CSRC Short Course on

Renormalization Group Methods and Applications

21-25 March, 2016

Conference Room II, Level 3, CSRC

	Mon, 21/3	Tues, 22/3	Wed, 23/3	Thur, 24/3	Fri, 25/3
8:30 -	Registration				
9:00 – 10:15	CSRC introduction (CSRC Central, Complex Sys)	Orland (phi-4 theory, epsilon expansion)	Chen Leiming (dynamic RG)	Chen Leiming (phase transition in incompressible flocks)	Conference on recent developments in nonequilibrium StatMech <u>Morning</u> a) Quantum quench and (pre)thermalization b) FDT violation in timescale separated systems <u>Afternoon</u> a) p-spin models (paper by Berthier and Kurchan, Nature Phys 2013) b) Hard sphere glass, colloids under shear, and active colloids, similarities and differences, phase diagram, scaling properties
10:15 – 10:45	tea break				
10:45 – 12:00	Tang (review of equilibrium StatMech, Legendre transforms, Ising model)	Orland (phi-4 theory, epsilon expansion)	Delamotte (NPRG)	Delamotte (NPRG)	
12:00 – 14:00	lunch break				
14:00 – 15:15	Lan Yueheng (nonlinear dynamics, phase space trajectory, fixed point analysis)	Delamotte (NPRG)	Delamotte (NPRG)	Delamotte (NPRG)	
15:15 – 15:45	tea break, discussions				
15:45 – 17:00	Special seminar by Guo Wenan (quantum criticality with two length scales)	Delamotte (NPRG)	Special seminar by Xing Xiangjun (Singular perturbations and the RG)	Group presentations	
18:30 – 21:00	Tutorials on generating fns, functional derivatives, gaussian integrals, etc., led by Lan Yueheng, Tang Leihan and Xing Xiangjun, project discussions				

14/3/16

For updated information, please visit

http://www.csrc.ac.cn/en/event/schools/2016-02-25/16.html

CSRC Short Course on

Renormalization Group Methods and Applications

21-25 March, 2016

List of lecture material (received 18 March 2016)

Lecturer	Topic	File
Lectures/tutorials		
	General background reading	ME_Fisher_98.pdf Kadanoff_Essay.pdf fourier.pdf
Leihan Tang	Phase transitions and critical phenomena, Ising model	LHT_scaling.pdf Delamotte_tutorial_material.pdf
Yueheng Lan	Nonlinear dynamics, phase space trajectory, fixed point analysis	Yueheng_Lan_renormimpv.pdf
Henri Orland	ϕ^4 theory, epsilon expansion	LHT_RG_notes.pdf
Bertrand Delamotte	Nonperturbative RG	Delamotte_cours2014-6lecture.pdf Delamotte_tutorial_material.pdf
Leiming Chen	Dynamic RG: Forster, Stephen and Nelson, PRA **16**, 732 (1977)	FSN_PhysRevA.16.732.pdf
	Phase transition in incompressible flocks	Leiming_Chen_NJP.pdf Leiming_Chen_SM_NJP.pdf
Special guest seminars		
Wenan Guo	Quantum criticality with two length scales	Wenan_Guo_Science_paper.pdf
Xiangjun Xing	Singular perturbations and the RG	Xiangjun_Xing_notes-ODE-RG.pdf

First course on StatMech: http://pages.physics.cornell.edu/~sethna/StatMech/

Renormalization group theory: Its basis and formulation in statistical physics*

Michael E. Fisher

Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742

The nature and origins of renormalization group ideas in statistical physics and condensed matter theory are recounted informally, emphasizing those features of prime importance in these areas of science in contradistinction to quantum field theory, in particular: critical exponents and scaling, relevance, irrelevance and marginality, universality, and Wilson's crucial concept of flows and fixed points in a large space of Hamiltonians.

CONTENTS

FOREWORD

''*In March 1996 the Departments of Philosophy and of Physics at Boston University cosponsored a Colloquium 'On the Foundations of Quantum Field Theory.' But in the full title, this was preceded by the phrase 'A Historical Examination and Philosophical Reflections,' which set the aims of the meeting. The participants were mainly high-energy physicists, experts in field theories, and interested philosophers of science.*[1] *I was called on to speak, essentially in a service role, presumably because I had witnessed and had some hand in the development of renormalization group concepts and because I have played a role in applications where these ideas really mattered. It is hoped that this article, based on the talk I presented in Boston, may prove of interest to a wider audience.*''

*Based on a lecture presented on 2 March 1996 at the Boston Colloquium for the Philosophy of Science: ''A Historical Examination and Philosophical Reflections on the Foundations of Quantum Field Theory,'' held at Boston University 1–3 March 1996.

[1]The proceedings of the conference are to be published under the title *Conceptual Foundations of Quantum Field Theory* (Cao, 1998): for details see the references collected in the Selected Bibliography.

I. INTRODUCTION

It is held by some that the ''Renormalization Group''—or, better, renormalization group*s* or, let us say, *Renormalization Group Theory* (or RGT) is ''one of the underlying ideas in the theoretical structure of Quantum Field Theory.'' That belief suggests the potential value of a historical and conceptual account of RG theory and the ideas and sources from which it grew, as viewed from the perspective of statistical mechanics and condensed matter physics. Especially pertinent are the roots in the theory of critical phenomena.

The proposition just stated regarding the significance of RG theory for Quantum Field Theory (or QFT, for short) is open to debate even though experts in QFT have certainly invoked RG ideas. Indeed, one may ask: How far is some concept only instrumental? How far is it crucial? It is surely true in physics that when we have ideas and pictures that are extremely useful, they acquire elements of reality in and of themselves. But, philosophically, it is instructive to look at the degree to which such objects are purely instrumental—merely useful tools—and the extent to which physicists seriously suppose they embody an essence of reality. Certainly, many parts of physics are well established and long precede RG ideas. Among these is statistical mechanics itself, a theory *not* reduced and, in a deep sense, *not* directly reducible to lower, more fundamental levels without the introduction of specific, new postulates.

Furthermore, statistical mechanics has reached a stage where it is well posed mathematically; many of the basic theorems (although by no means all) have been proved with full rigor. In that context, I believe it is possible to view the renormalization group as merely an instrument or a computational device. On the other hand, at one extreme, one might say: ''Well, the partition function itself is really just a combinatorial device.'' But most practitioners tend to think of it (and especially its logarithm, the free energy) as rather more basic!

Now my aim here is not to instruct those field theorists who understand these matters well.[2] Rather, I hope to convey to nonexperts and, in particular, to any with a philosophical interest, a little more about what Renor-

[2]Such as D. Gross and R. Shankar (see Cao, 1998, and Shankar, 1994). Note also Bagnuls and Bervillier (1997).

 0034-6861/98/70(2)/653(29)/$20.80

malization Group Theory is[3]—at least in the eyes of some of those who have earned a living by using it! One hopes such information may be useful to those who might want to discuss its implications and significance or assess how it fits into physics more broadly or into QFT in particular.

II. WHENCE CAME RENORMALIZATION GROUP THEORY?

This is a good question to start with: I will try to respond, sketching the foundations of RG theory in the *critical exponent relations* and crucial *scaling concepts*[4] of Leo P. Kadanoff, Benjamin Widom, and myself developed in 1963–66[5]—among, of course, other important workers, particularly Cyril Domb[6] and his group at King's College London, of which, originally, I was a member, George A. Baker, Jr., whose introduction of Padé approximant techniques proved so fruitful in gaining quantitative knowledge,[7] and Valeri L. Pokrovskii and A. Z. Patashinskii in the Soviet Union who were, perhaps, the first to bring field-theoretic perspectives to bear.[8] Especially, of course, I will say something of the genesis of the full RG concept—the systematic integrating out of appropriate degrees of freedom and the resulting RG flows—in the inspired work of Kenneth G. Wilson[9] as I saw it when he was a colleague of mine and Ben Widom's at Cornell University in 1965–1972. And I must point also to the general, clarifying formulation of RG theory by Franz J. Wegner (1972a) when he was associated with Leo Kadanoff at Brown University: their focus on *relevant, irrelevant* and *marginal* 'operators' (or perturbations) has played a central role.[10]

But, if one takes a step back, two earlier, fundamental theoretical achievements must be recognized: the first is the work of Lev D. Landau, who in reality, is the founder of systematic *effective field theories*, even though he might not have put it that way. It is Landau's *invention*—as it may, I feel, be fairly called—of the *order parameter* that is so important but often underappreciated.[11] To assert that there exists an order parameter in essence says: "I may not understand the microscopic phenomena at all" (as was historically, the case for superfluid helium), "but I recognize that there is a microscopic level and I believe it should have certain general, overall properties especially as regards locality and symmetry: those then serve to govern the most characteristic behavior on scales greater than atomic." Landau and Ginzburg (a major collaborator and developer of the concept[12]) misjudged one or two of the important general properties, in particular the role of fluctuations and singularity; but that does not alter the deep significance of this way of looking at a complex, condensed matter system. Know the nature of the order parameter—suppose, for example, it is a complex number and like a wave function—then one knows much about the macroscopic nature of a physical system!

Significantly, in my view, Landau's introduction of the order parameter exposed a novel and unexpected *foliation* or level in our understanding of the physical world. Traditionally, one characterizes statistical mechanics as directly linking the *microscopic* world of nuclei and atoms (on length scales of 10^{-13} to 10^{-8} cm) to the *macroscopic* world of say, millimeters to meters. But the order parameter, as a dynamic, fluctuating object in many cases intervenes on an intermediate or *mesoscopic* level characterized by scales of tens or hundreds of angstroms up to microns (say, $10^{-6.5}$ to $10^{-3.5}$ cm). The advent of Wilson's concept of the renormalization group gave more precise meaning to the effective ("coarse-grained") Hamiltonians that stemmed from the work of Landau and Ginzburg. One now pictures the LGW—for Landau-Ginzburg-Wilson—Hamiltonians as true but significantly renormalized Hamiltonians in which finer microscopic degrees of freedom have been integrated-out. (See below for more concrete and explicit expressions.) Frequently, indeed, in modern condensed matter theory one *starts* from this intermediate level with a physically appropriate LGW Hamiltonian *in place* of a true (or, at least, more faithful or realistic) microscopic Hamiltonian; and *then* one brings statistical mechanics

[3]It is worthwhile to stress, at the outset, what a "renormalization group" is *not*! Although in many applications the particular renormalization group employed may be invertible, and so constitute a continuous or discrete, group of transformations, it is, in general, only a *semigroup*. In other words a renormalization group is not necessarily invertible and, hence, cannot be 'run backwards' without ambiguity: in short it is *not* a "group." The misuse of mathematical terminology may be tolerated since these aspects play, at best, a small role in RG theory. The point will be returned to in Secs. VIII and XI.

[4]Five influential reviews antedating renormalization-group concepts are Domb (1960), Fisher (1965, 1967b), Kadanoff *et al.* (1967) and Stanley (1971). Early reviews of renormalization group developments are provided by Wilson and Kogut (1974b) and Fisher (1974): see also Wilson (1983) and Fisher (1983). The first texts are Pfeuty and Toulouse (1975), Ma (1976), and Patashinskii and Pokrovskii (1979). The books by Baker (1990), Creswick *et al.* (1992), and Domb (1996) present retrospective views.

[5]See Essam and Fisher (1963), Widom (1965a, 1965b), Kadanoff (1966), and Fisher (1967a).

[6]Note Domb (1960), Domb and Hunter (1965), and the account in Domb (1996).

[7]See Baker (1961) and the overview in Baker (1990).

[8]The original paper is Patashinskii and Pokrovskii (1966); their text (1979), which includes a chapter on RG theory, appeared in Russian around 1975 but did not then discuss RG theory.

[9]Wilson (1971a, 1971b), described within the QFT context in Wilson (1983).

[10]Note the reviews by Kadanoff (1976) and Wegner (1976).

[11]See Landau and Lifshitz (1958) especially Sec. 135.

[12]In particular for the theory of superconductivity: see V. L. Ginzburg and L. D. Landau, 1959, "On the Theory of Superconductivity," Zh. Eksp. Teor. Fiz. **20**, 1064; and, for a personal historical account, V. L. Ginzburg, 1997, "Superconductivity and Superfluidity (What was done and what was not)," Phys. Usp. **40**, 407–432.

to bear in order to understand the macroscopic level. The derivation and validity of the many types of initial, LGW Hamiltonians may then be the object of separate studies to relate them to the atomic level.[13]

Landau's concept of the order parameter, indeed, brought light, clarity, and form to the general theory of phase transitions, leading eventually, to the characterization of multicritical points and the understanding of many characteristic features of ordered states.[14] But in 1944 a bombshell struck! Lars Onsager, by a mathematical *tour de force*, deeply admired by Landau himself,[15] computed exactly the partition function and thermodynamic properties of the simplest model of a ferromagnet or a fluid.[16] This model, the *Ising model*, exhibited a sharp critical point: but the explicit properties, in particular, the nature of the critical singularities disagreed profoundly—as I will explain below—with essentially all the detailed predictions of the Landau theory (and of all foregoing, more specific theories). From this challenge, and from experimental evidence pointing in the same direction,[17] grew the ideas of *universal* but nontrivial *critical exponents*,[18] special *relations* between different *exponents*,[19] and then, *scaling descriptions* of the region of a critical point.[20] These insights served as stimulus and inspiration to Kenneth Wilson in his pursuit of an understanding of quantum field theories.[21] Indeed, once one understood the close mathematical analogy between doing statistical mechanics with effective Hamiltonians and doing quantum field theory (especially with the aid of Feynman's path integral) the connections seemed almost obvious. Needless to say, however, the realization of the analogy did not come overnight: in fact, Wilson himself was, in my estimation, the individual who first understood clearly the analogies at the deepest levels. And they are being exploited, to mutual benefit to this day.

In 1971, then, Ken Wilson, having struggled with the problem for four or five years,[22] was able to cast his renormalization group ideas into a conceptually effective framework—effective in the sense that one could do certain calculations with it.[23] And Franz Wegner, very soon afterwards,[24] further clarified the foundations and exposed their depth and breadth. An early paper by Kadanoff and Wegner (1971) showing when and how universality could *fail* was particularly significant in demonstrating the richness of Wilson's conception.

So our understanding of "anomalous," i.e., nonLandau-type but, in reality, standard critical behaviour was greatly enhanced. And let me stress that my *personal aim* as a theorist is to *gain* insight and understanding: What that may truly mean is, probably, a matter for deep philosophical review: After all, "What constitutes an *explanation*?" But, on the other hand, if you work as a theoretical physicist in the United States, and wish to publish in The Physical Review, you had better *calculate* something concrete and interesting with your new theory pretty soon! For *that* purpose, the *epsilon expansion*, which used as a small, perturbation parameter the deviation of the spatial dimensionality, d, from four dimensions, namely, $\epsilon=4-d$, provided a powerful and timely tool.[25] It had the added advantage, if one wanted to move ahead, that the method looked something like a cookbook—so that "any fool" could do or check the calculations, whether they really understood, at a deeper level, what they were doing or not! But in practice that also has a real benefit in that a lot of calculations do get done, and some of them turn up new and interesting things or answer old or new questions in instructive ways. A few calculations reveal apparent paradoxes and problems which serve to teach one and advance understanding since, as Arthur Wightman has observed, one asks: "Maybe we should go back and think more carefully about what we are actually doing in implementing the theoretical ideas?" So that, in outline, is what I want to convey in more detail, in this exposition.

III. WHERE STANDS THE RENORMALIZATION GROUP?

Beyond sketching the origins, it is the breadth and generality of RG theory that I wish to stress. Let me, indeed, say immediately that the full RG theory should no more be regarded as based on QFT perturbative expansions—despite that common claim—than can the magnificent structure of Gibbsian statistical mechanics be viewed as founded upon ideal classical gases, Boltzmannian kinetic theory, and the virial and cluster expansions for dilute fluids! True, this last route was still frequently retravelled in textbooks more than 50 years after Gibbs' major works were published; but it deeply misrepresents the power and range of statistical mechanics.

The parallel mischaracterizations of RG theory may be found, for example, in the much cited book by Daniel Amit (1978), or in Chapter 5 of the later text on *Statis-*

[13]These issues have been discussed further by the author in "Condensed Matter Physics: Does Quantum Mechanics Matter?" in *Niels Bohr: Physics and the World*, edited by H. Feshbach, T. Matsui and A. Oleson, 1988 (Harwood Academic, Chur), pp. 177–183.

[14]See Landau and Lifshitz (1958).

[15]As I know by independent personal communications from Valeri Pokrovskii and from Isaak M. Khalatnikov.

[16]Onsager (1944), Kaufman and Onsager (1949), Onsager (1949).

[17]See, e.g. Fisher (1965), Stanley (1971).

[18]Domb (1960, 1996) was the principal pioneer in the identification and estimation of critical exponents: see also the preface to Domb (1996) by the present author.

[19]Advanced particularly in Essam and Fisher (1963).

[20]Widom (1965a, 1965b), Domb and Hunter (1965), Kadanoff (1966), and Patashinskii and Pokrovskii (1966).

[21]Wilson (1971a, 1971b; 1983).

[22]See below and the account in Wilson (1983).

[23]As we will explain: see Wilson (1971a, 1971b).

[24]Wegner (1972a, 1972b).

[25]Wilson and Fisher (1972).

tical Field Theory by Itzykson and Drouffe (1989), or, more recently, in the lecture notes entitled *Renormalization Group* by Benfatto and Gallavotti (1995), "dedicated to scholars wishing to reflect on some details of the foundations of the modern renormalization group approach." There we read that the authors aim to expose how the RG looks to them as physicists, namely: "this means the achievement of a coherent perturbation theory based on second order (or lowest-order) calculations." One cannot accept that! It is analogous to asking "What does statistical mechanics convey to a physicist?" and replying: "It means that one can compute the second-virial coefficient to correct the ideal gas laws!" Of course, historically, that is not a totally irrelevant remark; but it is extremely misleading and, in effect, insults one of America's greatest theoretical physicists, Josiah Willard Gibbs.

To continue to use Benfatto and Gallavotti as strawmen, we find in their preface that the reader is presumed to have "some familiarity with classical quantum field theory." That surely, gives one the impression that, somehow, QFT is necessary for RG theory. Well, it is totally unnecessary![26] And, in particular, by implication the suggestion overlooks entirely the so-called "real space RG" techniques,[27] the significant Monte Carlo RG calculations,[28] the use of *functional* RG methods,[29] etc. On the other hand, if one wants to do certain types of calculation, then familiarity with quantum field theory and Feynmann diagrams can be very useful. But there is *no necessity*, even though many books that claim to tell one about renormalization group theory give that impression.

I do not want to be unfair to Giovanni Gallavotti, on whose lectures the published notes are based: his book is insightful, stimulating and, accepting his perspective as a mathematical physicist[30] keenly interested in field theory, it is authoritative. Furthermore, it forthrightly acknowledges the breadth of the RG approach citing as examples of problems implicitly or explicitly treated by RG theory:[31]

(i) The KAM (Kolmogorov-Arnold-Moser) theory of Hamiltonian stability
(ii) The constructive theory of Euclidean fields
(iii) Universality theory of the critical point in statistical mechanics
(iv) Onset of chaotic motions in dynamical systems (which includes Feigenbaum's period-doubling cascades)
(v) The convergence of Fourier series on a circle
(vi) The theory of the Fermi surface in Fermi liquids (as described by Shankar (1994; and in Cao, 1998))

To this list one might well add:

(vii) The theory of polymers in solutions and in melts
(viii) Derivation of the Navier-Stoker equations for hydrodynamics
(ix) The fluctuations of membranes and interfaces
(x) The existence and properties of 'critical phases' (such as superfluid and liquid-crystal films)
(xi) Phenomena in random systems, fluid percolation, electron localization, etc.
(xii) The Kondo problem for magnetic impurities in nonmagnetic metals.

This last problem, incidentally, was widely advertised as a significant, major issue in solid state physics. However, when Wilson solved it by a highly innovative, *numerical* RG technique[32] he was given surprisingly little credit by that community. It is worth noting Wilson's own assessment of his achievement: "This is the most exciting aspect of the renormalization group, the part of the theory that makes it possible to solve problems which are unreachable by Feynman diagrams. The Kondo problem has been solved by a nondiagrammatic computer method."

Earlier in this same passage, written in 1975, Wilson roughly but very usefully divides RG theory into four parts: (a) the formal theory of fixed points and linear and nonlinear behavior near fixed points where he especially cites Wegner (1972a, 1976), as did I, above; (b) the diagrammatic (or field-theoretic) formulation of the RG for critical phenomena[33] where the ϵ expansion[34] and its many variants[35] plays a central role; (c) QFT methods, including the 1970–71 Callan-Symanzik equations[36] and the original, 1954 Gell-Mann-Low RG theory—restricted to systems with only a single, marginal

[26]See, e.g., Fisher (1974, 1983), Creswick, Farach, and Poole (1992), and Domb (1996).

[27]See the reviews in Niemeijer and van Leeuwen (1976), Burkhardt and van Leeuwen (1982).

[28]Pioneered by Ma (1976) and reviewed in Burkhardt and van Leeuwen (1982). For a large scale calculation, see: Pawley, Swendsen, Wallace, and Wilson (1984).

[29]For a striking application see: Fisher and Huse (1985).

[30]The uninitiated should note that for a decade or two the term 'mathematical physicist' has meant a theorist who provides *rigorous proofs* of his main results. For an account of the use of the renormalization group *in* rigorous work in mathematical physics, see Gawędski (1986).

[31]Benfatto and Gallavotti (1995), Chap. 1.

[32]Wilson (1975); for the following quotation see page 776, column 1.

[33]Wilson (1972), Brézin, Wallace, and Wilson (1972), Wilson and Kogut (1974), Brézin, Le Guillou and Zinn-Justin (1976).

[34]Wilson and Fisher (1972), Fisher and Pfeuty (1972).

[35]Especial mention should be made of $1/n$ expansions, where n is the number of components of the vector order parameter (Abe, 1972, 1973; Fisher, Ma, and Nickel, 1972; Suzuki, 1972; and see Fisher, 1974, and Ma, 1976a) and of coupling-constant expansions in fixed dimension: see Parisi (1973, 1974); Baker, Nickel, Green, and Meiron (1976); Le Guillou and Zinn-Justin (1977); Baker, Nickel, and Meiron (1978): For other problems, dimensionality expansions have been made by writing $d=8-\epsilon$, $6-\epsilon$, $4+\frac{1}{2}m-\epsilon$ $(m=1, 2, \cdots)$, $3-\epsilon$, $2+\epsilon$, and $1+\epsilon$.

[36]The Callan-Symanzik equations are described, e.g., in Amit (1978) and Itzykson and Drouffe (1989). The coupling-constant expansions in fixed dimension (Parisi, 1973, 1974; Baker *et al.*, 1976) typically use these equations as a starting point and are usually presented purely formally in contrast to the full Wilson approach (b).

variable[37]—from which Wilson drew some of his inspiration and which he took to name the whole approach.[38] Wilson characterizes these methods as efficient calculationally—which is certainly the case—but applying only to Feynman diagram expansions and says: ''They completely hide the physics of many scales.'' Indeed, from the perspective of condensed matter physics, as I will try to explain below, the chief drawback of the sophisticated field-theoretic techniques is that they are safely applicable only when the basic physics is already well understood. By contrast, the general formulation (a), and Wilson's approach (b), provide insight and understanding into quite fresh problems.

Finally, Wilson highlights (d) ''the construction of nondiagrammatic RG transformations, which are then solved numerically.'' This includes the real-space, Monte Carlo, and functional RG approaches cited above and, of course, Wilson's own brilliant application to the Kondo problem (1975).

IV. EXPONENTS, ANOMALOUS DIMENSIONS, SCALE INVARIANCE AND SCALE DEPENDENCE

If one is to pick out a single feature that epitomizes the power and successes of RG theory, one can but endorse Gallavotti and Benfatto when they say ''it has to be stressed that the *possibility of nonclassical critical indices* (i.e., *of nonzero anomaly* η) *is probably the most important achievement of the renormalization group*.'' [39] For nonexperts it seems worthwhile to spend a little time here explaining the meaning of this remark in more detail and commenting on a few of the specialist terms that have already arisen in this account.

To that end, consider a locally defined microscopic variable which I will denote $\psi(\mathbf{r})$. In a ferromagnet this might well be the local magnetization, $\vec{M}(\mathbf{r})$, or spin vector, $\vec{S}(\mathbf{r})$, at point $\mathbf{r}$ in ordinary d-dimensional (Euclidean) space; in a fluid it might be the deviation $\delta\rho(\mathbf{r})$, of the fluctuating density at $\mathbf{r}$ from the mean density. In QFT the local variables $\psi(\mathbf{r})$ are the basic *quantum fields* which are 'operator valued.' For a magnetic system, in which quantum mechanics was important, $\vec{M}(\mathbf{r})$ and $\vec{S}(\mathbf{r})$ would, likewise, be operators. However, the distinction is of relatively minor importance so that we may, for ease, suppose $\psi(\mathbf{r})$ is a simple classical variable. It will be most interesting when ψ is closely related to the order parameter for the phase transition and critical behavior of concern.

By means of a scattering experiment (using light, x rays, neutrons, electrons, etc.) one can often observe the corresponding *pair correlation function* (or basic 'two-point function')

$$G(\mathbf{r}) = \langle \psi(\mathbf{0})\psi(\mathbf{r})\rangle, \tag{1}$$

where the angular brackets $\langle\cdot\rangle$ denote a statistical average over the thermal fluctuations that characterize all equilibrium systems at nonzero temperature. (Also understood, when $\psi(\mathbf{r})$ is an operator, are the corresponding quantum-mechanical expectation values.)

Physically, $G(\mathbf{r})$ is important since it provides a direct measure of the influence of the leading microscopic fluctuations at the origin $\mathbf{0}$ on the behavior at a point distance $r=|\mathbf{r}|$ away. But, almost by definition, in the vicinity of an appropriate critical point—for example the Curie point of a ferromagnet when $\psi \equiv \vec{M}$ or the gas-liquid critical point when $\psi = \delta\rho$—a strong ''ordering'' influence or correlation spreads out over, essentially, macroscopic distances. As a consequence, precisely *at* criticality one rather generally finds a *power-law decay*, namely,

$$G_c(\mathbf{r}) \approx D/r^{d-2+\eta} \quad \text{as} \quad r\to\infty, \tag{2}$$

which is characterized by the *critical exponent* (or critical *index*) $d-2+\eta$.

Now all the theories one first encounters—the so-called 'classical' or Landau-Ginzburg or van der Waals theories, etc.[40]—predict, quite unequivocally, that η *vanishes*. In QFT this corresponds to the behavior of a free massless particle. Mathematically, the reason underlying this prediction is that the basic functions entering the theory have (or are assumed to have) a smooth, *analytic, nonsingular* character so that, following Newton, they may be freely differentiated and, thereby expanded in Taylor series with positive integral powers[41] even at the critical point. In QFT the classical exponent value $d-2$ (implying $\eta=0$) can often be determined by naive dimensional analysis or 'power counting': then $d-2$ is said to represent the 'canonical dimension' while η, if nonvanishing, represents the 'dimensional anomaly.' Physically, the prediction $\eta=0$ typically results from a neglect of fluctuations or, more precisely as Wilson emphasized, from the assumption that only fluctuations on much smaller scales can play a significant role: in such circumstances the fluctuations can be safely incorporated into *effective* (or *renormalized*) parameters (masses, coupling constants, etc.) with no change in the basic character of the theory.

---

[37]See Wilson (1975), page 796, column 1. The concept of a ''marginal'' variable is explained briefly below: see also Wegner (1972a, 1976), Fisher (1974, 1983), and Kadanoff (1976).

[38]See Wilson (1975, 1983).

[39]See Benfatto and Gallavotti (1995) page 64.

[40]Note that 'classical' here, and in the quote from Benfatto and Gallavotti above means 'in the sense of the ancient authors'; in particular, it is *not* used in contradistinction to 'quantal' or to allude in any way to quantum mechanics (which has essentially no relevance for critical points at nonzero temperature: see the author's article cited in Footnote 13).

[41]The relevant expansion variable in scattering experiments is the square of the scattering wave vector, $\mathbf{k}$, which is proportional to $\lambda^{-1}\sin\frac{1}{2}\theta$ where θ is the scattering angle and λ the wavelength of the radiation used. In the description of near-critical thermodynamics, Landau theory assumes (and mean-field theories lead to) Taylor expansions in powers of $T-T_c$ and $\Psi=\langle\Psi(\mathbf{r})\rangle$, the equilibrium value of the order parameter.

But a power-law dependence on distance implies a *lack* of a definite length scale and, hence, a *scale invariance*. To illustrate this, let us rescale distances by a factor b so that

$$\mathbf{r} \Rightarrow \mathbf{r}' = b\mathbf{r}, \tag{3}$$

and, at the same time, rescale the order parameter ψ by some "covariant" factor b^{ω} where ω will be a critical exponent characterizing ψ. Then we have

$$G_c(\mathbf{r}) = \langle \psi(\mathbf{0})\psi(\mathbf{r})\rangle_c \Rightarrow G_c'(b\mathbf{r}) = b^{2\omega}\langle \psi(\mathbf{0})\psi(b\mathbf{r})\rangle_c \approx b^{2\omega} D/b^{d-2+\eta} r^{d-2+\eta}. \tag{4}$$

Now, observe that if one has $\omega = \frac{1}{2}(d-2+\eta)$, the factors of b drop out and the form in Eq. (2) is recaptured. In other words $G_c(\mathbf{r})$ is *scale invariant* (or covariant): its variation reveals no characteristic lengths, large, small, or intermediate!

Since power laws imply scale invariance and the *absence* of well separated scales, the classical theories should be suspect at (and near) criticality! Indeed, one finds that the "anomaly" η does *not* normally vanish (at least for dimensions d less than 4, which is the only concern in a condensed matter laboratory!). In particular, from the work of Kaufman and Onsager (1949) one can show analytically that $\eta = \frac{1}{4}$ for the $d=2$ Ising model.[42] Consequently, the analyticity and Taylor expansions presupposed in the classical theories are *not* valid.[43] Therein lies the challenge to theory! Indeed, it proved hard even to envisage the nature of a theory that would lead to $\eta \neq 0$. The power of the renormalization group is that it provides a conceptual and, in many cases, a computational framework within which anomalous values for η (and for other exponents like ω and its analogs for all local quantities such as the energy density) arise naturally.

In applications to condensed matter physics, it is clear that the power law in Eq. (2) can hold only for distances relatively large compared to atomic lengths or lattice spacings which we will denote a. In this sense the scale invariance of correlation functions is only *asymptotic*—hence the symbol $\approx$, for "asymptotically equals,"[44] and the proviso $r \to \infty$ in Eq. (2). A more detailed description would account for the effects of nonvanishing a, at least in leading order. By contrast, in QFT the microscopic distance a represents an "ultraviolet" cutoff which, since it is in general unknown, one normally wishes to remove from the theory. If this removal is not done with surgical care—which is what the renormalization program in QFT is all about—the theory remains plagued with infinite divergencies arising when $a \to 0$, i.e., when the "cutoff is removed." But in statistical physics one always anticipates a short-distance cutoff that sets certain physical parameters such as the value of T_c; infinite terms *per se* do not arise and certainly do *not* drive the theory as in QFT.

In current descriptions of QFT the concept of the *scale-dependence of parameters* is often used with the physical picture that the typical properties of a system measured at particular length (and/or time) scales change, more-or-less slowly, as the scale of observation changes. From my perspective this phraseology often represents merely a shorthand for a somewhat simplified view of RG *flows* (as discussed generally below) in which only one variable or a single trajectory is followed,[45] basically because one is interested only in one, unique theory—the real world of particle physics. In certain condensed matter problems something analogous may suffice or serve in a first attack; but in general a more complex view is imperative.

One may, however, provide a more concrete illustration of scale dependence by referring again to the power law Eq. (2). If the exponent η vanishes, or equivalently, if ψ has its canonical dimension, so that $\omega = \omega_{\rm can} = \frac{1}{2}(d-2)$, one may regard the amplitude D as a fixed, measurable parameter which will typically embody some real physical significance. Suppose, however, η does *not* vanish but is nonetheless relatively *small*: indeed, for many ($d=3$)-dimensional systems, one has $\eta \simeq 0.035$.[46] Then we can introduce a "renormalized" or "scale-dependent" parameter

$$\widetilde{D}(R) \approx D/R^{\eta} \quad \text{as} \quad R \to \infty, \tag{5}$$

and rewrite the original result simply as

$$G_c(r) = \widetilde{D}(r)/r^{d-2}. \tag{6}$$

Since η is small we see that $\widetilde{D}(R)$ varies slowly with the scale R on which it is measured. In many cases in QFT the dimensions of the field ψ (*alias* the order parameter) are subject only to marginal perturbations (see below) which translate into a $\log R$ dependence of the renormalized parameter $\widetilde{D}(R)$; the variation with scale is then still weaker than when $\eta \neq 0$.

V. THE CHALLENGES POSED BY CRITICAL PHENOMENA

It is good to remember, especially when discussing theory and philosophy, that physics is an experimental science! Accordingly, I will review briefly a few experimental findings[47] that serve to focus attention on the

[42] Fisher (1959): see also Fisher (1965, Sec. 29; 1967b, Sec. 6.2), Fisher and Burford (1967).

[43] Precisely the same problem undermines applications of catastrophe theory to critical phenomena; the assumed expressions in powers of $(T-T_c)$ and $\Psi = \langle\psi\rangle$ are simply not valid.

[44] See the Appendix for a discussion of appropriate conventions for the symbols $\simeq$, $\approx$, and $\sim$.

[45] See below and, e.g., Wilson and Kogut (1974), Bagnuls and Bervillier (1997).

[46] See, e.g., Fisher and Burford (1967), Fisher (1983), Baker (1990), and Domb (1996).

[47] Ideally, I should show here plots of impressive experimental data and, in particular, dramatic color pictures of carbon dioxide passing though its critical point. [See Stanley (1971) for black and white photographs.] It is not, however, feasible to reproduce such figures here; instead the presentation focuses on the conclusions as embodied in the observed power laws, etc.

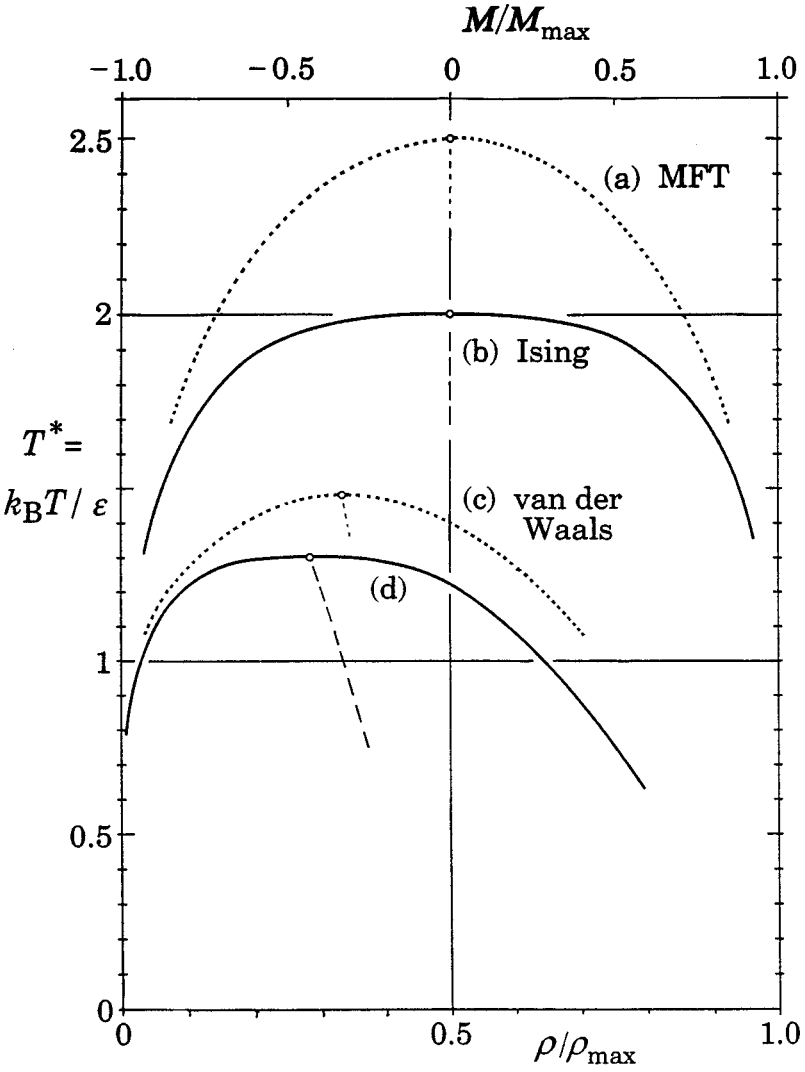


FIG. 1. Temperature variation of gas-liquid coexistence curves (temperature, T, *versus* density, ρ) and corresponding spontaneous magnetization plots (magnetization, M, *versus* T). The solid curves, (b) and (d), represent (semiquantitatively) observation and modern theory, while the dotted curves (a) and (c) illustrate the corresponding "classical" predictions (mean-field theory and van der Waals approximation). These latter plots are parabolic through the critical points (small open circles) instead of obeying a power law with the universal exponent $\beta \simeq 0.325$: see Eqs. (9) and (11). The energy scale ε, and the maximal density and magnetization, ρ_{max} and M_{max}, are nonuniversal parameters particular to each physical system; they vary widely in magnitude.

principal theoretical challenges faced by, and rather fully met by RG theory.

In 1869 Andrews reported to the Royal Society his observations of carbon dioxide sealed in a (strong!) glass tube at a mean overall density, ρ, close to 0.5 gm cm^{-3}. At room temperatures the fluid breaks into two phases: a liquid of density $\rho_{liq}(T)$ that coexists with a lighter vapor or gas phase of density $\rho_{gas}(T)$ from which it is separated by a visible meniscus or interface; but when the temperature, T, is raised and reaches a sharp critical temperature, $T_c \simeq 31.04\,^\circ$C, the liquid and gaseous phases become identical, assuming a common density $\rho_{liq} = \rho_{gas} = \rho_c$ while the meniscus disappears in a "mist" of "critical opalescence." For all T above T_c there is a complete "continuity of state," i.e., no distinction whatsoever remains between liquid and gas (and there is no meniscus). A plot of $\rho_{liq}(T)$ and $\rho_{gas}(T)$—as illustrated somewhat schematically in Fig. 1(d)—represents the so-called gas-liquid *coexistence curve*: the two halves, $\rho_{liq} > \rho_c$ and $\rho_{gas} < \rho_c$, meet smoothly at the *critical point* (T_c, ρ_c)—shown as a small circle in Fig. 1: the dashed line below T_c represents the *diameter* defined by $\overline{\rho(T)} = \frac{1}{2}[\rho_{liq}(T) + \rho_{gas}(T)]$.

The same phenomena occur in all elemental and simple molecular fluids and in fluid mixtures. The values of T_c, however, vary widely: e.g., for helium-four one finds 5.20 K while for mercury $T_c \simeq 1764$ K. The same is true for the critical densities and concentrations: these are thus "nonuniversal parameters" directly reflecting the atomic and molecular properties, i.e., the physics on the scale of the cutoff a. Hence, in Fig. 1, ρ_{max} (which may be taken as the density of the corresponding crystal at low T) is of order $1/a^3$, while the scale of $k_B T_c$ is set by the basic microscopic potential energy of attraction denoted ε. While of considerable chemical, physical, and engineering interest, such parameters will be of marginal concern to us here. The point, rather, is that the *shapes* of the coexistence curves, $\rho_{liq}(T)$ and $\rho_{gas}(T)$ *versus* T, become asymptotically *universal* in character as the critical point is approached.

To be more explicit, note first an issue of symmetry. In QFT, symmetries of many sorts play an important role: they may (or must) be built into the theory but can be "broken" in the physically realized vacuum state(s) of the quantum field. In the physics of fluids the opposite situation pertains. There is no real physical symmetry

between coexisting liquid and gas: they are just different states, one a relatively dense collection of atoms or molecules, the other a relatively dilute collection—see Fig. 1(d). However, if one compares the two sides of the coexistence curve, gas and liquid, by forming the ratio

$$R(T)=[\rho_c-\rho_{\rm gas}(T)]/[\rho_{\rm liq}(T)-\rho_c], \tag{7}$$

one discovers an extraordinarily precise *asymptotic* symmetry. Explicitly, when T approaches T_c from below or, introducing a convenient notation,

$$t\equiv(T-T_c)/T_c\to 0-, \tag{8}$$

one finds $R(T)\to 1$. This simply means that the physical fluid builds for itself an exact mirror symmetry in density (and other properties) *as* the critical point is *approached*. And this is a universal feature for all fluids near criticality. (This symmetry is reflected in Fig. 1(d) by the high, although *not* absolutely perfect, degree of asymptotic *linearity* of the coexistence-curve diameter, $\overline{\rho}(T)$—the dashed line described above.)

More striking than the (asymptotic) symmetry of the coexistence curve is the universality of its shape close to T_c—visible in Fig. 1(d) as a flattening of the graph relative to the parabolic shape of the corresponding classical prediction—see plot (c) in Fig. 1, which is derived from the famous van der Waals equation of state. Rather generally one can describe the shape of a fluid coexistence curve in the critical region via the power law

$$\Delta\rho\equiv\tfrac{1}{2}[\rho_{\rm liq}(T)-\rho_{\rm gas}(T)]\approx B|t|^{\beta}\quad\text{as}\quad t\to 0-, \tag{9}$$

where B is a *non*universal amplitude while the critical exponent β takes the *universal* value

$$\beta\simeq 0.325, \tag{10}$$

(in which the last figure is uncertain). To stress the point: β is a nontrivial number, not known exactly, but it is the *same* for all fluid critical points! This contrasts starkly with the classical prediction $\beta=\frac{1}{2}$ [corresponding to a parabola: see Fig. 1(c)]. The value in Eq. (10) applies to ($d=3$)-dimensional systems. Classical theories make the same predictions for all d. On the other hand, for $d=2$, Onsager's work (1949) on the square-lattice Ising model leads to $\beta=\frac{1}{8}$. This value has since been confirmed experimentally by Kim and Chan (1984) for a "two-dimensional fluid" of methane (CH_4) adsorbed on the flat, hexagonal-lattice surface of graphite crystals.

Not only does the value in Eq. (10) for β describe many types of fluid system, it also applies to anisotropic *magnetic materials*, in particular to those of Ising-type with one "easy axis." For that case, in vanishing magnetic fields, H, below the Curie or critical temperature, T_c, a ferromagnet exhibits a spontaneous magnetization and one has $M=\pm M_0(T)$. The sign, + or −, depends on whether one lets H approach zero from positive or negative values. Since, in equilibrium, there is a full, natural physical symmetry under $H\Rightarrow -H$ and $M\Rightarrow -M$ (in contrast to fluid systems) one clearly has $M_c=0$: likewise, the asymptotic symmetry corresponding to Eq. (7) is, in this case exact for all T: see Fig. 1, plots (a) and (b). Thus, as is evident in Fig. 1, the *global shape* of a spontaneous magnetization curve does not closely resemble a normal fluid coexistence curve. Nevertheless, in the asymptotic law

$$M_0(T)\approx B|t|^{\beta}\quad\text{as}\quad t\to 0-, \tag{11}$$

the exponent value in Eq. (10) still applies for $d=3$: see Fig. 1(b); the corresponding classical "mean-field theory" in plot (a), again predicts $\beta=\frac{1}{2}$. For $d=2$ the value $\beta=\frac{1}{8}$ is once more valid!

And, beyond fluids and anisotropic ferromagnets many other systems belong—more correctly their critical behavior belongs—to the "*Ising universality class*." Included are other magnetic materials (antiferromagnets and ferrimagnets), binary metallic alloys (exhibiting order-disorder transitions), certain types of ferroelectrics, and so on.

For each of these systems there is an appropriate order parameter and, via Eq. (2), one can then define (and usually measure) the correlation decay exponent η which is likewise universal. Indeed, essentially any measurable property of a physical system displays a universal critical singularity. Of particular importance is the exponent $\alpha\simeq 0.11$ (Ising, $d=3$) which describes the divergence to infinity of the specific heat via

$$\mathcal{C}(T)\approx A^{\pm}/|t|^{\alpha}\quad\text{as}\quad t\to 0\pm, \tag{12}$$

(at constant volume for fluids or in zero field, $H=0$, for ferromagnets, etc.). The amplitudes A^+ and A^- are again *non*universal; but their *dimensionless ratio*, A^+/A^-, is *universal*, taking a value close to 0.52. When $d=2$, as Onsager (1944) found, $A^+/A^-=1$ and $|t|^{-\alpha}$ is replaced by $\log|t|$. But classical theory merely predicts a jump in specific heat, $\Delta\mathcal{C}=\mathcal{C}_c^-\mathcal{C}_c^+>0$, for all d!

Two other central quantities are a divergent isothermal compressibility $\chi(T)$ (for a fluid) or isothermal susceptibility, $\chi(T)\propto(\partial M/\partial H)_T$ (for a ferromagnet) and, for all systems, a *divergent correlation length*, $\xi(T)$, which measures the growth of the 'range of influence' or of correlation observed say, via the decay of the correlation function $G(R;T)$—see Eq. (1) above—to its long-distance limit. For these functions we write

$$\chi(T)\approx C^{\pm}/|t|^{\gamma}\quad\text{and}\quad\xi(T)\approx\xi_0^{\pm}/|t|^{\nu}, \tag{13}$$

as $t\to 0\pm$, and find, for $d=3$ Ising-type systems,

$$\gamma\simeq 1.24\quad\text{and}\quad\nu\simeq 0.63 \tag{14}$$

(while $\gamma=1\frac{3}{4}$ and $\nu=1$ for $d=2$).

As hinted, there are other universality classes known theoretically although relatively few are found experimentally.[48] Indeed, one of the early successes of

[48]See e.g., the survey in Fisher (1974b) and Aharony (1976).

RG theory was delineating and sharpening our grasp of the various important universality classes. To a significant degree one found that only the vectorial or tensorial character of the relevant order parameter (e.g., scalar, complex number *alias* two-component vector, three-component vector, etc.) plays a role in determining the universality class. But the whys and the wherefores of this self-same issue represent, as does the universality itself, a prime challenge to any theory of critical phenomena.

VI. EXPONENT RELATIONS, SCALING AND IRRELEVANCE

By 1960–62 the existence of universal critical exponents disagreeing sharply with classical predictions may be regarded as well established theoretically and experimentally.[49] The next theoretical step was the discovery of *exponent relations*, that is, simple algebraic equations satisfied by the various exponents *independently* of the universality class. Among the first of these were[50]

$$\gamma=(2-\eta)\nu \quad \text{and} \quad \alpha+2\beta+\gamma=2. \tag{15}$$

As the reader may check from the values quoted above, these relations hold *exactly* for the $d=2$ Ising models and are valid when $d=3$ to within the experimental accuracy or the numerical precision (of the theoretical estimates[51]). They are even obeyed exactly by the classical exponent values (which, today, we understand[52] as valid for $d>4$).

The first relation in Eq. (15) pertains just to the basic correlation function $G(\mathbf{r};T)=\langle\psi(\mathbf{0})\psi(\mathbf{r})\rangle$ as defined previously. It follows from the assumption,[53] supported in turn by an examination of the structure of Onsager's matrix solution to the Ising model[54] *that in the critical region all lengths* (much larger than the lattice spacing a) *scale like the correlation length* $\xi(T)$—introduced in Eq. (13). Formally one expresses this principle by writing, for $t\to 0$ and $r\to\infty$,

$$G(T;\mathbf{r})\approx \frac{D}{r^{d-2+\eta}}\,\mathcal{G}\left(\frac{r}{\xi(T)}\right), \tag{16}$$

where, for consistency with (2), the *scaling function*, $\mathcal{G}(x)$, satisfies the normalization condition $\mathcal{G}(0)=1$. Integrating $\mathbf{r}$ over all space yields the compressibility/susceptibility $\chi(T)$ and, thence, the relation $\gamma=(2-\eta)\nu$. This *scaling law* highlights the importance of the correlation length ξ in the critical region, a feature later stressed and developed further, especially by Widom

[49]This retrospective statement may, perhaps, warrant further comment. First, the terms "universal" and "universality class" came into common usage only after 1974 when (see below) the concept of various types of renormalization-group fixed point had been well recognized (see Fisher, 1974b). Kadanoff (1976) deserves credit not only for introducing and popularizing the terms but especially for emphasizing, refining, and extending the concepts. On the other hand, Domb's (1960) review made clear that all (short-range) Ising models should have the same critical exponents irrespective of lattice structure but depending strongly on dimensionality. The excluded-volume problem for polymers was known to have closely related but *distinct* critical exponents from the Ising model, depending similarly *on* dimensionality but *not* lattice structure (Fisher and Sykes, 1959). And, as regards the Heisenberg model—which possesses what we would now say is an $(n=3)$-component vector or O(3) order parameter—there were strong hints that the exponents were again different (Rushbrooke and Wood, 1958; Domb and Sykes, 1962). On the experimental front matters might, possibly be viewed as less clear-cut: indeed, for ferromagnets, nonclassical exponents were unambiguously revealed only in 1964 by Kouvel and Fisher. However, a striking experiment by Heller and Benedek (1962) had already shown that the order parameter of the *anti*ferromagnet MnF_2, namely, the *sublattice magnetization* $M_0^{\dagger}(T)$, vanishes as $|t|^{\beta}$ with $\beta\simeq 0.33_5$. Furthermore, for fluids, the work of the Dutch school under Michels and the famous analysis of coexistence curves by Guggenheim (1949) allowed little doubt—see Rowlinson (1959), Chap. 3, especially, pp. 91–95—that all reasonably simple atomic and molecular fluids displayed the same but *non*classical critical exponents with $\beta\simeq\frac{1}{3}$: And, also well before 1960, Widom and Rice (1955) had analyzed the *critical isotherms* of a number of simple fluids and concluded that the corresponding critical exponent δ (see, e.g., Fisher, 1967b) took a value around 4.2 in place of the van der Waals value $\delta=3$. In addition, evidence was in hand showing that the *consolute point* in binary fluid mixtures was similar (see Rowlinson, 1959, pp. 165–166).

[50]See Fisher (1959; 1962; 1964, see Eq. (5.7); 1967b) for the first relation here; the second relation was advanced in Essam and Fisher (1963) where the now generally accepted notation for the thermodynamic critical exponents was also introduced. See, in addition, Fisher (1967a) based on a lecture given in March 1965. Actually the initial proposal was written as $\alpha'+2\beta+\gamma'=2$, where the primes denote exponents defined *below* T_c. This distinction, although historically important, is rarely made nowadays since, in general, scaling (see below) implies the $T\gtrless T_c$ equalities $\alpha'=\alpha$, $\gamma'=\gamma$, $\nu'=\nu$, etc. [also mentioned in Essam and Fisher and Fisher (1967a)]. Moved by the suggested thermodynamic exponent equality, Rushbrooke (1963) quickly showed that for magnetic systems (with $H\Rightarrow -H$ symmetry) the positivity of specific heats implied by the Second Law of Thermodynamics could be used to prove rigorously the *inequality* $\alpha'+2\beta+\gamma'\geqslant 2$. His proof was soon extended to fluid systems (Fisher 1964), see Eq. (2.20). Corresponding to the first equality in Eq. (15), the inequality $\gamma\leqslant(2-\eta)\nu$ was proven rigorously in (Fisher, 1969). Other valuable exponent inequalities encompassing "scaling laws" for the exponents as the limiting case of equality were proved by Griffiths (1965, 1972) for thermodynamic exponents and Buckingham and Gunton (1969) for correlation exponents.

[51]See e.g., Fisher (1967b), Baker (1990), Domb (1996).

[52]See Wilson and Fisher (1972), Wilson and Kogut (1974), Fisher (1974, 1983).

[53]See Fisher (1959, 1962).

[54]Onsager (1944), Kaufman and Onsager (1949).

(1965), Kadanoff (1966, 1976), and Wilson (1983).[55] It is worth remarking that in QFT the inverse correlation-length ξ^{-1}, is basically equivalent to the *renormalized mass* of the field ψ: *masslessness* then equates with *criticality* since $\xi^{-1}\to 0$.

The next theoretical question was: "How can one construct an *equation of state* for a system which has non-classical critical exponents?" The "equation of state"—for concreteness let us say, for a ferromagnet—is an equation relating the magnetization, M, the temperature T, the magnetic field, H, and perhaps, some further variable, say P, like, for example, the overall pressure or, more interestingly, the strength of the direct electromagnetic, dipole-dipole couplings. More generally, one wants to know the free energy $F(T,H,P)$ from which all the thermodynamic properties follow[56]—or, better still, the full correlation function $G(\mathbf{r};T,H,P)$ (where previously we had supposed $H = 0$ and $P = P_0$, fixed) since this gives more insight into the "structure" of the system.

The equation of state is crucial knowledge for any applications but, at first sight, the question appears merely of somewhat technical interest. Classical theory provides a simple answer—basically just a power series expansion in $(T-T_c)$, $(M - M_c)$, and $(P - P_c)$, etc.; but that always *enforces* classical exponent values! It transpires, therefore, that the mathematical issues are much more delicate: For convenience, let us focus on the *singular part* of the free energy density, namely,[57]

$$f_s(t,h,g)\equiv -\Delta F(T,H,P)/Vk_BT, \tag{17}$$

as a function of the physically appropriate reduced variables

$$t=(T-T_c)/T_c, \quad h=\mu_B H/k_BT, \quad g=P/k_BT. \tag{18}$$

Now, not only must $f(t,h,g)$ reproduce all the correct critical singularities when $t \to 0$ (for $h=0$, etc.), it must *also* be *free* of singularities, i.e. "analytic," *away* from the critical point (and the phase boundary $h=0$ below T_c).

The solution to this problem came most directly via Widom's (1965b) *homogeneity* or, as more customarily now called, *scaling hypothesis* which *embodies* a minimal number of the critical exponents. This may be written

$$f_s(t,h,g)\approx |t|^{2-\alpha}\mathcal{F}\left(\frac{h}{|t|^{\Delta}},\frac{g}{|t|^{\phi}}\right), \tag{19}$$

where α is the specific heat exponent introduced in Eq. (12) while the new exponent, Δ, which determines *how h scales with t*, is given by

$$\Delta=\beta+\gamma. \tag{20}$$

Widom observed, incidentally, that the classical theories themselves obey scaling: one then has $\alpha=0$, $\Delta=1\frac{1}{2}$, $\phi=-\frac{1}{2}$.

The second new exponent, ϕ, did *not* appear in the original critical-point scaling formulations;[58] neither did the argument $z=g/|t|^{\phi}$ appear in the *scaling function* $\mathcal{F}(y,z)$. It is really only with the appreciation of RG theory that we know that such a dependence should in general be present and, indeed, that a full spectrum $\{\phi_j\}$ of such higher-order exponents with $\phi\equiv\phi_1>\phi_2>\phi_3>\cdots$ must normally appear![59]

But how could such a spectrum of exponents be overlooked? The answer—essentially as supplied by the general RG analysis[60]—is that g and all the higher-order "coupling constants," say g_j, are *irrelevant* if their associated exponents ϕ_j are *negative*. To see this, suppose, as will typically be the case, that $\phi\equiv\phi_1=-\theta$ is negative (so $\theta>0$). Then, on approach to the critical point we see that

$$z=g/|t|^{\phi}=g|t|^{\theta}\to 0. \tag{21}$$

Consequently, $\mathcal{F}(y,z)$, in Eq. (19) can be replaced simply by $\mathcal{F}(y,0)$ which is a function of just a *single variable*. Furthermore, asymptotically when $T \to T_c$ we get the *same* function *whatever* the actual value of g — clearly[61] this is an example of *universality*.

Indeed, within RG theory this is the general mechanism of universality: in a very large (generally infinitely large) space of Hamiltonians, parametrized by t, h, and all the g_j, there is a controlling critical point (later seen to be a *fixed point*) about which each variable enters with a characteristic exponent. All systems with Hamiltonians differing only through the values of the g_j (within suitable bounds) will exhibit the *same critical behavior* determined by the same free-energy scaling function $\mathcal{F}(y)$, where now we drop the irrelevant argu-

---

[55]See also Wilson and Kogut (1974).

[56]Thus, for example, the equation of state is given by $M=-(\partial F/\partial H)_{T,P}$; the specific heat is $C=-T(\partial^2 F/\partial T^2)_{H=0,P}$.

[57]The "singular part," ΔF in Eq. (17), is found by subtracting from F analytic terms: $F_0(T,H,P)=F_c+F_1(T-T_c)+F_2H+\cdots$. In Eq. (17) the volume V of the physical system is shown but a conceptually crucial theoretical issue, namely the *taking* of the *thermodynamic limit*, $V\to\infty$, has, for simplicity, been ignored. In Eq. (18), μ_B denotes the Bohr magneton, so that h is dimensionless.

[58]Widom (1965), Domb and Hunter (1965), Kadanoff (1966), Patashinskii and Pokroskii (1966); and see Fisher (1967b) and Stanley (1971).

[59]See Wilson (1971a) and, for a very general exposition of scaling theory, Fisher (1974a).

[60]Wegner (1972, 1976), Fisher (1974a), Kadanoff (1976).

[61]Again we slide over a physically important detail, namely, that T_c for example, will usually be a function of any irrelevant parameter such as g. This comes about because, in a full scaling formulation, the variables t, h, and g appearing in Eq. (19) must be replaced by *nonlinear scaling fields* $\tilde{t}(t,h,g)$, $\tilde{h}(t,h,g)$ and $\tilde{g}(t,h,g)$ which are smooth functions of t, h, and g (Wegner, 1972, 1976; Fisher, 1983). By the same token it is usually advantageous to introduce a prefactor A_0 in Eq. (19) and "metrical factors" E_j in the arguments $y\equiv z_0$ and z_j (see, e.g., Fisher, 1983).

ment(s). Different universality classes will be associated with different controlling critical points in the space of Hamiltonians with, once one recognizes the concept of RG *flows*, different "domains of attraction" under the flow. All these issues will be reviewed in greater detail below.

In reality, the expectation of a general form of scaling[62] is frequently the most important consequence of RG theory for the practising experimentalist or theorist. Accordingly, it is worth saying more about the meaning and implications of Eq. (19). First, (i) it very generally *implies* the thermodynamic exponent relation Eq. (15) connecting α, β and γ; and (ii) since all leading exponents are determined entirely by the two exponents α and Δ $(=\beta+\gamma)$, it predicts similar exponent relations for any other exponents one might define—such as δ specified *on* the critical isotherm[63] by $H\sim M^{\delta}$. Beyond that, (iii) if one fixes P (or g) and similar parameters and observes the free energy or, in practice, the equation of state, the data one collects amount to describing a function, say $M(T,H)$, of *two variables*. Typically this would be displayed as *sets of isotherms*: i.e., many plots of M vs. H at various closely spaced, fixed values of T near T_c. But according to the scaling law Eq. (19) if one plots the *scaled variables* $f_s/|t|^{2-\alpha}$ or $M/|t|^{\beta}$ vs. the scaled field $h/|t|^{\Delta}$, for appropriately chosen exponents and critical temperature T_c, one should find that all these data "collapse" (in Stanley's (1971) picturesque terminology) onto a single curve, which then just represents the scaling function $x=\mathcal{F}(y)$ itself!

