(n \to \infty\) gives \(D^2 = \lim_{n \to \infty} \Delta^2_{(n)}\). Other aspects of \(D\) will be discussed in Sec. [XXI including the related derivative approximation Eq. [53], but enough has been said for now.

While these connections to derivative approximations allow us to export some intuitions from the continuum theories into these discrete theories, we must resist this (at least for now). In particular, I should stress again that we should not be thinking of any of H1, H2 and H3 as coming from the continuum theory under some approximation of the derivative. Rather, let us pretend these theories “came from nowhere” and let us see what sense we can make of them.

In addition to H1-H3, I will consider four more discrete heat equations (two of these have already been introduced above). First, consider a field \(\phi(t)\) which under some convenient relabeling of the lattice sites, \(\ell \mapsto (n, m)\), satisfies Eq. [12], namely,

\[
\text{H4:} \quad \frac{d}{dt} \phi_{n,m}(t) = \alpha \left[ \phi_{n-1,m}(t) - 2\phi_{n,m}(t) + \phi_{n+1,m}(t) + \phi_{n,m-1}(t) - 2\phi_{n,m}(t) + \phi_{n,m+1}(t) \right]
\]

As before, it will be convenient to organize this theory’s field values into a vector. We can handle the two indices by introducing a tensor product structure into the vector space, \(\phi(t) \in \mathbb{R}^L = \mathbb{R}^2 \otimes \mathbb{R}^2\), with the first tensor factor corresponding to the first index \(n\) and the second tensor factor corresponding to the second index \(m\). In these terms the dynamics of H4 given above is given by,

**Heat Equation 4 (H4):**

\[
\frac{d}{dt} \phi(t) = \alpha \left( \Delta^2_{(1),n} + \Delta^2_{(1),m} \right) \phi(t),
\]

where the notation \(A_n := A \otimes 1\) and \(A_m := 1 \otimes A\) mean \(A\) acts only on the first or second tensor space respectively. Thus, \(\Delta^2_{(1),n}\) is just \(\Delta^2_{(1)}\) applied to the first index, \(n\).

Likewise \(\Delta^2_{(1),m}\) is just \(\Delta^2_{(1)}\) applied to the second index, \(m\).
A similar treatment of Eq. (13) gives us,

**Heat Equation 5 (H5):**

\[
\frac{d}{dt} \phi(t) = \frac{\alpha}{3} \left[ \Delta^2_{(1),n} + \Delta^2_{(1),m} + \frac{1}{2} \left\{ \Delta^+_m - \Delta^-_m, \Delta^-_n - \Delta^+_n \right\} \right] \phi(t),
\]

where the curly brackets \( \{A, B\} = AB + BA \) indicate the anticommutator. While the third term looks complicated, it is just the analog of \( \Delta^2_{(1),n} \) and \( \Delta^2_{(1),m} \) but in the \( n - m \) direction.

At the risk of repeating myself endlessly, in both H4 and H5 these \( n, m \in \mathbb{Z} \) are just labels of lattice sites. The fact that our labels \( n, m \in \mathbb{Z} \) can be thought of as a square lattice does not force us to think of the lattice sites as being arranged in a square lattice. If we come to think anything like this, it should be by investigating the dynamics of these theories. Indeed, for H5 we will find this is not the case.

Finally, in addition to H4 and H5 I consider the following two theories:

**Heat Equation 6 (H6):**

\[
\frac{d}{dt} \phi(t) = \frac{\alpha}{2} \left( D_n^2 + D_m^2 \right) \phi(t)
\]

**Heat Equation 7 (H7):**

\[
\frac{d}{dt} \phi(t) = \frac{\alpha}{3} \left( D_n^2 + D_m^2 + (D_m - D_n)^2 \right) \phi(t).
\]

Having introduced these seven theories, let us next solve their dynamics.

### A. Solving Their Dynamics

To gain some intuition about H1-H7, let us next solve their dynamics. Conveniently, each of H1-H7 admit planewave solutions. Moreover, in each case these planewave solutions form a complete basis of solutions.

Considering first H1-H3 we have solutions,

\[
\phi_n(t; k) = e^{-ikn} e^{-\Gamma(k)t}.
\]

with \( k \in [-\pi, \pi] \). Note that extending \( k \) outside of this range does not yield new planewave solutions due to Euler’s identity, \( \exp(2\pi i) = 1 \). This is related to the phenomena of aliasing in digital image processing. For H1-H3 the wavenumber-dependent decay rate, \( \Gamma(k) \), for each theory is given by:

- **H1:** \( \Gamma(k) = \alpha (2 - 2\cos(k)) \)
- **H2:** \( \Gamma(k) = \frac{\alpha}{6} \left( \cos(2k) - 16 \cos(k) + 15 \right) \)
- **H3:** \( \Gamma(k) = \alpha k^2 \).

Note that \( \Gamma(k) \) for H3 follows Eq. (27), essentially from the definition of \( D \).

**Fig. 2.** The decay rates for the planewave solutions to the discrete heat equations are plotted as a function of wavenumber for H1, H2 and H3 (bottom to top).

Fig. 2 shows these decay rates as a function of wavenumber. Thus, what sets these theories apart is the rate at which high frequency planewaves decay. H1-H3 more-or-less agree at low frequencies, with each saying that if \( k \ll \pi \) then \( \Gamma(k) \approx \alpha k^2 \). If we consider only solutions with all or most of their wavenumber support with \( k \ll \pi \), we have an approximate one-to-one correspondence between the solutions to these theories. This is roughly why each of these theories have the same continuum limit, namely H00 defined in Sec. [II].

Note that the decay rate \( \Gamma(k) \) for H3 exactly matches the decay rate of the continuum theory, at least for \( k \in [-\pi, \pi] \). Note also that H2 is a nearer match to the continuum theory for \( k \ll \pi \) than H1 is.

\[
\begin{align*}
H1 & : \Gamma(k) = \alpha \left( k^2 - \frac{k^4}{12} + \mathcal{O}(k^6) \right) \\
H2 & : \Gamma(k) = \alpha \left( k^2 - \frac{k^6}{90} + \mathcal{O}(k^8) \right) \\
H3 & : \Gamma(k) = \alpha k^2.
\end{align*}
\]

This is due to its longer range coupling giving a better approximation of the derivative. In terms of converging to the continuum limit, one can expect H3 to outpace H2 which outpaces H1.

While interesting in their own right, these relationships with the continuum theory are not directly helpful in helping us understand H1-H3 in their own right as discrete-native theories.

Before discussing the dynamics of H4-H7, another word of warning. As mentioned above, due to Euler’s identity, \( \exp(2\pi i) = 1 \), the discrete planewave solutions with wavenumbers \( k \) and \( k + 2\pi \) are identical. Thus, it is best not to think of the x-axis of Fig. 2 abruptly ending at \( k = \pm \pi \) but rather as wrapping around on itself.

The planewave solutions to H4-H7 are,

\[
\phi_{n,m}(t; k_1, k_2) = e^{-ik_1n - ik_2m} e^{-\Gamma(k_1, k_2)t}
\]
with \(k_1, k_2 \in [-\pi, \pi]\). As before, extending \(k_1\) and \(k_2\) outside of this range does not yield new planewave solutions due to Euler’s identity, \(\exp(2\pi i) = 1\). The wavenumber-dependent decay rate \(\Gamma(k_1, k_2)\) for each theory is given by:

\[
\begin{align*}
\text{H4: } \Gamma(k_1, k_2) &= \alpha \left(2 - \cos(k_1) - \cos(k_2)\right) \\
\text{H5: } \Gamma(k_1, k_2) &= \frac{2\alpha}{3} \left(3 - \cos(k_1) - \cos(k_2) - \cos(k_2 - k_1)\right) \\
\text{H6: } \Gamma(k_1, k_2) &= \frac{\alpha}{2} \left(k_1^2 + k_2^2\right) \\
\text{H7: } \Gamma(k_1, k_2) &= \frac{\alpha}{3} \left(k_1^2 + k_2^2 + (k_2 - k_1)^2\right).
\end{align*}
\]

Note that \(\Gamma(k_1, k_2)\) for H6 and H7 follow from Eq. (27), essentially from the definition of \(D\). Unlike H1-H3, these theories do not all agree with each other in the small \(k_1\) and \(k_2\) regime. H4 and H6 agree that for \(k_1, k_2 \ll \pi\) we have \(\Gamma(k_1, k_2) \approx \frac{\alpha}{2}(k_1^2 + k_2^2)\). Moreover, H5 and H7 agree with each other in this regime, but not with H4 and H6. Do we have two different results in the continuum limit here?

Closer examination reveals that we do not. The key to realizing this is to note that under the transformation,

\[
\begin{align*}
k_1 &\mapsto k_1 \\
k_2 &\mapsto \frac{1}{2} k_1 + \frac{\sqrt{3}}{2} k_2,
\end{align*}
\]

we have \(\Gamma(k_1, k_2)\) for H7 mapping onto \(\Gamma(k_1, k_2)\) for H6. Applying this transformation to H5 does not map H5 onto H4, but it does bring their \(k_1, k_2 \ll \pi\) behavior into agreement.

Thus, if we consider only solutions with all or most of their wavenumber support with \(k_1, k_2 \ll \pi\) (or the appropriately transformed regime for H5 and H7) we have an approximate one-to-one correspondence between the solutions to these theories. Within this regime we can define their common continuum limit. Repeating our analysis of H1-H3 here, we expect H6 and H7 to converge in the continuum limit faster than H4 and H5.

This paper will make three attempts at interpreting these discrete heat equations. A first important point of comparison between these interpretations is what sense they make of these different convergence rates in the continuum limit.

Note also that for H6 and H7 the above discussed one-to-one correspondence is not approximate, nor is it restricted to the small \(k_1\) and \(k_2\) regime. We have an exact unrestricted one-to-one correspondence between the solutions of H6 and H7. Given any solution to H7 we can decompose it into planewaves, map these onto new planewaves using an invertible transformation (i.e., Eq. (37)) and then add them up to get a corresponding solution to H6. Similarly vice versa from H6 to H7.

A second important point of comparison between these three interpretations put forward in this paper is what sense they make of this one-to-one correspondence between H6 and H7. Such a correspondence does not automatically mean that these theories are identical or even equivalent.

Next consider the fact that H6 is manifestly rotation invariant in the \(k_1, k_2\) plane whereas H4, H5 and H7 are not. However, given the one-to-one correspondence between the solutions of H6 and H7, there is some (skewed) sense in which H7 is rotation invariant as well. A third important point of comparison between the coming interpretations is what sense they make of H6 and H7 being rotation invariant (at least in Fourier space).

Having introduced these theories and solved their dynamics in an interpretation-neutral way. We can now make a first (ultimately misled) attempt at interpreting them.

IV. A FIRST ATTEMPT AT INTERPRETING SOME DISCRETE SPACETIME THEORIES

Now that we have introduced these seven discrete theories and solved their dynamics, let’s get on to interpreting them. Let us begin by following our first intuitions and analyze these seven theories concerning their underlying manifold, locality properties, and symmetries. Ultimately however, as I will discuss later, much of the following is misled and will need to be revisited and revised later. Luckily, retracing where we went wrong here will be instructive later.

Let’s start by taking the initial formulation of the above theories seriously, i.e. Eq. (11), Eq. (12) and Eq. (13). Taken literally what are these theories about? These theories are intuitively about a field \(\phi(t)\) which maps lattice sites \((\ell \in L \cong \mathbb{Z}\cong \mathbb{Z}^2)\) and times \((t \in \mathbb{R})\) into temperatures \((\phi(t) \in \mathbb{R})\). That is a field \(\phi: \mathbb{Q} \to \mathbb{V}\) with manifold \(\mathbb{Q} \cong \mathbb{L} \times \mathbb{R}\) and value space \(\mathbb{V} = \mathbb{R}\). Taking \(\phi: \mathbb{Q} \to \mathbb{V}\) seriously as a fundamental field leads us to thinking of \(\mathbb{Q} = \mathbb{L} \times \mathbb{R}\) as the theory’s underlying manifold. It is important to note that in this interpretation, \(\mathbb{Q}\) is the entire manifold, it is not being thought of as embedded in some larger manifold. (However, a view like this will be considered in our third interpretation in Sec. [VII].)

Let’s see what consequences these interpretive moves have for locality and symmetry.

A. Intuitive Locality

Firstly, let’s develop a sense of comparative locality for H1, H2, and H3 taking \(\mathbb{Q}\) to be the underlying manifold. In a highly intuitive sense, theory H1 is the most local in that it couples together the fewest lattice sites: the instantaneous rate of change of \(\phi_n(t)\) only depends on itself, \(\phi_{n-1}(t)\), and \(\phi_{n+1}(t)\). It is this sense of locality which justifies us calling these sites its “nearest neighbors”. After this, H2 is the next most local in the same sense: it couples next-to-nearest neighbors. Finally, in this sense H3 is the least local, it has an infinite range
coupling: the instantaneous rate of change of \( \phi_n(t) \) depends on the current value at every lattice site. Thus at least on this intuitive notion of locality, \( H_1 > H_2 > H_3 \) with higher rated theories being more local. Similarly, assessing \( H_4-H_7 \) on this intuitive notion of locality gives the ratings, \( H_4, H_5 > H_6, H_7 \).

There is some tension however with these intuitive locality ratings and the rate we expect each theory to converge at in the continuum limit. For \( H_1-H_3 \) our intuitive locality ratings are \( H_1 > H_2 > H_3 \) but we expect convergence speeds in the continuum limit to be \( H_3 > H_2 > H_1 \). Similar tension exists for \( H_4-H_7 \). How is it that our most non-local theory is somehow the nearest to our perfectly local continuum theory?

In one sense there is no mystery here, when we make our derivative approximation longer range (more non-local) they can clearly get more accurate. But the question remains how exactly does an increasingly non-local operation give us an increasingly good approximation of a perfectly local operation (differentiation)? This tension will be dissolved and resolved in our second and third interpretations respectively. In particular, as I will discuss, these later interpretations negate or reverse all of the above intuitive locality judgements.

### B. Intuitive Symmetries

What discrete spacetime manifold can we intuitively read off of \( H_1-H_7 \)? As discussed in Sec. [II.C] intuitively the manifold underlying each of these theories is \( Q := L \times \mathbb{R} \). As discussed there, taking \( Q \) to be the theory’s underlying manifold limits our theories’ possible symmetries. Accounting for both external and internal symmetries (gauge symmetries are not relevant here) we found the widest scope of symmetry transformations possible were Eq. [13]. Namely, we found the possibilities are permutations of the lattice sites, time reparametrizations, and linear-affine rescalings. It is convenient to translate this in terms of the vector, \( \phi(t) \), as

\[
s : \quad \phi(t) \mapsto c_1 P \phi(\tau(t)) + c_2 \mathbf{1},
\]

for some permutation matrix, \( P \), and some monotone smooth function \( \tau(t) \) and for some \( c_1, c_2 \in \mathbb{R} \) where \( \mathbf{1} = (\ldots, 1, 1, 1, \ldots)^T \) is the constant vector. (Note that the permutation \( P \) cannot depend on time or else this transformation will be discontinuous).

