Taking Reductionism to the Limit: How to Rebut the Anti-reductionist Argument from Infinite Limits

Juha Saatsi and Alexander Reutlinger*†

October 6, 2017

Abstract

This paper analyses the anti-reductionist argument from renormalisation group explanations of universality, and shows how it can be rebutted if one assumes that the explanation in question is captured by the counterfactual dependence account of explanation.

1 Introduction: The anti-reductionist challenge

Statistical and condensed matter physics have always been a rich source for anti-reductionist arguments. One prominent anti-reductionist argument turns on the theoretical role played by infinite limits, such as the thermodynamic and continuum limits, in the context of renormalization group (RG) explanations. Anti-reductionists have appealed to RG explanations of, e.g., the occurrence of phase transitions and the universality of critical exponents. They have argued that the indispensable explanatory usage of continuum limits in these explanations speaks against reductionism, because such usage reveals a significant limitation of a more fundamental reductive theory that describes the atomic constituents of ultimately finite (albeit micro-physically absolutely huge) bits of matter. (Batterman 2000,

*To contact the authors please write to: Juha Saatsi, Woodhouse Lane, Leeds, LS2 9JT, United Kingdom; e-mail: J.T.Saatsi@leeds.ac.uk. Alexander Reutlinger, Munich Center For Mathematical Philosophy, Ludwig-Maximilians-Universitat Munchen, Ludwigstr. 31, 80539 Munchen, Germany.

†We benefitted from comments on earlier drafts of this work presented at University of Guelph and at workshops at LMU and Leeds, and especially from the extensive and very helpful comments by the two anonymous referees. The support of the Durham Emergence Project and John Templeton Foundation is gratefully acknowledged.

2002a, 2010, 2011, 2015; Morrison 2012, 2015) In particular, anti-reductionists have pointed to the indispensable explanatory role of *the fixed points* of renormalization group transformations. These fixed points presuppose limit assumptions that are arguably in tension with reductionism. (Menon and Callender (2013, 197) nicely summarize : “The fixed point only appears when the system has no characteristic length scale. This is why the infinite particle limit is crucial for the renormalization group approach.") Thus, taken at face value, the fixed points (incorporating such limit assumptions) contradict the finiteness of the physical systems exhibiting universality – a finiteness that is assumed by the fundamental physical theory to which reductionists allude. The upshot is that the explanatory indispensability of these fixed points is thus seen to reveal a philosophically significant limitation of a more fundamental theory.

Let us call the anti-reductionist argument sketched above “the argument from infinite limits". This argument can be seen to underwrite a specific challenge for the reductionists:

Anti-reductionist Challenge: The reductionists ought to show how the fixed points involved in RG explanations of critical phenomena can be (a) explanatorily indispensable and, at the same time, (b) compatible with reductionism.

Our main goal in this paper is to show how a reductionist can meet this challenge. We will assume for the sake of the argument that fixed points (and the presupposed limit assumptions) are indeed indispensable for RG explanations of universality. By making this assumption we will try to make the strongest possible case for anti-reductionism. Notwithstanding this assumption, we will argue that the supposed indispensability does not lead to any ontological commitments threatening reductionism. We will do so by arguing that a particular account of explanation – the counterfactual dependence account – captures the explanatory character of RG explanations, and that in the light of this understanding of RG explanations the indispensability of fixed points is not ontologically committing.

Our response to the anti-reductionist challenge addresses two clear lacunae with regard to the recent debate.

The first lacuna is that the reductionist analyses of the infinite limits (including the fixed-points of RG transformations) have focused on explanations of the *occurrence of phase transitions* (see for instance, Earman 2004; Butterfield 2011a, 2011b; Norton 2012; Menon and Callender 2013). These analyses do not address one of the key points emphasised by the anti-reductionists: namely, the idea that reductionists cannot capture RG explanations of *universality* of macro-behaviour of physical systems undergoing (second-order) phase-transitions. (Batterman 2015, for instance, points this out forcefully.)[1] To address the first lacuna, we focus on meeting the anti-reductionist challenge with respect to the RG explanation of universality.

The second lacuna concerns the fact that relatively little work has been done to explicate the *explanatory character* of RG explanations of universality. On the one hand, the reductionists' focus has been on inter-theoretic reduction relations (mostly framed as neo-Nagelian reduction). RG explanations, insofar as these have been discussed at all, have been portrayed – without much of an argument – in terms of the deductive-nomological (DN) account of explanations, as befitting a Nagelian approach to reduction (Butterfield 2011, Norton 2012). Portraying RG explanations as exemplifying the DN-model is a controversial and somewhat surprising claim, since most philosophers of science today agree that the DN account of scientific explanation is deeply problematic. So regarding it as an adequate explication of a particular scientific explanation requires a good rationale.

On the other hand, although the anti-reductionists do not advocate the covering law account, they have not provided a convincing philosophical account of RG explanations either. Most prominently, Batterman has advocated a 'minimal models' account according to which RG explanations are explanatory (roughly) by virtue of showing that the explanandum is completely *independent* of all micro-details (Batterman 2000, 2002a, 2002b; Batterman and Rice 2014). However, this approach faces serious objections (that are independent of the problems of the DN-model) and we do not regard it as convincing (see Lange 2015; Reutlinger 2017a; see also Jansson and Saatsi forthcoming).

In effort to address the second lacuna, we will explicate the explanatory character of the RG explanation of universality in relation to the counterfactual dependence account of explanation.

The plan of the paper is as follows: Section 2 reviews the (anti-)reductionism debate surrounding RG fixed points. In particular, we highlight a connection between the anti-reductionist argument from infinite limits and explanatory indispensability arguments. In Section 3, we clarify the *explanandum* at stake, emphasising the theoretical context of physics of critical phenomena that preceded the

[1]Exceptions: Reutlinger (2016, 2017a,b) and Hüttemann et al. (2015) address the explanandum of universality. The present paper builds on this line of work.

RG analysis. In Section 4, we provide a detailed exposition of the relevant physics of RG *explanans* and the role of fixed points (§4.1), and we argue that the counterfactual dependence account of scientific explanations captures RG explanations of universality (§4.2). In Section 5, we respond to the anti-reductionist challenge on the basis of the results of Section 4: we argue that the explanatory appeal to fixed points and limits is (merely) instrumentally indispensable and, therefore, it does not lead to an ontological commitment threatening reductionism. In the end we also discuss the limitations of our argument vis-à-vis explanatory (as opposed to ontological) anti-reductionism.

2 Anti-reductionism and Explanatory Indispensability

Let us now examine the structure of the argument from infinite limits more closely.

The anti-reductionists suggest that RG explanations undermine reductionism due to commitment to the fixed points of renormalization group transformations (and the presupposed limit assumptions). These are often said to 'control' critical phenomena, the universality of which is said to 'rely on the existence' of fixed points. The following passage from a leading text book is typical:

> Since [the behaviour of correlation functions] depends only on the fixed-point Hamiltonian, the correlation functions corresponding to all Hamiltonians that converge after RG transformations toward the same fixed point, have the same critical behaviour. Such a universality property thus divides the space of Hamiltonians into universality classes. Universality, beyond the quasi-Gaussian approximation, *relies on the existence* of large-distance (IR) fixed points of the RG in the space of Hamiltonians. (Zinn-Justin 2007, 226. Our emphasis).

Indeed, the RG framework is, in a sense, *all about* the fixed points of RG transformations: their properties, their classification, and the conditions under which they exist. In as far as this framework furnishes genuine explanations of critical phenomena that turns on these fixed points and their properties, and insofar as these fixed points involve limits of modelling parameters that transcend the finitude associated with the more fundamental theory, there is a clear *prima facie* challenge to reductionism here.

Batterman argues in this spirit that 'there are very good reasons to deny that [critical] phenomena are reducible to "fundamental" theory' (2011, 1034).[2]

> The renormalization group explains the universal behavior at criticality essentially by exploiting the divergence (blow up to infinity) of the correlation length. [...] Most crucial to the renormalization group explanation is, as noted, the ineliminable appeal to the thermodynamic limit and to the singularities that emerge in that limit. (2011, 1043)

The sense of anti-reductionism that Batterman supports by pointing to the explanatory indispensability of the fixed points primarily concerns explanation, not ontology. (The 'fundamental' theory, Batterman explicitly says, "gets the ontology of blobs of gases and fluids right" (2011, 1034).)