Indeed, this dramatic collapse is precisely found in fitting experimental data. Furthermore, the same "collapse" occurs for different systems since the scaling function $\mathcal{F}(y)$ itself, *also* proves to be *universal* (when properly normalized), as first stressed by Kadanoff (1976). A particularly striking example of such data collapse yielding the same scaling function for a range of irrelevant parameter values, may be found in the recent work by Koch *et al.* (1989).[64] They studied a quite different physical problem, namely, the proposed "vortex-glass" transition in the high-T_c superconductor YBCO. There the voltage drop, E, across the specimen, measured over 4 or 5 decades, plays the role of M; the current density J, measured over a similar range, stands in for h, while the external magnetic field, H, acting on the sample, provides the irrelevant parameter P. The scaling function was finally determined over 10 decades in value and argument and seen to be universal!

[62]Allowing for irrelevant variables, nonlinear scaling fields, and universality, as indicated in Eq. (19) and the previous footnote.

[63]See also Footnote 49 above.

[64]The scaling function, as plotted in this reference, strikes the uninitiated as two distinct functions, one for $T\geqslant T_c$, another for $T\leqslant T_c$. However, this is due just to the presentation adopted: scaling functions like $\mathcal{F}(y)$ in Eq. (19) are typically single functions *analytic through* $T=T_c$ for $y<\infty$ (i.e., $h\neq 0$) and can be re-plotted in a way that exhibits that feature naturally and explicitly.

VII. RELEVANCE, CROSSOVER, AND MARGINALITY

As mentioned, the scaling behavior of the free energy, the equation of state, the correlation functions, and so on, always holds only in some *asymptotic sense* in condensed matter physics (and, indeed, in most applications of scaling). Typically, scaling becomes valid when $t\sim(T-T_c)$ becomes small, when the field H is small, and when the microscopic cut-off a is much smaller than the distances of interest. But one often needs to know: "How small is small enough?" Or, put in other language, "What is the nature of the leading corrections to the dominant power laws?" The "extended scaling" illustrated by the presence of the second argument $z=g/|t|^{\phi}$ in Eq. (19) provides an answer via Eq. (21)—an answer that, phenomenologically, can be regarded as independent of RG theory *per se*[65] but which, in historical fact, essentially grew from insights gained via RG theory.[66]

Specifically, if the physical parameter $P\propto g$ is irrelevant then, by definition, $\phi=-\theta$, is negative and, as discussed, $z=g|t|^{\theta}$ becomes small when $|t|\to 0$. Then one can, fairly generally, hope to expand the scaling function $\mathcal{F}(y,z)$ in powers of z. From this one learns, for example, that the power law Eq. (11) for the spontaneous magnetization of a ferromagnet should, when t is no longer very small, be modified to read

$$M_0(T)=B|t|^{\beta}(1+b_{\theta}|t|^{\theta}+b_1 t+\cdots), \tag{22}$$

where b_{θ} $(\propto g)$ and b_1 are nonuniversal.[67] The exponent θ is often called the *"correction-to-scaling" exponent*—of course, it is universal.[68] It is significant because when θ is smaller than unity and b_{θ} is of order unity, the presence of such a singular correction hampers the reliable estimation of the primary exponent, here β, from experimental or numerical data.

Suppose, on the other hand, that ϕ is *positive* in the basic scaling law Eq. (19). Then when $t\to 0$ the scaled variable $z=g/|t|^{\phi}$ grows larger and larger. Consequently the behavior of $\mathcal{F}(y,z)$ for z small or vanishing becomes of less and less interest. Clearly, the previous discussion of asymptotic scaling fails! When that happens one says that the physical variable P represents a *relevant perturbation* of the original critical behavior.[69] Two possibilities then arise: *Either* the critical point may be *destroyed* altogether! This is, in fact, the effect of the magnetic field, which must itself be regarded as a relevant perturbation since $\phi_0\equiv\Delta=\beta+\gamma>0$. *Alternatively*, when z grows, the true, asymptotic critical behavior may

[65]See Fisher (1974a).

[66]See Wegner (1972) and Fisher (1974).

[67]See Wegner (1972, 1976) and Fisher (1974, 1983).

[68]For $d=3$ Ising-type systems one finds $\theta\simeq 0.54$: see Chen *et al.* (1982), Zinn and Fisher (1996).

[69]Wegner (1972, 1976), Kadanoff (1976): see also Fisher (1983).

crossover[70] to a new, quite *distinct* universality class with different exponents and a new asymptotic scaling function, say, $\mathcal{F}_\infty(y)$.[71]

The crossover scenario is, in fact, realized when the physical system is a ferromagnet with microscopic spin variables, say $\vec{S}(r)$, coupled by *short-range* "exchange" interactions while P measures the strength of the additional, *long-range* magnetic dipole-dipole coupling mediated by the induced electromagnetic fields.[72] Interested theorists had *felt* intuitively that the long-range character of the dipole-dipole coupling *should* matter, i.e., P should be *relevant*. But theoretically there seemed no feasible way of addressing the problem and, on the other hand, the experimentally observed critical exponents (for an important class of magnetic materials) seemed quite independent of the dipole-dipole coupling P.

The advent of RG theory changed that: First, it established a general framework within which the relevance or irrelevance of some particular perturbation P_j could be judged—essentially by the positive or negative sign of the associated exponent ϕ_j, with especially interesting *nonscaling* and *nonuniversal* behavior likely in the *marginal* case $\phi_j = 0$.[73] Second, for many cases where the $P_j=0$ problem was well understood, RG theory showed how the *crossover exponent* ϕ could be determined exactly or perturbatively. Third, the ϵ expansion allowed calculation of ϕ *and* of the new critical behavior to which the crossover occurred.[74] The dipole-dipole problem for ferromagnets was settled via this last route: the dipole perturbation is *always* relevant; *however*, the new, dipolar critical exponents for typical ferromagnets like iron, nickel and gadolinium are numerically so close in value to the corresponding short-range exponents[75] that they are almost indistinguishable by experiment (or simulation)!

On the other hand, in the special example of *anisotropic*, easy-axis or Ising-type ferromagnets in $d = 3$ dimensions the dipolar couplings behave as *marginal* variables at the controlling, *dipolar* critical point.[76] This leads to the prediction of *logarithmic* modifications of the classical critical power laws (by factors diverging as $\log|T - T_c|$ to various powers). The predicted logarithmic behavior has, in fact, been verified experimentally by Ahlers *et al.* (1975). In other cases, especially for $d=2$, marginal variables lead to continuously variable exponents such as $\alpha(g)$, and to quite different thermal variation, like $\exp(\tilde{A}/|t|^{\tilde{\nu}})$; such results have been checked both in exactly solved statistical mechanical models and in physical systems such as superfluid helium films.[77]

I have entered into these relatively detailed and technical considerations—which a less devoted reader need only peruse—in order to convey something of the flavor of how the renormalization group is used in statistical physics and to bring out those features for which it is so valued; because of the multifaceted character of condensed matter physics these are rather different and more diverse than those aspects of RG theory of significance for QFT.

VIII. THE TASK FOR RENORMALIZATION GROUP THEORY

Let us, at this point, recapitulate briefly by highlighting, from the viewpoint of statistical physics, what it is one would wish RG theory to accomplish. First and foremost, (i) it should explain the ubiquity of power laws at and near critical points: see Eqs. (2), (9), (11)–(13). I sometimes like to compare this issue with the challenge to atomic physics of explaining the ubiquity of sharp spectral lines. Quantum mechanics responds, crudely speaking, by saying: "Well, (a) there is some wave—or a *wave function* ψ — needed to describe electrons in atoms, and (b) to fit a wave into a confined space the wave length must be quantized: hence (c) only certain definite energy levels are allowed and, thence, (d) there are sharp, spectral transitions between them!"

Of course, that is far from being the whole story in quantum mechanics; but I believe it captures an important essence. Neither is the first RG response the whole story: but, *to anticipate*, in Wilson's conception RG theory crudely says: "Well, (a) there is a *flow* in some *space*, $\mathbb{H}$, *of Hamiltonians* (or "coupling constants"); (b) the critical point of a system is associated with a *fixed point* (or stationary point) of that flow; (c) the flow operator—technically the *RG transformation*,[78] $\mathbb{R}$—can

[70]See the extensive discussion of crossover in Fisher (1974b) and Aharony (1976).

[71]Formally, one might write $\mathcal{F}_\infty(y)=\mathcal{F}(y, z\to z_\infty)$ where z_∞ is a critical value which could be ∞; but a more subtle relationship is generally required since the exponent α in the prefactor in Eq. (19) changes.

[72]A "short-range" interaction potential, say $J(\mathbf{r})$, is usually supposed to decay with distance as $\exp(-r/R_0)$ where R_0 is some microscopic range, but certainly must decay *faster* than $1/r^{d+2}$; the dipole-dipole potential, however, decays more slowly, as $1/r^d$, and has a crucially important angular dependence as well.

[73]See the striking analysis of Kadanoff and Wegner (1971).

[74]Fisher and Pfeuty (1972), Wegner (1972b).

[75]Fisher and Aharony (1973).

[76]Aharony (1973, 1976).

[77]See Kadanoff and Wegner (1971) and, for a review of the extensive later developments—including the Kosterlitz-Thouless theory of two-dimensional superfluidity and the Halperin-Nelson-Kosterlitz-Thouless-Young theory of two-dimensional melting—see Nelson (1983).

[78]As explained in more detail in Secs. XI and XII below, a specific renormalization transformation, say $\mathbb{R}_b$, acts on some 'initial' Hamiltonian $\mathcal{H}^{(0)}$ in the space $\mathbb{H}$ to transform it into a new Hamiltonian, $\mathcal{H}^{(1)}$. Under repeated operation of $\mathbb{R}_b$ the initial Hamiltonian "flows" into a sequence $\mathcal{H}^{(l)}$ $(l=1,2,\cdots)$ corresponding to the iterated RG transformation $\mathbb{R}_b\cdots\mathbb{R}_b$ (l times) which, in turn, specifies a new transformation $\mathbb{R}_{b^l}$. These "products" of repeated RG operations serve to define a *semigroup* of transformations that, in general, does *not* actually give rise to a group: see Footnote 3 above and the discussion below in Sec. XI associated with Eq. (35).

be *linearized* about that fixed point; and (d) typically, such a linear operator (as in quantum mechanics) has a spectrum of discrete, but nontrivial eigenvalues, say λ_k; then (e) each (asymptotically independent) exponential term in the flow varies as $e^{\lambda_k l}$, where l is the *flow* (or renormalization) *parameter* and corresponds to a physical power law, say $|t|^{\phi_k}$, with critical exponent ϕ_k proportional to the eigenvalue λ_k." How one may find suitable transformations $\mathbb{R}$ and why the flows matter, are the subjects for the following chapters of our story.

Just as quantum mechanics does much more than explain sharp spectral lines, so RG theory should also explain, at least in principle, (ii) the values of the leading thermodynamic and correlation exponents, α, β, γ, δ, ν, η, and ω (to cite those we have already mentioned above) and (iii) clarify why and how the classical values are in error, including the existence of borderline dimensionalities, like $d_\times = 4$, above which classical theories become valid. Beyond the leading exponents, one wants (iv) the correction-to-scaling exponent θ (and, ideally, the higher-order correction exponents) and, especially, (v) one needs a method to compute crossover exponents, ϕ, to check for the relevance or irrelevance of a multitude of possible perturbations. Two central issues, of course, are (vi) the understanding of universality with nontrivial exponents and (vii) a derivation of *scaling*: see (16) and (19).

And, more subtly, one wants (viii) to understand the *breakdown* of universality and scaling in certain circumstances—one might recall continuous spectra in quantum mechanics—and (ix) to handle effectively logarithmic and more exotic dependences on temperature, etc.

An important further requirement as regards condensed matter physics is that RG theory should be firmly related to the science of statistical mechanics as perfected by Gibbs. Certainly, there is no need and should be no desire, to replace standard statistical mechanics as a basis for describing equilibrium phenomena in pure, homogeneous systems.[79] Accordingly, it is appropriate to summarize briefly the demands of statistical mechanics in a way suitable for describing the formulation of RG transformations.

We may start by supposing that one has a set of microscopic, fluctuating, mechanical variables: in QFT these would be the various quantum fields, $\psi(\mathbf{r})$, defined—one supposes—at all points in a Euclidean (or Minkowski) space. In statistical physics we will, rather, suppose that in a physical system of volume V there are N discrete "degrees of freedom." For classical fluid systems one would normally use the coordinates $\mathbf{r}_1, \mathbf{r}_2, \cdots, \mathbf{r}_N$ of the constituent particles. However, it is simpler mathematically—and the analogies with QFT are closer—if we consider here a set of "*spins*" $s_{\mathbf{x}}$ (which could be vectors, tensors, operators, etc.) associated with discrete lattice sites located at uniformly spaced points $\mathbf{x}$. If, as before, the lattice spacing is a, one can take $V = Na^d$ and the density of degrees of freedom in d spatial dimensions is $N/V = a^{-d}$.

In terms of the basic variables $s_{\mathbf{x}}$, one can form various "local operators" (or "physical densities" or "observables") like the local magnetization and energy densities

$$M_{\mathbf{x}} = \mu_B s_{\mathbf{x}}, \quad \mathcal{E}_{\mathbf{x}} = -\tfrac{1}{2} J \sum_{\boldsymbol{\delta}} s_{\mathbf{x}} s_{\mathbf{x}+\boldsymbol{\delta}}, \quad \cdots, \tag{23}$$

(where μ_B and J are fixed coefficients while $\boldsymbol{\delta}$ runs over the nearest-neighbor lattice vectors). A physical system of interest is then specified by its *Hamiltonian* $\mathcal{H}[\{s_{\mathbf{x}}\}]$—or energy function, as in mechanics—which is usually just a spatially uniform sum of local operators. The crucial function is the *reduced Hamiltonian*

$$\bar{\mathcal{H}}[s;\, t,\, h,\, \cdots,\, h_j, \cdots] = -\mathcal{H}[\{s_{\mathbf{x}}\};\, \cdots,\, h_j, \cdots]/k_B T, \tag{24}$$

where s denotes the set of all the microscopic spins $s_{\mathbf{x}}$ while $t, h, \cdots, h_j, \cdots$ are various "*thermodynamic fields*" (in QFT—the coupling constants): see Eq. (18). We may suppose that one or more of the thermodynamic fields, in particular the temperature, can be controlled directly by the experimenter; but others may be "given" since they will, for example, embody details of the physical system that are "fixed by nature."

Normally in condensed matter physics one thus focuses on some specific form of $\mathcal{H}$ with at most two or three variable parameters—the Ising model is one such particularly simple form with just two variables, t, the reduced temperature, and h, the reduced field. An important feature of Wilson's approach, however, is to regard any such "physical Hamiltonian" as merely specifying a subspace (spanned, say, by "coordinates" t and h) in a very large space of possible (reduced) Hamiltonians, $\mathbb{H}$: see the schematic illustration in Fig. 2. This change in perspective proves crucial to the proper formulation of a renormalization group: in principle, it enters also in QFT although in practice, it is usually given little attention.

Granted a microscopic Hamiltonian, statistical mechanics promises to tell one the thermodynamic properties of the corresponding macroscopic system! First one must compute the partition function

$$Z_N[\bar{\mathcal{H}}] = \mathrm{Tr}_N^s\{e^{\bar{\mathcal{H}}[s]}\}, \tag{25}$$

where the *trace operation*, $\mathrm{Tr}_N^s\{\cdot\}$, denotes a summation

[79] One may, however, raise legitimate concerns about the adequacy of customary statistical mechanics when it comes to the analysis of random or impure systems—or in applications to systems far from equilibrium or in metastable or steady states—e.g., in fluid turbulence, in sandpiles and earthquakes, etc. And the use of RG ideas in chaotic mechanics and various other topics listed above in Sec. III, clearly does *not* require a statistical mechanical basis.

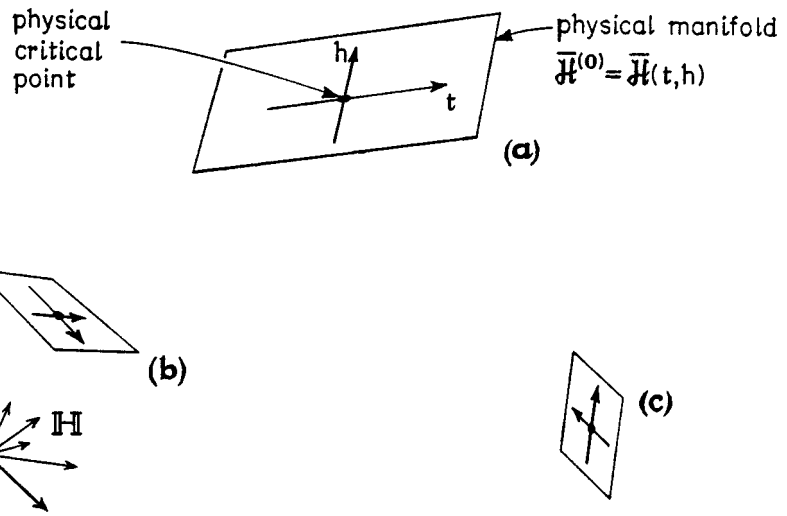


FIG. 2. Schematic illustration of the space of Hamiltonians, $\mathbb{H}$, having, in general, infinitely many dimensions (or coordinate axes). A particular physical system or model representing, say, the ferromagnet, iron, is specified by its reduced Hamiltonian $\bar{\mathcal{H}}(t,h)$, with $t=(T-T_c)/T_c$ and $h=\mu_B H/k_B T$ defined for *that* system: but in $\mathbb{H}$ this Hamiltonian specifies only a submanifold—the physical manifold, labelled (a), that is parametrized by the 'local coordinates' t and h. Other submanifolds, (b), $\cdots$ (c), $\cdots$ located elsewhere in $\mathbb{H}$, depict the physical manifolds for Hamiltonians corresponding to other particular physical systems, say, the ferromagnets nickel and gadolinium, etc.

or integration[80] over all the possible values of all the N spin variables $s_{\mathbf{x}}$ in the system of volume V. The *Boltzmann factor*, $\exp(\bar{\mathcal{H}}[s])$, measures, of course, the probability of observing the microstate specified by the set of values $\{s_{\mathbf{x}}\}$ in an equilibrium ensemble at temperature T. Then the thermodynamics follow from the total free energy density, which is given by[81]

$$f[\bar{\mathcal{H}}] \equiv f(t, h, \cdots h_j \cdots) = \lim_{N,V\to\infty} V^{-1} \log Z_N[\bar{\mathcal{H}}]; \tag{26}$$

this includes the singular part $f_s[\bar{\mathcal{H}}]$ near a critical point of interest: see Eq. (17). Correlation functions are defined similarly in standard manner.

To the degree that one can actually perform the trace operation in Eq. (25) for a particular model system and take the "thermodynamic limit" in Eq. (26) one will obtain the precise critical exponents, scaling functions, and so on. This was Onsager's (1944) route in solving the $d=2$, spin-$\frac{1}{2}$ Ising models in zero magnetic field. At first sight one then has no need of RG theory. That surmise, however, turns out to be far from the truth. The issue is "simply" one of understanding! (Should one ever achieve truly high precision in simulating critical systems on a computer—a prospect which still seems some decades away—the same problem would remain.) In short, while one knows for sure that $\alpha = 0$ (log), $\beta=\frac{1}{8}$, $\gamma=1\frac{3}{4}$, $\nu=1$, $\eta=\frac{1}{4}$,$\cdots$ for the planar Ising models one does not know *why* the exponents have these values or *why* they satisfy the exponent relations Eqs. (15) or why the scaling law Eq. (16) is obeyed. Indeed, the seemingly inevitable mathematical complexities of solving even such physically oversimplified models exactly[82] serve to conceal almost all traces of general, underlying mechanisms and principles that might "explain" the results. Thus it comes to pass that even a rather crude and approximate solution of a two-dimensional Ising model by a real-space RG method can be truly instructive.[83]

IX. KADANOFF'S SCALING PICTURE

The year from late-1965 through 1966 saw the clear formulation of scaling for the thermodynamic properties in the critical region and the fuller appreciation of scaling for the correlation functions.[84] I have highlighted Widom's (1965) approach since it was the most direct and phenomenological—a bold, new thermodynamic hypothesis was advanced by generalizing a particular feature of the classical theories. But Domb and Hunter (1965) reached essentially the same conclusion for the thermodynamics based on analytic and series-expansion considerations, as did Patashinskii and Pokrovskii (1966)

[80]Here, for simplicity, we suppose the $s_{\mathbf{x}}$ are classical, commuting variables. If they are operator-valued then, in the standard way, the trace must be defined as a sum or integral over diagonal matrix elements computed with a complete basis set of N-variable states.

[81]In Eq. (26) we have explicitly indicated the thermodynamic limit in which N and V become infinite maintaining the ratio $V/N = a^d$ fixed: in QFT this corresponds to an infinite system with an ultraviolet lattice cutoff.

[82]See the monograph by Rodney Baxter (1982).

[83]See Niemeijer and van Leeuwen (1976), Burkhardt and van Leeuwen (1982), and Wilson (1975, 1983) for discussion of real-space RG methods.

[84]Although one may recall, in this respect, earlier work (Fisher, 1959, 1962, 1964) restricted (in the application to ferromagnets) to zero magnetic field.

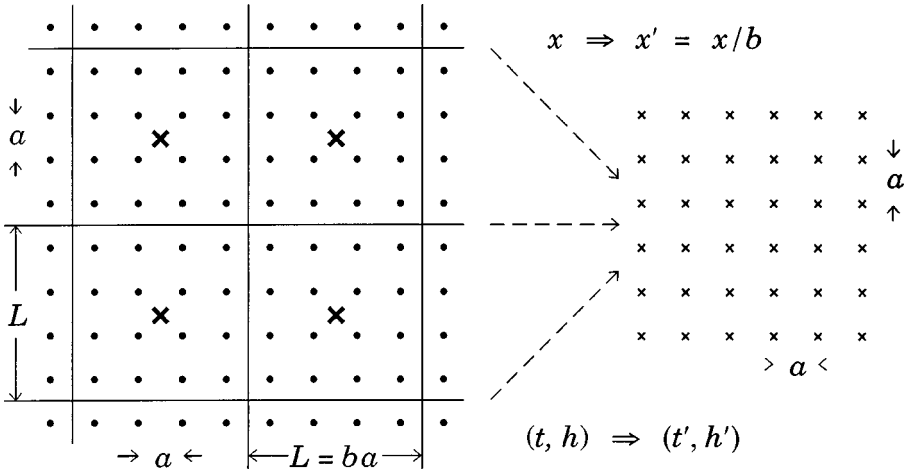


FIG. 3. A lattice of spacing a of Ising spins $s_{\mathbf{x}}=\pm 1$ (in $d=2$ dimensions) marked by solid dots, divided up into Kadanoff blocks or cells of dimensions $(L=ba)\times(L=ba)$ each containing a block spin $s'_{\mathbf{x}'}=\pm 1$, indicated by a cross. After a rescaling, $\mathbf{x}\Rightarrow\mathbf{x}'=\mathbf{x}/b$, the lattice of block spins appears identical with the original lattice. However, one supposes that the temperature t, and magnetic field h, of the original lattice can be renormalized to yield appropriate values, t' and h', for the rescaled, block-spin lattice: see text. In this illustration the spatial rescaling factor is $b = 4$.

using a more microscopic formulation that brought out the relations to the full set of correlation functions (of all orders).[85]

Kadanoff (1966), however, derived scaling by introducing a completely new concept, namely, the *mapping* of a critical or near-critical system onto itself by a reduction in the effective number of degrees of freedom.[86] This paper attracted much favorable notice since, beyond obtaining all the scaling properties, it seemed to lay out a direct route to the actual *calculation* of critical properties. On closer examination, however, the implied program seemed—as I will explain briefly—to run rapidly into insuperable difficulties and interest faded. In retrospect, however, Kadanoff's scaling picture embodied important features eventually seen to be basic to Wilson's conception of the full renormalization group. Accordingly, it is appropriate to present a sketch of Kadanoff's seminal ideas.

For simplicity, consider with Kadanoff (1966), a lattice of spacing a (and dimensionality $d>1$) with $S=\frac{1}{2}$ Ising spins $s_{\mathbf{x}}$ which, by definition, take only the values $+1$ or -1: see Fig. 3. Spins on nearest-neighbor sites are coupled by an energy parameter or coupling constant, $J>0$, which favors parallel alignment [see, e.g., Eq. (23) above]. Thus at low temperatures the majority of the spins point "up" $(s_{\mathbf{x}}=+1)$ or, alternatively, "down" $(s_{\mathbf{x}}=-1)$; in other words, there will be a spontaneous magnetization, $M_0(T)$, which decreases when T rises until it vanishes at the critical temperature $T_c>0$: recall (11).

Now divide the lattice up into (disjoint) blocks, of dimensions $L\times L\times\cdots\times L$ with $L=ba$ so that each block contains b^d spins: see Fig. 3. Then associate with each block, $\mathcal{B}_{\mathbf{x}'}$ centered at point $\mathbf{x}'$, a new, effective *block spin*, $s'_{\mathbf{x}'}$. If, finally, we *rescale* all spatial coordinates according to

$$\mathbf{x}\Rightarrow\mathbf{x}'=\mathbf{x}/b, \tag{27}$$

the new lattice of block spins $s'_{\mathbf{x}'}$ looks just like the original lattice of spins $s_{\mathbf{x}}$. Note, in particular, the density of degrees of freedom is unchanged: see Fig. 3.

[85]It was later seen (Kiang and Stauffer, 1970; Fisher, 1971, Sec. 4.4) that thermodynamic scaling with general exponents (but particular forms of scaling function) was embodied in the "droplet model" partition function advanced by Essam and Fisher (1963) from which the exponent relations $\alpha'+2\beta+\gamma'=2$, etc., were originally derived. (See Eq. (15), Footnote 49, and Fisher, 1967b, Sec. 9.1; 1971, Sec. 4.)

[86]Novelty is always relative! From a historical perspective one should recall a suggestive contribution by M. J. Buckingham, presented in April 1965, in which he proposed a division of a lattice system into cells of geometrically increasing size, $L_n=b^nL_0$, with controlled intercell couplings. This led him to propose "the *existence* of an asymptotic 'lattice problem' such that the description of the nth order in terms of the $(n-1)$th is the same as that of the $(n+1)$th in terms of the nth." This is practically a description of "scaling" or "self similarity" as we recognize it today. Unfortunately, however, Buckingham failed to draw any significant, correct conclusions from his conception and his paper seemed to have little influence despite its presentation at the notable international conference on *Phenomena in the Neighborhood of Critical Points* organized by M. S. Green (with G. B. Benedek, E. W. Montroll, C. J. Pings, and the author) and held at the National Bureau of Standards, then in Washington, D.C. The Proceedings, complete with discussion remarks, were published, in December 1966, under the editorship of Green and J. V. Sengers (1966). Nearly all the presentations addressed the rapidly accumulating experimental evidence, but many well known theorists from a range of disciplines attended including P. W. Anderson, P. Debye, C. de Dominicis, C. Domb, S. F. Edwards, P. C. Hohenberg, K. Kawasaki, J. S. Langer, E. Lieb, W. Marshall, P. C. Martin, T. Matsubara, E. W. Montroll, O. K. Rice, J. S. Rowlinson, G. S. Rushbrooke, L. Tisza, G. E. Uhlenbeck, and C. N. Yang; but B. Widom, L. P. Kadanoff, and K. G. Wilson are *not* listed among the participants.

But if this appearance is to be more than superficial one must be able to relate the new or "renormalized" coupling J' between the block spins to the original coupling J, or, equivalently, the renormalized temperature deviation t' to the original value t. Likewise one must relate the new, renormalized magnetic field h' to the original field h.

To this end, Kadanoff supposes that b is large but less than the ratio, ξ/a, of the *correlation length*, $\xi(t,h)$, to the lattice spacing a; since ξ diverges at criticality—see Eq. (13)—this allows, asymptotically, for b to be chosen *arbitrarily*. Then Kadanoff notes that the total coupling of the magnetic field h to a block of b^d spins is equivalent to a coupling to the average spin

$$\overline{s}_{\mathbf{x}'} \equiv b^{-d} \sum_{\mathbf{x} \in \mathcal{B}_{\mathbf{x}'}} s_{\mathbf{x}} \cong \zeta(b) s'_{\mathbf{x}'}, \tag{28}$$

where the sum runs over all the sites $\mathbf{x}$ in the block $\mathcal{B}_{\mathbf{x}'}$, while the "asymptotic equivalence" to the new, Ising block spin $s'_{\mathbf{x}'}$ is, Kadanoff proposes, determined by some "spin rescaling or renormalization factor" $\zeta(b)$. Introducing a similar thermal renormalization factor, $\vartheta(b)$, leads to the *recursion relations*

$$t' \approx \vartheta(b) t \quad \text{and} \quad h' \approx \zeta(b) h. \tag{29}$$

Correspondingly, the basic correlation function—compare with Eqs. (1), (4), and (16)—should renormalize as

$$G(\mathbf{x};\, t,\, h) \equiv \langle s_{\mathbf{0}}\, s_{\mathbf{x}} \rangle \approx \zeta^2(b) G(\mathbf{x}';\, t', h'). \tag{30}$$

In summary, under a spatial scale transformation and the integration out of all but a fraction b^{-d} of the original spins, the system asymptotically *maps back into itself* although at a renormalized temperature and field! However, the map is *complete* in the sense that *all* the statistical properties should be related by similarity.

But how should one choose—or, better, determine—the renormalization factors ζ and ϑ? Let us consider the basic relation Eq. (30) *at* criticality, so that $t = h = 0$ and, by Eq. (29), $t' = h' = 0$. Then, if we accept the observation/expectation Eq. (2) of a power law decay, i.e., $G_c(\mathbf{x}) \sim 1/|\mathbf{x}|^{d-2+\eta}$ one soon finds that $\zeta(b)$ must be just a power of b. It is natural, following Kadanoff (1966), then to propose the forms

$$\zeta(b) = b^{-\omega} \quad \text{and} \quad \vartheta(b) = b^{\lambda}, \tag{31}$$

where the two exponents ω and λ characterize the critical point under study while b is an essentially unrestricted *scaling parameter*.

By capitalizing on the freedom to choose b as $t,\, h \to 0$, or, more-or-less equivalently, by *iterating* the recursion relations Eqs. (29) and (30), one can, with some further work, show that all the previous scaling laws hold, specifically, Eqs. (15), (16), and (19) although with $g \equiv 0$. Of course, all the exponents are now determined by ω and λ: for example, one finds $\nu = 1/\lambda$ and $\beta = \omega\nu$. Beyond that, the analysis leads to new exponent relations, namely, the so-called *hyperscaling laws*[87] which explicitly involve the spatial dimensionality: most notable is[88]

$$d\nu = 2 - \alpha. \tag{32}$$

But then Kadanoff's scaling picture is greatly strengthened by the fact that this relation holds *exactly* for the $d = 2$ Ising model! And also for all other exactly soluble models when $d < 4$.[89]

Historically, the careful numerical studies of the $d = 3$ Ising models by series expansions[90] for many years suggested a small but significant deviation from Eq. (32) as allowed by pure scaling phenomenolgy.[91] But, in recent years, the accumulating weight of evidence critically reviewed has convinced even the most cautious skeptics![92]

Nevertheless, all is not roses! Unlike the previous exponent relations (all being independent of d) hyperscaling fails for the classical theories unless $d = 4$. And since one knows (rigorously for certain models) that the classical exponent values are valid for $d > 4$, it follows that hyperscaling cannot be generally valid. Thus something is certainly missing from Kadanoff's picture. Now, thanks to RG insights, we know that the breakdown of hyperscaling is to be understood via the second argument in the "fuller" scaling form Eq. (19): when d exceeds the appropriate borderline dimension, $d_\times$, a "dangerous irrelevant variable" appears and must be allowed for.[93] In essence one finds that the scaling function limit $\mathcal{F}(y,\, z \to 0)$, previously accepted without question, is no longer well defined but, rather, diverges as a power of z: asymptotic scaling survives but $d^* \equiv (2-\alpha)/\nu$ sticks at the value 4 for $d > d_\times = 4$.

However, the issue of hyperscaling was *not* the main road block to the analytic development of Kadanoff's picture. The principal difficulties arose in explaining the *power-law* nature of the rescaling factors in Eqs. (29)–(31) and, in particular, in justifying the idea of a *single*, effective, renormalized coupling J' between adjacent block spins, say $s'_{\mathbf{x}'}$ and $s'_{\mathbf{x}'+\boldsymbol{\delta}'}$. Thus the interface between two adjacent $L \times L \times L$ blocks (taking $d = 3$ as an

[87]See (Fisher, 1974a) where the special character of the hyperscaling relations is stressed.

[88]See Kadanoff (1966), Widom (1965a), and Stell (1965, unpublished, quoted in Fisher, 1969, and 1968).

[89]See, e.g., Fisher (1983) and, for the details of the exactly solved models, Baxter (1982).

[90]For accounts of series expansion techniques and their important role see: Domb (1960, 1996), Baker (1961, 1990), Essam and Fisher (1963), Fisher (1965, 1967b), and Stanley (1971).

[91]As expounded systematically in (Fisher, 1974a) with hindsight enlightened by RG theory.

[92]See Fisher and Chen (1985) and Baker and Kawashima (1995, 1996).

[93]See Fisher in (Gunton and Green, 1974, p. 66) where a "dangerous irrelevant variable" is characterized as a "hidden relevant variable;" and (Fisher, 1983, App. D).

example) separates two block faces each containing b^2 strongly interacting, original lattice spins $s_{\mathbf{x}}$. Well below T_c all these spins are frozen, ''up'' or ''down,'' and a single effective coupling could well suffice; but at and above T_c these spins must fluctuate on many scales and a single effective-spin coupling seems inadequate to represent the inherent complexities.[94]

One may note, also that Kadanoff's picture, like the scaling hypothesis itself, provides no real hints as to the origins of universality: the rescaling exponents ω and λ in Eq. (31) might well change from one system to another. Wilson's (1971a) conception of the renormalization group answered both the problem of the ''lost microscopic details'' of the original spin lattice and provided a natural explanation of universality.

X. WILSON'S QUEST

Now because this account has a historical perspective, and since I was Ken Wilson's colleague at Cornell for some twenty years, I will say something about how his search for a deeper understanding of quantum field theory led him to formulate renormalization group theory as we know it today. The first remark to make is that Ken Wilson is a markedly independent and original thinker and a rather private and reserved person. Secondly, in his 1975 article, in Reviews of Modern Physics, from which I have already quoted, Ken Wilson gave his considered overview of RG theory which, in my judgement, still stands well today. In 1982 he received the Nobel Prize and in his Nobel lecture, published in 1983, he devotes a section to ''Some History Prior to 1971'' in which he recounts his personal scientific odyssey.

He explains that as a student at Caltech in 1956–60, he failed to avoid ''the default for the most promising graduate students [which] was to enter elementary-particle theory.'' There he learned of the 1954 paper by Gell-Mann and Low ''which was the principal inspiration for [his] own work prior to Kadanoff's (1966) formulation of the scaling hypothesis.'' By 1963 Ken Wilson had resolved to pursue quantum field theories as applied to the strong interactions. Prior to summer 1966 he heard Ben Widom present his scaling equation of state in a seminar at Cornell ''but was puzzled by the absence of any theoretical basis for the form Widom wrote down.'' Later, in summer 1966, on studying Onsager's solution of the Ising model in the reformulation of Lieb, Schultz, and Mattis,[95] Wilson became aware of analogies with field theory and realized the applicability of his own earlier RG ideas (developed for a truncated version of fixed-source meson theory[96]) to critical phenomena. This gave him a scaling picture but he discovered that he ''had been scooped by Leo Kadanoff.'' Thereafter Ken Wilson amalgamated his thinking about field theories on a lattice and critical phenomena learning, in particular, about Euclidean QFT[97] and its close relation to the transfer matrix method in statistical mechanics—the basis of Onsager's (1944) solution.

That same summer of 1966 I joined Ben Widom at Cornell and we jointly ran an open and rather wide-ranging seminar loosely centered on statistical mechanics. Needless to say, the understanding of critical phenomena and of the then new scaling theories was a topic of much interest. Ken Wilson frequently attended and, perhaps partially through that route, soon learned a lot about critical phenomena. He was, in particular, interested in the series expansion and extrapolation methods for estimating critical temperatures, exponents, amplitudes, etc., for lattice models that had been pioneered by Cyril Domb and the King's College, London group.[98] This approach is, incidentally, still one of the most reliable and precise routes available for estimating critical parameters. At that time I, myself, was completing a paper on work with a London University student, Robert J. Burford, using high-temperature series expansions to study in detail the correlation functions and scattering behavior of the two- and three-dimensional Ising models.[99] Our theoretical analysis had already brought out some of the analogies with field theory revealed by the transfer matrix approach. Ken himself undertook large-scale series expansion calculations in order to learn and understand the techniques. Indeed, relying on the powerful computer programs Ken Wilson developed and kindly made available to us, another one of my students, Howard B. Tarko, extended the series analysis of the Ising correlations functions to temperatures below T_c and to all values of the magnetic field.[100] Our results have lasted rather well and many of them are only recently being revised and improved.[101]

Typically, then, Ken Wilson's approach was always ''hands on'' and his great expertise with computers was ever at hand to check his ideas and focus his thinking.[102]

---

[94]In hindsight, we know this difficulty is profound: in general, it is *impossible* to find an adequate single coupling. However, for certain special models it does prove possible and Kadanoff's picture goes through: see Nelson and Fisher (1975) and (Fisher, 1983). Further, in defense of Kadanoff, the condition $b \ll \xi/a$ was supposed to ''freeze'' the original spins in each block sufficiently well to justify their replacement by a simple block spin.

[95]See Schultz *et al.* (1964).

[96]See Wilson (1983).

[97]As stressed by Symanzik (1966) the Euclidean formulation of quantum field theory makes more transparent the connections to statistical mechanics. Note, however, that in his 1966 article Symanzik did not delineate the special connections to critical phenomena *per se* that were gaining increasingly wide recognition; see, e.g., Patashinskii and Pokrovskii (1966), Fisher (1969, Sec. 12) and the remarks below concerning Fisher and Burford (1967).

[98]See the reviews Domb (1960), Fisher (1965, 1967b), Stanley (1971).

[99]Fisher and Burford (1967).

[100]Tarko and Fisher (1975).

[101]See Zinn and Fisher (1996), Zinn, Lai, and Fisher (1996), and references therein.

[102]See his remarks in Wilson (1983) on page 591, column 1.

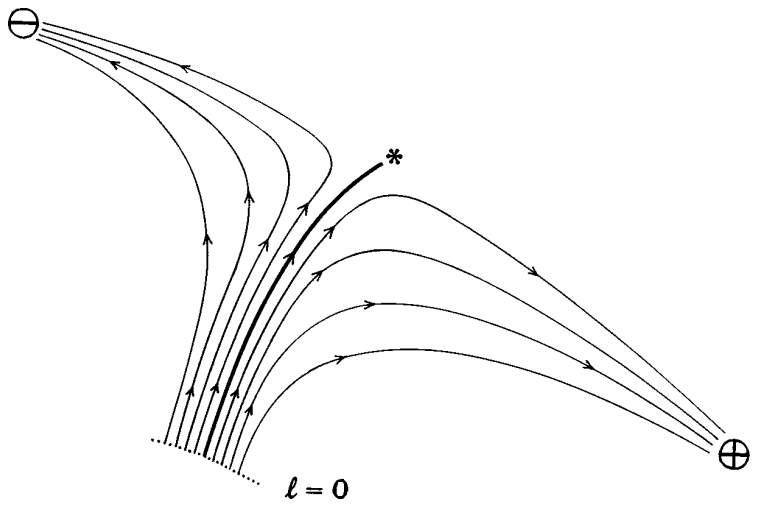


FIG. 4. A ''vision'' of flows in some large space inspired by a seminar of K. G. Wilson in the period 1967–1970. The idea conveyed is that initially close, smoothly connected points at the start of the flow—the locus $l=0$ — can eventually separate and run to far distant regions representing very different ''final'' physical states: the essence of a phase transition. In modern terms the flow is in the space $\mathbb{H}$ of Hamiltonians; the intersection of the separatrix, shown bolder, with the initial locus ($l=0$) represents the physical critical point; $*$ denotes the controlling fixed point, while $\oplus$ and $\ominus$, represent asymptotic high-T, disordered states and low-T, ordered states, respectively.

From time to time Ken would intimate to Ben Widom or myself that he might be ready to tell us where his thinking about the central problem of explaining scaling had got to. Of course, we were eager to hear him speak at our seminar although his talks were frequently hard to grasp. From one of his earlier talks and the discussion afterwards, however, I carried away a powerful and vivid picture of *flows*—flows in a large space. And the point was that at the initiation of the flow, when the ''time'' or ''flow parameter'' l, was small, two nearby points would travel close together; see Fig. 4. But as the flow developed a point could be reached—a bifurcation point (and hence, as one later realized, a stationary or fixed point of the flow)—beyond which the two originally close points could separate and, as l increased, diverge to vastly different destinations: see Fig. 4. At the time, I vaguely understood this as indicative of how a sharp, nonanalytic phase transition could grow from smooth analytic initial data.[103]

But it was a long time before I understood the nature of the space—the space $\mathbb{H}$ of Hamiltonians—and the *mechanism* generating the flow, that is, a renormalization group transformation. Nowadays, when one looks at Fig. 4, one sees the locus of initial points, $l=0$, as identifying the manifold corresponding to the original or 'bare' Hamiltonian (see Fig. 2) while the trajectory leading to the bifurcation point represents a locus of critical points; the two distinct destinations for $l\to\infty$ then typically, correspond to a high-temperature, fully disordered system and to a low-temperature fully ordered system: see Fig. 4.

In 1969 word reached Cornell that two Italian theorists, C. Di Castro and G. Jona-Lasinio, were claiming[104] that the ''multiplicative renormalization group,'' as expounded in the field-theory text by Bogoliubov and Shirkov (1959), could provide ''a microscopic foundation'' for the scaling laws (which, by then, were well established phenomenologically). The formalism and content of the field-theoretic renormalization group was totally unfamiliar to most critical-phenomena theorists: but the prospect of a microscopic derivation was clearly exciting! However, the articles[105] proved hard to interpret as regards concrete progress and results. Nevertheless, the impression is sometimes conveyed that Wilson's final breakthrough was somehow anticipated by Di Castro and Jona-Lasinio.[106]

Such an impression would, I believe, be quite misleading. Indeed, Di Castro was invited to visit Cornell where he presented his ideas in a seminar that was listened to attentively. Again I have a vivid memory: walking to lunch at the Statler Inn after the seminar I checked my own impressions with Ken Wilson by asking: ''Well, did he really say anything new?'' (By ''new'' I meant some fresh insight or technique that carried the field forward.) The conclusion of our conversation was ''No''. The point was simply that none of the problems then outstanding—see the ''tasks'' outlined above (in Section VIII)—had been solved or come under effective attack. In fairness, I must point out that the retrospective re-

[103]See the (later) introductory remarks in Wilson (1971a) related to Fig. 1 there.

[104]The first published article was Di Castro and Jona-Lasinio (1969).

[105]See the later review by Di Castro and Jona-Lasinio (1976) for references to their writings in the period 1969–1972 prior to Wilson's 1971 papers and the ϵ-expansion in 1972.

[106]See, for example, Benfatto and Gallavotti (1995) on page 96 in *A Brief Historical Note*, which is claimed to represent only the authors' personal ''cultural evolution through the subject.''

view by Di Castro and Jona-Lasinio themselves (1976) is reasonably well balanced: One accepted a scaling hypothesis and injected that as an ansatz into a general formalism; then certain insights and interesting features emerged; but, in reality, only scaling theory had been performed; and, in the end, as Di Castro and Jona-Lasinio say: "Still one did not see how to perform explicit calculations." Incidentally, it is also interesting to note Wilson's sharp criticism[107] of the account presented by Bogoliubov and Shirkov (1959) of the original RG ideas of Stueckelberg and Petermann (who, in 1953, coined the phrase "groupes de normalization") and of Gell-Mann and Low (1954).

One more personal anecdote may be permissible here. In August 1973 I was invited to present a tutorial seminar on renormalization group theory while visiting the Aspen Center for Physics. Ken Wilson's thesis advisor, Murray Gell-Mann, was in the audience. In the discussion period after the seminar Gell-Mann expressed his appreciation for the theoretical structure created by his famous student that I had set out in its generality, and he asked: "But tell me, what has all that got to do with the work Francis Low and I did so many years ago?" [108] In response, I explained the connecting thread and the far-reaching intellectual inspiration: certainly there is a thread but—to echo my previous comments—I believe that its length is comparable to that reaching from Maxwell, Boltzmann, and ideal gases to Gibbs' general conception of ensembles, partition functions, and their manifold inter-relations.

XI. THE CONSTRUCTION OF RENORMALIZATION GROUP TRANSFORMATIONS: THE EPSILON EXPANSION

In telling my story I have purposefully incorporated a large dose of hindsight by emphasizing the importance of viewing a particular physical system—or its reduced Hamiltonian, $\bar{\mathcal{H}}(t,h,\cdots)$: see Eq. (24)—as specifying only a relatively small manifold in a large space, $\mathbb{H}$, of possible Hamiltonians. But why is that more than a mere formality? One learns the answer as soon as, following Wilson (1975, 1983), one attempts to implement Kadanoff's scaling description in some concrete, computational way. In Kadanoff's picture (in common with the Gell-Mann-Low, Callan-Symanzik, and general QFT viewpoints) one *assumes* that after a "rescaling" or "renormalization" the new, renormalized Hamiltonian (or, in QFT, the Lagrangean) has the *identical form* except for the renormalization of a single parameter (or coupling constant) or—as in Kadanoff's picture—of at most a small *fixed* number, like the temperature t and field h. That assumption is the dangerous and, unless one is especially lucky,[109] the *generally false* step! Wilson (1975, p. 592) has described his "liberation" from this straight jacket and how the freedom gained opened the door to the systematic design of RG transformations.

To explain, we may state matters as follows: Gibbs' prescription for calculating the partition function—see Eq. (25)—tells us to sum (or to integrate) over the allowed values of *all* the N spin variables $s_{\mathbf{x}}$. But this is very difficult! Let us, instead, adopt a strategy of "divide and conquer," by separating the set $\{s_{\mathbf{x}}\}$ of N spins into two groups: first, $\{s_{\mathbf{x}}^{<}\}$, consisting of $N'=N/b^d$ spins which we will leave as untouched fluctuating variables; and, second, $\{s_{\mathbf{x}}^{>}\}$ consisting of the remaining $N-N'$ spin variables over which we will integrate (or sum) so that they drop out of the problem. If we draw inspiration from Kadanoff's (or Buckingham's[110]) block picture we might reasonably choose to integrate over all but one central spin in each block of b^d spins. This process, which Kadanoff has dubbed "decimation" (after the Roman military practice), preserves translational invariance and clearly represents a concrete form of "coarse graining" (which, in earlier days, was typically cited as a way to derive, "in principle," mesoscopic or Landau-Ginzburg descriptions).

Now, after taking our partial trace we must be left with some new, *effective Hamiltonian*, say, $\bar{\mathcal{H}}_{\rm eff}[s^{<}]$, involving only the preserved, unintegrated spins. On reflection one realizes that, in order to be faithful to the original physics, such an effective Hamiltonian must be defined via its Boltzmann factor: recalling our brief outline of statistical mechanics, that leads directly to the explicit formula

$$e^{\bar{\mathcal{H}}_{eff}[s^{<}]}=\mathrm{Tr}^{s^{>}}_{N-N'}\{e^{\bar{\mathcal{H}}[s^{<}\cup s^{>}]}\}, \tag{33}$$

where the 'union', $s^{<}\cup s^{>}$, simply stands for the full set of original spins $s\equiv\{s_{\mathbf{x}}\}$. By a spatial rescaling, as in Eq. (27), and a relabelling, namely, $s_{\mathbf{x}}^{<}\Rightarrow s'_{\mathbf{x}'}$, we obtain the "renormalized Hamiltonian," $\bar{\mathcal{H}}'[s']\equiv\bar{\mathcal{H}}_{eff}[s^{<}]$. Formally, then, we have succeeded in defining an *explicit renormalization transformation*. We will write

$$\bar{\mathcal{H}}'[s']=\mathbb{R}_b\{\bar{\mathcal{H}}[s]\}, \tag{34}$$

where we have elected to keep track of the spatial rescaling factor, b, as a subscript on the RG operator $\mathbb{R}$.

Note that if we complete the Gibbsian prescription by taking the trace over the renormalized spins we simply get back to the desired partition function, $Z_N[\mathcal{H}]$. (The formal derivation for those who might be interested is set out in the footnote below.[111]) Thus nothing has been lost: the renormalized Hamiltonian retains all the ther-

[107] See, especially, Wilson (1975) on page 796, column 1, and Footnote 10 in Wilson (1971a).

[108] That is, in Gell-Mann and Low (1954).

[109] See Footnote 94 above and Nelson and Fisher (1975) and Fisher (1983).

[110] Recall Footnote 86 above.

[111] We start with the definition Eq. (33) and recall Eq. (25) to obtain

$$Z_{N'}[\bar{\mathcal{H}}']\equiv\mathrm{Tr}^{s'}_{N'}\{e^{\bar{\mathcal{H}}[s']}\}=\mathrm{Tr}^{s^{<}}_{N'}\{e^{\bar{\mathcal{H}}_{eff}[s^{<}]}\}$$

$$=\mathrm{Tr}^{s^{<}}_{N'}\,\mathrm{Tr}^{s^{>}}_{N-N'}\{e^{\bar{\mathcal{H}}[s^{<}\cup s^{>}]}\}=\mathrm{Tr}^{s}_{N}\{e^{\bar{\mathcal{H}}[s]}\}=Z_N[\bar{\mathcal{H}}],$$

from which the free energy $f[\bar{\mathcal{H}}]$ follows via Eq. (26).

modynamic information. On the other hand, experience suggests that, rather than try to compute Z_N directly from $\bar{\mathcal{H}}'$, it will prove more fruitful to *iterate* the transformation so obtaining a sequence, $\bar{\mathcal{H}}^{(l)}$, of renormalized Hamiltonians, namely,

$$\bar{\mathcal{H}}^{(l)} = \mathbb{R}_b[\bar{\mathcal{H}}^{(l-1)}] = \mathbb{R}_{b^l}[\bar{\mathcal{H}}], \tag{35}$$

with $\bar{\mathcal{H}}^{(0)} \equiv \bar{\mathcal{H}}$, $\bar{\mathcal{H}}^{(1)} = \bar{\mathcal{H}}'$. It is these iterations that give rise to the *semigroup* character of the RG transformation.[112]

But now comes the crux: thanks to the rescaling and relabelling, the microscopic variables $\{s'_{\mathbf{x}'}\}$ are, indeed, completely equivalent to the original spins $\{s_{\mathbf{x}}\}$. However, when one proceeds to determine the nature of $\bar{\mathcal{H}}_{eff}$, and thence of $\bar{\mathcal{H}}'$, by using the formula (33), one soon discovers that one *cannot* expect the original form of $\bar{\mathcal{H}}$ to be reproduced in $\bar{\mathcal{H}}_{eff}$. Consider, for concreteness, an initial Hamiltonian, $\mathcal{H}$, that describes Ising spins ($s_{\mathbf{x}} = \pm 1$) on a square lattice in zero magnetic field with just nearest-neighbor interactions of coupling strength $K_1 = J_1/k_B T$: in the most conservative Kadanoff picture there must be *some* definite recursion relation for the renormalized coupling, say, $K'_1 = \mathcal{T}_1(K_1)$, embodied in a definite function $\mathcal{T}(\cdot)$. But, in fact, one finds that $\bar{\mathcal{H}}_{eff}$ must actually contain *further* nonvanishing spin couplings, K_2, between second-neighbor spins, K_3, between third-neighbors, and so on up to *indefinitely* high orders. Worse still, four-spin coupling terms like $K_{\square 1} s_{\mathbf{x}_1} s_{\mathbf{x}_2} s_{\mathbf{x}_3} s_{\mathbf{x}_4}$ appear in $\bar{\mathcal{H}}_{eff}$, again for *all* possible arrangements of the four spins! And also six-spin couplings, eight-spin couplings, $\cdots$. Indeed, for any given set Q of $2m$ Ising spins on the lattice (and its translational equivalents), a nonvanishing coupling constant, K'_Q, is generated and appears in $\mathcal{H}'$!

The only saving grace is that further iteration of the decimation transformation Eq. (33) cannot (in zero field) lead to anything worse! In other words the space $\mathbb{H}_{\mathrm{Is}}$ of Ising spin Hamiltonians in zero field may be specified by the infinite set $\{K_Q\}$, of all possible spin couplings, and is *closed* under the decimation transformation Eq. (33). Formally, one can thus describe $\mathbb{R}_b$ by the full set of *recursion relations*

$$K'_P = \mathcal{T}_P(\{K_Q\}) \quad (\text{all } P). \tag{36}$$

Clearly, this answers our previous questions as to what becomes of the complicated across-the-faces-of-the-block interactions in the original Kadanoff picture: They actually carry the renormalized Hamiltonian *out* of the (too small) manifold of nearest-neighbor Ising models and introduce (infinitely many) further couplings. The resulting situation is portrayed schematically in Fig. 5: the renormalized manifold for $\mathcal{H}'(t',h')$ generally has no overlap with the original manifold. Further iterations, and *continuous* [see Eq. (40) below] as against discrete RG transformations, are suggested by the flow lines or "trajectories" also shown in Fig. 5. We will return to some of the details of these below.

In practice, the naive decimation transformation specified by Eq. (33) generally fails as a foundation for useful calculations.[113] Indeed, the design of effective RG transformations turns out to be an art more than a science: there is no standard recipe! Nevertheless, there are guidelines: the general philosophy enunciated by Wilson and expounded well, for example, in a recent lecture by Shankar treating fermionic systems,[114] is to attempt to *eliminate* first those microscopic variables or degrees of freedom of "least direct importance" to the macroscopic phenomenon under study, while *retaining* those of most importance. In the case of ferromagnetic or gas-liquid critical points, the phenomena of most significance take place on long length scales—the correlation length, ξ, diverges; the critical point correlations, $G_c(\mathbf{r})$, decay slowly at long-distances; long-range order sets in below T_c.

Thus in his first, breakthrough articles in 1971, Wilson used an ingenious "phase-space cell" decomposition for continuously variable scalar spins (as against ± 1 Ising spins) to treat a lattice Landau-Ginzburg model with a general, single-spin or 'on-site' potential $V(s_{\mathbf{x}})$ acting on each spin, $-\infty < s_{\mathbf{x}} < \infty$. Blocks of cells of the smallest spatial extent were averaged over to obtain a single, renormalized cell of twice the linear size (so that $b=2$). By making sufficiently many simplifying approximations Wilson obtained an explicit *nonlinear, integral recursion relation* that transformed the l-times renormalized potential, $V^{(l)}(\cdot)$, into $V^{(l+1)}(\cdot)$. This recursion relation could be handled by computer and led to a *specific numerical estimate* for the exponent ν for $d=3$ dimensions that was *quite different* from the classical value $\frac{1}{2}$ (and from the results of any previously soluble models like the spherical model[115]). On seeing that result, I knew that a major barrier to progress had been overcome!

---

[112] Thus successive decimations with scaling factors b_1 and b_2 yield the quite general relation

$$\mathbb{R}_{b_2}\mathbb{R}_{b_1} = \mathbb{R}_{b_2 b_1},$$

which essentially defines a unitary *semigroup* of transformations. See Footnotes 3 and 78 above, and the formal algebraic definition in MacLane and Birkhoff (1967): a unitary semigroup (or 'monoid') is a set M of elements, u, v, w, x,$\cdots$ with a binary operation, $xy = w \in M$, which is associative, so $v(wx) = (vw)x$, and has a unit u, obeying $ux = xu = x$ (for all $x \in M$)—in RG theory, the unit transformation corresponds simply to $b=1$. Hille (1948) and Riesz and Sz.-Nagy (1955) describe semigroups within a continuum, functional analysis context and discuss the existence of an infinitesimal generator when the flow parameter l is defined for continuous values $l \geqslant 0$: see Eq. (40) below and Wilson's (1971a) introductory discussion.

[113] See Kadanoff and Niemeijer in Gunton and Green (1974), Niemeijer and van Leeuwen (1976), Fisher (1983).

[114] See R. Shankar in Cao (1998) and Shankar (1994).

[115] For accounts of the critical behavior of the spherical model, see Fisher (1966a), where long-range forces were also considered, and, e.g., Stanley (1971), Baxter (1982), and Fisher (1983).

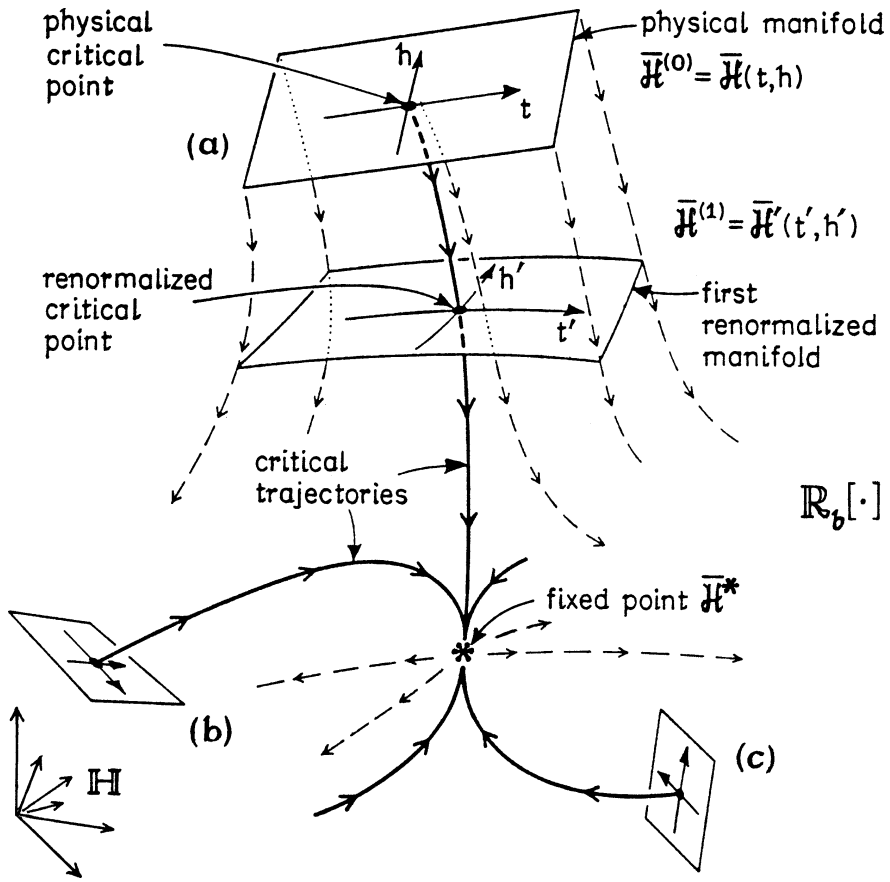


FIG. 5. A depiction of the space of Hamiltonians $\mathbb{H}$ — compare with Fig. 2—showing initial or physical manifolds [labelled (a), (b), $\cdots$, as in Fig. 2] and the flows induced by repeated application of a discrete RG transformation $\mathbb{R}_b$ with a spatial rescaling factor b (or induced by a corresponding continuous or differential RG). Critical trajectories are shown bold: they all terminate, in the region of $\mathbb{H}$ shown here, at a fixed point $\overline{\mathcal{H}}^*$. The full space contains, in general, other nontrivial, critical fixed points, describing multicritical points and distinct critical-point universality classes; in addition, trivial fixed points, including high-temperature "sinks" with no outflowing or relevant trajectories, typically appear. *Lines of fixed points* and other more complex structures may arise and, indeed, play a crucial role in certain problems. [After Fisher (1983).]