It should be stressed that according to this interpretation this is the largest space of symmetries that \( H_1-H_7 \) could possibly have. Indeed, I have been charitable considering the lattice sites structured only as a set (perhaps artificially) increasing the size of \( \text{Auto}(Q) \). Given this, it would be highly surprising if we found \( H_1-H_7 \) to have symmetries outside of this set. (Indeed, such a surprise is coming in the next section.)

#### Symmetries of \( H_1-H_7 \): First Attempt

What then are the symmetries of \( H_1-H_7 \) according to this interpretation? The detailed technical of this evaluation are in Appendix [A] but ultimately for \( H_1-H_3 \) the symmetries are:

1) discrete shifts which map lattice site \( (n,m) \mapsto (n + d_1, m + d_3) \) for some integers \( d_1, d_3 \in \mathbb{Z} \)

2) two negation symmetries which map lattice site \( (n,m) \mapsto (-n,m) \) and \( (n,m) \mapsto (n,-m) \) respectively.

3) a 4-fold symmetry which maps lattice site \( (n,m) \mapsto (m,-n) \)

4) constant time shifts which map \( t \mapsto t + \tau \) for some real \( \tau \in \mathbb{R} \)

5) and linear-affine rescaling which maps \( \phi(t) \mapsto c_1 \phi(t) + c_2 \) for some \( c_1, c_2 \in \mathbb{R} \)

These are the symmetries of a uniform one-dimensional lattice \( x_n = n \in \mathbb{R} \) (plus time shifts and linear-affine rescaling). The above negation symmetry corresponds to mirror reflection. Previously I had warned against prematurely interpreting the lattice sites underlying \( H_1-H_3 \) as being organized into a uniform grid. Now, however, having investigated this theory’s dynamical symmetries we have some motivation to do so.

What about \( H_4 \) and \( H_6 \)? The technical details of this evaluation are in Appendix [A] but ultimately for \( H_4 \) and \( H_6 \) the symmetries are:

1) discrete shifts which map lattice site \( (n,m) \mapsto (n + d_1, m + d_3) \) for some integers \( d_1, d_3 \in \mathbb{Z} \)

2) two negation symmetries which map lattice site \( (n,m) \mapsto (-n,m) \) and \( (n,m) \mapsto (n,-m) \) respectively.

3) a 4-fold symmetry which maps lattice site \( (n,m) \mapsto (m,-n) \)

4) constant time shifts which map \( t \mapsto t + \tau \) for some real \( \tau \in \mathbb{R} \)

5) and linear-affine rescaling which maps \( \phi(t) \mapsto c_1 \phi(t) + c_2 \) for some \( c_1, c_2 \in \mathbb{R} \)

These are the symmetries of a square lattice \( \tau_{n,m} = (n,m) \in \mathbb{R}^2 \) (plus time shifts and linear-affine rescaling). The above 4-fold symmetry corresponds to quarter rotation. Previously I had warned against prematurely interpreting the lattice sites underlying \( H_4-H_7 \) as being organized into a square lattice. Now, however, having investigated these theories’ dynamical symmetries we have some motivation to do so at least for \( H_4 \) and \( H_6 \).

What about \( H_5 \) and \( H_7 \)? The technical details of this evaluation are in Appendix [A] but ultimately for \( H_5 \) and \( H_7 \) the symmetries are:
1) discrete shifts which map lattice site \((n, m) \mapsto (n + d_2, m + d_3)\) for some integers \(d_2, d_3 \in \mathbb{Z}\).

2) an exchange symmetry which maps lattice site \((n, m) \mapsto (m, n)\).

3) a 6-fold symmetry which maps lattice site \((n, m) \mapsto (-m, n + m)\). (Roughly, this permutes the three terms in Eq. 29 for \(H5\) and Eq. 31 for \(H7\).)

4) constant time shifts which map \(t \mapsto t + \tau\) for some real \(\tau \in \mathbb{R}\).

5) and linear-affine rescaling which maps \(\phi(t) \mapsto c_1\phi(t) + c_2\) for some \(c_1, c_2 \in \mathbb{R}\).

These are the symmetries of a hexagonal lattice \(L\), which we were able to find what sort of lattice structure \(L\) has for each theory (e.g. a uniform grid, square lattice and hexagonal lattice). Recall the discussion of matching dynamical symmetries with spacetime symmetries in Sec. II. For continuum spacetime theories, we can only discover their fixed spacetime structures by investigating the dynamics. (Of course, we have no hope of discovering them directly through dynamical means, they are dynamically-fixed.) The same is true here, we started with a bare lattice, \(L\), investigated the dynamics, and now we have good candidates for what lattice structures each of these theories have in addition to \(L\).

Finally, in this interpretation what sense can be made of \(H6\) and \(H7\) having a one-to-one correspondence between their solutions discussed after Eq. 37? While this correspondence between solutions certainly exists, little sense can be made of it here. As the above discussion has revealed this interpretation associates very different lattice structures to \(H6\) and \(H7\) and correspondingly very different lattice symmetries. While there is nothing wrong per se with this assessment our later interpretations will make better sense of this correspondence.

To summarize, this interpretation has the benefit of being highly intuitive. Taking the fields given to us, \(\phi : \mathcal{Q} \rightarrow \mathbb{R}\), seriously we identified the underlying manifold as \(\mathcal{Q}\). From this we got some intuitive notions of locality. Moreover, by finding these theories’ dynamical symmetries we were able to grant some more structure to their lattice sites (e.g. a uniform grid, square lattice and hexagonal lattice).

However, there are three issues with this interpretation which will become clear in light of our later interpretations. Firstly, our locality assessments are in tension with the rates at which these theories converge to the (perfectly local) continuum theory in the continuum limit. Secondly, despite the one-to-one correspondence between the solutions to \(H6\) and \(H7\), this interpretation regards them as significantly different theories: with different lattice structures and (here consequently) different symmetries. The final issue (which will become clear in the next section) is that this interpretation drastically under predicts the kinds of symmetries which \(H1-H7\) can and do have. In fact, each of \(H1-H7\) have a hidden continuous translation symmetry. Moreover, \(H6\) and \(H7\) have a hidden continuous rotation symmetry.

As I will discuss, the root of all of these issues is taking the theory’s lattice structure to be a piece of fixed background structure and moreover taking the lattice itself to be a fundamental part of the underlying manifold. Our second attempt at interpreting these theories will fix these issues.

V. A SECOND ATTEMPT AT INTERPRETING SOME DISCRETE SPACETIME THEORIES

In the previous section, I claimed that \(H1-H7\) have hidden continuous translation and rotation symmetries. But how can this be? How can discrete spacetime theories have such continuous symmetries? As discussed above, if we take our underlying manifold to be \(Q\) then our symmetries are limited to \(\text{Auto}(Q)\) which clearly cannot support continuous translation and rotation symmetries.

In order to avoid this conclusion we must deny the premise, \(Q\) must not be the underlying manifold. What led us to believe \(Q\) was the underlying manifold? We arrived at this conclusion by taking the real scalar field \(\phi : Q \rightarrow \mathcal{V}\) to be fundamental. \(Q\) is the underlying manifold because it is where our fundamental field maps from. In order to avoid this conclusion we must deny the premise, the field \(\phi : Q \rightarrow \mathcal{V}\) must not be fundamental.

But if \(\phi : Q \rightarrow \mathcal{V}\) is not fundamental then what is? Fortunately, our above discussion has already provided us with another field which we might take as fundamental. Namely, \(\phi(t)\) defined in Eq. 18.

On this second interpretation I will be taking the formulations of \(H1-H7\) in terms of \(\phi(t)\) seriously: namely Eq. (19), Eq. (23), Eq. (24), and Eqs. (28)-(31). Taken literally what are these theories about (in this formulation)? These theories are intuitively about a field \(\phi : \mathbb{R} \rightarrow \mathbb{R}^L\) which maps times \((t \in \mathbb{R})\) into infinite dimensional vectors \((\phi \in \mathbb{R}^L)\). That is a field \(\phi : \mathcal{M} \rightarrow \mathcal{V}\) with manifold \(\mathcal{M} = \mathbb{R}\) and value space \(\mathcal{V} = \mathbb{R}^L\). Taking \(\phi : \mathcal{M} \rightarrow \mathcal{V}\) seriously as a fundamental field leads us to thinking of \(\mathcal{M} = \mathbb{R}\) as the theory’s underlying manifold.

Notice that in this interpretation the lattice sites, \(L\), are no longer a part of our manifold. They have been “internalized” into the value space. Indeed, in this interpre-
tation H1-H7 are classical continuum spacetime theories (albeit with an abnormally large value space).

Let’s see what consequences these interpretive moves have for locality and symmetry. To preview: this second interpretation either dissolves or resolves all of our issues with the first interpretation. To preview: the tension between locality and convergence in the continuum limit is dissolved. H6 and H7 are seen to be equivalent in a stronger sense. And, perhaps most importantly, this interpretation reveals H1-H7’s hidden continuous translation and rotation symmetries. However, as I will discuss, this interpretation has some issues of its own which will ultimately require us to make a third pass over these theories.

A. Internalized Locality

Before discussing the effect this move has on the theories’ possible symmetries, let’s think briefly about what it does to our sense of locality. I claimed above that this interpretation dissolves the tension between convergence in the continuum limit and the sense locality developed in Sec. IV A. It does this by dissolving the possibility of any notion of locality stemming from the lattice sites.

In this interpretation the lattice sites have been internalized, they are no longer part of the manifold and therefore we no longer have a right to extract intuitions about locality from them. In this interpretation, the manifold consists only of times, \( t \in \mathbb{R} \). Consequently our only notion of locality is locality in time. The dynamics of each of H1-H7 are local in time and are therefore local in every possible sense. There is no longer any tension concerning how the differences in locality line up with the differences in continuum convergence rate; there simply are no differences in locality anymore.

If this seems unsatisfying to you I agree. One might feel that internalization’s ban on extracting notions of locality from the lattice sites is far too extreme. Intuitively, more strongly coupled lattice sites ought to be in some sense closer together. Moreover, one might rightly hope for an interpretation which not only dissolves the tension between a theory’s locality and its convergence in the continuum limit, but resolves it by bringing them into harmony. Indeed, if we have no notion of locality between lattice sites it is difficult to see how we get a notion of locality in the continuum limit.

These are all valid complaints which will be addressed in Sec. VII as I make a third attempt at interpreting H1-H7.

B. Internalized Symmetries

But how does this internalization move affect a theory’s capacity for symmetry? At first glance, this may appear to have made things worse. If our manifold is now just times \( t \in \mathbb{R} \) then our only possible dynamical symmetries are time-reparametrizations (i.e., not continuous translations and rotations). However, while there are certainly less possible dynamical/spacetime symmetries, we are now open to a wider range of internal symmetries. It is among these internal symmetries that we will find H1-H7’s hidden continuous translation and rotation transformations. As I will argue these symmetries can reasonably be given these names despite being internal symmetries. (In the following section I will present a third attempt at interpreting these theories which “externalizes” these symmetries, making them genuinely spacial translations and rotations.)

With our focus now on \( \phi : M \to \mathcal{V} \), let us consider its possibilities for symmetries. As discussed above, associated with the manifold we have only time reparametrizations, \( \text{Diff}(M) = \text{Diff}(\mathbb{R}) \). However, associated with the value space (i.e., an infinite dimensional vector space) we now have the full range of invertible linear-affine transformations over \( \mathbb{R}^L \), namely \( \text{Aff}(\mathbb{R}^L) \). (In principle these could be tied together into gauge transformations, but this is not relevant here.)

Thus, taken together the possibly symmetries for our theories are \( r \in \text{Diff}(\mathbb{R}) \times \text{Aff}(\mathbb{R}^L) \) which act on \( \phi(t) \) as,

\[
\begin{align*}
    r : \quad \phi(t) &\mapsto \Lambda \phi(\tau(t)) + c
\end{align*}
\]

for some monotone smooth function \( \tau(t) \), any invertible linear transformation \( \Lambda \), and some vector \( c \in \mathbb{R}^L \). Contrast this with the symmetries available to us on our first interpretation, namely Eq. (38). The present interpretation has a significantly wider class of symmetries than before.

Symmetries of H1-H7: Second Attempt

Which of the above transformations are symmetries for H1-H7? The technical details of this evaluation are in Appendix [A] but the results are the following. For H1-H3 the symmetries in this interpretation are:

1) Action by \( T^\epsilon \) sending \( \phi(t) \mapsto T^\epsilon \phi(t) \) where \( T^\epsilon \) is defined below.

2) negation symmetry which maps lattice site \( n \mapsto -n \).

3) constant time shifts which map \( t \mapsto t + \tau \) for some real \( \tau \in \mathbb{R} \)

4) and linear-affine rescaling which maps \( \phi_\ell(t) \mapsto c_1 \phi_\ell(t) + c_2 \) for some \( c_1, c_2 \in \mathbb{R} \)

These are exactly the same symmetries that we found on the previous interpretation with one difference: discrete shifts have been replaced with action by

\[
T^\epsilon := \exp(\epsilon D). \quad (40)
\]
with $\epsilon \in \mathbb{R}$. As I will now discuss, $T^c$ can be thought of as a continuous translation operator despite it being here classified as an internal symmetry.

First note that $T^c$ is a generalization of discrete shift operation in the sense that taking $\epsilon = d_1 \in \mathbb{Z}$ reduces action by $T^c$ to the map $n \mapsto n + d_1$. Moreover, note that $T^c$ is additive in the sense that $T^c_1 T^c_2 = T^c_{1+2}$. In particular, this means $T^{1/2} T^{1/2} = T^1$: there is something we can do twice to move one space forward. The same is true for all fractions adding to one. Finally, recall from the discussion following Eq. (27) that $h(x+\epsilon) = \exp(\epsilon \partial_x)h(x)$. In this sense also $T^c = \exp(\epsilon D)$ is a translation operator. More will be said about $T^c$ in Sec. VI.

Thus we have our first big lesson: discrete spacetime theories can have continuous translation symmetries. The fact that our discrete theories at first appeared on some lattice with some lattice structure does nothing to forbid this.

Next let’s consider H4-H7. Previously the symmetries of H4 and H6 matched and the symmetries of H5 and H7 also matched. Here however, these pairings are broken up and a new matching pair is formed between H6 and H7. More will be said about this momentarily.