The step from 'explanatory indispensability' to ontology is relatively short, however, and various philosophers are willing to take it. This willingness can be rooted, in general, in a venerable tradition in the philosophy of science that associates scientific realists' ontological commitments tightly to explanatory indispensability. Thus, Psillos (2011), for instance, follows Sellars (1963) in adopting an 'explanatory criterion of reality', according to which "something is real if its positing plays an indispensable role in the explanation of well-founded phenomena." (Psillos, 2011, p. 15) More generally, philosophers in the Quinean and Putnamian tradition have argued for realist commitment to mathematical and other abstracta on the basis of their explanatory indispensability to our best theories of empirical phenomena (see e.g. Baker and Colyvan 2011).

Morrison (2012, 2015) has defended ontological anti-reductionism in this spirit – both in general, and with respect to RG fixed-points especially. Regarding 'the explanatory power of fixed points', in particular, Morrison argues that the reductionists "ignore a crucial feature of emergence, specifically the ability to properly explain universal behaviour and [...] the role of RG in that context." (2015,110) Namely:

> The calculation of values for critical indices and the cooperative behaviour defined in terms of fixed points is the foundation of universality. RG is the only means possible for explaining that behaviour;

[2] Batterman puts the term 'fundamental' in scare quotes, because he regards it as ambiguous: "a theory may be fundamental in that it properly characterizes the detailed constitution of the systems it studies, but can fail to be fundamental in its ability to explain and provide understanding of the systems it correctly describes." (Ibid.)

> what happens at finite N is, in many ways, irrelevant. Finite systems can be near the fixed point in the RG space and linearization around a fixed point will certainly tell you about finite systems, but the fixed point itself requires the limit. (Ibid.).[3]

Turning now to the reductionists, the indispensability of RG fixed points is not by any means denied by them; rather, the issue concerns their status as *(non-)representational elements* of the less fundamental theory. Thus, Norton (2012, 222), for example, characterizes them as "points in a diagram: mathematical pegs on which to hang limit properties." Norton draws an apt distinction between a meaningful and well-defined limit of a sequence of *systems*, on the one hand, and a limit of a sequence of *properties*, on the other. The crux of the distinction is that the latter may not correspond to any possible system, in which case it cannot function as an 'idealisation', but at best as a useful 'approximation'. According to Norton, RG fixed points are such approximations, for they "do not arise from an investigation of the properties of infinite limit systems. They are not idealizations" (Ibid.).

However, regardless of its status as an 'idealisation' or 'approximation', an anti-reductionist may respond that the very fact that a mathematical limit that plays an indispensable explanatory role is still puzzling from the point of view of the more fundamental theory. *Why* is the use of such limits indispensable for explaining the phenomena? *Why* doesn't the indispensability of such limits indicate a feature of critical phenomena that transcends the ontology of the more fundamental theory?

Moreover, problems arise for Norton regarding his construal of an RG explanation as a covering law explanation (viz. DN-explanation):

> "Renormalization group methods take the theoretical framework of statistical mechanics as the covering law. They select as the particular conditions a broad class of Hamiltonians pertinent to the materials. They then derive universality under conditions close to criticality. The renormalization group analysis simply is a covering law explanation." (2012, 227)

There are two difficulties with Norton's appeal to the DN-model here: first, regarding the DN-model's assumption that the explanans statements are (approximately) true, and secondly, regarding general objections to the DN-model.

[3] See also Morrison (2012).

The DN-model, as formulated by Hempel and Oppenheim, incorporates the truth requirement according to which the explanans of an actual explanation must be true (otherwise we have at best a potential explanation). If the RG explanans involves an ineliminable appeal to limits, how is this compatible with the truth requirement, since it is not literally true that, for instance, the number of molecules in an actual gas is infinite (as per thermodynamic limit)? One could try to argue in response (as one referee suggested) that assumptions about the relevant limits in the context of RG explanations, while not literally true, are nevertheless *approximately* true in some sense. Admittedly there is a certain amount of leeway in the truth requirement: approximately true explanans can also support actual explanations. And, indeed, it seems reasonable to relax the truth requirement from literal truth to approximate truth, because demanding that all explanantia are true simpliciter would render all of known physics outright non-explanatory. But this raises the question of how to make sense of the infinite limits in the RG explanation as being 'approximately true' with respect to finite systems.

It is far from clear, in particular, how the *indispensability* of these limits meshes with the DN-model. From the perspective of the DN-model it is natural to expect, as Hempel (1962) does, that explanations improve if the explanantia are closer to the truth, and this is what happens in the case of de-idealizable 'Galilean' idealisations, of course. (For example, an explanatory deduction that truthfully incorporates a small but non-vanishing friction term may be more cumbersome, but such de-idealization improves predictive accuracy and, if anything, also provides for a better explanation.) Providing a sense in which approximately true explanans can underwrite actual explanations hinges on the availability of this kind of story of how the approximations involved are, strictly speaking, *not* indispensable for deducing the explanandum. And arguably this kind of story is not forthcoming in the case of RG explanations.

So, although we deem Norton's approximation/idealisation distinction justified and appropriate in relation to the RG fixed points, it does not in and of itself fully respond to the Anti-Reductionist Challenge. The sense in which the fixed points of the RG explanation are merely instrumental 'mathematical pegs' needs to be further elucidated, with reference to an appropriate understanding of the nature of the RG explanation. Norton's suggestion to regard the RG explanation as a covering law explanation is problematic, not only for the reasons given, but also due to the more generic criticisms of the DN-model. Hempel identifies explanatory understanding with nomic expectability (provided by a suitable inductive or deductive argument), but well-known, extended critiques of the DN-model have shown that the provision of nomic expectability is neither necessary nor sufficient for having an explanation (e.g. Salmon 1989, 46–50). Consequently it is increasingly popular to identify explanatory understanding with knowledge of explanatory dependence relations that are taken to differ from nomic expectability, even in cases that are structurally similar to DN-deductions (Woodward 2003, Strevens 2008). These prominent viewpoints challenge the DN-model on rather general grounds in a way that really puts the onus on those who insist that the DN-model is nevertheless appropriate in the specific context of RG explanations.

In a spirit similar to Norton, Menon and Callender (2013) admit that RG fixed points are an indispensable part of the explanatory resources needed to account for critical phenomena, and also that fixed points transcend the representational resources of the more fundamental, finite, reducing theory. Yet they go on to suggest that we can account in reductive terms for why the explanatory appeal to fixed points is warranted. Their discussion makes little connection with physicists' explanations of universality, however, and they offer no analysis of the nature of physicists' RG explanations. Their account thus unfortunately leaves open the indispensable role that RG fixed points play in actual scientific accounts of universality, and exactly why this role can be regarded as ontologically innocuous. These issues that are not addressed by Menon and Callender are precisely the target of this paper.[4]

[4]More specifically, Menon and Callender point to finite-size crossover theory as explaining why a *particular* finite system can be treated as an infinite system, indicating that it is very difficult or perhaps impossible to empirically distinguish between a system flowing *to* the critical point, as opposed to flowing *close* to it. Regarding the indispensable use of infinite limits in explanations of universality, they summarise their argument:

> "When we try to explain the universality of critical behavior in finite systems, we do have to employ the infinite idealization, but as we have seen, this idealization is not irreducible if we can use the topological structure of system space in our reductive explanation. We can de-idealize for particular systems, and see why they can be treated as if they flow to the critical point. Understanding the behavior of infinite systems is crucial to explaining the behavior of finite systems, since we only get the fixed points by examining the behavior of infinite systems, but this in itself does not imply emergence. We agree with Batterman (2011) that mathematical singularities in the renormalization group method are information sources, not information sinks. We disagree with his contention that the renormalization group explanation requires the infinite idealization, and is thus emergent. It requires consideration of the behavior of infinite systems, but it does not require us to idealize any finite system as an infinite system. Any actual infinite idealizations in a renormalization group explanation can be de-idealized using finite-size cross-over theory. Locating fixed points does not require an infinite idealization, it just requires that our microscopic theory can talk about infinite systems, and indeed it can." (2013, 221-2)

There are also some other insightful reductionist commentaries on the theoretical role of fixed points in the RG analysis of critical phenomena, but we find these similarly lacking in perspicuity regarding the fixed points' explanatory role. For example, while we are largely in agreement with Hüttemann et al. (2015), we don't think they go far enough in responding to the anti-reductionist challenge by virtue of leaving the explanatory role of RG fixed points unanalysed. And we see little reason to analyse RG explanations simply as DN explanations, as Butterfield (2011a, 2011b) and Norton (2012) do. It is against this context that we now aim to do better.

3 The RG Explanandum

The RG framework furnishes a number of explanations regarding critical phenomena. The first order of business is to precisify the explanandum that we have chosen to focus on: the universality of critical exponents. Making this explanandum more precise contributes to addressing the 'first lacuna' presented in the introduction.