I returned from a year's sabbatic leave at Stanford University in the summer of 1971, by which time Ken Wilson's two basic papers were in print. Shortly afterwards, in September, again while walking to lunch as I recall, Ken Wilson discussed his latest results from the nonlinear recursion relation with me. Analytical expressions could be obtained by expanding $V^{(l)}(s)$ in a power series:

$$V^{(l)}(s) = r_l\, s^2 + u_l\, s^4 + v_l\, s^6 + \cdots . \tag{37}$$

If truncated at quadratic order one had a soluble model—the Gaussian model (or free-field theory)—and the recursion relation certainly worked *exactly* for that! But to have a nontrivial model, one had to start not only with r_0 (as, essentially, the temperature variable) but, as a minimum, one also had to include $u_0 > 0$: the model then corresponded to the well known $\lambda\varphi^4$ field theory. Although one might, thus, initially set $v_0 = w_0 = \cdots = 0$, all these higher order terms were immediately generated under renormalization; furthermore, there was no reason for u_0 to be small and, for this reason and others, the standard field-theoretic perturbation theories were ineffective.

Now, I had had a long-standing interest in the effects of the *spatial dimensionality* d on singular behavior in various contexts:[116] so that issue was raised for Ken's recursion relation. Indeed, d appeared simply as an explicit parameter. It then became clear that $d=4$ was a special case in which the leading order corrections to the Gaussian model vanished. Furthermore, above $d=4$ dimensions classical behavior reappeared in a natural way (since the parameters u_0, v_0, ... all then became irrelevant). These facts fitted in nicely with the known special role of $d=4$ in other situations.[117]

For $d=3$, however, one seemed to need the infinite set of coefficients in Eq. (37) which all coupled together

[116]Fisher and Gaunt (1964), Fisher (1966a, 1966b; 1967c; 1972).

[117]See references in the previous footnote and Larkin and Khmel'nitskii (1969), especially Appendix 2.

under renormalization. But I suggested that maybe one could treat the dimensional deviation, $\epsilon = 4 - d$, as a small, *nonintegral* parameter in analyzing the recursion relations for $d < 4$. Ken soon showed this was effective! Furthermore, the recursion relations proved to be *exact* to leading order in ϵ (so that if one replaced $b = 2$ by a general value, the expected universal results were indeed, independent of b). A paper, entitled by Ken, "Critical Exponents in 3.99 Dimensions" was shortly written, submitted, and published:[118] it contained the first general formula for a nonclassical exponent, namely, $\gamma = 1 + \frac{1}{6}\epsilon + O(\epsilon^2)$.

It transpired, however, that the perturbation parameter ϵ provided more—namely, a systematic way of ordering the infinite set of discrete recursion relations not only for the expansion coefficients of $V^{(l)}(s)$ but also for further terms in the appropriate full space $\mathbb{H}$, involving spatial gradients or, equivalently but more usefully, the momenta or wave vectors $\mathbf{q}_i$ labelling the spin variables $\hat{s}_{\mathbf{q}}$, now re-expressed in Fourier space. With that facility in hand, the previous approximations entailed in the phase-space cell analysis could be dispensed with. Wilson then saw that he could precisely implement his *momentum-shell renormalization group*[119]—subsequently one of the most-exploited tools in critical phenomena studies!

In essence this transformation is like decimation[120] except that the division of the variables in Eq. (33) is made in momentum space: for ferromagnetic or gas-liquid-type critical points the set $\{\hat{s}_q^<\}$ contains those 'long-wavelength' or 'low-momentum' variables satisfying $|\mathbf{q}| \leqslant q_\Lambda / b$, where $q_\Lambda = \pi/a$ is the (ultraviolet) momentum cutoff implied by the lattice structure. Conversely, the 'short-wavelength', 'high-momentum' spin components $\{\hat{s}_{\mathbf{q}}^>\}$ having wave vectors lying in the momentum-space *shell*: $q_\Lambda / b < |\mathbf{q}| \leqslant q_\Lambda$, are integrated out. The spatial rescaling now takes the form

$$\mathbf{q} \Rightarrow \mathbf{q}' = b\mathbf{q}, \tag{38}$$

as follows from Eq. (27); but in analogy to $\zeta(b)$ in Eq. (28), a *nontrivial spin rescaling factor* ("multiplicative-wave function renormalization" in QFT) is introduced via

$$\hat{s}_{\mathbf{q}} \Rightarrow \hat{s}'_{\mathbf{q}'} = \hat{s}_{\mathbf{q}} / \hat{c}[b, \bar{\mathcal{H}}]. \tag{39}$$

The crucially important rescaling factor $\hat{c}$ takes the form $b^{d-\omega}$ and must be *tuned* in the critical region of interest [which leads to $\omega = \frac{1}{2}(d - 2 + \eta)$: compare with Eq. (4)]. It is also worth mentioning that by letting $b \to 1+$, one can derive a *differential* or continuous flow RG and rewrite the recursion relation Eq. (34) as[121]

$$\frac{d}{dl}\bar{\mathcal{H}} = \mathbb{B}[\bar{\mathcal{H}}]. \tag{40}$$

Such continuous flows are illustrated in Figs. 4 and 5. (If it happens that $\mathcal{H}$ can be represented, in general only approximately, by a single coupling constant, say, g, then $\mathbb{B}$ reduces to the so-called beta-function $\beta(g)$ of QFT.)

For deriving ϵ expansions on the basis of the momentum shell RG, Feynman-graph perturbative techniques as developed for QFT prove very effective.[122] They enter basically because one can take $u_0 = O(\epsilon)$ small and they play a role both in efficiently organizing the calculation and in performing the essential integrals (particularly for systems with simple propagators and vertices).[123] Capitalizing on his field-theoretic expertise, Wilson obtained, in only a few weeks after submitting the first article, *exact expansions* for the exponents ν, γ, and ϕ to order ϵ^2 (and, by scaling, for all other exponents).[124] Furthermore, the anomalous dimension—defined in Eq. (2) at the beginning of our story—was calculated exactly to order ϵ^3: I cannot resist displaying this striking result, namely,

$$\eta = \frac{(n+2)}{2(n+8)^2}\epsilon^2 + \frac{(n+2)}{2(n+8)^2}\left[\frac{6(3n+14)}{(n+8)^2} - \frac{1}{4}\right]\epsilon^3 + O(\epsilon^4), \tag{41}$$

where the symmetry parameter n denotes the number of components of the microscopic spin vectors, $\vec{s}_{\mathbf{x}} \equiv (s_{\mathbf{x}}^\mu)_{\mu=1,\ldots,n}$, so that one has just $n = 1$ for Ising

[118] Wilson and Fisher (1972). The first draft was written by Ken Wilson who graciously listed the authors in alphabetical order.

[119] See Wilson and Fisher (1972) Eq. (18) and the related text.

[120] A considerably more general form of RG transformation can be written as

$$\exp(\bar{\mathcal{H}}'[s']) = \mathrm{Tr}_N^s\{\mathcal{R}_{N',N}(s'; s)\exp(\bar{\mathcal{H}}[s])\},$$

where the trace is taken over the full set of original spins s. The $N' = N/b^d$ renormalized spins $\{s'\}$ are introduced via the RG kernel $\mathcal{R}_{N',N}(s'; s)$ which incorporates spatial and spin rescalings, etc., and which should satisfy a trace condition to ensure the partition-function-preserving property (see Footnote 111) which leads to the crucial free-energy flow equation: see Eq. (43) below. The decimation transformation, the momentum-shell RG, and other transformations can be written in this form.

[121] See Wilson (1971a) and Footnote 112 above: in this form the RG semigroup can typically be extended to an Abelian group (MacLane and Birkhoff, 1967); but as already stressed, this fact plays a negligible role.

[122] See Wilson (1972), Brézin, Wallace, and Wilson (1972), Wilson and Kogut (1974), the reviews Brézin, Le Guillou, and Zinn-Justin (1976), and Wallace (1976), and the texts Amit (1978) and Itzykson and Drouffe (1989).

[123] Nevertheless, many more complex situations arise in condensed matter physics for which the formal application of graphical techniques without an adequate understanding of the appropriate RG structure can lead one seriously astray.

[124] See Wilson (1972) which was received on 1 December 1971 while Wilson and Fisher (1972) carries a receipt date of 11 October 1971.

spins.[125] Over the years these expansions have been extended to order ϵ^5 (and ϵ^6 for η)[126] and many further related expansions have been developed.[127]

XII. FLOWS, FIXED POINTS, UNIVERSALITY AND SCALING

To complete my story—and to fill in a few logical gaps over which we have jumped—I should explain how Wilson's construction of RG transformations in the space $\mathbb{H}$ enables RG theory to accomplish the "tasks" set out above in Sec. VIII. As illustrated in Fig. 5, the recursive application of an RG transformation $\mathbb{R}_b$ induces a *flow* in the space of Hamiltonians, $\mathbb{H}$. Then one observes that "sensible," "reasonable," or, better, "well-designed" RG transformations are *smooth*, so that points in the original physical manifold, $\bar{\mathcal{H}}^{(0)}=\bar{\mathcal{H}}(t,h)$, that are close, say in temperature, remain so in $\bar{\mathcal{H}}^{(1)}\equiv\bar{\mathcal{H}}'$, i.e., under renormalization, and likewise as the flow parameter l increases, in $\bar{\mathcal{H}}^{(l)}$. Notice, incidentally, that since the spatial scale renormalizes via $\mathbf{x}\Rightarrow\mathbf{x}'=b^l\mathbf{x}$ one may regard

$$l=\log_b(|\mathbf{x}'|/|\mathbf{x}|), \tag{42}$$

as measuring, logarithmically, the scale on which the system is being described—recall the physical *scale dependence of parameters* discussed in Sec. IV; but note that, in general, the *form* of the Hamiltonian is also changing as the "scale" is changed or l increases. Thus a partially renormalized Hamiltonian can be expected to take on a more-or-less generic, mesoscopic form: Hence it represents an appropriate candidate to give meaning to a Landau-Ginzburg or, now, LGW effective Hamiltonian: recall the discussion of Landau's work in Sec. II.

Thanks to the smoothness of the RG transformation, if one knows the free energy $f_l\equiv f[\bar{\mathcal{H}}^{(l)}]$ at the l-th stage of renormalization, then one knows the *original* free energy $f[\bar{\mathcal{H}}]$ and its critical behavior: explicitly one has[128]

$$f(t,h,\cdots)\equiv f[\bar{\mathcal{H}}]=b^{-dl}\,f[\bar{\mathcal{H}}^{(l)}]\equiv b^{-dl}f_l(t^{(l)},h^{(l)},\cdots). \tag{43}$$

Furthermore, the smoothness implies that all the universal critical properties are preserved under renormalization. Similarly one finds[129] that the critical point of $\bar{\mathcal{H}}^{(0)}\equiv\bar{\mathcal{H}}$ maps on to that of $\bar{\mathcal{H}}^{(1)}\equiv\bar{\mathcal{H}}'$, and so on, as illustrated by the bold flow lines in Fig. 5. Thus it is instructive to follow the *critical trajectories* in $\mathbb{H}$, i.e., those RG flow lines that emanate from a physical critical point. In principle, the topology of these trajectories could be enormously complicated and even chaotic: in practice, however, for a well-designed or "apt" RG transformation, one most frequently finds that the critical flows terminate—or, more accurately, come to an asymptotic halt—at a *fixed point* $\bar{\mathcal{H}}^*$, of the RG: see Fig. 5. Such a fixed point is defined, via Eqs. (34) or (40), simply by

$$\mathbb{R}_b[\bar{\mathcal{H}}^*]=\bar{\mathcal{H}}^* \quad \text{or} \quad \mathbb{B}[\bar{\mathcal{H}}^*]=0. \tag{44}$$

One then searches for fixed-point solutions: the role of the fixed-point equation is, indeed, roughly similar to that of Schrödinger's Equation $\mathcal{H}\Psi=E\Psi$, for stationary states Ψ_k of energy E_k in quantum mechanics.

Why are the fixed points so important? Some, in fact, are *not*, being merely *trivial*, corresponding to *no interactions* or to *all spins frozen*, etc. But the *non*trivial fixed points represent critical states; furthermore, the nature of their criticality, and of the free energy in their neighborhood, must, as explained, be *identical* to that of all those distinct Hamiltonians whose critical trajectories converge to the same fixed point! In other words, a particular fixed point *defines* a *universality class* of critical behavior which "governs," or "attracts" all those systems whose critical points eventually map onto it: see Fig. 5.

Here, then we at last have the natural explanation of *universality*: systems of quite different physical character may, nevertheless, belong to the domain of attraction of the *same* fixed point $\bar{\mathcal{H}}^*$ in $\mathbb{H}$. The distinct sets of inflowing trajectories reflect their varying physical content of associated irrelevant variables and the corresponding nonuniversal rates of approach to the asymptotic power laws dicated by $\mathcal{H}^*$: see Eq. (22).

From each critical fixed point, there flow at least two "unstable" or outgoing trajectories. These correspond to one or more *relevant* variables, specifically, for the case illustrated in Fig. 5, to the temperature or thermal field, t, and the magnetic or ordering field, h. See also Fig. 4. If there are further relevant trajectories then, as discussed in Sec. VII, one can expect *crossover* to different critical behavior. In the space $\mathbb{H}$, such trajectories will then typically lead to distinct fixed points describing (in general) completely new universality classes.[130]

---

[125] See, e.g., Fisher (1967b, 1974b, 1983), Kadanoff *et al.* (1967), Stanley (1971), Aharony (1976), Patashinskii and Pokrovskii (1979).

[126] See Gorishny, Larin, and Tkachov (1984) but note that the $O(\epsilon^5)$ polynomials in n are found accurately but some coefficients are known only within uncertainties.

[127] Recall Footnote 35.

[128] Recall the partition-function-preserving property set out in Footnote 111 above which actually implies the basic relation Eq. (43).

[129] See Wilson (1971a), Wilson and Kogut (1974), and Fisher (1983).

[130] A skeptical reader may ask: "But what if no fixed points are found?" This can well mean, as it has frequently meant in the past, simply that the chosen RG transformation was poorly designed or "not apt." On the other hand, a fixed point represents only the simplest kind of asymptotic flow behavior: other types of asymptotic flow may well be identified and translated into physical terms. Indeed, near certain types of trivial fixed point, such procedures, long ago indicated by Wilson (1971a, Wilson and Kogut, 1974), *must* be implemented: see, e.g., Fisher and Huse (1985).

But what about *power laws* and *scaling*? The answer to this question was already sketched in Sec. VIII; but we will recapitulate here, giving a few more technical details. However, trusting readers or those uninterested in the analysis are urged to *skip to the next section*!

That said, one must start by noting that the smoothness of a well-designed RG transformation means that it can always be expanded locally—to at least some degree—in a Taylor series.[131] It is worth stressing that it is this very property that fails for free energies in a critical region: to regain this ability, the large space of Hamiltonians is crucial. Near a fixed point satisfying Eq. (43) we can, therefore, rather generally expect to be able to *linearize* by writing

$$\mathbb{R}_b[\bar{\mathcal{H}}^* + g\mathcal{Q}] = \bar{\mathcal{H}}^* + g\mathbb{L}_b\mathcal{Q} + o(g) \tag{45}$$

as $g \to 0$, or in differential form,

$$\frac{d}{dl}(\bar{\mathcal{H}}^* + g\mathcal{Q}) = g\mathbb{B}_1\mathcal{Q} + o(g). \tag{46}$$

Now $\mathbb{L}_b$ and $\mathbb{B}_1$ are *linear operators* (albeit acting in a large space $\mathbb{H}$). As such we can seek eigenvalues and corresponding "eigenoperators", say $\mathcal{Q}_k$ (which will be "partial Hamiltonians"). Thus, in parallel to quantum mechanics, we may write

$$\mathbb{L}_b\mathcal{Q}_k = \Lambda_k(b)\mathcal{Q}_k \quad \text{or} \quad \mathbb{B}_1\mathcal{Q}_k = \lambda_k\mathcal{Q}_k, \tag{47}$$

where, in fact, (by the semigroup property) the eigenvalues must be related by $\Lambda_k(b) = b^{\lambda_k}$. As in any such linear problem, knowing the spectrum of eigenvalues and eigenoperators or, at least, its dominant parts, tells one much of what one needs to know. Reasonably, the $\mathcal{Q}_k$ should form a basis for a general expansion

$$\bar{\mathcal{H}} \cong \bar{\mathcal{H}}^* + \sum_{k\geq 1} g_k\mathcal{Q}_k. \tag{48}$$

Physically, the expansion coefficient g_k ($\equiv g_k^{(0)}$) then represents the thermodynamic field[132] conjugate to the "critical operator" $\mathcal{Q}_k$ which, in turn, will often be close to some combination of *local* operators. Indeed, in a characteristic critical-point problem one finds two *relevant operators*, say $\mathcal{Q}_1$ and $\mathcal{Q}_2$ with $\lambda_1, \lambda_2 > 0$. Invariably, one of these operators can, say by its symmetry, be identified with the local energy density, $\mathcal{Q}_1 \cong \mathcal{E}$, so that $g_1 \cong t$ is the thermal field; the second then characterizes the order parameter, $\mathcal{Q}_2 \cong \Psi$ with field $g_2 \cong h$. Under renormalization each g_k varies simply as $g_k^{(l)} \approx b^{\lambda_k l} g_k^{(0)}$.

Finally,[133] one examines the flow equation (43) for the free energy. The essential point is that the degree of renormalization, b^l, can be *chosen* as large as one wishes. When $t \to 0$, i.e., in the critical region which it is our aim to understand, a good choice proves to be $b^l = 1/|t|^{1/\lambda_1}$, which clearly diverges to ∞. One then finds that Eq. (43) leads to the *basic scaling relation* Eq. (19) which we will rewrite here in greater generality as

$$f_s(t, h, \cdots, g_j, \cdots) \approx |t|^{2-\alpha}\mathcal{F}\left(\frac{h}{|t|^{\Delta}}, \cdots, \frac{g_j}{|t|^{\phi_j}}, \cdots\right). \tag{49}$$

This is the essential result: recall, for example, that it leads to the "collapse" of equation-of-state data as described in Sec. VI.

Now, however, the critical exponents can be expressed directly in terms of the RG eigenexponents λ_k (for the fixed point in question). Specifically one finds

$$2 - \alpha = \frac{d}{\lambda_1}, \quad \Delta = \frac{\lambda_2}{\lambda_1}, \quad \phi_j = \frac{\lambda_j}{\lambda_1}, \quad \text{and} \quad \nu = \frac{1}{\lambda_1}. \tag{50}$$

Then, as already explained in Secs. VI and VII, the sign of a given ϕ_j and, hence, of the corresponding λ_j determines the *relevance* (for $\lambda_j > 0$), *marginality* (for $\lambda_j = 0$), or *irrelevance* (for $\lambda_j < 0$) of the corresponding critical operator $\mathcal{Q}_j$ (or "perturbation") and of its conjugate field g_j: this field might, but for most values of j will *not*, be under direct experimental control. As explained previously, all exponent relations (15), (20), etc., follow from scaling, while the first and last of the equations (50) yield the *hyperscaling relation* Eq. (32).

When there are no marginal variables and the least negative ϕ_j is larger than unity in magnitude, a simple scaling description will usually work well and the Kadanoff picture almost applies. When there are *no* relevant variables and only one or a few *marginal variables*, field-theoretic perturbative techniques of the Gell-Mann-Low (1954), Callan-Symanzik[134] or so-called "parquet diagram" varieties[135] may well suffice (assuming the dominating fixed point is sufficiently simple to be well understood). There may then be little incentive for specifically invoking general RG theory. This seems, more or less, to be the current situation in QFT and it applies also in certain condensed matter problems.[136]

XIII. CONCLUSIONS

My tale is now told: following Wilson's 1971 papers and the introduction of the ϵ-expansion in 1972 the significance of the renormalization group approach in

[131] See Wilson (1971a), Wilson and Kogut (1974), Fisher (1974b), Wegner (1972, 1976), Kadanoff (1976).

[132] Reduced, as always, by the factor $1/k_BT$: see e.g., Eq. (18).

[133] See references in Footnote 131.

[134] See Wilson (1975), Brézin *et al.* (1976), Amit (1978), Itzykson and Drouffe (1989).

[135] Larkin and Khmel'nitskii (1969).

[136] See, e.g., the case of dipolar Ising-type ferromagnets in $d = 3$ dimensions investigated experimentally by Ahlers, Kornblit, and Guggenheim (1975) following theoretical work by Larkin and Khmel'nitskii (1969) and Aharony (see 1976, Sec. 4E).

statistical mechanics was soon widely recognized[137] and exploited by many authors interested in critical and multicritical phenomena and in other problems in the broad area of condensed matter physics, physical chemistry, and beyond. Some of these successes have already been mentioned in order to emphasize, in particular, those features of the full RG theory that are of general significance in the wide range of problems lying beyond the confines of quantum field theory and fundamental high-energy physics. But to review those developments would go beyond the mandate of this Colloquium.[138]

A further issue is the relevance of renormalization group concepts to quantum field theory. I have addressed that only in various peripheral ways. Insofar as I am by no means an expert in quantum field theory, that is not inappropriate; but perhaps one may step back a moment and look at QFT from the general philosophical perspective of understanding complex, interacting systems. Then, I would claim, statistical mechanics is a central science of great intellectual significance—as just one reminder, the concepts of ''spin-glasses'' and the theoretical and computational methods developed to analyze them (such as ''simulated annealing'') have proved of interest in physiology for the study of neuronal networks and in operations research for solving hard combinatorial problems. In that view, even if one focuses only on the physical sciences, the land of statistical physics is broad, with many dales, hills, valleys and peaks to explore that are of relevance to the real world and to our ways of thinking about it. Within that land there is an island, surrounded by water: I will not say ''by a moat'' since, these days, more and broader bridges happily span the waters and communicate with the mainland! That island is devoted to what was ''particle physics'' and is now ''high-energy physics'' or, more generally, to the deepest lying and, in that sense, the ''most fundamental'' aspects of physics. The reigning theory on the island is quantum field theory—the magnificent set of ideas and techniques that inspired the symposium[139] that lead to this Colloquium. Those laboring on the island have built most impressive skyscrapers reaching to the heavens!

Nevertheless, from the global viewpoint of statistical physics—where many degrees of freedom, the ever-present fluctuations, and the diverse spatial and temporal scales pose the central problems—quantum field theory may be regarded as describing a rather special set of statistical mechanical models. As regards applications they have been largely restricted to $d=4$ spatial dimensions [more physically, of course to (3+1) dimensions] although in the last decade *string theory* has dramatically changed that! The practitioners of QFT insist on the preeminence of some pretty special symmetry groups, the Poincaré group, $SU(3)$, and so on, which are not all so ''natural'' at first sight—even though the role of guage theories as a unifying theme in modeling nature has been particularly impressive. But, of course, we know these special features of QFT are not matters of choice—rather, they are forced on us by our explorations of Nature itself. Indeed, as far as we know presently, there is only one high-energy physics; whereas, by contrast, the ingenuity of chemists, materials scientists, and of Life itself, offers a much broader, multifaceted and varied panorama of systems to explore both conceptually and in the laboratory.

From this global standpoint, renormalization group theory represents a theoretical tool of depth and power. It first flowered luxuriantly in condensed matter physics, especially in the study of critical phenomena. But it is ubiquitous because of its potential for linking physical behavior across disparate scales; its ideas and techniques play a vital role in those cases where the fluctuations on many different physical scales truly interact. But it provides a valuable perspective—through concepts such as 'relevance,' 'marginality' and 'irrelevance,' even when scales are well separated! One can reasonably debate how vital renormalization group concepts are for quan-

[137]Footnote 86 drew attention to the first international conference on critical phenomena organized by Melville S. Green and held in Washington in April 1965. Eight years later, in late May 1973, Mel Green, with an organizing committee of J. D. Gunton, L. P. Kadanoff, K. Kawasaki, K. G. Wilson, and the author, mounted another conference to review the progress in theory in the previous decade. The meeting was held in a Temple University Conference Center in rural Pennsylvania. The proceedings (Gunton and Green, 1974) entitled *Renormalization Group in Critical Phenomena and Quantum Field Theory*, are now mainly of historical interest. The discussions were recorded in full but most papers only in abstract or outline form. Whereas in the 1965 conference the overwhelming number of talks concerned experiments, now only J.M.H. (Anneke) Levelt Sengers and Guenter Ahlers spoke to review experimental findings in the light of theory. Theoretical talks were presented, in order, by P. C. Martin, Wilson, Fisher, Kadanoff, B. I. Halperin, E. Abrahams, Niemeijer (with van Leeuwen), Wegner, Green, Suzuki, Fisher and Wegner (again), E. K. Riedel, D. J. Bergman (with Y. Imry and D. Amit), M. Wortis, Symanzik, Di Castro, Wilson (again), G. Mack, G. Dell-Antonio, J. Zinn-Justin, G. Parisi, E. Brézin, P. C. Hohenberg (with Halperin and S.-K. Ma) and A. Aharony. Sadly, there were no participants from the Soviet Union but others included R. Abe, G. A. Baker, Jr., T. Burkhardt, R. B. Griffiths, T. Lubensky, D. R. Nelson, E. Siggia, H. E. Stanley, D. Stauffer, M. J. Stephen, B. Widom and A. Zee. As the lists of names and participants illustrates, many active young theorists had been attracted to the area, had made significant contributions, and were to make more in subsequent years.

[138]Some reviews already mentioned that illustrate applications are Fisher (1974b), Wilson (1975), Wallace (1976), Aharony (1976), Patashinskii and Pokrovskii (1979), Nelson (1983), and Creswick *et al.* (1992). Beyond these, attention should be drawn to the notable article by Hohenberg and Halperin (1977) that reviews dynamic critical phenomena, and to many articles on further topics in the Domb and Lebowitz series *Phase Transitions and Critical Phenomena*, Vols. 7 and beyond (Academic, London, 1983).

[139]See Cao (1998).

tum field theory itself. Certain aspects of the full theory do seem important because Nature teaches us, and particle physicists have learned, that quantum field theory is, indeed, one of those theories in which the different scales are connected together in nontrivial ways via the intrinsic quantum-mechanical fluctuations. However, in current quantum field theory, only certain facets of renormalization group theory play a pivotal role.[140] High energy physics did not have to be the way it is! But, even if it were quite different, we would still need renormalization group theory in its fullest generality in condensed matter physics and, one suspects, in further scientific endeavors in the future.

ACKNOWLEDGMENTS

I thank Professor Alfred I. Tauber, Director, and, especially, Professor Tian Yu Cao of the Center for Philosophy and History of Science at Boston University for their roles in arranging and organizing the symposium on the foundations of quantum field theory and for inviting me to participate. Daniel M. Zuckerman kindly read and commented on the draft manuscript (based on my lecture at the symposium) and assisted with checking and completing the Bibliography; Caricia Claremont patiently typed and retyped the manuscript. Comments from Stephen G. Brush, N.D. Mermin, R. Shankar, David J. Wallace, B. Widom, and Kenneth G. Wilson have been appreciated. Much of the scientific work by the author and his associates described or cited in this review has been supported by the National Science Foundation: for that support,[141] now entering its 32nd year, I am truly grateful.

APPENDIX. ASYMPTOTIC BEHAVIOR

In specifying critical behavior (and asymptotic variation more generally) a little more precision than normally used is really called for. Following well-established custom, I use $\simeq$ for "approximately equals" in a rough and ready sense, as in $\pi^2 \simeq 10$. But to express "$f(x)$ varies like x^λ when x is small and positive," i.e., just to specify a critical exponent, I write:

$$f(x) \sim x^\lambda \quad (x \to 0+). \tag{A1}$$

Then the precise implication is

$$\lim_{x\to 0+}[\ln|f(x)|/\ln x] = \lambda; \tag{A2}$$

see Fisher (1967b, Sec. 1.4; 1983, Sec. 2.4). To give more information, specifically a *critical amplitude* like D in Eq. (2), I define $\approx$ as "asymptotically equals" so that

$$f(x) \approx g(x) \tag{A3}$$

as $x \to 0+$ implies

$$\lim_{x\to 0+} f(x)/g(x) = 1. \tag{A4}$$

Thus, for example, one has

$$(1-\cos x) \approx \tfrac{1}{2}x^2 \sim x^2, \tag{A5}$$

when $x \to 0$. See Fisher (1967b, Secs. 6.2, 6.3, and 7.2) but note that in Eqs. (6.2.6)–(6.3.5) the symbol $\simeq$ should read $\approx$; note also De Bruijn's (1958) discussion of $\approx$ in his book *Asymptotic Methods in Analysis*. The AIP and APS "strong recommendation" to use $\approx$ as "approximately equals" is to be, and has been strongly decried![142] It may also be remarked that few physicists, indeed, use $\sim$ in the precise mathematical sense originally introduced by Poincaré in his pioneering analysis of asymptotic series: see, e.g., Jeffreys and Jeffreys (1956) Secs. 17·02, 23·082, and 23·083. De Bruijn and the Jeffreys also, of course, define the $O(\cdot)$ symbol which is frequently misused in the physics literature: thus $f = O(g)$ $(x \to 0)$, should mean $|f| < c|g|$ for some constant c and $|x|$ small enough so that, e.g. $(1-\cos x) = O(x)$ is correct even though less informative than $(1-\cos x) = O(x^2)$.

SELECTED BIBLIOGRAPHY

The reader is cautioned that this article is not intended as a systematic review of renormalization group theory and its origins. Likewise, this bibliography makes no claims of completeness; however, it includes those contributions of most significance in the personal view of the author. The reviews of critical phenomena and RG theory cited in Footnote 4 above contain many additional references. Further review articles appear in the series *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green (later replaced by J. L. Lebowitz) and published by Academic Press, London (1972): some are listed below. Introductory accounts in an informal lecture style are presented in Fisher (1965, 1983).

[140]It is interesting to look back and read in Gunton and Green (1973) pp. 157–160, Wilson's thoughts in May 1973 regarding the "Field Theoretic Implications of the Renormalization Group" at a point just before the ideas of *asymptotic freedom* became clarified for non-Abelian gauge theory by Gross and Wilczek (1973) and Politzer (1973).

[141]Most recently under Grant CHE 96-14495.

[142]More recently, an internal AIP/APS memorandum (H-3, revised 6/94) states: "*Approximate equality:* $\simeq$ is preferred." Unfortunately, however, the latest published guidelines (e.g., for Rev. Mod. Phys. as revised October 1995) do not recognize this well advised change of policy. [Ed. Note: the third edition of the RMP Style Guide, currently in press, reflects this and other style changes.]

Abe, R., 1972, "Expansion of a Critical Exponent in Inverse Powers of Spin Dimensionality," Prog. Theor. Phys. **48**, 1414–1415.

Abe, R., 1973, "Expansion of a Critical Exponent in Inverse Powers of Spin Dimensionality," Prog. Theor. Phys. **49**, 113–128.

Aharony, A., 1973, "Critical Behavior of Magnets with Dipolar Interactions. V. Uniaxial Magnets in d Dimensions," Phys. Rev. B **8**, 3363–3370; 1974, erratum *ibid.* **9**, 3946.

Aharony, A., 1976, "Dependence of Universal Critical Behavior on Symmetry and Range of Interaction," in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green (Academic, London), Vol. 6, pp. 357–424.

Ahlers, G., A. Kornblit, and H. J. Guggenheim, 1975, "Logarithmic Corrections to the Landau Specific Heat near the Curie Temperature of the Dipolar Ising Ferromagnet $LiTbF_4$," Phys. Rev. Lett. **34**, 1227–1230.

Amit, D. J., 1978, *Field Theory, the Renormalization Group and Critical Phenomena* (McGraw-Hill Inc., London); see also 1993, the expanded Revised 2nd Edition (World Scientific, Singapore).

Bagnuls, C., and C. Bervillier, 1997, "Field-Theoretic Techniques in the Study of Critical Phenomena," J. Phys. Studies **1**, 366–382.

Baker, G. A., Jr., 1961, "Application of the Padé Approximant Method to the Investigation of some Magnetic Properties of the Ising Model," Phys. Rev. **124**, 768–774.

Baker, G. A., Jr., 1990, *Quantitative Theory of Critical Phenomena* (Academic, San Diego).

Baker, G. A., Jr., and N. Kawashima, 1995, "Renormalized Coupling Constant for the Three-Dimensional Ising Model," Phys. Rev. Lett. **75**, 994–997.

Baker, G. A., Jr., and N. Kawashima, 1996, "Reply to Comment by J.-K. Kim," Phys. Rev. Lett. **76**, 2403.

Baker, G. A., Jr., B. G. Nickel, M. S. Green, and D. I. Meiron, 1976, "Ising Model Critical Indices in Three Dimensions from the Callan-Symanzik Equation," Phys. Rev. Lett. **36**, 1351–1354.

Baker, G. A., Jr., B. G. Nickel, and D. I. Meiron, 1978, "Critical Indices from Perturbation Analysis of the Callan-Symanzik Equation," Phys. Rev. B **17**, 1365–1374.

Baxter, R. J., 1982, *Exactly Solved Models in Statistical Mechanics* (Academic, London).

Benfatto, G., and G. Gallavotti, 1995, *Renormalization Group*, Physics Notes, edited by P. W. Anderson, A. S. Wightman, and S. B. Treiman (Princeton University, Princeton, NJ).

Bogoliubov, N. N., and D. V. Shirkov, 1959, *Introduction to the Theory of Quantized Fields* (Interscience, New York), Chap. VIII.

Brézin, E., J. C. Le Guillou, and J. Zinn-Justin, 1976, "Field Theoretical Approach to Critical Phenomena," in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green (Academic, London), Vol. 6, pp. 125–247.

Brézin, E., D. J. Wallace, and K. G. Wilson, 1972, "Feynman-Graph Expansion for the Equation of State near the Critical Point," Phys. Rev. B **7**, 232–239.

Buckingham, M. J., and J. D. Gunton, 1969, "Correlations at the Critical Point of the Ising Model," Phys. Rev. **178**, 848–853.

Burkhardt, T. W. and J. M. J. van Leeuwen, 1982, Eds., *Real Space Renormalization* (Springer, Berlin).

Cao, T. Y., 1998, Ed., *The Conceptual Foundations of Quantum Field Theory* (Cambridge University Press, Cambridge).

Chen, J.-H., M. E. Fisher, and B. G. Nickel, 1982, "Unbiased Estimation of Corrections to Scaling by Partial Differential Approximants," Phys. Rev. Lett. **48**, 630–634.

Creswick, R. J., H. A. Farach, and C. P. Poole, Jr., 1992, *Introduction to Renormalization Group Methods in Physics* (John Wiley & Sons, Inc., New York).

De Bruijn, N. G., 1958, *Asymptotic Methods in Analysis* (North Holland, Amsterdam).

Di Castro, C., and G. Jona-Lasinio, 1969, "On the Microscopic Foundation of Scaling Laws," Phys. Lett. **29A**, 322–323.

Di Castro, C., and G. Jona-Lasinio, 1976, "The Renormalization Group Approach to Critical Phenomena," in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green (Academic, London), Vol. 6, pp. 507–558.

Domb, C., 1960, "On the Theory of Cooperative Phenomena in Crystals," Adv. Phys. (Philos. Mag. Suppl.) **9**, 149–361.

Domb, C., 1996, *The Critical Point: A Historical Introduction to the Modern Theory of Critical Phenomena* (Taylor and Francis, London).

Domb, C., and D. L. Hunter, 1965, "On the Critical Behavior of Ferromagnets," Proc. Phys. Soc. (London) **86**, 1147–1151.

Domb, C., and M. F. Sykes, 1962, "Effect of Change of Spin on the Critical Properties of the Ising and Heisenberg Models," Phys. Rev. **128**, 168–173.

Essam, J. W., and M. E. Fisher, 1963, "Padé Approximant Studies of the Lattice Gas and Ising Ferromagnet below the Critical Point," J. Chem. Phys. **38**, 802–812.

Fisher, D. S., and D. A. Huse, 1985, "Wetting Transitions: A Functional Renormalization Group Approach," Phys. Rev. B **32**, 247–256.

Fisher, M. E., 1959, "The Susceptibility of the Plane Ising Model," Physica (Amsterdam) **25**, 521–524.

Fisher, M. E., 1962, "On the Theory of Critical Point Density Fluctuations," Physica (Amsterdam) **28**, 172–180.

Fisher, M. E., 1964, "Correlation Functions and the Critical Region of Simple Fluids," J. Math. Phys. **5**, 944–962.

Fisher, M. E., 1965, "The Nature of Critical Points," in *Lectures in Theoretical Physics*, Vol. VIIC (University of Colorado Press, Boulder), pp. 1–159.

Fisher, M. E., 1966a, "Notes, Definitions and Formulas for Critical Point Singularities," in M. S. Green and J. V. Sengers, 1966, Eds., cited below.

Fisher, M. E., 1966b, "Shape of a Self-Avoiding Walk or Polymer Chain," J. Chem. Phys. **44**, 616–622.

Fisher, M. E., 1967a, "The Theory of Condensation and the Critical Point," Physics (Long Island City, NY) **3**, 255–283.

Fisher, M. E., 1967b, "The Theory of Equilibrium Critical Phenomena," Rep. Prog. Phys. **30**, 615–731.

Fisher, M. E., 1967c, "Critical Temperatures of Anisotropic Ising Lattices II. General Upper Bounds," Phys. Rev. **162**, 480–485.

Fisher, M. E., 1969, "Phase Transitions and Critical Phenomena," in *Contemporary Physics*, Vol 1, Proc. International Symposium, Trieste, 7–28 June 1968 (International Atomic Energy Agency, Vienna), pp. 19–46.

Fisher, M. E., 1971, "The Theory of Critical Point Singularities," in *Critical Phenomena, Proc. 1970 Enrico Fermi Internat. Sch. Phys., Course No. 51, Varenna, Italy*, edited by M. S. Green (Academic, New York), pp. 1–99.

Fisher, M. E., 1972, "Phase Transitions, Symmetry and Dimensionality," *Essays in Physics* (Academic, London), Vol. 4, pp. 43–89.

Fisher, M. E., 1974a, "General Scaling Theory for Critical Points," *Proceedings of the Nobel Symp. XXIV, Aspenäsgården, Lerum, Sweden, June 1973*, in *Collective Properties of Physical Systems*, edited by B. Lundqvist and S. Lundqvist (Academic, New York), pp. 16–37.

Fisher, M. E., 1974b, "The Renormalization Group in the Theory of Critical Phenomena," Rev. Mod. Phys. **46**, 597–616.

Fisher, M. E., 1983, "Scaling, Universality and Renormalization Group Theory," in Lecture Notes in Physics, Vol. 186, *Critical Phenomena*, edited by F. J. W. Hahne (Springer, Berlin), pp. 1–139.

Fisher, M. E., and A. Aharony, 1973, "Dipolar Interactions at Ferromagnetic Critical Points," Phys. Rev. Lett. **30**, 559–562.

Fisher, M. E., and R. J. Burford, 1967, "Theory of Critical Point Scattering and Correlations. I. The Ising Model," Phys. Rev. **156**, 583–622.

Fisher, M. E., and J.-H. Chen, 1985, "The Validity of Hyperscaling in Three Dimensions for Scalar Spin Systems," J. Phys. (Paris) **46**, 1645–54.

Fisher, M. E., and D. S. Gaunt, 1964, "Ising Model and Self-Avoiding Walks on Hypercubical Lattices and 'High Density' Expansions," Phys. Rev. **133**, A224–A239.

Fisher, M. E., S.-K. Ma, and B. G. Nickel, 1972, "Critical Exponents for Long-Range Interactions," Phys. Rev. Lett. **29**, 917–920.

Fisher, M. E., and P. Pfeuty, 1972, "Critical Behavior of the Anisotropic n-Vector Model," Phys. Rev. B **6**, 1889–91.

Fisher, M. E., and M. F. Sykes, 1959, "Excluded Volume Problem and the Ising Model of Ferromagnetism," Phys. Rev. **114**, 45–58.

Gawędski, K., 1986, "Rigorous Renormalization Group at Work," Physica **140A**, 78–84.

Gell-Mann, M., and F. E. Low, 1954, "Quantum Electrodynamics at Small Distances," Phys. Rev. **95**, 1300–1312.

Gorishny, S. G., S. A. Larin, and F. V. Tkachov, 1984, "ϵ-Expansion for Critical Exponents: The $O(\epsilon^5)$ Approximation," Phys. Lett. **101A**, 120–123.

Green, M. S., and J. V. Sengers, 1966, Eds., *Critical Phenomena: Proceedings of a Conference held in Washington, D.C. April 1965*, N.B.S. Misc. Publ. 273 (U.S. Govt. Printing Off., Washington, 1 December 1966).

Griffiths, R. B., 1965, "Thermodynamic Inequality near the Critical Point for Ferromagnets and Fluids," Phys. Rev. Lett. **14**, 623–624.

Griffiths, R. B., 1972, "Rigorous Results and Theorems," in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green (Academic, London), Vol. 1, pp. 7–109.

Gross, D., and F. Wilczek, 1973, "Ultraviolet Behavior of Non-Abelian Gauge Theories," Phys. Rev. Lett. **30**, 1343–46.

Guggenheim, E. A., 1949, *Thermodynamics* (North Holland, Amsterdam), Fig. 4.9.

Gunton, J. D., and M. S. Green, 1974, Eds., *Renormalization Group in Critical Phenomena and Quantum Field Theory: Proceedings of a Conference*, held at Chestnut Hill, Pennsylvania, 29–31 May 1973 (Temple University, Philadelphia).

Heller, P., and G. B. Benedek, 1962, "Nuclear Magnetic Resonance in MnF_2 near the Critical Point," Phys. Rev. Lett. **8**, 428–432.

Hille, E., 1948, *Functional Analysis and Semi-Groups*, 1948 (American Mathematical Society, New York).

Hohenberg, P. C., and B. I. Halperin, 1977, "Theory of Dynamic Critical Phenomena," Rev. Mod. Phys. **49**, 435–479.

Itzykson, D., and J.-M. Drouffe, 1989, *Statistical Field Theory* (Cambridge University Press, Cambridge).

Jeffreys, H., and B. S. Jeffreys, 1956, *Methods of Mathematical Physics*, 3rd ed. (Cambridge University, London).

Kadanoff, L. P., 1966, "Scaling Laws for Ising Models near T_c," Physics (Long Island City, NY) **2**, 263–272.

Kadanoff, L. P., 1976, "Scaling, Universality and Operator Algebras," in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green (Academic, London), Vol. 5a, pp. 1–34.

Kadanoff, L. P., *et al.*, 1967, "Static Phenomena near Critical Points: Theory and Experiment," Rev. Mod. Phys. **39**, 395–431.

Kadanoff, L. P., and F. J. Wegner, 1971, "Some Critical Properties of the Eight-Vertex Model," Phys. Rev. B **4**, 3989–3993.

Kaufman, B., and L. Onsager, 1949, "Crystal Statistics II. Short-Range Order in a Binary Lattice," Phys. Rev. **76**, 1244–52.

Kiang, C. S., and D. Stauffer, 1970, "Application of Fisher's Droplet Model for the Liquid-Gas Transition near T_c," Z. Phys. **235**, 130–139.

Kim, H. K., and M. H. W. Chan, 1984, "Experimental Determination of a Two-Dimensional Liquid-Vapor Critical-Point Exponent," Phys. Rev. Lett. **53**, 170–173.

Koch, R. H., V. Foglietti, W. J. Gallagher, G. Koren, A. Gupta, and M. P. A. Fisher, 1989, "Experimental Evidence for Vortex-Glass Superconductivity in Y-Ba-Cu-O," Phys. Rev. Lett. **63**, 1511–1514.

Kouvel, J. S., and M. E. Fisher, 1964, "Detailed Magnetic Behavior of Nickel near its Curie Point," Phys. Rev. **136**, A1626–A1632.

Landau, L. D., and E. M. Lifshitz, 1958, *Statistical Physics*, Vol. 5 of *Course of Theoretical Physics* (Pergamon, London), Chap. XIV.

Larkin, A. I., and D. E. Khmel'nitskii, 1969, "Phase Transition in Uniaxial Ferroelectrics," Zh. Eksp. Teor. Fiz. **56**, 2087–2098 [Sov. Phys. JETP **29**, 1123–1128].

Le Guillou, J. C., and J. Zinn-Justin, 1977, "Critical Exponents for the n-Vector Model in Three Dimensions from Field Theory," Phys. Rev. Lett. **39**, 95–98.

Ma, S.-K., 1976a, *Modern Theory of Critical Phenomena* (W. A. Benjamin, Inc., Reading, MA).

Ma, S.-K., 1976b, "Renormalization Group by Monte Carlo Methods," Phys. Rev. Lett. **37**, 461–474.

MacLane, S. and G. Birkhoff, 1967, *Algebra* (MacMillan, New York), Chap. II, Sec. 9, Chap. III, Secs. 1, 2.

Nelson, D. R., 1983, "Defect-mediated Phase Transitions" in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green (Academic, London), Vol. 7, pp. 1–99.

Nelson, D. R., and M. E. Fisher, 1975, "Soluble Renormalization Groups and Scaling Fields for Low-Dimensional Spin Systems," Ann. Phys. (N.Y.) **91**, 226–274.

Niemeijer, Th., and J. M. J. van Leeuwen, 1976, "Renormalization Theory for Ising-like Systems," in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green (Academic, London), Vol. 6, pp. 425–505.

Nickel, B. G., 1978, "Evaluation of Simple Feynman Graphs," J. Math. Phys. **19**, 542–548.

Onsager, L., 1944, "Crystal Statistics I. A Two-dimensional Model with an Order-Disorder Transition," Phys. Rev. **62**, 117–149.

Onsager, L., 1949, "Discussion remark following a paper by G. S. Rushbrooke at the International Conference on Statistical Mechanics in Florence," Nuovo Cimento Suppl. (9) **6**, 261.

Parisi, G., 1973, "Perturbation Expansion of the Callan-Symanzik Equation," in Gunton and Green (1974), cited above.

Parisi, G., 1974, "Large-Momentum Behaviour and Analyticity in the Coupling Constant," Nuovo Cimento A **21**, 179–186.

Patashinskii, A. Z., and V. L. Pokrovskii, 1966, "Behavior of ordering systems near the transition point," Zh. Eksp. Teor. Fiz. **50**, 439–447 [Sov. Phys. JETP **23**, 292–297].

Patashinskii, A. Z., and V. L. Pokrovskii, 1979, *Fluctuation Theory of Phase Transitions* (Pergamon, Oxford).

Pawley, G. S., R. H. Swendsen, D. J. Wallace, and K. G. Wilson, 1984, "Monte Carlo Renormalization Group Calculations of Critical Behavior in the Simple-Cubic Ising Model," Phys. Rev. B **29**, 4030–40.

Pfeuty, P., and G. Toulouse, 1975, *Introduction au Group des Renormalisation et à ses Applications* (Univ. de Grenoble, Grenoble); translation, *Introduction to the Renormalization Group and to Critical Phenomena* (Wiley and Sons, London, 1977).

Politzer, H. D., 1973, "Reliable Perturbative Results for Strong Interactions?" Phys. Rev. Lett. **30**, 1346–49.

Riesz, F. and B. Sz.-Nagy, 1955, *Functional Analysis* (F. Ungar, New York) Chap. X, Secs. 141, 142.

Rowlinson, J. S., 1959, *Liquids and Liquid Mixtures* (Butterworth Scientific, London).

Rushbrooke, G. S., 1963, "On the Thermodynamics of the Critical Region for the Ising Problem," J. Chem. Phys. **39**, 842–843.

Rushbrooke, G. S., and P. J. Wood, 1958, "On the Curie Points and High-Temperature Susceptibilities of Heisenberg Model Ferromagnetics," Mol. Phys. **1**, 257–283.

Schultz, T. D., D. C. Mattis, and E. H. Lieb, 1964, "Two-Dimensional Ising Model as a Soluble Problem of Many Fermions," Rev. Mod. Phys. **36**, 856–871.

Shankar, R., 1994, "Renormalization-Group Approach to Interacting Fermions," Rev. Mod. Phys. **66**, 129–192.

Stanley, H. E., 1971, *Introduction to Phase Transitions and Critical Phenomena* (Oxford University, New York).

Stell, G., 1968, "Extension of the Ornstein-Zernike Theory of the Critical Region," Phys. Rev. Lett. **20**, 533–536.

Suzuki, M., 1972, "Critical Exponents and Scaling Relations for the Classical Vector Model with Long-Range Interactions," Phys. Lett. **42A**, 5–6.

Symanzik, K., 1966, "Euclidean Quantum Field Theory. I. Equations for a Scalar Model," J. Math. Phys. **7**, 510–525.

Tarko, H. B., and M. E. Fisher, 1975, "Theory of Critical Point Scattering and Correlations. III. The Ising Model below T_c and in a Field," Phys. Rev. B **11**, 1217–1253.

Wallace, D. J., 1976, "The ϵ-Expansion for Exponents and the Equation of State in Isotropic Systems," in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green (Academic, London), Vol. 6, pp. 293–356.

Wegner, F. J., 1972a, "Corrections to Scaling Laws," Phys. Rev. B **5**, 4529–4536.

Wegner, F. J., 1972b, "Critical Exponents in Isotropic Spin Systems," Phys. Rev. B **6**, 1891–93.

Wegner, F. J., 1976, "The Critical State, General Aspects," in *Phase Transitions and Critical Phenomena*, edited by C. Domb and M. S. Green (Academic, London), Vol. 6, pp. 7–124.

Widom, B., 1965a, "Surface Tension and Molecular Correlations near the Critical Point," J. Chem. Phys. **43**, 3892–97.

Widom, B., 1965b, "Equation of State in the Neighborhood of the Critical Point," J. Chem. Phys. **43**, 3898–3905.

Widom, B., and O. K. Rice, 1955, "Critical Isotherm and the Equation of State of Liquid-Vapor Systems," J. Chem. Phys. **23**, 1250–1255.

Wilson, K. G., 1971a, "Renormalization Group and Critical Phenomena I. Renormalization Group and Kadanoff Scaling Picture," Phys. Rev. B **4**, 3174–83.

Wilson, K. G., 1971b, "Renormalization Group and Critical Phenomena II. Phase-Space Cell Analysis of Critical Behavior," Phys. Rev. B **4**, 3184–3205.

Wilson, K. G., 1972, "Feynman-Graph Expansion for Critical Exponents," Phys. Rev. Lett. **28**, 548–551.

Wilson, K. G., 1975, "The Renormalization Group: Critical Phenomena and the Kondo Problem," Rev. Mod. Phys. **47**, 773–840.

Wilson, K. G., 1983, "The Renormalization Group and Critical Phenomena" (1982, Nobel Prize Lecture) Rev. Mod. Phys. **55**, 583–600.

Wilson, K. G., and M. E. Fisher, 1972, "Critical Exponents in 3.99 Dimensions," Phys. Rev. Lett. **28**, 240–243.

Wilson, K. G., and J. Kogut, 1974, "The Renormalization Group and the ϵ Expansion," Phys. Rep. **12**, 75–200.

Zinn, S.-Y., and M. E. Fisher, 1996, "Universal Surface-Tension and Critical-Isotherm Amplitude Ratios in Three Dimensions," Physica A **226**, 168–180.

Zinn, S.-Y., S.-N. Lai, and M. E. Fisher, 1996, "Renormalized Coupling Constants and Related Amplitude Ratios for Ising Systems," Phys. Rev. E **54**, 1176–1182.

From Order to Chaos, II
Essays: Critical, Chaotic, and Otherwise
by
Leo P. Kadanoff

General Introduction:
The Worlds of Science

This book is a selection of my research and popular essays, with particular emphasis on works which review or discuss in a general way some scientific or technical question. The papers are all about the world of science, or rather about the different worlds in which a scientist works. In my own work I can see at least four different kinds of things which might be meant when one talks about the worlds of a scientist.

First, I might point to a little society or social grouping composed of scientists and a few associates. This **social world** defines the group in which we work and exchange ideas. This is the audience for our papers, the source of our applause, and our critics and competition. A scientist can go to different places all over the world and see mostly just the usual group of associates. A scientist can be thrown into a new little group, formed in an allied field of scientific endeavor, and immediately recognize the society and the social norms. This little world is close and closed. It defines our successes and failures.

But there is in addition a more intimate social group which defines our work. Much scientific work is done in direct collaboration with other scientists. Many of the papers in this volume have several authors. Typically each author brings a slightly different experience and point of view to the joint effort, so that the eventual product is much better than would have been produced by any single person.

This fact came home to me at the very beginning of my career as an 'independent' scientist. Gordon Baym and I had both been trained at Harvard, he under Schwinger and myself under Paul Martin and Roy Glauber. He had learned how to apply variational methods to the derivation of Green's function approximations. I was working on the development of approximations which built in some thermodynamic and conservation laws. With huge effort over a period of months, I

had derived one or two approximations which fit my criteria. A day after I had described to him what I had done, he showed me how to construct an infinite number of new approximations which fit into the general scheme. The results appeared in part in our book *Quantum Statistical Mechanics* and in part in our paper, which appears as #5 in my publication list[1]. Two heads had done a lot more than one.

My scientific life has contained many other very fruitful collaborations. I describe some of these in the introductory essays which head the various sections of this book.

But, we scientists also work in a very different kind of tight little world, the artificial little world constructed by our ideas. Some of my recent work has been related to the development of models for the behavior of avalanches or sand slides. To construct this **model world** we considered a simplified example in which square or cubic grains of sand were stacked in neat piles. If a given pile overtopped its neighbors by more than a specified amount, then several grains would fall onto the neighboring stacks. Clearly, this model represented a totally artificial oversimplification of any picture of the behavior of real sand. Nonetheless, a whole group of us threw ourselves quite wholeheartedly into the study of this little model. For several years at a time, we took this artificial example and pretended it was the whole universe.[2] We examined this tiny world with the same seriousness that one might examine the history of the British Empire, or a science of the human mind. Our goal was to develop and understand the laws which governed behavior in this tiny closed-off cosmos.

Why should serious people study such inadequate toys? Clearly these toys cannot accurately represent the third type of scientific world, the **real world** in which we work and live. Nonetheless it might be profitable to study such hermetic little model worlds because perhaps the experience developed in the little world can be extended and applied to our real world. Maybe there is something in our model avalanche which can be carried over and give some deep

---

[1] The publication list can be found at the end of this book. In general, items in this list will be given by a number prefixed by the sign '#'

[2] In this case, the collaborative team included Ashvin B. Chhabra, Mitchell Feigenbaum, Amy Kolan, Sidney Nagel, Itamar Procaccia, Lei Wu, and Su-Min Zhou.

***

insight into how avalanches work. Perhaps these ideas might even have some practical use in the protection of Swiss mountain villages or to the design of particle detectors. Maybe not. Probably not. But one cannot tell what might be carried to the real world until the model world is examined and understood.

Science also sees another version of the real world, the world of people. People and the society support science. Naturally some return is demanded. One demand is that science generates ideas and concepts which can be meaningful to the public at large and can catch its imagination or satisfy its curiosity. To realize this goal, we scientists must be teachers in the broadest sense Another demand is that we, from time to time, satisfy the aspirations of society for better technology or for a better understanding of the applications or limits of technology. We can only occasionally help the society in this direction, but our help can be quite crucial. We have served in the development of weapons, of communications, and of health care. We cannot be sure where we will be needed in the future, but we are required to be alert to ways in which we can serve.

The fourth version of the world of the scientist is the most important and the one which we scientists most vividly experience. What is it that one really transfers from the oversimplified model world to the complex world of reality? Clearly the medium of exchange is ideas. One carries some concept of how an avalanche must work. Some idea. And then one takes the idea and asks how well that idea agrees with mundane reality. (But reality is in most cases richer, more beautiful, and more fertile than our imaginations. So in most cases, this comparison enriches rather than just checks our ideas.) The results of all the model building, all the comparison with reality, and all the up and back of scientific exchange is a set of concepts which can then be applied to other situation. Our final outcome then is something which can be added to the **world of ideas.**

The series of essays in this volume relates in some degree to all four worlds.

I cannot imagine anyone who would wish to go through all this material from beginning to end. So let me take the reader for a walking tour through this material so that he or she might plan a particular path which might proving pleasing or useful.

This book is divided into four parts, which I shall describe in inverse order. The last part, **Turbulence and Chaos**, is connected with my most recent work aimed at describing and explaining how chaos and

complexity arises in physical systems. This subject is in the process of development. However, some of the important ideas in it have already become apparent. In one view, this subject starts from a question: Given that the laws of physics are simple and predictive, how can we have a world which is so complex and apparently unpredictable. The question is clearly in the world of ideas. To answer it, one turns to the development of mathematical models and of real physical systems, in both cases looking for simultaneous simplicity and complexity.

This volume's third section describes my introduction to complexity. During the late 1960s, it became fashionable for scientists to look away from the traditional applications of their research to military systems, and to focus instead upon problems which might be relevant to the broader needs of society. In the U.S. National Science Foundation, this relevance boom even gave rise to a new program RANN, Research Applied to National Needs[3]. In any case, the same social forces which pushed the NSF toward RANN pushed me toward studying the complex of forces which shape the physical and social environment of our urban areas. This part of the book, **Simulations, Urban Studies, and Social Systems**, includes the outcome of this effort. It also includes some editorial pieces written for Physics Today which report upon the health and decay in another kind of social system, the one of physics itself.

The second section in this book come from my best and most important contribution to science, my work on the understanding of **Scaling and Phase Transitions.** Here I played a part in the invention of a tiny world, the critical system, and then devoted considerable effort to studying the detailed properties of that world. In this period, many different and very intelligent people focused an amazing amount of effort upon a very closed and partial model of reality. Despite the limited focus of the work, it has had its consequences. For more details about what happened, I ask you to look at the introduction to that section.