Let’s consider H4 first. For H4 the symmetries in this interpretation are:

1) Action by $T_n^c$ sending $\phi(t) \mapsto T_n^c \phi(t)$ where by convention $T_n^c = T^c \otimes 1$. Similarly for $T_m^c = 1 \otimes T^c$.

2) two negation symmetries which map lattice site $(n, m) \mapsto (-n, m)$ and $(n, m) \mapsto (n, -m)$ respectively.

3) a 4-fold symmetry which maps lattice site $(n, m) \mapsto (m, -n)$

4) constant time shifts which map $t \mapsto t + \tau$ for some real $\tau \in \mathbb{R}$

5) and linear-affine rescaling which maps $\phi(t) \mapsto c_1 \phi(t) + c_2$ for some $c_1, c_2 \in \mathbb{R}$

These are the symmetries which we found on our first interpretation but with action by $T_n^c$ and $T_m^c$ replacing the discrete shifts. The same discussion following Eq. (40) applies here, justifying us calling $T_n^c$ and $T_m^c$ continuous translation operations.

Before moving on to analyze the symmetries of H6 and H7, let’s first see what this interpretation has to say about the one-to-one correspondence between them. As noted following Eq. (37), H6 and H7 have a one-to-one correspondence between their solutions. Given a solution to H7 I can decompose it into planewaves, map these onto new planewaves using the invertible transformation Eq. (37) and then add them up to get a corresponding solution to H6. Similarly vice versa from H7 to H6. In our first interpretation, despite this, H6 and H7 were judged to be different theories because they had different lattice structures and (there consequently) different symmetries.

Things are substantially different in this interpretation. The transformations between solutions of H6 and H7 described above is a linear transformation on $\phi$. Thus, in this interpretation the only difference between H6 and H7 is a change of basis for the vector space $\mathbb{R}^L$. Because this is a transformation of the form Eq. (39) (but notably not Eq. (38)) for any symmetry transformation for H6 there is a corresponding symmetry transformation for H7.

This is in strong contrast to the results of our previous analysis in Sec. IV. There H6 was seen to have the symmetries of a square lattice and H7 was seen to have the symmetries of a hexagonal lattice. By contrast, in the present interpretation H6 and H7 are thoroughly equivalent.

Thus we have our second big lesson: discrete theories which are presented to us with very different lattice structures (i.e., a square lattice versus a hexagonal lattice), may nonetheless turn out to be completely equivalent theories. This not only in terms of a one-to-one mapping between their solutions as we saw before, but in terms of their symmetries as well. In this interpretation, the process for switching between lattice structures is simply a change of basis in the value space.

In the rest of this subsection I will only discuss the symmetries H6, analogous conclusions are true for H7 after applying Eq. (37). The symmetries for H6 in this interpretation are:

1) Action by $T_n^c$ sending $\phi(t) \mapsto T_n^c \phi(t)$ where by convention $T_n^c = T^c \otimes 1$. Similarly for $T_m^c = 1 \otimes T^c$.

2) an exchange symmetry which maps lattice site $(n, m) \mapsto (m, n)$.

3) a 6-fold symmetry which maps lattice site $(n, m) \mapsto (-n, m + n)$. (Roughly, this permutes the three terms in Eq. (29) for H5 and Eq. (31) for H7.)

4) constant time shifts which map $t \mapsto t + \tau$ for some real $\tau \in \mathbb{R}$.

5) and linear-affine rescaling which maps $\phi(t) \mapsto c_1 \phi(t) + c_2$ for some $c_1, c_2 \in \mathbb{R}$.

These are the symmetries which we found on our first interpretation but with action by $T_n^c$ and $T_m^c$ replacing the discrete shifts. The same discussion following Eq. (40) applies here, justifying us calling $T_n^c$ and $T_m^c$ continuous translation operations.

In the rest of this subsection I will only discuss the symmetries H6, analogous conclusions are true for H7 after applying Eq. (37). The symmetries for H6 in this interpretation are:

1) Action by $T_n^c$ sending $\phi(t) \mapsto T_n^c \phi(t)$ where by convention $T_n^c = T^c \otimes 1$. Similarly for $T_m^c = 1 \otimes T^c$.

2) an exchange symmetry which maps lattice site $(n, m) \mapsto (m, n)$. 

3) Action by $R^\theta$ sending $\phi(t) \mapsto R^\theta \phi(t)$ with $R^\theta$ defined below.

4) constant time shifts which map $t \mapsto t + \tau$ for some real $\tau \in \mathbb{R}$

5) and linear-affine rescaling which maps $\phi(t) \mapsto c_1 \phi(t) + c_2$ for some $c_1, c_2 \in \mathbb{R}$.

As with H4, we have here action by $T_n^a$ and $T_n^m$ replacing the discrete shifts from before. However, additionally we have the quarter rotations replaced with action by $R^\theta$ where

$$R^\theta = \exp(\theta(N D_m - M D_n))$$  \hspace{1cm} (41)

with $\theta \in \mathbb{R}$ and where $N$ and $M$ are position operators which return the first and second index.

As I will now discuss, $R^\theta$ can be thought of as a continuous rotation operator despite it being here an internal symmetry. First note that $R^\theta$ is a generalization of quarter rotation operation in the sense that taking $\theta = \pi/2$ reduces action by $R^\theta$ to the map $(n, m) \mapsto (m, -n)$. Moreover, note that $R^\theta$ is cyclically additive in the sense that $R^{\theta_1} R^{\theta_2} = R^{\theta_1 + \theta_2}$ with $R^{2\pi} = \mathbb{I}$. In particular, this means $R^{\pi/4} R^{\pi/4} = R^{\pi/2}$. There is something we can do twice to make a quarter rotation. Similarly for all fractional rotations. Finally, recall from the discussion following Eq. (27) that $D$ is closely related to the continuum derivative operator, exactly matching its spectrum for $k \in [-\pi, \pi]$. Recall also that rotations are generated through the derivative as $h(R(x, y)) = \exp(\theta(x \partial_x - y \partial_y))h(x, y)$. In this sense also $R^\theta$ is a rotation operator. More will be said about $R^\theta$ in Sec. VI.

This adds to our first big lesson: discrete spacetime theories can have not only continuous translation but also continuous rotation symmetries. The fact that our discrete theories at first appeared with lattice structure does nothing to forbid this.

To summarize: this second attempt at interpreting H1-H7 has fixed all of the issues with our previous interpretation. Firstly, there is no longer any tension between these theories differing locality properties and the rates at which they converge to the (perfectly local) continuum theory in the continuum limit. (There are no longer any differences in locality.) Secondly, the fact that we have a one-to-one correspondence between the solutions to H6 and H7 is properly reflected in their matching symmetries. Finally, this interpretation has exposed the fact that H1-H7 have a hidden continuous translation and rotation symmetries.

These are all substantial improvements, but ultimately this interpretation has its own issues. The way that the tension is dissolved between locality and convergence in the continuum limit is unsatisfying. We ought to be able to extract intuitions about locality from the lattice sites. Moreover, while this interpretation has indeed exposed H1-H7’s hidden continuous translation and rotation symmetries, the way it classifies them seems wrong. They are here classified as internal symmetries (i.e., symmetries on the value space) whereas intuitively they should be external symmetries (i.e., symmetries on the manifold).

The root of both of these issues is taking the theory’s lattice structure to be internalized into the theory’s value space. Our third attempt at interpreting these theories will fix this. Before making this third attempt, let me complete the current interpretation by applying our continuum notion of general covariance to H1-H7.

C. Internalized General Covariance

Rewriting H1-H7 in the coordinate-free language of differential geometry we have:

H1-H7: $\langle \mathcal{M}, t_{ab}, T^a, \Delta^2_{H1-H7}, \phi \rangle$  \hspace{1cm} (42)

DPMs: $T^a \nabla_x \phi = \alpha \Delta^2_{H1-H7} \phi$

where $\mathcal{M}$ is a 0 + 1 dimensional manifold, $t_{ab}$ is a fixed metric field with signature (1), $\nabla_x$ is the unique covariant derivative operator compatible with $t_{ab}$, and $T^a$ is a fixed constant time-like unit vector field, $\nabla_a T^b = 0$ and $t_{ab} T^a T^b = 1$. Here $\phi : \mathcal{M} \to V$ is a dynamical infinite-dimensional vector field and $\Delta^2_{H1-H7}$ is whichever operator appears in the relevant dynamical equation: Eq. (19), Eq. (23), Eq. (24) and Eqs. (28)-(31). Notice that on this second interpretation, the lattice has disappeared from the manifold, having been fully-internalized.

Thus we have our third big lesson: given a discrete spacetime theory with some lattice structure we can always reformulate it in such a way that it has no lattice structure whatsoever. In fact, there is no longer even a lattice here. In this interpretation, this is done by internalizing the lattice structure.

Before making a third attempt at interpreting H1-H7, I need to introduce some mathematical tools regarding bandlimited functions and Nyquist-Shannon sampling theory.

VI. BRIEF REVIEW OF BANDBLIMITED FUNCTIONS AND NYQUIST-SHANNON SAMPLING THEORY

This section will provide a thorough but informal mathematical overview of the primary tools used in the third attempt to interpret H1-H7, namely bandlimited functions and Nyquist-Shannon sampling theory. My intention is not to prove these theorems of sampling theory in any technical sense but rather to make them intuitive. For a selection of introductory texts on sampling theory see [40-42].

To introduce the topic I will restrict our attention to the one-dimensional case with uniform sample lattice before generalizing to higher dimensions and non-uniform samplings later on.
A. One Dimension Uniform Sample Lattices

A bandlimited function is one whose Fourier transform has compact support. Consider a generic bandlimited function, $f_B(x)$, with a bandwidth of $K$. That is, a function $f_B(x)$ such that its Fourier transform,

$$\mathcal{F}[f_B(x)](k) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f_B(x) e^{-ikx} \, dx,$$  \hspace{1cm} (43)

has support only for wavenumbers $|k| < K$.

Suppose that we know the value of $f_B(x)$ only at the regularly spaced sample points, $x_n = na + b$, with some spacing, $0 \leq a \leq a^* := \pi/K$, and offset, $b \in \mathbb{R}$. Let $f_n = f_B(x_n)$ be these sample values. Having only the discrete sample data, $\{(x_n, f_n)\}_{n \in \mathbb{Z}}$, how well can we approximate the function?

The Nyquist-Shannon sampling theorem \[43\] tells us that from this data we can reconstruct $f_B$ exactly everywhere! That is, from this discrete data, $\{(x_n, f_n)\}_{n \in \mathbb{Z}}$, we can determine everything about the function $f_B$ everywhere. In particular, the following reconstruction is exact,

$$f_B(x) = \sum_{n=-\infty}^{\infty} S\left(\frac{x-x_n}{a}\right) f_n$$ \hspace{1cm} (44)

$$= \sum_{n=-\infty}^{\infty} S_n\left(\frac{x-b}{a}\right) f_n,$$

where

$$S(y) = \frac{\sin(\pi y)}{\pi y}, \quad \text{and} \quad S_n(y) = S(y-n),$$ \hspace{1cm} (45)

are the normalized and shifted sinc functions. Note that $S_n(m) = \delta_{nm}$ for integers $n$ and $m$. Moreover, note that each $S_n(x)$ is both $L_1$ and $L_2$ normalized and that taken together the set $\{S_n(x)\}_{n \in \mathbb{Z}}$ forms an orthonormal basis with respect to the $L_2$ inner product. The fact that any bandlimited function can be reconstructed in this way is equivalent to the fact that this orthonormal basis spans the space of bandlimited functions with bandwidth of $K = \pi$.

As a concrete example, let us consider the function $f_B(x) = 1 + S(x-1/2) + x S(x)/2$, shown in Fig.\[3\]. This function has a bandwidth of $K = \pi$ and so has a critical sample spacing of $a^* = \pi/K = 1$. Thus, we can fully reconstruct $f_B(x)$ knowing only its values at $x_n = na + b$ for any spacing $a \leq a^* = 1$. In particular the sample values at $x_n = n/2$ are sufficient to exactly reconstruct the function, see Fig.\[3\]. So too are the sample values at the integers $x_n = n$ and at $x_n = n + 1/3$, see Fig.\[3\] and Fig.\[3\]. In each of these cases the reconstruction is given by Eq.\[44\].

Everything about this function can be reconstructed from any uniform sample lattice with $a \leq a^* = 1$. In particular, the value of $f_B$ at a third of the way between sample point, $f_B(2/3)$, is fixed by $\{(n, f_B(n))\}_{n \in \mathbb{Z}}$ even though we have no sample at or even near $x = 2/3$. The derivative of $f_B$ at zero, $f'_B(0)$, is fixed by $\{(n, f_B(n))\}_{n \in \mathbb{Z}}$ even though the only sample point we have in this neighborhood is $f_B(0)$. Moreover, the derivative at $x = 2/3$, namely $f'_B(2/3)$, is fixed by $\{(n, f_B(n))\}_{n \in \mathbb{Z}}$ even we have no sample points in the neighborhood.

On first exposure this may be shocking: how can a function’s behavior everywhere be fixed by its value at a discrete set of points? When $f_B$ is represented discretely, where has all of the information gone? Where is the information about the derivative at $x = 1/3$ stored in the discrete representation? One may feel that any such discretely-determined function must belong to a very restricted class.
Classes of functions fixed by their values at discrete points are not uncommon in mathematics. For instance, polynomials of finite degree are fixed knowing their values at only finitely many places. More surprisingly, the Identity Theorem of complex analysis tells us that any entire function, \( g_E(x) \), is fixed by the values \{\( g_E(1), g_E(1/2), g_E(1/3), \ldots \)\}. Recall that entire functions are those functions whose Taylor series based at any point converges everywhere. The class of entire functions include all polynomials, all trig-functions, all hyperbolic trig-function, and a great many elementary combinations of these.

Before discussing bandlimited functions further, it will be instructive to take a brief detour into the land of entire functions.

An Entire Detour

A generic entire function, \( g_E(x) \), is fully determined by knowing all of its derivatives at any given location. For instance, any \( g_E(x) \) can be equivalently represented by the infinite-dimensional vector, \( \mathbf{g} \), which collects together all of \( g_E \)'s derivative at \( x = 0 \). Namely, the \( r \)th entry of \( \mathbf{g} \) is \( g_r = \partial^r g_E(0) \) for integers \( r \geq 0 \).

Every property of \( g_E(x) \) can be represented in terms of this vector. By construction, the derivatives of \( g_E \) at \( x = 0 \) are “stored” in this representation trivially as,

\[
\partial^r g_E(0) = g_r = v^r \mathbf{g},
\]

where \( v_r = (0, \ldots, 0, 1, 0, \ldots) \) with the 1 in the \( r \)th place (counting from zero).