Critical phenomena involve continuous (second-order) phase transitions in macroscopic systems near the critical point, where large-scale collective behaviour becomes significant. Standard examples of systems exhibiting critical phenomena include liquid-vapour and magnets. Dynamically generated collective behaviour is quantified by the correlation length, characterising the scale at which a collective behaviour is observed. At the critical point the correlation length diverges (in the models of critical phenomena), indicating that near this point it becomes very large, capturing system-wide macroscopic properties. Near the critical point macroscopic, thermodynamic properties obey characteristic power laws as a function of reduced temperature t, proportional to the distance from critical temperature: $t = \frac{t-t_c}{t_c}$. It is remarkable that micro-physically very different systems, such as liquid-vapour and ferromagnets, can have similar power laws, with identical critical exponents. This is an instance of the kind of universality that comprises the explanandum at stake.

Consider, for example, the scaling laws obeyed by ferromagnets, on the one hand, and simple liquids, on the other. (Here we have magnetic susceptibility χ, heat capacity C_H, and magnetisation M; compressibility κ, heat capacity C_V, and liquid and gas densities ρ_l, ρ_g.)[5]

$$
\begin{aligned}
\chi &\propto |t|^{-\gamma} & \qquad \kappa &\propto |t|^{-\gamma}\\
C_H &\propto |t|^{-\alpha} & C_V &\propto |t|^{\alpha}\\
M &\propto |t|^{\beta} & (\rho_l - \rho_g) &\propto |t|^{\beta}
\end{aligned}
$$

It is crucial to be clear on the precise nature of universality in question. It is not the case that *all* different systems exhibiting critical phenomena are exactly similar in this way. Rather, simple liquid-vapour and ferromagnetic systems have the same critical exponents by virtue of belonging to the same *universality class* (viz. the 'Ising class', also containing the theoretical spin-1/2 Ising model). Other universality classes describe systems with different critical exponents. In general, a specific universality class, identified by its critical exponents, *depends* on the following variables: spatial dimensionality, the symmetry of the order parameter (also called the 'spin dimensionality'), and the range of the microscopic interactions. We will return to this central issue regarding what universality *depends on* below.

The real explanandum of the RG analysis is this kind of *curtailed* universality, with systems falling within a relatively small number of distinct universality classes. A blunt notion of universality – that micro-physically different systems can obey power laws with identical critical exponents – is not at issue, as it can be established by 'classical' (non-RG) methods of mean-field theory and Landau, and it indeed had already been established prior to the development of RG analysis (see e.g. Als-Nielsen and Birgeneau, 1977; Kopietz et al. 2010, ch. 2). These classical methods yield estimates of critical exponents that do not fare well empirically, however, and they failed to indicate the dependence of the critical exponents on systems' spatial dimensionality in particular. The celebrated explanatory contribution of the RG analysis must be appreciated and understood in this (pre-RG) theoretical context. This is something that many expositions of the RG analysis emphasise quite explicitly:

> The starting point is mean field theory which allows us to describe phase transitions and explore the neighbourhood of the critical temperature. In the case of second order phase transitions, continuous

[5] The critical exponents α, β, γ furthermore obey a simple (Rushbrooke) inequality $\alpha+2\beta+\gamma \geq 2$, showing that they are not independent from one another. This is another explanandum for the RG framework.

> phase transitions where the correlation length diverges, this leads to the concept of *super-universality*. The latter is summarized in Landau's theory of critical phenomena. A number of quantities, like the exponents which characterize the singular behaviour of physical observables near the critical temperature, are universal, i.e. independent of the system (provided it has only short range interactions), and even the dimensions of space. However, empirical evidence, exact solutions of 2D models, and finally an analysis of corrections to mean field theory, had shown that a universality of such general nature could not be true. [...] The existence of even a more restricted universality was puzzling. It took many years to develop the [RG] which explains the origin of universality: it relies on the existence of IR fixed points of RG transformations. (Zinn-Justin, 2006, 218)

What we are emphasising here, along with Zinn-Justin, is the fact that in the pre-RG context of mean field and Landau theories what really needed explaining was not universality *per se* – how microscopically very different systems could be similar in their macro properties – but the observed *dependence* of critical exponents and universality classes on the systems' spatial dimensions and the other features that carve the nature into these 'universality kinds'. So, the question was: how does it follow from the laws of statistical mechanics, including the dynamical laws and the partition function connecting the micro- and macro-levels, that the properties exhibiting universality depend on the variables outlined above. The explanatory contribution of the RG analysis has been to answer this question by deriving the values of the critical exponents for large classes of Hamiltonians in a way that brings out the explanatory dependencies as their logico-mathematical consequence for systems of sufficiently many components.[6] Furthermore, this framework provides an understanding of the nature of the dependence in question as a collective probabilistic matter, having to do with the way in which chancy fluctuations across a huge range of scales relate and contribute to the macroscopic properties.[7]

[6]One should not overplay the rigour and precision here: due to the level of abstraction and mathematical intractability, RG 'derivations' involve various approximation schemes and plausibility considerations, yielding approximate values for the measured critical exponents.

[7]Ignoring the importance of fluctuations is where the mean field theoretical approaches to critical phenomena go wrong. This is particularly critical for systems with the spatial dimensionality below 'upper critical dimensionality', which is typically 4. It can be shown that for systems of larger dimensionality mean field theories, despite their crude approximations, yield the correct universal critical exponents.

4 The RG Explanans

In this section, we will reconstruct the relevant physics of the explanans of RG explanations in detail (§4.1). Then, in response to the 'second lacuna' presented in the introduction, we will suggest that the counterfactual dependence account of explanation nicely captures the explanatory character of RG explanations (§4.2).

4.1 The Physics of the RG Explanans

After some preliminaries (in subsection (a)), we will focus on: (b) the sense of *coarse-graining* associated with the renormalisation operation; (c) the notion of *fixed-point* of the renormalisation operation and its explanatory role; and (d) the notion of *universality class* of Hamiltonians.

(a) Preliminaries. An RG analysis of collective behaviour near criticality brings out a network of dependencies between the critical exponents, on the one hand, and (i) spatial dimensionality, (ii) the dimensionality of the spin parameter, and (iii) certain qualitative features of the Hamiltonians that characterise systems micro-physically, on the other.[8] An RG analysis explains by showing how the critical exponents, viz. the explanandum, depends on (i)-(iii), and how this network of dependencies mathematically follows from statistical mechanics. The RG framework accomplishes this by various means. First, there are extremely general RG analyses that treat spatial dimensionality and the dimensionality of the spin parameter as variables, and determine how RG fixed points depend on these variables. (e.g. Zinn-Justin, 2007) Second, there are more circumscribed RG analyses of specific classes of Hamiltonians, regarding systems of specific spatial dimensions and spin parameters. We will focus on the latter.

Recall that a Hamiltonian, or energy function, characterizes the energy of the interactions between the system's components, and also the energetic effect of

[8] This set of dependencies is unlike those represented by the ideal gas law, for instance, in that the explanans variables cannot be grasped independently of the microlevel description. Even spatial dimensionality has to be understood in a particular way, as it matters only insofar as it tracks the connectivity of systems' elementary degrees of freedom. For example, a three-dimensional 'magnet' made of effectively two dimensional slabs, the spins of which do not interact across the slabs, does not count as three dimensional in this respect. Similarly, an anisotropic d-dimensional lattice where the energy parameters connecting the lattice points in the direction of one axis tend to zero is, from the point of the connectivity of elementary degrees of freedom, effectively $(d-1)$-dimensional. The relevance of spatial dimensionality can be qualitatively understood in probabilistic terms. Classical, pre-RG approaches of critical phenomena underestimate the importance of statistical fluctuations, the probability of which increases along with spatial dimensionality (see e.g. Stanley, 1999, 365).

the external conditions (e.g. magnetic field) on the system. For example, for a very simple spin-1/2 Ising model, the Hamiltonian function is given by

$$H = -J \sum_{<ij>} S_i S_j - h \sum_i S_i \tag{1}$$

Here S_i is the 'spin' parameter ($S_i = \pm 1$) ranging over all lattice sites; the first summation is over all interacting pairs of spins, and the coupling constant J gives the interaction energy. The energetic contribution of an external magnetic field is represented by h. The lattice of spins can be a 1-dimensional string, 2-dimensional square lattice, 3-dimensional cubic lattice, or (more abstractly) d-dimensional hyper-cubic lattice.