The book's first section, **Fundamentals Issues in Hydrodynamics, Condensed Matter and Field Theory**, is

---

[3] This program was in large measure designed and put into place by Joel Snow. Our recent Presidents have had many unkind words for civil servants. In my experience, I have found governmental science administrators to be thoughtful hard-working people. As a group, they have contributed a lot to science without getting much thanks.

***

devoted to describing the relationship among different models, either in general or in particular examples. Physics contains many different models, which might describe different aspects of the very same physical system. Clearly, one should ask how the different realities caught by the different models fit together. Some of the essays in this section directly confront this general issue, others ask how it s resolved in a particular physical example. A major theme of this section is that physics is really about how models, which give different levels of physical description, may fit together.

Each section in the book is headed by a specific introduction which goes into more details on the questions mentioned here, and outlines some of the contents of the papers.

There is an overall theme to the whole book. Each of the first three sections is devoted to 'old' scientific questions, questions which have been asked and mostly answered. One can expect that each field of science starts from questions built upon small piece if the world, and that in time these questions will become answered as well as the times and means available permit. Then, the subfield gets mined out. Naturally, the scientist must then direct his or her work away from these particular aspects of reality. Naturally, also, when that happens there is some temptation to see and bemoan 'the end of science'. Perhaps this is particularly tempting right now for a physicist since some of our most exciting problems have been either solved or run into major technical barriers.

However, in the introduction to the last section, I shall argue that we have in front of us a mostly uncharted territory, concerning the development of complexity in the world. The ideas in this part of science is connected with understanding the relationship and linkage among the worlds we have explored. Right now, I see not an end for physics but a beginning. Now is an exciting time to be working in physics.

Section A. Fundamental Issues in Hydrodynamics, Condensed Matter and Field Theory. From Level to Level

This section contains a collection of essays about apparently disconnected subjects in field theory, condensed matter physics, and hydrodynamics. There is, however, a thread of connection among all these different essays. In each case, we ask some kind of question about the relation between different levels of description of the physical world. This is a very natural question for a person trained

as I am, in the area of condensed matter physics. My thesis advisers, Paul Martin and Roy Glauber, continually directed my attention at the relation between a microscopic description of reality and a macroscopic description. Thus a gas is composed of molecules, but it also obeys the laws of fluid mechanics. A microwave cavity contains photons but also an electric field. Or again, a fluid near its critical point is a bunch of molecules, but they also be described by a scale-invariant field theory.

In my imagination I see every physics problem as a kind of little world. Each world has its own rules, which apply to the description at that level. This idea is brought out in the paper[4] A1, *On two Levels*, which looks for a more macroscopic, hydrodynamic, level of description within a computer game in which simplified particles go through a kind of dance. Naturally, one sees the dancers by looking very closely, while a more lumped description shows the hydrodynamic flow. In this paper the logic is one in which the more lumped description is built up from the more microscopic one. This same point is made again, backwards, in the next essay. Here Paul Martin and I start from the equations of motion of hydrodynamics, and look at local fluctuations to gain a more microscopic description of fluid flow. In this case, one and the same physical description covers two quite different ranges of physical size and physical phenomena.

There is something more to say about this paper. A parallel computation was done by Landau and Placzek, long before Paul Martin and I wrote our piece. I do not recall ever having seen this parallel effort. Nonetheless its existence serves as a reminder that rarely do we produce something completely new in science. Every piece is based upon predecessors, and if we did not do the work, it still probably would be produced by others in substantially similar form. This point was brought home to me in my work on electrons and phonons. Essay A3 in this volume is a review piece, based upon the research described in paper #14 on my publication list. The latter is a joint work by Richard Prange and myself. It arose in that form because Prange and I did substantially identical, independent research on this topic. His work was written up first, and appeared one morning in my mailbox. After I called him and described my

---

[4] Each reference to a paper which appears in this volume is given as a letter followed by a number. The letter denotes the section is this volume while the following number gives the placement of the paper in the section.

***

own thinking about this subject, he generously suggested that we write it up together. During these many years, I have felt in his debt for this fine courtesy.

Paper A4 is partially a reprise of the ideas of A3 but now applied to superfluid motion rather than the normal state[5]. It is illustrative of an important idea about the relation between different levels of experience. The basic microscopic forces and interactions are quite the same in normal materials as in superfluids. However, the nature of the physical state of the two systems are vastly different. The superfluid forms a condensate in which all the particles in the system cooperate to produce a single quantum state. This condensation extends across the entire spatial extent of the system and produces a change in behavior which spans both the micro and the macro levels of description. The superfluid is thus essentially different from its normal state counterpart in both worlds of description.

A qualitative change in the collective behavior of a large group of particles is called a phase transition. Often a phase transition is accompanied by a change in the symmetries shown by the system. The paper on superfluidity shows in one example how phase transitions may manifest themselves equally in the microscopic and macroscopic domains. Paper A5 reviews Kenneth Wilson's work on symmetry changes and phase transitions in systems of quarks and strings. The world is strong interactions is described by a theory, *quantum chromodynamics*, is which quarks are the fundamental particles. The basic question here how can one have a description of quantum chromodynamics in terms of almost free quarks, and yet not directly see quarks in our world. The answer is that the confinement is a concept which describes the macro level, the world of ordinary nuclear physics, while the micro-level is correctly given by quarks only weakly coupled by gluons. This work that Wilson started has grown into the subbranch of particle and nuclear physics called lattice gauge theory.

---

[5] Another important thread of my research effort in the earlier part of my career was concerned with understanding the relationship between the microscopic and the macroscopic properties of superconductors and superfluids. These include, for example, the collaborative works with P.C. Martin (#1, #2, and #6) on my publication list, with Vinay Ambegaokar (#4), with Brian Pippard (#20) and with my students D. Markowitz (#10), Igor I. Falko (#16), G. Laramore (#31), and Jack Swift (#32).

***

This review, A5, is one of many papers I have written to describe and perhaps help explain the work of other physicists. I have always felt enlarged when I could work with the beautiful and deep ideas of others. In fact, many of my papers were intended to be in part or in whole explanations or descriptions of beautiful ideas of others. Thus, for example, my paper #57 on my publication list extends the work of A.A. Migdal[6], #59, written with J. Jose, S. Kirkpatrick, and D. Nelson is partially built to convince the reader of the correctness of a theory due to Kosterlitz and Thouless[7], #133 addresses the ideas of Bak, Tang, and Wiesenfeld[8], while in #69 Mahito Kohmoto and I review the work of Sato, Miwa, and Jimbo. Of course, this recognition of others is not entirely unselfish. I have always held the opinion that people are more likely to recognize your work if your recognize theirs.

In paper A6, Eduardo Fradkin and I explain ideas from the theory of critical phenomena and phase transitions in a way which has proved fruitful in a variety of other fields. This paper shows how 'particles' with very peculiar behavior arises from a treatment of two coupled Ising models. It is one of the first treatments of fractional statistics particles, the so-called anyons, which have had a vogue in condensed matter physics and field theory. This paper has perhaps received a bit less attention than it deserved. It is part of a long series of papers aimed at describing the mathematical structure of two-dimensional statistical mechanics. These will be discussed in more detail in the introductory essay for section B.

In my work, as in the work of most physicists, a major goal has been the building of bridges between different worlds of experience. Science has two traditional windows for looking at the world, through experimental study or theoretical construction. In recent decades a third window has opened, which approaches reality through the construction of computer models. Paper A7 is an attempt to describe some of the great accomplishments of computer-based physics. It argues that the importance of this relatively new tool is that it can be

---

[6] A.A. Migdal, Z. Eksper. Theoret. Fiz **69** 810, 1475 (1975).

[7] J. M. Kosterlitz and D. Thouless, J. Phys. C. **6** 1181 (1973), J. M. Kosterlitz J. Phys. C. **7** 1046 (1974). See also the work of V. L. Berezinskii JETP **32** 493 (1971) and JETP **34** 610 (1971).

[8] P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. Lett. **59,** 381 (1987), Phys. Rev **A38**, 364 (1988).

***

effectively used in conjunction with the older tools of analysis and experiment. If I were writing this column now, I would say the same thing in a more negative way by pointing out that computer-based physics can be ineffective when it constructs totally closed worlds, which then cannot be related to the richer perspectives of theory and of experiment. However, now as then, I believe that this new tool can be properly used to either explore some totally new area of behavior, or alternatively to check a precisely formulated idea.

From the earliest period of my work I have made use of computers. For example, an applied paper[9] on heat transfer in semi-transparent materials contained a solution to a radiation transfer equation obtained with the aid of a computer. Later on Jack Swift and I, see paper B4, described a computer algorithm which might be used to simulate the approach to equilibrium in a system obeying statistical mechanics[10]. Abdullah Sadiq later implemented this idea in his PhD thesis. In more recent years, my students, my coworkers and I have made a major effort to use the computer as a device to generate ideas which could be checked against theory and experimental reality. We have concentrated on efforts on small computers, because we found these to be truly flexible tools for studying a tiny portion of reality.

Paper A7 was written for the Reference Frames column in *Physics Today*. Gloria Lubkin, the editor, had the idea of such a column, describing personal views of the worlds of physics. She and I worked together in producing these columns. I have had much good feeling about them. I think that they are rather successful, in large measure because of her enthusiasm and good sense. This section ends with two columns, A8 and A9, intended to assess some of the worlds constructed by scientists and how they are related to one another. The first does this by looking at physics as a whole; the second looks at an example from hydrodynamics, condensed matter physics, and elementary particle experiment.

---

[9] This joint work with Henry Hidalgo is # 8 on my publication list.

[10] This approach use the method of *cellular automata*. Later on, this approach was very much improved by others. See paper A1 for a discussion of some of its implications. After a while I returned to this subject in work done jointly with McNamara and G. Zanetti, #125. We contributed to the further development and checking of the applicability of such models to describe equilibrium and nonequilibrium situations. The avalanche work already mentioned uses another kind of automaton.

***

B. Introduction to Section on Scaling and Phase Transitions

On the joys of creation.

This section describes the development of ideas of scaling and universality as they relate to phase transitions and critical phenomena. In the late 1960s and early 1970s, a group of physicists and chemists changed the way scientists look at problems in statistical mechanics and related fields. Looking back at this work from almost thirty years later, I still feel proud, pleased, and somewhat surprised that I could play a role in such an achievement.

My story starts when I was finishing up as a graduate student at Harvard in 1960. Kurt Gottfried and Paul Martin both pointed out to me that the problem of 'second order phase transitions' were was quite interesting and not understood at all. Kurt and I even did a calculation, never published, in which we looked for critical fluctuations in three dimensional superconductors. We got, and were discouraged by what is in retrospect the correct answer which is that it is next door to impossible to observe critical fluctuation in three dimensional superconductors. Then, I put the problem aside for a while.

I came back to it during a nine-month period I was spending in the Cavendish laboratory in Cambridge University, invited by Neville Mott to take part in what he called a theoretical jamboree. This kind of jamboree permitted one to have lots of free time for research, so I began to struggle once again with the critical behavior which occurs very near second order phase transitions. Previously many scientists had believed that the Van der Walls-Weiss-Landau mean field theory of this transition was essentially correct. But, by this time, the experimental work of Sasha Voronel and others had shown that the mean field theory was wrong for the classical liquid gas phase transitions. C. F. Keller's experiment[11] and Brian Pippard's[12] analysis had shown that it did not work for the transitions in Helium. On the theoretical side, Cyril Domb, Michael Fisher, and the King's college school had proven that the mean field theory did not apply to Ising models either. The time seemed ripe for a new approach.

---

[11] See C.F. Kellers (Thesis, Duke University, 1960).

[12] A.B. Pippard, Proc. Roy. Soc. (London) **A216** , 547 (21953).

I entered the problem by studying Lars Onsager's solution of the two dimensional Ising model. This exact solution of this model of two dimensional magnetism had been announced in the 1940s, but it had never been fully analyzed. Onsager and C.N. Yang had calculated some of the thermodynamic properties, but there was really no explanation of what physics might be demonstrated by Onsager's solution. Here was a tiny little world, just waiting to be explored and perhaps even captured. As far as I could tell, nobody had looked at the spatial correlations built into the Onsager solution. My own background pushed me toward looking these correlations. The work on the connection between hydrodynamics and correlations (A2) were all about the relation between thermodynamic behavior and space-time correlations, especially in the long-wavelength limit. My Green's function studies and much of the other work described in section A was aimed at understanding correlations over large regions in space and time. I was certainly ready to attack correlations in Ising models.

I began to calculate the spin-spin correlation function of the Ising model near its critical point. (See paper B1.) The calculation was long and difficult. It involved the Weiner-Hopf technique for solving integral equations, which made extensive use of complex variable techniques. Fortunately, I had received an excellent training in this method in graduate school with the applied math courses I had taken from George Carrier and Arthur Bryson and with the complex variable course I had taken from George Mackey. After about six months, I had a finished work which I sent off to the Journal of Mathematical Physics.

This paper was potentially very important for this subarea of physics. It contained the germs of much of what was to prove to be the correct theory of second order phase transitions, all worked out in a particular example. The paper also had defects. It was hard to read. The boundary conditions at the edge of the material were handled in a sloppy and incorrect fashion. (I did not believe that the boundary conditions were really relevant to the solution, so I did not inquire very deeply into their correct handling.) And the paper certainly did not proclaim why it was potentially important. It was rejected, twice. And in each rejection it crossed the Atlantic once or twice by slow boat. This paper was written before the days of extensive use of preprints. So, for a period of almost a year, there were only a few people who knew the contents of this paper. It was eventually published, after the second rejection, in Nuovo Cimento.

***

In the meantime, the field had progressed considerably. There was a conference on second order phase transitions, held at the U.S. National Bureau of Standards, in which the field was reviewed. Several people-notably Fisher, Ben Widom, and Michael Buckingham-described some conjectures about spatial correlations. Widom, in particular, had introduced some 'magic' (the word was used by Martin when he mentioned the work to me) relations among the critical indices. These critical indices are numbers used to describe the order of magnitudes or the various quantities which can be used to describe the near-critical behavior of a system near a second order phase transition. My spin correlation work also involved these critical indices. There were many of them, and I had a hard time remembering them all. So I had developed a mnemonic device, which involved expressing all orders of magnitudes in terms of two independent magnitudes, the natural fluctuations in the spin and in the energy density.

Then, in Christmas week of 1965 I had a sudden vision. A gift from the gods. I had a simple view of how these magnitude relations might be true and general. In modern terms, I had developed a scaling analysis of the critical behavior of Ising models based upon the idea of running coupling constants, i.e. couplings which depended upon the distance scale. (Some idea of this kind was also present in field theory in the work of Stückelberg and Peterman and of Murray Gell-Mann and Francis Low[13]. Unfortunately, I was only barely aware of this earlier work.) The awfully complex and convoluted extension of the Onsager solution which I had previously done could now be explained in terms of a few simple and appealing ideas. And better yet, these ideas could be extended to understanding most problems involving second order phase transitions.

Many people have written about the difficulties which they experienced in getting their ideas accepted by the scientific community. I have never had much problem in this direction. I have always been an establishment figure, and most of my good work has gotten all the recognition it deserved. I had this 'Christmas vision' while I was on the faculty at Urbana. My colleagues there said they liked the work. My collaborator, Gordon Baym, had particularly warm words, while my seniors, John Bardeen and David Pines, nodded approval. I am told my preprint elicited seminars on

---

[13] E.C.G. Stückelberg and A. Peterman, Helv. Phys. Acta. **26**, 499 (1953). Murray Gell-Mann and Francis Low Phys. Rev. **95** 1300 (1954).

the work at Harvard and Cornell. My joy was increased when I spoke to Mel Green, who told me about data that had been presented at the Bureau of Standards conference, which supported my conclusions. I also recalled Paul Martin's statement about 'magic' relations, and then discovered Ben Widom's papers which had just been published in the Journal of Chemical Physics[14]. Widom had developed scaling arguments which gave many of my conclusions. Later on I found out that Patashinskii and Pokrovskii had also gotten the right answers at about the same time.[15]

My own work was published in the journal *Physics.* It is paper B2 in this volume. My answers in this paper were the same as that of Widom and of Patashinskii and Pokrovskii, but my point of view was somewhat different. There were, in fact, quite a large number of contributors to the development of these new ideas. In hindsight, it is easy to see that we--all of us-- had discovered and invented a wonderful new world. Near critical points, thermodynamic systems show a very special behavior. If you examine the system on distance scales which are much larger than the typical microscopic distances, you see correlations which extend over these very long distances. The system has to chose to fall into one or another thermodynamic phases, and is showing the vacillations in its decision-making. This process of choosing produces a beautiful, closed world with its own rules and its own, internally consistent explanations. The world had been penetrated and could be explained.

The next job was to see whether the ideas that had been developed all jibed with the experimental facts. To do this, a group of us got together at Urbana and ran a seminar aimed at looking at all the known experimental and theoretical data about critical phenomena, to see whether it fit into the newly proposed pattern. This group, W. Gotze, D. Hamblen, R. Hecht, E. A. S. Lewis, V. V. Palciauskas, M. Rayl, J. Swift, D. Aspnes, J. W. Kane and myself, looked at all the available literature and gave seminars on what we had learned. We convinced ourselves that all the data was consistent with the new point of view, and that much of it could be explained by the new scaling theory of critical phenomena. We reported our conclusions in the review paper, which is number B3 in this volume.

---

[14]B. Widom J Chem. Phys. **43**, 3892,3898 (1965)

[15] A.Z.Patashinskii and V.L. Pokrovskii, Soviet Phys. JETP **23** 292 (1966).

***

At some point in the development of this review paper, I am not sure when, a new idea was added to the concepts of scaling and running couplings. This idea has the modern name 'universality', which I first heard applied in a conversation with Sasha Polyakov and Sasha Migdal in a dollar bar in Moscow. Their use of the word came from descriptions of field theoretical solutions to problems of this kind. The basic idea is that there are only a few different solutions to near-critical problems. Using the solutions, one could group the problems into 'universality classes'. Many apparently different problems have the same solution and belong in the same class. In critical phenomena, the universality classes are in large measure defined by giving the dimensionally of the system and by describing the kind of information which the system must transfer over long distances. As a specific example, liquid gas phase transitions and Ising-model-magnets show exactly the same behavior near their respective critical points.

This universality idea provides a truly new way of looking at problems in statistical mechanics, field theory, hydrodynamics, etc. Instead of solving each problem, one looks for classes of problem each class having a common solution. We essentially used the idea (before it had a name) in organizing our review paper. It had been explicitly used in the previous work on critical phenomena in Helium by Pippard, and had been implicit in the work on Ising models of the King's college school. Bob Griffiths added considerably to the concept by pointing out its relation to the geometry of thermodynamic surfaces. By looking back on my Ising model work, I could see that the universality idea was clearly and closely related to the scaling concepts. The universality concept said that one could classify the different microscopic worlds produced by critical phenomena. This idea made critical behavior itself into a ideal world, beautiful and self-contained.

Over the next few years, I had very considerably pleasure in exploring the corners of our new world. Paper B4, by Swift and myself, describes how to set up a cellular automaton, i.e. a parallel processing computer, so as simulate hydrodynamic equations. Such simulations have been much developed since then. (See, for example, paper E6, which makes use of another automaton.) Paper B4 with Jack Swift is one of a pair done at roughly the same time we were aiming at developing a theory of dynamical critical phenomena. Paper B4 was interesting, maybe even important, but not of much use in critical phenomena. On the other hand B5 describes the work

Jack Swift and I did on developing the beginnings of a theory of transport behavior near critical points. As always, new ideas worked from solid older ones. Here the original ideas were creations of Marshall Fixman and Kiozi Kawasaki. My own development toward these ideas go back to work I had done on transport theory in my thesis (paper 6 with Paul Martin, on my publication list) and to a long series of papers I had done on transport with a wide variety of collaborators.[16]

Papers B6 and B7 explain and describe the world of scaling. In particular, paper B6 look back to the older work of Fisher and of Buckingham and describes how their ideas can be incorporated in the new viewpoint. Paper B7 summarizes what we had learned about the phenomenological theory of critical systems, from a slightly more mature perspective than that of the 1967 review paper.

One important result in the theory of critical phenomena is what is called the operator product expansion or short distance expansion (See paper B8). This expansion, due to K. Wilson and myself, is based upon the idea that, at criticality, there are a limited number of different local fluctuating quantities. This idea permits one to express products of fluctuating quantities at criticality in terms of an expansion in the fundamental critical quantities. This idea provides a mechanism for understanding the critical symmetry in a rather deep way. In paper B9, H. Ceva and I explore this expansion for the two-dimensional Ising model.

Meanwhile the field had exploded. Ken Wilson had absorbed the new viewpoint, and magnificently extended it. He introduced two rich new concepts: the fixed point and the space of coupling constants. With these concepts the scaling and universality point of view became a theory: the renormalization theory of critical phenomena (and of much else). The remaining papers in this chapter trace out how the scaling-university theory and the renormalization group work together. Paper B11 attempts to be a formal review of my contributions to the whole area. I should also point back to paper A5, which describes how the running coupling constant concept found

---

[16] During the early part of my career, I spent a lot of time working on transport properties of various many particle systems. All this work formed a basis for my understanding of the relation between the macro and micro levels of description. The resulting papers, especially #13 (with D.C. Langreth), #15 (with M. Revzen), and #25 (with J.W. Kane) formed an important part of my education.

***

application in quantum chromodynamics. Paper B10, a joint effort with Humphrey Maris, describes how some of the new ideas could be presented at a level which is accessible to undergraduates.

I should emphasize once again my tremendous personal gratification in the accomplishments of the whole field. The story I have told has many actors and many excellent accomplishments, among them my own. In the end, we can say that a group of scientists created something beautiful, which was not there before. I saw it happen. I saw understanding and nature come into correspondence. It was great to have been there.

But science does not stand still to admire past accomplishments. Immediately after Wilson had invented the renormalization group, Franz Wegner showed how to phase the renormalization arguments into a formal algebra of coupling constants. Wilson and I independently (see papers B8 and B9) described the complementary algebra of fluctuating quantities near the critical point. My own work, which I did jointly with many collaborators[17] carried forward the understanding of these algebras in the context of two-dimensional critical behavior. Then Polyakov invented conformal field theory, which proved to be a tool which would enable one to understand all the behavior of almost any critical theory in two dimensions. The algebraic technique was refined and perfected by the field theorists. The scientific area of critical behavior exploded and indirectly contributed to the development of string theory.

In the meantime, with many collaborators, I contributed to the solidification of the gains that had been made in understanding critical phenomena. For example, Franz Wegner and I (see #43) developed concepts related to continuously varying critical indices.[18] N. Berker, R. Ditzian, Gary Grest, Michael Widom and I applied these concepts to other statistical mechanical systems. This work was

---

[17] Among these coworkers are H. Ceva, Alan Brown, Michael Widom, Ad Pruisken, and Gary Grest

[18] I lost a bet involving a bottle of bottle of Scotch (or Vodka) to A.A. Midgal on this issue. I had bet that all two dimensional critical indices were rational numbers. I figured I could not lose, since it is usually very hard to prove that a number which arises in the context of some physical problem is not rational. But then Rodney Baxter proved that in the eight-vertex model, indices vary continuously. I paid off. The Midgal's junior (A.A.) and senior (A.B.) generously repaid with an 'irrational' vodka bottle which had been bent into an impossible shape in their home workshop.

***

subsequently carried forward and extended with many other people[19] In another series of papers, A. Houghton, M.C. Yalabik and I explored the consequences of the real space analysis of renormalization near critical points.

The understanding our little community has generated has, in the course of time, spread far beyond the study of second order phase transitions. The scaling and correlation ideas form a crucial part of the phenomenology of particle physics and reappears in astrophysics in, for example, analyses of fractal distribution of matter in the universe. Classical Mechanics and hydrodynamics contain many scalings, with the behaviors in turbulence and near the onset of chaos being most like those in critical behavior. In materials science, since the work of P. de Gennes, B. Mandelbrot, and T. Witten, people are always looking for some sort of scaling or universality. We have come far by looking at the small corner of physical behavior very near critical points. In part, we have accomplished our trip by following Onsager's solution and squeezing out every generalization that could possible be found in that rich reservoir.

Beyond the export of our specific technical methods, we have also exported a point of view, encompassing the way in which one might look at the structure of physics. One image of this structure is that each little world of phenomena is really based upon the physical laws which describe a more fundamental level of reality. This image leads one to a reductionist outlook. Then one would say that the true goal of physics should be to reach deeper and deeper toward the basic laws which describe the fundamental interactions in the world. However, the study of critical phenomena and other topics in condensed matter physics pushes one toward another and complementary image of nature. This view was expressed in P.W. Anderson's thoughtful paper 'More is Different'[20] in which he pointed out that every level of reality can have its own deep and fundamental laws. In studying a particular little model world, the goal of the scientist is threefold: First, to expose the fundamental laws in their most general form and to show how they work out in the specific system in hand. Second, to show how this particular

---

[19] My collaborators during this period include most of a generation of Dutch Statistical Mechanics-- Ad Pruisken, Marcel den Nijs, and Bernard Nienhuis -- and other coworkers including J. Banavar and Morgan Grover.

[20] P. W. Anderson Science 177 393 (1972)

level of experience is related to other closely connected parts of reality. And third, to take the ideas generated in the study of this one particular part of the world and apply them to other portions of the world. The study of critical phenomena is not in any way unique. However, it is a particularly successful and beautiful example of the generation of deep ideas about the simplified world of critical fluctuations, about how those fluctuations are defined by the microscopic behavior, and about how these ideas manifest themselves in macroscopic behavior. The ideas generated were exported widely and illuminated, in an unexpected fashion, many other areas of science.

C. Simulations, Urban Studies, and Social Systems Models and Arguments

Science includes the critical application of ideas to real world situations. In the 1950s, 60s, and 70s I found many applied problems very interesting. While I was still in college, I worked for the Guided Missile Division of Republic Aviation Corporation, helping with the proposal stage of the design of robot aircraft. In graduate school and thereafter, I worked for the AVCO corporation doing heat transfer calculations related to the design of guided missiles and space probes. I found the latter job extremely instructive. It enabled me to interact with some scientists whose paths I would not normally have crossed, including Hans Bethe, Jim Keck, and Arthur Kantrowitz. I was very impressed by the intellectual vitality of the work at AVCO, particularly the branch in Everett, Massachusetts.

But after a while, I became convinced that this kind of work was socially unproductive. During the Eisenhower years I had been quite uncritical about most of our (U.S.) foreign policy. However, while I was a postdoc in Copenhagen, our government did things I felt that I could not explain to the 'foreigners' around me. The Bay of Pigs invasion of Cuba was the example that I best remember. Having no explanation for others, I became less convinced that our role was always a good one. One defining event for me was sitting in the cafeteria at lunch time at AVCO's Missile Division in Wilmington, Massachusetts discussing the dangers to the world of nuclear weapons. I even wondered whether it was prudent to pick up stakes and go someplace like Brazil. This discussion was taking place during the Cuban Missile Crisis. On the way to work that day, I had driven by Boston's Logan airport which had a small section covered with Strategic Air Command bombers, all probably loaded with nuclear weapons. At that time, I could not see why the US could not tolerate

the same kind of proximity of enemies that the Soviets had to endure. So I began to lose sympathy with US Cold War policy. This progression in my thinking continued through the period of the beginning of the Vietnam War. Naturally, I could no longer work in the missile industry.[21]

So, for a while, my applied interests had no good outlet. But then, starting in about 1967 I began work on the use of scientific techniques to describe and control urban social and economic phenomena.

I got started in this work though the efforts of Dale Compton who was then head of the Coordinated Sciences Laboratory at Urbana. He introduced me to Jerrold Voss, an urban studies specialist. I worked on their project for studying tendencies in urban real estate prices and land use in the little town of Kankakee, Illinois. I do not know what I or they expected to gain from my involvement in this work. I brought essentially no knowledge whatsoever to the project. I do not know any economics or sociology. I got involved in the computer side of the project, supervising the computer specialist in the Urbana laboratory. In this computer work, we absorbed great piles of historical data about real estate prices in Kankakee and tried to display them in some way which would enable us to understand what was going on. Little by little I gained some rudiments of information about the urban scene, and some small knowledge of computers. (The actual computer at the Coordinated Sciences Laboratory[22] was large and far too complex for me to program.) I did not contribute much but I was proud, and somewhat flattered to be included in their work, and its writeup, C1 in this section.

Then, I moved to Providence, Rhode Island, and Brown University. At that time there was a great national push toward understanding the dynamics of urban development. Jay Forrester of MIT had developed a computer model of urban change, which took a very

---

[21] To complete my story: The lunch-table discussion ended when one of the executives suggested, with some annoyance, that we should all get back to work. We did.

[22] This Coordinates Sciences Laboratory (CSL) work eventually replaced my work in the missile industry. Life does have its ironies. This laboratory was almost fully supported on Department of Defense contracts. As a result of complaints about the Vietnam war by people like myself, after a while it became impossible for labs like CSL to use military funds to contribute to civilian research.

***

simplified view of a urban society, produced a quantification of that view in terms of a computer model, and then used the output of that model to prescribe social policy. I did not like the policy prescribed. He basically suggested that by removing housing for poor people the city could free up land which could be used for industrial expansion, to--in Forrester's view--the long-range benefit of all. In my view, this policy made the poorer section of society pay for the economic development which would occur in the city, and tried to sanctify that policy by giving it the blessing of apparently-objective computer output. Consequently, I set out to use the modeling tools that Forrester had developed to reach conclusions which were more to my liking. Brown University was a good environment for this because they had a computer system which was set up so that anyone could use it. I went to work reproducing Forrester's model in a more flexible computer environment.

Brown University contained a group of people interested in urban public policy. Eventually, the group of urban research coworkers at Brown would include Benjamin Chinitz, Graham Crampton, Bennett Harrison, Susan Jacobs, John Tucker, and Herbert Weinblatt. We took as one of our tasks the incisive criticism of Forrester's work. The parameters and structure of his model had never been verified against real data. We could have tried to disprove the model by looking at the world. However, that would be difficult or impossible since many of the social concepts in the model are very hard to quantify. Instead, we chose another route. (See C2 and C3). Our group changed the focus of the model from a one-city application to a national application. The nation was viewed as being completely made up of cities. We insisted on keeping track of the poor people. This social group would be tend to be squeezed out of the city by a local application of the housing-removal strategy, and would be just squeezed by a national application of that policy. We also tried to develop a more objective (or perhaps just different) set of criteria for the success of the policy. Our first result was that while not changing the model at all, we could reach opposite conclusions from that of the Forrester group. By our criteria, the poor population was highly squeezed with little overall gain to any population segment in the urban society.

This first work can be construed in two ways. One truth was that you got out of such models just about what you put into them. They were mostly a way of recording your preconceptions. The second truth was that such preconceptions could be fluently incorporated

into models of this type and one could, in fact, reach any conclusions that one wanted. Hence, at one level, this work could be construed as a criticism of all model-building in which ones preconceptions were not tested against data.

I am not sure that we listened to our own criticism fully. We went on to build other models which more accurately recorded our own prejudices (see paper C4) and points of view. But, after a while the point we had made began to sink in. If these models really represented little more than we could say in words, why not leave out the computer? The construction of this sort of computer model seemed to be a rather pointless endeavor. For this reason and others, I moved away from urban studies.

Nonetheless the experience had taught me much. My collaborators had valuable insights into the functionings of a city and I was pleased to absorb some of these. And I had learned a new technique: the use of the computer to model experience. The computer would prove crucial in the next stage of my career, in which I would work on hydrodynamics and chaos.

My experience in urban work extended beyond research. I worked (unpaid) for the Rhode Island State Planning Department as chair of a committee which constructed criticisms of all proposals for federal funding of public programs. The point of such a committee is to see that everyone talks to everyone so that there is no unnecessary unhappiness engendered by such proposals. Since I enjoy talking, this was a fun job for me.

At the same time, I taught courses for undergraduate majors in urban studies at Brown. We discussed modeling, and the spatial structure of cities, and social policy and lots of other things. Such teaching was a rewarding experience for me, particularly because I could no longer do the kind of teaching which dominated my early career. At Urbana, I spent a lot of time teaching graduate students either in large classes or one-on-one as their graduate thesis adviser. But by 1971 or so, this kind of teaching became much less attractive. Within a few years, a large fraction of the jobs which had been available to physics graduate students dried up. Our students were not finding appropriate work. Consequently, graduate-school teachers like myself felt their work to be unnecessary.

The climate for employment of advanced-degree students in physics had basically changed because of the Vietnam War. Many physicists and other scientists were critical of the policy of the government.

***

Military contractors looked for potential workers who would not criticize governmental aims and policies. Gradually engineers replaced physicists in the 'defense' industries. Fewer physicists hired meant fewer graduates needed, meant fewer teachers needed to train the graduate students, means fewer physicists hired .... In any case, there was little satisfaction in training large numbers of graduate students in this period.

However, the job of training urban studies specialists is not very satisfactory either. As pointed out by Alan Altshuler[23], one problem of the urban specialist is that there exists no large base of real knowledge, information, or technique which would enable the trained person to better prescribe urban planning policy than the ordinary intelligent citizen. The specialist has nothing special to add. As you can imagine this situation is rather discouraging to the teacher, who then has nothing special to teach.

So my urban public policy phase gradually withered away. When I accepted a job at the University of Chicago in 1978, I knew that the City of Chicago and its University were both far too professional to allow an easy outlet for my urban interests. After I came to Chicago, I devoted myself single-mindedly to teaching and research in the physical and mathematical sciences. But some residue of my social interests remained. Within the framework provided by an occasional column in *Physics Today* I tried to comment upon the social and economic context in which physicists work[24]. Three more columns are included at the end of this Section on the urban scene and public policy. These columns are indeed about the relation between the physics community and public policy. I enjoy writing the columns, and I even think that many people enjoy reading them. But Altshuler's criticism of the urban planning specialist can equally well be extended to most columnists. The question to ask is: "By what process have you become qualified to offer us advice?"

D. Turbulence and Chaos
Questions without Answers

In the last dozen or so years, my interests have turned to yet another field of science: the description of complexity. My own view of this

---

[23]Alan Altshuler, The Urban Planning Process, Cornell University Press, Ithica (1965).

[24] The reader has probably noted that I have also taken advantage of these introductions to pontificate in this direction.

***

field starts with a major intellectual problem: We know that the laws of physics are rather simple in structure. Newton's laws or the Schrodinger equation or even string theory is described by a rather simple system of equations. Ones expectation might be that such simplicity in formulation should lead to simplicity in outcome. However, all our experience in life contradicts any expectation of simple outcomes. The world is wondrously complicated and bewilderingly diverse. How can it be that from simple beginnings one gets complex endings.

This problem is not illuminated by much of the traditional practice of physics. Most systems picked for study by physicists are picked precisely because they have simple outcomes. Kepler's orbits or the quantum harmonic oscillator or the simple pendulum have been extensively studied and are used as examples just because they have simply predictable outcomes. Even the many-body systems studied in critical phenomena have simply predictable outcomes. But the world tends not to be predictable in the same way. Once again our life experience suggests opposite outcomes from those of the exemplary physical systems. A pot of boiling water, or a double pendulum, or any person will not behave in expectable or predictable ways. Why should our professional experience with physics belie our civilian experience as human observers of the world.

But the conundrum is worse yet. Most of the systems traditionally studied by physicists remain equally complicated throughout their history. A microwave cavity, any pendulum, or any electrical circuit with passive linear elements does not gain extra complexity as time goes on. However, if you take a hot mass of rock of an appropriate size and throw it into orbit around the sun, you might observe it to gain all kinds of organizing features. Oceans and continents will develop and move around. Mountain ranges will grow and decay. Great deserts will spread out and then be covered with ice. Little cells of organization will emerge, become more complex, and form into trees and us. Is all this development of complexity and organization outside the laws of physics and of science?

Our prejudice as scientists is to say *no! Physics encompasses all.* But our job as scientists is to see and understand. How is it that simple laws can give complex outcomes? How can organization develop from blind law and naked chance? A major thread of development in the modern physical, mathematical, and biological sciences has been devoted to answering this class of questions. My own work on this area has been modest in outcome, but quite exciting and meaningful

to me. These two last sections of the volume are devoted to work on the borders of this great subject.

For me, the story starts in about 1982. I am working at the University of Chicago, doing phase transitions and critical phenomena, and being somewhat dissatisfied with my work. In the great days of the 1960s I was all excited with this subject, and its creative possibilities. Later on, I got considerable joy in seeing what Fisher, Wilson, Wegner, Nelson, Polyakov, and many others had created based upon the earlier work. But, by this time the freshest joy of discovery, either directly or though the work of others, has begun to go out of the field. Then, at a crucial moment, Bob Gomer, a colleague in Chemistry, asks why I am devoting so much work to a particular model system. His implication is that the model is not so real as to be of practical interest, and perhaps not so deep as to have real intellectual interest. My reply seems superficially convincing, but I know that his implied criticism is right. So, I resolve to learn something new.

The new subject I find is *dynamical systems theory* the study of the time development of relatively simple systems. There are portions of the subject which are absolutely beautiful closed topics, not as deep as critical phenomena but roughly similar in structure. Two of these beautiful topics are *iterated one-dimensional maps* and the *onset of chaos.* The first is about problems in which a system is described by a single variable, say its degree of excitation, and jumps from one value of the excitation to another via discrete time steps. The second is about a kind of continuous transition which might occur when a system first shows chaotic behavior. In 1982 or so, my students Albert Zisook, Michael Widom, Scott C. Shenker and I set out to learn about dynamical systems theory. Our work is was filled with great thrills of invention and discovery and then the partial disappointments of finding that our creations had been preempted by previous workers. Some of the structure of this subject in reviewed in papers D1 and D2.

Eventually, we found our own area which could serve as an entry to this research field. We built upon the work of Mitchell Feigenbaum, who had first understood the onset of chaos via a renormalization group calculation. His work was built upon the tools I had previously known: scaling, universality, and renormalization. When we began, there were two kinds of onsets which had been studied both of which involved a period doubling route to chaos. We looked to other routes

and learned about them through the development of renormalization calculations. Some treatment of this subject can be found in paper D3.

Thus, our strategy was to enter this new field by considering systems which produces time-sequences of data-- like annual populations, or economic outputs, or daily stock prices-- and to see how such a sequence may develop chaotic attributes. These systems were picked to be just complicated enough so that we could fully encompass them with the computers we then had available. (Here was the point at which I would see the computer knowledge gained by studying urban problems pay off for me in genuine, hard science.) Papers D3 and D4 describe the results of this kind of analysis. All the examples studied are just simple mathematical toy systems picked to be analyzed with computer and then with techniques like the renormalization group. In this way, we began to understand the development of complexity in toy examples.

However, these examples also proved to be unsatisfactory in the long run. They were both too simple and too complicated. At onset of chaos they were simple. In their chaotic regimes they could show a behavior which, when studied in detail, was bewildering diverse and did not seem to lead to useful and interesting general principles. We were stuck once again.

The best source of theoretical inspiration is the real world, and the best window to that world is provided by experiments. The systems studied by experimentalists have a much better chance of organizing themselves than computer toys. The real world has much, much more time to organize than our toy examples. A computer might analyze 10^8 or so very simple events per second. (In looking at our models, we looked at much fewer events, perhaps a few thousand time steps in every second of computer analysis.) However, a real system--say a cubic centimeter of gas--will have 10^{32} or so collision events per second. Naturally, the real system will have a greater opportunity to 'get organized'. Conversely, our computer models will have to be picked with very great care if they are to give any inkling of true natural phenomena.

However, with luck or skill, one can relate model systems to real-work outcomes. One very pleasing example of how to do so is given in paper D5, authored by Jensen, Libchaber, Procaccia, and Stavans. This work is devoted to showing how a real system, a pot of liquid mercury heated from below, first develops chaos. The onset in this real system follows one of the universal routes to chaos, the so called

quasiperiodic route, which was extensively studied by the Chicago theoretical group. This paper itself is mostly experimental in content. The system studied is the brainchild of Albert Libchaber, its actualization is due in large part to Joel Stavans. Libchaber understood how to adjust the system to push it into the domain investigated by the theorists. And pleasingly enough, after adjustment the system behaved in exactly the fashion predicted by theory.

It is important to recognize that our understanding of the physical world is due in very large measure to the experimental scientist. Pure thought gets us only so far. The worlds which the theorist may construct are only powerful and rich because they are modeled on the real worlds that nature produces. And our only access to such real worlds is through the studies of the experimental scientist. Without experimental check, the universality of the onset of chaos is a theorists' possibility. The experiment shows that it is a reality.

In addition to illustrating the craft of the experimentalist, this paper also has a pleasing theoretical content. According to the renormalization theory, near onset dynamical systems produce a characteristic structure in which very delicate and small structures are produced of excitation versus time. The incipient chaos is, in fact, built into structural details which persist to arbitrary small scales. The word for such a delicate and scale invariant geometrical structure is a *fractal*. This word was coined by Benoit Mandelbrot, who brilliantly argued that such structures were rather pervasive in nature, and then introduced many clever and incisive techniques for producing and analyzing them. Our description of the experiment was based upon a study of the properties of the fractal produced at onset in this hydrodynamic system. This in turn pointed back to a previous paper on the nature of fractals of this type, D6 and to my concern about the proper use of the fractal concept. (See D7). (See D8 for a summary of some research on multifractals.)

Paper D6 is in some ways an embarrassment to me. It is an excellent expository work which in fact helped make a particular form of analysis of fractal behavior very fashionable. The form of analysis is called multifractal or multiscaling analysis and follows from the work of Mandelbrot[25], Hentschel and Procaccia[26] and others. The analysis

---

[25]B. Mandelbrot J. Fluid Mech **62** 331 (1974).

[26] H.G.E. Hentschel and I. Procaccia, Physica **8D**, 435 (1983).

***

closest to the one in our paper was done by Georgio Parisi and Uriel Frisch, published in a slightly obscure place[27], but explained to me in some detail by Parisi. Unfortunately I either forgot or did not understand his explanation, so that we did not appropriately cite this earlier work in ours. Fortunately, an erratum[28] and a bit of publicity[29] set the record straight.

I was involved in a few more research projects in dynamical systems theory, one with Charles Amick, Emily Ching and Vered Rom-Kedar, another with Mahito Kohmoto and Chao Tang (see #91), and still another with Oreste Piro and Mario Feingold (see # 127). This work was pleasing, but we were never again able to use dynamical systems theory to achieve the generality or experimental relevance of the work on the mercury cell. We were stuck once more.

But we did follow another theoretically motivated line of research, based upon simple mathematical models of extended physical systems. In 1981, Thomas Witten and Leonard Sander[30] developed a models of the dynamical aggregation of cluster called DLA. This resulting clusters exhibited a kind of universality along with scaling and fractal behavior. At Chicago, we followed up on these ideas and applied them to other model systems. [31]

But, then we turned back to experimental systems. With David Bensimon, who was my student, Boris Shraiman, a postdoc, and Albert Libchaber I started looking at fluid flow, which can produce a genuine and rich complexity in both space and time. The example they suggested is the flow of two liquids confined in a very narrow space between two glass plates. This setup is called a Hele-Shaw cell. It is capable of producing both a very simple behavior and also a rich complexity. A portion of the theory of such systems is outlined in the review paper D9 in this section. This paper was followed by

---

[27]Georgio Parisi and Uriel Frisch On Turbulence and Predictability in Geophysical Fluid Dynamics and Climate Dynamics M. Ghil, R. Benzi, and G. Parisi, editors (North Holland, Amsterdam ,1985).

[28] T.C. Halsey, M.H. Jensen, L.P. Kadanoff, I. Procaccia, B. Shraiman, Phys. Rev A**34** 1601 E (1986).

[29] Barbara G. Levi, Physics Today April 1986 p. 17.

[30]Thomas Witten and Leonard Sander, Phys. Rev. Letters, **47**, 1400 (1981).

[31] Some of the original work can be found in the publication list. See,#89 (Alexander, Domany, LPK, and Bensimon) and #102 (LPK and Liang) as well as #98 (Bensimon, Shraiman, and LPK).

***

many others[32] as we once again constructed theory in response to the experiments of the Libchaber laboratory.

Another closely related experiment is discussed in paper D10.

Paper D11 describes how concepts drawn from dynamical systems and other chaotic situations can be come part of undergraduate teaching. I have spent a very considerable period of time designing courses which use computers to teach about chaos. This teaching is a fitting and satisfactory outcome for the research directed at chaos and dynamical systems.

E. Complex Patterns

Section introduction: From Correlation to Complexity.

In the previous section, I discussed how apparently complex patterns could be produced by dynamical systems. My own study of complexity has been going through a kind of transition in recent years. I have begun to realize that the dynamical systems usually studied often show only a very limited and partial complexity. Often these systems produce complex time patterns, but the complexity is limited to one or a few chaotic functions of time. The true richness of natural systems go well beyond few-variable chaos. This richness has been emphasized by the people at the Santa Fe Institute, and most particular Murry Gell-Mann[33] and Stuart Kauffman[34]. Specifically, Stuart argued that a sufficiently complex dynamical system would naturally develop a rapidly increasing complexity. Structures would emerge with rapidly proliferating complexity. Details of history would determine exactly which structures were produced. However--in this view--the production of structure was an inevitable occurrence in a system with a sufficient number of variables and interconnections among them.

---

[32] This work eventually grew into a very big activity, finally involving extensive interactions with people in the Math Department at Chicago. So far, I had the pleasure of collaborating with Bensimon, Shraiman, Libchaber, Giovanni Zocchi, Bruce Shaw, Wei-shen Dai, Su-min Zhou, Peter Constantin, Todd Dupont, Ray Goldstein, Michael Shelley, and Andrea Bertozzi in papers on the Hele Shaw cell.

[33] M. Gell-Mann. The Quark and the Jaguar, W.H. Freeman, New York, 1994.

[34] Stuart Kauffman, At Home in the Universe, Oxford University Press, New York, 1995. Stuart Kauffman, The Origins of Order Oxford University Press, New York, 1993.

***

Real extended systems encourage complexity. Most truly interesting dynamical systems (you and me for example) involve dynamics in both space and time. The spatial dimension permits a structured complexity of two different sorts. First there is translational invariance. This invariance permits structures to repeat again and again at different point in space. But in a chaotic system, nothing quite repeats--instead it is reproduced with different variations each time. For example, all the people in the world are in one sense similar structures, but they have interesting and important individuality that in the end, make life interesting. Another aspect of this structured complexity is that it exists on all length scales. There are people, and organ systems, and tissues, and individual cells, and the working parts of cells, and individual molecules. Each of these different scales shows its own, unique, organized behavior-Each scale's behavior is repetitive and complex and history dependent.

Recently, condensed matter physics (following biology and many other disciplines) has moved into the study of such multi-structured complexity.

As part of one kind of study of this subject,_I have been working on turbulence. In my whole career, and particularly at Chicago, I have been blessed by many coworkers who have led and immeasurably enriched various aspects of my research. Theoretical visitors, especially Procaccia, and a whole rich collection of graduate students and postdocs have made my research what it is. But, probably, I learned most from my collaboration with Albert Libchaber.

For several years, Libchaber and I worked as a team in studying complexity. He led the way in visualizing and realizing complexity, I tried to apply theoretical tools to understanding what he had done. We studied three systems together. I have already mentioned the onset of chaos in the mercury cell and the Hele Shaw system. But, the richest case which we studied together was convective turbulence. Turbulence is chaos in both space and time. When a relatively large volume of fluid is heated from below, the fluid on the bottom expands and tends to rise. Thus the fluid is set into motion. At high heating rates the motion becomes unstable and turbulence results. This kind of system is called a Rayleigh-Bénard cell. Our work is summarized in E1, E2 and E3.. My work in this area has been the product of collaboration with many people including Bernard Casting, G. Gunaratne, F. Heslot, S. Thomae, Xiao-Zhong Wu, S. Zaleski, and S. Zanetti, Anton Kast, and Masaki Sano. An additional study was done by Dan Rothman and myself as paper E4. This paper involved using

***

a very simple kind of computer model, called a cellular automaton, to produce the kinds of swirling motion seen in the Rayleigh-Bénard cell. It was a very pleasant kind of *tour de force* and produced lovely pictures and patterns.

This turbulence work is and was a dream shared by Libchaber and myself. We hoped we could and would really understand how a hydrodynamic system developed a very rich complexity. Something about this dream came to reality. Our work did expose many aspects of the behavior in the cell, including many aspects of the complicated geometry produced by the motion. (The cover of this book shows four different pictures of this geometrical structure.) The experimental work cast into doubt the general applicability of the Kolmogorov theory to systems of this kind. This theory involves a cascade of energy from larger scales to smaller ones. The cascade produces some kind of fractal structure. But then, as always, the work reached a barrier. We had gotten all the knowledge we could from the experiments on convective turbulence, and it was time to move on. Albert and I turned to different approaches to complexity, he in biological systems, and me through the mathematical sciences.

In papers E5 and E6, I turned away from real turbulence and instead studied a simplified model of the cascade process built into the Kolmogorov theory. This work involved a system simpler than fluids, but one which nevertheless showed a rich structure on a wide range of length scales. Of course, because this model does not begin to show the richest kind of complexity and structure, it is in the end slightly unsatisfactory. (Please see #168, #170, and #171 for more detailed writeups of my collaborations on this subject with Roberto Benzi, Detlef Lohse, Jane Wang, and Norbert Schörghofer.)

One kind of structure repeated again and again in hydrodynamic systems arise from the propensity of these systems to produce mathematical singularities. Papers E7, E8, and E9 describe the universal structures which arise from this kind of singularity mechanism. But once again one gets nothing like the complexity shown by the natural world, so the analysis feels incomplete and somewhat unsatisfactory. At least this is somewhat the feeling I get from rereading the papers. When the chase is on, investigations like these are always great fun and wonderfully satisfying. There is a neatness and completeness in the classification of the universal properties of relatively simple similarity solutions, which stands in contrast to the complexity and messiness of the real world.)

***

Systems in Statistical Equilibrium can only show a small amount of complexity. This limitation on complexity is built into the equilibrium statistical mechanics of Boltzmann, Gibbs, and Maxwell. Their statistical mechanics formalism is one in which the probability of a given configuration is proportional to an exponential of the Hamiltonian divided by minus the temperature times the Boltzmann constant. Such a weight prevents the most complex configurations from dominating the system. However, if a system is not in statistical equilibrium, this exponential weight no longer holds. Something much more complex may emerge. The next series of papers, is based upon particles which cannot go to true equilibrium because their interparticle collisions do not conserve energy. Among the physical systems which show these properties are glass balls or ordinary sand. We use the phrase 'granular material' to describe these systems. Papers E10. E11, and E12 are all aimed at elucidating the properties of these theoretical models of 'sand' in the simplest possible geometries. In fact, we observe that the granular material does have a far richer and more structured behavior than the usual equilibrium systems. Their main characteristic is that portions of them may freeze or slow down into a glassy state. This slowing or freezing then serves as a memory of the history of the system. Sand castles are examples of such memory.

Memory if the subject of the next paper, E13, which aims at explaining how non-equilibrium phenomena can serve as part of a solid state memory device. None of these works actually reach to the levels of complexity familiar from everyday life, but they are all essays in the right direction.

Scientific work on turbulence and complexity is far from completed. Science cannot yet answer, or even properly formulate, questions related to how complexity does actually arise in the world. I started saying that computers can do 10^8 or so calculations per second while a bit of gas sees 10^{32} events/second. But biological systems are even more richly complex. In the course of evolution there might have been 10^{55} or so collisions among atoms in biological systems. This provides lots and lots of steps in which different levels of complexity might have developed. We certainly do not understand the development of complexity. In fact, we do not even understand how the patterned formed by the clumping of matter as it is pushed

***

around by the motion of a fluid.[35] But, unanswered questions and partially formulated problems are the bread and butter of science. I am pleased to see that my table is still full.

[35]This has been a subject of study for me in recent years with my principal collaborators being Mario Feingold, Oreste Piro, and Peter Constantin, Itamar Procaccia, and Emily Ching.

***

Notes on Fourier transforms on lattice and continuum

Fourier transform is a powerful technique often used in the analysis of collective excitations in macroscopic systems. A simple reason why it is so successful is the existence of translational symmetry in such systems. Thus each Fourier mode often behaves independently, and the analysis of a many-body problem becomes much simplified after the Fourier decomposition. Photons, phonons, and plane-wave states of electrons are good examples of such analysis. The notes below, taken from the textbook *Principles of Condensed Matter Physics* by P. Chaikin and T. Lubensky, explains the commonly used conventions and properties when performing Fourier transforms (for condensed matter physicists).

Appendix 2A Fourier transforms

In this appendix, we will review Fourier transforms for functions of one- and d-dimensional continuous variables and for functions defined at lattice sites on one- and d-dimensional lattices.

1 One dimension

We begin with a function $f(x)$ of a single variable x in the interval $[-L/2, L/2]$ (i.e. $-L/2 \leq x \leq L/2$). If $f(x)$ satisfies reasonable continuity and boundedness conditions (e.g. it does not have an infinite number of zeros in some finite interval of x), it can be expanded in a uniformly convergent Fourier series:

$$f(x) = \sum_q \psi_q(x) f(q) \,, \tag{2A.1}$$

where $\psi_q(x)$ satisfies the same boundary conditions as $f(x)$. Common boundary conditions on $f(x)$ are $f(x = \pm L/2) = 0$ or $f'(x = \pm L/2) = 0$. In condensed matter physics, one is often interested in bulk systems in the thermodynamic limit, $L \to \infty$, for which most physical properties of interest do not depend on the boundary conditions. In this case, any physically reasonable boundary condition can be imposed. The periodic boundary condition requiring $f(x)$ to be a periodic function of period L,

$$f(x) = f(x + L) \,, \tag{2A.2}$$

is computationally the simplest and is almost universally used in situations where surface properties are not relevant. The condition (2A.2) is equivalent to wrapping the line of length L on a circle of circumference L and tying the two ends together. The functions $\psi_q(x)$ must satisfy the periodic boundary condition and can be chosen to be

$$\psi_q(x) = Ae^{iqx} \,, \tag{2A.3}$$

where

$$q = \frac{2\pi}{L} n, \qquad n = 0, \pm 1, \pm 2, \ldots \tag{2A.4}$$

and where A is an arbitrary normalization constant. The functions e^{iqx} satisfy the *orthogonality condition,*

$$\int_{-L/2}^{L/2} dx e^{i(q-q')x} = \frac{\sin[(q-q')L/2]}{[(q-q')/2]} = L\delta_{n,n'} = L\delta_{q-q',0} \,, \tag{2A.5}$$

where $\delta_{a,b}$ is the Kronecker delta ($\delta_{a,b} = 1$ if $a = b$ and $\delta_{a,b} = 0$ otherwise) and the *completeness condition,*

$$\sum_q e^{-iqx} = \lim_{N\to\infty} \sum_{N-1}^{N-1} e^{-i(2\pi n/L)x} = \lim_{N\to\infty} \frac{\sin[2\pi(N - 1/2)x/L]}{\sin(\pi x/L)} = L\delta(x) \,, \tag{2A.6}$$

where $\delta(x)$ is the Dirac delta that is zero except at $x = 0$ but whose integral over x is unity. Thus, for periodic boundary conditions,

$$\left.\begin{aligned} f(x) &= A\textstyle\sum_q e^{iqx} f(q), \\ f(q) &= \frac{1}{AL} \textstyle\int_{-L/2}^{L/2} e^{-iqx} f(x) dx. \end{aligned}\right\} \tag{2A.7}$$

To treat systems in the limit $L \to \infty$, one takes the continuum limit in which $q = (2\pi/L)n$ is treated as a continuous variable and

$$\sum_q \equiv \sum_n \Delta n = \frac{L}{2\pi} \sum_q \Delta q \to L \int_{-\infty}^{\infty} \frac{dq}{2\pi} \,, \tag{2A.8}$$

where $\Delta n = 1$. Thus Eqs. (2A.7) can be rewritten as

$$f(x) = AL \int_{-\infty}^{\infty} \frac{dq}{2\pi} e^{iqx} f(q) \overset{LA=1}{\to} \int_{-\infty}^{\infty} \frac{dq}{2\pi} e^{iqx} f(q), \tag{2A.9}$$

$$f(q) = \frac{1}{AL} \int_{-\infty}^{\infty} dx e^{-iqx} f(q) \overset{LA=1}{\to} \int_{-\infty}^{\infty} dx e^{-iqx} f(q) \,. \tag{2A.10}$$

The normalization constant A is often chosen to be equal to L^{-1} so that the factors LA and $(LA)^{-1}$ become unity as shown in the final form on the right hand side of Eqs. (2A.9) and (2A.10). Other choices, such as $A = L^{-1/2}$ so that $L^{-1/2}$ appears as a factor in both Eqs. (2A.9) and (2A.10), are also used. In the continuum limit, the orthogonality and completeness relations (2A.5) and (2A.6) become

$$\int_{-\infty}^{\infty} dx e^{i(q-q')x} = \lim_{L\to\infty} L\delta_{q-q',0} \equiv 2\pi\delta(q-q') \tag{2A.11}$$

and

$$\int_{-\infty}^{\infty} \frac{dq}{2\pi} e^{-iq(x-x')} = \delta(x-x'). \tag{2A.12}$$

The identification of $L\delta_{q,q'}$ with $2\pi\delta(q-q')$ can be seen from

$$\sum_q \delta_{q,0} = 1 = \frac{L}{2\pi}\int dq\delta_{q,0} \to \int_{-\infty}^{\infty} dq\delta(q)\,. \tag{2A.13}$$

2 *d dimensions*

The generalization of the above formulae to d dimensions is straightforward. Let $f(\mathbf{x})$ be a function of a d-component vector $\mathbf{x} = (x_1, x_2, ..., x_d)$ and impose periodic boundary conditions on each of the components of $\mathbf{x}$:

$$f(x_1, ..., x_i, ..., x_d) = f(x_1, ..., x_i + L_i, ..., x_d)\,, \quad i = 1, 2, ..., d\,. \tag{2A.14}$$

Then $f(\mathbf{x})$ can be expanded in a Fourier series similar to Eqs. (2A.7):

$$f(\mathbf{x}) = A\sum_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{x}} f(\mathbf{q}) \tag{2A.15}$$

$$f(\mathbf{q}) = \frac{1}{AV}\int d^d x e^{-i\mathbf{q}\cdot\mathbf{x}} f(\mathbf{x})\,, \tag{2A.16}$$

where $V = L_1 L_2 ... L_d$ and

$$\mathbf{q} = \left(\frac{2\pi}{L_1}n_1, \frac{2\pi}{L_2}n_2, ..., \frac{2\pi}{L_d}n_d\right)\,, \tag{2A.17}$$

where the coefficients n_i are integers. In the infinite volume limit, these relations become

$$f(\mathbf{x}) = A\sum_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{x}} f(\mathbf{q}) = AV\int \frac{d^d q}{(2\pi)^d} e^{i\mathbf{q}\cdot\mathbf{x}} f(\mathbf{q})$$

$$\overset{AV=1}{\to} \int \frac{d^d q}{(2\pi)^d} e^{i\mathbf{q}\cdot\mathbf{x}} f(\mathbf{q}) \tag{2A.18}$$

$$f(\mathbf{q}) = \frac{1}{AV}\int d^d x e^{-i\mathbf{q}\cdot\mathbf{x}} f(\mathbf{x}) \overset{AV=1}{\to} \int d^d x e^{-i\mathbf{q}\cdot\mathbf{x}} f(\mathbf{x}), \tag{2A.19}$$

where again the normalization factor A is often chosen to be equal to V^{-1} as indicated by the final form of these equations. It is understood that the $\mathbf{x}$- and $\mathbf{q}$-integrals in Eqs. (2A.18) and (2A.19) are over all space. Finally, in the infinite volume, continuum limit, the orthogonality and completeness conditions become

$$\int d^d x e^{i(\mathbf{q}-\mathbf{q}')\cdot\mathbf{x}} = V\delta_{\mathbf{q},\mathbf{q}'} = (2\pi)^d \delta^{(d)}(\mathbf{q}-\mathbf{q}') \tag{2A.20}$$

and

$$\int \frac{d^d q}{(2\pi)^d} e^{-i\mathbf{q}\cdot(\mathbf{x}-\mathbf{x}')} = \delta^{(d)}(\mathbf{x}-\mathbf{x}')\,, \tag{2A.21}$$

where $\delta^{(d)}(\mathbf{x})$ is a d-dimensional Dirac delta function.