The values of \( g_E \) away from \( x = 0 \) can be recovered via \( g_E \)'s Taylor series.

\[
g_E(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} g_n.
\]

The derivatives of \( g_E \) away from \( x = 0 \) can be recovered by taking derivatives of the above formula. Ultimately one finds,

\[
\partial^r g_E(a) = \sum_{n=r}^{\infty} \frac{a^{n-r}}{(n-r)!} g_n = v^r T_E^x \mathbf{g},
\]

where the entries of the matrix \( T_E^x \) are \([T_E^x]_{i,j} = a^{-j}/(j-i)! \) if \( j \geq i \) and 0 otherwise. This recovers the above formula when \( r = 0 \).

\[
T_E^x = \begin{pmatrix}
1 & a & a^2/2! & a^3/3! & a^4/4! & \ldots \\
0 & 1 & a & a^2/2! & a^3/3! & \ldots \\
0 & 0 & 1 & a & a^2/2! & \ldots \\
0 & 0 & 0 & 1 & a & \ldots \\
\cdots & \cdots & \cdots & \cdots & \cdots & \ldots 
\end{pmatrix}
\]

\( T_E^x \) acts as the translation operator for this representation of entire functions. Indeed, \( T_E^x \) is a representation of the translation group, satisfying \( T_E^x T_E^y = T_E^{x+y} \). If \( \mathbf{g} \) represents \( g_E(x) \) then \( T_E^x \mathbf{g} \) represents \( g_E(x+a) \).

Thus, we have seen how any derivative of \( g_E \) away from \( x = 0 \) can be recovered when we represent \( g_E \) with \( \mathbf{g} \). In this representation, the derivatives at \( x = 0 \) are stored “primitively” while everything else requires some unpacking. Of course, there is nothing special about this representation. We might alternatively choose to store the derivatives of \( g_E(x) \) at \( x = a \) primitively and recover everything else from there. More creatively, recalling the identity theorem, we could store the values of \( g_E(x) \) at \( x = 1/n \) for integer \( n \) and recover everything else from there. Each of the above representations are related to each other by a simple change of basis.

Of course, one can go wrong in designing a representation for \( g_E(x) \): knowing only the even derivatives at \( x = 0 \) won’t fix \( g_E \). However, we still have great freedom in how to represent \( \mathbf{g} \). Any derivative of \( g_E \) at any location can appear “primitively” in our designer representation. Moreover, any wish-list of derivatives of \( g_E \) can appear primitively in our designer representation. To see this, note that if this wish-list does not already give us a complete basis we can simply begin adding in derivatives of \( g_E \) at \( x = 0 \) until the space is spanned. Eventually this will complete the basis.

Consider the representation of \( g_E \) given by,

\[
\hat{\mathbf{g}} = (g_E(0), g_E(1), g_E'(0), g_E'(1), g_E''(0), g_E''(1), \ldots)^T
\]

This is obviously enough information to recover \( g_E \); it contains all of the derivatives of \( g_E \) at \( x = 0 \) and \( x = 1 \). However, it is an interesting question exactly how much of this information can be deleted without losing our ability to recover \( g_E \) exactly. I have no answer to this question, but an analogous question will soon arise for bandlimited functions.

Let’s now return to a discussion of bandlimited functions.

Return to One Dimensional Uniform Sample Lattices

Much of the above discussion about entire functions carries over unchanged to bandlimited functions.

To begin, let’s see how the values and derivatives of \( f_B \) everywhere are encoded into its values at a sufficiently dense uniform sample lattice. The values of \( f_B \) at the sample points \( x = x_n \) are stored trivially as,

\[
f_B(x_n) = f_n = w_n^T \mathbf{f},
\]

where \( w_n = (\ldots, 0, 1, 0, \ldots)^T \) with the 1 in the \( n \)th position. Note that \( w \) is infinite in both directions.

The values of \( f_B \) away from the sample points (at \( x = x_n + \epsilon a \) for \( \epsilon \in \mathbb{R} \)) can be reconstructed as,

\[
f_B(x_n + \epsilon a) = \sum_{m=-\infty}^{\infty} S_m(n + \epsilon) \ f_m = w_n^T T_B^\epsilon \mathbf{f}
\]
where the entries of the matrix $T_B^ε$ are $[T_B^ε]_{i,j} = S_i(j + ε)$. Note $T_B^ε$ acts as the translation operator for this representation of bandlimited functions. Indeed, $T_B^ε$ is a representation of the translation group, satisfying $T_B^{α+β} = T_B^α T_B^β$. If $f$ represents $f_B(x)$ then $T_B^ε f$ represents $f_B(x + ε)$.

From this translation operator we can identify the derivative operator for bandlimited functions, $D_B$, as

$$D_B := \lim_{ε→0} T_B^ε - \mathbb{1}/ε.$$  \hfill (53)

It should be noted that $D_B$ and $T_B^ε$ commute and moreover we have the usual relationship between derivatives and translations, $T_B^ε = \exp(ε D_B)$.

From the above definition of $D_B$ one can easily work out its matrix entries as $[D_B]_{i,j} = (-1)^{|i-j|}/(i-j)$ when $i \neq j$ and 0 when $i = j$. Note that $D_B$ acts as the derivative operator for this representation of bandlimited functions. If $f$ represents $f_B(x)$ then $\frac{1}{a} D_B f$ represents $f_B'(x)$.

$$D_B = \text{Toeplitz}(\frac{1}{4}, \frac{1}{3}, \frac{2}{3}, 1, 0, 1, \frac{1}{2}, \frac{3}{4}, \frac{1}{4}, \ldots)$$ \hfill (54)

Comparing this with the $D$ operator introduced in Eq. (25) we see that they are numerically identical. Indeed, $D_B$ is diagonal in the Fourier basis with spectrum $\lambda_D_B(k) = ik$ for $k \in [-π, π]$. This is exactly the defining property of the $D$ operator introduced earlier, see Eq. (27).

Indeed, $D_B = D$ and moreover $T_B = T$. If we were to extend our discussion to two-dimensional functions we could find a discrete representation of the rotation operator for bandlimited functions. $R_B$. This would come out numerically equal to the $R$ operator introduced earlier in Eq. (41), namely $R_B = R$. Thus, the discrete notions of derivative, translation, and rotation that we have been using up until now are intimately connected with bandlimited functions.

It should be noted that $D = D_B$ gives us the following remarkable derivative approximation (which is exact for bandlimited functions):

$$\partial_x f(x) \approx 2 \sum_{m=1}^{∞} (-1)^{m+1} \frac{f(x + ma) - f(x - ma)}{2ma}.$$ \hfill (55)

Namely, when $h$ is bandlimited with bandwidth of $K$ and $a \leq π/K$ then this formula is exact. Moreover, if the Fourier transform of $h$ is mostly supported in $[-K, K]$ with thin tails (e.g., Gaussian tails) outside this region, then this is a very good derivative approximation.

Ultimately we can compute any derivative of $f_B$ anywhere from our sample data as

$$\partial_x^r f_B(x_n + ε a) = \frac{1}{a^r} w_n^T D_B^r T_B^ε f.$$ \hfill (56)

Thus, we can recover any value or derivative of $f_B$ from its values on any sufficiently dense uniform sample lattice. We can translate between any two uniform sample lattices with the same spacing by using the bandlimited translation operator, $T_B^ε$. Translating between uniform sample lattices with different spacings is more difficult, but it can be done (as long as both are sufficiently dense with spacings $a \leq π/K$). Ultimately, each of these re-descriptions can be accomplished by a simple change of basis.

### B. Non-Uniform Sample Lattices

The previous subsection showed how any value or derivative of $f_B$ can be recovered from its values on any sufficiently dense uniform sample lattice. Moreover, it showed how changing between representing $f_B$ with different uniform sample lattices is ultimately just a change of basis.

However, just as I discussed in the case of entire functions, we can be more creative with how we try to represent $f_B$ than this. In particular, we can begin designing a customized representation by picking out any wish-list of values or derivatives of $f_B$ to appear “primitively” in our representation. If this wish-list does not give us a complete spanning of the space of bandlimited function, then we can just begin adding in samples off of a uniform lattice until it does. What results is a non-uniform sample lattice (albeit one with a uniform sub-lattice).

Alternatively we might consider beginning from an overly dense sampling and trimming down from there. For example, figure Fig. 3b shows $f_B$ sampled at twice the necessary frequency. This is a representation of $f_B$ in an overcomplete basis. Imagine oversampling by a factor of ten with a spacing of $a = a^*/10$. Intuitively, this sample lattice has ten times the information needed to recover the function exactly. If we were to delete all but every tenth data point we would still be able to recover the function. But what if we just half of the sample points, but did so randomly? This would result in a non-uniform sample lattice. See for instance Fig. 3c. Hopefully, the reader has some intuition that at least some non-uniform sample lattices are sufficient to exactly reconstruct $f_B$.

The answers to such questions are given by the various non-uniform sampling theorems [40][41]. The details of these theorems are not important here; They can all be summarized as saying that reconstruction is possible when our non-uniform sample points are “sufficiently dense” in some technical sense. The sampling shown in Fig. 3e is sufficiently dense. The reconstruction in the non-uniform case is significantly more complicated than it is in the uniform case, recall Eq. (44). In the non-uniform case it is generally of the form

$$f_B(x) = \sum_{m=-∞}^{∞} G_m(x; \{z_n\}_{n\in\mathbb{Z}}) f_B(z_m)$$ \hfill (57)

for some reconstruction functions, $G_m$, which depend in a complicated way on the location of all of the other sample points, $\{z_n\}_{n\in\mathbb{Z}}$. 


C. Higher Dimensional Sampling

The same story about bandlimited functions is largely true in higher dimensions as well. A two-dimensional function \( f_B(x, y) \) is bandlimited if its Fourier transform \( \mathcal{F}[f_B(x, y)](k_x, k_y) \) is compactly supported in the \((k_x, k_y)-\)plane. Specifying the value of the bandwidth is less straightforward in the high dimensional case as the Fourier transform’s support may have different extents in different directions. However, any compact region can be bounded in a square. We can thus always imagine \( f_B(x, y) \) as being bandlimited with \((k_x, k_y) \in [-K, K] \times [-K, K] \) for some \( K > 0 \). As such, we can represent \( f_B(x, y) \) with a (sufficiently dense) uniform sample lattice in both the \( x \) and \( y \) directions. That is we can represent \( f_B(x, y) \) in terms of its sample values on a sufficiently dense square lattice.

Once we have such a uniform sampling, the reasoning carried out above applies unchanged. We can include any values or derivatives of \( f_B(x, y) \) as part of our representation, as long as this is part or supplemented by a sufficiently dense (in some technical sense) sample lattice.

For a concrete example consider the bandlimited function shown in Fig. 4a, namely,

\[
f_B(x, y) = J_1(\pi r)/(\pi r)
\]

where \( J_1 \) is the first Bessel function and \( r = \sqrt{x^2 + y^2} \).

This function is bandlimited with \( \sqrt{k_x^2 + k_y^2} < K = \pi \) and hence critical spacing \( a^* = \pi/K = 1 \). Moreover, this function is rotationally symmetric.

Given this function’s bandwidth of \( K = \pi \), we can represent it via its sample values taken on a square lattice with spacing \( a = 1/2 \leq a^* = 1 \), see Fig. 4a. We can also use a coarser square lattice with a spacing of \( a = a^* = 1 \), see Fig. 4b. We could also use a rotated square lattice, see Fig. 4c. Sampling the function on a hexagonal lattice also works, see Fig. 4d. Finally we can use a non-uniform lattice of sample points, see Fig. 4e.

From each of these discrete representations, we could recover the original bandlimited function everywhere exactly via some generalization of Eq. (44).

Thus, there is no conceptual barrier to representing a rotationally invariant bandlimited function on a square or hexagonal lattice. Indeed, there is no issue with representing such a function on any sufficiently dense lattice. In light of the analogy proposed in this paper, we can see this as analogous to the unsurprising fact that there is no conceptual barrier to representing rotationally invariant functions in Cartesian coordinates. There is no requirement that our representation (be it a choice of coordinates or a choice of sample points) latches onto the symmetries of what is being represented.

Thus we have a non-uniform sampling theory for higher dimensions. But what about a sampling theory on curved spaces? While such things are not relevant for the aims of this paper, recently notable progress has been made on developing a sampling theory for curved manifolds [30, 44].

VII. A THIRD ATTEMPT AT INTERPRETING SOME DISCRETE SPACETIME THEORIES

In Sec. [IV] it was revealed that H1-H7 have hidden continuous symmetry transformations which intuitively correspond to spacial translations and rotation. In our first attempt at interpreting H1-H7 the possibility of such symmetries were outright denied, see Sec. [IV]. In our second attempt, these hidden symmetries were exposed, but they were classified (unintuitively) as internal symmetries, see Sec. [V]. This is due to an “internalization” move made in our second interpretation. This move also undercut our ability to use the lattice sites to reason about locality.

In this section (using the tools introduced in the previous section) I will show how we can externalize these symmetries by 1) inventing a continuous manifold for them to live on and 2) embedding our states/dynamics onto this manifold as bandlimited functions.

A perspective similar to this third interpretation has been put forward in the physics literature by Achim Kempf [24, 34] among others [35–37]. For an overview see [41].

A. Choice of Manifold, Embedding and Sample Points

If we are going to externalize these symmetries then we need to have a big enough manifold on which to do the job. Clearly neither the manifold in our first interpretation, \( Q \), or in our second, \( \mathbb{R} \), is up to the task. What manifold \( \mathcal{M} \) might be up to the task?

The first thing we must do is pick which of our theory’s symmetries we would like to externalize (there may be some symmetries we want to keep internal). For H1-H7 we want to externalize the following symmetries: continuous translations, continuous rotations, mirror reflections, and constant time shifts. In any case, we collect these symmetries together in a group \( G_{\text{ext}} \). Clearly, our choice of manifold \( \mathcal{M} \) needs to be big enough to have \( G_{\text{ext}} \) as a subgroup of \( \text{Diff}(\mathcal{M}) \). Of course this doesn’t uniquely specify the manifold we ought to use. If \( \mathcal{M} \) works, then so does any \( \mathcal{M}' \) with \( \mathcal{M} \) as a sub-manifold. For standard Occamistic reasons, it is natural to go with the smallest manifold which gets the job done. The larger the gap between \( G_{\text{ext}} \) and \( \text{Diff}(\mathcal{M}) \) the more fixed spacetime structures will need to be introduced later on. One might proceed by trying to formalize the “size” of this gap and prove something about it minimum. However, here I prefer to just get building.

