Emergence Without Limits: the Case of Phonons

Alexander Franklin and Eleanor Knox

August 2017

Abstract

Recent discussions of emergence in physics have focussed on the use of limiting relations, and often particularly on singular or asymptotic limits. We discuss a putative example of emergence that does not fit into this narrative: the case of phonons. These quasi-particles have some claim to be emergent, not least because the way in which they relate to the underlying crystal is almost precisely analogous to the way in which quantum particles relate to the underlying quantum field theory. But there is no need to take a limit when moving from a crystal lattice based description to the phonon description. Not only does this demonstrate that we can have emergence without limits, but also provides a way of understanding cases that do involve limits.

Introduction

‘Emergence’ is a tricky term, but an important one. Leaving aside its traditional philosophical use, the term has been widely used in the physics community at least since Anderson’s “More is Different” [1]. And there’s reason to believe that the term latches on to something, and that that something is a philosophically interesting trait which may be ascribed to a reasonably well-defined cluster of phenomena. Philosophers of physics have thus rightly offered various analyses of emergence. As none of these analyses cover the full range of examples to which the term applies, we suggest that they fail to capture an important feature of emergence. In what follows, we’ll look at one such example, that of the emergence of phonons in a crystal, argue that it isn’t well-captured by accounts of emergence that rely on the presence of limits or essential idealisations, and suggest an account that might do better.

Perhaps the broadest definition of emergence in the philosophy of physics literature comes from Jeremy Butterfield, who tells us that emergent behaviour
is behaviour that is “novel and robust relative to some comparison class” [13, p.1065]. We’ll take this as a jumping off point for our analysis, but as Butterfield is perfectly aware, the wide applicability of this definition lies in its lack of specificity. It provides the scaffolding for a full account of emergence; an analysis of novelty and robustness is needed to finish the construction. While robustness is reasonably well understood and amounts to invariance under relevant perturbations (see section 3.1), novelty has proved harder to spell out. However, an account based on asymptotic limits has received particular attention in recent years.

This account suggests that behaviour is novel (and robust) relative to some more fine-grained (lower-level) behaviour when the higher-level mathematics is derived from the lower-level mathematics via an asymptotic limit; see e.g. [6, 9, 13, 14]. Phase transitions have been held to be emergent in this sense, because relating this thermodynamic behaviour to a statistical mechanical description requires us to take the thermodynamic limit. In Butterfield’s view, this kind of relationship allows a novelty that is compatible with (Nagel-Schaffner) reduction:

my meanings of ‘emergence’ and ‘reduction’ are in tension with each other: since logic teaches us that valid deduction gives no new “content”, how can one ever deduce novel behaviour? (Of course, this tension is also shown by the fact that many authors who take emergence to involve novel behaviour thereby take it to also involve irreducibility.) My answer to this ‘how?’ question, i.e. my reconciliation, will lie in the use of limits. ... the idea is that one performs the deduction after taking a limit of some parameter: so the main morals will be that in such a limit there can be novelty, compared with what obtains away from the limit, and that (pace some authors) this should count as reduction, not irreducibility.

[12, p.1068]

Butterfield further notes both that the limits of interest need not be asymptotic and that the account is unlikely to capture all putative examples of emergence. But, despite this disclaimer, a casual reader of the literature would be forgiven for thinking that limits (or even singular limits) were the only game in town.

The fan of an asymptotic limits analysis faces a dilemma. On the one hand, one can, with Batterman, hold that the use of an asymptotic limit indicates a failure of reduction. But this makes the use of an asymptotic limit look mysterious at the lower level. Otherwise, one might agree with Butterfield that emergence is compatible with reduction. But, in order to make this gel with

1Batterman, notably, disagrees with Butterfield on this point, taking the infinite idealisations required in deriving the asymptotic limit to imply explanatory irreducibility; see [9, p.1033].
the asymptotic analysis, we need to give some explanation of the applicability of the limit in question. And these explanations often seem to dissolve the novelty that made the asymptotic analysis look emergent in the first place. It’s therefore helpful to look explicitly at an example of emergence that does not fit the asymptotic limit mold: our case study will provide an example of a kind of emergence that maintains explanatory novelty even in the face of reduction, escaping the tension engendered by the asymptotic analysis.

Phonons are the remarkably particle-like vibrational modes of a crystal. Their behaviour, and the way in which they are modelled, bears a startling similarity to that of particles as described by quantum field theory. The relation of quantum particles to the underlying quantum fields is often taken to be a paradigm case of emergence, claims that phonons are emergent therefore carry considerable weight – such claims are further defended in section 2.2. But, as we’ll argue, it’s impossible to shoehorn the phonon example into the standard narrative of asymptotic limits; although there are approximations involved, they do not involve singular limits. Instead, we’ll argue that the example is better captured by an account of novelty that focusses on novel explanation. We’ll argue that the phonon description counts as novel because the change to phonon variables makes novel abstractions available which allow for novel explanations. As such the example fits well with an account of novelty previously advocated in [24].

Section 1 discusses the relevant physics, starting with a simple example of normal modes in masses on springs, which will serve to illustrate the way in which a change of variable can facilitate explanation. We’ll then move on to discuss the more complex example of phonons, emphasising the various approximations and idealizations needed to move to the phonon description.

Section 2 examines how this physics fits with the asymptotic account. We’ll argue first that the move to the phonon description does not involve limits in any important sense, let alone singular ones. It does involve approximations and idealizations, but these are not essential in the way that essential idealisations accounts of emergence require. But phonons should nonetheless be thought of as emergent phenomena; their connection to the derivation of particles in quantum field theory makes this particularly compelling.