3 Transforms on a lattice

Often one is interested in functions that are defined only at points on a regular periodic lattice rather than at all points in space. The Fourier transformation of these functions is the subject of this sub-section.

One-dimensional lattices Let f_l be a function of the integer l indexing the lattice site located at position $R_l = la$ of a one-dimensional lattice with lattice spacing a (see Sec. 2.5). The function f_l can be expanded in a discrete Fourier series

$$f_l = \sum_q \tilde{\psi}_q(l) f_q, \tag{2A.22}$$

where $\tilde{\psi}_q(l)$ satisfies the same boundary conditions as f_l. Again, we choose the periodic boundary condition,

$$f_l = f_{l+N}, \tag{2A.23}$$

where N is an integer. In this case, we can choose

$$\tilde{\psi}_q(l) = Ae^{iqR_l} = \tilde{\psi}_q(l+N)\,, \tag{2A.24}$$

where

$$q = \frac{2\pi}{Na} n, \tag{2A.25}$$

where n is an integer. Because R_l is an integral multiple of the lattice spacing a, the function $\tilde{\psi}_q(x)$ in Eq. (2A.23) is periodic in q as well as in l:

$$\tilde{\psi}_q(l) = \tilde{\psi}_{q+(2\pi)/a}\,(l). \tag{2A.26}$$

Thus all the functions $\tilde{\psi}_q(l)$ and $f(q)$ are completely characterized by q in the interval $[-\pi/a, \pi/a]$, i.e., by q in the first *Brillouin zone* (BZ) of the one-dimensional lattice.

The number of points in the first BZ is equal to the number of sites N in the direct lattice. This follows because the number of points in some region of space is equal to its "volume" divided by the "volume" per point. The volume of the first BZ of a one-dimensional lattice is $2\pi/a$ and volume per point is simply $\Delta q = (2\pi)/Na$, so that

$$\text{number of points in first BZ} = \frac{2\pi}{a}\frac{1}{\Delta q} = \frac{2\pi/a}{(2\pi)/Na} = N. \tag{2A.27}$$

The functions e^{iqR_l} satisfy an orthogonality condition similar to that of the functions e^{iqx}:

$$\sum_{l=0}^{N} e^{i(q-q')R_l} = N\delta_{q,q'} \overset{N\to\infty}{\to} \frac{2\pi}{a}\delta(q-q'), \tag{2A.28}$$

where Eq. (2A.11) with $L = Na$ was used to relate the Kronecker delta to the Dirac delta. The completeness condition is

$$\sum_{q\in \text{1st BZ}} e^{iqR_l} = \frac{1-e^{iNl}}{1-e^{il}} = N\delta_{l,0}. \tag{2A.29}$$

In the continuum limit, this equation becomes

$$a\int_{-\pi/a}^{\pi/a} \frac{dq}{2\pi} e^{iqR_l} = \delta_{l,0}. \tag{2A.30}$$

When the above results are combined, the lattice Fourier transforms can be written as

$$f_l \quad = \quad A\sum_{q\in \text{1st BZ}} e^{iqR_l} f_q \overset{N\to\infty}{\to} A(Na)\int_{-\pi/a}^{\pi/a} \frac{dq}{2\pi} e^{iqR_l} f_q$$

$$\overset{ANa=1}{\rightarrow} \quad \int_{-\pi/a}^{\pi/a} \frac{dq}{2\pi} e^{iqR_l} f_q \tag{2A.31}$$

$$f_q = \frac{1}{NA} \sum_l e^{-iqR_l} f_l \overset{ANa=1}{\rightarrow} a \sum_l e^{-iqR_l} f_l. \tag{2A.32}$$

Again, the choice of A is arbitrary. Often the choice $A = (1/Na)$ is made as shown on the far right hand side of these equations. In this case, $(NA)^{-1} = a$, and the sum over l in Eq. (2A.32) could be replaced by an integral over R_l in a spatial continuum limit.

d-dimensional lattices The generalization of lattice Fourier transforms to d-dimensional lattices is again straightforward. If $f_\mathbf{l}$ is a function of the lattice index $\mathbf{l}$ satisfying periodic boundary conditions, $f_\mathbf{l} = f_{\mathbf{l}+\mathbf{N}}$, where $\mathbf{N} = (N_1, N_2, ..., N_d)$, then

$$f_\mathbf{l} = \sum_\mathbf{q} \tilde{\psi}_\mathbf{q}(\mathbf{l}) f_\mathbf{q}, \tag{2A.33}$$

where, since $\mathbf{R}_{\mathbf{l}+\mathbf{N}} = \mathbf{R}_\mathbf{l} + \mathbf{R}_\mathbf{N}$,

$$\tilde{\psi}_\mathbf{q}(\mathbf{l}) = A e^{i\mathbf{q}\cdot\mathbf{R}_\mathbf{l}} \tag{2A.34}$$

with

$$\mathbf{q} = \left(\frac{2\pi}{N_1 a} n_1, \frac{2\pi}{N_2 a} n_2, ..., \frac{2\pi}{N_d a} n_d \right) . \tag{2A.35}$$

The restriction of $\mathbf{R}_\mathbf{l}$ to lattice points leads to

$$\tilde{\psi}_{\mathbf{q}+\mathbf{G}}(\mathbf{l}) = \tilde{\psi}_\mathbf{q}(\mathbf{l}) , \tag{2A.36}$$

where $\mathbf{G}$ is a reciprocal lattice vector. Thus, as in the one-dimensional case, only wave vectors $\mathbf{q}$ in the first Brillouin zone need be considered. The number of points in the first Brillouin zone is again equal to the number of points, $N = N_1 N_2 ... N_d$ in the lattice. The orthogonality and completeness conditions are now

$$\mathrm{v}_0 \sum_\mathbf{l} e^{i(\mathbf{q}-\mathbf{q}')\cdot\mathbf{R}_\mathbf{l}} = V \delta_{\mathbf{q},\mathbf{q}'} = (2\pi)^d \delta^{(d)}(\mathbf{q} - \mathbf{q}'), \tag{2A.37}$$

and

$$\frac{1}{N} \sum_\mathbf{q} e^{-i\mathbf{q}\cdot(\mathbf{R}_\mathbf{l} - \mathbf{R}_{\mathbf{l}'})} \rightarrow \mathrm{v}_0 \int \frac{d^d q}{(2\pi)^d} e^{-i\mathbf{q}\cdot(\mathbf{R}_\mathbf{l} - \mathbf{R}_{\mathbf{l}'})} = \delta_{\mathbf{l},\mathbf{l}'}, \tag{2A.38}$$

where $\mathrm{v}_0 = V/N$ is the volume of a unit cell and the q-integral is over the first BZ. The Fourier transform equations are

$$f_\mathbf{l} = A \sum_\mathbf{q} e^{i\mathbf{q}\cdot\mathbf{R}_\mathbf{l}} f_\mathbf{q} \rightarrow AV \int \frac{d^d q}{(2\pi)^d} e^{i\mathbf{q}\cdot\mathbf{R}_\mathbf{l}} f_\mathbf{q}$$

$$\overset{AV=1}{\rightarrow} \quad \int \frac{d^d q}{(2\pi)^d} e^{i\mathbf{q}\cdot\mathbf{R}_\mathbf{l}} \tag{2A.39}$$

and

$$f_\mathbf{q} = \frac{1}{NA} \sum_\mathbf{l} e^{-i\mathbf{q}\cdot\mathbf{R}_\mathbf{l}} f_\mathbf{l} \overset{AV=1}{\rightarrow} \mathrm{v}_0 \sum_\mathbf{l} e^{-i\mathbf{q}\cdot\mathbf{R}_\mathbf{l}} . \tag{2A.40}$$

Bibliography

FLUIDS

J.P. Boon and S. Yip, *Molecular Hydrodynamics* (McGraw-Hill, New York, 1980).

Jean Pierre Hansen, *Theory of Simple Liquids* (Academic Press, New York, 1986).

Appendix A(翻译)

2016 年 3 月 15 日

1 生成函数和关联函数

1.1 概率-配分函数

假定，我们研究的所有系统都放进一个热源温度可调的热浴[1]中，此时，平衡态的概率分布为

$$p_l = \frac{1}{Z} e^{-\frac{E_l}{kT}} \tag{1}$$

其中E_l 是态l 的能量，k为玻尔兹曼常数，T为温度，归一化因子

$$Z = \sum_l e^{-E_l/kT} \tag{2}$$

并且，在大多数情况下，它并不是一个可测量量。但它很重要，因为它可以求出响应系数，如：比热、磁化率等物理量。实际上，我们对$\ln Z$更有兴趣。定义

$$F \equiv -kT \ln Z, \quad (\Omega \equiv \ln Z) \tag{3}$$

F为系统的自由能。

对于伊辛模型

$$Z = \sum_{\{S_i = \pm 1\}} e^{\beta J \sum\limits_{<i,j>} S_i \cdot S_j + \beta \sum\limits_i B_i \cdot S_i}, \ \beta = 1/kT \tag{4}$$

$Z = Z(B_i, T)$是磁场和温度的函数。

注：对于10×10正方晶格的二维伊辛模型，系统态的个数（$S_1 = \pm 1, S_2 = \pm 1, \cdots, S_{100} = \pm 1$）是$2^{100}(\approx 10^{30})$，因此，即使对于尺寸很小的系统，求对上式强行求和也是不可行的，对于三维情况或者海森堡模型更加困难。

1.2 与热力学的联系

把自由能的统计学定义($F = -kTlnZ$)与热力学定义($F = U - TS$)相联系很重要。Z还可以写为：

$$Z = \sum_E \omega(E) e^{-\beta E} \tag{5}$$

[1]对于大尺寸系统，其自身就是一个热浴

其中，$\omega(E)$能量E的微观状态数，这些位型的熵是

$$S(E) = k\ln\omega(E) \tag{6}$$

因此

$$Z = \sum_E e^{-\beta(E-TS(E))} \tag{7}$$

E和$S(E)$一般都是广延量，因此

$$e^{-\beta(E-TS(E))} = e^{-\beta N f(E)} \tag{8}$$

这里N是自由度数，（不要和磁性系统中$\overrightarrow{S}$的维度相混淆），则$f(E) = N^{-1}(E - TS(E))$，$f(E)$是一个强度量。并且(8)式在使得$f(E)$最小的E值附近有一个尖峰($E = U$)。当$N \to \infty$ 时，求和式$\sum_E$ 中除了$E = U$ 之外，其所有项都可以忽略不计

$$Z \approx e^{-\beta(U-TS(U))}, (U =< E >) \tag{9}$$

$$F \approx U - TS(U), (N \to \infty) \ \ (\text{with } \frac{\partial F}{\partial E}|_{E=U} = 0) \tag{10}$$

这是自由能的热力学定义。

$$\frac{\partial F}{\partial E}|_{E=U} = 0 \quad \longrightarrow \quad \frac{\partial S(U)}{\partial U} = 1/T \tag{11}$$

例如，对流体来说

$$U = U(S, V, N) \quad (V\text{是系统的体积}) \tag{12}$$

$$dU = TdS - PdV + \mu dN \tag{13}$$

F是U的勒让德变换,由于

$$F = U - TS \tag{14}$$

$$dF = -SdT - PdV + \mu dN \tag{15}$$

$$F = F(T, V, N) \tag{16}$$

因此，内能是熵的函数，而自由能是温度的函数，当然，也可以对另外两个量(V或N)做勒让德变换定义自由能，我们会看到，这很重要。

1.3 伊辛模型的关联函数

在大多数系统中，自由能本身并不是一个可测量量，而系统的响应函数是可以测量的，它们描述系统对外部参量变化所做出的响应：比热是内能对温度变化的响应 $\frac{\partial U}{\partial T}$, 膨胀系数（流体） $\frac{\partial V}{\partial P}$, 磁化率（磁性系统） $\frac{\partial M}{\partial B}$，等等。有些关联函数（如下）实验上也很重要的[2]。

所有这些量，都可以表达成自由能F的导数，所以称F是这些量的生成函数，这就是它的有趣之处。为了证明这一点，考虑伊辛模型，其一个重要的物理量就是自发磁化强度

$$M = \sum_i < S_i > \tag{17}$$

[2]在粒子物理中，与关联函数类似的是格林函数，其直接与散射截面（S矩阵元）相联系。

它在研究伊辛模型的相变中扮演着一个很特别的角色，因为它随温度变化的行为表示$B=0$时系统相的特征，如图1所示,其中，T_c 是转变温度，因此把$\overrightarrow{M}$称为序参量。

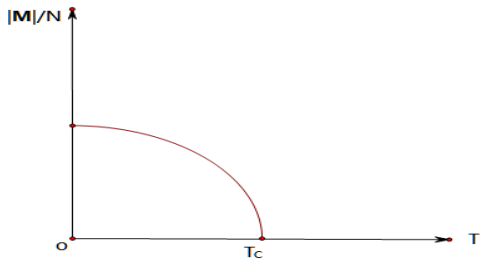


图 1:

当B均匀时，由(4)式可以得到：

$$\frac{\partial Z}{\partial B}=\beta\sum_{\{S_i\}}(\sum_i S_i)e^{-\beta H} \tag{18}$$

$$\overrightarrow{M}=\frac{1}{\beta Z}\frac{\partial Z}{\partial B}=-(\frac{\partial F}{\partial B})_T \tag{19}$$

对于一个磁系统，与(13)类似：

$$dF=-SdT-MdB \tag{20}$$

下面，重新定义$B\to\beta B$，则

$$Z=\sum_{\{S_i\}}\exp[\beta J\sum_{<i,j>}S_i\cdot S_j+\sum_i B_i\cdot S_i] \tag{21}$$

因此

$$M=(\frac{\partial\Omega}{\partial B})_T,\ \Omega=\ln Z \tag{22}$$

定义$\Gamma+\Omega=MB$, 这是对关于B的Ω做勒让德变换

$$d\Gamma=-MdB+MdB+BdM-\frac{\partial\Omega}{\partial T}|_B dT \tag{23}$$

$$\to\quad \Gamma=\Gamma[M,T] \tag{24}$$

Γ 是吉布斯自由能，它是磁化强度的函数，而不是磁场的函数。当然，在实际中，很容易加外磁场，而加一个磁化强度是很难的，但是在$N\to\infty$时，两者等价，这是因为在大尺寸下，吉布斯自由能的在平均值附近的涨落非常小：确定的外部磁场对应确定的磁化强度(等于$\sum_i S_i$的热力学平均值)，而$\sum_i S_i$)的涨落与自发磁化强度M相比非常之小。

使用Γ 而不是Ω 的好处是Γ是序参量的函数，而序参量正是物理上重要的量（而非B）。由(23) 可以推出：

$$B=\frac{\partial\Gamma}{\partial M}|_T,\qquad (tobecompared\ with\ M=\frac{\partial\Omega}{\partial B}|_T) \tag{25}$$

Γ 称作“吉布斯自由能”或“有效作用量”，也可以称作“顶点函数的生成泛函”和“单粒子不可约关联函数的产生泛函”，不同的叫法有不同的起因，视具体情况而定[3]。

[3]量子场论里面也有一个与Γ类似的基本物理量

分析出Z和Γ之间的关系是很重要的

$$Z = e^{-\Gamma + MB} \tag{26}$$

和下式相比

$$Z = \sum_{\{S_i\}} e^{-\beta H + B\sum\limits_i S_i} \tag{27}$$

根据如下替换，两个式子非常相似

$$\begin{aligned} \sum_i S_i &\to M \\ \beta H &\to \Gamma \end{aligned} \tag{28}$$

一旦做这样的替换，(26)和(27) 式的主要不同当然是(27)那样对所有位型求和$\sum\limits_{\{S_i\}}$，换句话说，考虑$\Gamma(M)$而不是$H(S_i)$, 相当于考虑了所有的涨落。我们后面将会看到，所谓的“平均场”近似，是在一个“最好的”位型附近，忽略所有涨落，即$\Gamma_{M.,F.} = H$. 因此Γ和H差异的正好体现了涨落的重要程度，我们可以这样说，Γ是完全的，精确的理论，而H 是忽略了所有涨落的近似理论。

关联函数有三种形式：

1. 定义关联函数

$$G^{(n)}(x_1, \cdots, x_n) \equiv < S(x_1) \cdots S(x_n) > \tag{29}$$

其中，x_i是D维系统第i个格点的坐标向量：$S_i = S(x_i), x_i = \begin{pmatrix} x_i^1 \\ \vdots \\ x_i^D \end{pmatrix}$，可以由$Z$ 求导得出下面这些公式：

$$\begin{aligned} m_i \equiv < S_i > &= \frac{1}{Z} \sum_{\{S_i\}} S_i e^{-\beta H + \sum\limits_k B_k S_k} \\ &= \frac{1}{Z} \frac{\partial Z}{\partial B_i} \end{aligned} \tag{30}$$

$$\begin{aligned} G^{(2)}(x_i, x_j) = < S_i \cdot S_j > &= \frac{1}{Z} \sum_{\{S_i\}} S_i \cdot S_j e^{-\beta H + \sum\limits_k B_k S_k} \\ &= \frac{1}{Z} \frac{\partial^2}{\partial B_i B_j} Z[B] \end{aligned} \tag{31}$$

$$G^{(n)}(x_{i1}, \cdots, x_{in}) = < S(x_{i1}) \cdots S(x_{in}) > = \frac{1}{Z} \frac{\partial^n}{\partial B_{i1} \cdots \partial B_{in}} Z[B] \tag{32}$$

$Z[B]$ 就是自旋关联函数的产生泛函。注：很多时候，我们对外场$B \to 0$时系统的关联函数更感兴趣，这时，我们要先假想一个“虚拟的”，一般的场B（与x有关）求导得出关联函数之后只需令$B = 0$即可。

$$< S_{i1} \cdots S_{in} > |_{B=0} = \frac{1}{Z} \frac{\partial^n}{\partial B_{i1} \cdots \partial B_{in}} Z[B]|_{B=0} \tag{33}$$

2. 定义厄塞尔函数(connected correlation function)：

$$G_c^{(n)}(x_{i1}, \cdots, x_{in}) \equiv < (S_{i1} - < S_{i1} >) \cdots (S_{in} - < S_{in} >) > \tag{34}$$

这是自旋在平均值附近涨落的关联函数。在$T < T_C$ 时，存在自发磁化强度，$< S_i >\neq 0$, 也因此了自旋之间存在了一种平庸的关联：

$$< S_iS_j > \xrightarrow[T\to 0]{} < S_i >< S_j > \tag{35}$$

所以，connected correlation 减掉了了不重要的部分

$$G_c^{(2)}(x_i, x_j) =< S_i \cdot S_j > - < S_i >< S_j > \xrightarrow{T\to 0} 0 \tag{36}$$

【作业1：证明(35)和(36)(一个很方便的办法是先证明$Z \approx e^{-\beta E_0}\ \ (\beta \to \infty)$)，其中$E_0$ 是基态能量。**】**

所有的厄塞尔函数(connected correlation)都可以由$\Omega(B)$ 得到，例如：

$$\begin{aligned} G_c^{(2)}(x_i, x_j) =& < S_i \cdot S_j > - < S_i >< S_j > \\ =& \frac{1}{Z}\frac{\partial^2 Z}{\partial B_i \partial B_j} - \frac{1}{Z^2}\frac{\partial Z}{\partial B_i}\frac{\partial Z}{\partial B_j} \\ =& \frac{\partial}{\partial B_j} \cdot (\frac{1}{Z}\frac{\partial Z}{\partial B_i}) \end{aligned} \tag{37}$$

由(30)和(22)得

$$G_c^{(2)}(x_i, x_j) = \frac{\partial m_i}{\partial B_j} \tag{38}$$

$$G_c^{(2)}(x_i, x_j) = \frac{\partial^2 \Omega}{\partial B_i \partial B_j} \tag{39}$$

因此，对任意n

$$G_c^{(n)}(x_{i1}, \cdots, x_{in}) = \frac{\partial^n \Omega}{\partial B_{i1} \cdots \partial B_{in}} \tag{40}$$

所以，把$\Omega(B)$ 称为关联函数的产生泛函。

(38)有一个很有趣的物理解释。磁化率定义为磁化强度对磁场变化的响应

$$\chi = \frac{\partial M}{\partial B}|_T \tag{41}$$

因此，$G_c^{(2)}(x_i, x_j)$是系统的局域磁化率：它表示系统的局域磁化强度m_i 在x_j 这一点对外部磁场变化的响应；同时也是磁场从传播子，因为它描述了磁场的变化如何从x_j传播到x_i.

3. 单粒子的不可约关联函数，Γ 是其生成泛函。首先，对一般的$m_i =< S_i >$ 定义Γ, 对一般场B中的$\Omega(B)$做勒让德变换可以得到

$$\Gamma(m) + \Omega(B) = \sum_i B_i m_i \tag{42}$$

做变换：

$$m_i = \frac{\partial \Omega}{\partial B_i}|_T \longrightarrow B_i = \frac{\partial \Gamma}{\partial m_i}|_T \tag{43}$$

我们可以定义：

$$\Gamma^{(n)}(x_{i1}, \cdots, x_{(in)}) \equiv \frac{\partial^n \Gamma}{\partial m_{i1}, \cdots, \partial m_{in}} \tag{44}$$

大多数时候，我们对$m_i = 0(\forall i)$时的函数感兴趣，Γ^n 称作“单粒子不可约关联函数的顶角函数”。

所有这些关联函数的定义都对应于其产生泛函的场展开，如Γ[4]

$$\Gamma = \sum_n \frac{1}{n!} \int d^D x_{i1} \cdots d^D x_{in}\ m(x_{i1}) \cdots m(x_{in}) \Gamma^{(n)}(x_{i1}, \cdots x_{in}) \tag{45}$$

所有这些关联函数都携带着系统相同的信息，并且可以根据其他关联函数(如：$G_c^{(n)}$)重新构建一种最感兴趣的关联函数(如：$\Gamma^{(n)}$))。

作业2：

- 证明：$\Gamma^{(2)}(x_i, x_j)$ 和$G_c^{(2)}(x_i, x_j)$ 其矩阵互为逆矩阵：$\sum_{x_j} \Gamma^{(2)}(x_i, x_j) \cdot G_c^{(2)}(x_j, x_k) = \delta_{i,k}$
- 证明：$G_c^{(x_e, x_p, x_q)} = \sum_{x_i, x_j, x_k} \Gamma^{(3)}(x_i, x_j, x_k) G^{(2)}(x_i, x_e) G^{(2)}(x_j, x_p) G^{(2)}(x_k, x_q)$
- 图示以上等式。

1.4 连续系统的关联函数

在欧几里得空间中考虑场论，将β 吸收进H里面

$$Z = \int D\phi e^{-H(\phi) + \int B\phi} \tag{46}$$

厄塞尔函数(connected关联函数)定义为：

$$G^{(n)}(x_1, \cdots, x_n) = < \phi(x_1) \cdots \phi(x_n) > = \frac{1}{Z} \frac{\delta^n Z}{\delta B(x_i) \cdots \delta B(x_n)} \tag{47}$$

$$C_x^{(n)}(x_1, \cdots, x_n) = < \prod_i^n (\phi(x_i) - < \phi(x_i) >) > = \frac{\delta^n \Omega}{\delta B(x_1) \cdots \delta B(x_n)} \tag{48}$$

由Ω 的泛函勒让德变换来定义Γ:

$$\Gamma(m) + \Omega(B) = \int d^D x\ B(x) m(x), \quad [m(x) = \frac{\delta \Omega}{\delta B(x)}] \tag{49}$$

仅当

$$\frac{\delta \Gamma}{\delta B(x)}|_m = -\frac{\delta \Omega}{\delta B(x)} + m(x) = 0 \tag{50}$$

时，Γ是$m(x)$ 的泛函。

要注意(50),它表示当$m(x)$有确定值时，Γ和B无关，例如

$$\Omega(B) = a \int B^n(x) d^D x \tag{51}$$

$$\Rightarrow m(x) = anB^{n-1}(x) \tag{52}$$

由(49)得到，

$$\Gamma = -\int d^D x\ aB^n(x) + \int d^D x\ B(x) m(x) \tag{53}$$

$$\frac{\delta \Gamma}{\delta B(x)}|_n = -anB^{n-1}(x) + m(x) \tag{54}$$

[4](45)是实际上是针对连续系统而不是晶格系统，因此是对x_i积分而不是对i求和。

由(52)可知，这一项完全消失。另一方面：

$$B(x) = (\frac{m(x)}{an})^{\frac{1}{n-1}} \tag{55}$$

$$\Rightarrow \Gamma(m) = \int d^D x \ (-\frac{a}{(an)^{\frac{n}{n-1}}} + \frac{1}{(an)^{\frac{1}{n-1}}}) m(x)^{\frac{n}{n-1}} \tag{56}$$

$$\Rightarrow \Gamma[m(B)] = (n-1)a \int d^D x \ B^n(x) \tag{57}$$

因此，

$$\frac{\delta \Gamma[m(B)]}{\delta B(x)} = n(n-1)aB^{n-1}(x) \tag{58}$$

也可以用下面的方式推导：

$$\begin{aligned} \frac{\delta \Gamma[m(B)]}{\delta B(x)} &= \int d^D y \ \frac{\delta \Gamma}{\delta m(y)} \frac{\delta m(y)}{\delta B(x)} \\ &= \int d^D y \ \frac{n-1}{n(an)^{\frac{1}{n-1}}} \frac{nm(y)^{\frac{1}{n-1}}}{n-1} an(n-1)B^{n-2}(y)\delta^0(x-y) \\ &= n(n-1)aB^{n-1}(x) \end{aligned} \tag{59}$$

现在考虑一般情况，

$$\frac{\delta \Gamma}{\delta m(x)} = -\int d^D y \ \frac{\delta \Omega}{\delta B(y)} \frac{\delta B(y)}{\delta m(x)} + \int_y \frac{\delta B(y)}{\delta m(x)} m(y) + B(x) \tag{60}$$

$$\Rightarrow \frac{\delta \Gamma}{\delta m(x)} = B(x) \tag{61}$$

定义

$$\Gamma^{(n)}(x_1, \cdots, x_n) = \frac{\delta^n \Gamma}{\delta m(x_1) \cdots \delta m(x_n)} \tag{62}$$

作业3：证明$\int d^D y \ \Gamma^{(2)}(x,y) G_c^{(2)}(y,z) = \delta^D(x-z)$. （与作业2.2类似）。

Scaling

1. Critical Point Exponents

Summary of various "singular behavior" at the critical point:

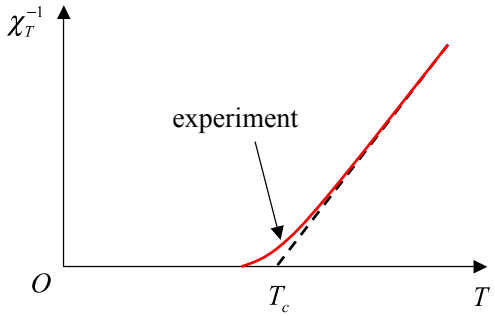
- order parameter vanishes
- magnetic susceptibility/compressibility diverges
- heat capacity jumps/diverges
- correlation length diverges
- …

The mean-field theories give specific predictions on how these quantities go to zero or infinity as $T \to T_c$, but these often do not agree with experiments.

Example: magnetic susceptibility

Mean-field:

$\chi_T \sim (T - T_c)^{-1}$



Better described by $\chi_T \sim \left(T - T_c\right)^{-\gamma}$, with $\gamma = 1.1$ to 1.4.

Definition of critical exponent:

If a function $f(t)$ diverges or goes to zero as $t \equiv \dfrac{T - T_c}{T_c} \to 0$,

one may define a critical exponent (assuming the limit exists)

$$\lambda \equiv \lim_{t \to 0} \frac{\ln f(t)}{\ln t}$$

and write, in the neighborhood of the critical point, $f(t) \sim t^{\lambda}$.

In general, the function $f(t) = At^{\lambda} + Bt^{\lambda_1} + \ldots$ $(\lambda_1 > \lambda)$ may contain corrections to the leading-order term. This may generate a lot of headache in experimental or numerical determination of the critical exponents.

Most commonly used critical exponents

Specific heat:	$C \sim t^{-\alpha}$
Order parameter:	$m \sim (-t)^{\beta}$
Susceptibility/compressibility:	$\chi \sim t^{-\gamma}$
Critical isotherm ($\varepsilon = 0$):	$h \sim m^{\delta}$
Correlation length:	$\xi \sim \lvert t \rvert^{-\nu}$
Pair correlation function ($t = 0$):	$c(r) \sim r^{-(d-2+\eta)}$

Scaling laws (critical exponent identities)

$$\alpha + 2\beta + \gamma = 2, \quad d\nu = 2 - \alpha$$
$$\beta(\delta - 1) = \gamma, \quad \gamma = (2 - \eta)\nu$$

2. Scaling hypothesis

Free energy per site: $g(t,h) = g_0(t,h) + g_s(t,h)$

At $h = 0$, $g_s(t,0) \sim |t|^{2-\alpha}$

"Scaling": $t \to bt,\ g_s \to b^{2-\alpha} g_s$

At $t = 0$, $g_s(0,h) \sim h^{1+1/\delta}$

"Scaling": $h \to bh,\ g_s \to b^{1+1/\delta} g_s$

Homogeneity hypothesis (Ben Widom, 1965):

$$g_s(t,h) = \lambda^{-d} g_s(\lambda^y t, \lambda^x h)$$

(x, y take values that are model specific)

With this assumption, we obtain,

$$m(t,h) = -\frac{\partial g}{\partial h} = \lambda^{-d+x} m(\lambda^y t, \lambda^x h)$$
$$c_h(t,h) = -T_c^{-1} \frac{\partial^2 g}{\partial t^2} = \lambda^{-d+2y} c_h(\lambda^y t, \lambda^x h)$$
$$\chi(t,h) = -\frac{\partial^2 g}{\partial h^2} = \lambda^{-d+2x} \chi(\lambda^y t, \lambda^x h)$$

Special cases:

i) $h = 0, \lambda = |t|^{-1/y}$; ii) $t = 0, \lambda = |h|^{-1/x}$

$$m(t,0) = (-t)^{(d-x)/y} m(-1,0)$$
$$c_h(t,0) = |t|^{(d/y)-2} c_h(\pm 1,0)$$
$$\chi(t,0) = |t|^{-(2x-d)/y} \chi(\pm,0)$$

$$m(0,h) = |h|^{(d/x)-1} m(0,\pm 1)$$

Hence,

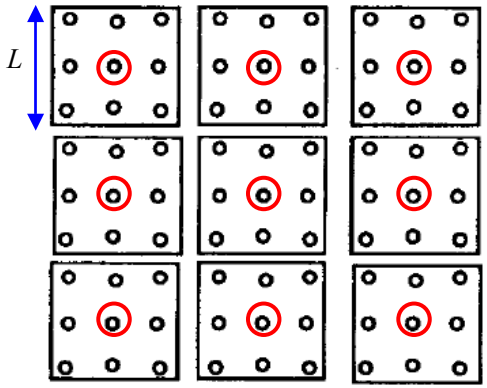
$$\alpha = 2 - (d/y)$$
$$\beta = (d-x)/y$$
$$\gamma = (2x-d)/y$$
$$\delta = x/(d-x)$$

$$\alpha + 2\beta + \gamma = 2$$
$$\beta(\delta - 1) = \gamma$$

But why?

3. Kadanoff block spins

For a historical account, read pages 10-18 of Leo Kadanoff's article *From Order to Chaos II*.



Block spin transformation (coarse-graining):

$$\sigma_i \to \tilde{\sigma}_J = \text{sign}\left(\sum_{i\in J}\sigma_i\right)$$

Known as the *majority rule*

i) The original and transformed system share the same critical point ⇒ At this special point, the system becomes scale invariant!

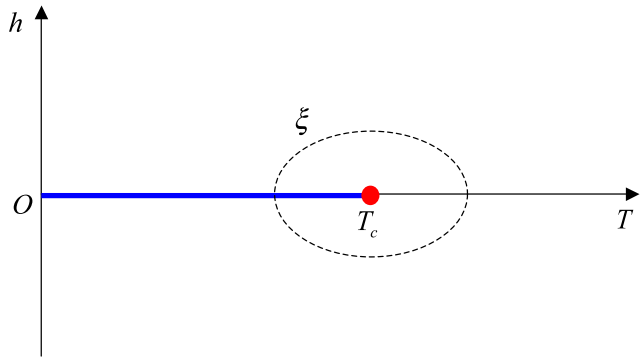
ii) Away from the critical point,

$$t \to \tilde{t} = tL^y, \quad h \to \tilde{h} = hL^x$$ (key assumption)

iii) Free energy per site: $g_s \to \tilde{g}_s = g_s(\tilde{t},\tilde{h}) = L^d g_s(t,h)$, or

$$g_s(t,h) = L^{-d} g_s(L^y t, L^x h)!$$

In addition, Kadanoff's scaling analysis yields a prediction of the spatial structure of critical fluctuations.



Scale of physical quantities in the critical region

Correlation length: ξ

(Free) Energy: kT_c

Upon the block transformation, $\xi(t,h) \to \tilde{\xi} = \xi(\tilde{t},\tilde{h}) = \xi(t,h)/L$

Hence: $\xi(t,h) = L\xi(L^y t, L^x h)$

Divergence of ξ

$h = 0$: $\xi \sim |t|^{-\nu}$, $\nu = 1/y$

$T = T_c$: $\xi \sim |h|^{-1/x}$

Hyperscaling relation: $\alpha = 2 - d/y = 2 - d\nu$

Evaluation and summary: testing the scaling laws

1. Critical point exponents are difficult to measure experimentally.

2. Experimental physicists have worked extremely hard to achieve the unachievable, and along the way improved their instruments as well as our understanding of Nature.

3. The scientific process of verifying a hypothesis is often biased despite the presumed honesty and integrity shown by most scientists in their investigations.

4. The scaling theory has been a tremendeous success when compared with experimental data. It is now part of the "established wisdom" of the scientific community. It has been applied to many different areas of science in addition to the equilibrium critical phenomena where the ideas were first developed.

5. Beyond the scaling theory, the renormalization group ideas were developed and implemented. The mean-field theories were shown to give a correct description above four dimensions, but fluctuations "renormalize" parameters in the Landau free energy functional as one

goes to larger and larger scales, yielding different set of values for the exponents.

6. With the availability of powerful computers, predictions of the scaling theory have been verified to great precision for many different types of phase transitions, and universality classes have been identified.

7. We have witnessed a beautiful example of scientific research in modern history. From van der Waals to Ken Wilson, it took a century for the correct ideas to be developed and subtleties of Nature understood.

The renormalization group analysis of connecting orbits

Yueheng Lan
Department of Physics
Beijing University of Posts and Telecommunications

March, 2016

Main contents

Main contents

Main contents

Main contents

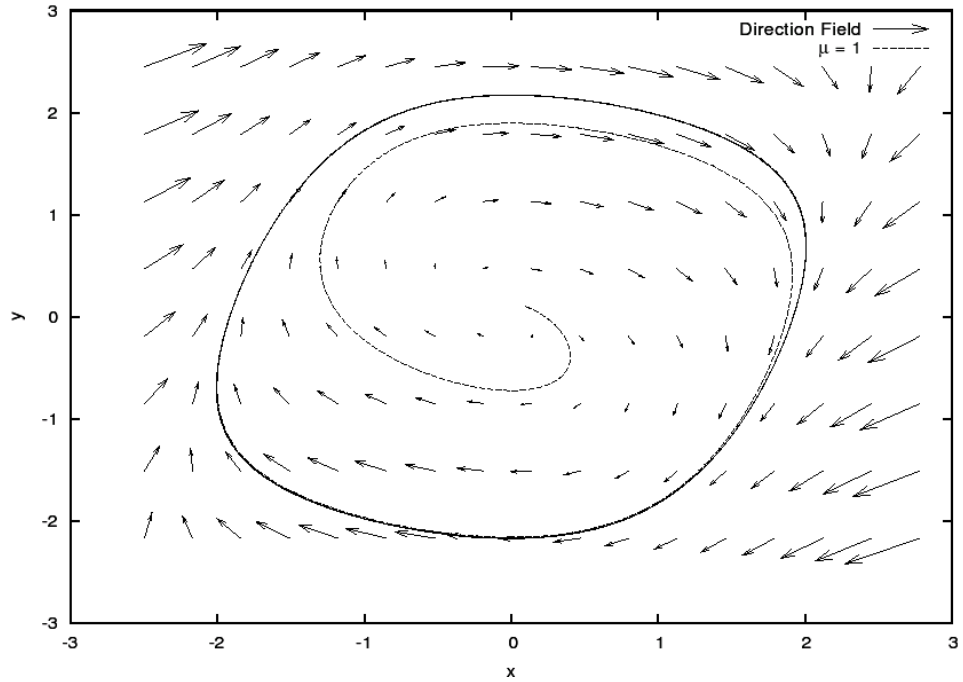
Main contents

Nonlinear dynamics and phase space

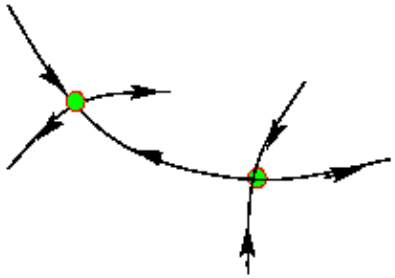
- State and dynamics

$$\begin{aligned} \dot{x}_1 &= f_1(x_1, x_2, \cdots, x_n) \\ \dot{x}_2 &= f_2(x_1, x_2, \cdots, x_n) \\ \cdots &= \cdots \\ \dot{x}_n &= f_n(x_1, x_2, \cdots, x_n) \end{aligned}$$

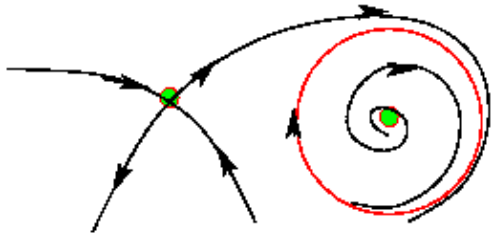
- The phase space - a geometric representation
- Vector field and trajectories
- Invariant set and organization of trajectories



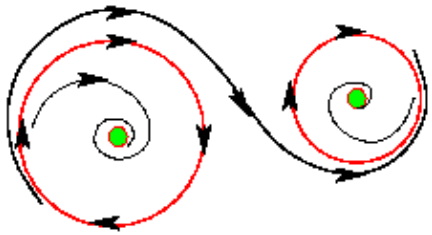
Connections



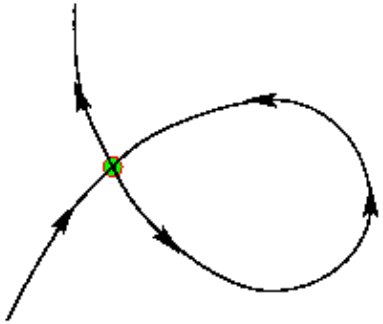
(a)



(b)

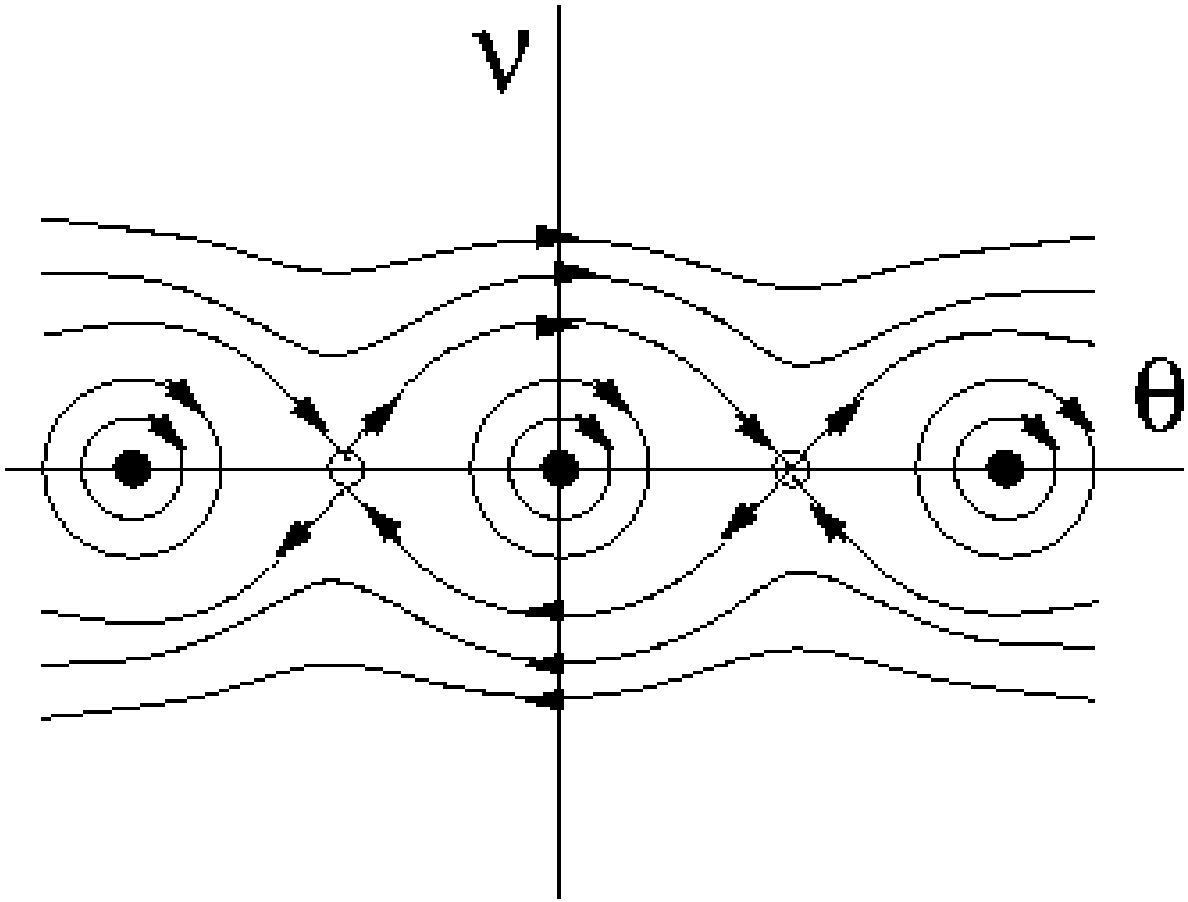


(c)

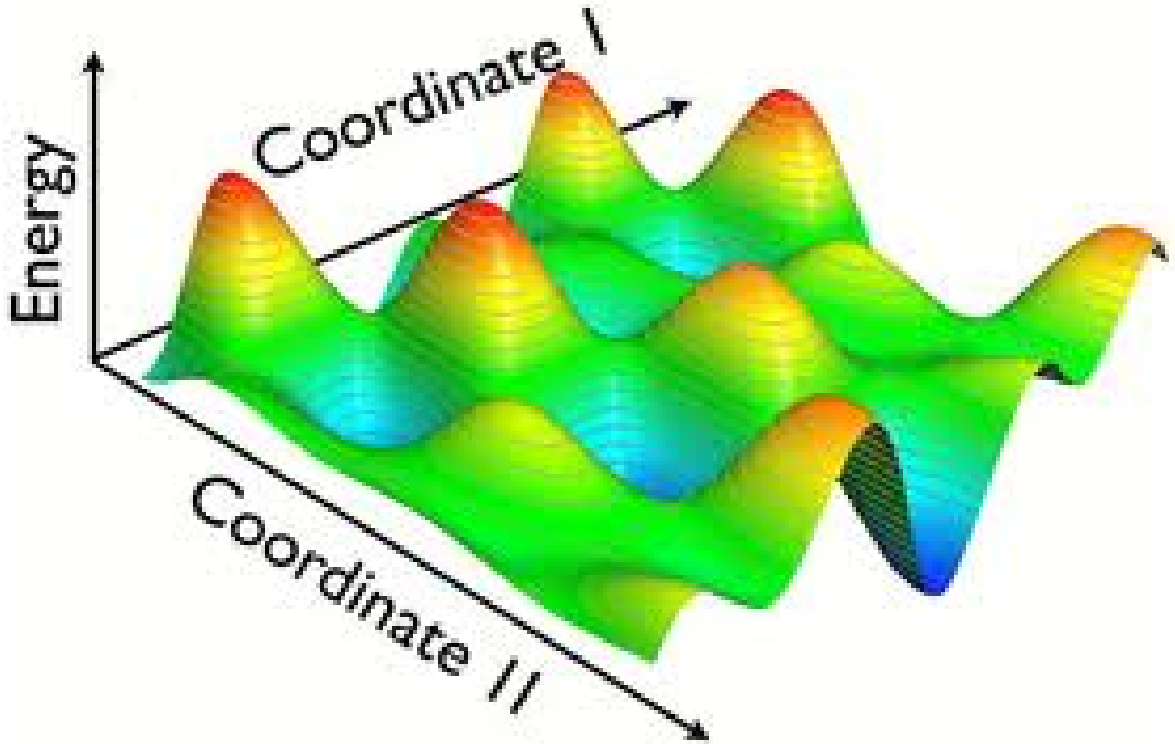


(d)

Pendulum orbits



State transition in chemical reactions



Computation of heteroclinic connections

- Exact analytic solutions under certain conditions: nonlinear integrable systems, partially integrable systems.
- Asymptotic methods for analytic approximations: local stability analysis plus interpolation or matching.
- Numerical methods: two-point boundary problem; shooting method; relaxation (variational) method.
- Challenge:
  (1) Need to know the existence and **both** ends;
  (2) Need to know the local behavior near **two** ends;
  (3) Hard to represent the dynamics on the connection;
  (4) Hard to derive analytic expressions.

Computation of heteroclinic connections

- Exact analytic solutions under certain conditions: nonlinear integrable systems, partially integrable systems.
- Asymptotic methods for analytic approximations: local stability analysis plus interpolation or matching.
- Numerical methods: two-point boundary problem; shooting method; relaxation (variational) method.
- Challenge:
  (1) Need to know the existence and **both** ends;
  (2) Need to know the local behavior near **two** ends;
  (3) Hard to represent the dynamics on the connection;
  (4) Hard to derive analytic expressions.

Computation of heteroclinic connections

- Exact analytic solutions under certain conditions: nonlinear integrable systems, partially integrable systems.
- Asymptotic methods for analytic approximations: local stability analysis plus interpolation or matching.
- Numerical methods: two-point boundary problem; shooting method; relaxation (variational) method.
- Challenge:
  (1) Need to know the existence and **both** ends;
  (2) Need to know the local behavior near **two** ends;
  (3) Hard to represent the dynamics on the connection;
  (4) Hard to derive analytic expressions.

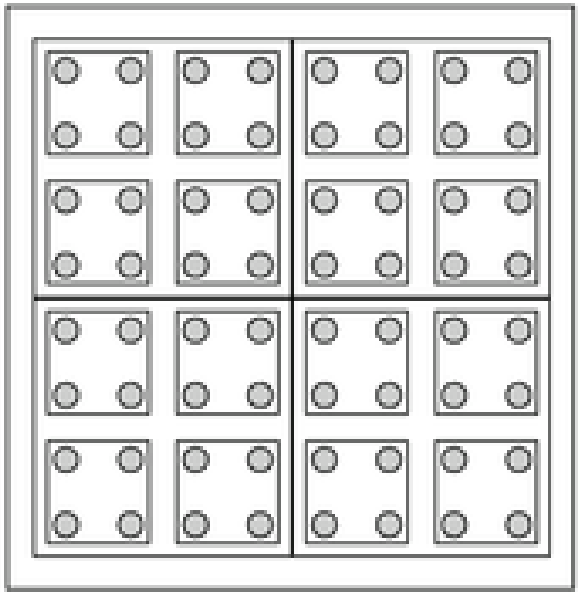
Computation of heteroclinic connections

- Exact analytic solutions under certain conditions: nonlinear integrable systems, partially integrable systems.
- Asymptotic methods for analytic approximations: local stability analysis plus interpolation or matching.
- Numerical methods: two-point boundary problem; shooting method; relaxation (variational) method.
- Challenge:
  (1) Need to know the existence and **both** ends;
  (2) Need to know the local behavior near **two** ends;
  (3) Hard to represent the dynamics on the connection;
  (4) Hard to derive analytic expressions.

Main contents

Renormalization group in physics

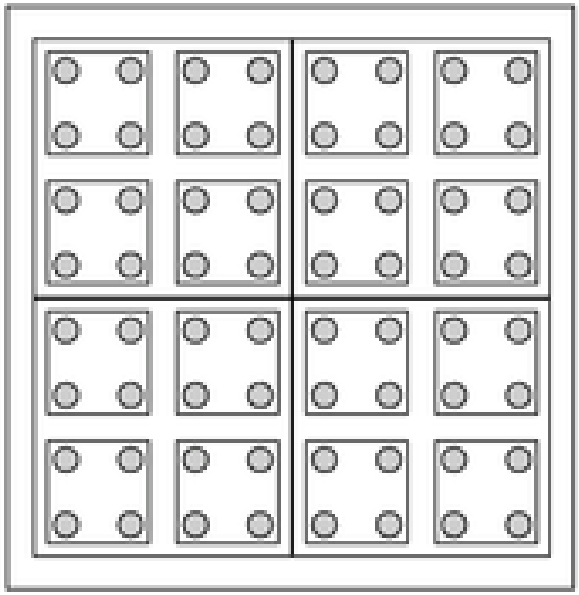
- RG investigates changes of physical laws at different scales.
- RG and scale invariance: a renormalizable system at one scale consists of self-similar copies of itself at a smaller scale, with convergent coupling parameters when scaled up.
- In statistical physics: block spin; In quantum physics: renormalization equation $\partial g/\partial \ln \mu = \beta(g)$; In nonlinear dynamics: the universal route to chaos.



Block spin renormalization group for a spin system described by $H(T, J):(T, J) \to (T', J')$, (T'', J''). **Resummation**

Renormalization group in physics

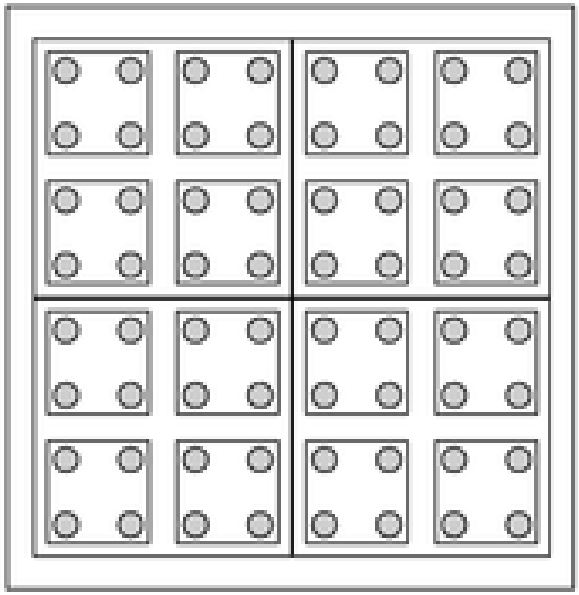
- RG investigates changes of physical laws at different scales.
- RG and scale invariance: a renormalizable system at one scale consists of self-similar copies of itself at a smaller scale, with convergent coupling parameters when scaled up.
- In statistical physics: block spin; In quantum physics: renormalization equation $\partial g/\partial \ln \mu = \beta(g)$; In nonlinear dynamics: the universal route to chaos.



Block spin renormalization group for a spin system described by $H(T, J)$:$(T, J) \to (T', J')$ (T'', J''). **Resummation**

Renormalization group in physics

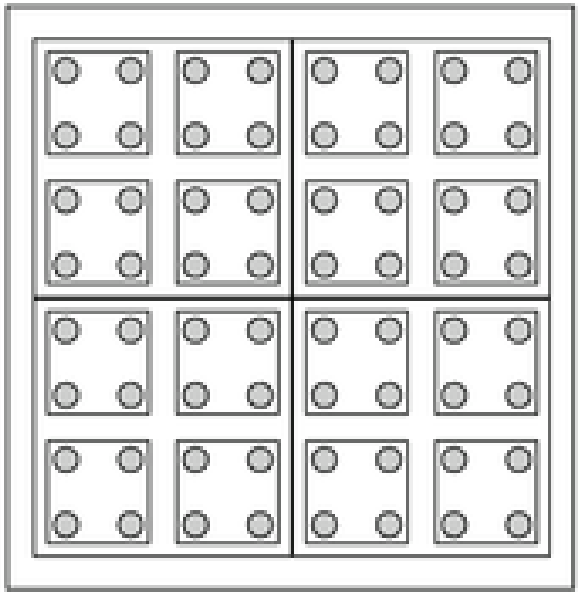
- RG investigates changes of physical laws at different scales.
- RG and scale invariance: a renormalizable system at one scale consists of self-similar copies of itself at a smaller scale, with convergent coupling parameters when scaled up.
- In statistical physics: block spin; In quantum physics: renormalization equation $\partial g/\partial \ln \mu = \beta(g)$; In nonlinear dynamics: the universal route to chaos.



Block spin renormalization group for a spin system described by $H(T, J)$:$(T, J) \to (T', J') \to (T'', J'')$. **Resummation**

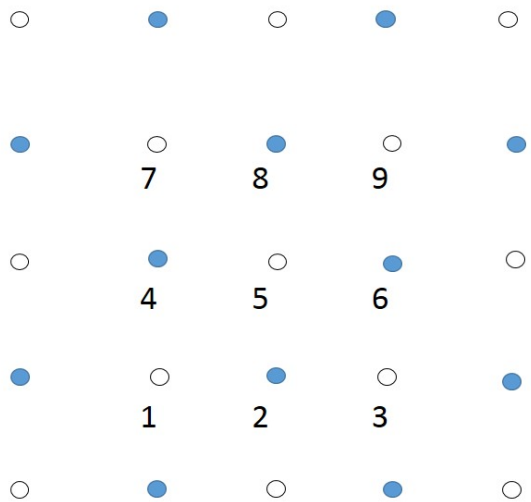
Renormalization group in physics

- RG investigates changes of physical laws at different scales.
- RG and scale invariance: a renormalizable system at one scale consists of self-similar copies of itself at a smaller scale, with convergent coupling parameters when scaled up.
- In statistical physics: block spin; In quantum physics: renormalization equation $\partial g/\partial \ln \mu = \beta(g)$; In nonlinear dynamics: the universal route to chaos.



Block spin renormalization group for a spin system described by $H(T, J)$:$(T, J) \to (T', J') \to (T'', J'')$. **Resummation?**

2D ising model as an example



One-step decimation is possible if
$K_{01} = \ln 2 + \frac{1}{2}\ln\cosh 2K + \frac{1}{8}\ln\cosh 4K$, $K_1 = \frac{1}{4}\ln\cosh 4K$,
$L_1 = \frac{1}{8}\ln\cosh 4K$,$M_1 = \frac{1}{8}\ln\cosh 4K - \frac{1}{2}\ln\cosh 2K$.