For this class of models, the spin parameter S_i has only one component, and the model has global $\mathbb{Z}_2$ symmetry: in the absence of an external magnetic field the Hamiltonian is invariant under $S_i \to -S_i$ $(\forall i)$. We can enlarge the set of possible interactions by allowing the spin parameter to have further possible values, and/or more components. For instance, in spin-1 Ising models $S_i \in \{1, 0, -1\}$, and in q-state Potts models $S_i \in \{0, 1, 2, \ldots, q\}$. The spin parameter can also be a vector, instead of scalar. For example, in XY-models the spin parameter has two components, $\mathbf{S}_i = (S_i^x, S_i^y)$, and in n-vector models $\mathbf{S}_i$ has n components. Changes in these features of the spin parameter can change the symmetry of the Hamiltonian (depending on how it has been defined in terms of the various spin-spin interactions), resulting in different critical exponents.

The dependence of the critical exponent on spatial dimensionality and the dimensionality of the spin parameter can be studied by RG analyses of specific models, involving specific (classes of) Hamiltonians, and specific spatial dimensions and spin parameters. Such model-specific RG analyses collectively contribute to showing how critical exponents depend on the specific dimensionalities and symmetries. One can, for example, compare the results of RG analyses of two-dimensional vs. three-dimensional spin-1/2 Ising models. Or one can compare RG analyses of n-dimensional spin-1/2 vs. spin-1 models. In the context of these models one can furthermore show that a specific Hamiltonian is *not* responsible for the value of the critical exponent, since there is much leeway in the exact form of the Hamiltonian, as long as the essential parameters – spatial dimensionality and dimensionality of the spin parameter – are kept fixed. This establishes a *local* universality in relation to such specific models: critical phenomena are independent of the details of the original microscopic interaction, since the specific modelling Hamiltonian can be perturbed without changing the features that factor into the calculation of critical exponents.[9]

(b) Coarse-graining and renormalisation. More generally, RG analyses show that the different Hamiltonians in a given universality class are similar in that they exhibit similar collective behaviors under sufficient 'coarse-graining': when it comes to long-distance physics near the critical point, the differences in the microphysical couplings wash out. The sense in which the different Hamiltonians are similar in this way is provided in terms of mathematical renormalization group (RG) transformations. The behaviour of Hamiltonians under iterated RG transformations can be used to determine the critical exponents near the critical point, as we will presently explain. Roughly speaking, universality with respect to variation in the specific Hamiltonian then follows from different systems' similarity in this respect, and the RG analysis provides a sense in which a given universality class depends on the Hamiltonians in that class having this feature (in addition to depending on spatial dimensionality and the dimensionality of the spin parameter).

In order to flesh out this sketch, and to pinpoint the role played by RG fixed points in finding out about this dependence, we now present a schematic outline of an RG analysis.[10] The gist of the (very broadly applicable) RG framework is to explore ways of re-expressing – 'renormalizing' – a set of relevant modelling parameters in terms of another (possibly simpler) set of parameters, and then rescaling the system, in a way that keeps unchanged some physical aspects of interest. In the context of critical phenomena, a renormalization transformation amounts to the coarse-graining of the short-distance degrees of freedom, while keeping a system's long-distance physics fixed.[11] This is achieved, in particular,

[9]There are limits how much a Hamiltonian can be changed without affecting critical phenomena. It matters, in particular, how short/long-range the micro-interactions are. Again, this can be studied by case-specific models, looking e.g. at spin-1/2 Ising models for interaction

$$H = -J \sum_{<ij>} \frac{1}{r_{ij}^{\omega}} S_i S_j$$

as a function of ω, with r_{ij} the distance between lattice sites i, j (e.g. Cannas, 1995).

[10]We only provide a schematic presentation of the key concepts involved in an RG analysis; for further details see e.g. Fisher (1983, 1998), Wilson (1983), Cardy (1996), McComb (2004), Zinn-Justin (2007), Sethna (2006), Pathria and Beale (2009), Nishimori and Ortiz (2010). Our presentation is mostly drawn from the Nishimori and Ortiz (2010). See also Appendix for further details.

[11]Much of the ingenuity in the development of the RG framework has gone into techniques that can be used to implement with sufficient rigour this kind of coarse-graining (e.g. real-space RG, momentum space RG, Monte Carlo RG).

by ensuring that the partition function is left intact by the re-parametrisation.

The RG analysis operates on a large (possibly infinitely dimensional) abstract space of possible 'models', or parameters, with a different dimension for each possible parameter of the Hamiltonian (e.g. couplings between immediate neighbours, next-neighbours, etc.), as well as for each 'control' parameter that can be tuned in an experiment (e.g. temperature, external magnetic field, chemical composition, etc.). The RG framework studies the way in which this high-dimensional space of parameters maps into itself under a renormalization operation R_b of the relevant parameters: $\mathbf{u}' = R_b(\mathbf{u})$.[12] Iteration of the renormalization operation induces a 'flow' in the parameter space, whereby a 'model' $\mathbf{u}$, specified by particular parameter values, gets mapped to a different point $\mathbf{u}'$.

$$\mathbf{u} \to R_b(\mathbf{u}) \to R_b^2(\mathbf{u}) \to \cdots \to R_b^n(\mathbf{u}) \to \cdots$$

Analysing the structure of such flow in the space of parameters is the essence of the RG analysis: it qualitatively explains why the long-distance physics (especially with respect to scaling laws) for microscopically different systems S_1, S_2, S_3,... is similar near their respective critical points, and it quantitatively allows for a calculation of critical exponents. (We will focus on the qualitative explanation below. See Appendix for comments on the quantitative aspects.)

A rough idea of qualitative RG understanding of universality near criticality can be given as follows. Let's assume that two Hamiltonians H_1 and H_2 in the space of parameters display similar flow behaviour and end up close to one another under repeated RG transformations; viz. the physics captured by those Hamiltonians can be modelled, at a sufficiently coarse-grained level, by *effective* Hamiltonians that are close to one another in the space of parameters. Then the two systems captured by H_1 and H_2 have a similar long-distance behaviour. It is an extraordinary fact that large classes of possible systems (viz. classes of possible Hamiltonians) in this way lead to similar long-distance behaviour near criticality. The RG analysis brings out this fact by revealing broad structural features exhibited by RG flows in the space of parameters. It is here that an indispensable theoretical role is played by the *fixed points* of RG flows.

[12] R_b is associated with a scaling factor b, which determines the rescaling of the system's length scale by $1/b$. A generic Hamiltonian is written as the sum of products of parameters u_n and 'operators' (viz. microscopic degrees of freedom) O_n: $H = \sum_n u_n O_n = \sum_n \mathbf{u} \cdot \mathbf{O}$. For example, in equation (1) the coupling constants J and h are parameters, and $S_i S_j$ and S_i are operators. Renormalisation transformations form a semi-group: $\mathbf{u}'' = R_{b_2}(\mathbf{u}') = R_{b_2} R_{b_1}(\mathbf{u}) = R_{b_1 b_2}(\mathbf{u})$.

(c) Fixed-points. A *fixed point* $\mathbf{u}^*$ (in the space of parameters) is defined as a point (or more generally, a sub-manifold) that gets mapped to itself under renormalization, thereby terminating the RG flow (since further RG transformations do not flow to a different point): $\mathbf{u}^* = R_b(\mathbf{u}^*)$. Prima facie, it seems possible that RG flows might exhibit wildly unstable, even chaotic behaviour, indicating very fine-tuned dependence of a system's large-scale behaviour on its microscopic couplings. It turns out that this isn't so (at least for very large classes of Hamiltonians of interest): instead, in the space of parameters there are points – the fixed points of RG transformations – towards which RG flows are 'attracted'. There is thus much structure and regularity to the way in which different Hamiltonians 'coarse-grain' so as to give rise to similar macroscopic properties. In particular, under repeated RG transformations the effective, coarse-grained Hamiltonians 'gravitate' close to a fixed point, as long as the starting point of the iterated RG transformations – fixed by an original microphysical Hamiltonian and some given values of the control parameters – is sufficiently close to a broad *basin of attraction* of the fixed point. RG fixed points and their associated basins of attraction thus give the abstract space of parameters an interesting topological structure.

The theoretical resources involving the fixed points and their basins of attraction are indispensable (in the current state of physics at least) for grasping this topological structure exhibited by the space of parameters, and for studying its repercussions on those models that, from the perspective of the more fundamental theory, can be taken to faithfully represent a system of interest approaching a critical point. When universal scaling phenomena is demonstrated and measured in the laboratory, it concerns (from the reductionist perspective, at least) finite systems. Not all points in the abstract space of parameters correspond to these finite systems: for the points at *criticality*, in particular, the correlation length ξ diverges, in blatant contradiction with the finitude of the actual systems of interest. Thus, for the reductionist these points are best construed as mathematical approximations of properties of sequences of corresponding finite models, having these points as limits (cf. Norton 2012).