In fact, as elaborated in section 3, the phonon example fits much better with an account of emergence that takes the relevant sense of novelty to be explanatory novelty; in this section, the sense in which the phonon description is robust is also evaluated. Thus we argue that phonons can be thought of as novel and robust.
1 Some Important Variable Changes

Phonons are often called quasi-particles, and their particle-like behaviour will be important to our presentation here. However, at their most basic, phonons are modes of vibration in crystal lattices, and the move to a phonon description is characterised by a change of variables that suitably simplifies the description of these modes of vibration. This section introduces the physics of phonons in a fairly standard way by first introducing another example – masses on springs – which also makes use of a change of variables to simplify the description of modes of vibration. Our aim here is not merely to summarise the physics, but also to emphasise a particular feature of these variable changes: both allow us to abstract in ways that are particularly helpful for explanation.

1.1 Masses on Springs

Consider the following, very simple, case of two particles of equal mass $m$ oscillating on springs with constants $k$ and $k'$ as shown in the diagram:

![Figure 1: Coupled masses on springs.](image)

Their motion is characterised by the following equations:

\[
\begin{align*}
m\ddot{x}_1 &= -kx_1 - k'(x_1 - x_2) \quad (1) \\
m\ddot{x}_2 &= -kx_2 - k'(x_2 - x_1) \quad (2)
\end{align*}
\]

Despite the simplicity of the model, the standard (and only straightforward) way to solve these equations involves transforming the variables,

\[
\begin{align*}
\eta_1 &= x_1 + x_2 \\
\eta_2 &= x_1 - x_2.
\end{align*}
\]

One can thus convert equations (1) and (2) into linear uncoupled differential equations for two simple harmonic oscillators:
\[ m\ddot{\eta}_1 = -k\eta_1 \]
\[ m\ddot{\eta}_2 = -(k + 2k')\eta_2. \]  

If one makes the substitutions \( \omega_1 = \sqrt{k/m}, \omega_2 = \sqrt{(k+2k')/m}, \) these equations have the following solutions:

\[ \eta_1 = \alpha_1 e^{i\omega_1 t} + \alpha_1^* e^{-i\omega_1 t} \]
\[ \eta_2 = \alpha_2 e^{i\omega_2 t} + \alpha_2^* e^{-i\omega_2 t} \]  

where \( \alpha_1, \alpha_1^*, \alpha_2 \) and \( \alpha_2^* \) are set by initial conditions. These two equations and their solutions characterise normal modes of the system, and general solutions of the equations are superpositions of these normal modes. We’ll call the variables \( \eta_1 \) and \( \eta_2 \) that define these modes ‘normal mode variables’.

This system is much simpler than the phonon case. In particular, in the phonon case, the move to normal mode variables will involve important approximations and, as we’ll see in section 3.2, it might be thought to exhibit more novelty. But the case in hand suffices to highlight some aspects of the interest and importance of normal modes. In addition, in writing down these equations, we implicitly incorporate assumptions which may be seen as analogous to the harmonic approximation discussed below.

The change of variables described in equations (3) is obviously of great calculational value. But it could be useful for other reasons too. For example, the change of variables might aid the explanation of certain phenomena: suppose (in a rather stretched example) that the central spring is composed of a material that glows only when compressed, and that we seek to explain the frequency with which the spring lights up with the system in steady state. If we remain in the original displacement variables, any explanation of this phenomenon will require two variables and two equations. But if we move to the normal mode variables, only \( \eta_2 \) and its associated equation are relevant. By leaving out reference to \( \eta_1 \), this explanation in terms of the normal mode variables abstracts away from irrelevant details in a way that an explanation based on displacement variables cannot. This feature of a well-chosen change of variables is appealed to in [24] and plays a role in our analysis of the emergence of phonons.
1.2 Phonons

Needless to say, a real crystal lattice is a more complicated system than a pair of masses on springs. Here we have a quantum system comprised of electrons and different kinds of nuclei, arranged in a 3-D lattice. A number of approximations are required if we are to get a tractable description of normal modes. We start with the adiabatic approximation, which assumes that we can treat electronic and nuclear degrees of freedom independently (because electron masses are small and their velocities are fast relative to nuclear masses and velocities). Using this approximation, the account that follows will focus only on nuclear degrees of freedom and not electronic ones.

Now let's outline the physics of a classical three-dimensional crystal with \( n \) atoms. We could try to describe the vibrations of this crystal using \( 3n \) displacement variables and \( 3n \) coupled differential equations. But, happily, a crystal is a highly symmetric system, and things are a little easier than this. We start by considering the unit cells of the crystal – the basic repeating unit with \( m \) atoms – defined by the primitive translation vectors \( a_1, a_2 \) and \( a_3 \); see figure 2. For a very simple monatomic lattice, \( m = 1 \), for a very simple diatomic lattice \( m = 2 \), and so on. We need to consider enough of these unit cells to model inter-cell interactions adequately, but we do little to change the physics if we impose periodic boundary conditions and consider only the cells in some larger supercell containing \( N \) cells. In a 3-D lattice, we can impose a boundary condition after \( N_1, N_2 \) and \( N_3 \) (where \( N = N_1N_2N_3 \)) repetitions of the translation vectors \( a_1, a_2 \) and \( a_3 \) respectively. Exploiting this translational symmetry means we now

---

2This section should be treated with the usual philosophy of physics proviso. Our aim here is not to teach the reader solid state physics, but merely to highlight some selected features of philosophical interest. The reader unfamiliar with the physics and interested in more detail should look at e.g. [22, Ch.4].
have $3Nm < n$ coordinates $u_{Lsa}$, where $L$ labels the cell, $s$ labels the atom within the cell and greek indices label the cartesian components of displacement vectors. Although in some cases this reduces the number of coordinates, in many cases the supercell is taken to be of the order of the size of the crystal itself. Moreover, merely reducing the number of coordinates involved in the problem does little to make it calculationally tractable: more simplification is needed.