2D ising model as an example (continued)

- The partition function is

$$Q_N = \sum_{\{\sigma_i\}} \exp(\sum_{n.n.} K\sigma_i\sigma_j) \ (K = \beta J)$$

Let's carry out the summation over spin σ_5

$$\begin{aligned} \sum_{\sigma_5=\pm 1} & \quad \Pi_{n.n.}...e^{K\sigma_5(\sigma_2+\sigma_4+\sigma_6+\sigma_8)}... \\ = & \quad \Pi_{n.n.}...2\cosh K(\sigma_2+\sigma_4+\sigma_6+\sigma_8)... \end{aligned}$$

- It is only possible that

$$2\cosh K(\sigma_2+\sigma_4+\sigma_6+\sigma_8) = \exp(K_{01}+L_1(\sigma_2\sigma_8+\sigma_4\sigma_6))$$
$$\exp(\frac{1}{2}K_1(\sigma_2\sigma_4+\sigma_2\sigma_6+\sigma_4\sigma_8+\sigma_6\sigma_8)+M_1\sigma_2\sigma_8\sigma_4\sigma_6)$$

2D ising model as an example (continued)

- The partition function is

$$Q_N = \sum_{\{\sigma_i\}} \exp(\sum_{n.n.} K\sigma_i\sigma_j) \ (K = \beta J)$$

Let's carry out the summation over spin σ_5

$$\begin{aligned} \sum_{\sigma_5=\pm 1} & \quad \Pi_{n.n.}...e^{K\sigma_5(\sigma_2+\sigma_4+\sigma_6+\sigma_8)}... \\ = & \quad \Pi_{n.n.}...2\cosh K(\sigma_2+\sigma_4+\sigma_6+\sigma_8)... \end{aligned}$$

- It is only possible that

$$2\cosh K(\sigma_2+\sigma_4+\sigma_6+\sigma_8) = \exp(K_{01}+L_1(\sigma_2\sigma_8+\sigma_4\sigma_6))$$
$$\exp(\frac{1}{2}K_1(\sigma_2\sigma_4+\sigma_2\sigma_6+\sigma_4\sigma_8+\sigma_6\sigma_8)+M_1\sigma_2\sigma_8\sigma_4\sigma_6)$$

General formulation

- With the coupling constants $K_1, K_2, \cdots$, the free energy is

$$\exp(-\beta A) = \sum_{\{\sigma_i\}} \exp[-\beta H(\{\sigma_i\}, \{K_\alpha\})] .$$

The scale change will induce

$$N_1 = \ell^{-d} N , \xi_1 = \ell^{-1} \xi$$

and the change of the coupling parameters

$$\exp(-\beta A) = \exp(N_1 K_{01}) \sum_{\sigma_{1j}} \exp(-\beta H(\{\sigma_{1i}\}, \{K_{1\alpha}\})) .$$

- The energy per spin is

$$f(\{K_\alpha\}) = \ell^{-d}[-K_{01} + f(\{K_{1\alpha}\})] .$$

General formulation

- With the coupling constants $K_1, K_2, \cdots$, the free energy is

$$\exp(-\beta A) = \sum_{\{\sigma_i\}} \exp[-\beta H(\{\sigma_i\}, \{K_\alpha\})] .$$

The scale change will induce

$$N_1 = \ell^{-d} N , \xi_1 = \ell^{-1} \xi$$

and the change of the coupling parameters

$$\exp(-\beta A) = \exp(N_1 K_{01}) \sum_{\sigma_{1j}} \exp(-\beta H(\{\sigma_{1i}\}, \{K_{1\alpha}\})) .$$

- The energy per spin is

$$f(\{K_\alpha\}) = \ell^{-d}[-K_{01} + f(\{K_{1\alpha}\})] .$$

The RG flow

- For the fixed point K^*, $K^* = \mathscr{R}_\ell(K^*)$. For a point in the neighborhood $K = K^* + k$, the RG equation can be linearized

$$K' = \mathscr{R}_\ell(K) = K^* + \nabla\mathscr{R}_\ell k\,.$$

  The eigenvalues λ_i and eigenvectors u_i of the linearization matrix $\mathscr{A}_\ell = \nabla\mathscr{R}_\ell(K^*)$ determines the critical behavior of the system.(relevant $\lambda_i > 1$,irrelevant $\lambda_i < 1$, marginal $\lambda_i = 1$)

- After n repeats of the transformation,

$$\begin{aligned} \xi(u_1\,, u_2\,, ...) &= \ell^n \xi(\lambda_1^n u_1\,, \lambda_2^n u_2, ...) \\ f_s(u_1\,, u_2\,, ...) &= \ell^{-nd} f_s(\lambda_1^n U_1, \lambda_2^n u_2, ...)\,. \end{aligned}$$

  Note that $\lambda_i = \ell^{y_i}$. We may obtain all the scaling relations

The RG flow

- For the fixed point K^*, $K^* = \mathscr{R}_\ell(K^*)$. For a point in the neighborhood $K = K^* + k$, the RG equation can be linearized

$$K' = \mathscr{R}_\ell(K) = K^* + \nabla\mathscr{R}_\ell k\,.$$

  The eigenvalues λ_i and eigenvectors u_i of the linearization matrix $\mathscr{A}_\ell = \nabla\mathscr{R}_\ell(K^*)$ determines the critical behavior of the system.(relevant $\lambda_i > 1$,irrelevant $\lambda_i < 1$, marginal $\lambda_i = 1$)

- After n repeats of the transformation,

$$\begin{aligned}\xi(u_1\,, u_2\,, ...) &= \ell^n \xi(\lambda_1^n u_1\,, \lambda_2^n u_2, ...)\\ f_s(u_1\,, u_2\,, ...) &= \ell^{-nd} f_s(\lambda_1^n U_1, \lambda_2^n u_2, ...)\,.\end{aligned}$$

  Note that $\lambda_i = \ell^{y_i}$. We may obtain all the scaling relations.

Main contents

Series expansion of a differential equation

- Suppose that we have a set of n-dimensional ODEs

$$\dot{\mathbf{x}} = L\mathbf{x} + \epsilon N(\mathbf{x})$$

- We may make the expansion

$$\mathbf{x} = \mathbf{u}_0 + \epsilon \mathbf{u}_1 + \epsilon^2 \mathbf{u}_2 + \cdots$$

which results in

$$\begin{aligned} \dot{\mathbf{u}}_0 &= L\mathbf{u}_0 \\ \dot{\mathbf{u}}_1 &= L\mathbf{u}_1 + N(\mathbf{u}_0) \\ \dot{\mathbf{u}}_2 &= L\mathbf{u}_2 + N_2(\mathbf{u}_0, \mathbf{u}_1) \\ &\cdots \end{aligned}$$

Series expansion of a differential equation

- Suppose that we have a set of n-dimensional ODEs

$$\dot{\mathbf{x}} = L\mathbf{x} + \epsilon N(\mathbf{x})$$

- We may make the expansion

$$\mathbf{x} = \mathbf{u}_0 + \epsilon \mathbf{u}_1 + \epsilon^2 \mathbf{u}_2 + \cdots$$

which results in

$$\begin{aligned}
\dot{\mathbf{u}}_0 &= L\mathbf{u}_0 \\
\dot{\mathbf{u}}_1 &= L\mathbf{u}_1 + N(\mathbf{u}_0) \\
\dot{\mathbf{u}}_2 &= L\mathbf{u}_2 + N_2(\mathbf{u}_0, \mathbf{u}_1) \\
&\cdots
\end{aligned}$$

From the naive solution to the RG equation

- This series of equations can be solved as

$$\begin{aligned}
\mathbf{u}_0(t,t_0) &= e^{L(t-t_0)}\mathbf{A}(t_0) \\
\mathbf{u}_1(t,t_0) &= e^{L(t-t_0)}\int_{t_0}^{t} e^{-L(\tau-t_0)}N(e^{L(t-t_0)}\mathbf{A})d\tau \\
\mathbf{u}_2(t,t_0) &= e^{L(t-t_0)}\int_{t_0}^{t} e^{-L(\tau-t_0)}N_2(e^{L(t-t_0)}\mathbf{A},\mathbf{u}_1(t,t_0))d\tau\ .
\end{aligned}$$

- The series expansion gives $\mathbf{x}=\tilde{\mathbf{x}}(t;t_0,\mathbf{A}(t_0))$.
- The RG equation is a set of equations for $d\mathbf{A}(t_0)/dt_0$ derived from

$$\frac{d\tilde{\mathbf{x}}(t;t_0,\mathbf{A}(t_0))}{dt_0}|_{t=t_0}=0$$

One simple examle

- Consider the simple example

$$\dot{x} = y\,,\ \dot{y} = -x\,,$$

which can be solved exactly with

$$x = R\sin(t - t_0 + \theta)\,,\ y = R\cos(t - t_0 + \theta)\,,$$

where $R = R(t_0)\,, \theta = \theta(t_0)$ specify the initial condition.

- In phase space, orbits of the equation are circles with radius R and azimuth angle θ.
- The RG equation derived from

$$\frac{\partial x(t; R(t_0), \theta(t_0), t_0)}{\partial t_0}|_{t=t_0} = 0, \frac{\partial y(t; R(t_0), \theta(t_0), t_0)}{\partial t_0}|_{t=t_0} = 0$$

is $dR(t_0)/dt_0 = 0, d\theta(t_0)/dt_0 = 1$ as expected.

One simple examle

- Consider the simple example

$$\dot{x} = y\,,\ \dot{y} = -x\,,$$

which can be solved exactly with

$$x = R\sin(t - t_0 + \theta)\,,\ y = R\cos(t - t_0 + \theta)\,,$$

where $R = R(t_0)\,, \theta = \theta(t_0)$ specify the initial condition.

- In phase space, orbits of the equation are circles with radius R and azimuth angle θ.
- The RG equation derived from

$$\frac{\partial x(t; R(t_0), \theta(t_0), t_0)}{\partial t_0}|_{t=t_0} = 0, \frac{\partial y(t; R(t_0), \theta(t_0), t_0)}{\partial t_0}|_{t=t_0} = 0$$

is $dR(t_0)/dt_0 = 0, d\theta(t_0)/dt_0 = 1$ as expected.

One simple examle

- Consider the simple example

$$\dot{x} = y\,,\ \dot{y} = -x\,,$$

which can be solved exactly with

$$x = R\sin(t - t_0 + \theta)\,,\ y = R\cos(t - t_0 + \theta)\,,$$

where $R = R(t_0)\,, \theta = \theta(t_0)$ specify the initial condition.

- In phase space, orbits of the equation are circles with radius R and azimuth angle θ.
- The RG equation derived from

$$\frac{\partial x(t; R(t_0), \theta(t_0), t_0)}{\partial t_0}|_{t=t_0} = 0, \frac{\partial y(t; R(t_0), \theta(t_0), t_0)}{\partial t_0}|_{t=t_0} = 0$$

is $dR(t_0)/dt_0 = 0, d\theta(t_0)/dt_0 = 1$ as expected.

The RG analysis as a coordinate transformation

- Hamiltonian dynamics: action-angle variables. For a harmonic oscillator $H = 1/2p^2 + 1/2q^2 = I$.
- In a general nonlinear dynamical system,

$$\dot{x} = f(x)$$

which has the general solution $x(t) = \phi(t; A_0(t_0), t_0)$. The equation

$$\frac{\partial \phi(t; A_0(t_0), t_0)}{\partial t_0}|_{t=t_0} = 0$$

gives an equation for $dA_0(t_0)/dt_0$, which governs the evolution of the new coordinates A_0.

- The RG analysis is equivalent to a coordinate transformation in this sense, but often associated with approximations in the nonlinear case.

The RG analysis as a coordinate transformation

- Hamiltonian dynamics: action-angle variables. For a harmonic oscillator $H = 1/2p^2 + 1/2q^2 = I$.
- In a general nonlinear dynamical system,

$$\dot{x} = f(x)$$

which has the general solution $x(t) = \phi(t; A_0(t_0), t_0)$. The equation

$$\frac{\partial \phi(t; A_0(t_0), t_0)}{\partial t_0}|_{t=t_0} = 0$$

gives an equation for $dA_0(t_0)/dt_0$, which governs the evolution of the new coordinates A_0.

- The RG analysis is equivalent to a coordinate transformation in this sense, but often associated with approximations in the nonlinear case.

The RG analysis as a coordinate transformation

- Hamiltonian dynamics: action-angle variables. For a harmonic oscillator $H = 1/2p^2 + 1/2q^2 = I$.
- In a general nonlinear dynamical system,

$$\dot{x} = f(x)$$

which has the general solution $x(t) = \phi(t; A_0(t_0), t_0)$. The equation

$$\frac{\partial \phi(t; A_0(t_0), t_0)}{\partial t_0}|_{t=t_0} = 0$$

gives an equation for $dA_0(t_0)/dt_0$, which governs the evolution of the new coordinates A_0.

- The RG analysis is equivalent to a coordinate transformation in this sense, but often associated with approximations in the nonlinear case.

Further development of the RG analysis

- It can be used for nonlinear partial equations and is able to derive the phase or amplitude equations.
- The invariance condition has been extended to the analysis of maps.
- It is also used to determine the center manifold near a bifurcation point.
- Problem: for the dynamics on a submanifold, the n invariance equations contain less than n unknowns!?

Further development of the RG analysis

- It can be used for nonlinear partial equations and is able to derive the phase or amplitude equations.
- The invariance condition has been extended to the analysis of maps.
- It is also used to determine the center manifold near a bifurcation point.
- Problem: for the dynamics on a submanifold, the n invariance equations contain less than n unknowns!?

Further development of the RG analysis

- It can be used for nonlinear partial equations and is able to derive the phase or amplitude equations.
- The invariance condition has been extended to the analysis of maps.
- It is also used to determine the center manifold near a bifurcation point.
- Problem: for the dynamics on a submanifold, the n invariance equations contain less than n unknowns!?

Further development of the RG analysis

- It can be used for nonlinear partial equations and is able to derive the phase or amplitude equations.
- The invariance condition has been extended to the analysis of maps.
- It is also used to determine the center manifold near a bifurcation point.
- Problem: for the dynamics on a submanifold, the n invariance equations contain less than n unknowns!?

Dynamics on a submanifold

- The initial vector $\mathbf{A}$ should be taken as $\mathbf{A} = (A_1\,, 0\,, 0\,, \cdots\,, 0)^t$.
- The i-th component of $\mathbf{u}_1$ can be computed as

$$u_{1,i}(t, t_0) = e^{\lambda_i (t-t_0)} \int^t e^{-\lambda_i (\tau - t_0)} N(e^{L(t-t_0)}\mathbf{A}) d\tau\,,$$

  where $\int^t$ denotes integration without constant term.
- The first component

$$\frac{d\tilde{x}_1(t; t_0, A_1(t_0))}{dt_0}|_{t=t_0} = 0 \qquad (1)$$

  is enough to derive the RG equation for $dA_1(t_0)/dt_0$, which also satisfies other component equations.

Dynamics on a submanifold

- The initial vector $\mathbf{A}$ should be taken as $\mathbf{A} = (A_1\,, 0\,, 0\,, \cdots\,, 0)^t$.
- The i-th component of $\mathbf{u}_1$ can be computed as

$$u_{1,i}(t, t_0) = e^{\lambda_i(t-t_0)} \int^t e^{-\lambda_i(\tau-t_0)} N(e^{L(t-t_0)}\mathbf{A}) d\tau\,,$$

where $\int^t$ denotes integration without constant term.

- The first component

$$\frac{d\tilde{x}_1(t; t_0, A_1(t_0))}{dt_0}|_{t=t_0} = 0 \qquad (1)$$

is enough to derive the RG equation for $dA_1(t_0)/dt_0$, which also satisfies other component equations.

Dynamics on a submanifold

- The initial vector $\mathbf{A}$ should be taken as $\mathbf{A} = (A_1\,, 0\,, 0\,, \cdots\,, 0)^t$.
- The i-th component of $\mathbf{u}_1$ can be computed as

$$u_{1,i}(t, t_0) = e^{\lambda_i(t-t_0)} \int^t e^{-\lambda_i(\tau-t_0)} N(e^{L(t-t_0)}\mathbf{A}) d\tau\,,$$

  where $\int^t$ denotes integration without constant term.
- The first component

$$\frac{d\tilde{x}_1(t; t_0, A_1(t_0))}{dt_0}|_{t=t_0} = 0 \qquad (1)$$

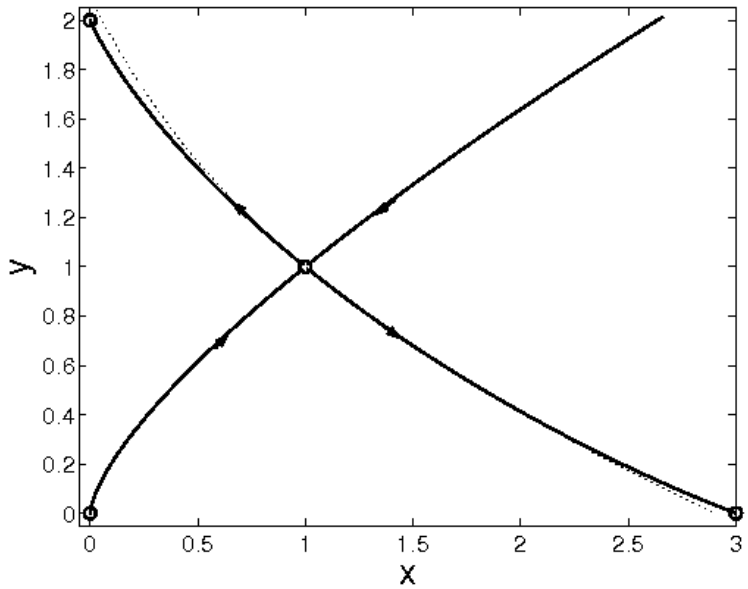
  is enough to derive the RG equation for $dA_1(t_0)/dt_0$, which also satisfies other component equations.

Main contents

The Lotka-Volterra model

- The Lotka-Volterra model of competition is

$$\begin{aligned} \dot{x} &= x(3-x-2y) \\ \dot{y} &= y(2-x-y) \end{aligned}$$

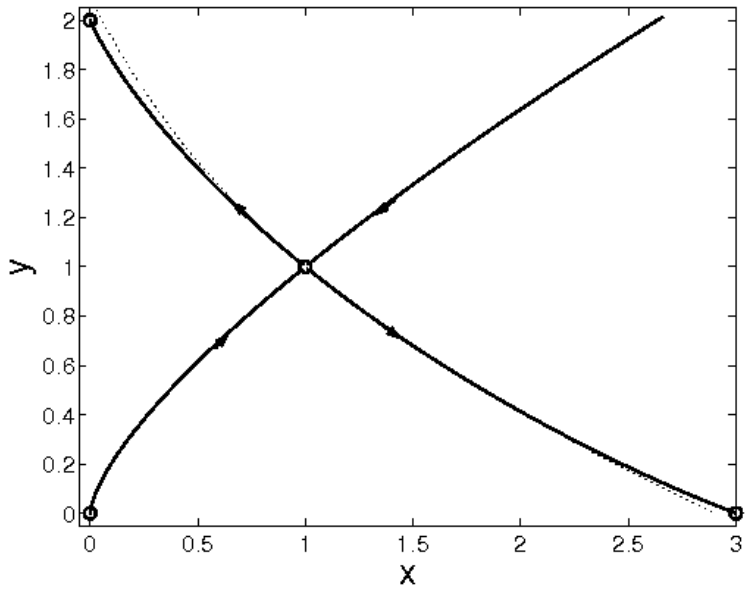


- The model describes the competition between the rabbits and the sheep fed on the grass of the same lawn.
- Vector field and trajectories
  - four equilibria
    $P_1 = (0,0)\,, P_2 = (0,2)\,, P_3 = (1,1)\,, P_4 = (3,0).$
  - Their approximation is
    $(1,1)\,, (2.908, -0.003)\,, (-0.113, 2.105).$

The Lotka-Volterra model

- The Lotka-Volterra model of competition is

$$\begin{aligned} \dot{x} &= x(3-x-2y) \\ \dot{y} &= y(2-x-y) \end{aligned}$$

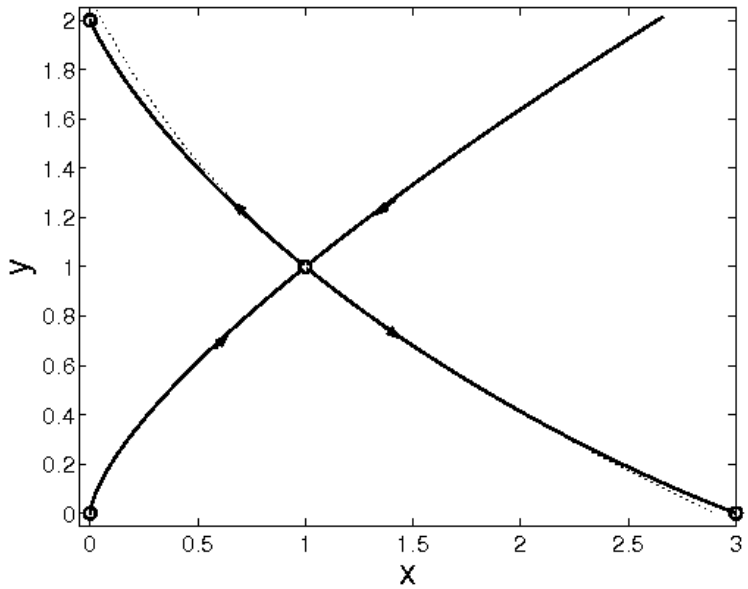


- The model describes the competition between the rabbits and the sheep fed on the grass of the same lawn.
- Vector field and trajectories
  - four equilibria
    $P_1 = (0,0)\,, P_2 = (0,2)\,, P_3 = (1,1)\,, P_4 = (3,0).$
  - Their approximation is
    $(1,1)\,, (2.908, -0.003)\,, (-0.113, 2.105).$

The Lotka-Volterra model

- The Lotka-Volterra model of competition is

$$\begin{aligned} \dot{x} &= x(3-x-2y) \\ \dot{y} &= y(2-x-y) \end{aligned}$$

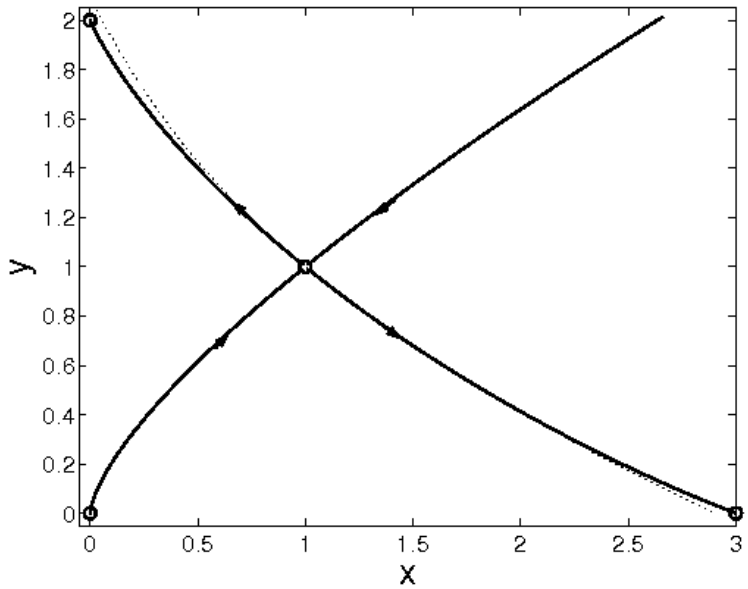


- The model describes the competition between the rabbits and the sheep fed on the grass of the same lawn.
- Vector field and trajectories
  - four equilibria $P_1 = (0,0)\,, P_2 = (0,2)\,, P_3 = (1,1)\,, P_4 = (3,0).$
  - Their approximation is $(1,1)\,, (2.908, -0.003)\,, (-0.113, 2.105).$

The Lotka-Volterra model

- The Lotka-Volterra model of competition is

$$\begin{aligned} \dot{x} &= x(3-x-2y) \\ \dot{y} &= y(2-x-y) \end{aligned}$$

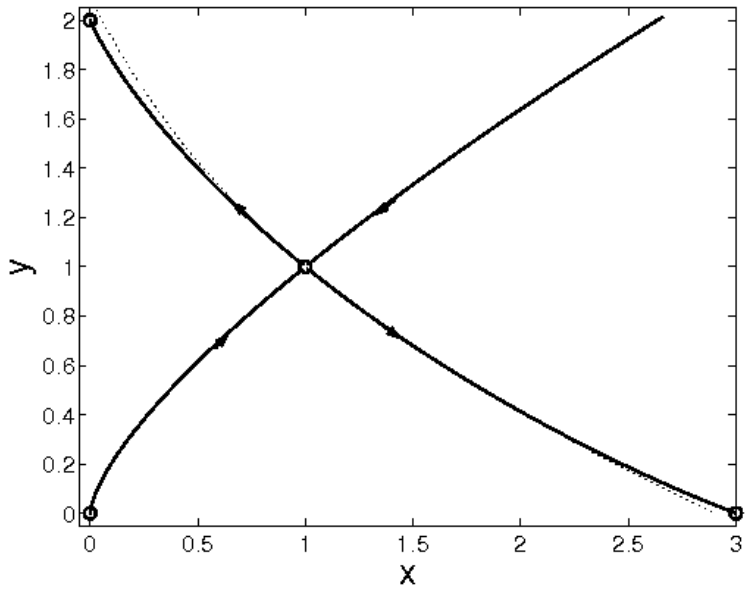


- The model describes the competition between the rabbits and the sheep fed on the grass of the same lawn.
- Vector field and trajectories
  - four equilibria
    $P_1 = (0,0)\,, P_2 = (0,2)\,, P_3 = (1,1)\,, P_4 = (3,0).$
  - Their approximation is
    $(1,1)\,, (2.908, -0.003)\,, (-0.113, 2.105).$

The Lotka-Volterra model

- The Lotka-Volterra model of competition is

$$\begin{aligned} \dot{x} &= x(3-x-2y) \\ \dot{y} &= y(2-x-y) \end{aligned}$$



- The model describes the competition between the rabbits and the sheep fed on the grass of the same lawn.
- Vector field and trajectories
  - four equilibria
    $P_1 = (0,0)\,, P_2 = (0,2)\,, P_3 = (1,1)\,, P_4 = (3,0).$
  - Their approximation is
    $(1,1)\,, (2.908, -0.003)\,, (-0.113, 2.105).$

Series solution

- The series expansion is

$$x = x_0 + \epsilon x_1 + \epsilon^2 x_2 + \ldots, y = y_0 + \epsilon y_1 + \epsilon^2 y_2 + \ldots.$$

- The solution is then

$$\begin{aligned} z &= \epsilon a(t_0) e^{(\sqrt{2}-1)(t-t_0)} + \frac{\sqrt{3}\epsilon^2 a^2(t_0)}{6}(\sqrt{2}-1) \\ &\quad (e^{2(\sqrt{2}-1)(t-t_0)} - e^{(\sqrt{2}-1)(t-t_0)}) + O(\epsilon^3) \\ w &= \frac{\sqrt{3}\epsilon^2 a^2(t_0)}{102}(1+3\sqrt{2}) e^{2(\sqrt{2}-1)(t-t_0)} + O(\epsilon^3). \end{aligned}$$

- From $\partial z(t, t_0)/\partial t_0 = 0$, we get

$$\frac{da(t_0)}{dt_0} = a\left(\sqrt{2} - 1 - \frac{17\sqrt{3}(3-2\sqrt{2})}{102}\epsilon a\right).$$

Series solution

- The series expansion is

$$x = x_0 + \epsilon x_1 + \epsilon^2 x_2 + \ldots, y = y_0 + \epsilon y_1 + \epsilon^2 y_2 + \ldots.$$

- The solution is then

$$\begin{aligned} z &= \epsilon a(t_0) e^{(\sqrt{2}-1)(t-t_0)} + \frac{\sqrt{3}\epsilon^2 a^2(t_0)}{6}(\sqrt{2}-1) \\ &\quad (e^{2(\sqrt{2}-1)(t-t_0)} - e^{(\sqrt{2}-1)(t-t_0)}) + O(\epsilon^3) \\ w &= \frac{\sqrt{3}\epsilon^2 a^2(t_0)}{102}(1+3\sqrt{2}) e^{2(\sqrt{2}-1)(t-t_0)} + O(\epsilon^3). \end{aligned}$$

- From $\partial z(t, t_0)/\partial t_0 = 0$, we get

$$\frac{da(t_0)}{dt_0} = a\left(\sqrt{2}-1-\frac{17\sqrt{3}(3-2\sqrt{2})}{102}\epsilon a\right).$$

Series solution

- The series expansion is

$$x = x_0 + \epsilon x_1 + \epsilon^2 x_2 + \ldots, y = y_0 + \epsilon y_1 + \epsilon^2 y_2 + \ldots.$$

- The solution is then

$$\begin{aligned} z &= \epsilon a(t_0) e^{(\sqrt{2}-1)(t-t_0)} + \frac{\sqrt{3}\epsilon^2 a^2(t_0)}{6}(\sqrt{2}-1) \\ &\quad (e^{2(\sqrt{2}-1)(t-t_0)} - e^{(\sqrt{2}-1)(t-t_0)}) + O(\epsilon^3) \\ w &= \frac{\sqrt{3}\epsilon^2 a^2(t_0)}{102}(1+3\sqrt{2}) e^{2(\sqrt{2}-1)(t-t_0)} + O(\epsilon^3). \end{aligned}$$

- From $\partial z(t, t_0)/\partial t_0 = 0$, we get

$$\frac{da(t_0)}{dt_0} = a\left(\sqrt{2} - 1 - \frac{17\sqrt{3}(3-2\sqrt{2})}{102}\epsilon a\right).$$

Main contents

The Kuramoto-Sivashinsky equation

- The Kuramoto-Sivashinsky equation is an important physics model

$$u_t = (u^2)_x - u_{xx} - \nu u_{xxxx} \,,$$

where $\nu > 0$ is the hyper-viscosity parameter.

- With periodic boundary condition on $[0\,, 2\pi]$, we may expand

$$u(t,x) = i \sum_{k=-\infty}^{\infty} a_k e^{ikx} \,.$$

- For the antisymmetric solution $u(t,-x) = -u(t,x)$, a_k is real and $a_{-k} = -a_k$. The equation becomes

$$\dot{a}_k = (k^2 - \nu k^4) a_k - k \sum_{m=-\infty}^{\infty} a_m a_{k-m} \,.$$

The Kuramoto-Sivashinsky equation

- The Kuramoto-Sivashinsky equation is an important physics model

$$u_t = (u^2)_x - u_{xx} - \nu u_{xxxx} \,,$$

  where $\nu > 0$ is the hyper-viscosity parameter.
- With periodic boundary condition on $[0\,, 2\pi]$, we may expand

$$u(t,x) = i \sum_{k=-\infty}^{\infty} a_k e^{ikx} \,.$$

- For the antisymmetric solution $u(t,-x) = -u(t,x)$, a_k is real and $a_{-k} = -a_k$. The equation becomes

$$\dot{a}_k = (k^2 - \nu k^4) a_k - k \sum_{m=-\infty}^{\infty} a_m a_{k-m} \,.$$

The Kuramoto-Sivashinsky equation

- The Kuramoto-Sivashinsky equation is an important physics model

$$u_t = (u^2)_x - u_{xx} - \nu u_{xxxx} \,,$$

  where $\nu > 0$ is the hyper-viscosity parameter.
- With periodic boundary condition on $[0\,, 2\pi]$, we may expand

$$u(t,x) = i \sum_{k=-\infty}^{\infty} a_k e^{ikx} \,.$$

- For the antisymmetric solution $u(t,-x) = -u(t,x)$, a_k is real and $a_{-k} = -a_k$. The equation becomes

$$\dot{a}_k = (k^2 - \nu k^4) a_k - k \sum_{m=-\infty}^{\infty} a_m a_{k-m} \,.$$

Perturbation analysis

- Assume
$$a_k = \epsilon a_{k,1} + \epsilon^2 a_{k,2} + \epsilon^3 a_{k,3} + \cdots .$$
- For the $1-d$ unstable manifold of the origin at $\nu < 1$, we may get
$$a_{1,1}(t,t_0) = r(t_0) e^{(1-\nu)(t-t_0)} , a_{k,1} = 0 \text{ for } k > 1 .$$
where $r(t_0)$ is the renormalization parameter.
- The RG equation for $r(t_0)$ is obtained from $da_1(t,t_0)/dt_0|_{t=t_0} = 0$:
$$\frac{dr_0}{dt_0} = (1-v)r_0 + \frac{2r_0^3}{1-7\nu} - \frac{6r_0^5}{(1-7\nu)^2(-1+13\nu)} + \cdots .$$

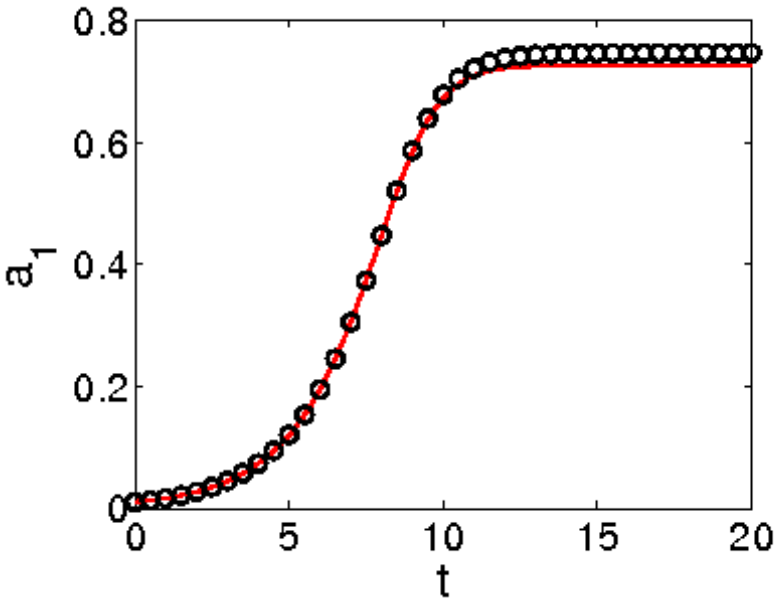
Perturbation analysis

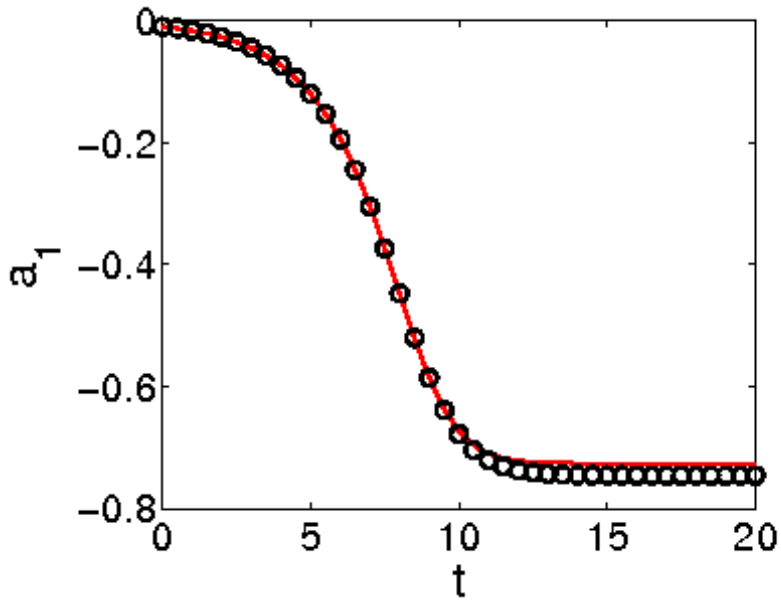
- Assume
$$a_k = \epsilon a_{k,1} + \epsilon^2 a_{k,2} + \epsilon^3 a_{k,3} + \cdots .$$
- For the $1-d$ unstable manifold of the origin at $\nu < 1$, we may get
$$a_{1,1}(t,t_0) = r(t_0) e^{(1-\nu)(t-t_0)} \, , a_{k,1} = 0 \text{ for } k > 1 \, .$$
where $r(t_0)$ is the renormalization parameter.
- The RG equation for $r(t_0)$ is obtained from $da_1(t,t_0)/dt_0|_{t=t_0} = 0$:
$$\frac{dr_0}{dt_0} = (1-v) r_0 + \frac{2r_0^3}{1-7\nu} - \frac{6r_0^5}{(1-7\nu)^2(-1+13\nu)} + \cdots .$$

Perturbation analysis

- Assume
$$a_k = \epsilon a_{k,1} + \epsilon^2 a_{k,2} + \epsilon^3 a_{k,3} + \cdots .$$
- For the $1-d$ unstable manifold of the origin at $\nu < 1$, we may get
$$a_{1,1}(t,t_0) = r(t_0) e^{(1-\nu)(t-t_0)} \, , a_{k,1} = 0 \text{ for } k > 1 \, .$$
where $r(t_0)$ is the renormalization parameter.
- The RG equation for $r(t_0)$ is obtained from $da_1(t,t_0)/dt_0|_{t=t_0} = 0$:
$$\frac{dr_0}{dt_0} = (1-v)r_0 + \frac{2r_0^3}{1-7\nu} - \frac{6r_0^5}{(1-7\nu)^2(-1+13\nu)} + \cdots .$$

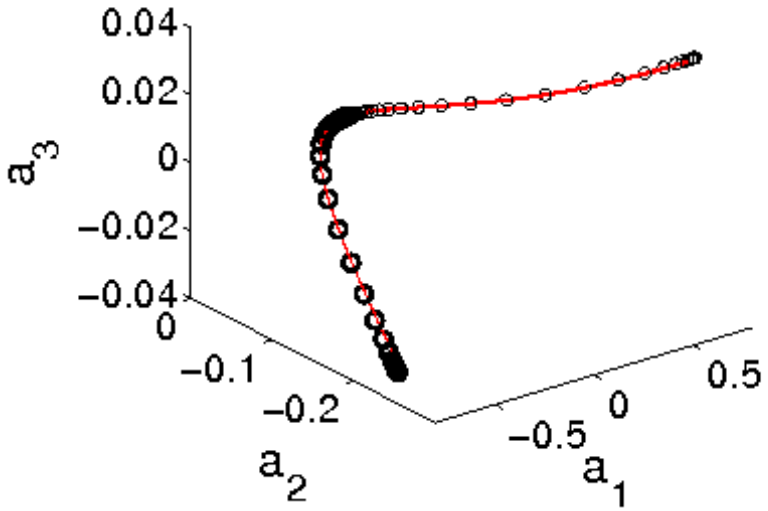
The time evolution on the connection

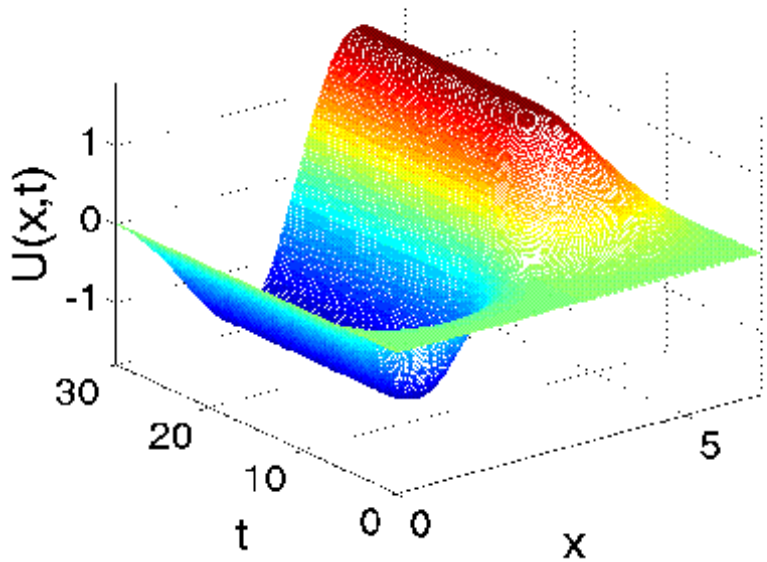




$$\nu = 0.5$$

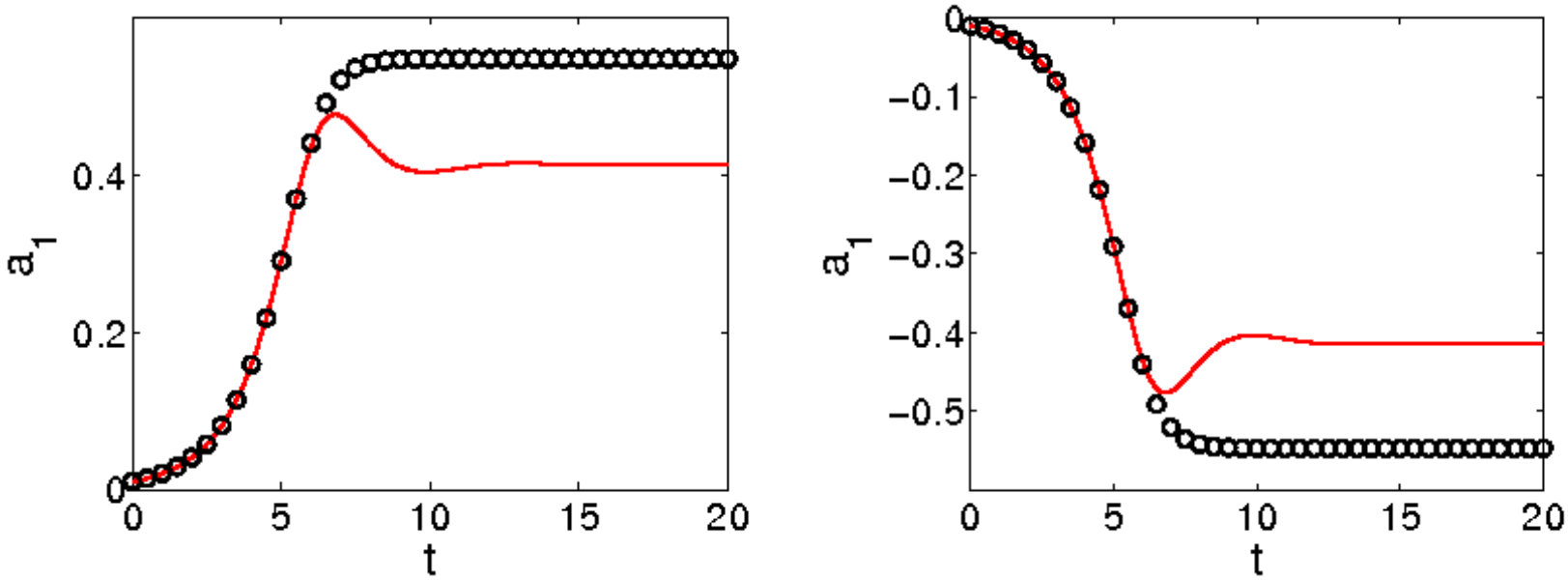
The manifold and physical observable





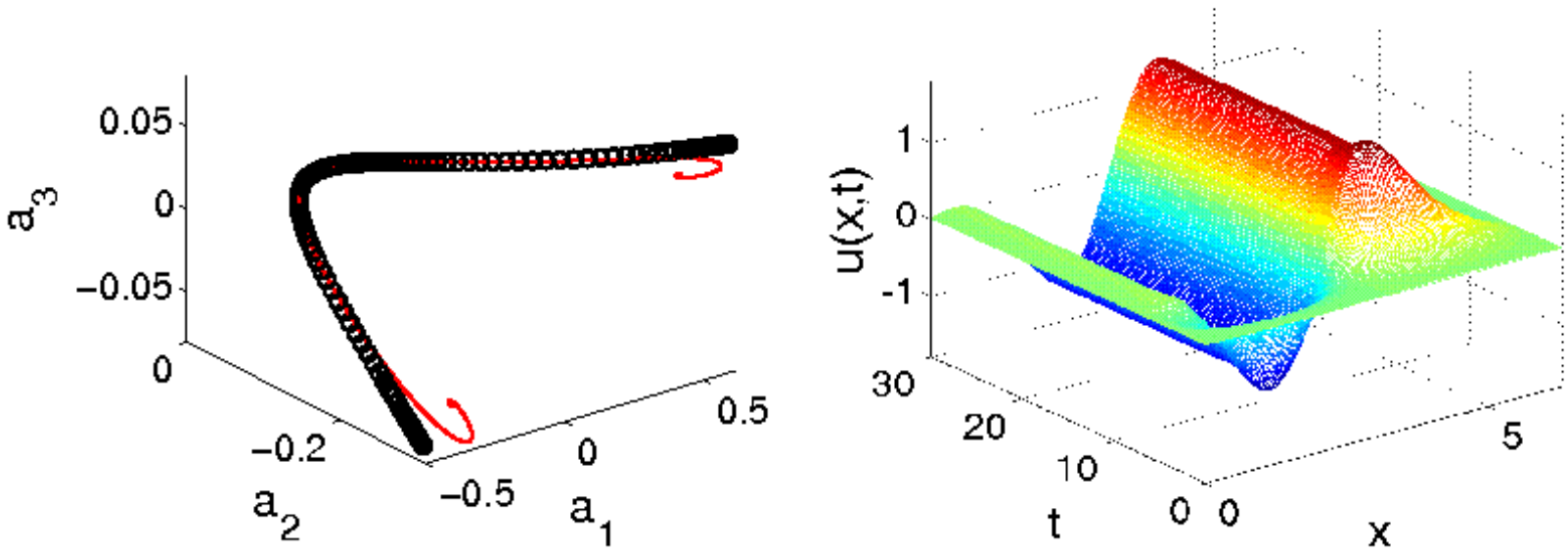
$$\nu = 0.5$$

The time evolution on the connection



$$\nu = 0.3$$

The manifold and physical observable



$$\nu = 0.3$$

[Y. Lan, Phys. Rev. E 87, 012914(2013)]

Problems and solutions

- Only simplest cases are treated. What if the ones with more complex structure? loop structure - homoclinic orbits; spirals.
- Can we match local dynamics at both ends ?
- What if multiple stable or unstable directions exist.
- Solution:
  (1) Choose a proper series expansion solutions;
  (2) Utilize the arbitrariness in the series solution;
  (3) Use the right parametrization.

Main contents

One simple ODE

- The equation is

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= x - \epsilon x^2 \,, \end{aligned}$$

where ϵ is a bookkeeping parameter. The equation is invariant under the transformation $t \to -t \,, y \to -y$.

- Along the unstable direction

$$\begin{aligned} x &= a e^{t-t_0} + \epsilon e^{t-t_0}((2+a_1)a^2 + \epsilon a_2 a^3 + \epsilon^2 a_3 a^4 - 2a^2 e^{(t-t_0)}) \\ y &= a e^{t-t_0} + \epsilon e^{t-t_0}((2+a_1)a^2 + \epsilon b_2 a^3 + \epsilon^2 b_3 a^4 - 4a^2 e^{(t-t_0)}) \,, \end{aligned}$$

while along the stable direction

$$\begin{aligned} x &= b e^{-(t-t_0)} + \epsilon((2+b_1)b^2 e^{-(t-t_0)} - 2b^2 e^{-2(t-t_0)}) + O(\epsilon^2) \\ y &= -b e^{-(t-t_0)} + \epsilon(-(2+b_1)b^2 e^{-(t-t_0)} + 4b^2 e^{-2(t-t_0)}) + O(\epsilon^2) \end{aligned}$$

One simple ODE

- The equation is

$$\begin{aligned} \dot{x} &= y \\ \dot{y} &= x - \epsilon x^2 \,, \end{aligned}$$

where ϵ is a bookkeeping parameter. The equation is invariant under the transformation $t \to -t \,, y \to -y$.

- Along the unstable direction

$$\begin{aligned} x &= a e^{t-t_0} + \epsilon e^{t-t_0}((2+a_1)a^2 + \epsilon a_2 a^3 + \epsilon^2 a_3 a^4 - 2a^2 e^{(t-t_0)}) \\ y &= a e^{t-t_0} + \epsilon e^{t-t_0}((2+a_1)a^2 + \epsilon b_2 a^3 + \epsilon^2 b_3 a^4 - 4a^2 e^{(t-t_0)}) \,, \end{aligned}$$

while along the stable direction

$$\begin{aligned} x &= b e^{-(t-t_0)} + \epsilon((2+b_1)b^2 e^{-(t-t_0)} - 2b^2 e^{-2(t-t_0)}) + O(\epsilon^2) \\ y &= -b e^{-(t-t_0)} + \epsilon(-(2+b_1)b^2 e^{-(t-t_0)} + 4b^2 e^{-2(t-t_0)}) + O(\epsilon^2) \,. \end{aligned}$$

The exact connection

- The dynamics on the connecting orbit is

$$\frac{da}{dt} = a - \frac{1}{6}(a_1 + 2)a^2 \,,$$

which has two equilibria: $a = 0$ and $a = \frac{6}{2+a_1}$.

- By setting $a \to \frac{6}{2+a_1} - b$ and comparing the above two solutions, we have
$a_1 = -1\,, a_2 = 0\,, a_3 = 0\,, b_1 = -1\,, b_2 = \frac{1}{18}\,, b_3 = 0$, which gives

$$\begin{aligned} x &= a - \frac{a^2}{6} \\ y &= a - \frac{a^2}{2} + \frac{a^3}{18}\,. \end{aligned}$$

- $a \to 6 - a$ results in $x \to x\,, y \to -y$. The solution is exact!

The exact connection

- The dynamics on the connecting orbit is

$$\frac{da}{dt} = a - \frac{1}{6}(a_1 + 2)a^2 \,,$$

  which has two equilibria: $a = 0$ and $a = \frac{6}{2+a_1}$.

- By setting $a \to \frac{6}{2+a_1} - b$ and comparing the above two solutions, we have
  $a_1 = -1\,, a_2 = 0\,, a_3 = 0\,, b_1 = -1\,, b_2 = \frac{1}{18}\,, b_3 = 0$, which gives

$$\begin{aligned} x &= a - \frac{a^2}{6} \\ y &= a - \frac{a^2}{2} + \frac{a^3}{18}\,. \end{aligned}$$

- $a \to 6 - a$ results in $x \to x\,, y \to -y$. The solution is exact!

The exact connection

- The dynamics on the connecting orbit is

$$\frac{da}{dt} = a - \frac{1}{6}(a_1 + 2)a^2 \,,$$

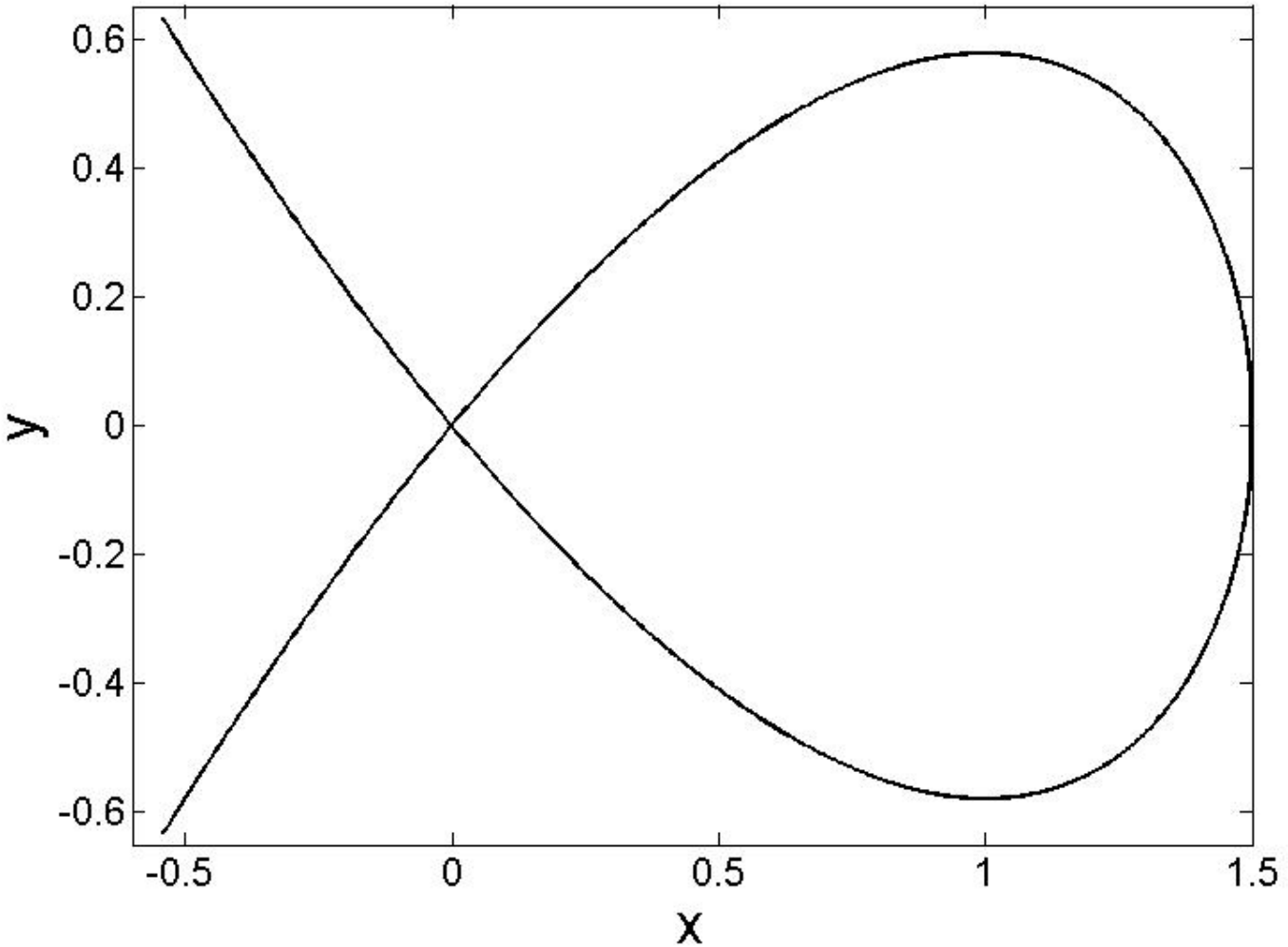
  which has two equilibria: $a = 0$ and $a = \frac{6}{2+a_1}$.

- By setting $a \to \frac{6}{2+a_1} - b$ and comparing the above two solutions, we have
  $a_1 = -1\,, a_2 = 0\,, a_3 = 0\,, b_1 = -1\,, b_2 = \frac{1}{18}\,, b_3 = 0$, which gives

$$\begin{aligned} x &= a - \frac{a^2}{6} \\ y &= a - \frac{a^2}{2} + \frac{a^3}{18}\,. \end{aligned}$$

- $a \to 6 - a$ results in $x \to x\,, y \to -y$. The solution is exact!

The homoclinic orbit of a reversible system



Main contents

The Lorenz equation

- The Lorenz equation is

$$\begin{aligned} \dot{x} &= \sigma(y-x) \\ \dot{y} &= rx - y - xz \\ \dot{z} &= xy - bz\,. \end{aligned}$$

When $\sigma = 10\,, r = 8/3\,, b = 7$, the equation has three equilibria $(0,0,0)\,, (4,4,6)\,, (-4,-4,6)$.

- The origin is a saddle and $(4,4,6)$ is a stable spiral. The solution with only one exponential $\exp(\lambda_0 t)$ has to match with the spiral solution

$$\begin{aligned} x &= 4 + A_1 r e^{-\lambda_1 t} + r e^{-\lambda_2 t}(B_1 \cos(\omega t) + C_1 \sin(\omega t)) + O(r^2) \\ y &= 4 + A_2 r e^{-\lambda_1 t} + r e^{-\lambda_2 t}(B_2 \cos(\omega t) + C_2 \sin(\omega t)) + O(r^2) \\ z &= 6 + A_3 r e^{-\lambda_1 t} + r e^{-\lambda_2 t}(B_3 \cos(\omega t) + C_3 \sin(\omega t)) + O(r^2) \end{aligned}$$

where A_i, B_j, C_k are linear functions of (a_1, a_2, a_3).

The Lorenz equation

- The Lorenz equation is

$$\begin{aligned} \dot{x} &= \sigma(y-x) \\ \dot{y} &= rx - y - xz \\ \dot{z} &= xy - bz\,. \end{aligned}$$

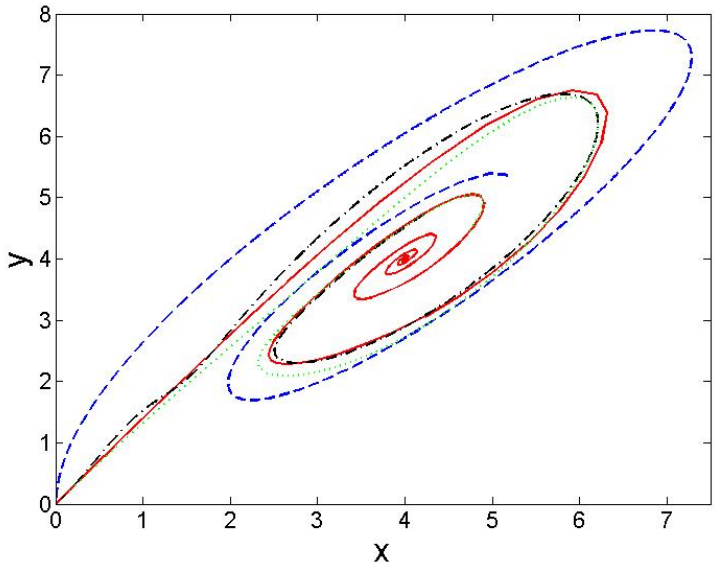
When $\sigma = 10\,, r = 8/3\,, b = 7$, the equation has three equilibria $(0,0,0)\,, (4,4,6)\,, (-4,-4,6)$.

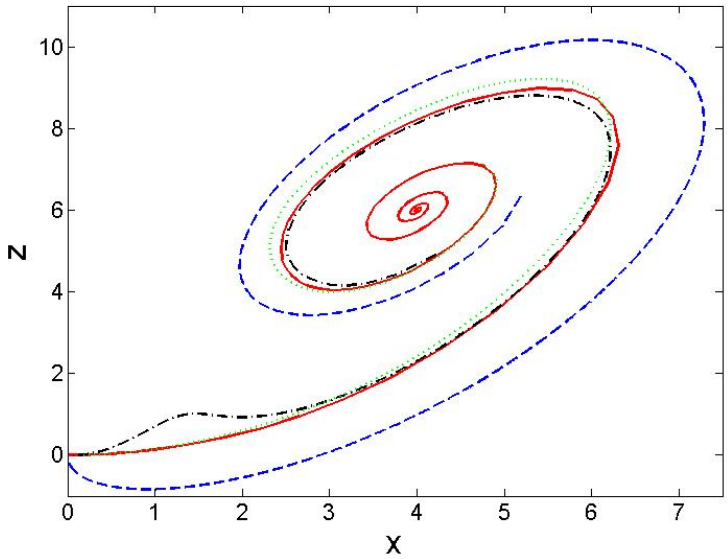
- The origin is a saddle and $(4,4,6)$ is a stable spiral. The solution with only one exponential $\exp(\lambda_0 t)$ has to match with the spiral solution

$$\begin{aligned} x &= 4 + A_1 r e^{-\lambda_1 t} + r e^{-\lambda_2 t}(B_1 \cos(\omega t) + C_1 \sin(\omega t)) + O(r^2) \\ y &= 4 + A_2 r e^{-\lambda_1 t} + r e^{-\lambda_2 t}(B_2 \cos(\omega t) + C_2 \sin(\omega t)) + O(r^2) \\ z &= 6 + A_3 r e^{-\lambda_1 t} + r e^{-\lambda_2 t}(B_3 \cos(\omega t) + C_3 \sin(\omega t)) + O(r^2) \end{aligned}$$

where A_i, B_j, C_k are linear functions of (a_1, a_2, a_3).

The spiral connection in the Lorenz system





Solid line: benchmark solution;
other lines: different approximations.

Summary

- An extension of the RG method has been proposed and was successfully used.
- The method was applied to four typical physical systems.
- Generalize to the treatment of dynamics on invariant submanifolds of dimension higher than one.
- Problems and challenges:
  (1) Treat the oscillatory case? With noise?
  (2) What is the best parameterization ?
  (3) No general rule for the convergence.
  (4) Connecting orbits between other invariants.

Summary

- An extension of the RG method has been proposed and was successfully used.
- The method was applied to four typical physical systems.
- Generalize to the treatment of dynamics on invariant submanifolds of dimension higher than one.
- Problems and challenges:
  (1) Treat the oscillatory case? With noise?
  (2) What is the best parameterization ?
  (3) No general rule for the convergence.
  (4) Connecting orbits between other invariants.

Summary

- An extension of the RG method has been proposed and was successfully used.
- The method was applied to four typical physical systems.
- Generalize to the treatment of dynamics on invariant submanifolds of dimension higher than one.
- Problems and challenges:
  (1) Treat the oscillatory case? With noise?
  (2) What is the best parameterization ?
  (3) No general rule for the convergence.
  (4) Connecting orbits between other invariants.

Summary

- An extension of the RG method has been proposed and was successfully used.
- The method was applied to four typical physical systems.
- Generalize to the treatment of dynamics on invariant submanifolds of dimension higher than one.
- Problems and challenges:
  (1) Treat the oscillatory case? With noise?
  (2) What is the best parameterization ?
  (3) No general rule for the convergence.
  (4) Connecting orbits between other invariants.