Let’s begin by picking out all of our theory’s continuous translation symmetries (in either space or time). As many of these as there are will give us a lower bound.
FIG. 4. Several different (but completely equivalent) graphical representations of the bandlimited function given by Eq. (58). This function has a bandwidth of \( \sqrt{k_x^2 + k_y^2} < K = \pi \) and so has a critical spacing of \( a^* = \pi/K = 1 \) in every direction. The scale of each subfigure is 5x5. In each subfigure, the colored regions are the Voronoi cells around the sample points (black).

Subfigure a) shows the function values for all \( x \). b) shows \( f_B \) sampled on a square lattice with \( z_{n,m} = (n/2, m/2) \). Since \( 1/2 < a^* = 1 \) this is an instance of oversampling. c) shows \( f_B \) sampled on a square lattice with \( z_{n,m} = (n, m) \). This is an instance of critical sampling since \( a = a^* = 1 \). d) shows \( f_B \) sampled on a square lattice with \( z_{n,m} = (n + m, n - m)/\sqrt{2} \). e) \( f_B \) sampled on a hexagonal lattice of with \( z_{n,m} = (n + m/2, \sqrt{3}m/2) \in \mathbb{R}^2 \). f) shows \( f_B \) sampled on an irregular lattice.

on the number of dimensions our manifold requires. Another guide to the necessary number of dimensions is the dimensionality of the lattice structure revealed in our first interpretation Sec. IV: e.g., a uniform grid, a square lattice and a hexagonal lattice. Either of these indicators suggest that for H1-H3 we have \( M_{1-3} \cong \mathbb{R}^2 \) and for H4-H7 we have \( M_{4-7} \cong \mathbb{R}^3 \). When the subscript on \( M \) is not relevant it will be dropped.

Once we have a manifold selected, we need to somehow embed \( \phi(t) \) (or equivalently \( \phi(t) \)) into it. While other embeddings are possible, given the tools developed in Sec. VI and how well they appear to suit our purposes there are substantial reasons to go for a bandlimited embedding. That is, we are going to think of each \( \phi(t) \) as a sample value which is drawn from some bandlimited field \( \phi_B : \mathcal{M} \rightarrow \mathbb{R} \) at some sample point \( z(t) \in \mathcal{M} \). That is,

\[
H1-H7: \quad \phi(t) = \phi_B(z(t)) \tag{59}
\]

But what points should we take for \( z(t) \)? In principle we here have complete freedom\(^{14}\) in selecting these sample points. However, perhaps surprisingly, if we make

---

\(^{14}\) One may feel some tension here with the point stressed in Sec. VI that a choice of sampling lattice must always be sufficiently dense in some technical sense. This is true when we already have in mind a fixed bandlimited function and manifold. To describe this function we need a sufficiently dense sampling, depending on its bandwidth. However, here we have no such function and manifold in mind. We are building a manifold and then associating certain values with certain points on the manifold. From these we will then construct a bandlimited function. The bandwidth of the resulting function will be compatible with our choice of sample points. By construction these sample values capture all of the information about the function.
natural choices about how the symmetries we have already identified fit onto $\mathcal{M}$ then our way forward here is more-or-less fixed.

Let us consider H4-H7 and suppose we make the following choices lining up our symmetry transformations $G$ with certain diffeomorphisms on $\mathcal{M}$. Specifically, I take:

1. our continuous translation symmetry $T^x_\epsilon$ to act on the manifold as $(t, x, y) \mapsto (t + \epsilon a, x, y)$ for some lattice spacing $a > 0$,
2. our continuous translation symmetry $T^y_\epsilon$ to act on the manifold as $(t, x, y) \mapsto (t, x + \epsilon a, y)$,
3. our constant time shifts $t \mapsto t + \tau$ act on the manifold as $(t, x, y) \mapsto (t + \tau, x, y)$,
4. and finally $\phi_{0,0}(0)$ to be a sample of $\phi_B$ at $z_{0,0}(0) = (0, 0, 0)$.

Given these choices after fixing everything else is fixed with

\begin{align*}
\text{H1-H3: } z_n(t) &= (t, na) \\
\text{H4-H7: } z_{n,m}(t) &= (t, na, ma) 
\end{align*}

with similar logic for applying for H1-H3. Fig. 5 shows for H1-H3 these sample points (vertical black lines.) as they lie on the spacetime manifold $\mathcal{M}_{1-3} \cong \mathbb{R}^2$. One can imagine an analogous figure for H4-H7 with the sample points forming a square lattice extended through time.

One may worry that we are here taking a square lattice for each of H4-H7 whereas for H5 and H7 we naturally ought to embed on a hexagonal lattice. This point will resolve itself naturally later.

Having chosen a manifold, embedding, and sample points, I will next reconstruct the bandlimited function $\phi_B(t)$ from these sample points.

---

sample points. We are completely free in how to do this. Conversely, we could also get new sample values by sampling our old function at the new sample points. Unlike before, here our new sample points are restricted to be sufficiently dense according to the bandwidth of the old function.

### B. Bandlimited Dynamics

The previous subsection motivated us to think of the discrete variables from H1-H7 as samples of some bandlimited function, $\phi_B$, as in Eq. (59). Using these sample values we can use the tools discussed in Sec. VII to reconstruct $\phi_B$ exactly. In particular making use of Eq. (44) we have,

\begin{align*}
\text{H1-H3: } \phi_B(t, x) &= \sum_{n=-\infty}^{\infty} S_n(x/a) \phi_n(t) \\
\text{H4-H7: } \phi_B(t, x,y) &= \sum_{n,m=-\infty}^{\infty} S_n(x/a) S_m(x/a) \phi_{n,m}(t)
\end{align*}

Note that by construction $\phi_B(t,x)$ and $\phi_B(t,x,y)$ are both bandlimited with bandwidth of $K = \pi/a$ for each time $t$.

Fig. 5 shows for H1-H3 what this bandlimited function $\phi_B$ might look like. In particular, this figure shows one of the planewave solutions decaying exponentially in time. (Note that at a fixed wavenumber, H1-H3 only differ by a time rescaling such that this figure represents them all equally well.) One can imagine an analogous figure for H4-H7.

In addition to translating the state-of-the-world at each time into the bandlimited setting, we can also translate over the dynamics. This translation is aided by the fact that the derivative is the generator of translations, i.e., $h(x + a) = \exp(a \partial_x) h(x)$. For H1 we have,

\[
\frac{\partial}{\partial t} \phi_B(t, x) = \sum_{n=-\infty}^{\infty} S_n(x/a) \frac{d}{dt} \phi_n(t) = \alpha \sum_{n=-\infty}^{\infty} S_n(x/a) \left[ \phi_{n+1}(t) - 2\phi_n(t) + \phi_{n-1}(t) \right]
\]

\[
= \alpha \sum_{n=-\infty}^{\infty} S_n(x/a) \left[ \phi_n(t) - 2\phi_B(t, x) + \phi_B(t, x + a) \right] - \alpha \left[ \exp(-a \partial_x) - 2 + \exp(a \partial_x) \right] \phi_B(t, x)
\]

\[
= \alpha (2\cosh(a \partial_x) - 2) \phi_B(t, x).
\]

Similarly for the other theories we have:

\begin{align*}
\text{H2: } \partial_t \phi_B &= \frac{\alpha}{6} [-\cosh(2a \partial_x) + 16\cosh(a \partial_x) - 15] \phi_B \\
\text{H3: } \partial_t \phi_B &= \alpha a^2 \partial_x^2 \phi_B \\
\text{H4: } \partial_t \phi_B &= \alpha [\cosh(a \partial_x) + \cosh(a \partial_y) - 2] \phi_B \\
\text{H5: } \partial_t \phi_B &= \alpha [\cosh(a \partial_x) + \cosh(a \partial_y)] + \cosh(a (\partial_x - \partial_y) - 3) \phi_B, \\
\text{H6: } \partial_t \phi_B &= \frac{\alpha a^2}{2} (\partial_x^2 + \partial_y^2) \phi_B \\
\text{H7: } \partial_t \phi_B &= \frac{\alpha a^2}{3} (\partial_x^2 + \partial_y^2 + (\partial_x - \partial_y)^2) \phi_B
\end{align*}

We can easily solve each of these dynamical equations. Just as in Sec. III these dynamics admit a complete basis
of planewave solutions. Here we have:

$$H1-H3: \phi(t, x; k) = e^{-ikx} e^{-\Gamma(k)t} \tag{69}$$

$$H4-H7: \phi(t, x, y; k) = e^{-ik_1x-i\bar{k}_2y} e^{-\Gamma(k_1, k_2)t}.$$ 

Each of these planewaves decays at the same rates given in Sec. III. There is however, one substantial difference here. Before the wavenumber was restricted to $k \in [-\pi, \pi]$ with solutions with $k$ outside of this range being identical solutions in this range. Here, the wavenumber is unrestricted. However, $\phi_{P}$ simply has no support over these solutions outside of $k \in [-K, K]$.

### C. Bandlimited Locality

Let’s next develop a sense of comparative locality for H1-H7 according to this interpretation. Recall that in Sec. III A we found an intuitive notion of locality such that H1 > H2 > H3 and H4, H5 > H6, H7 with higher rated theories being more local. Viewed from a bandlimited perspective however, a different notion of locality becomes natural. Indeed, as I will now discuss, all of these locality comparisons are here reversed.

In general, differential equations are considered local when they only involve derivatives up to a finite order. Each of these derivatives is a local operation and there is no way to build from a finite set of them something non-local. However, when one is allowed an infinite number of derivatives one can create non-local dynamics. Recall that $h(x + a) = \exp(a \partial_x) h(x)$. Indeed, this is exactly what is going on in the dynamical equations of H1, H2, H4 and H5. From a bandlimited perspective, these are highly non-local theories despite previously being the most local. The bandlimited function $\phi_{B}$ is instantaneously coupled to the value it takes a distance of $a$ or even $2a$ away. By contrast, H3, H6 and H7 are perfectly local from the bandlimited perspective. On this new notion of locality we have H3 > H1, H2 and H6, H7 > H4, H5. These are essentially the reverse judgements of what we had before.

But which of these two notions of locality should we care about? This depends on which view we take of the spacetime manifold underlying these theories. Indeed, it’s not surprising that changes in the underlying manifold as drastic as $Q \rightarrow M_{1-7}$ will have drastic consequences for our intuitive notions of locality. As I have discussed previously, (unlike with $M_{1-7}$) taking $Q$ to be the manifold underlying H1-H7 systematically underpredicts these theories’ symmetries. Thus we find substantial reason to prefer the bandlimited notion of locality associated with $M_{1-7}$.

One may still be puzzled, however. Suppose we stick with viewing $M$ as the underlying manifold. One may reason (poorly) as follows: When we view the dynamics of H3 (or H6 or H7) in terms of $\phi_{P}$ it is local. However, when we view the dynamics in terms of its sample points (which are after all extremely local: samples at a point) we find dynamics like Eq. (24) which couples the sample points to each other at an infinite range via $D$. What gives?

The oversight in the above line-of-thought is thinking that the sample points correspond to localized degrees of freedom of $\phi_{B}$. They do not. Yes, the sample point of $f_{B}$ at $x_{0}$ can be understood as

$$f_{B}(x_{0}) = \int dx f_{B}(x) \delta(x-x_{0}). \tag{70}$$

However, it is also true for every bandlimited function $f_{B}$ and for every $x_{0}$ that,

$$f_{B}(x_{0}) = \frac{1}{a} \int dx f_{B}(x) S \left( \frac{x-x_{0}}{a} \right). \tag{71}$$

for $a < \pi/K$. This is because $\delta(x)$ and $S(x/a)$ have identical Fourier transforms for $k \in [K,-K]$. For bandlimited functions, these two kernels are identical.

We thus have two mathematical representations for what it means to evaluate a bandlimited function at a point. As I will now discuss, the second representation is more in line with the nature of bandlimited function than the first. Mathematically, this is because its kernel is bandlimited whereas the other’s (the Dirac delta) is not. In Eq. (70), we project the bandlimited function onto a kernel outside of the bandlimited universe, whereas in Eq. (71) we stay within the bandlimited universe.

Standardly, functions are thought of as something which maps some “point” in an input space to some “point” in an output space. Thus, standardly, at the core of being a function is the notion of “evaluating a function at a point”. This is often taken as a primitive unanalyzable operation: it’s just what functions do. However, pretend for a moment that we meet an alien species who have never thought of functions in this way. Rather, they take as an primitive unanalyzable operation: it’s just what functions do. Thus, they would likely answer them by pointing to an equation like Eq. (70).

To this they may ask, how do we know that such distributions as the Dirac delta exist when and where we need them? In a bandlimited context they do not. Thus, bizarrely, for bandlimited functions the supposedly basic notion of “evaluating a function at a point” breaks down. This likely has consequences for how we think of the spacetime manifold (if points aren’t the sort of things we can evaluate functions at, what are they?) but this is a question for another paper.

There is a physical story which runs parallel to this mathematics regarding the localization of degrees of freedom and counterfactuals. My claim is that if we restrict our attention to bandlimited functions, then bandlimited functions have no localized degrees of freedom. I am
here understanding degrees of freedom as things which can vary independently from each other. This is a context sensitive notion in that it depends on both what the other candidate degrees of freedom are and how we are allowed to vary them. One cannot (while keeping \( f_B \) bandlimited) change the value of \( f_B \) only at one point or even only in a compact region. Suppose you could. The difference between the function before and after the change would itself have to be bandlimited (the set of bandlimited functions is closed under subtraction). But this is impossible since no bandlimited function can be compactly supported. Every compactly supported function has non-zero support over all wavenumbers.

To be clear: whether or not the sample value \( f_B(x_n) \) is a local degree of freedom of \( f_B \) depends on context. Suppose we fix \( f_B \) by giving its values at some (potentially non-uniform) sample lattice \( x_n \). In one sense, all of these sample values are degrees of freedom because we can vary them all independently. Changing each of these would change \( f_B \) almost everywhere, but its values at all the other sample points would remain the same. One cannot however, vary one of these sample values while only changing the function locally. Thus, for both physical and mathematical reasons it is improper to associate the sample values of a bandlimited function with the sample point (besides as a mere label). Nor can one associate the sample values with the weighted average of the function over some compact region\(^\text{15}\). If the sample value is to be associated with some weighted average it must be over the whole domain, e.g., as in Eq. (71).

In light of this we may want to clarify what exactly is meant by Eq. (59). We ought not think of pinning the bandlimited function down at these points \( \phi(t) \) on the manifold. Bandlimited functions don’t know what points are. Rather, this must be understood in a softer way as fixing a certain weighted average of the function, along the lines of Eq. (71).

Ultimately, for \( H3, H6, \) and \( H7, \) the apparent tension in between the locality of the dynamics for \( \phi_B(t, x) \) namely and the non-locality of the dynamics for its sample values, \( \phi_n(t) = \phi_B(t, x_n) \) is resolved as thus. The sample values themselves are to be understood as non-local objects. Hence, it is unsurprising if these non-local things obey non-local dynamics.