We therefore make a further approximation: the harmonic approximation. We assume that atomic displacements are small relative to interatomic distances. This allows us to expand the potential energy function in terms of the interatomic distances and drop terms above second order. If we now derive equations of motion from the Lagrangian in the usual way, we have a (very!) large set of coupled differential equations. The key to uncoupling these is – in close analogy with the method used for the masses on springs – to look for wave solutions, and then look for variables distinguished by wavevector $k$ rather than by the individual atoms. This effectively moves us from the real lattice to the reciprocal lattice, related by a discrete Fourier transform to the original. The result is a set of normal mode variables $q_{kj}$ labelled by their wavevectors $k$, with an additional index $j$ that tells us whether the mode described is optical or acoustic. When redescribed in terms of these variables, our original problem becomes one of a series of uncoupled harmonic oscillators; each normal mode variable obeys the equation of motion:

$$\ddot{q}_{kj} = -\omega^2_{kj} q_{kj}$$

Each of these variables corresponds to a particular mode of vibration of the crystal, and the atomic displacements naturally described by our original coordinates can be captured as linear combinations of the vibrational modes. Our normal mode variables are sometimes called phonons, but there’s an ambiguity in the terminology here: physics texts use the term ‘phonon’ to refer both to the variables themselves, and to the normal modes of vibration associated with them. This becomes particularly confusing when we talk about numbers of phonons – the number of coordinates remains the same for a given model! We’ll henceforth use the terms ‘phonon coordinate’ or ‘phonon variable’ to refer to $q_{kj}$ and reserve the term ‘phonon’ for the physical modes themselves.

With the new variables in hand, we naturally describe the crystal not in terms of individual atomic displacements, but in terms of aggregate properties of the whole supercell. This is not because we’ve moved to fewer variables; as noted above, there are $3Nm$ phonon coordinates just as there were $3Nm$ displacement coordinates. Whilst we still have as many variables available it turns out that certain behaviours (e.g. heat transport) can be characterised very simply using only relevant phonon variables. This, to foreshadow, is an abstraction which was unavailable before the variable change: had we con-
continued to work in the displacement variables, such behaviour may have been inexplicable.

The phonon coordinates are just coordinates – linear combinations of the displacement coordinates. It’s only in the harmonic approximation that they give us perfectly uncoupled equations of motion, and hence only the harmonic approximation that allows us to derive them as a particularly interesting class of variable, but there’s nothing to stop us using them to describe systems where the approximation doesn’t apply. Once we have derived the phonon variables using the harmonic approximation, we are free to relax the approximation, and reintroduce terms beyond second order in the atomic displacement. These anharmonicities are essential to explaining the properties of real crystals, and, indeed, are essential to some of the most characteristically particle-like features of phonons (although some of these will only become apparent once we move to a quantum description). As temperatures increase, the harmonic approximation will become less and less accurate until, at some high temperature, normal modes of vibration are no longer distinguishable even approximately. At this point, we might say that the phonons themselves disappear, even though the phonon variables still exist. But most interesting cases will lie somewhere in-between; the ontology of phonons, like that of most physical phenomena, is a fuzzy and approximate matter.

The description thus far has been entirely classical, and has given us no reason to think of phonons as resembling particles. For this we need quantum mechanics, specifically, we’d like to perform second quantisation on the phonon mode variables $q_{kj}$ and their associated momenta $p_{kj}$. Recall that our phonon variables feature in independent harmonic oscillator equations. The quantization procedure for a harmonic oscillator is part of elementary quantum mechanics. Given some set of independent harmonic oscillators, we can define operators that add or subtract a quantum of energy – these are sometimes known as creation and annihilation operators. A creation operator, $\hat{a}^\dagger$, adds a quantum of energy, and increases the occupation number for a given oscillator. The annihilation operator $\hat{a}$ removes a quantum of energy and decreases the occupation number. Since our phonon problem is just a harmonic oscillator problem, we can introduce creation and annihilation operators for each vibrational mode picked out by a choice of $k$ and $j$:

$$q_{kj} = \sqrt{\frac{\hbar}{2\omega_{kj}}} (\hat{a}^\dagger_{-kj} + a_{kj})$$

$$p_{kj} = i \sqrt{\frac{\hbar \omega_{kj}}{2}} (\hat{a}^\dagger_{kj} - a_{kj})$$

We can now think of phonons as quanta of vibration with energy $\hbar \omega$, and of states of the crystal as characterised by phonon number state $|n_{kj}\rangle$, which is
raised and lowered by the creation and annihilation operators.\(^3\) The parallels with the physics of photons will now be obvious to some readers; the above equations could have been written for photons rather than phonons. And indeed, once we reintroduce anharmonic terms, the mathematics of phonon-phonon interaction is represented by diagrams which look remarkably particle-like:

![Figure 3: Diagrams for three phonon processes: for example the top-left diagram represents 
\(\hat{a}_{-k_1j_1} \hat{a}_{k_2j_2} \hat{a}_{k_3j_3} \).](image)

2 Phonons and Limits

Various authors have seen phonons as emergent [34, 36, 29, 16, 27]. In section 2.2, we’ll defend this claim more explicitly, but for the moment let’s simply note that the claim seems relatively natural in the current context; we started with an atomic lattice description, and moved to a description of particles which could be thought of as interacting, both with other phonons and with other particles such as neutrons and photons. The phonon description certainly seems novel, and it’s robust insofar as it survives relaxations of the harmonic approximation.