Under renormalization the correlation length transforms as $\xi[\mathbf{u}'] = b^{-1}\xi[\mathbf{u}]$ due to rescaling by factor b. Thereby the RG flow is away from criticality upon each successive operation of R_b, assuming ξ is finite to begin with (see Figure 1). Therefore, in order for the renormalisation flow to terminate, a fixed point must have a divergent correlation length (critical fixed point), or else it must vanish (trivial fixed point). Clearly the correlation length must also be divergent for all the points in the basin of attraction of a critical fixed point. This basin is the critical manifold. The points in the parameter space that correspond to finite models are *not* on this manifold – they do not flow to a (critical) fixed point under RG transformations. Rather, the basin of attraction of a (critical) fixed point comprises points in the space of parameters which are *at criticality*, featuring control parameters (e.g. temperature, or external magnetic field) that have been taken to the critical point in the corresponding phase space. Since the correlation length diverges at a (critical) fixed-point and everywhere on its associated critical manifold, and since the correlation length cannot diverge without taking a thermodynamic limit, the latter is needed to connect statistical mechanics to the RG fixed point.

For concreteness' sake, consider a trajectory in the abstract space of models, induced by smoothly changing one of the control parameters, t. (See figure 1.) This 'physical line' captures how a system modelled by a given Hamiltonian changes as the control parameter moves ever closer to the critical point T_c, where the correlation length diverges. This point T_c is sometimes called the 'physical critical point', but one shouldn't read too much into this label: a reductionist takes the divergence of the correlation length to indicate, of course, that this point in the space of parameters is at best an idealisation of the finite physical system, or perhaps merely a vehicle for approximating some of the properties represented by the points outside of criticality $T < T_c$. This 'physical line' in the space of models is not an RG flow, and the changing macroscopic properties of a system that tracks such trajectory can be measured in the laboratory, e.g. when critical indices are measured. But these macroscopic changes along a physical line cannot be studied theoretically due to the intractability of the huge number of correlations and interactions at different scales due to fluctuations near criticality.

The RG framework deals with this intractability by renormalizing the relevant modelling parameters, yielding more and more coarse-grained *effective* models in a way that keeps the macroscopic physics unchanged. Any point on the 'physical line' can be taken as the starting point of iterated RG transformations, which induce a corresponding flow in the space of models. Unlike the 'physical line', these flows do not correspond to any physical change of the system, but rather capture *equivalence classes* of models that share the same long-distance physics. Of all the points on the physical line, the 'physical critical point' is special, since it (and only it) flows to a (critical) fixed point upon successive iterations of the renormalization transformation. (That is, this point belongs to the critical manifold.)

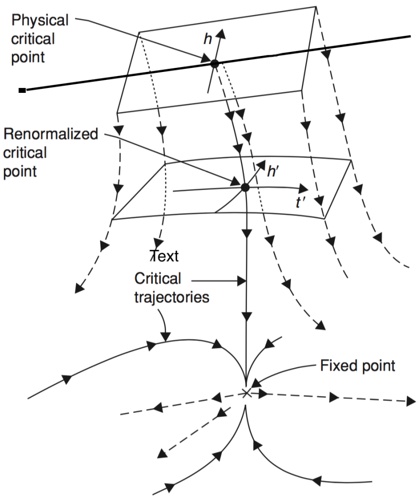

Figure 1: Schematic depiction of RG flow. The thicker, straight line through the 'physical critical point' h,t_c is the 'physical line'.

(d) Universality classes of Hamiltonians. Although the basin of attraction only comprises points in the critical manifold, there is also a broader RG flow *towards* (although not into) a critical fixed point, exhibited by points outside the critical manifold. (Cf. Figure 1.) That is, a flow emanating from a point that is sufficiently close to a basin of attraction will end up in the vicinity of the critical fixed point after some finite number of renormalisation transformations, before veering away from it towards a trivial fixed point. In particular, systems modelled by different off-critical parameters in the 'critical domain' – the neighbourhood of a critical point where correlation length is very large with respect to the microscopic scale – will end up flowing close enough to the critical fixed point for the flow to be examined in terms of *linearised* RG. This examination formally reveals the aspects of the parameter space that are relevant for the value of critical exponents, as well as the aspects that are irrelevant, in the sense that change in the irrelevant parameters is inconsequential to the value of critical exponents. (See Appendix for details.) Systems with different Hamiltonians that only differ in the irrelevant parameters therefore exhibit the same long-distance behaviour near their critical point. They belong to the same universality class.

With the distinction between relevant and irrelevant parameters we can capture an important mathematical fact about the behaviour of a large class of Hamiltonians under a given renormalization transformation. Once we fix the laws of statistical physics and an appropriate renormalisation transformation, the fact that two systems A and B are in the same universality class follows with mathematical necessity. For a want of a better analogy, consider, for example, composition of forces. Assume that two different sets of force vectors $\{\mathbf{f}_1, \mathbf{f}_2\}$ and $\{\mathbf{f}_3, \mathbf{f}_4\}$ result in the same total force $\mathbf{f}_1 + \mathbf{f}_2 = \mathbf{F} = \mathbf{f}_3 + \mathbf{f}_4$. The fact that both sets are similar in this way – they both belong to the same 'universality' class of component vectors that add up to $\mathbf{F}$ – follows with mathematical necessity, once we fix the law of force composition. Similarly, a given universality class depends on its Hamiltonians in the same way: the fact that two systems with different Hamiltonians only vary in the irrevant parameters follows with mathematical necessity, once we fix a renormalisation transformation.[13]

4.2 The Counterfactual Dependence Account of RG Explanations

Having summarised the key concepts of the RG explanans, let us now consider the philosophical issue at stake in the second lacuna: which philosophical account of explanation best captures RG explanations? We think a promising approach to RG explanations is the counterfactual dependence account of scientific explanation. This approach takes as its starting point the key idea behind Woodward's counterfactual account of causal explanation:

> "An explanation ought to be such that it enables us to see what sort of difference it would have made for the explanandum if the factors cited in the explanans had been different in various possible ways." (Woodward 2003: 11)

Understanding explanatory relevance thus in terms of counterfactual dependence is not necessarily tied to a causal interpretation. The basic idea can be extended from causal to non-causal explanations, as Woodward himself indicates in terms of "what-if-things-had-been-different questions":

[13] Some philosophers have classified this kind of dependence as clearly not causal, and RG explanation as a kind of 'non-causal' explanation (this point has been elaborated by Lange 2009, 2013). Lange's notion of 'distinctively mathematical' explanations is one possible way of interpreting this mathematical aspect of RG explanations. (Cf. Lange 2013) However, although we think that the RG explanation is non-causal and mathematical, we disagree with Lange's analysis of the *explanatoriness* of these kinds of 'distinctly mathematical' explanations. As we will argue in Section 4.2, one can also view this explanation from the perspective of counterfactual dependence account of explanation, according to which it is the counterfactual dependences between the explanandum and the explanans that drives the explanation, not the fact that the explanandum is mathematically necessary given the explanans. (See also Jansson and Saatsi, forthcoming)

> “[T]he common element in many forms of explanation, both causal and non-causal, is that they must answer what-if-things-had-been-different questions." (Woodward 2003: 221)

Proponents of a counterfactual dependence account of explanation have developed and made precise the idea Woodward expresses. They hold that both causal and non-causal explanations are explanatory by virtue of exhibiting how the explanandum counterfactually depends on the explanans. The global motivation for defending this approach to scientific explanation stems from its unificationist or monist prospects – that is, it is attractive to have one single theory of explanation for two types of explanation (causal and non-causal). In the recent literature, the counterfactual dependence account of explanation has been articulated and explored in application to various examples of non-causal explanations (Frisch 1998; Bokulich 2008; Kistler 2013; Saatsi and Pexton 2013; Pexton 2014; Reutlinger 2016; Saatsi 2017; French and Saatsi forthcoming; Jansson and Saatsi forthcoming; Woodward forthcoming).

We will now argue in more detail why we think the counterfactual dependence account applies to RG explanations. To do so, we will focus on the core of the account consisting of two necessary conditions for being a scientific explanation (we follow the exposition in Reutlinger 2016). We ignore other necessary conditions here on which proponents of this account of explanation differ.

First, the counterfactual dependence account requires that one can infer the explanandum from the explanans (where this inference may be deductive or statistical-inductive). In the case of RG explanations, this condition is satisfied because the RG explanans (consisting of Hamiltonians and the theoretical framework of statistical mechanics, RG transformations, the determination of fixed point, and so on, as described in §4.1) deductively entails the RG explanandum. We take the satisfaction of this condition to be the kernel of truth in Butterfield’s and Norton’s claim that RG explanations are DN-explanations (see §2).