We thus have good reason to favor the bandlimited notion of locality introduced here over the intuitive one introduced in Sec. 4. Another such reason is given in the next section: unlike the bandlimited notion of locality, the intuitive notion of locality discussed above is fragile and not preserved under resampling.

### D. Bandlimited Nyquist-Shannon Resampling

Let’s next see how changing between different lattice representations affects the dynamics.

#### Equivalence of \( H6 \) and \( H7 \) via Resampling

First, let’s see what this third interpretation has to say about the one-to-one correspondence between the solutions to \( H6 \) and \( H7 \) noted following Eq. (37). In our first interpretation, \( H6 \) and \( H7 \) were seen as different theories with different symmetries despite this correspondence. In our second interpretation, however, \( H6 \) and \( H7 \) were seen as equivalent via a change of basis on the value space.

As I will now discuss, here \( H6 \) and \( H7 \) are still seen as equivalent, but now via a change of coordinate and a change of sample points. Namely, we can make sense of the skew transformation Eq. (37) in terms of a coordinate transformation as:

\[
\begin{align*}
x & \mapsto x + \frac{1}{2} y \\
y & \mapsto \sqrt{\frac{3}{2}} y.
\end{align*}
\]  

(72)

None of our previous interpretations were able to make sense of Eq. (37) in terms of a coordinate transformation because they had no continuous manifold on which to define it.

Applying this transformation to \( H7 \)'s dynamics, namely Eq. (68), we find

\[
H7 : \partial_t \phi_B = \frac{\alpha}{2} \frac{a^2}{2} (\partial_x^2 + \partial_y^2 ) \phi_B.
\]

(73)

Thus, in their bandlimited formulation, \( H6 \) and \( H7 \) are just a change of coordinates away from each other. This is a much stronger notion of equivalence than just having a one-to-one correspondence between solutions.

Beginning from this unified bandlimited description of \( H6 \) and \( H7 \), how should we understand the two (seemingly different) dynamical equations we started from, namely Eq. (30) and Eq. (31)? As I will now discuss, these discrete dynamical equations result from describing the single \( \phi \) with different sample points.

Note that in Sec. 4.1.4, we have embedded both \( H6 \) and \( H7 \) onto the manifold via a square lattice, \( \tau_n, m(t) = (t, n a, m a) \). However, applying the coordinate transformation which maps \( H7 \) onto \( H6 \), namely Eq. (72), transforms a square lattice onto a hexagonal one. See Fig. 6. Thus, taking into account this coordinate change, we have effectively embedded \( H7 \) onto our manifold using a hexagonal lattice.
Indeed, after applying Eq. (72) to H7, the only remaining difference between it and H6 is that H6’s sample points form a square lattice and H7’s form a hexagonal lattice. Thus, in our third interpretation H6 and H7 are seen as describing the same bandlimited function just using different sample points. We have thus, not only shown in what ways these theories are identical (as the second interpretation also did) but we have also shed light on what is going on behind the scenes in our first interpretation.

Our second big lesson holds true in this interpretation just as it did in the second one: discrete theories which are presented to us with very different lattice structures (i.e., a square lattice versus a hexagonal lattice), may nonetheless turn out to be completely equivalent theories. In this interpretation, the process for switching between lattice structures is simply reformulating as a bandlimited function, and then resampling.

**Boosted Resampling of H1-H3**

In order to better see how this process or resampling works in general, let’s work through another example. In particular, I will first recover the discrete dynamics for H1, namely Eq. (11), from its bandlimited dynamics, namely Eq. (62). Then I will discuss how one might resample H1-H3 using boosted sample points.

In Sec. VII A we embedded H1-H3 onto a manifold via the sample point \( z_n(t) = (t, n a) \). Following this in Sec. VII B we reconstructed the bandlimited field \( \phi_B(t, x) \) and solved its dynamics. In particular, we found exponentially decaying plane wave solutions Eq. (60). One of these plane wave solutions is shown in Fig. 4 along with its original sample points (vertical black lines). Note that at a fixed wavenumber, H1-H3 only differ by a time rescaling such that this figure represents them all equally well.

Before considering the boosted sample points (red lines in Fig. 5) let’s first cast the bandlimited dynamics for H1 down onto these stationary sample points.

Using the identity Eq. (71) for bandlimited functions we have

\[
\frac{d}{dt} \phi_n(t) = \frac{1}{a} \int dx S_n(x/a) \partial_t \phi_B(t, x) \tag{74}
\]

From here we would like to get \( (d/dt) \phi_n(t) \) in terms of the other sample values, \( \phi_m(t) \). To do this we can rewrite \( \partial_t \phi_B(t, x) \) as follows:

\[
\partial_t \phi_B(t, x) = a[2\cosh(a \partial_x) - 2] \phi_B(t, x) = a[\exp(-a \partial_x) - 2 + \exp(a \partial_x)] \phi_B(t, x) = a[\phi_B(t, x - a) - 2\phi_B(t, x) + \phi_B(t, x + a)]
\]

\[
= \alpha \sum_m [S_m(x/a) - 2S_m(x/a) + S_{m+1}(x/a)] \phi_m(t)
\]

where in the last step we have used Eq. (61). Plugging this into Eq. (74) and using the fact that \( \{S_n(x)\}_{n \in \mathbb{Z}} \) form an orthonormal basis in the \( L^2 \) norm we have,

\[
\frac{d}{dt} \phi_n(t) = \alpha \sum_m [\delta_{n, m - 1} - 2\delta_{n, m} + \delta_{n, m + 1}] \phi_m(t)
\]

\[
= \alpha [\phi_{n+1}(t) - 2\phi_n(t) + \phi_{n-1}(t)].
\]

Thus we have recovered the discrete dynamics of H1, Eq. (11) from its bandlimited dynamics, Eq. (62). Using a similar process we can recover the discrete dynamics for H2-H7 from their bandlimited dynamics.

We can do more than this however. We can not only recover the original discrete dynamics from the bandlimited dynamics, but new discrete dynamics as well. We can do this by describing the bandlimited dynamics on a new set of sample points.

For instance, let’s consider H1 sampled on the boosted sample points (slanted red lines) shown in Fig. 5 with

\[
\text{Boosted: } z_n^{\text{Boost}}(t) = (t, n a + v t)
\]

for some speed \( v \). Let \( \varphi_n(t) \) be the sample values at these new sample points, that is

\[
\varphi_n(t) = \phi_B(z_n^{\text{Boost}}(t)).
\]

What are the dynamics which these new sample values obey?

Using the identity Eq. (71) for bandlimited function we have,

\[
\frac{d}{dt} \varphi_n(t) = \frac{1}{a} \int dx S_{n+v t/a}(x/a) \frac{d}{dt} \phi_B(t, x)
\]

\[
= \frac{1}{a} \int dx S_n(x/a) \frac{d}{dt} \phi_B(t, x - v t)
\]

\[
= \frac{1}{a} \int dx S_n(x/a) \left[ \partial_t \phi_B - v \partial_x \phi_B \right]_{(t, x-v t)}
\]

Repeating our previous derivation Eq. (75) we can simplify the first term. This leads us to

\[
\frac{d}{dt} \varphi_n(t) = \alpha [\varphi_{n+1}(t) - 2\varphi_n(t) + \varphi_{n-1}(t)]
\]

\[
- \frac{v}{a} \int dx S_n(x/a) \partial_x \phi_B(t, x - v t)
\]
Making use of the derivative approximation (which is exact for bandlimited functions) given by Eq. (55) and collecting these sample values into a vector $\varphi(t) = (\ldots, \varphi_{-1}(t), \varphi_0(t), \varphi_1(t), \ldots)$ we have

$$\text{H1:} \quad \frac{d}{dt} \varphi(t) = \alpha \Delta^2(1) \varphi(t) - \frac{v}{a} D \varphi(t). \quad (81)$$

Repeating this process for H2 and H3 we would find,

$$\text{H2:} \quad \frac{d}{dt} \varphi(t) = \alpha \Delta^2(2) \varphi(t) - \frac{v}{a} D \varphi(t) \quad (82)$$

$$\text{H3:} \quad \frac{d}{dt} \varphi(t) = \alpha D^2 \varphi(t) - \frac{v}{a} D \varphi(t). \quad (83)$$

Note that the appearance of this new term in the dynamics means that none of these theories are Galilean boost invariant.

Also note how the infinite range discrete derivative operator $D$ appears in each of these equations, even when we start off with only finite range discrete derivative approximation. Moreover, note that while before this resampling H1-H3 were local in the intuitive sense of Sec. IV A (i.e., nearest-neighbor couplings only), they are no longer once we have resample them. Thus, this intuitive notion of locality is uncomfortably representation dependent and hence unphysical. Another example of this loss of intuitive locality under resampling is described in the next section.

Resampling H5 on a Square Lattice

Before going on to discuss the symmetry of these theories under this third interpretation, one final resampling should be discussed regarding H5.

Noted that applying the coordinate transformation Eq. (72) to H5 changes its dynamics from Eq. (66) to:

$$\text{H5:} \quad \partial_t \phi_B = \frac{\alpha}{3} \left[ 2 \cosh(a \partial_x) - 2 \right] + 2 \cosh(a \sqrt{3} \partial_y + \partial_z)/2 - 2$$

Note that this dynamics manifestly has a one-sixth rotation symmetry.

Like with H7, it is this version of H5’s dynamics which we can think of as being sampled on a hexagonal lattice to give back (13). However, we do not have to sample this theory on a hexagonal lattice. Sampling it on a square lattice has the effect of taking $\partial_x \rightarrow D_n$ and $\partial_y \rightarrow D_m$ resulting in the discrete dynamics,

$$\text{H5:} \quad \partial_t \phi_B = \frac{\alpha}{3} \left[ 2 \cosh(D_n) - 2 \right] + 2 \cosh((\sqrt{3} D_m + D_n)/2) - 2$$

This is equivalent to Eq. (29) under a change of sample points from hexagonal to square.

Note that before this resampling H5 was local in the intuitive sense of Sec. IV A (i.e., nearest-neighbor couplings only), it is no longer. Thus, this intuitive notion of locality is uncomfortably representation dependent and hence unphysical.

E. Bandlimited Symmetries

Now that we have translated the dynamics of our seven heat equations into a bandlimited setting, we can now discuss their dynamical symmetries. While no new symmetries have been revealed in moving from our second to our third interpretation, the symmetries are represented and classified differently. In particular, all of the symmetries (except for the linear-affine rescaling) are now represented as external symmetries. In particular, the continuous translation and rotation symmetries identified earlier are now honest-to-goodness manifold symmetries, represented by diffeomorphisms $d \in \text{Diff}(M)$.

Our first big lesson holds true here as well: discrete spacetime theories can have (external) continuous translation and rotation symmetries. The fact that our discrete theories at first appeared on some lattice with some lattice structure does nothing to forbid this.

I could end this section here, but I think it is helpful to see the independence of dynamical symmetries on the lattice and lattice structure explicitly. The symmetries of a dynamics has nothing to do with the symmetries of the lattice it is represented on. Just as we can represent any bandlimited state on any lattice, so too can we represent any bandlimited dynamics on any lattice. To see this consider Figs. 7, 8 and 9.

In each of these figures we begin from some initial heat distribution with a bandlimited representation,

$$\phi_B(0,x,y) = \frac{J_1(\pi r)}{\pi r} + \frac{J_0(\pi r) - J_2(\pi r)}{2} \quad (86)$$

where $J_n(r)$ is the $n^{th}$ Bessel function and $r = \sqrt{x^2 + y^2}$. This function is bandlimited with bandwidth of $K = \pi$ and is rotationally invariant. This function is shown in the first columns of Figs. 7, 8 and 9.

We can therefore represent this function with sampling on a square lattice with $a = 1$. We could equivalently represent this function on a hexagonal lattice or even an irregular lattice. For each of H4, H5 and H6=H7 such representations are shown in the second columns of Figs. 7, 8 and 9.

For each of these theories, we then have a choice of which representation to carry out the dynamics in. I here consider four options: as a bandlimited function, as samples on a square lattice, as samples on a hexagonal lattice or as samples on an irregular lattice. The various options for H4, H5 and H6=H7 are shown in Figs. 7, 8 and 9 respectively.

Let’s begin with the dynamics of H5 represented on a hexagonal lattice. This is shown in the middle row of Fig. 9. The bandlimited representation of the initial heat
The dynamics of H4 is here shown being carried out in a variety of lattice representations. In the leftmost column the initial condition is shown in its bandlimited representation, given by Eq. (86). In the rightmost column the final evolved state is shown in its bandlimited representation. Here the evolution time is $t = 0.8$ and the diffusion rate is $\alpha = 1$. This state can be found in four different ways. Firstly by applying the dynamics Eq. (65) to of Eq. (86). The other three ways are shown in the three rows of this figure. The first row shows the initial condition being sampled onto a square lattice. This is then evolved forward in time via Eq. (12). The bandlimited representation of the final state is then recovered through the methods discussed in Sec. VI. The second and third rows show the same process carried out on a hexagonal lattice and an irregular lattice. Notice that the final state has a 4-fold symmetry regardless of how the dynamics is represented. Notice that the final state is the same regardless of how the dynamics is represented.

Alternatively, we could have carried out this evolution with no lattice representation at all. That is, we could have skipped from Fig. 8b1 directly to Fig. 8b4. We could do this by applying the dynamics Eq. (84) directly to the bandlimited initial condition Eq. (86). It is in this sense that the bandlimited and discrete representations of our dynamics are equivalent. The first and third rows of Fig. 8 show the exact same evolution via H5 represented on different lattices, namely a square lattice and an irregular lattice. In the first row the evolution is carried out by a resampled version of Eq. (13), namely Eq. (85). In the third row the evolution is carried out by whatever resampling of Eq. (13) corresponds to this irregular lattice.

Notice that the final state has a 6-fold symmetry regardless of how the dynamics is represented. Moreover, notice that the final state is the same regardless of how the dynamics is represented. Just as we can represent any bandlimited state on any lattice, so too can we represent any bandlimited dynamics on any lattice.

Fig. 7 makes the same demonstration for H4. Notice that the final state has a 4-fold symmetry regardless of how the dynamics is represented. Notice that the final state is the same regardless of how the dynamics is represented.
FIG. 8. The dynamics of H5 is here shown being carried out in a variety of lattice representations. In the leftmost column the initial condition is shown in its bandlimited representation, given by Eq. (86). In the rightmost column the final evolved state is shown in its bandlimited representation. Here the evolution time is $t = 8.3$ and the diffusion rate is $\alpha = 1$. This state can be found in four different ways. Firstly by applying the dynamics Eq. (84) to Eq. (86). The other three ways are shown in the three rows of this figure. The second row shows the initial condition being sampled onto a hexagonal lattice. This is then evolved forward in time via Eq. (13). The bandlimited representation of the final state is then recovered through the methods discussed in Sec. VI. The second and third rows show the same process carried out on a square lattice and an irregular lattice. Notice that the final state has a 6-fold symmetry regardless of how the dynamics is represented. Notice that the final state is the same regardless of how the dynamics is represented.