One example

- The Van der Pol equation is

$$\frac{d^2y}{dt^2} + y = \epsilon[\frac{dy}{dt} - \frac{1}{3}(dy/dt)^3] .$$

- A naive expansion

$$y = y_0 + \epsilon y_1 + \epsilon^2 y_2 + \cdots$$

gives

$$\begin{aligned} y(t) &= R_0 \sin(t + \Theta_0) + \epsilon[-\frac{R_0^3}{96}\cos(t + \Theta_0) + \\ &\frac{R_0}{2}(1 - \frac{R_0^2}{4})(t - t_0)\sin(t + \Theta_0) + \frac{R_0^3}{96}\cos 3(t + \Theta_0)] + O(\epsilon^2) \end{aligned}$$

where R_0 , Θ_0 are determined by the initial conditions.

- The expansion breaks down when $\epsilon(t - t_0) > 1$. The arbitrary initial time t_0 may be treated as the ultraviolet cutoff in the usual field theory.

One example

- The Van der Pol equation is

$$\frac{d^2y}{dt^2} + y = \epsilon[\frac{dy}{dt} - \frac{1}{3}(dy/dt)^3] .$$

- A naive expansion

$$y = y_0 + \epsilon y_1 + \epsilon^2 y_2 + \cdots$$

gives

$$\begin{aligned} y(t) &= R_0 \sin(t + \Theta_0) + \epsilon[-\frac{R_0^3}{96}\cos(t + \Theta_0) + \\ &\frac{R_0}{2} \quad (1 - \frac{R_0^2}{4})(t - t_0)\sin(t + \Theta_0) + \frac{R_0^3}{96}\cos 3(t + \Theta_0)] + O(\epsilon^2) \end{aligned}$$

where R_0 , Θ_0 are determined by the initial conditions.

- The expansion breaks down when $\epsilon(t - t_0) > 1$. The arbitrary initial time t_0 may be treated as the ultraviolet cutoff in the usual field theory.

One example

- The Van der Pol equation is

$$\frac{d^2y}{dt^2} + y = \epsilon[\frac{dy}{dt} - \frac{1}{3}(dy/dt)^3] \,.$$

- A naive expansion

$$y = y_0 + \epsilon y_1 + \epsilon^2 y_2 + \cdots$$

gives

$$\begin{aligned} y(t) \quad &= \quad R_0 \sin(t + \Theta_0) + \epsilon[-\frac{R_0^3}{96}\cos(t + \Theta_0) + \\ &\frac{R_0}{2} \quad (1 - \frac{R_0^2}{4})(t - t_0)\sin(t + \Theta_0) + \frac{R_0^3}{96}\cos 3(t + \Theta_0)] + O(\epsilon^2) \end{aligned}$$

where R_0 , Θ_0 are determined by the initial conditions.

- The expansion breaks down when $\epsilon(t - t_0) > 1$. The arbitrary initial time t_0 may be treated as the ultraviolet cutoff in the usual field theory.

Renormalization?

- Split $t-t_0$ as $t-\tau+\tau-t_0$ and absorb the terms containing $\tau-t_0$ into the renormalized counterparts $R\,,\Theta$ of R_0 and Θ_0.
- Assume $R_0(t_0)=Z_1(t_0,\tau)R(\tau)\,,\Theta_0(t_0)=\Theta(\tau)+Z_2(t_0,\tau)$ where $Z_1=1+\sum_1^\infty a_n\epsilon^n\,,Z_2=\sum_1^\infty b_n\epsilon^n$. The choice $a_1=-(1/2)(1-R^2/4)(\tau-t_0)\,,b_1=0$ removes the secular term to order ϵ:

$$\begin{aligned} y(t) &= [R+\epsilon\frac{R}{2}(-\frac{R^2}{4})(t-\tau)]\sin(t+\Theta)- \\ &\quad \epsilon\frac{R^3}{96}\cos(t+\Theta)+\epsilon\frac{R^3}{96}\cos 3(t+\Theta)+O(\epsilon^2)\,, \end{aligned}$$

where $R\,,\Theta$ are functions of τ.

Renormalization?

- Split $t-t_0$ as $t-\tau+\tau-t_0$ and absorb the terms containing $\tau-t_0$ into the renormalized counterparts $R\,,\Theta$ of R_0 and Θ_0.
- Assume $R_0(t_0)=Z_1(t_0,\tau)R(\tau)\,,\Theta_0(t_0)=\Theta(\tau)+Z_2(t_0,\tau)$ where $Z_1=1+\sum_1^\infty a_n\epsilon^n\,,Z_2=\sum_1^\infty b_n\epsilon^n$. The choice $a_1=-(1/2)(1-R^2/4)(\tau-t_0)\,,b_1=0$ removes the secular term to order ϵ:

$$\begin{aligned} y(t) &= [R+\epsilon\frac{R}{2}(-\frac{R^2}{4})(t-\tau)]\sin(t+\Theta)- \\ &\quad \epsilon\frac{R^3}{96}\cos(t+\Theta)+\epsilon\frac{R^3}{96}\cos 3(t+\Theta)+O(\epsilon^2)\,, \end{aligned}$$

where $R\,,\Theta$ are functions of τ.

Renormalization?

- The solution should not depend on τ. Therefore $(\partial y/\partial\tau)_t = 0$:

$$\frac{dR}{d\tau} = \epsilon\frac{R}{2}(1 - \frac{R^2}{4}) + O(\epsilon^2)\,, \frac{d\Theta}{d\tau} = O(\epsilon^2)\,.$$

- The initial condition $R(0) = 2a\,, \Theta(0) = 0$ gives

$$y(t) = R(t)\sin(t) + \frac{\epsilon}{96}R(t)^3[\cos(3t) - \cos(t)] + O(\epsilon^2)\,.$$

Renormalization?

- The solution should not depend on τ. Therefore $(\partial y/\partial \tau)_t = 0$:

$$\frac{dR}{d\tau} = \epsilon \frac{R}{2}(1 - \frac{R^2}{4}) + O(\epsilon^2)\,, \frac{d\Theta}{d\tau} = O(\epsilon^2)\,.$$

- The initial condition $R(0) = 2a\,, \Theta(0) = 0$ gives

$$y(t) = R(t)\sin(t) + \frac{\epsilon}{96} R(t)^3[\cos(3t) - \cos(t)] + O(\epsilon^2)\,.$$

Renormalization group transformation

Ref: K. G. Wilson, Rev. Mod. Phys. **55**, 583-595 (1983)

1. Fixed Points

— RG transformation: general scheme

— Analysis of RG flow in the space of coupling constants

- Fixed points and critical surface
- Linear analysis around the fixed point
- Scaling fields, relevant and irrelevant, universality
- Multiple fixed points

— Caveats (implicit assumptions)

(a) RG transformation

Dimensionless Hamiltonian

Coupling constants

$$H = -\beta\tilde{H} = \sum_{\alpha=1}^{n} K_\alpha \psi_\alpha\left(\{\sigma_i\}\right)$$

Terms with different form, symmetry or spatial structure

Examples:

Nearest neighbor interaction: $\psi_1 = \sum_{\langle ij\rangle} \sigma_i\sigma_j$

2nd nearest neighbor interaction: $\psi_2 = \sum_{\{ij\}} \sigma_i\sigma_j$

Coupling to a field: $\psi_3 = \sum_i \sigma_i$

"Fundamental" versus "effective": Looks a bit scary! See Wilson for a discussion on what to include and how to handle the ever expanding list of terms under the RG transformation!

RG transformation: perform the partition sum under the constraint that a subset of variables σ_I are fixed.

$$\sum_{\{\sigma_i\} \text{ at a fixed } \{\sigma_I\}} e^{H(\{\sigma_i\})} \equiv e^{H'(\{\sigma_I\})}$$

Renormalized Hamiltonian:

$$H' = Ng\left(\{K\}\right) + \sum_{\alpha=1}^{n} K_\alpha' \psi_\alpha\left(\{\sigma_I\}\right)$$

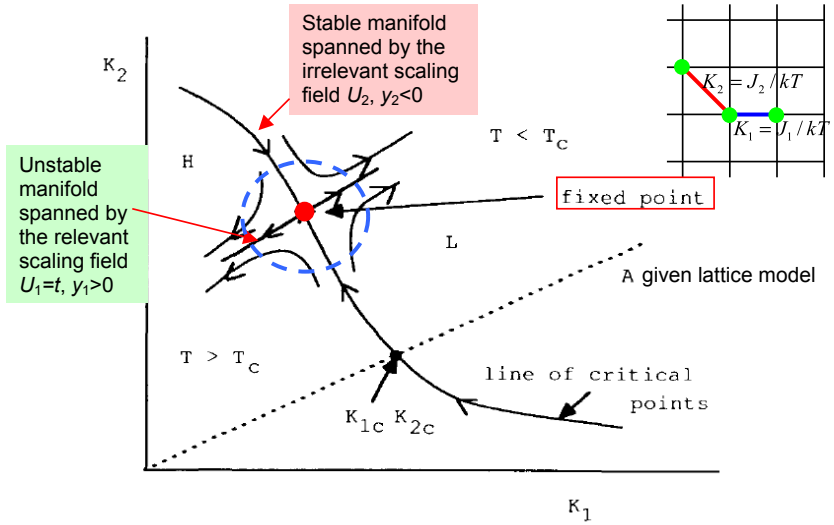
$$K_\alpha' = R_\alpha\left(K_1, K_2, \ldots, K_n\right)$$

Defines a flow in the space of coupling parameters

"Lattice constant" of the transformed system increases by a factor *b*.

(b) RG flow: topological features

Consider the example of an Ising model with nearest and 2nd nearest neighbor interactions in zero field.



i) Fixed points:

$$K_1^* = R_1\left(K_1^*, K_2^*\right)$$
$$K_2^* = R_2\left(K_1^*, K_2^*\right)$$

Classified as stable, unstable, and hyperbolic (with stable and unstable directions)

ii) Critical surface:

- "Separatrix" of the RG flow, with $\xi = \infty$ everywhere.
- Contains the hyperbolic fixed point.

iii) Linear analysis near the fixed point

$$\begin{aligned} K_1' &= R_1\left(K_1^* + \delta K_1, K_2^* + \delta K_2\right) \\ &= K_1^* + \delta K_1 \left.\frac{\partial R_1}{\partial K_1}\right|_{K_1^*, K_2^*} + \delta K_2 \left.\frac{\partial R_1}{\partial K_2}\right|_{K_1^*, K_2^*} \\ &= K_1^* + \delta K_1' \end{aligned}$$

$$\Rightarrow \begin{pmatrix} \delta K_1' \\ \delta K_2' \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} \delta K_1 \\ \delta K_2 \end{pmatrix}, \qquad M_{ij} = \left.\frac{\partial R_i}{\partial K_j}\right|_{K_1^*, K_2^*}$$

Left eigenvectors of matrix **M**:

$$\sum_i \varphi_{\alpha i} M_{ij} = \lambda_\alpha \varphi_{\alpha j}, \qquad \lambda_\alpha = b^{y_\alpha}$$

block size

Define

$$U_\alpha = \varphi_\alpha \cdot \delta K = \sum_i \varphi_{\alpha i} \delta K_i, \qquad \text{scaling fields}$$

Under the RG transformation,

$$U_\alpha' = \varphi_\alpha \cdot \delta K' = \varphi_\alpha \cdot \mathbf{M} \delta K = \lambda_\alpha \varphi_\alpha \cdot \delta K = \lambda_\alpha U_\alpha$$

Changes by a factor set by the eigenvalue.

Note that the scaling field describing flow along the critical surface, say U_2, must correspond to a contracting direction with y_2<0. The other scaling field U_1, which measures the distance to the critical surface, must have a y_1>0. U_1 can in fact be identified with $t = (T - T_c)/T_c$.

iv) Widom scaling, relevant scaling fields

Singular part of the free energy under the RG transformation:

$$f_s(U_1,U_2) = b^{-d} f_s(U_1',U_2') = b^{-d} f_s(b^{y_1}U_1, b^{y_2}U_2)$$

reminds us the Widom scaling

Identify $U_1 = t$, and choose $b = |t|^{-1/y_1}$,

$$\Rightarrow \quad f_s(t,U_2) = |t|^{d/y_1} f_s(\pm 1, |t|^{-y_2/y_1} U_2)$$

Since $y_2 < 0$, the second argument of the function on the right vanishes as $t \to 0$. Therefore the dependence of f_s on t (at zero field) is given by the prefactor $|t|^{d/y_1} = |t|^{2-\alpha}$, with $\alpha = 2 - d/y_1$.

More importantly, models with different U_2 flow to the same fixed point.

$\Rightarrow$ U_2 is an irrelevant scaling field.

Universality: models that differ microscopically (say with a different initial U_2) exhibit exactly the same critical behavior and share the same set of critical indices. In the language of the RG, these models differ in terms of the irrelevant scaling fields.

v) General case:

The above discussion can be easily generalized to the case with *n* scaling fields. See textbook for details. Scaling fields with a positive y_α are known as relevant and must be included in the discussion. Those with a negative y_α are irrelevant and can be dropped from the discussion.

Critical exponents of a given model can be obtained from the y_α's once correct associations are made.

vi) Multiple fixed points

The critical surface may have several internal dimensions and contains more than one fixed point.

Example: the *n*-component ferromagnetic spin model

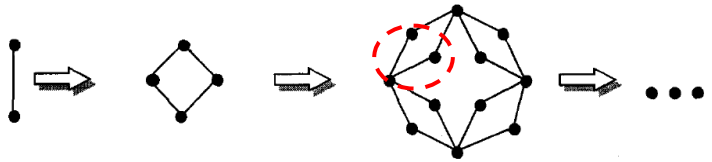
$$\tilde{H} = -\sum_{\langle ij \rangle, \alpha} J_\alpha S_i^\alpha S_j^\alpha$$

In the space spanned by the $K_\alpha = J_\alpha / kT$, the critical manifold contains at least *n* fixed points.

n = 1:	Ising fixed point
n = 2:	XY fixed point
n = 3:	Heisenberg fixed point

2. Ising Spins on the Diamond Fractal Lattice

Lattice generated through an iterative procedure



$l=1$ $l=2$ $l=3$

Placing Ising spins on the vertices of the lattice.

Summation of spins on the middle vertices of a diamond renormalizes the interaction and field strength on the upper and lower vertices:

$$Z(N,J,H) = e^{Ng(J,h)} Z(N/4, J', h')$$

$$g = \frac{1}{8}\ln[16\cosh(2J+h)\cosh(2J-h)\cosh^2 h]$$

$$J' = \frac{1}{2}\ln\left[\frac{\cosh(2J+h)\cosh(2J-h)}{\cosh^2 h}\right] \equiv R_1(J,h)$$

$$h' = h + \ln\frac{\cosh(2J+h)}{\cosh(2J-h)} \equiv R_2(J,h) \ .$$

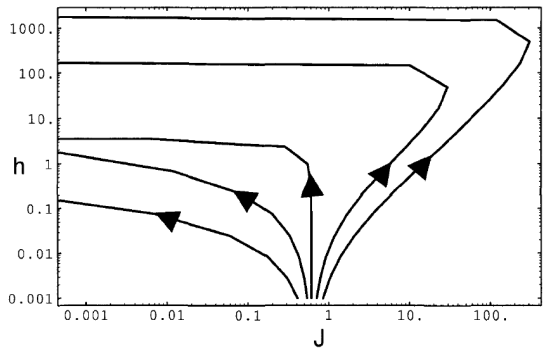
Same as the 1D Ising chain except each time two branches contribute, so the results differ by a factor of 2.

RG flow

Critical point at

$h^* = 0$

$J^* = 0.60938$



singular part of the free energy must thus be of the scaling form

$$f_s(t,h) = \frac{1}{4} f_s(\frac{\partial R_1}{\partial J}t, \frac{\partial R_2}{\partial h}h) = \frac{1}{\lambda} f_s(\lambda^y t, \lambda^x h)$$

with $\lambda = 4$ and

$$y = \frac{\ln 1.67857}{\ln 4} = 0.373618$$

$$x = \frac{\ln 2.67857}{\ln 4} = 0.710732$$

from which we find the critical exponents

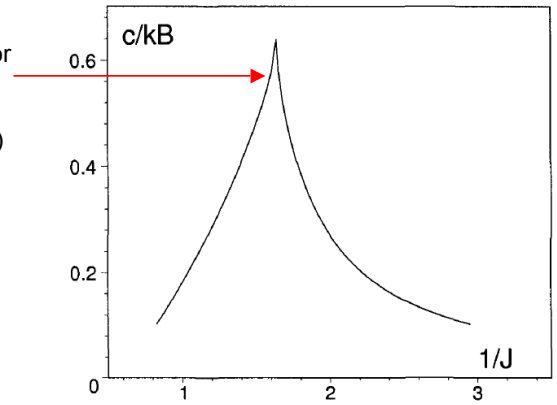
$$\alpha = \frac{2y-1}{y} = -0.676532$$

$$\beta = \frac{1-x}{y} = 0.774234$$

$$\gamma = \frac{2x-1}{y} = 1.12806$$

$$\delta = \frac{x}{1-x} = 2.45701 \ .$$

Cusp singularity for the specific heat (with a negative α)



3. The ε-expansion

Wilson and Fisher, 1972.

For simplicity in notation, let's consider the scalar model (n=1).

Spin configuration represented by a continuous field $S_{\mathbf{r}} \in (-\infty, \infty)$.

Hamiltonian: $$H = -\int d^d r \left[\frac{1}{2} r S_{\mathbf{r}}^2 + u S_{\mathbf{r}}^4 + \frac{K}{2} \left| \nabla S_{\mathbf{r}} \right|^2 \right]$$

Fourier transforms

$$S_{\mathbf{q}} = \int d^d r e^{-i\mathbf{q}\cdot\mathbf{r}} S_{\mathbf{r}} \qquad \text{transform}$$

$$S_r = \int_{|\mathbf{q}|\le\Lambda} \frac{d^d q}{(2\pi)^d} e^{i\mathbf{q}\cdot\mathbf{r}} S_{\mathbf{q}} = \int_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}} S_{\mathbf{q}} \qquad \text{inverse transform}$$

$$\int d^d r e^{-i\mathbf{q}\cdot\mathbf{r}} = (2\pi)^d \delta^d(\mathbf{q}) \qquad \text{orthogonality}$$

$$\int_{\mathbf{q}} e^{i\mathbf{q}\cdot\mathbf{r}} = \delta^d(\mathbf{r}) \qquad \text{completeness}$$

$$H = H_0 + H_1$$

$$H_0 = -\frac{1}{2} \int_{\mathbf{q}} \left(r + Kq^2 \right) S_{\mathbf{q}} S_{-\mathbf{q}}$$

$$H_1 = -u \int_{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3,\mathbf{q}_4} S_{\mathbf{q}_1} S_{\mathbf{q}_2} S_{\mathbf{q}_3} S_{\mathbf{q}_4} (2\pi)^d \delta^d(\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 + \mathbf{q}_4)$$

Energy in a magnetic field:

$$h \int d^d r S_{\mathbf{r}} = h S_{\mathbf{q}=0}$$

3.1 The Gaussian model (u = 0)

i) momentum shell integration: elimination of modes $S_{\mathbf{q}}$ with $\mathbf{q}$ in the momentum shell $\Lambda' < |\mathbf{q}| \le \Lambda$, does not change r and h.

ii) rescaling: $\mathbf{q} \to \tilde{\mathbf{q}} = \left(\frac{\Lambda}{\Lambda'} \right) \mathbf{q} = l\mathbf{q}, \quad S_{\mathbf{q}} \to \tilde{S}_{\mathbf{q}} = l^{-(2+d)/2} S_{\mathbf{q}}$

$$r' = r l^2, \quad h' = l^{(2+d)/2} h, \quad K' = K$$

Identifying r with $t = (T - T_c)/T_c$, the singular part of the free energy transforms as,

$$g(t,h) = l^{-d} g(t l^2, h l^{1+d/2})$$

$\Rightarrow$ $\alpha = 2 - d/2, \quad \beta = d/4 - 1/2$ (Gaussian exponents)

Identical to the mean field exponents when d = 4.

Gaussian averages:

Define $$\langle O \rangle = \frac{\prod_{\mathbf{q}} \int dS_{\mathbf{q}} O e^{-\frac{1}{2}\int_{\mathbf{q}} a_{\mathbf{q}} S_{\mathbf{q}} S_{-\mathbf{q}}}}{\prod_{\mathbf{q}} \int dS_{\mathbf{q}} e^{-\frac{1}{2}\int_{\mathbf{q}} a_{\mathbf{q}} S_{\mathbf{q}} S_{-\mathbf{q}}}}$$

$$\left\langle S_{\mathbf{q}_1} S_{\mathbf{q}_2} \right\rangle = \frac{1}{a_{\mathbf{q}_1}} (2\pi)^d \delta^d(\mathbf{q}_1 + \mathbf{q}_2)$$

Wick's theorem (classical fields)

$$\left\langle S_{\mathbf{q}_1} S_{\mathbf{q}_2} \cdots S_{\mathbf{q}_n} \right\rangle = \sum_{j \ne 1} \left\langle S_{\mathbf{q}_1} S_{\mathbf{q}_j} \right\rangle \left\langle S_{\mathbf{q}_2} \cdots S_{\mathbf{q}_{j-1}} S_{\mathbf{q}_{j+1}} \cdots S_{\mathbf{q}_n} \right\rangle$$

3.2 The ϕ^4 model

To derive the RG transformation, we

i) Separate H into a part $H_<$ which only depends on $S_{\mathbf{q}}$ with $q \le \Lambda - d\Lambda$ and the rest $\delta H = \delta H_0 + \delta H_1$, with

$$\delta H_0 = -\frac{1}{2}\int_{\Lambda - d\Lambda < |\mathbf{q}| \le \Lambda} \left(r + Kq^2\right) S_{\mathbf{q}} S_{-\mathbf{q}}$$

$$\delta H_1 = -u \int_{\substack{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3,\mathbf{q}_4 \\ \text{with at least one in} \\ \text{the momentum shell}}} S_{\mathbf{q}_1} S_{\mathbf{q}_2} S_{\mathbf{q}_3} S_{\mathbf{q}_4} (2\pi)^d \delta^d\left(\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 + \mathbf{q}_4\right)$$

ii) Perform the momentum shell integration:

$$\prod_{\Lambda - d\Lambda < |\mathbf{q}| \le \Lambda} \int dS_{\mathbf{q}} e^{H_< + \delta H_0 + \delta H_1} = e^{H_<} \prod_{\Lambda - d\Lambda < |\mathbf{q}| \le \Lambda} \int dS_{\mathbf{q}} e^{\delta H_0} e^{\delta H_1} = A e^{H_<} \left\langle e^{\delta H_1} \right\rangle_{\delta H_0}$$

Here $A = \prod_{\Lambda - d\Lambda < |\mathbf{q}| \le \Lambda} \int dS_{\mathbf{q}} e^{\delta H_0}$ is some constant. The average is done over the modes in the momentum shell under the Gaussian model δH_0.

Cumulant expansion:

$$\ln \left\langle e^{\delta H_1} \right\rangle_{\delta H_0} = \left\langle \delta H_1 \right\rangle_{\delta H_0} + \frac{1}{2}\left[\left\langle \delta H_1^2 \right\rangle_{\delta H_0} - \left\langle \delta H_1 \right\rangle_{\delta H_0}^2 \right] + \ldots$$

1st cumulant 2nd cumulant

Computation of $\left\langle \delta H_1 \right\rangle_{\delta H_0}$:

Relevant contribution comes from the case when two of the four q's are in the momentum shell.

$$\begin{aligned}
\left\langle \delta H_1 \right\rangle_{\delta H_0} &= const - 6u \int_{|\mathbf{q}_3|,|\mathbf{q}_4| \le \Lambda - d\Lambda} \int_{\Lambda - d\Lambda < |\mathbf{q}_1|,|\mathbf{q}_2| \le \Lambda} \left\langle S_{\mathbf{q}_1} S_{\mathbf{q}_2} \right\rangle S_{\mathbf{q}_3} S_{\mathbf{q}_4} (2\pi)^d \delta^d\left(\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 + \mathbf{q}_4\right) \\
&= const - 6u \int_{|\mathbf{q}_3|,|\mathbf{q}_4| \le \Lambda - d\Lambda} \int_{\Lambda - d\Lambda < |\mathbf{q}_1|,|\mathbf{q}_2| \le \Lambda} \frac{1}{r + K\Lambda^2} (2\pi)^d \delta^d\left(\mathbf{q}_1 + \mathbf{q}_2\right) \\
&\qquad \times S_{\mathbf{q}_3} S_{\mathbf{q}_4} (2\pi)^d \delta^d\left(\mathbf{q}_1 + \mathbf{q}_2 + \mathbf{q}_3 + \mathbf{q}_4\right) \\
&= const - 6u \int_{|\mathbf{q}_3|,|\mathbf{q}_4| \le \Lambda - d\Lambda} S_{\mathbf{q}_3} S_{\mathbf{q}_4} (2\pi)^d \delta^d\left(\mathbf{q}_3 + \mathbf{q}_4\right) \int_{\Lambda - d\Lambda < |\mathbf{q}_1| \le \Lambda} \frac{1}{r + K\Lambda^2} \\
&= const - \frac{6u}{r + K\Lambda^2} K_d \Lambda^{d-1} d\Lambda \int_{|\mathbf{q}_3| \le \Lambda - d\Lambda} S_{\mathbf{q}_3} S_{-\mathbf{q}_3}
\end{aligned}$$

Here $K_d = A_d / (2\pi)^d = 2^{1-d}\pi^{-d/2} / \Gamma(d/2)$ is a constant.

The above equation yields a change in *r* under the RG transformation:

$$r \to \tilde{r} = r + \frac{12u}{r + K\Lambda^2} K_d \Lambda^{d-1} d\Lambda$$

Computation of $\left\langle \delta H_1^2 \right\rangle_{\delta H_0}$:

$$\delta H_1^2 = u^2 \int_{\substack{\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3,\mathbf{q}_4 \\ \text{with at least one in} \\ \text{the momentum shell}}} S_{\mathbf{q}_1} S_{\mathbf{q}_2} S_{\mathbf{q}_3} S_{\mathbf{q}_4} (2\pi)^d \delta^d(\mathbf{q}_1+\mathbf{q}_2+\mathbf{q}_3+\mathbf{q}_4)$$
$$\times \int_{\substack{\mathbf{q}_5,\mathbf{q}_6,\mathbf{q}_7,\mathbf{q}_8 \\ \text{with at least one in} \\ \text{the momentum shell}}} S_{\mathbf{q}_5} S_{\mathbf{q}_6} S_{\mathbf{q}_7} S_{\mathbf{q}_8} (2\pi)^d \delta^d(\mathbf{q}_5+\mathbf{q}_6+\mathbf{q}_7+\mathbf{q}_8)$$

Relevant contributions come from

i) One $\mathbf{q}$ from each group is in the momentum shell $\Rightarrow S^6$ term, irrelevant.

ii) Four of the $\mathbf{q}$'s are in the momentum shell, with at least one in each group $\Rightarrow$ renormalizes *u*

iii) Six of the $\mathbf{q}$'s are in the momentum shell, $\Rightarrow$ renormalizes *r* (higher order in *u*, not considered at this order)

Essentially, we only need to worry about terms from category ii).

Due to "kinematics", i.e., the constraints on the $\mathbf{q}$'s from the δ-functions and the infinitesimal width of the momentum shell, terms that contribute involve pairing of two $\mathbf{q}$'s from one group each with a partner in the second group. The number of such combinations is 72.

Consider a particular example:

$$\int_{|\mathbf{q}_3|,|\mathbf{q}_4|,|\mathbf{q}_7|,|\mathbf{q}_8| \le \Lambda - d\Lambda} S_{\mathbf{q}_3} S_{\mathbf{q}_4} S_{\mathbf{q}_7} S_{\mathbf{q}_8} \int_{\Lambda - d\Lambda < |\mathbf{q}_1|,|\mathbf{q}_2|,|\mathbf{q}_5|,|\mathbf{q}_6| \le \Lambda} \left\langle S_{\mathbf{q}_1} S_{\mathbf{q}_5} \right\rangle \left\langle S_{\mathbf{q}_2} S_{\mathbf{q}_6} \right\rangle$$
$$\times (2\pi)^d \delta^d(\mathbf{q}_1+\mathbf{q}_2+\mathbf{q}_3+\mathbf{q}_4)(2\pi)^d \delta^d(\mathbf{q}_5+\mathbf{q}_6+\mathbf{q}_7+\mathbf{q}_8)$$
$$= \int_{|\mathbf{q}_3|,|\mathbf{q}_4|,|\mathbf{q}_7|,|\mathbf{q}_8| \le \Lambda - d\Lambda} S_{\mathbf{q}_3} S_{\mathbf{q}_4} S_{\mathbf{q}_7} S_{\mathbf{q}_8} (2\pi)^d \delta^d(\mathbf{q}_3+\mathbf{q}_4+\mathbf{q}_7+\mathbf{q}_8)$$
$$\times \frac{1}{\left(r + K\Lambda^2\right)^2} \int_{\Lambda - d\Lambda < |\mathbf{q}_1|,|\mathbf{q}_2| \le \Lambda} (2\pi)^d \delta^d(\mathbf{q}_1+\mathbf{q}_2-\mathbf{q}_7-\mathbf{q}_8)$$

The double integral in the box above depends on both $d\Lambda$ and $\mathbf{q} = \mathbf{q}_7 + \mathbf{q}_8$. For $|\mathbf{q}| < d\Lambda$, which is the case of interest here, the double integral yields $K_d \Lambda^{d-1} d\Lambda$. Hence correction to the *u*-term takes the form,

$$u \to \tilde{u} = u - \frac{36u^2}{\left(r + K\Lambda^2\right)^2} K_d \Lambda^{d-1} d\Lambda$$

Rescaling

As in the Gaussian case, we now perform the rescaling

$$\mathbf{q} \to \tilde{\mathbf{q}} = \left(\frac{\Lambda}{\Lambda'}\right)\mathbf{q}, \quad S_{\mathbf{q}} \to \tilde{S}_{\mathbf{q}} = \left(\frac{\Lambda}{\Lambda'}\right)^{-(2+d)/2} S_{\mathbf{q}}.$$

Let $dl = d\Lambda / \Lambda$, the combined change of parameters takes the form,

$$r' = \tilde{r}\left(1+dl\right)^2 = r + \left(2r + \frac{12u}{r+K\Lambda^2}C\right)dl$$

$$u' = \tilde{u}\left(1+dl\right)^{4-d} = u + \left[\left(4-d\right)u - \frac{36u^2}{\left(r+K\Lambda^2\right)^2}C\right]dl$$

$$h' = \left(1+dl\right)^{(2+d)/2} h = h + \frac{2+d}{2}hdl$$

$$K' = K$$

where $C = K_d \Lambda^d$

RG flow equations:

$$\frac{dr}{dl} = 2r + \frac{12u}{r+K\Lambda^2}C, \qquad \frac{dh}{dl} = \frac{2+d}{2}h$$

$$\frac{du}{dl} = \left(4-d\right)u - \frac{36u^2}{\left(r+K\Lambda^2\right)^2}C, \qquad \frac{dK}{dl} = 0$$

Fixed points:

i) Gaussian: $h^* = r^* = u^* = 0$

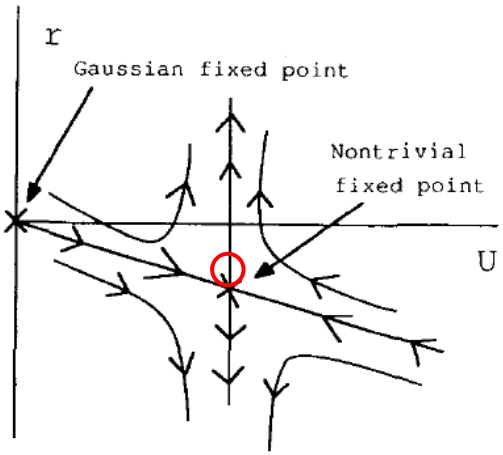
ii) nontrivial: $h^* = 0, r^* = -\frac{\varepsilon}{6+\varepsilon}K\Lambda^2, u^* = \frac{\varepsilon}{\left(6+\varepsilon\right)^2}\frac{K^2\Lambda^4}{C}$

where $\varepsilon = 4-d$. The two merge at d = 4.

RG flow for d<4:

Location of the fixed point and the RG flow are correct to the first order in ε.

Additional terms in the Landau expansion need to be considered if higher order terms in ε are desired in the RG flow equation.



Critical exponents

Linearize around the nontrival fixed point, and keep terms to the first order in ε:

$$\frac{d\delta r}{dl} = \left(2 - \frac{\varepsilon}{3}\right)\delta r + \frac{12C}{K\Lambda^2}\delta u$$

$$\frac{d\delta u}{dl} = -\varepsilon \delta u$$

In matrix form, $$\frac{d}{dl}\begin{pmatrix} \delta r \\ \delta u \end{pmatrix} = \begin{pmatrix} 2 - \frac{\varepsilon}{3} & \frac{12C}{K\Lambda^2} \\ 0 & -\varepsilon \end{pmatrix}\begin{pmatrix} \delta r \\ \delta u \end{pmatrix}$$

Hence the two eigenvalues are given by

$y_1 = 2 - \frac{\varepsilon}{3} > 0,$ scaling of t

$y_2 = -\varepsilon < 0,$ irrelevant but needed to obtain the order ε term.

In addition, $y_h = 1 + \frac{d}{2} = 3 - \frac{\varepsilon}{2}.$

$$\alpha = 2 - \frac{d}{y_1} = \frac{\varepsilon}{6} + O\left(\varepsilon^2\right)$$

$$\beta = \frac{d - y_h}{y_1} = \frac{1}{2} - \frac{\varepsilon}{6} + O\left(\varepsilon^2\right)$$

$$\gamma = \frac{2y_h - d}{y_1} = 1 + \frac{\varepsilon}{6} + O\left(\varepsilon^2\right)$$

$$\delta = \frac{y_h}{d - y_h} = 3 + \varepsilon + O\left(\varepsilon^2\right)$$

	α		β		γ		δ	
		RG (1-loop)		RG (1-loop)		RG (1-loop)		RG (1-loop)
d ≥ 4	0	0	1/2	1/2	1	1	3	3
d = 3	0.10	0.17	0.33	0.33	1.24	1.17	4.8	4
d = 2	0 (log)	0.33	1/8	0.17	7/4	1.33	15	5

3.3 Conclusions

- Derivation can be made more systematic by using Feynman diagram techniques for book-keeping. This is a must when one pushes the calculation to higher orders in ε. References can be found in the textbook (e.g., S-K Ma, Zinn-Justin)
- Demonstration of the notion of universality: which aspects of the problem affect and do not affect critical properties: dimensionality, symmetry, etc.
- A universal language and framework to study physical systems that exhibit scale invariance, with applications outside equilibrium stat mech.

An introduction to the non-perturbative renormalization group

B. Delamotte (LPTMC, UPMC, CNRS)
Course at IPhT, CEA-Saclay, consisting of 7 sessions (2h each).
Lecture notes taken and typed by: X. Cao, L. Fister, J.J. Lopez-Villarejo (IPhT, CEA-Saclay)
A comprehensive review on the subject of the course is in preparation by the lecturer.

(Dated: February 19, 2014)

CONTENTS

"Il y a un autre monde mais il est dans celui-ci."
P. Eluard

I. INTRODUCTION

The subject of this course is the Non-Perturbative Renormalization Group (NPRG). Wilson was the pioneer in this field with his Renormalization Group (RG) ideas. Although he devised a conceptual framework that was valid at the non-perturbative level — and which he actually applied as such to the Kondo problem —, his RG calculation techniques were mainly used in a perturbative framework. The modern form that we will be studying is mainly due to Wetterich[1]. Other nomenclatures to refer to the NPRG in the literature include the names 'exact RG' (ERG) or 'functional RG' (FRG).

In its full form the NPRG equation is an exact description, hence, it is applicable without conceptual limitations. In practice, however, truncations are inevitable. They have to be constructed such that the dominating effects of the physics at hand are captured well. However, by this freedom in choosing truncations the flow equations maintains its flexibility, such that a variety of fields has been covered so far. In particular in statistical systems or quantum field theories with (second order) phase transitions the flow equation is a powerful method, as comparatively simple approximations suffice to obtain quantitatively satisfying results. Moreover, the NPRG has been applied to as complex systems as quantum chromodynamics or quantum gravity, where in the latter, Weinberg's idea of asymptotic safety, meaning the existence of a non-trivial ultraviolet fixed-point, is realised.

II. STATISTICAL MECHANICS – O(N) MODELS

We present equilibrium statistical models defined on the lattice for various ferromagnetic systems with a $O(N)$ symmetry. They are related to euclidean field theories

[1] U. Ellwanger and, later, T. Morris also developed closely related ideas.

based on the same $O(N)$ symmetry[2], and share with them *universal* properties.

For systems at thermal equilibrium, the stationary probability distribution is entirely determined by the hamiltonian H of the system and is Gibbsian: $P_{\rm eq}(\phi) \propto \exp(-H(\phi)/kT)$.

A. The O(N) models on the lattice

Let us present the ferromagnetic O(N) models defined on the lattice (d-dimensional, with lattice spacing a).

$$\begin{cases} O(3) \to \text{Heisenberg model} \\ O(2) \to \text{XY model} \\ O(1) \approx \mathcal{Z}_2 \to \text{Ising model} \end{cases}$$

We consider classical spins $\mathbf{S}_i$ defined on lattice sites i and having N components. The simplest $O(N)$-invariant hamiltonian is

$$H = -J \sum_{\langle ij \rangle} \mathbf{S}_i \cdot \mathbf{S}_j \,, \tag{1}$$

where $J > 0$ (ferromagnetic interaction), $\mathbf{S}_i^2 = 1$ and $\langle ij \rangle$ means summation on nearest neighbor pairs of spins. The coupling to an external source $\mathbf{h}$ (magnetic field) contributes to the hamiltonian by a term:

$$- \sum_i \mathbf{h}_i \cdot \mathbf{S}_i \ . \tag{2}$$

The partition function is

$$\mathcal{Z}[\mathbf{h}] = \sum_{\{\mathbf{S}_i\}} \exp\{-H + \sum_i \mathbf{h}_i \cdot \mathbf{S}_i\} \,, \tag{3}$$

where the $\beta = 1/k_B T$ coefficient in front of the hamiltonian has been absorbed into the definition of H for simplicity; the symbolic sum on $\{\mathbf{S}_i\}$ adds up all spin configurations for all lattice sites (it may contain integrals).

We can define an average magnetization per lattice site,

$$\mathbf{M} = \langle \mathbf{S}_i \rangle \,, \tag{4}$$

which plays the role of an *order parameter* in this system, characterizing the phase transition (see below). For now, we will assume the thermodynamic limit[3] in the following, because strictly-speaking there is no phase transition for a system with a finite number of degrees of freedom.

[2] Field theories whose potential terms are powers of $\boldsymbol{\varphi}^2$.

[3] infinite number of degrees of freedom associated to an infinite volume; not the continuum limit.

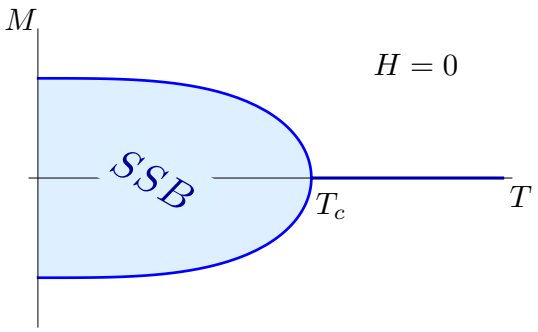


FIG. 1 Diagram magnetization-temperature. For T larger than the critical temperature T_c, $M = 0$ and the system in its symmetric phase while for $T < T_c$, $M \neq 0$ and the system is in its spontaneously symmetry broken (SSB) phase.

There are two competing effects as regards the value of the average magnetization: the kinetic / vibrational energy associated to the temperature and the tendency of spins to align among themselves, given by the attractive spin-coupling term in the hamiltonian. At high temperature, the thermal disordering effect dominates, and we have a symmetrical phase where the average magnetization is zero. At small temperature, the spin-couplings dominate and there is a spontaneous breaking of the symmetry $O(N)$ to $O(N-1)$: spins align preferentially in one direction. These two distinct phases are separated by a second order phase transition with the critical temperature T_c of the order of $kT_c \sim J$. Figure 1 is a sketch of the behaviour of M (magnitude of the magnetization) as a function of temperature.

The *correlation function*[4] $G^{(2)}_{\alpha\beta}(\mathbf{r}_i - \mathbf{r}_j) \equiv \left\langle S_i^\alpha S_j^\beta \right\rangle$ depends on the distance between the lattice sites i, j. For generic $T > T_c$, the correlation function shows an exponentionally-decaying asymptotic behavior:

$$G^{(2)}_{\alpha\beta}(\mathbf{r}) \underset{|\mathbf{r}|>>a}{\sim} \frac{e^{-|\mathbf{r}|/\xi(T)}}{|\mathbf{r}|^{\#}} \tag{5}$$

with $\xi(T)$, the *correlation length*[5]. For generic temperatures much larger than T_c, $\xi(T) \sim a$ and the system is weakly correlated. It behaves as clusters of correlated degrees of freedom of typical size ξ that are almost uncorrelated from each other. [Thus, the Central Limit

[4] The *connected correlation function*, that we will see later on, is equal to the correlation function at zero external field and for temperatures above T_c.

[5] At distances of the order of the lattice spacing, the system does not show rotation invariance ($O(d)$-symmetry) and the correlation length depends in general on the direction of $\mathbf{r}$. When the system is close to the phase transition, $\xi \gg a$ and at large distances rotation invariance is effectively restored: The correlation length is no longer anisotropic and one can speak of a unique, direction-independent, correlation length.

Theorem (CLT) applies to the distributions of mean values: for example, the magnetization mode: $\sum_i S_i$ is gaussian distributed at large temperature (and field theory is useless)][could we refer to the average magnetization? what's the statement about field theory?, we have not introduced it yet, anyway]. On the other hand, for $T \simeq T_c$, the correlation length diverges as a power law:

$$\xi(T) \underset{T\to T_c^+}{\sim} (T - T_c)^{\nu} \tag{6}$$

with ν being a *critical exponent*. The system is in this case *strongly correlated* and the two-point correlation function decreases algebraically so that, up to very large distances, the spins cannot be considered independent. For $T < T_c$, and for $N > 1$, $N-1$ massless Goldstone bosons exist and the system remains strongly correlated even far away from T_c. For $N = 1$, the correlation length in the low temperature phase is finite and it is only around T_c that the system is strongly correlated.

B. The O(N) models in the continuum

Two models obtained in the continuum from the lattice $O(N)$ models will be interesting in the following: the $\boldsymbol{\varphi}^4$ and the non-linear sigma models. These two models belong to the same *universality class* (see section III below) as the lattice model. They are formulated in terms of either a N-component real field $\boldsymbol{\varphi} = (\varphi_1, \cdots, \varphi_N)$ or a $N-1$-component real field $\boldsymbol{\pi}$ both defined in the continuum.

1. The Non-linear sigma (NLσ) model.

Starting from (3), and using delta functions, we can rewrite the sum over the configurations of the spins as integrals over unconstrained real variables:

$$\mathcal{Z}[\mathbf{h}] = \prod_i \left(\int d\boldsymbol{\varphi}_i \delta(\boldsymbol{\varphi}_i^2 - 1) \right) \exp\Big\{ -\frac{H}{T} + \sum_j \mathbf{h}_j \cdot \boldsymbol{\varphi}_j \Big\}. \tag{7}$$

Here, we have now called $\mathbf{S}_i \to \boldsymbol{\varphi}_i$. Furthermore, we have trivially redefined H such that the temperature is shown explicitly in order to elucidate the low-temperature expansion below. Now, we can use the relation $\boldsymbol{\varphi}_i \cdot \boldsymbol{\varphi}_j = -{}^1\!/_2 (\boldsymbol{\varphi}_i - \boldsymbol{\varphi}_j)^2 + cte$ to take the continuum limit, which leads to kinetic terms in the action,

$$(\boldsymbol{\varphi}_i - \boldsymbol{\varphi}_j)^2 \underset{a\to 0}{\rightsquigarrow} (\partial_\mu \boldsymbol{\varphi})^2 \,, \tag{8}$$

hence,

$$\mathcal{Z}[\mathbf{h}] = \int \mathcal{D}[\boldsymbol{\varphi}] \delta(\boldsymbol{\varphi}^2(\mathbf{r}) - 1) \exp\Big\{ -\frac{1}{2T} \int_{\mathbf{r}} (\partial_\mu \boldsymbol{\varphi})^2 + \int_{\mathbf{r}} \mathbf{h} \cdot \boldsymbol{\varphi} \Big\}, \tag{9}$$

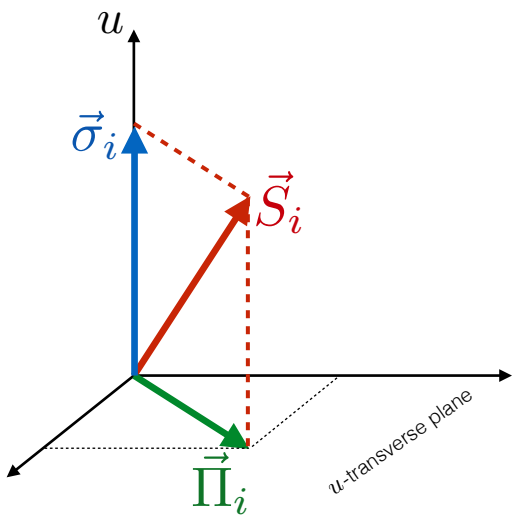


FIG. 2 Decomposition of the spin vector along the magnetization direction. The $N-1$-dimensional π vector corresponds to the Goldstone modes.

where

$$\int_{\mathbf{r}} = \int \mathrm{d}^d \mathbf{r}. \tag{10}$$

In the low temperature phase of the O($N > 1$) model in $2d$ spontaneous symmetry breaking occurs (O(N)→ O($N-1$)). Here, the $\boldsymbol{\varphi}(\mathbf{r})$ field can be decomposed into a longitudinal component $\sigma(\mathbf{r})$, that is, a component parallel to the direction $\mathbf{u}$ of the spontaneous magnetization and a transverse component $\boldsymbol{\pi}(\mathbf{r})$ such that $\mathbf{u} \cdot \boldsymbol{\pi}(\mathbf{r}) = 0$ (which implies that $\boldsymbol{\pi}(\mathbf{r})$ is a $(N-1)$-component field):

$$\boldsymbol{\varphi}(\mathbf{r}) = \sigma(\mathbf{r})\mathbf{u} + \boldsymbol{\pi}(\mathbf{r}) \quad \text{with} \quad \sigma^2(\mathbf{r}) + \boldsymbol{\pi}^2(\mathbf{r}) = 1 \,. \tag{11}$$

The projection is sketched in figure 2.

At low temperature and in the broken phase, $\langle \boldsymbol{\pi}^2(\mathbf{r}) \rangle$ is small and, after the rescaling $\boldsymbol{\varphi}(\mathbf{r}) \to \sqrt{T}\boldsymbol{\varphi}(\mathbf{r})$, one can take $\sigma(\mathbf{r}) = \sqrt{1/T - \boldsymbol{\pi}^2(\mathbf{r})}$ and neglect the constraint $|\boldsymbol{\pi}(\mathbf{r})| < 1$. In the presence of a magnetic field, $\mathbf{u}$ and $\mathbf{h}$ are parallel and the partition function reads:[6]

$$\mathcal{Z}[\mathbf{h}] = \int \mathcal{D}[\boldsymbol{\pi}] \exp\Big\{ -H[\boldsymbol{\pi}] + \int_{\mathbf{r}} h\sigma \Big\}, \tag{12}$$

with the hamiltonian $H[\boldsymbol{\pi}]$ of the $\boldsymbol{\pi}$ fields,

$$H[\boldsymbol{\pi}] = \frac{1}{2} \int_{\mathbf{r}} \Big[(\partial_\mu \boldsymbol{\pi})^2 + T \frac{(\boldsymbol{\pi}.\partial_\mu \boldsymbol{\pi})^2}{1 - T\boldsymbol{\pi}^2} \Big]. \tag{13}$$

[6] Let us notice that the validity of the NLσ model goes beyond the low temperature phase contrary to what could be believed from the derivation given here.

2. The φ^4 model.

We start from (3), and its re-writted (equivalent) form in the precedent section (7). We take again the continuum limit to obtain (9). Now, we change the delta into a smoothed (double) Gaussian term:

$$\delta(\boldsymbol{\varphi}^2 - 1) \to \exp\{-\lambda(\boldsymbol{\varphi}^2 - 1)^2\}. \tag{14}$$

The delta being a functional delta we obtain in this replacement a product at all points $\mathbf{r}$ of a double gaussian which means an exponential of the sum (in fact, of the integral) of $-\lambda(\boldsymbol{\varphi}^2-1)^2$. With this replacement, we obtain the $\boldsymbol{\varphi}^4$ model,

$$H[\boldsymbol{\varphi}] = \int d^d r \left\{ \frac{1}{2}(\boldsymbol{\partial}\boldsymbol{\varphi})^2 + \frac{r_0}{2}\boldsymbol{\varphi}^2 + \frac{u_0}{4!}(\boldsymbol{\varphi}^2)^2 \right\}, \tag{15}$$

$$\mathcal{Z}[\mathbf{h}] = \int \mathcal{D}[\boldsymbol{\varphi}] \exp\Big\{ -H[\boldsymbol{\varphi}] + \int_{\mathbf{r}} \mathbf{h}\cdot\boldsymbol{\varphi} \Big\}, \tag{16}$$

where u_0, r_0 are trivial functions of λ. The gradient term is reminiscent of the scalar product $\mathbf{S}_i \cdot \mathbf{S}_j$ of the lattice model. The potential term $\frac{r_0}{2}\boldsymbol{\varphi}^2 + \frac{u_0}{4!}(\boldsymbol{\varphi}^2)^2$ smoothly replaces the constraint $\mathbf{S}_i^2 = 1$.

Another, exact way to the $O(N)$ model defined in terms of an unconstrained vector is given by the *Hubbard-Stratonovich transformation*. The potential thus obtained behaves as $\log(\cosh(|\boldsymbol{\varphi}|)$. Once expanded at order 4 in the field this potential gives back the $\boldsymbol{\varphi}^4$ model. However, the true potential involves all powers and they are a priori all important if one wants to compute non-universal quantities, e.g. T_c.

C. The free energies (generating functionals)

In the following, we are mostly interested in the continuous models, in particular in the $\boldsymbol{\varphi}^4$ model. Thus, we define the quantities we used for these models. The generalization to lattice models is straightforward and we do not give them here.

We can construct two free energies from the partition function: The Helmoltz and the Gibbs free energy that are obtained from each other by a Legendre transform.

The Helmoltz free energy, $\mathcal{W}$, is a functional of the external source $\mathbf{h}$ and reads:

$$\mathcal{W}[\mathbf{h}] = \log \mathcal{Z}[\mathbf{h}], \tag{17}$$

where we have absorbed the factor $-k_B T$ in front of the $\log \mathcal{Z}[\mathbf{h}]$ term. $\mathcal{Z}[\mathbf{h}]$ is the generating functional of the correlation functions while $\mathcal{W}[\mathbf{h}]$ generates the *connected* correlation functions. For the one-point function:

$$\phi_i(\mathbf{r}) = \langle \varphi_i(\mathbf{r}) \rangle = \frac{\delta \mathcal{W}[\mathbf{h}]}{\delta h_i(\mathbf{r})}. \tag{18}$$

For a vanishing external source ($\mathbf{h} = 0$), $\boldsymbol{\phi}(\mathbf{r})$ is nothing but the order parameter. It vanishes in the high temperature phase while it is finite in the low temperature phase where spontaneous symmetry breaking occurs.

The second functional derivative of $\mathcal{W}[\mathbf{h}]$ is:

$$\frac{\delta^2 \mathcal{W}[\mathbf{h}]}{\delta h_i(\mathbf{r}) \delta h_j(\mathbf{r}')} = \langle \boldsymbol{\varphi}_i(\mathbf{r}) \boldsymbol{\varphi}_j(\mathbf{r}') \rangle - \langle \boldsymbol{\varphi}_i(\mathbf{r}) \rangle \langle \boldsymbol{\varphi}_j(\mathbf{r}') \rangle. \tag{19}$$

All connected correlation functions $G_c^{(n)}{}_{i_1 \dots i_n}[\mathbf{r}_1, \dots \mathbf{r}_n; \mathbf{h}]$ can be obtained from $\mathcal{W}[\mathbf{h}]$ by taking n functional derivatives w.r.t. the external source $h_{i_1}(\mathbf{r}_1), \dots, h_{i_n}(\mathbf{r}_n)$:

$$G_c^{(n)}{}_{i_1 \dots i_n}[\mathbf{r}_1, \dots, \mathbf{r}_n; h] = \frac{\delta^n \mathcal{W}[\mathbf{h}]}{\delta h_{i_1}(\mathbf{r}_1) \dots \delta h_{i_n}(\mathbf{r}_n)}. \tag{20}$$

Notice that, at this stage, the $G^{(n)}$ are really functional of the field $\mathbf{h}$. When evaluated in a uniform field configuration they become functions of the field momenta.

The Gibbs free energy – also called the effective action in the context of high energy physics – is obtained from $\mathcal{W}$ by a Legendre transform[7]

$$\Gamma[\boldsymbol{\phi}] + \mathcal{W}[\mathbf{h}] = \int_{\mathbf{r}} \mathbf{h}\cdot\boldsymbol{\phi}. \tag{21}$$

It is important to realize that in (21), the computation of $\Gamma[\boldsymbol{\phi}]$ requires to eliminate $\mathbf{h}$ for $\boldsymbol{\phi}$. This is achieved by inverting the equation of state (18) yielding $\boldsymbol{\phi} = \boldsymbol{\phi}[\mathbf{h}]$ to get $\mathbf{h} \equiv \mathbf{h}[\boldsymbol{\phi}]$.[8]

From (21) we obtain

$$\frac{\delta \Gamma[\boldsymbol{\phi}]}{\delta \phi_i(\mathbf{r})} = h_i(\mathbf{r}), \tag{22}$$

which is the reciprocal of (18). This relation shows that at vanishing external source the equlibrium states $\boldsymbol{\phi}$ are given by the minima of $\Gamma[\boldsymbol{\phi}]$.

The effective action $\Gamma[\boldsymbol{\phi}]$ is the generating functional of one-particle-irreducible (1PI) correlation functions (also called the vertex functions):

$$\Gamma^{(n)}_{i_1 \cdots i_n}[\mathbf{r}_1, \cdots, \mathbf{r}_n; \boldsymbol{\phi}] = \frac{\delta^n \Gamma[\boldsymbol{\phi}]}{\delta \phi_{i_1}(\mathbf{r}_1) \cdots \delta \phi_{i_n}(\mathbf{r}_n)}. \tag{23}$$

The connected correlation functions can be expressed in terms of the $\Gamma^{(n)}$'s. Let us show this for $n = 2$. Consider the equation of state (18) that we differentiate with

---

[7] Let us notice that the Legendre transform is often defined by: $\Gamma[\boldsymbol{\phi}] = \sup_{\mathbf{h}} \left(\int_{\mathbf{r}} \mathbf{h}\cdot\boldsymbol{\phi} - \mathcal{W}[\mathbf{h}] \right)$. For what follows, it is useless to compute the supremum with respect to $\mathbf{h}$ and we explain in the following how to proceed without this requirement.

[8] Notice that in the broken phase $\mathbf{h} = 0$ corresponds to infinitely many $\boldsymbol{\phi}$ fields having all the same modulus and the equation $\boldsymbol{\phi} = \boldsymbol{\phi}[\mathbf{h}]$ is thus ambiguous in the limit $\mathbf{h} \to 0$. However, even in this case, (22) remains well-defined.

respect to the field $\mathbf{h}$. We obtain

$$\begin{aligned}\delta(1-2) &= \frac{\delta}{\delta h(2)}\frac{\delta\Gamma[\boldsymbol{\phi}]}{\delta\phi(1)} \\ &= \int d3 \frac{\delta^2\Gamma[\boldsymbol{\phi}]}{\delta\phi(1)\delta\phi(3)}\frac{\delta\phi(3)}{\delta h_(2)} \\ &= \int d3 \frac{\delta^2\Gamma[\boldsymbol{\phi}]}{\delta\phi(1)\delta\phi_(3)}\frac{\delta^{(2)}\mathcal{W}[\mathbf{h}]}{\delta h(3)\delta h_(2)} \\ &= \int d3\, \Gamma^{(2)}_{i_1 i_3}[\mathbf{r}_1,\mathbf{r}_3;\boldsymbol{\phi}] G^{(2)}_{i_3 i_2}[\mathbf{r}_3,\mathbf{r}_2;\mathbf{h}] \qquad (24)\end{aligned}$$

(we use the notations $1 \equiv (\mathbf{r}_1, i_1)$, $\int d1 = \int d^d r_1 \sum_{i_1}$, etc.). In matrix form this relation is expressed as

$$\Gamma^{(2)} = {G_c}^{-1}\,. \qquad (25)$$

The inverse of the two-point function is the full (field-dependent) propagator, G_c. Note that here we drop the superscript (2) in G_c.

In the following, we will need this function evaluated in a uniform, that is, constant, field configuration $\boldsymbol{\phi}(\mathbf{r}) = \boldsymbol{\phi}$. In general, a uniform field configuration ensures translational invariance in position space. In momentum space this leads to momentum conservation in any n-point functions. Hence, an n-point function depends only on $(n-1)$ momenta. When $\boldsymbol{\phi}$ is a constant, we call ρ the $O(N)$ invariant

$$\rho = \boldsymbol{\phi}^2/2\,. \qquad (26)$$

The functions $\Gamma^{(n)}$ are called one-particle-irreducible (1PI) because their perturbative expansion involves only graphs that are 1PI. This means that they remain connected when any of their internal line is cut.

III. UNIVERSALITY

A natural question in strongly correlated systems is to know whether their statistical properties are insensitive to their small distance features[9].

Strikingly, some properties are indeed independent of the microscopic details. As a consequence, vastly different physical systems can exhibit similar behaviour close to criticality. According to this feature of *universality* these systems are said to be in the same *universality class*. Determining the specific universality class of the physics at hand is non-trivial. It has to be studied for each particular case. However, symmetries and dimensionality can provide a practical guideline.

A salient feature of the behavior of the thermodynamic quantities close to a second order phase transition is the fact that they are all power laws either as functions of the temperature, the size of the system or the length scale at which it is studied. The *critical exponents* of these power laws are among the best known universal quantities: the correlation length, specific heat and the correlator of the Ising model at T_c behave with the same power law as those of water around its critical point or that of the demixing (i.e. phase separation) transition: They lie in the same universality class.