Section 2.1 looks at whether this putative emergence can be captured by the asymptotic limits account, and argues that it cannot. But section 2.2 argues

\[^3\]|n_{kj}\rangle = \frac{1}{\sqrt{n_{kj}!}} (\hat{a}_{kj}^\dagger)^{n_{kj}} |0\rangle.

9
that there are good reasons to consider phonons to be emergent nonetheless. Finally, in section 3, we’ll argue that the novelty exhibited by the phonon description is best thought of as explanatory novelty.

2.1 Asymptotic Limits

What relation could the example of phonons possibly have to the asymptotic limits literature? On the face of it, our derivation of the phonon picture involved the taking of no limits, let alone singular ones. But it did involve a series of approximations or idealizations: the harmonic approximation, the imposition of a periodic boundary condition, and the adiabatic approximation. Given that the asymptotic limits analysis is related to discussions of essential idealizations, it’s worth taking a look at the approximations used in the derivation of the phonon description.

We start with a very basic presentation of what has become the paradigm example of an asymptotic limit that signifies emergence: the use of the thermodynamic limit in connecting a thermodynamic description of phase transitions with a statistical mechanical account. Thermodynamics tells us that phase transitions occur when there is a discontinuity in a derivative of the free energy. The statistical mechanical free energy function is analytic and can have no such discontinuity. One solves the problem by taking the thermodynamic limit of the free energy function – the limit of this function as the number of particles tends to infinity. Real systems, of course, are composed of finitely many particles. There is a reasonable philosophical consensus [7, 9, 13, 14] that phase transitions are novel phenomena of the kind required for emergence (where the hallmark of emergence is novelty coupled with robustness). And many commentators locate the source of the novelty in the use of the thermodynamic limit.

Batterman thinks that the need for the limit shows that finite statistical mechanics cannot account for the phase transition. He claims that the limit is required in order to describe the emergent behaviour corresponding to the phase transition: that is, one may locate the phase transition only after the limit has been taken. Butterfield thinks that the limit flags a novel behaviour that in fact emerges, albeit weakly, in finite systems before the limit. But both agree that novelty springs from two related features: that limits are used in deriving phase transitions, and that the relevant limit is singular. All agree that it is crucial that there is a contrast between behaviour in large but finite systems and infinite systems; a large gradient in the derivative of the free energy is not the same thing as an actual non-analyticity.

Perhaps basing our asymptotic account on this one example is unfair. There is a broader but connected literature on essential idealizations that might provide the basis for something like the asymptotic limits account. The essential
idealizations literature stresses that the representation of some phenomena requires the incorporation of false assumptions into our physical models; such assumptions cannot be de-idealised in the sense discussed in McMullin [28]. There has been much discussion of phase transitions and essential idealizations, but no explicit connection of essential idealizations with the novelty required for emergence, although the work of Sorin Bangu (e.g. [3]) seems sometimes to hint at this. In fact, we find it difficult to view anything in the essential idealizations literature as an account of emergence – essential idealizations involve representational failures exhibited by our models; this does not look like a good recipe for novelty.

Nonetheless, for the sake of completeness here, we’ll consider the idea that novelty somehow arises whenever the derivation of a phenomenon from the lower level involves an essential idealization, that is, an idealization that cannot be de-idealized without losing the phenomenon in question. We’ll see that even if essential idealizations did lead to novelty, the derivation of the phonon description does not involve such idealizations.

Much subtlety has been glossed over in the above, and there is no univocal account in the literature of exactly how to characterise the purported novelty of phenomena like phase transitions. But we can use this rough sketch as a comparator for our phonon physics: can anything like the accounts above capture the sense in which phonon behaviour is novel? Do any of the approximations required for the derivation either look like essential idealizations or involve singular limits?

In order to evaluate this we should first make clear the distinction between approximation and idealization. Here, we will follow John Norton [30] in defining approximation as involving an “inexact description of a target system” and idealization as involving reference to “another system whose properties provide an inexact description of the target system”.4

Let us move on to our evaluation of the approximations involved in the derivation of the phonon description. Might some of these involve singular limits, or be idealizations in disguise? We start with the harmonic approximation. Recall that this approximation depends on assuming that the displacement of the atoms in the lattice is very small in comparison to the lattice spacing: \( \langle u^2 \rangle \ll a^2 \) where \( a \) is the lattice spacing and \( u \) is the atomic displacement.5 This breaks down as temperatures approach the melting temperature of the solid. Once the harmonic approximation has been made, and we have transformed to a set of decoupled equations in phonon variables, so-called anharmonic terms may be re-introduced – these are crucial for the modelling of

4A third category, abstraction, which has sometimes been called “Aristotelian idealization” [28], will be relevant in section 3. This involves the stripping away, or throwing out, of features in our explanations and representations.

5There are systems for which this assumption doesn’t lead to the harmonic approximation: for example, quantum hydrogen and helium crystals.
phonon-phonon interactions and in order for phonons to have finite mean free
paths, these thus play crucial explanatory roles. Such anharmonicities may be
expressed as terms of the adiabatic potential beyond second order in the dis-
placement variable expansion.