Likewise, Fig. 9 makes the same demonstration for H6. Notice that the final state is rotation invariant regardless of how the dynamics is represented. Notice that the final state is the same regardless of how the dynamics is represented.

These figures demonstrate clear as can be that a theory’s lattice structure has nothing to do with its dynamical symmetries. We can represent any bandlimited dynamics on any lattice.

F. Bandlimited General Covariance

As the above discussion has shown, giving our discrete theory a bandlimited representation has had many of the same benefits one expects from a generally covariant formulation. Namely, we have exposed certain parts of our theory as merely representational artifacts and in the process we have come to a better understanding of our theory’s symmetries and background structures. This is the work of the titular discrete analog of general covariance. This analogy will be spelled out in detail in the following section.

Now, however, I show how to combine this discrete analog with our usual continuum notion of general covariance. To do this, one simply takes the dynamical equations of H1-H7 (i.e., Eq. (62)-Eq. (68)) and recast them in the coordinate-free language of differential geometry. For simplicity, however, I will just consider H4 and H6=H7 here.

Beginning with H6=H7 we should first note that its dynamics Eq. (67) are nearly identical to those of H0 the
FIG. 9. The dynamics of $H_6$ is here shown being carried out in a variety of lattice representations. In the left most column the initial condition is shown in its bandlimited representation, given by Eq. (86). In the rightmost column the final evolved state is shown in its bandlimited representation. Here the evolution time is $t = 1$ and the diffusion rate is $\alpha = 1$. This state can be found in four different ways. Firstly by applying the dynamics Eq. (67) to Eq. (86). The other three ways are shown in the three rows of this figure. The first row shows the initial condition being sampled onto a square lattice. This is then evolved forward in time via Eq. (30). The bandlimited representation of the final state is then recovered through the methods discussed in Sec. VI. The second and third rows show the same process carried out on a hexagonal lattice and an irregular lattice. Notice that the final state is rotation invariant regardless of how the dynamics is represented. Notice that the final state is the same regardless of how the dynamics is represented.

We can rewrite $H_6=H_7$ in the coordinate-free language of differential geometry as follows.

$$H_6=H_7: \begin{align*}
&\text{KPMs: } \langle \mathcal{M}, t_{ab}, h^{ab}, \nabla_a, T^a, \phi_B \rangle \\
&\text{DPMs: } T^a \nabla_a \psi = \frac{\alpha_0}{2} h^{bc} \nabla_b \nabla_c \phi_B.
\end{align*} (87)$$

The geometric objects used in this formulation are all just as defined following Eq. (10) except that $\phi_B$ here is bandlimited. In order for this reformulation to really be coordinate-free, we need some geometric way of understanding that $\phi_B$ is bandlimited. But how can this be expressed geometrically, i.e., in terms of $h^{ab}$ and $\nabla_a$?

Consider any space-like hypersurface in this spacetime. That is, consider any surface $H \in \mathcal{M}$ such that all of its tangent vectors $x^a$ have $t_{ab} x^b = 0$. Consider the eigen-problem for functions $f : H \to \mathbb{R}$ defined on this surface, $h^{bc} \nabla_b \nabla_c f = -\lambda f$. Because the spacetime is flat, we know that $H$ is flat and therefore the eigensolutions are all planewaves, with frequency $k = \sqrt{\lambda}$. We can now say what it means for $\phi_B$ to be bandlimited.

$\phi_B$ is bandlimited if and only if for any space-like hypersurface $H$ if we restrict $\phi_B$ to $H$ and then expanded it in the above discussed eigenbasis, then only eigensolutions with eigenvalues in some fixed finite range are...
needed. The extent of this range is the bandwidth of \( \phi_B \).

Note that spelling out what it means for \( \phi_B \) to be bandlimited did not require talking about \( T^a \). Thus, this geometric definition of being bandlimited can be applied in Galilean spacetimes as well. We did however make use of the flatness of the spacetime. As such, another geometric definition of being bandlimited will need to be developed for curved spacetimes [44].

Before further analyzing \( H_6=H_7 \), let’s next consider \( H_4 \). Rewritten in the coordinate-free language of differential geometry, \( H_4 \) becomes:

\[
\text{H4: KPMs: } \langle M, t_{ab}, h^{ab}, \nabla_a, T^a, X^a, Y^a, \phi_B \rangle \quad (88)
\]

\[
\text{DPMs: } T^a \nabla_a \psi = \frac{\alpha}{2} F(X^b \nabla_b, Y^c \nabla_c) \phi_B,
\]

where

\[
F(x, y) = 2\cosh(a x) + 2\cosh(a y) - 4. \quad (89)
\]

Here \( X^a \) and \( Y^a \) are a pair of fixed constant space-like unit vectors which are orthogonal to each other. That is,

\[
\nabla_a X^b = 0 \quad \nabla_a Y^b = 0 \quad (90)
\]

\[
t_{ab} X^a = 0 \quad t_{ab} Y^a = 0
\]

\[
h_{ab} X^a X^b = 1 \quad h_{ab} Y^a Y^b = 1
\]

\[
h_{ab} X^a Y^b = 0
\]

Note that the inverse space metric \( h_{ab} \) is only well defined for spacelike vectors, see [8]. Roughly, \( X^a \) and \( Y^a \) here serve to pick out the directions for the rotational anomalies appearing in Fig. 6.

Now that we have applied both our discrete and continuous notion of general covariance to \( H_4 \) and \( H_6=H_7 \), we should be in a position to identify their background structures and symmetries.

The fixed fields \( X^a \) and \( Y^a \) appearing in \( H_4 \) count as additional background structures and limit the spacetime symmetries of \( H_4 \). Namely they forbid rotation invariance.

Turning our attention towards \( H_6=H_7 \) we see that, perhaps surprisingly, its background structures and symmetries are exactly the same as \( H_0 \)'s. The only difference between these \( H_6=H_7 \) and \( H_0 \) is that here \( \phi_B \) is bandlimited whereas there \( \psi \) is not. This is a restriction at the level of KPMs as to which dynamical fields are allowed. In fact, in either theory the dynamics guarantees that if the temperature field starts off bandlimited it will stay bandlimited. Thus this restriction of the allowed dynamical fields is really just a restriction on the allowed initial conditions. Thus, ultimately the only difference between \( H_6=H_7 \) and \( H_0 \) is a restriction on the initial conditions.

As innocent as this restriction on initial conditions may seem, it has serious implications for counterfactual reasoning. As discussed following Eq. (71), when restricted to bandlimited functions we can no longer ask “What would have happened, if things had been different only in this compact region?” Any bandlimited counter-instance must be globally different.

However, other than this restriction there is no substantial difference between \( H_6=H_7 \) and \( H_0 \). Any lattice structure suggested by our original formulation of \( H_6 \) and \( H_7 \), Eq. (30)-(31) has been revealed to be nothing more than a coordinate-like representational artifact. Our third big lesson is visible here in full force: given a discrete spacetime theory with some lattice structure we can always reformulate it in such a way that it has no lattice structure whatsoever. In fact, there is no longer even a lattice here. In this interpretation, this is done by reformulating it as a bandlimited function on some manifold.

\section{VIII. Two Discrete Analogs of General Covariance}

Three lessons have been repeated throughout this paper. Each of these lessons is visible in both our second and third attempts at interpreting \( H_1-H_7 \). Combined these lessons give us a rich analogy between lattice structures and coordinate systems: Lattice structure is rather less like a fixed background structure and rather more like a coordinate system, i.e., merely a representational artifact.

These three lessons run counter to the three first intuitions one is likely to have regarding lattice structure discussed in Sec. [4]. Namely, that lattices and lattice structure: restrict our symmetries, distinguish our theories, and are fundamentally “baked-into” the theory. As we have seen, they do not restrict our symmetries, they do not distinguish our theories and they are representational not fundamental. In particular, we have learned the following three lessons.

Our first lesson was that taking the lattice and/or lattice structure seriously as a fixed background structure or as a fundamental part of the underlying manifold systematically under predicts the symmetries that discrete theories can and do have. Indeed, discrete theories can have significantly more symmetries than our first intuitions might allow for. As Sec. [V] and Sec. [VII] have shown each of \( H_1-H_7 \) has a continuous translation symmetry despite being introduced with discrete lattice structures. Moreover, \( H_6 \) and \( H_7 \) even have a continuous rotation symmetry. The fact that a lattice structure was used in the initial statement of these theory’s dynamics does not in any way restrict their symmetries. There is no conceptual barrier to having a theory with continuous symmetries formulated on a discrete lattice.

In light of the proposed analogy between lattice structure and coordinate systems this first lesson is not mysterious. Coordinate systems are neither background structure nor a fundamental part of the manifold. The use of a certain coordinate system does not in any way restrict a theory’s symmetries. Indeed, it is a familiar fact that there is no conceptual barrier to having a rotationally invariant theory formulated on a Cartesian coordinate system.
Our second lesson was that discrete theories which are presented to us with very different lattice structures may nonetheless turn out to be completely equivalent theories. Indeed, as we have seen, two of our discrete theories (H6 and H7) have a one-to-one correspondence between their solutions. This despite the fact that these theories were initially presented to us with different lattice structures (i.e., a square lattice and a hexagonal lattice respectively).

However, when in Sec. [V] we took these lattice structures seriously as a fixed background structure, we found that despite having one-to-one correspondence H6 and H7 were inequivalent; they were here judged to have different symmetries. Only in Sec. [V] and Sec. [VII] when stopped take the lattice structure so seriously did we ultimately see H6 and H7 as having the same symmetries. Indeed, in these later two interpretations H6 and H7 were seen to be identical, simply re-descriptions of a single theory. In Sec. [V] this re-description is a change of basis in the theory’s value space, whereas in Sec. [VII] this re-description is a change of the sample points we are using to describe the bandlimited field state.

Moreover, as I have discussed in Sec. [VII] our ability to switch between two different lattice structures for H6=H7 holds more generally. For any discrete theory, we can always re-described it using any16 lattice structure we wish.

In light of the proposed analogy between lattice structure and coordinate systems this second lesson is not mysterious. Unsurprisingly, continuum theories presented to us in different coordinate systems may turn out to be equivalent. Moreover, we can always re-describe any continuum theory in any coordinates we wish.

Our third lesson was that, in addition to being able to switch between lattice structures, we can also reformulate any discrete theory in such a way that it has no lattice structure whatsoever and indeed no lattice whatsoever. I have shown two ways of doing this. In Sec. [V] this was done by internalizing the lattice structure into the theory’s value space. In Sec. [VII] this was done by embedding the discrete theory onto a continuous manifold using bandlimited functions. Adopting a lattice structure and switching between them was then handled using Nyquist-Shannon sampling theory, see Sec. [VII].

In light of the proposed analogy between lattice structure and coordinate systems this third lesson is not mysterious. This is analogous to the familiar fact, discussed in Sec. [II] that any continuum theory can be written in a generally covariant (i.e., coordinate-free) way. Thus, the two above-discussed ways of reformulating a discrete theory to be lattice-free are each analogous to reformulating a continuum theory to be coordinate-free (i.e., a generally covariant reformulation). Thus we have not one but two discrete analogs of general covariance. See Fig. 10

16 There is some subtlety here which was discussed in Sec. VII

Internal Discrete General Covariance:

- Coordinate Systems ↔ Lattice Structure
- Changing Coordinates ↔ Change of Basis in Value Space
- Gen. Cov. Formulation ↔ Internalized Formulation
  (i.e., coordinate-free) (i.e., lattice-free)

External Discrete General Covariance:

- Coordinate Systems ↔ Lattice Structure
- Changing Coordinates ↔ Nyquist-Shannon Resampling
- Gen. Cov. Formulation ↔ Bandlimited Formulation
  (i.e., coordinate-free) (i.e., lattice-free)

FIG. 10. A schematic of the two notions of discrete general covariance introduced in this paper. These are compared in Sec. VII. The internal strategy is applied to H1-H7 in Sec. V whereas the external strategy is applied to H1-H7 in Sec. VII.

Before contrasting these two analogies, let’s recap what they agree on. In either case, as one would hope, our discrete analog helps us to disentangle a discrete theory’s substantive content from its merely representational artifacts. In particular, in both cases, lattice structure is revealed to be non-substantive and merely representational as is the lattice itself. Lattice structure is no more attached or baked-into our discrete spacetime theories than coordinate systems are to our continuum theory. In either case, getting clear about this has helped us to expose our discrete theory’s hidden continuous symmetries.

What distinguishes these two notions of discrete general covariance is how they treat the lattice and lattice structure after it has been revealed as being coordinate-like and so merely representational. The approach in Sec. [V] was to internalize the lattice structure into the theory’s value space. By contrast, the approach in Sec. [VII] was to keep the lattice structure external, but to flesh it out into a continuous manifold such that it is no longer fundamental. Let us therefore call these two notions of discrete general covariance internal and external respectively.

While for H1-H7 these internal and external approaches have agreed on what symmetries there are, they have disagreed about how they are to be classified. Moreover, these two approaches pick out very different underlying manifolds for our discrete theories. As a consequence, they license very different conclusions about locality.

In each of these differences I find reason to favor the external approach. To briefly overview my feelings: It is more natural for the continuous translation and rotation symmetries of H1-H7 to be classified as external. Moreover, keeping the lattice structure external as a part of the manifold, allows us to draw intuitions about locality from it. However, neither of these reasons are decisive and I think either approach is likely to be fruitful for further investigation/use.
There is some subtlety here which was discussed in Sec. VII.

covariant (i.e., coordinate-free) way.

I have discussed, this is analogous to the familiar fact that any continuum theory can be written in a generally coordinate system. These lessons serve to undermine the strong analogy between lattice structures and coordinate systems. These lessons serve to undermine the three first intuitions about lattice structure laid out in Sec. I.

Firstly, as I have shown, taking lattice structure seriously as a fixed background structure (or as a fundamental part of the underlying manifold) systematically under predicts the symmetries that discrete theories can and do have. Indeed, as I have shown, lattice structure does not in any way restrict a discrete theory’s possible symmetries. Discrete theories can and do have significantly more symmetries than our first intuitions might allow for. There is no conceptual barrier to having a theory with continuous symmetries formulated on a discrete lattice. As I have discussed, this is analogous to the familiar fact that there is no conceptual barrier to having a continuum theory with rotational symmetry formulated on a Cartesian coordinate system.