Second, the counterfactual dependence account also requires that the explanans allows us to evaluate counterfactuals of the following form as true: “if some variables figuring in the explanans had specific different values (which typically corresponds to assuming that the initial conditions of a physical system are different than they actually are), then the explanandum phenomenon would also be different in some specific way". RG explanations satisfy this condition, since the RG explanans enables us to determine whether a physical system S would be in a different universality class, if certain features of S were different than they actually are. In other words, the RG explanans conveys what being in a specific universality class *counterfactually depends on*. For instance, the RG explanans supports the following explanatorily relevant counterfactuals:

- If a physical system S had a different spatial dimensionality than it actually has, then S would be in such-and-such a different universality class than it actually is in.
- If a physical system S had a different symmetry of the order parameter than it actually has, then S would be in such-and-such a different universality class than it actually is in.
- If a physical system S had a (sufficiently) different range of the microscopic interactions than it actually has, then S would be in such-and-such a different universality class than it actually is in.

As we have seen in previous sections, it is a central purpose of the RG framework to underwrite such conditionals. First, as discussed in Section 3, the key explanandum regarding universality that was left outstanding in the pre-RG context of mean field and Landau theories, was the observed dependence of critical exponents and universality classes on those features of the world that seem to carve the nature into these broad kinds. Secondly, as discussed in detail in Section 4.1, the RG framework provides the means to bring out the relevant dependencies, by virtue of showing *exactly how* critical exponents depend on spatial dimensionality, dimensionality of the spin parameter, and the range of micro-interactions.

Let us highlight the crucial point encoded in the two necessary conditions of the counterfactual dependence account: the RG framework does not only deductively entail that many physical systems with different original microphysical Hamiltonians display the same macro-behavior (as required by the first condition). In addition, the RG framework also provides a wealth of modal information regarding what being in a specific universality class depends on (thereby satisfying the second necessary condition).

In sum, we take it that the counterfactual dependence account of explanation has a good claim to capture the explanatory character of RG explanations (see Reutlinger 2016 for an in-depth discussion). This addresses the 'second lacuna'. We are now in a position to use the assumption that the counterfactual dependence account applies to RG explanations to rebut the anti-reductionist challenge.

5 How to Meet the Anti-reductionist Challenge

Recall the anti-reductionist challenge motivated by the argument from infinite limits: to show how the fixed points involved in RG explanations of critical phenomena can be both (a) explanatorily indispensable and, at the same time, (b) compatible with reductionism? How should the reductionist respond?

Our response, in the light of our analysis of the RG explanation, is to argue that even if reference to fixed points is indispensable for RG explanations, it plays a merely instrumental role and does not lead to an ontological commitment in tension with reductionism. This argument rebuts ontological anti-reductionism of the sort most explicitly defended by Morrison (2012, 2015) in particular. At the end of this section we will consider the residual issue of explanatory (anti-)reductionism.

Let us now go through this argument in detail. To begin with, we grant the anti-reductionists (at least for the sake of the argument) that reference to fixed points is indispensable for RG explanations of universality. Indeed, fixed points are in a profound sense at the heart of our best understanding of critical phenomena, which in many ways turns on the classification of fixed points and examination of their properties.

However, the actual physical systems of interest – being finite both spatially and with respect to the number of microphysical components – cannot be represented by the fixed point Hamiltonian itself, or the points on the critical manifold for that matter, since the correlation length diverges for these points. But although the fixed point Hamiltonian cannot be taken to represent the finite target system, reference to the fixed point can be explanatorily indispensable. In particular, fixed points are *instrumentally indispensable* for finding out and expressing facts about those 'models' in the space of parameters that lie in the critical domain but outside of criticality – facts that are explanatorily relevant for critical phenomena. It is these off-critical 'models' that represent the features of the world on which the critical exponents and universality classes depend, and RG analyses explain by virtue of (and to the extent they succeed in) providing correct information about such dependencies (see below).

Although fixed points do not, in and of themselves, represent anything about the finite systems exhibiting critical phenomena, reference to fixed points is nevertheless indispensable. There are at least two reasons for this. Firstly, the RG analysis provides a method for calculating the critical exponents through an analysis of the nature of the RG flow in the neighbourhood of the fixed point. Obviously, we cannot speak of the neighbourhood of a fixed point without speaking of the fixed point itself, so analysing the structure of the equivalence classes of Hamiltonians (viz. the RG flows) in the neighbourhood of the fixed point naturally involves a reference to the fixed point (cf. §4.1 (c) and (d)). But here the reference to the fixed point is merely playing the role of determining that we are operating in the *critical domain* where the correlation length is large enough with respect to the microscopic scale for the linearised RG analysis to be valid. (See Appendix.) This kind of reference to critical fixed points in connection with specific models is indispensable for doing the calculations that contribute to showing how the critical exponents depend on the relevant explanatory parameters.

Secondly, reference to critical fixed points is indispensable for expressing the explanatorily relevant feature shared by all possible Hamiltonians in a given universality class. This is done by reference to the mathematical fact that upon renormalization they all end up in the vicinity of the same fixed point, where their further coarse-graining is similarly dependent only on a few relevant variables (cf. §4.1 (d)). Although it is necessary to make reference to the fixed point in expressing and theorising about this feature that the Hamiltonians share, one can again adopt an instrumentalist attitude to the fixed point itself, while using it to express the explanatorily relevant feature regarding the physical Hamiltonians' behaviour under coarse-graining.

If we thus grant that reference to fixed points is indispensable, how do we avoid an ontological commitment that is in tension with reductionism, and how do we justify the claim that fixed points are merely *instrumentally* indispensable? Our response relies on the assumption that the counterfactual dependence account captures the explanatory character of the RG explanation. If this is correct, then we naturally relate RG explanation's ontological commitments with those (and only those) variables on which the explanandum depends. Relative to the framework of the counterfactual dependence account, we can draw a distinction between those aspects of an explanation that feature in (or represent) explanatory counterfactual dependence relations, on the one hand, and those aspects that play some other role, e.g. in communicating or facilitating the explanation, on the other. Given such a distinction, an explanation's ontological commitments are naturally associated with the former aspects. That is, only those factors on which the explanandum counterfactually depends carry ontological import.

Much more has been said about this broad philosophical stance towards explanations' realist commitments, e.g. in Saatsi (2016b), which provides a more general analysis of explanatory indispensability with a view to distinguishing between ontologically committing ('thick') and instrumental ('thin') explanatorily indispensable assumptions. According to Saatsi, the precise content of the thin/thick distinction is relative to a given account of scientific explanation. Saatsi then argues, in the context of different counterfactual dependence and modal accounts of explanation, that the explanatory indispensability of mathematics is insufficient for ontological commitment to mathematics.[14] (This satisfies the fairly minimal veridicality criterion that Woodward and others have associated with counterfactual dependence accounts: actual (as opposed to merely potential) explanations should get right the explanatory modal facts. Yet at the same time it leaves room for instrumentalism about explanations' other aspects, regardless of their 'indispensability' or otherwise.)

Hence, assumptions about RG fixed points can be indispensable without being ontologically committing in a way that poses a threat to reductionism, if the RG explanandum does not counterfactually depend on the fixed points. We will now argue that this is the case.

In the light of the counterfactual dependence account of explanation, we see that the RG explanation works by bringing out how the explanandum (viz. critical exponents) counterfactually depends on features such as the spatial dimensionality, the symmetry of the order parameter, and the range of the microscopic interactions. And these are all the explanatory dependencies involved in RG explanations. In particular, there is no analogous counterfactual dependence of the RG explanandum on the fixed points. In the context of the RG explanations, a fixed point is simply not considered to be a 'variable', with different possible values, that we can associate, via RG transformations, with different possible states of the explanandum variable. Instead, if one accepts that fixed points are indispensable for RG explanations of universality (as anti-reductionists do), then one is committed to the claim that it is impossible to consider counterfactual variations of the relevant assumptions about fixed points (for instance, through de-idealization) without losing explanatory power. Indeed, everyone in the debate agrees that it is not part of an RG explanation to exhibit what would happen if the fixed points were different.[15] Similarly, it is not a part of RG explanations to li-

[14] In a similar spirit, Saatsi (2017) argues for an instrumentalist attitude towards 'strange attractors' and certain other mathematical features that are indispensable in explanations provided by dynamical systems theory. See also Woodward (2003b).