One could scarcely find a more standard style of approximation in physics.
And on the face of it, there are no limits applied at all. But might one be able to
think of the approximation as a limit of some sort nonetheless? It’s hard to see
how. The value of this approximation relies on not taking the limit of the ratio
that we take to be small; in the limit, there would be no reason to reintroduce
higher order terms, and most equations would reduce to triviality. And even if
there is some way of shoehorning this kind of approximation into limits talk,
there are two reasons why the friend of singular limits should not go down this
route. For one thing, there’s no reason to think there are singular limits around,
and most commenters agree that it is the singularity of the thermodynamic
limit that is relevant to novelty. And for another, these kinds of approximation
are endemic in physics – if these lead to emergence, then emergence would be
so ubiquitous as to be utterly un-interesting.

What of idealization, essential or otherwise? The case above seems to be a
classic case of approximation. The description depends on a ratio being small,
and thus certain higher order terms being negligible. These things are literally
true of the target system; no further idealized system is required. The point
at which approximation comes in is the point at which we set higher order
terms to zero; this, of course, is not literally true. But all reference is to the
original system, hence, in Norton’s terms, we are discussing approximation,
and a humdrum approximation at that. It’s worth emphasising that while this
approximation is central to standard derivations of the phonon description,
limited de-approximation is essential to recover phonons with realistic proper-
ties. Thus it would be a mistake to regard the harmonic approximation as an
essential idealization.

Next, consider the periodic boundary condition: imposing a boundary condi-
tion like this involves no limits, singular or otherwise. Might it be an ideal-
ization? Perhaps. But it is certainly not an essential one. There is no particular
size of supercell at which we must set the periodic boundary; pragmatic con-
siderations are the only issue here. We can set the boundaries as wide as we
like – indeed, they could be of the order of the size of the crystal. There’s no
question that normal modes can be defined for large boundaries.

However, the imposition of periodic boundary conditions does involve an
idealization of sorts, because it allows us to deal with a model of the crys-
tal without edges which might have otherwise affected the description of the
phonons. This idealization is related to the more basic assumption that each
cell of the crystal has identical properties – this is not true of any real crystals.
Essential idealization claims in this context can be dispelled by observing that
we can, here, de-idealize: small defects, imperfections and edge effects can
be introduced without undermining the physical picture developed above.\textsuperscript{6} Large defects or widespread imperfections may lead to modelling failures, but the behaviour of the real systems thus represented would deviate from standard phonon behaviour and, as such, the assumption of no large defects is justified.

The adiabatic approximation is also invoked in the derivation of the phonon description. Recall that this means that the electron states and nuclear motion dynamically decouple. This allows us to start from the assumption that the total system wavefunction is the product of the nuclear and electronic wavefunctions, to solve for this assumption, and to reintroduce interaction effects into the potential perturbatively. But it is an approximation in just the same sense as the harmonic approximation. Again, if limits were introduced to model this approximation they would be both gerrymandered and non-singular, and no idealization is involved.

2.2 Are Phonons Really Emergent?

So the phonon example does not fall neatly under an asymptotic limit analysis of emergence. Should phonons nevertheless be considered emergent? We think so, at least if one is amenable to the notion of emergence more generally. Granted, there are those who will define emergence in such a way that phonons do not fall under the definition; any definition (for example Batterman’s) that precludes the compatibility of emergence and reduction will do this. But, if we stick with our original analysis of emergence as novel and robust behaviour, there is reason to think that phonons are not merely a phenomenon that has been called emergent, but rather a phenomenon that stands to be particularly revealing as to the nature of the novelty involved in emergence. In section 3 we articulate the kind of novelty that phonons display. But, first, in this section, we defend the claim that phonons deserve to be called ‘emergent’, even before we give an analysis of their novelty.

The relation that phonons hold to the underlying crystal description is almost identical to the relation that quantum particles hold to the underlying quantum field. And if any inter-theoretic relation betokens an interesting emergence, surely the relationship between quantum particles and the field does.

Phonons are bosons and behave as such. The superficial resemblance between the physics of phonons and the physics of photons is striking (though many other bosons would allow for an equally striking analogy). Both obey Bose-Einstein statistics. Both contribute quanta of energy $\hbar \omega$, and the energy of

\textsuperscript{6}[22, §1.5.2] is a brief taxonomy of crystal defects, Kantorovich later discusses how crystal defects can be taken into account as anharmonicities, and can otherwise lead to shorter mean free paths of phonons. In the limit, defects will lead to an amorphous (non-crystalline) solid; see [33] for a discussion of techniques for modelling phonons in amorphous solids.
the quantum system depends on the number of photons or phonons. One can even build a saser (sound amplification by stimulated emission of radiation) much like a laser. But emergence is a relational property – phonons, if they are emergent, are emergent by virtue of the relation between their physics and that of the underlying crystal. It’s therefore much more compelling to look at the way in which we derive a photon description from the electromagnetic field in the standard model, and compare this to the way in which we derive the phonon description. And here the similarities are much more than skin-deep.

It is no coincidence that quantum field theory (QFT) textbooks\textsuperscript{7} sometimes begin with solid state physics and the physics of phonons. The process by which we derive a particle description from a quantum field is much like the one described in section 1.2. In particular, one looks for a normal modes description that allows one to describe the system in terms of independent harmonic oscillators, and then derives a description with creation and annihilation operators, and particle occupation numbers.

If one wants to deny that phonons are emergent, one must either deny that the two cases are appropriately analogous, or deny that particles emerge from quantum field theory. Start with the first option: how might one challenge the analogy? At a formal level, it is hard to drive a wedge between, say, phonons and photons: the mathematics involved in both is strikingly similar, as is the derivation if we simplify our cases a little, and compare phonons in a monatomic crystal with particle states in massless scalar field theory.\textsuperscript{8}

A challenge to the analogy will therefore need to come from elsewhere; and if the inter-theoretic relation itself is analogous, any difference must somehow lie in the nature of the phonons, photons, crystals and quantum fields themselves. Doubtless some readers’ intuitions will push them to deny that phonons should be taken as ontologically seriously as quantum particles. But such attempts to block the analogy clearly beg the question; our analysis here aims to demonstrate that phonons are a novel and robust phenomenon, and thus, by the lights of the emergence programme, belong in the ontology at relevant energies and length scales.