Secondly, as I have shown, discrete theories which are presented to us with very different lattice structures (e.g., a square lattice versus a hexagonal lattice), may nonetheless turn out to be completely equivalent theories. Moreover, given any discrete theory with some lattice structure always re-describe it using a different lattice structure. As I have discussed, this is analogous to the familiar fact that our continuum theories can be described in different coordinates, and moreover we can switch between these coordinate systems freely.

Thirdly, as I have discussed, in addition to being able to switch between lattice structures, we can also reformulate any discrete theory in such a way that it has no lattice structure (and indeed no lattice) whatsoever. As I have discussed, this is analogous to the familiar fact that any continuum theory can be written in a generally covariant (i.e., coordinate-free) way.

While the details of switching between lattice structures and of lattice-free reformulation differ between the two notions of discrete general covariance mentioned above (see Fig. 10) the above three lessons are clear in either case. Lattice structure is very much coordinate-like and consequently ought to be viewed as a merely representational artifact.

This result is significant for two reasons. Firstly, it has consequences for other issues in the philosophy of space and time. More on this in Sec. X. However, this alone stands as a shocking conclusion.

One might have an intuition that the world could be fundamentally set on a lattice. This lattice might be square or hexagonal and we might discover which by probing the world at the smallest possible scales looking for violations of rotational symmetry, or other lattice artifacts. Many serious efforts at quantum gravity assume that the world is set on a lattice of some sort at the smallest scales. However, as this paper clearly demonstrates this just cannot be the case. The world cannot be “fundamentally set on a square lattice” (or any other lattice) any more than it could be “fundamentally set in a certain coordinate system”. Lattice structures are just not the sort of thing that can be fundamental; they are thoroughly representational. Spacetime cannot be discrete (even when it might be representable as such).

IX. CONCLUSION

This paper has introduced two discrete analogs of general covariance (see Fig. 10) and demonstrated their usefulness. In either case, as hoped, when applied to a discrete spacetime theory (i.e., a lattice theory) this discrete analog helps us disentangle the theory’s substantive content from its representational artifacts. Indeed, my analysis has shown that lattice structure is rather less like a fixed background structure or part of an underlying manifold and rather more like a coordinate system, i.e., merely a representational artifact. Ultimately, as I have shown, the lattice structure supposedly underlying any discrete “lattice” theory has the same level of physical import as coordinates do, i.e., none at all. Namely, lattice structure is no more attached or baked-into to our discrete spacetime theories than coordinate systems are to our continuum theory.

Three lessons learned throughout this paper support this strong analogy between lattice structures and coordinate systems. These lessons serve to undermine the three first intuitions about lattice structure laid out in Sec. I.

Firstly, as I have shown, taking lattice structure seriously as a fixed background structure (or as a fundamental part of the underlying manifold) systematically under predicts the symmetries that discrete theories can and do have. Indeed, as I have shown, lattice structure does not in any way restrict a discrete theory’s possible symmetries. Discrete theories can and do have significantly more symmetries than our first intuitions might allow for. There is no conceptual barrier to having a theory with continuous symmetries formulated on a discrete lattice. As I have discussed, this is analogous to the familiar fact that there is no conceptual barrier to having a continuum theory with rotational symmetry formulated on a Cartesian coordinate system.

Secondly, as I have shown, discrete theories which are presented to us with very different lattice structures (e.g., a square lattice versus a hexagonal lattice), may nonetheless turn out to be completely equivalent theories. Moreover, given any discrete theory with some lattice structure always re-describe it using a different lattice structure. As I have discussed, this is analogous to the familiar fact that our continuum theories can be described in different coordinates, and moreover we can switch between these coordinate systems freely.

Thirdly, as I have discussed, in addition to being able to switch between lattice structures, we can also reformulate any discrete theory in such a way that it has no lattice structure (and indeed no lattice) whatsoever. As I have discussed, this is analogous to the familiar fact that any continuum theory can be written in a generally covariant (i.e., coordinate-free) way.

As conclusive as the above discussion is, it opens a number of questions which require further investigation. Firstly, the above work can be extended in a number of directions: to Lorentzian theories, to theories with non-linear dynamics, to gauge theories, to gravitational theories, to first quantized theories, to second quantized theories, etc. Some interesting work has already been done in the physics literature about bandlimited quantum field theory.

Moreover, the above work raises some interesting questions about the nature of locality in a bandlimited world. As discussed in Sec. VII C, there are no local degrees of freedom in a bandlimited world. Indeed, to change a bandlimited field somewhere requires that we change it...
everywhere. As discussed in Sec. [VII], this non-locality
only shows at the level of KPs, that is as a restriction
on what worlds are possible before dynamics are con-
sidered, off-shell. The dynamics of bandlimited theories,
however, can be totally local however, see H3 and H6-H7.
What are the philosophical consequences of this new sort
of counterfactual non-dynamical non-locality?

Another set of interesting questions arises in connection
with the status of the manifold in the spacetime
manifold in the above discussion. Consider the following
in light of the dynamical vs geometrical spacetime de-
bate. Roughly, which of dynamical and spacetime sym-
metries are explanatorily prior. Are spacetime structures
merely codifications of the dynamical behavior of mat-
ter? Or do they have an independent existence and act
to govern the dynamical behavior of matter (by for in-
stance restricting its possible symmetries)? Moreover,
consider Norton’s complaint [45] that proponents of the
dynamical approach must assume some prior spacetime
symmetry or structure (name the spacetime manifold itself) to even
begin talking about the dynamics of matter let alone its
dynamical symmetries.

As I have shown in this paper, we can make sense of
dynamics and dynamical symmetries without (much of)
a manifold. In particular, in my second interpretation
given in Sec. [V] I moved (much of) the spacetime mani-
fold into the theory’s value space, leaving only time be-
in principle time could be internalized as well. I was
then able to analyze the theory’s symmetries and
decide which of these to externalize. In particular, in
Sec. [VII] I was able to design a manifold specifically to
fit with the theory’s already-studied dynamics. Thus the
spacetimes discussed in Sec. [VII] and the spacetime struc-
tures placed on top of them, are very much in line with
the dynamical approach.

This is (potentially) a very different situation to the
manifold which underlies our continuum heat equation
(H0 in Sec. [II]) if it is understood along the lines of the
dynamical approach. What underpins this difference? As
the above discussion has shown, H0 and H6=H7 only
differ as to whether the dynamical fields are bandlim-
ited. In the first case, the spacetime manifold seems to
be an ineliminable part of the theory, along the lines of
Norton’s complaint. However, in the second case the
manifold seems far from necessary. Indeed, the manifold
underlying H6=H7 was invented in Sec. [VII] as a means
of better codifying the dynamics of our theories (in full
accordance with the dynamical approach). Indeed, there
had been much discussion of the dynamical symmetries
of H6=H7 long prior to finding a suitable manifold for
these theories.

This is suggestive of the possibility that the manifold
underlying H0 is not so necessary as it first appears.
Indeed, if the manifold underlying H0 is necessary to even
begin describing the dynamics, its necessity is contingent
in the following sense: if the fields under consideration
had been bandlimited they could have been described
without a manifold. That is, the descriptive necessity of
the manifold is contingent upon the existence of arbitrar-
ily high frequencies in the physical fields.

However, more needs to be done to develop this point.
Principally, the above work ought to be extended to
Lorentzian theories as these serve as the main stage for
the dynamical vs geometric spacetime debate.

ACKNOWLEDGMENTS

The authors thanks James Read, Jason Pye, and Nick
Menicucci for their helpful feedback.

[1] O. Pooley, “Background Independence, Diffeomorphism
Invariance, and the Meaning of Coordinates,” (2015),
arXiv:1506.03512 [gr-qc]
[2] J. D. Norton, Reports on Progress in Physics 56, 791
(1993)
[3] M. Friedman, “Space-time theories,” in Foundations
of Space-Time Theories: Relativistic Physics and Philos-
ophy of Science (Princeton University Press, 1983) pp.
32–70.
[4] L. Freidel and N. Tah, “Substantive general covariance
and the einstein-klein dispute: A noetherian approach,”
(2021), arXiv:2109.08516 [hist-ph]
[5] J. Earman, ‘World enough and space-time: absolute ver-
sus relational theories of space and time’ (MIT Press,
Cambridge ; Mass ; London, 1989).
[6] J. Brian Pitts, Studies in History and Philosophy of Mod-
ern Physics 37, 347 (2006).
[7] O. Pooley, “Space-time theories,” in [EPSA philosopi-
ical issues in the sciences], edited by M. Suárez, M. Do-
rato, and M. Rédei (Springer, Dordrecht ; London, 2010)
Chap. 19, pp. 197–210.
[8] J. Read, Background independence in classical and quan-
tum gravity, Master’s thesis, University of Oxford (2016).
[9] T. Teitel, Studies In History And Philosophy Of Modern
Physics 65 , pp41 (2019).
[10] G. Belot, General Relativity and Gravitation 43 , 2865
(2011).
[11] P. Janotta and H. Hinrichsen, Journal of Physics A:
Mathematical and Theoretical 47 , 323001 (2014).
[12] M. Müller, SciPost Physics Lecture Notes , 028 (2021),
hits://scipost.org/10.21468/SciPostPhysLectNotes.28/pdf.
[13] V. J. Wright and S. Weigert, Quantum 5 , 588
(2021) https://quantum-journal.org/papers/q-2021-11-25-588/pdf/.
[14] F. Shahandeh, PRX Quantum 2 (2021), 10.1103/prxquantum.2.010330,
http://link.aps.org/pdf/10.1103/PRXQuantum.2.010330.
[15] L. Henderson, Studies in History and Philosophy of Sci-
ence Part A 72 , 292 (2020).
[16] J. Bub, Entropy 17 , 7374 (2015).
Appendix A: Analysis of Symmetries for H1-H7

This appendix will identify the symmetries for H1-H7 under the three interpretations put forward in the main paper. However, it is convenient to do this in the reverse order in which these interpretations were introduced in the main text. In general, the more work we do reformulating our theories the more clearly we can see their symmetries.

1. Symmetries in the Third Interpretation

Let’s begin by analyzing the symmetries of H1-H7 in our third attempt at interpreting them, in Sec. VII.

For H6=H7 we can look to Eq. (87) to determine their symmetries. For H6=H7 we find the same symmetries as we did for H0 following Eq. (10). Namely, the two-dimensional Euclidean group (spatial translations, rotations, and reflections) plus constant time shifts and a linear-affine rescaling.

Similarly, for H4 we can look to Eq. (88) to identify its symmetries. The extra spacetime structures restrict our rotation symmetries to just quarter turns and allow us only a restricted set of reflection symmetries. Applying the same analysis to H5 one would find we are restricted to one-sixth turns and only a restricted set of reflection symmetries (different from H4’s).

We could also cast H1-H3, (namely Eq. (62)-(64)) into a generally covariant form as well. Doing so we would find their symmetries are the one-dimensional Euclidean symmetries (different from H4’s).

2. Symmetries in the Second Interpretation

Let us next analyze the symmetries of H1-H7 in our second attempt at interpreting them, in Sec. VI.

For each of H1-H7, viewed through each of our three interpretations, we have a one-to-one correspondence between its solutions across these interpretations. However, as we learned moving from the first to second interpretation, this does not mean that these theories have the same symmetries in each interpretation. Granted, yes, it is true that using this correspondence between solutions, we can transfer any symmetry transformation in one interpretation over into some transformation in the other interpretations. However, what counts as a symmetry transformation in one theory might not in another. Contrast Eq. (38) with Eq. (39): The scope of possible symmetry transformation varies from interpretation to interpretation.

Thus every symmetry revealed above for our third interpretation gives us a candidate symmetry for our other two transformations, but more must be done. In particular, we need to check whether these transformations are of the forms allowed in Eq. (38) and Eq. (39).
Let’s begin with the second interpretation. As revealed in Sec. VI translation of a bandlimited function \( f_B(x) \rightarrow f_B(x + \epsilon) \) is represented in terms of its vector of sample values \( f \) as \( f \rightarrow T^\epsilon_B f \), see Eq. (52). Moreover, as discussed following Eq. (54), the operator \( T^\epsilon_B \) in Sec. VI is numerically identical to the operator \( T^\epsilon \) appearing in Sec. V.

Thus for H1-H3 our candidate symmetry for continuous translation is \( \phi(t) \rightarrow T^\epsilon \phi \). Likewise for H4-H7 the candidate symmetries are \( \phi(t) \rightarrow T^\epsilon_n \phi \) and \( \phi(t) \rightarrow T^\epsilon_m \phi \). These are all of the form Eq. (39) and are thus viable symmetries under our second interpretation. Indeed, this is a symmetry of H1-H7 under our second interpretation.

Similar considerations apply for constant time shifts, reflection symmetries and for linear-affine rescalings. It is easy to find the symmetry candidates here and check that they are of the form Eq. (39). The only other non-trivial symmetry to transfer over is the continuous rotation symmetry. To see this we need the following facts.

For functions on \( \mathbb{R}^2 \) rotations are generated through the derivative as \( h(R(x, y)) = \exp(\theta(x\partial_y - y\partial_x)) \). Suppose that \( h = h_B \) is bandlimited and we sample it in two ways. Once on some square lattice, and once on another lattice identical to the first but rotated around \( n = 0 \) and \( m = 0 \) by an angle \( \theta \). The sample values in these two case, \( h, h' \in \mathbb{R}^2 \otimes \mathbb{R}^2 \) are related as \( h' = R^\theta_B h \) where

\[
R^\theta_B = \exp(\theta(ND_{B,m} - MD_{B,n})) \tag{A1}
\]

with \( \theta \in \mathbb{R} \) and where \( N \) and \( M \) are position operators which return the first and second index. Finally, since \( D_B \) is numerically identical to \( D \) we have \( R^\theta_B = R^\theta \) as defined in Eq. (41).

This transformation is of the form Eq. (39) and is thus a viable symmetry under our second interpretation. Indeed, this is a symmetry of H6 under our second interpretation. Moreover, with a slight change of basis, it is for H7 as well.

3. Symmetries in the First Interpretation

Let us now check whether the above discussed symmetries are still symmetries in our first interpretation. To do this we just need to check which are of the form Eq. (38). That is, when are the linear transformations \( \Lambda \) in Eq. (39) permutations?

It is not hard to check when \( T^\epsilon \) and \( R^\theta \) reduce to permutations. Similarly for \( T^\epsilon_n \) and \( T^\epsilon_m \) reduce to permutations.

It is not hard to check when \( T^\epsilon \) and \( R^\theta \) reduce to permutations. Similarly for \( T^\epsilon_n \) and \( T^\epsilon_m \) reduce to permutations after the basis change which adapts it to H7. Ultimately, this shows the symmetries under the first interpretation are just those claimed in Sec. IV.