[15] Note that we restrict this claim to RG explanations of universality of critical phenomena. Given

cense counterfactual assertions about what would happen if the laws of statistical mechanics were different.

If this reasoning is sound, then the ontological reductionist is able to maintain an instrumentalist attitude towards RG fixed points; reference to them in the RG explanations of universality does not have an ontological commitment. In particular, the kind of explanatory indispensability at stake does not suggest any ontological commitment to facts that are beyond the domain of statistical mechanics.

Admittedly, there is still room here for an *explanatory* anti-reductionist to protest that we haven't given a full reductionist story, purely in terms of the 'nuts and bolts' of the micro-constituents and their interactions, *why* in the first place the mathematical space of parameters has an interesting topological structure shaped by the fixed points and critical surfaces.[16] This raises deep issues that require not only a solid grasp of the nature of RG explanations, but also a careful analysis of the nature of explanatory (anti-)reductionism itself. What exactly does it take to underwrite, purely in terms of the 'nuts and bolts' of the reducing theory, an explanatory use of novel theoretical concepts like RG fixed points? What exactly is the explanatory anti-reductionist (thus construed) asking from the reductionist? These issues require further work, but in our view there is a real risk that explanatory irreducibility turns out to be a thesis that is so undeniable and widespread that it seriously reduces the interest of specific considerations turning on RG explanations of universality. For instance, consider a popular example from the current literature on explanation: Koenigsberg's bridges' property of being *non-Eulerian.* This is a novel theoretical concept that is explanatorily indispensable (with respect to the bridges' traversibility), and it is not at all clear how we could hope to give a full reductionist account of the property of being non-Eulerian, purely in terms of the 'nuts and bolts' of the physics that describes the bridges at the microlevel.[17] If one takes this as an argument for some kind of explanatory anti-reductionism, then some kind explanatory anti-reductionism

the extremely broad-ranging applicability of RG methods in statistical physics and quantum field theory, we consider it an open question whether there are other RG theories which underwrite explanatory counterfactuals with an RG fixed-point as a variable, such that they cannot be thought of as capturing an explanatory dependence of the explanandum in question on features of finite physical systems in the vicinity of the fixed-point.

[16]We appreciate an anonymous referee pressing us on this.

[17]A connected graph G is Eulerian iff every vertex has an even degree. In connection with this now well worn example the philosophical discussion has taken an explicitly ontological turn (e.g. Pincock 2007). Indeed, explanatory anti-reductionism is more or less taken as for granted, the only issue of real interest being ontological.

seems undeniable given how quickly examples of this sort multiply, and it is no longer clear what particular contribution is made by the RG explanation of universality.

6 Conclusion

We have provided a strategy for meeting the anti-reductionist challenge and for rebutting the argument for ontological anti-reductionism from infinite limits in the context of RG explanations of universality. This strategy is broadly in agreement with Norton's metaphor of fixed points as a 'mathematical peg', and also with Menon and Callender (2013). But we think it is critical that an analysis of the explanatory role of fixed points is properly couched in the context of a fitting philosophical account of the explanation at stake. In as far as the counterfactual dependence account of explanation provides a fitting account, a reductionist can happily admit the explanatory indispensability of the limits involved in RG fixed points.

The challenge of *explanatory* anti-reductionism, raised by Batterman in particular, requires more discussion than we are able to provide here. In our view the issue at stake is clouded by the current lack of clarity as to the exact nature of explanatory anti-reductionism – this is why we have chosen to focus on an explicitly ontological construal of the anti-reductionist challenge. For example, Menon and Callender (2013, 210) characterise explanatory irreducibility, reasonably, as taking place "when the explanation of a higher-level phenomenon requires a conceptual novelty, yet the reducing theory does not have the resources to explain why the conceptual novelty is warranted." In the context of the RG explanation of universality the question then is whether or not the more fundamental theory has the resources to explain why the explanatory appeal to fixed points and critical surfaces of RG space is warranted. As discussed at the end of the previous section, it is not clear to us what this question amounts to exactly – what is being asked by way of providing reductionist warrant for novel theoretical concepts. But we do maintain that this question cannot be answered without taking properly into account the nature of the RG explanation, and the preceding analysis contributes to this task. To recall, we claim to have shown how the reductionist can understand how RG fixed points and critical surfaces are instrumental in approximating aspects of finite physical systems near criticality, and how they function as indispensable instruments in bringing out explanatory dependencies between physical variables. But their explanatory indispensability notwithstanding, the reductionist can consistently maintain that the fixed points and critical surfaces do not represent anything at all in the finite systems that exhibit universality.

Appendix: Linearised RG

In the vicinity of the fixed point RG transformations can be linearised in a way that supports the distinction between relevant and irrelevant parameter (e.g. Nishimori and Ortiz (2010). (Linearised RG analysis also supports quantitative calculation of critical exponents.)

Consider a renormalization transformation from $\mathbf{u}$ to $\mathbf{u}' = R_b(\mathbf{u})$ in the vicinity of the fixed point. Writing these two points in terms of small deviations from the fixed point we have $\mathbf{u} = \mathbf{u}^* + \delta\mathbf{u}$, and $\mathbf{u}' = \mathbf{u}^* + \delta\mathbf{u}'$. In the neighbourhood of the fixed point $R_b(\mathbf{u})$ can be expanded to first order:

$$\mathbf{u}' = \mathbf{u}^* + \delta\mathbf{u}' = R_b(\mathbf{u}^* + \delta\mathbf{u}) = R_b(\mathbf{u}^*) + \frac{\partial R_b}{\partial u}|_{u^*} \cdot \mathbf{u} + \cdots$$

This yields a linearised renormalization equation

$$\delta\mathbf{u}' = T_b(\mathbf{u}^*) \cdot \delta\mathbf{u}$$

where $T_b(\mathbf{u}^*)$ is a matrix with components

$$[T_b(\mathbf{u}^*)]_{ij} = \frac{\partial u_i'}{\partial u_j}|_{u^*}$$

It turns out that critical exponents can be calculated from the eigenvalues $\{\lambda_i\}$ and eigenvectors $\{\phi_i\}$ of this linear transformation T_b. The small deviations from the fixed point, $\delta\mathbf{u}$ and $\delta\mathbf{u}'$ can be written as:

$$\mathbf{u} = \mathbf{u}^* + \sum_i g_i\phi_i, \ \mathbf{u}' = \mathbf{u}^* + \sum_i g_i'\phi_i$$

where g_i are the *scaling variables* that characterise the properties of the parameter space near the fixed point. An eigenvalue can be expressed as a power of the rescaling factor b, as

$$\lambda_i(b) = b^{y_i}$$

where the exponent y_i characterises the parameter flow near the fixed point. Remarkably, critical phenomena can be determined from these exponents and the scaling variables.

In the vicinity of the fixed point $\mathbf{u}*$, where the linearised RG theory is valid, the nature of the parameter space is characterised by the scaling variables g_i, as well as the exponents y_i (associated with the eigenvalues λ_i). In relation to the local axes given by ϕ_i, the scaling variables identify certain directions of the parameter space as relevant for its critical behaviour: namely, the directions of the eigenvectors ϕ_i for which $y_i > 0$. The scaling variables in these directions are the *relevant* variables. The scaling variables with negative exponents ($y_i < 0$) are *irrelevant*, and the variables with $y_i = 0$ are *marginal*. *Critical* fixed points must have both relevant and irrelevant scaling variables in order to be associated with phase-transitions and critical points: the critical manifold is spanned by the eigenvectors associated with the irrelevant scaling variables, forming the basin of attraction of the fixed point. The relevant scaling variables can in turn be identified with the control parameters, the tuning of which is relevant to critical phenomena (e.g. t and h in magnetic systems).

This distinction between relevant and irrelevant scaling variables delineates the explanatorily critical features of non-renormalised, physical Hamiltonians – the features on which the critical exponents depend. The critical exponents depend only on the relevant variables, in the following sense. In the vicinity of the fixed point the relevance of these variables amounts to the fact that the RG flow veers away from the fixed point, in directions that are "orthogonal" to the critical manifold. A universality class depends on the way in which RG flows from different starting points (different physical Hamiltonians) are similar in this way. (See e.g. Pathria and Beale, 2009. p. 436 ff.) The irrelevance of the irrelevant variables, on the other hand, amounts to the fact that in the vicinity of the fixed point changes in the irrelevant variables do not determine a system's scaling properties under coarse-graining; only changes in the relevant variables matter. The only explanatory relevance that we can associate with the irrelevant variables turns on the fact that systems that differ only in the irrelevant variables, within a given universality class, are similar the sense that their coarse-grained descriptions asymptotically coincide: if the relevant scaling variables vanish – viz. if the relevant control parameters are tuned to criticality – then any change in the remaining, irrelevant parameters is immaterial to the large-distance properties, since iterated renormalisation of any Hamiltonian on the critical manifold flows to the same critical fixed point.