So we should look to the more fundamental level of description for any prospective disanalogy. The most striking difference between the phonon case and the case of quantum particles lies not in the description of the particles, but in our access to the underlying field or lattice from which they are derived. In the phonon case we have direct access to the physics of the crystal – we can experimentally confirm the lattice physics itself. In the case of the quantum field, our experimental evidence is often mediated via the particle description; our major sources of evidence come from, for example, particle accelerators.\textsuperscript{9}

\textsuperscript{7}E.g. [26].
\textsuperscript{8}See [34, p.7].
\textsuperscript{9}One could deny even this distinction and claim that observation of, say, the arrangement of iron filings around a bar magnet is fairly direct experimental evidence for the underlying quantum
It is, however, difficult to see how one might move from this legitimate disanalogy between phonons and quantum particles to an argument against the emergence of phonons. Worries about the epistemic status of the quantum field could shake our faith that there is a quantum field for particles to emerge from; as a consequence one might become wary of an emergentist account of quantum particles. But even if we grant such concerns, one could still mount a case that phonons are an emergent phenomenon, and one might likewise grant that if the status of the quantum field were settled satisfactorily, quantum particles would look equally novel and robust with respect to that field.

The above style of argument does not, therefore, undermine the analogy itself. But it might threaten our argument that we should accept phonons as emergent because quantum particles clearly are. Some interpreters of quantum field theory would deny that particles emerge from QFT at all. These deniers could fall into two camps. The first camp (exemplified by Doreen Fraser [19]), holds that there are no particle states in realistic quantum field theories. The second camp contains those who hold that particles, rather than fields, are fundamental to quantum field theory.

Doreen Fraser has argued against the existence of exact particle-states, and hence by her lights, the existence of particles in quantum field theory. Her 2008 paper argues that the kind of particle states derivable from non-interacting fields are not precisely definable in interacting systems. But a lack of precise definability of particle states is no obstacle to viewing particles as emergent; it is characteristic of emergent phenomena that they involve approximation and idealization. Fraser implicitly acknowledges this compatibility when discussing Wallace’s [34] emergentism:

For this to be a viable response, the cogency of the distinction between fundamental and less fundamental entities must be defended and a case must be made for admitting additional, non-fundamental entities into our ontology. [19, p.858]

Although we have not much discussed the distinction between fundamental and non-fundamental entities here (though we mention it again at the end of section 3), we take it that the kind of emergence literature we are building on does take the distinction to be cogent. And the account of novelty we develop here and elsewhere is, in part, intended to form part of a case for admitting non-fundamental entities into our ontology.

We’ve so far argued that once one accepts that QFT is a theory of quantum fields, an account of emergent particles is not far behind. But must one accept this? Some sources (for example [25]) discuss ‘particle interpretations’ of field. Indeed, as [15] points out, Fraser and Wallace are really in agreement about the non-fundamentality of particles.
quantum field theory. Such a view would hold that particles are in fact fundamental to QFT. Even in the absence of a settled definition of emergence, we can agree that if particles are fundamental to a theory they are not emergent with respect to that theory! But it’s hard to find a worked out version of such a view in the contemporary literature, although Paul Teller [32] is sometimes held up as an example. There are good reasons why particle interpretation defenders are hard to find; Doreen Fraser’s arguments, mentioned above, preclude exact particle states for interacting systems, and the presence of interesting vacuum state physics (when particle number is zero), also speaks against QFT being fundamentally about particles. Credible defenders of the existence of particles in QFT (e.g. [34] and [35]) think of them in precisely the emergent way that we advocate here.

We therefore stand by our claim that particles in QFT are emergent if anything is. Granted, one must interpret QFT as being fundamentally about a field, and one must meet Fraser’s challenge by putting some flesh on the bones of a general account of emergence. Our main concern here is not to defend the thesis of emergentism in general, but rather to counter certain assumptions about emergence among the emergentists. Once one is working within such a framework, it is natural to say that particles emerge from quantum field theory, and, as a result of the analogy, to claim that phonons emerge from the crystal lattice.

3 Emergence as robustness with novel explanatory value

So phonons seem to be well described as an emergent phenomenon, and hence a good candidate for demonstrating novel and robust behaviour. What can they teach us about the right way to analyse robustness and novelty?

3.1 Robustness

Robustness is important, but, unlike in the case of novelty, the ingredients of an analysis have been provided in the literature. To show that a phenomenon is robust is to show that it survives perturbations of the underlying physics; in order to be emergent, a phenomenon must not be too fragile or too fleeting. The robustness we are concerned with here is robustness as it applies to an actual concrete system: might perturbations of the physics of this crystal arising from, say, small temperature changes, displacements, or knocks, destroy the phenomenon of interest? If the answer is ‘yes’ for ordinary small changes to the system, then the phenomenon would be insufficiently robust to form an interesting part of higher level physics.
Phonons are robust in this sense. They exist over long periods of time in real crystals under different conditions; the nature and use of approximations in their derivation gives us clues as to the source of their stability. For one thing, the harmonic approximation holds in a wide range of conditions; Debye temperatures are 200 – 500 K for many common elements, and are much higher in some cases. Below these temperatures, we can apply the low temperature Debye model in which the harmonic approximation holds and phonons do not interact. But the phonon description survives even when we relax the harmonic approximation; at higher temperatures, we reintroduce anharmonicitites and these lead to phonon-phonon interaction. Phonons remain a useful description at very high temperatures – indeed, they can be used to describe the system more or less as long as the system remains crystalline.