References

Als-Nielsen, J. and R. J. Birgeneau (1977). Mean field theory, the Ginzburg criterion, and marginal dimensionality of phase transitions. *American Journal of Physics 45*, 554–560.

Baker, A. and M. Colyvan (2011). Indexing and mathematical explanation. *Philosophia Mathematica 19*(3), 323–334.

Batterman, R. W. (2000). Multiple realizability and universality. *The British Journal for the Philosophy of Science 51*(1), 115–145.

Batterman, R. W. (2002a). *The devil in the details: Asymptotic reasoning in explanation, reduction, and emergence.* New York: Oxford University Press.

Batterman, R. W. (2002b). Asymptotics and the role of minimal models. *The British Journal for the Philosophy of Science 53*(1), 21–38.

Batterman, R. W. (2010). Reduction and renormalization. In A. Hüttemann and G. Ernst (Eds.), *Time, Chance, and Reduction: Philosophical Aspects of Statistical Mechanics*, pp. 159–179. Cambridge: Cambridge University Press.

Batterman, R. W. (2011). Emergence, singularities, and symmetry breaking. *Foundations of Physics 41*(6), 1031–1050.

Batterman, R. W. (2015). Reduction and multiple realizability. Unpublished manuscript.

Batterman, R. W. and C. C. Rice (2014). Minimal model explanations. *Philosophy of Science 81*(3), 349–376.

Bokulich, A. (2008). Can classical structures explain quantum phenomena. *The British Journal for the Philosophy of Science 59*(2), 217–235.

Butterfield, J. (2011a). Emergence, reduction and supervenience: A varied landscape. *Foundations of Physics 41*(6), 920–959.

Butterfield, J. (2011b). Less is different: Emergence and reduction reconciled. *Found Phys 41*(6), 1065–1135.

Butterfield, J. (2014). Reduction, emergence, and renormalization. *The Journal of Philosophy 111*(1), 5–49.

Cannas, S. A. (1995). One-dimensional ising model with long-range interactions: a renormalization-group treatment. *Physical Review B 52*(5), 3034.

Cardy, J. (1996). *Scaling and renormalization in statistical physics*, Volume 5. Cambridge: Cambridge university press.

Earman, J. (2004). Curie's principle and spontaneous symmetry breaking. *International Studies in the Philosophy of Science 18*(2-3), 173–198.

Fisher, M. E. (1983). *Scaling, university and renormalization group theory*, Volume 186. Berlin: Springer.

Fisher, M. E. (1998). Renormalization group theory: Its basis and formulation in statistical physics. *Reviews of Modern Physics 70*(2), 653.

French, S. and J. Saatsi (Forthcoming). Symmetries and explanatory dependencies. In A. Reutlinger and J. Saatsi (Eds.), *Explanation Beyond Causation.* Oxford: Oxford University Press.

Frisch, M. (1998). *Theories, Models, and Explanation.* Ph. D. thesis.

Hempel, C. (1962). Explanation in Science and History. In R.G. Colodny (Ed.), *Frontiers of Science and Philosophy*, pp. 9–19. London: Allen and Unwin.

Hüttemann, A., R. Kuhn, and O. Terzidis (2015). Stability, emergence and part-whole reduction. In B. Falkenburg and M. Morrison (Eds.), *Why More Is Different: Philosophical Issues in Condensed Matter Physics and Complex Systems*, pp. 169–200. Springer.

Jansson, L. and J. Saatsi (Forthcoming). Explanatory abstractions. *British Journal for the Philosophy of Science.*

Kistler, M. (2013). The interventionist account of causation and non-causal association laws. *Erkenntnis 78*(1), 65–84.

Kopietz, P., L. Bartosch, and F. Schütz (2010). *Introduction to the functional renormalization group*, Volume 798. Dordrecht: Springer.

Lange, M. (2009). A tale of two vectors. *Dialectica 63*(4), 397–431.

Lange, M. (2013). What makes a scientific explanation distinctively mathematical. *The British Journal for the Philosophy of Science 64*(3), 485–511.

Lange, M. (2015). On “minimal model explanations”: A reply to Batterman and Rice. *Philosophy of Science 82*(2), 292–305.

McComb, W. D. (2004). *Renormalization methods: a guide for beginners.* Oxford: Oxford University Press.

Menon, T. and C. Callender (2013). Turn and face the strange ch-ch-changes: Philosophical questions raised by phase transitions. In R. Batterman (Ed.), *The Oxford Handbook of Philosophy of Physics*, pp. 189–223. Oxford: Oxford University Press.

Morrison, M. (2012). Emergent physics and micro-ontology. *Philosophy of Science 79*(1), 141–166.

Morrison, M. (2014). Complex systems and renormalization group explanations. *Philosophy of Science 81*(5), 1144–1156.

Morrison, M. (2015a). *Reconstructing reality: Models, mathematics, and simulations.* New York: Oxford University Press.

Morrison, M. (2015b). Why is more different? In B. Falkenburg and M. Morrison (Eds.), *Why More Is Different: Philosophical Issues in Condensed Matter Physics and Complex Systems*, pp. 91–114. Springer.

Nishimori, H. and G. Ortiz (2010). *Elements of phase transitions and critical phenomena.* Oxford: Oxford University Press.

Norton, J. D. (2012). Approximation and idealization: Why the difference matters. *Philosophy of Science*, 207–232.

Pathria, R., P. Beale, and D. (2009). *Statistical mechanics.* Boston: Academic Press.

Pexton, M. (2014). How dimensional analysis can explain. *Synthese 191*(10), 2333–2351.

Pincock, C. (2007). A Role for Mathematics in the Physical Sciences. *Noûs 41*, 253–75.

Psillos, S. (2011). Living with the abstract: realism and models. *Synthese 180*(1), 3–17.

Reutlinger, A. (2016). Is there a monist theory of causal and non-causal explanations? The counterfactual theory of scientific explanation. *Philosophy of Science 83*, 733–745.

Reutlinger, A. (2017a). Do renormalization group explanations conform to the commonality strategy? *Journal of General Philosophy of Science*, *143*, 143–150.

Reutlinger, A. (2017b). Are causal facts really explanatorily emergent? Ladyman and Ross on higher-level causal facts and renormalization group explanation. *Synthese 194*, 2291–2305.

Saatsi, J. (2016a). On explanations from geometry of motion. *The British Journal for the Philosophy of Science*, Advance Access, 2016. DOI:10.1093/bjps/axw007.

Saatsi, J. (2016b). On the 'indispensable explanatory role' of mathematics. *Mind*, 125, 1045–1070.

Saatsi, J. (2017). Dynamical systems theory and explanatory indispensability. *Philosophy of Science*, 84(5).

Saatsi, J. and M. Pexton (2013). Reassessing Woodward's account of explanation: regularities, counterfactuals, and noncausal explanations. *Philosophy of Science 80*(5), 613–624.

Salmon, W. (1989). *Four Decades of Scientific Explanation.* Minneapolis:University of Minnesota Press.

Sellars, W. (1963). *Science, perception, and reality.* Atascadero, CA: Ridgeview Publishing Company.

Sethna, J. (2006). *Statistical mechanics: entropy, order parameters, and complexity*, Volume 14. Oxford: Oxford University Press.

Stanley, H. (1999). Scaling, universality, and renormalization: Three pillars of modern critical phenomena. *Reviews of modern physics*, 71, S358.

Wilson, K. G. (1983). The renormalization group and critical phenomena. *Reviews of Modern Physics 55*(3), 583–600.

Woodward, J. (2003a). *Making things happen: A causal theory of explanation.* Oxford: Oxford University Press.

Woodward, J. (2003b). Experimentation, causal inference, and instrumental realism. In H. Radder (Ed.), *The Philosophy of Scientific Experimentation*, pp. 87–118. Pittsburgh: University of Pittsburgh Press.

Woodward, J. (Forthcoming). Toward a theory of non-causal explanation. In A. Reutlinger and J. Saatsi (Eds.), *Explanation Beyond Causation*. Oxford: Oxford University Press.

Zinn-Justin, J. (2006). Determination of critical exponents and equation of state by field theory methods. In P. H. Damgaard and J. Jurkiewicz (Eds.), *New developments in quantum field theory*, pp. 217–232. Dordrech: Springer.

Zinn-Justin, J. (2007). *Phase transitions and renormalization group*. Oxford: Oxford University Press.