A great deal of work in the philosophy of physics goes into particular robustness demonstrations; both Butterfield’s [13] and parts of Batterman’s work (notably his discussion of the rainbow in his [6]) can be seen as providing such demonstrations. In fact, we think that the importance of the asymptotic limits analysis lies much more in its relation to robustness than its connection to novelty. Consider, for example, the use of the thermodynamic limit in explaining a phenomenon like a phase transition. If we think that the use of the thermodynamic limit demonstrates novelty, we fall on the horns of the dilemma mentioned on p.2: either the use of the limit indicates a failure of reduction, in which case the phase transition is novel but the novelty is mysterious, or use of the limit is reducible and it is hard to see where the novelty lies. By contrast, the use of the thermodynamic limit, properly justified, can easily be seen as a demonstration of the robustness of phase transition behaviour.\(^\text{11}\) If one demonstrates the appropriateness of the thermodynamic limit, in which the number of particles, \(N\), goes to infinity, then one has demonstrated that the exact number of particles doesn’t matter, so long as it is large. That is, one has demonstrated that phase transitions are robust under changes to the size of the system (as well as many other changes). Of course, there is much work to be done in justifying the use of the limit. One can see the work of Jeremy Butterfield [13] as providing such a justification and hence giving an argument for the robustness of phase transition phenomena.

One must be careful when discussing robustness, for it is easy to conflate two related ideas connected to perturbations of the underlying physics.\(^\text{12}\) On the one hand, we can consider varying the exact conditions of a particular system – i.e. changes which the system might actually undergo. On the other, we can consider varying the very nature and make-up of the system itself – these are counterfactual or imagined changes which bring out features such as common structure. There has been considerable interest in the invariance of higher level theories under changes of the second kind; it goes under the title of ‘uni-

\(^{11}\)This point was originally made to one of us by Adam Caulton; this is also a claim with which we believe Batterman would agree.

\(^{12}\)Note that a similar distinction is made in [20] between actual and counterfactual stability.
versality’ or ‘multiple realizability’. It is demonstrated, for example, by phase transition behaviour, which is similar in ferromagnets and fluids; see [18, 10]. But we do not think that this kind of behaviour is an essential trait of emergence, despite the historical discussion of ‘multiple realizability’. For suppose some novel phenomenon is only demonstrated in a specific class of materials – particular kinds of metal, for example. Now suppose that the phenomenon persists in a range of circumstances – it survives many knocks, movements and environmental changes. Is the restriction to metals really relevant to its emergent status? The literature on multiple realizability suggests that it is, but it’s also famously difficult to understand exactly why this should be the case. By contrast, the robustness discussed above is obviously relevant to emergence, precisely because it’s related to the usefulness, lifetime, and observability of the phenomenon in question.

Why then, is multiple realizability often focussed on in lieu of robustness? The confusion turns out to be understandable; the features that lead to robustness often lead to multiple realizability as well. In order to demonstrate that some phenomenon is immune to changes in a given system, one will often, but not always, highlight features of the phenomenon that render it immune to changes of the material make-up of a system. The features that make phase transitions insensitive to particle number may help to make them insensitive to other details as well. But the connection is not guaranteed. Although phonon behaviour is demonstrated in a wide range of kinds of crystal, and is thus stable under counterfactual or imagined changes, we claim that phonon phenomena would count as emergent even if phonons were only realised in, say, aluminium lattices.

In addition, multiple realisability has repeatedly (although in our view mistakenly – see for example, [37]) featured in arguments against reduction; see e.g. [17]. In this paper we defend the view that emergent phenomena are best understood as novel and robust, but a well-established tradition in philosophy regards claims to emergence as shorthand for the denial of reduction. As such, multiple realisability is seen to serve as an argument for emergence insofar as it undermines reductionist claims. Thus multiple realisability and emergence have often been linked despite the fact that this link detracts from a clear understanding of both concepts.

3.2 Novelty

The sense in which phonons are novel requires more explanation. We have established that, if phonons are to be considered novel, their novelty does not seem to come via the use of approximations and idealizations involving singular limits. But, to our minds, they do fall very nicely under an alternative

---

13Novelty might otherwise be associated with unpredictability; in-principle unpredictability would preclude reduction and thus be inappropriate to the phonons case. Bedau, in e.g. [11], has
account of novelty. [24] argues that novelty might be analysed as explanatory novelty, and that this explanatory novelty might come about as the result of particular kinds of changes of variable or quantity between theoretical levels. In this section, we’ll demonstrate how the phonon example fits into this account.

Let us start with the simple normal modes example of section 1.1. In this case, we noted that not only did a move to normal mode variables help us to solve the equations of motion, but that it might also help with explanatory goals. It is quite obviously possible for a phenomenon to be linked to the normal mode variables rather than the displacement variables; our (admittedly rather contrived) example involved a light that flashed when the central spring was compressed. The most concise explanation of this phenomenon appeals only to the normal mode variable $\eta_2$ and equation (5). That is, after the variable change we can make a novel abstraction whereby the explanation does not refer to $\eta_1$. While the phenomenon is explicable in the original displacement variables, both such variables must be appealed to in order to provide the explanation. As long as we think that good abstractions lead to better explanations, the normal mode variables offer a better, more powerful explanation by virtue of allowing an explanatory abstraction.14

It’s this kind of abstraction that we think is the key to the novelty displayed by phonons. But the example above might leave one unconvinced, for our normal mode variable is a function of just two displacement variables, and the explanation is not terribly complex when translated back into the original variables. One might very well be able to understand the value of an explanation that appeals to an equation that calculated $x_1 - x_2$ without actually changing variables. So the sense in which our abstractive explanation is novel may seem very weak indeed.

The case is different when we move to a more complex example. Just as in the normal modes example above, we can highlight just a small number of phonon variables in the explanation of a phenomenon, abstracting away from other phonon variables. Our phonon variables are linear combinations of a vast number of atomic displacement variables. So one can imagine that it will be much harder to ‘see’ the abstraction offered by the move to phonons in the physics as described by the displacement variables before the variable change.
The introduction of phonon variables allows for information to be aggregated in new ways and facilitates the selection of a few salient variables to explain a given phenomenon. Phonons are not merely *calculationally* powerful but also *explanatorily* powerful, at least if we agree that the right abstraction enhances explanatory power.

However, the use of the harmonic approximation and the number of variables involved mean that the value of an explanation which appeals only to a small number of phonon variables may be completely obscure from the perspective of the displacement physics. Indeed, from the perspective of the displacement physics, the ‘abstraction’, with its vast number of variables, will not look like an abstraction at all. It is in this sense that we think that the phonon description has novel explanatory power.

But this novel explanatory power does not indicate any kind of failure of theoretical reduction. We know perfectly well what function of the displacement variables leads to the phonon variables, and we can back-translate the physics into our original description. But that does not mean we can understand *why* the description thus obtained is explanatorily useful without appealing to the phonon description. Even though the phonon description is translatable into the displacement variables, its explanatory value is invisible unless we choose the right class of variables. One might think that the explanatory value of two mathematically intertranslatable sets of variables would be equivalent; however, consider the counterfactual account of explanation: it’s apparent that different variable choices would invoke different sets of counterfactuals and thus different explanations would be available.

This account of novel explanatory power follows that given in [24], but that account focused on the relationship between thermodynamics and statistical mechanics, where the mathematical relationships between quantities (the bridge laws) involve summations, coarse-grainings, and limiting procedures. These relationships mean that the thermodynamic description possesses less information than the statistical mechanical one – details are washed out, or abstracted away from, when we move between descriptive levels. So even if we have a good reduction in the sense that we can derive the thermodynamic variables from the statistical mechanical ones, we won’t be able to recover the exact statistical mechanical description from the thermodynamic one. [24] proposed that this “mathematical irreversibility” was important to novelty.

The phonon example is importantly different. The phonon variables, \( q_{kj} \), are just linear functions of the displacement coordinates \( u_{Ls\alpha} \). And, of course, the latter are also linear combinations of the former. So a complete translation between descriptions is possible in either direction. In this sense, the two descriptions seem to express a *duality*, rather than a standard reductive relationship. This leads to a question that has been pressed on us by David Wallace: if the relationship here is really one of duality, can one nonetheless talk about novelty and emergence?
We think (contra [24]) that the answer, at least to the question with regards to novelty, is yes: explanatory novelty can be displayed even when the descriptive change is entirely reversible. The phonon case demonstrates this. But emergence is plausibly a relation that is, by definition, asymmetric; one cannot both think of phonons as emerging from the crystal lattice and of the crystal lattice as emerging from the phonons. This sounds right to us, and suggests that mere robustness and novelty may not be enough for emergence. We thus may wish to define emergence as a relation that holds between less and more fundamental phenomena.15

4 Conclusion

This paper has aimed to show that a physical phenomenon can be emergent despite its description not requiring appeal to an asymptotic limit. Phonon behaviour is novel and robust, but the phonon variables are not derived by any limiting procedure, and are, in fact, exactly translatable into the displacement variables. Nonetheless, we think that the phonon description allows for explanatory abstractions that are not available in the underlying description, and therefore involves novel explanations. Insofar as they are also robust under perturbations of the underlying crystal physics, and are described by a theory that is less fundamental than the basic theory of the crystal, they are emergent.

Thus we draw the broader conclusion that, while asymptotic limits play a role in robustness analysis, pointing to such limits does not help explain the novelty of emergent phenomena. Moreover the example of phonons illustrates that recognisably emergent phenomena neither require limits nor essential idealisation to figure in their derivation from the more fundamental description.

It’s worth, by way of summing up, considering what kind of account of emergence is on offer in this paper. Strong emergentists often claim that emergent phenomena are in-principle irreducible, and, as such, their account is rightly regarded as metaphysical; while weak emergentists may claim in-practice irreducibility and thus a merely epistemological account. Our account, by contrast, holds that emergence and reduction are compatible and features both metaphysical and epistemological aspects. On the one hand, one reason to accept phonons as emergent is that phonon variables are useful and allow us to understand phenomena that would be incomprehensible were we forced only to appeal to displacement variables. On the other hand, the fact that phonons allow for such good explanations tells us something about the world – phonon variables pick up on real dependency relations. Despite the fact that we can translate the phonon description directly into an account in terms of atomic displacements, phonons are emergent both because they allow for otherwise

15Of course, this leaves open the question of defining a fundamentality relation sufficient to provide the requisite directionality to emergence, but that project is beyond the scope of this paper.
unavailable explanations and because, in some sense, they’re really out there.

Acknowledgements

We’d like to thank audiences at the Oxford Philosophy of Physics seminar, BSPS 2016 and ‘Emergence and the Limit’ at the LSE for their comments and suggestions on earlier versions of this paper, and David Wallace for helpful conversations. A.F. is supported by an LAHP studentship.

References