For this set of systems, for $T \to T_c^+$, we have:

$$\xi(T) \sim (T - T_c)^{-\nu}\,, \qquad (27)$$

$$G^{(2)}(\mathbf{r}) \underset{\xi \gg |\mathbf{r}| \gg a}{\sim} \frac{e^{-|\mathbf{r}|/\xi(T)}}{|\mathbf{r}|^{d-2+\eta}} \sim \frac{1}{|\mathbf{r}|^{d-2+\eta}} \xrightarrow{F.T.} \frac{1}{q^{2-\eta}}\,, \qquad (28)$$

$$\chi(T) \equiv \frac{\partial M}{\partial h}\Big|_{h=0} \sim (T - T_c)^{-\gamma}\,, \qquad (29)$$

where $\chi(T)$ is the (magnetic, for Ising) susceptibility and η is the anomalous dimension. Other universal quantities at a second order phase transition are amplitude ratios.

On the other hand, properties such as the existence of a finite T_c (in the negative case, there should still exist a "transition" at $T = 0$) and its value, are non-universal but dependent on the microscopic details.

IV. MEAN FIELD THEORY

Landau's idea of mean field (MF) theory is to approximate the functional integral itself by proposing an ansatz for the effective action $\Gamma[\phi]$. MF theory constitutes a classical approximation of a quantum field theory in the sense that fluctuations around the expectation value of the field are not taken into account. Instead, only the classical field configuration is considered. Indeed, refering to the integrand (argument) of the functional integral, MF theory is also known as the method of steepest descent, the saddle-point approximation or simply the classical approximation (see next section).

The ansatz is constructed such that

- $\Gamma[\boldsymbol{\phi}]$ retains the $O(N)$ symmetry,
- $T \approx T_c$ is assumed and, hence, $\boldsymbol{\phi} = \langle\boldsymbol{\varphi}\rangle$ is small,
- $\Gamma[\boldsymbol{\phi}]$ derives from a density of Gibbs free energy, and

[9] The situation looks somewhat involved for strongly correlated systems especially when they show *damage spreading* (the reader is encouraged to look this term up). When it is the case, the system is chaotic and one could naively think that all microscopic details matter. Although true for the time evolution of the system, universality can, nevertheless, emerge for averaged quantities

- $\Gamma[\boldsymbol{\phi}]$ is analytic and we can expand and retain only the smallest terms, which corresponds to assuming only mild fluctuations around the mean field (MF) value.

The easiest non-trivial ansatz is given by an expansion up to fourth order in the field[10],

$$\begin{aligned}\Gamma[\boldsymbol{\phi}] &= \int d^d x \left\{ \frac{1}{2} \left(\partial_\mu \phi\right)^2 + U\left(\phi^2\right) \right\} \\ &\approx \int d^d x \left\{ \frac{1}{2} (\partial_\mu \boldsymbol{\phi})^2 + \frac{r_0}{2} \boldsymbol{\phi}^2 + \frac{u_0}{4!} (\boldsymbol{\phi}^2)^2 \right\} . \quad (30)\end{aligned}$$

Note that the $O(N)$ symmetry is satisfied by $U\left(\phi^2\right)$ and the original space-time dimensional symmetry is preserved by the kinetic term.

MF theory often serves as a starting point in the investigation of new models, because of its advantageous features: it does certainly exhibit phase transitions, shows universality and yields power laws as criticality is approached. However, as it stands, this approximation is often too crude to describe the physics at hand accurately. With regard to $O(N)$-models it incorrectly predicts that neither the existence of a phase transition, nor the values of the critical exponents, depend on N or d. Clearly, this is in contradiction with the $1d$ Ising model, which shows no phase transition, or the Mermin–Wagner theorem, which dictates that continuous symmetries cannot be broken spontaneously in 2 dimensions.

Contrary to our last hypothesis, the long wavelength excitations present in the vicinity of the 2nd order phase transition produce non-analiticities (in the thermodynamic limit), .

V. THE PERTURBATIVE RG

One way to improve the MF approach is to approximate the integrand of the functional integral $Z[h]$, instead of the effective action itself. We will consider a perturbative expansion around the dominating classical trajectory[11]. Hence, this procedure is usually refered to as perturbation theory (PT). The zeroth order of this approximation corresponds to the classical approximation or MF theory.

A convenient way to introduce PT follows: for the generating function defined in (16), we decompose the Hamiltonian into its Gaussian components,

$$H_0(\varphi) = \frac{1}{2} \int_x \left(\left(\partial_\mu \varphi\right)^2 + r_0 \varphi^2 \right) , \quad (31)$$

and Taylor expand[12] the interaction term H_I,

$$H = H_0 + H_I \approx H_0 + \int_x \frac{u_0}{4!} \varphi^4(x) . \quad (32)$$

The generating functional is now given as a series in the bare coupling, u_0. This allows for a computation of correlation functions of arbitrary order. The resulting expressions can be illustrated with Feynman diagrams. After the perturbative expansion, we still have to perform the functional integral itself, so that an actual sum over different field configurations is carried out, unlike in MF theory.

On its part, the idea of *perturbative renormalisation* is to include the effect of fluctuations which perturb the system mildly via a change of parameters defining the theory. Hence, we reparametrise $u_0 \to u_R$, $r_0 \to r_R$ (this coupling corresponds to the square of the mass, m_R, of the field ϕ) and $\varphi_0 \to \varphi_R$, as we know that the bare expressions are ill-defined (they lead to divergent terms in the UV cut-off). The reparametrisation leads to 'renormalised' correlation functions, $\Gamma_R^{(n)}\left(\{p_i\}, m_R, u_R\right)$, which are expressed in terms of the renormalised quantities and the momenta p_i carried by the external legs of the vertex.

Naturally, in a series expansion, it is necessary to check for its convergence. Actually, the perturbative expansion has bad properties, as it leads to asymptotic series at best[13]. As an example where standard PT predicts the wrong behaviour, we may consider the beta-function of the coupling of φ^4 theory in $d = 3$ (the beta function describes the change of the coupling under a change of renormalization scale). In fact, PT yields a divergent series. Nevertheless, in this case we can still apply the idea of a small perturbation around the free theory as we know the perturbative expansion up to high orders. Via a reordering of the series, a Borel resummation can be successfully employed: The resulting series of approximants is well-behaved and approaches the correct result with increasing expansion order.

[10] As the mean of φ vanishes at the phase transition, going to higher order in powers of ϕ does not improve the description of the critical phenomena.

[11] Note that the assumption of a small coupling is the major difference with the NPRG, which can be applied for strong couplings too.

[12] Note that the interchange of the functional integral with the series may yield (insurmountable) problems.

[13] The reason why perturbative expansions are applied is indeed a practical one, not a mathematical one: In many physical systems, PT has been successfully applied even to high orders in the coupling. Most prominently in QED the perturbative expansion yields astonishingly good agreement with experiments.

More generally, we can summarise the cases in which PT is expected to work or fail, respectively. The latter situations motivate a non-perturbative description in terms of the NPRG.
The perturbative approach is expected to work well, if either

- the coupling is small; in this case the results are supposably qualitatively and potentially even quantitatively good even at low order: This is the case of QED; or
- the coupling is not small, many terms of the perturbative expansion have been computed and the series is Borel-summable. In this case, the Borel transformed expansion can be resummed (e.g. by Padé approximants) and one can build a sequence of approximants that become more and more accurate. This is the case of the φ^4 theory in three dimensions where the series expansion of the β function is known at six loops.

In contrast, we can encounter the pitfalls of perturbation theory in the following cases:

- The system is strongly coupled and the series are either non Borel summable or not known at high enough orders to be resummed. This is the case of the non-linear σ-model in $d = 3$, where the perturbative renormalisation in $d = 2 + \epsilon$ with $\epsilon \to 1$ yields series expansion that cannot be resummed and that are almost useless quantitatively for the $O(N)$ models (for other non-linear σ-models, they can even be qualitatively wrong).
- The system is strongly coupled but the series are Borel summable and known at high enough orders to be resummed but the resummed series are quantitatively very far from the exact result. This is the case of the φ^4 theory in $d = 2$ ($N = 1$) where the critical exponents have been computed at five loops order but do not match the exact Onsager's results: The anomalous dimension found perturbatively is $\eta_{\text{PT}} = 0.145(15)$, and disagrees with the exact result, $\eta_{\text{exact}} = .25$. The mismatch can even be of a qualitative nature, e.g. the $O(4)$-model in $d = 2$ predicts an $\eta_{\text{PT}} \simeq 0.1$, but due to the Mermin–Wagner theorem there is no phase transition, and $\eta_{\text{exact}} = 0$.
- Another possible failure of perturbation theory may occur when genuinely non-perturbative phenomena exist, that is, when one computes physical quantities that are not Taylor expandable.[14] Such a situation is encountered in the RG context when a fixed point is not connected to the gaussian fixed point by any RG trajecory: Perturbation theory cannot find it since it amounts to an expansion of the RG flow around the gaussian fixed point.

In all of the latter cases the NPRG improves the description of or even solves the physics at hand.

VI. NPRG

The main difference between the non-perturbative and the perturbative formulations of the RG is that it does not rely on an expansion around the Gaussian model[15]. Again, the idea is to improve the (MF) approach of Landau, however, in this occasion, extending the RG idea of Wilson.

We will be interested in calculating the effective action $\Gamma[\boldsymbol{\phi}]$, (21). Doing so requires summing all the *fluctuations i.e.*, all possible values that the *fluctuating field* $\boldsymbol{\varphi}$ can assume.[16]

Before we embark on the derivation of the underlying NPRG equation, we outline the distinct scales that occur in the computation of the generating functional when starting from the microscopic interactions encoded in H. We will pick the example of the φ^4-model.

A. Scales

We introduce an artificial ultraviolet cutoff in momentum space, Λ. With this choice we only allow for modes with momenta $p < \Lambda$ to contribute. The choice of the cutoff is motivated by lattice field theory, where the lattice spacing, a, constitutes a smallest resolvable distance in position space. This is equivalent to the choice of a largest resolvable momentum, $\Lambda \sim 1/a$, due to the relation of the two representations by a Fourier-transformation. Dropping modes with momenta $p > \Lambda$ is exact in the discretised spacetime. In contrast to this, in the continuum, we can think of the effects from large momenta being absorbed into the definition of H within a renormalisation procedure.
By construction, when looking at the theory at the scale Λ we are not sensitive to the effects that arise from smaller momenta. Therefore, at this scale we can identify the free energy with the Hamiltonian of the

[14] As a toy model, consider the function $f(x) = \exp(-1/x)$. This function is not expandable around $x = 0$ since $f^{(}n)(0) = 0$ for all n and is therefore found to be 0 at all orders in the expansion in x.

[15] Hence, the NPRG does not suffer from the potential problems that arise in the interchange of the series and functional integral as performed in the perturbative RG.

[16] According to the physics under study, the fluctuations can be either quantum or statistical. We need not specify since the field theory formalism treats the two in the same way.

FIG. 3 Relevant scales between the microscopic dynamics given by the hamiltonian and the free energy.

system, i.e. $\Gamma|_{k=\Lambda} = H$. As already stressed, we want to compute the macroscopic theory, given by full effective action, Γ, from the microscopic dynamics. In the course of this computation we encounter particular scales that will dominate and which will make the properties of the system change qualitatively. We will now sketch the distinct scales and outline their effects.

The relevant scales are set by Λ, the couplings in the action (mass and coupling constant) and, for a finite-size system, the inverse of the size of the system, L^{-1}. For an illustration of the hierarchy of scales in momentum space see figure 3. Note that the identity $\Gamma = H$ holds only at the MF approximation and the different scales remain being given by the bare parameters, r_0 and u_0 only at this approximation. Beyond MF they have to be read off from the effective action. Notice also that bringing the system close to criticality requires to fine tune the bare parameters since, otherwise, the renormalized mass of the order of the large ultra violet scales.

With the finite scale Λ we have defined the resolution of our system in the ultraviolet[17]. For many systems, all bare length scales are given in terms of Λ which is the fundamental UV scale and, in these cases, all microscopic scales are of order Λ, e.g. $u_0^{1/(4-d)}$. (this is the case of the $O(N)$ model derived from the lattice ferromagnetic model with the Hubbard-Stratonovich transformation). It can happen that it is not so and that $u_0^{1/(4-d)}$ is independent of Λ and can therefore be much smaller (by orders of magnitude).

In order to avoid considering finite-size effects we take in the following $L \to \infty$.

Below Λ, the first smaller scale in momentum space is the Ginzburg scale. For the theory at hand it is defined by the coupling, $u_0^{1/(4-d)}$. As we aim at studying critical physics, the mass, which is the next distinct scale, has to be small compared to the other scales. Therefore, it must be well separated from the Ginzburg scale. Note that only in the limit $u_0 \to \infty$ and $\Lambda \to \infty$ we can have a scale invariant theory for $T \to T_c$ for all momenta.

The behaviour of the system is qualitatively different in the regions given in figure 3. As an example we study the two-point function, $\Gamma^{(2)}(p)$, defined in (25). We have implicitly assumed a uniform field configuration and hence, because of translation invariance, the correlator only depends on (the absolute value of) one momentum, p, see section II.C.

In region 3, $u_0^{1/(4-d)} < p \ (\ll \Lambda)$, the mass-scale is negligible and, thus, the system is dominated by the scale $u_0^{1/(4-d)}$. In this domain no universal scaling behaviour related to critical physics emerges and PT works well.

In region 2, $m \ll p \ll u_0^{1/(4-d)}$, we find (close-to-)critical behaviour: compared to the value of the momentum, the mass is negligible, and, at the same time, the Ginzburg scale is large. As a result, the propagator exhibits a scaling behaviour according to its anomalous dimension, cf. (28). In the critical domain MF fails because the contributions of fluctuations on all scales belonging to this region add up coherently and their effects become strong. PT resummed by means of the RG finds the right scaling behavior for the 2-point function with an accurate determination of η if computed at large orders (at least three loops).

In region 1, $p \ll m$ and, even though close to criticality, the system looks weakly correlated: the comparatively large mass supresses deviations from MF and non-analyticities do not show up. As a consequence, MF theory (with possibly perturbative corrections), $\Gamma^{(2)}(p) = c_1 \left(p^2 + m^2\right) + c_2\, p^4 + ...$, works well. In a certain sense, MF theory is taylored for this regime. Obviously, if $m \to 0$, this region vanishes and Landau's idea of MF theory is not applicable.

B. NPRG flow: General idea

In order to study the critical phenomena we have to remove the infrared regulator, i.e. we have to take the limit $m \to 0$, or equivalently $\xi \to \infty$ or $T \to T_c$. There are different ways to regularise the non-analyticities that show up in this procedure.

The first possibility is to consider a non-vanishing mass: We bring the system out of criticality, and then study the behaviour approaching the critical limit. This amounts to analysing the change of the system with respect to a change in the mass. In a differential formulation, i.e. taking the derivative $\frac{\partial}{\partial m}$, this leads to the Callan–Symanzik RG equation.

Another possibility is to put the system in a finite box of spatial extend $L < \infty$. By means of a scaling analysis with respect to L we can identify the behaviour as

[17] If necessary, we assume that an UV renormalisation has been performed and the renormalised parameters at scale Λ define the Hamiltonian.

the system approaches the critical limit, i.e. $L \to \infty$. By the nature of the simulational setup, this procedure is most conveniently applied in lattice Monte-Carlo computations.

The third way to regularise the infrared non-analyticities is realised in the NPRG formalism. The idea is to sum over the fluctuations existing on all scales bewteen Λ and 0 in a better way than perturbatively. To this aim, we construct a family of models that interpolate smoothly and in the most convenient way between Λ and 0[18]. As we will see later, the NPRG has the form of an evolution equation with respect to a momentum scale k, which we introduce as an artificial scale. The initial condition is set in the ultraviolet, $k = \Lambda$, where $\Gamma|_\Lambda = H$. By slightly lowering the scale k we sum over fluctuations between Λ and $\Lambda - dk$. These fluctuations modify the effective action and, by iterating this step we finally reach the limit $k \to 0$, where we all fluctuations on all scales have been taken into account. We are left with the (full) effective action, $\Gamma_{k=0} = \Gamma$. In summary, at a finite scale k, the running effective action Γ_k is a precursor of the effective action satisfying

$$\begin{cases} \Gamma_{k=\Lambda} = H \\ \Gamma_{k=0} = \Gamma \end{cases} . \tag{33}$$

Due to this interpolation the NPRG equation is also known as flow equation: The effective action flows through the momentum interval. The trajectory in this interpolation in between $k = \Lambda$ and $k = 0$ depends on the details of the way fluctuations are summed over (choice of regulator function) as we will see below. However, the limits (33) are unique. This method of considering only part of the interval in figure 3 in each step relates to Wilson's idea of integrating fluctuations momentum shells by momentum shells.

In order to specify the idea of the NPRG, we construct a deformation of the original model, Z_k, by adding a term to the Hamiltonian,

$$Z_k = \int \mathcal{D}\varphi e^{-H - \Delta H_k + \int_x h\cdot\varphi} . \tag{34}$$

The deformation, ΔH_k, can be chosen to be quadratic in the fields. In principle, also terms of higher order in the fields would be possible. However, the quadratic term is the easiest one: in this case the NPRG equation has one-loop structure, as we will see below. Thus, we define

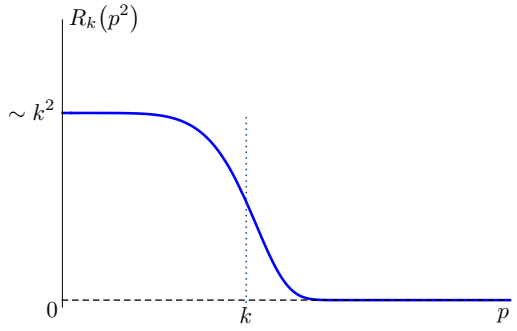


FIG. 4 Typical form of the regulator function $R_k(q)$: R_k vanishes for $q \gg k$, i.e. it does not modify the action at scales larger than k, and acts like an additional mass term in the infrared.

$$\begin{aligned} \Delta H_k &= \frac{1}{2}\int_x \varphi(x) R_k(x-y)\varphi(y) \\ &\xrightarrow{F.T.} \frac{1}{2}\int_q \varphi(q) R_k(q)\varphi(-q) , \end{aligned} \tag{35}$$

with the so-called regulator, $R_k(q)$, which can be chosen freely as long as it does not contradict the conditions imposed on Γ_k, (33). We choose the regulator as a momentum-dependent mass-like term that drives the system away from criticality when $k > 0$. In the limit, $k \to 0$, the regulator must not modify the theory at any scale. As a consequence, it must strictly vanish in this limit $\Delta H_{k=0} = 0$. These properties are met by taking the regulator of a shape similar to figure 4.[19] In fact, it is usually expressed as q^2 times a dimensionless function: $R_k(q) = q^2 r(q^2/k^2)$. The small value of $R_k(q)$ for $q^2/k^2 \gg 1$ reflects the fact that it does not change the theory at large momentum scales.

The modes with large momenta compared to k are refered to as 'rapid' modes, in analogy with the Brownian motion. Instead, the 'slow' modes in the infrared, $q \ll k$, are frozen by the regulator. For $k \sim \Lambda$ the regulator term is of the order of Λ^2 for all q and all fluctuations are frozen.

Remarks:

- In principle, one could take $R_k(p)$ to be a step function that vanishes for $p > k$ and takes a positive infinite value for $p < k$. This is known to be a ultra sharp cut-off and corresponds to the "block-spin" version of RG. In practice, it is a potential source

[18] Note that the NPRG provides an exact equation for the effective action, however, in practical applications approximations are inevitable. Hence, the 'most convenient' way of regularisation has to be chosen with respect to a given truncation scheme.

[19] Note that in the standard nomenclature the scale is written as an index, however, the regulator is a function of both k and momentum, $R_k(q) = R(q,k)$.

of difficulties when approximations are to be made and could lead to bad results when calculating, for example, the anomalous dimension.

- In the (original version of the) Callan–Symanzik RG approach one modifies the Hamiltonian by adding the a mass, $\int_x \Delta m^2 \varphi^2(x)$, which would correspond to $R_k(q) \equiv \Delta m^2$ in NPRG language. Therefore, it suppresses the fluctuations of all scales at once. In this sense one may see Wetterich NPRG as a refined version of Callan–Symanzik RG.

C. Wetterich's equation

As a next step we derive the equation that describes the evolution of Γ_k. Before doing this in detail, we shall precise the definition of Γ_k. Starting from the partition function of modified model $\mathcal{Z}_k$, see above (34, 35), we define as usual

$$\mathcal{W}_k = \log \mathcal{Z}_k \tag{36}$$

and its Legendre transform

$$\Gamma_k^{Leg}[\boldsymbol{\phi}] + \mathcal{W}_k[\mathbf{h}] = \int_{\mathbf{r}} \mathbf{h}\cdot\boldsymbol{\phi}\,. \tag{37}$$

However, we shall define Γ_k using the following modified Legendre transform

$$\Gamma_k[\boldsymbol{\phi}] + \mathcal{W}_k[\mathbf{h}] = \int_{\mathbf{r}} \mathbf{h}\cdot\boldsymbol{\phi} - \Delta H_k[\boldsymbol{\phi}]\,, \tag{38}$$

or, equivalently,

$$\Gamma_k[\boldsymbol{\phi}] = \Gamma_k^{Leg}[\boldsymbol{\phi}] - \Delta H_k[\boldsymbol{\phi}]\,. \tag{39}$$

This modification is necessary in order that Γ_k satisfies the limiting conditions (33). Indeed, at $k = 0$ no difference has been made: $\Gamma_0 = \Gamma_0^{Leg} = \Gamma$ since ΔH_0 vanishes; While at $k = \Lambda$, it is easy to see that $\Gamma_\Lambda^{Leg}[\boldsymbol{\phi}] \sim H[\boldsymbol{\phi}] + \Delta H_k[\boldsymbol{\phi}]$ (because at this scale, the MF approximation is almost exact, thanks to the regulator ΔH_k which presents a mass $m \sim \Lambda$ large enough to suppress all fluctuations); Therefore, only after subtracting $\Delta H_k[\boldsymbol{\phi}]$ would we have $\Gamma_\Lambda \sim H$ as desired [20]. Now, we state the cornerstone of the NPRG approach, the equation of Wetterich. This is a central piece of this course.

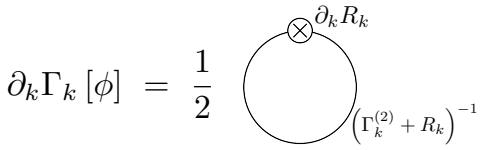


FIG. 5 Diagrammatic representation of Wetterich's equation (40). The line denotes the (full) propagator , $(\Gamma_k^{(2)} + R_k)^{-1}$, the crossed circle illustrates the derivative of the regulator function, $\partial_k R_k$.

Theorem (Wetterich) $\Gamma_k[\boldsymbol{\phi}]$ satisfies the following equation

$$\partial_k \Gamma_k[\boldsymbol{\phi}] = \frac{1}{2}\int_{x,y} \partial_k R_k(x,y)\left(\Gamma_k^{(2)}[\boldsymbol{\phi}] + R_k\right)^{-1}_{(x,y)} \tag{40}$$

where

$$\Gamma_k^{(2)}[\boldsymbol{\phi}](x,y) = \left.\frac{\delta^2\Gamma_k}{\delta\boldsymbol{\phi}(x)\delta\boldsymbol{\phi}(y)}\right|_{\boldsymbol{\phi}}$$

is the 1PI propagator of the modified effective action. $R_k(x,y) = R_k(x-y)$ is the regulator, and $\left(\Gamma_k^2 + R_k\right)^{-1}$ is the inverse of $\Gamma_k^{(2)}(x,y) + R_k(x,y)$ in the sense of integral kernel linear operator, i.e. $B(x,y) = A^{-1}_{(x,y)} \Leftrightarrow \int_z B(x,z)A(z,y) = \delta^d(x-y)\,.$

Its derivation is conceptually straightforward and we leave it to VI.C.2. Another, convenient form of the equation is:

$$\partial_k \Gamma_k = \frac{1}{2}\,\tilde{\partial}_k \,\mathrm{Tr}\log\left(\Gamma_k^{(2)} + R_k\right) \tag{41}$$

where $\tilde{\partial}_k$ acts only on the k-dependence of R_k and not on $\Gamma_k^{(2)}$:

$$\tilde{\partial}_k = \frac{\partial R_k}{\partial k}\frac{\partial}{\partial R_k}\,, \tag{42}$$

and the trace means integral over spatial or (momentum) indices (and for more complex theories, summation over any internal index).

Wetterich's equation (40) can be expressed in a diagrammatic form[21]: Under the integral the full propagator is multiplied by the derivative of the regulator function and in the language of diagrams this can be expressed as a closed loop representing the propagator with the insertion of a cross representing the derivative of the regulator. This is given in figure 5.

Below, we make several remarks on Wetterich's equation.

[20] In fact, it may be necessary for the equality $\Gamma_\Lambda = H$ to hold strictly, e.g. in the computation for non-universal quantities. Then, the regulator must diverge at Λ. As a result, Γ_Λ^{Leg} diverges also while Γ_Λ stays finite.

[21] Albeit their graphical similarities, the diagrammatic representations of flow equations do not involve Feynman diagrams, because the latter ones refer to perturbative quantities in a strict sense.

- Wetterich's approach can be generalised to $O(N)$ models with $N > 1$ and fermionic field theories without conceptual problems.

- Wetterich's equation (40), like any other RG equation, is to be interpreted as a dynamical system: The role of time is played by k which runs from Λ to 0[22]. In this sense, it must be completed by the initial condition (33).

- Non-polynomial forms of the initial condition H do not lead to technical difficulties contrary to the perturbative approach where it is difficult (and even often impossible) to deal with infinitely many interaction terms. For example, as already mentioned in section II.B.2, the potential of the Ising model has the complicated form $V(\varphi) \sim \log\cosh\varphi$. Obviously, a low-order polynomial expansion in PT fails to approximate this form and thus non universal quantities cannot be computed perturbatively at least in a reliable way.[23]

- It is a partial differential equation (PDE) in the sense that $\Gamma_k[\boldsymbol{\phi}] = \Gamma[k;\boldsymbol{\phi}]$ depends on two variables. As a result, one has to deal with difficulties intrinsic to PDE study (*e.g.* stability, reliability of the numerical solutions). Even worse, the right hand side of (40) is *non-linear* and is *functional* in the field variable. Nevertheless, it is better posed in mathematical terms than a path-integral in the continuum.

- The quantity $G_{c,k} = \left(\Gamma_k^{(2)}[q,\boldsymbol{\phi}] + R_k(q)\right)^{-1}$ is the full (functional) propagator at scale k (in particular, it is the full functional propagator of the theory for $k = 0$). This is made clear in the derivation VI.C.2.

- When the background field $\boldsymbol{\phi}$ is uniform, the (running) effective action $\Gamma_k[\boldsymbol{\phi}(x) = \boldsymbol{\phi}]$ becomes the (running) effective potential (up to volume factor) and the right hand side of (40) Fourier transforms into an integral over a single momentum. In the $O(N)$ case where $\Gamma_k^{(2)}$ is a scalar, the inverse of $\Gamma_k^{(2)}(q,\boldsymbol{\phi}) + R_k(q)$ boils down to $1/(\Gamma_k^{(2)}(q,\boldsymbol{\phi}) + R_k(q))$ and

$$\partial_k\Gamma_k[\boldsymbol{\phi}(x) = \boldsymbol{\phi}] = \frac{1}{2}\int_q \frac{\partial_k R_k(q)}{\Gamma_k^{(2)}(q,\boldsymbol{\phi}) + R_k(q)}.$$

The integrand is under control because of the presence of $\partial_k R_k$ in the numerator and of $R_k(q)$ in the denominator (see figure 4): Only the region where $q \sim k$ is really contributing[24]. This implements Wilson's RG idea of momentum shell integration of fluctuations.

- The (linearly realized) symmetries of the bare action are preserved by the RG flow as long as the regulator term ΔH_k is invariant. This is the case of the $O(N)$ symmetry for instance. In gauge theories, however, the regulator breaks gauge invariance, because it acts like a mass term in the infrared. In this case, the breaking of gauge invariance along the RG flow can be controlled by modified Slavnov–Taylor identities which take into account the regulator term[25].

- To make contact with PT, it is convenient to reintroduce the $\hbar$ factors and to expand ((41)) at the lowest, non trivial order in $\hbar$. Taking into account that the right-hand-side of the flow equation is proportional to $\hbar$, it is sufficient to replace $\Gamma_k^{(2)}$ by its mean-field approximation $H^{(2)}$. In this case, $\tilde{\partial}_k$ becomes ∂_k and thus:

$$\partial_k\Gamma_k = \frac{1}{2}\,\hbar\,\partial_k\,\mathrm{Tr}\log\left(H^{(2)} + R_k\right). \tag{43}$$

As a consequence, the flow becomes a total derivative and we can (trivially) integrate between 0 and Λ:

$$\Gamma - H = \frac{1}{2}\,\hbar\,\mathrm{Tr}\log\left(H^{(2)}\right) + \mathrm{const} + \mathcal{O}(\hbar^2). \tag{44}$$

This is the 1-loop result for the effective action. Equation (44) implies that any sensible approximation of Γ_k will lead to a flow equation which is at least 1-loop exact. Equation (43) can also be read the other way around: it is remarkable that substituting in this equation the bare inverse propagator $H^{(2)} + R_k$ by the full inverse functional propagator $\Gamma_k^{(2)} + R_k$ turns this one-loop equation into the exact flow equation.

- We shall see in the following that for $N \geq 2$, any sensible approximation of Γ_k leads to flows that are also one-loop exact in $d = 2 + \epsilon$, that is, the low-temperature expansion obtained from the non-linear σ model is also retrieved automatically *with*

[22] This is the direction to follow in statistical physics, i.e. where one tries to deduce the macroscopic behavior from the microscopic description of the system. To deal with other problems of physics, e.g. quantum gravity, it may be the other way around.

[23] The non polynomial character of the bare potential is actually not the only difficulty encountered perturbatively in the computation of non universal quantities.

[24] Indeed, for $q \gg k$, $R_k(q)$ almost vanishes and, hence, so does $\partial_k R_k(q)$. Moreover, the $R_k(q)$ term in the denominator regularizes the infrared divergences, if any.

[25] It is however non trivial to preserve the modified Slavnov–Taylor identities as well as unitarity when approximations are performed.

the same set of flow equations that are one-loop exact in $d = 4 - \epsilon$. This means that the critical exponent ν, for instance, computed with the simplest sensible ansatz for Γ_k (see the following) will automatically be one-loop exact around four *and* around two dimensions. We can therefore expect that the flow equations for the $O(N)$ models obtained from very simple approximations of Γ_k will be at least clever interpolations between a one-loop result in $d = 4 - \epsilon$ and another one-loop result in $d = 2 + \epsilon$. We show in the following that very accurate results can indeed be obtained with NPRG. Let us also emphasize that this property is very non-trivial (but perhaps in the limit $N \to \infty$) and is not recovered at any finite order of PT: even computed at large orders, the perturbative RG equations of the $\boldsymbol{\varphi}^4$ model for $N > 1$ are not able to reproduce Mermin-Wagner's theorem for instance and those obtained from the non-linear σ model cannot predict that four is the upper critical dimension of the $O(N)$ systems.

- In the same vein, we shall show that any sensible approximation of Γ_k leads to flows of the coupling constant that are exact in the limit $N \to \infty$. This implies in particular that the critical exponents obtained from any sensible approximation of the NPRG equations are exact in the limit $N \to \infty$.

In summary, the NPRG approach transforms a (quantum/statistical) field theory problem, which is usually formulated as a (functional) integral problem, into a (non-linear functional integro partial) differential equation. In the general form this equation is exact, *i.e.*, the integration of (40) with the initial condition $\Gamma_\Lambda = H$ gives the full solution $\Gamma = \Gamma_{k=0}$ of the theory. However, in practice approximations have to be employed. It is certainly possible to expand (40) around its Gaussian solution: this leads to PT. The key-point is to construct approximation schemes that go beyond PT. We shall discuss in the next section two such schemes: the derivative expansion (DE) and the Blaizot–Mendez–Wschebor (BMW) scheme. The rest of this section is devoted to the derivation of Wetterich's equation (40) and a historical remark on the Wilson-Polchinski approach to RG; neither of them is necessary to understand the rest of the lecture.

1. Derivation of (33): $\Gamma_{k=\Lambda} \sim H$

The meaning of the equation $\Gamma_{k=\Lambda} \sim H$ is that at scale Λ all fluctuations are (almost) frozen and thus the mean-field approximation is valid. Let us show this in more details.

We start from the definition of Z_k, (34), and of Γ_k, (38). We find:

$$h_x = \frac{\delta \Gamma_k}{\delta \phi_x} + \int_y R_k(x-y)\phi_y \ . \tag{45}$$

Thus, by substituting (38) and (45) into the defintion of W_k we obtain:

$$e^{-\Gamma_k[\phi]} = \int \mathcal{D}\varphi \exp\left\{ -H[\varphi] + \int_x \frac{\delta \Gamma_k}{\delta \phi_x} (\varphi - \phi)_x + \right. \\ \left. -\frac{1}{2} \int_{x,y} (\varphi - \phi)_x \, R_k(x-y) \, (\varphi - \phi)_y \right\} . \tag{46}$$

If we choose a function $R_k(q)$ that diverges for all q as $k \to \Lambda$ we find :

$$\exp\left(-\frac{1}{2} \int (\varphi_x - \phi_x) \, R_{k=\Lambda}(x-y) \, (\varphi_y - \phi_y) \right) \sim \delta(\varphi - \phi) \tag{47}$$

that is, it behaves as a functional Dirac delta. Therefore,

$$\Gamma_k[\phi] \to H[\varphi = \phi] \quad \text{as} \quad k \to \Lambda \, , \tag{48}$$

if the cut-off R_k is such that it diverges in this limit. If R_k does not diverge and is only very large,

$$\Gamma_{k=\Lambda} \approx H \ . \tag{49}$$

2. Derivation of (40)

We recall from (34, 35) that

$$\mathcal{Z}_k[\mathbf{h}] = \int \mathcal{D}\varphi \exp\left(-H - \Delta H_k + \int_x \mathbf{h} \cdot \varphi \right) .$$

Taking ∂_k of both sides, we have

$$\partial_k \mathcal{Z}_k[\mathbf{h}] = -\int \mathcal{D}\varphi \, \partial_k \Delta H_k[\varphi] e^{-H - \Delta H_k + \mathbf{h}\cdot\varphi},$$

where

$$\partial_k \Delta H_k[\varphi] = \frac{1}{2} \int_{x,y} \partial_k R_k(x,y) \varphi(x) \varphi(y)$$

since $\Delta H_k = \frac{1}{2} \int_x \varphi(x) R_k(x,y) \varphi(y)$, see (35). In terms of $\mathcal{W}_k = \log \mathcal{Z}_k$, this implies

$$\begin{aligned} &\partial_k \mathcal{W}_k[\mathbf{h}] \\ &= -\frac{1}{\mathcal{Z}_k[\mathbf{h}]} \int \mathcal{D}\varphi \int_{x,y} \frac{1}{2} \partial_k R_k(x,y) \varphi(x) \varphi(y) e^{-H - \Delta H_k + \mathbf{h}\cdot\varphi} \\ &= -\frac{1}{2} \int_{x,y} \partial_k R_k(x,y) \langle \varphi(x) \varphi(y) \rangle_{k,\mathbf{h}} \, , \end{aligned} \tag{50}$$

where $\langle \varphi(x)\varphi(y) \rangle_{k,\mathbf{h}}$ is the (full) 2-point correlator of the modified theory ($H \to H + \Delta H_k$) with external source. In terms of connected correlators,

$$\langle \varphi(x)\varphi(y) \rangle_{k,\mathbf{h}} = \langle \varphi(x) \rangle_{k,\mathbf{h}} \langle \varphi(y) \rangle_{k,\mathbf{h}} + G_{c,k}[x,y;\mathbf{h}]$$

where $G_{c,k}[x,y;\mathbf{h}] := \frac{\delta^2 \mathcal{W}_k}{\delta \mathbf{h}(x)\delta \mathbf{h}(y)}\Big|_{\mathbf{h}}$ is the connected 2-point function of the modified theory. Plugging back to (50) and noting that $\langle \varphi(x)\rangle_{k,\mathbf{h}} = \boldsymbol{\phi}(x)$, we have

$$\partial_k \mathcal{W}_k[\mathbf{h}] = -\frac{1}{2}\int_{x,y} \partial_k R_k(x,y)\left(\boldsymbol{\phi}(x)\boldsymbol{\phi}(y) + G_{c,k}[x,y;\mathbf{h}]\right) \tag{51}$$

To obtain the evolution of Γ_k we need to Legendre transform this equation. Notice that the left hand side $\partial_k \mathcal{W}_k[\mathbf{h}]$ needs more than a simple Legendre transform. Indeed, the partial derivative ∂_k is taken while keeping $\mathbf{h}$ fixed: $\partial_k \mathcal{W}_k[\mathbf{h}] = \partial_{k|\mathbf{h}} \mathcal{W}_k[\mathbf{h}]$ while in (40), $\partial_k \Gamma_k := \partial_{k|\boldsymbol{\phi}}\Gamma_k$ is taken keeping $\boldsymbol{\phi}$ fixed. They are not the same partial derivative but are related by $\partial_{k|\mathbf{h}} = \partial_{k|\boldsymbol{\phi}} + \int_x \partial_k \boldsymbol{\phi}(x) \frac{\delta}{\delta \boldsymbol{\phi}(x)}$. Therefore,

$$\begin{aligned}
\partial_k \mathcal{W}_k[\mathbf{h}] &= -\partial_{k|\mathbf{h}}\Gamma_k^{Leg}[\boldsymbol{\phi}] + \partial_{k|\mathbf{h}}\left(\int_x \mathbf{h}(x)\boldsymbol{\phi}(x)\right) \\
&= -\partial_{k|\boldsymbol{\phi}}\Gamma_k^{Leg}[\boldsymbol{\phi}] - \int_x \partial_k\boldsymbol{\phi}(x)\frac{\delta\Gamma_k^{Leg}[\boldsymbol{\phi}]}{\delta\boldsymbol{\phi}(x)} \\
&\quad + \int_x \mathbf{h}(x)\partial_k\boldsymbol{\phi}(x)\,.
\end{aligned}$$

But the two integrals cancel each other since $\delta\Gamma_k^{Leg}/\delta\boldsymbol{\phi}(x) = \mathbf{h}(x)$, so

$$\begin{aligned}
\partial_{k|\mathbf{h}}\mathcal{W}_k[\mathbf{h}] &= -\partial_{k|\boldsymbol{\phi}}\Gamma_k^{Leg}[\boldsymbol{\phi}] \\
&= -\partial_k\Gamma_k[\boldsymbol{\phi}] - \partial_k \Delta H_k[\boldsymbol{\phi}] \\
&= -\partial_k\Gamma_k[\boldsymbol{\phi}] - \frac{1}{2}\int_{x,y}\partial_k R_k(x,y)\boldsymbol{\phi}(x)\boldsymbol{\phi}(y)\,.
\end{aligned}$$

Comparing this with (51) we get

$$\partial_k\Gamma_k[\boldsymbol{\phi}] = \frac{1}{2}\int_{x,y}\partial_k R_k(x,y) G_{c,k}[x,y;\mathbf{h}]\,. \tag{52}$$

Now, since $\mathcal{W}_k$ and Γ_k^{Leg} are related by a Legendre transformation, we have

$$G_{c,k}[\mathbf{h}](x,y) = \left(\frac{\delta^2\Gamma_k^{Leg}}{\delta\boldsymbol{\phi}\delta\boldsymbol{\phi}}\Big|_{\boldsymbol{\phi}}\right)^{-1}_{(x,y)}\,.$$

Here, the inverse is in the integral kernal sense[26]. Finally, differentiating $\Delta H_k = \frac{1}{2}\int_x \varphi(x) R_k(x,y)\varphi(y)$, see (35), gives

$$\begin{aligned}
&\frac{\delta^2\Gamma_k^{Leg}}{\delta\boldsymbol{\phi}(x)\delta\boldsymbol{\phi}(y)} \\
=&\frac{\delta^2\Gamma_k}{\delta\boldsymbol{\phi}(x)\delta\boldsymbol{\phi}(y)} + \frac{\delta^2\Delta H_k}{\delta\boldsymbol{\phi}(x)\delta\boldsymbol{\phi}(y)} \\
=&\Gamma_k^{(2)}[\boldsymbol{\phi}](x,y) + R_k(x,y)
\end{aligned}$$

where $\Gamma_k^{(2)}[\boldsymbol{\phi}](x,y) = \frac{\delta^2\Gamma_k}{\delta\boldsymbol{\phi}(x)\delta\boldsymbol{\phi}(y)}\Big|_{\boldsymbol{\phi}}$. So

$$G_{c,k}[\mathbf{h}](x,y) = \left(\Gamma_k^{(2)}[\boldsymbol{\phi}] + R_k\right)^{-1}_{(x,y)} \tag{53}$$

Plugging this back to (52) gives (40).

VII. APPROXIMATION SCHEME I: THE DERIVATIVE EXPANSION

We will study two approximation schemes for the NPRG equation (40). The Blaizot–Mendez–Wschebor (BMW) method (Section VIII) at order n aims at computing the full momentum dependence of the functions $\Gamma_k^{(m)}(\mathbf{p}_1,\mathbf{p}_2,...)$ with $m \le n$. The derivative expansion (DE) is less ambitious and aims at describing the physics at zero external momentum only. This is enough to capture the thermodynamics of the system, e.g., critical exponents, phase diagrams and so on.

The scale k in the NPRG formalism acts as an infrared regulator (for $k \neq 0$) somewhat similar to a box of finite size $\sim k^{-1}$. Thus, for $k > 0$, there is no phase transition and thus no singularity in the free energy Γ_k, which can therefore be power-expanded safely[27]. We can therefore conclude that

(i) the singularities of Γ build up as k is lowered and are thus smoothened by k in Γ_k,

(ii) the precursor of the critical behavior should already show up at finite k for $|p| \gg k$. Here and below, p is momentum of the probe (say the argument in the two–point function $\Gamma_k^{(2)}(p)$).

An important consequence of the regularity of Γ_k at $k > 0$ is that it can be expanded in a power series of $\partial\phi(x)$. For slowly varying fields $\phi(x)$ this expansion is expected to be well-behaved (see the following for a discussion of the validity of the DE). This is the basis of the *derivative expansion* that consists in proposing an *ansatz* for Γ_k involving only a finite number of derivatives of the field.

The underlying idea is that we are mostly interested (for the study of critical phenomena)[This seems inconsistent with something said in a previous paragraph: the precursor of critical behaviour shows up for $|p| \gg k$, and here we are saying that we only care about $|p| \to 0$ for the study of critical phenomena] [Stricto sensu we should first take $k \to 0$ and then $p \to 0$ but it is too hard by (ii) above. So let us assume something about the double limit

[26] Reminder: This is a general property of Legendre transform. If $f(x)$ and $g(p)$ are Lengendre transform of each other, *i.e.*, $p = f'(x)$ and $x = g'(p)$ are inverse to each other, $f''(x) = \mathrm{d}p/\mathrm{d}x$ and $g''(p) = \mathrm{d}x/\mathrm{d}p$ so $f''(x) = (g''(p))^{-1}$.

[27] There can actually exist nonanalyticities showing up in the flow at finite k that do not result from the existence of a second order phase transition. This is what occurs for instance in the Random Field Ising Model at sufficiently low dimension or in the Pair Contact Process with Diffusion.

$(p,k)\to 0$(scaling hypothesis enters here) and approach it by first taking $p\to 0$ and then $k\to 0$. Yes if we are not careful or nature is not kind (say letting scaling hypothesis break down) we may miss completely the point by reversing limit order] in the long distance physics, that is the $|p|\to 0$ region of the correlation functions. Thus, we keep only the lowest orders of the expansion of Γ_k in $\partial\phi$ while we keep, for the moment being, all orders in the field ϕ. For the Ising case:

$$\Gamma_k = \int d^dx \left(U_k(\phi(x)) + \frac{1}{2} Z_k(\phi(x)) \left(\partial\phi\right)^2 \right) + \mathcal{O}(\partial^4) \tag{54}$$

The term U_k is the potential. The function $Z_k(\phi)$ is called the field renormalisation or wave-function renormalisation function[28]. When $k\to 0$, U_k will become the *effective potential*.

For several fields in the $O(N)$ models, because of the $O(N)$ symmetry, any function of local fields can only depends on its ($O(N)$-invariant) norm

$$\rho(x) := \frac{1}{2}\boldsymbol{\phi}(x)^2 = \frac{1}{2}\sum_{i=1}^{N} \phi_i(x)\phi_i(x)\,. \tag{55}$$

The approximation of Γ_k at order ∂^2 of the DE is

$$\Gamma_k = \int d^dx \left(U_k(\rho) + \frac{1}{2} Z_k(\rho) \left(\partial\boldsymbol{\phi}\right)^2 + \frac{1}{4} Y_k(\rho) \left(\boldsymbol{\phi}.\partial\boldsymbol{\phi}\right)^2 + O(\partial^4) \right). \tag{56}$$

Notice that the term with coefficient $Y_k(\rho)$ is specific to $O(N)$-models with $N\neq 1$: when $N=1$ it can be absorbed into the term $\sim Z_k(\rho)\dots$.

To derive flow equations for the coefficients $(U_k, Z_k, \dots)$, we need to read them off from Γ_k. As it turns out, they can be derived from Γ_k or its functional derivatives evaluated at a uniform field configuration (with given norm ρ), as we can see below. Moreover, due to $O(N)$-symmetry, the 'direction' of the constant vector in field-space is ad libitum. A particularly simple choice is to take only the zero-component in field-space being non-vanishing, i.e,

$$\boldsymbol{\phi}_{unif.}(\rho) := \begin{pmatrix} (2\rho)^{1/2} \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \text{with} \quad \rho = \text{constant}\,. \tag{57}$$

- Evaluating Γ_k directly in a uniform field configuration, (57), kills all ∂'s in (56) and yields

$$\Omega_d U_k(\rho) = \Gamma_k[\boldsymbol{\phi}_{unif.}(\rho)] \tag{58}$$

with Ω_d the volume of space-time.

- The DE ansatz (56) implies that the 1PI 2-point function in the uniform configuration (57) becomes[29]

$$\Gamma^{(2)}_{k,ij}(q) = U_k'(\rho)\delta_{ij} + \phi_i\phi_j U_k''(\rho) + Z_k(\rho)q^2\delta_{ij} + \frac{1}{2}Y_k(\rho)\phi_i\phi_j q^2\,. \tag{59}$$

For components $i=j>2$ in the above equation, only the first and the third terms contribute. Finally, deriving with respect to q^2 and evaluating at $q=0$, we extract Z_k as the (running) coefficient of the quadratic term in q:

$$Z_k(\rho) = \partial_{q^2}\, \Gamma^{(2)}_{k,jj}[\boldsymbol{\phi}_{unif.}(\rho)](q)\Big|_{q=0}\,. \tag{60}$$

Effectively, the derivative with respect to q^2 removes the terms of lower order in momentum, whereas by evaluation at vanishing momentum the terms of higher powers are neglected. Note that the above is valid up to any order in the DE.

- Generalisation to further coefficients follows the same recipe.

At this point, the derivation of a flow equation for any of the coefficients $U_k(\rho)$, $Z_k(\rho)$, etc. is straightforward (although, perhaps, tedious). For example, for U_k: take ∂_k of (58), apply Wetterich's equation (40) to $\partial_k\Gamma_k$ at the right hand side, and apply the ansatz (56) to the right hand side of Wetterich's equation. This yields

$$\partial_k U_k = \frac{1}{2}\int_q \partial_k R_k(q) \Big((N-1)\left(U_k' + Z_k q^2 + R_k(q)\right)^{-1} + \left(U_k' + 2\rho U_k'' + Z_k q^2 + Y_k\rho q^2 + R_k(q)\right)^{-1} \Big)\,. \tag{61}$$

We show the derivation in subsection VII.A.1. As can be imagined, the calculation of $\partial_k Z_k$ is more involved and we will again relegate it to subsection VII.A.2. The set of flow equations for U_k, Z_k (and Y_k) is a closed system constituting the DE truncation of Wetterich's equations up to order $O(\partial^4)$. We emphasise that the procedure involves no conceptual difficulty and can be automatized and generalized to higher derivative orders.

The remaining of this section is organised as follows. VII.A explains technical details in the derivation of flow equations in DE at order $O(\partial^4)$ and can be skipped without affecting the understanding of the rest of the lecture. VII.B introduces the Local Potential Approximation (LPA) and discusses its application to the Ising

[28] The function $Z_k(\phi)$ has, of course, nothing to do with the partition function $Z_k[h]$, although it is customary to use the same symbol for both functions.

[29] For the derivation of it see VII.A.1, (64).

model. VII.C discusses an important refinement of LPA that is related to the anomalous dimension. We conclude with a brief discussion of field expansion approximations on top of the DE, in VII.F.

A. Derivation of flow equations

1. Flow equation for U_k, (61)

Recall the DE ansatz (56) in terms of the $O(N)$-invariant (55),

$$\Gamma_k = \int d^d u \left(U_k(\rho) + \frac{1}{2} Z_k(\rho) \left(\partial \boldsymbol{\phi}\right)^2 + \frac{1}{4} Y_k(\rho) \left(\boldsymbol{\phi}.\partial\boldsymbol{\phi}\right)^2 + O(\partial^4) \right) , \tag{62}$$

where $\rho = \rho(u) = \frac{1}{2}\sum_l \phi_l(u)\phi_l(u)$. We now calculate its order–2 functional derivatives and evaluate in a uniform field configuration: Therefore, the only way for a term containing two $\partial\phi$ factors to survive is that each of them is "hit" by one functional derivative. So we can proceed as if $Z_k(\rho)$ and $Y_k(\rho)$ were field independent, which simplifies the calculation considerably. With this in mind, we formally apply the functional derivative $\delta\phi_l(u)/\delta\phi_i(x) = \delta_{il}\delta^d(x-u)$ and usual rules of calculus (in particular linearity and the chain rule) to each term of the integrand in (56)

$$\begin{aligned}
\left.\frac{\delta^2 U_k(\rho(u))}{\delta\phi_i(x)\delta\phi_j(y)}\right|_{\boldsymbol{\phi}_{unif.}} &= \left(U_k'\delta_{ij} + U_k''\phi_i\phi_j\right)\delta_{ux,uy} \\
\left.\frac{\delta^2 Z_k(\rho)\partial\boldsymbol{\phi}(u).\partial\boldsymbol{\phi}(u)}{2\delta\phi_i(x)\delta\phi_j(y)}\right|_{\boldsymbol{\phi}_{unif.}} &= Z_k(\rho)\delta_{ij}\partial_u\delta_{ux}\partial_u\delta_{uy} \\
\left.\frac{\delta^2 Y_k(\rho)(\boldsymbol{\phi}(u).\partial\boldsymbol{\phi}(u))^2}{4\delta\phi_i(x)\delta\phi_j(y)}\right|_{\boldsymbol{\phi}_{unif.}} &= \frac{1}{2}Y_k(\rho)\phi_i\phi_j\partial_u\delta_{ux}\partial_u\delta_{uy} ,
\end{aligned} \tag{63}$$

with $\delta_{xy} = \delta^d(x-y)$ and $\delta_{ux,yu} = \delta_{ux}\delta_{yu}$. Then we integrate each of them with respect to u[30] and sum the individual terms to obtain

$$\begin{aligned}
\left.\frac{\delta^2\Gamma_k}{\delta\phi_i(x)\delta\phi_j(y)}\right|_{\boldsymbol{\phi}_{unif.}} &= \left(U_k'(\rho)\delta_{ij} + U_k''(\rho)\phi_i\phi_j\right)\delta_{xy} \\
&\quad - \left(Z_k(\rho)\delta_{ij} + \frac{1}{2}Y_k(\rho)\phi_i\phi_j\right)\partial_x^2\delta_{xy}
\end{aligned}$$

Taking Fourier transform gives

$$\begin{aligned}
\Gamma^{(2)}_{k,ij}(q) &= U_k'(\rho)\delta_{ij} + \phi_i\phi_j U_k''(\rho) \\
&\quad + Z_k(\rho)q^2\delta_{ij} + \frac{1}{2}Y_k(\rho)\phi_i\phi_j q^2 .
\end{aligned} \tag{64}$$

[30] In the last two cases we need to integrate by parts: Since ρ, ϕ_i, ϕ_j are norm and components of $\boldsymbol{\phi}_{unif.}$ so are u independent they reduce to the same calculation $\int_u \partial_u\delta^d(u-x)\partial_u\delta_{uy} = -\int_u \partial_u^2\delta_{ux}\delta_{uy} = -\int_u \partial_x^2\delta_{ux}\delta_{uy} = -\partial_x^2\delta_{xy}$.

Thanks to $O(N)$ symmetry, we can take $\boldsymbol{\phi}_{unif.}$ to be $\boldsymbol{\phi}_{unif.}(\rho) = (\sqrt{2\rho}, 0, \ldots, 0)^T$. We assume also that the regulator is $O(N)$ invariant, *i.e.*, a scalar matrix $R_{k,ij}(q) = R_k(q)\delta_{ij}$. Then $\left(\Gamma_k^{(2)}(q) + R_k(q)\right)_{ij}$ becomes a diagonal matrix whose only non-zero entries are

$$\begin{aligned}
\left(\Gamma_k^{(2)}(q) + R_k(q)\right)_{11} &= U_k'(\rho) + 2\rho U_k''(\rho) + \\
&\quad + q^2(Z_k(\rho) + \rho Y_k(\rho)) + R_k(q), \\
\left(\Gamma_k^{(2)}(q) + R_k(q)\right)_{jj,j>1} &= U_k'(\rho) + q^2 Z_k(\rho) + R_k(q) .
\end{aligned} \tag{65}$$

Plugging its inverse into the right hand side of Wetterich's equation (40) gives

$$\begin{aligned}
&\partial_k\Gamma_k[\boldsymbol{\phi}_{unif.}(\rho)] = \\
&= \frac{\Omega_d}{2}\int_q \partial_k R_k(q) \left\{ (N-1)\left(U_k' + Z_k q^2 + R_k\right)^{-1} \right. \\
&\qquad \left. + \left(U_k' + 2\rho U_k'' + Z_k q^2 + Y_k\rho q^2 + R_k\right)^{-1} \right\} .
\end{aligned} \tag{66}$$

According to (58) the left hand side is just $\Omega_d\partial_k U_k(\rho)$, so we have obtained (61).

2. Flow equation for Z_k

Since Z_k is read off from the 1PI 2-point function, see (60), we need to calculate $\partial_k\Gamma_k^{(2)}$ to obtain its flow equation. We shall do this from Wetterich's equation in its compact form (41)

$$\Gamma_k[\boldsymbol{\phi}] = \frac{1}{2}\mathrm{Tr}\left(\partial_k R_k G_k[\boldsymbol{\phi}]\right) \tag{67}$$

where

$$G_k[\boldsymbol{\phi}] := \left(\Gamma_k^{(2)}[\boldsymbol{\phi}] + R_k\right)^{-1} ,$$

where its diagrammatic representation, see figure (5), will proove particularly useful.

We shall regard G_k, R_k and $\Gamma_k^{(2)}$ as "super" matrices whose rows and columns are indexed by $\{(p,i) : p \in \mathbf{R}^d, i = 1, \ldots, N\}$ (and for which multiplication or contraction means integrating over p and summing over i) and apply formally calculus for finite matrices. For example, we can apply $\partial(A^{-1}) = -A^{-1}(\partial A)A^{-1}$ to derive $G_k[\boldsymbol{\phi}]$ respect to $\phi_i(p)$:

$$\frac{\delta G_k}{\delta\phi_i(p)} = -G_k\frac{\delta\Gamma_k^{(2)}}{\delta_i(x)}G_k = -G_k\Gamma^{(3)}_{i,k}G_k. \tag{68}$$

Similarly, we can take functional derivative of (67) to obtain the flow of one-point function:

$$\partial_k\frac{\delta\Gamma_k}{\delta\phi_i(p)} = -\frac{1}{2}\mathrm{Tr}\left(\partial_k R_k\, G_k\Gamma^{(3)}_{i,k}G_k\right) , \tag{69}$$

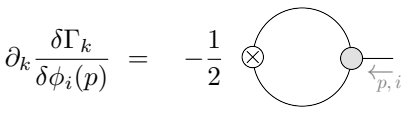


FIG. 6 Diagrammatic representation of the flow equation for the one-point function, (69). The shaded blob denotes the three-point vertex, $\Gamma^{(3)}$.

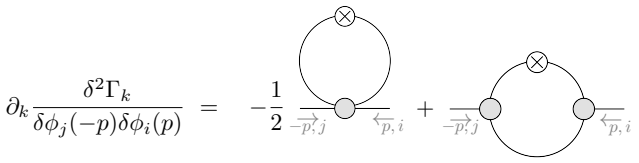


FIG. 7 Diagrammatic representation of the flow of the two-point function, cf. (70). Shaded blobs denote vertex functions, $\Gamma^{(3)}$ and $\Gamma^{(4)}$.

which is given diagrammatically in figure 6.

Now taking another functional derivative yields the flow equation for the two-point functions, $\Gamma^{(2)}_{k,ij}$. This time, the functional derivative can hit one of the two propagators, hence, relation (68) can be applied again. This results in diagram involving two three-point functions, where the overall-sign is positive and the combinatorical factor of 2 emerges from symmetry. In addition, the derivative can act on the $\Gamma^{(3)}_k$ to give a four-point function, $\Gamma^{(4)}_k$. As a consequence, there are two distinct diagrams contibuting to the flow of the two-point function. Omitting the $O(N)$-indices and momenta for brevity, the equation is schematically given by

$$\Gamma^{(2)} = -\frac{1}{2}\mathrm{Tr}\left(\Gamma^{(4)}_k G_k R_k G_k\right) + \mathrm{Tr}\left(\Gamma^{(3)}_k G_k \Gamma^{(3)}_k G_k R_k G_k\right), \tag{70}$$

which is diagrammatically represented in figure 7. We can read off from each of the diagrams the corresponding contribution to $\Gamma^{(2)}_k(p)$. For example, the second diagram gives the term

$$\sum \int_q \Gamma^{(3)}_k(\{p,i\},\{-q,i'\},\{q-p,i''\})$$
$$\Gamma^{(3)}_k(\{-p,j\},\{q,j'\},\{p-q,j''\})$$
$$G_{k,i'j'}(q) G_{k,i''l}(q-p) G_{k,lj''}(q-p)$$
$$\partial_k R_{k,ll}(q-p).$$

Here $\sum$ denotes a summation over all indices other than i and j. (Note that in order to preserve $O(N)$ symmetry, R_k is always taken to be a diagonal(indeed scalar) with respect to inner indices. Hence we sum only over $R_{k,ll}$.)

So far we have been doing *exact* calculation for $\Gamma^{(2)}(q)$. Now we shall replace Γ by its DE ansatz (56) and then invoke (60) to extract the flow of Z_k.

B. Local potential approximation: Application to the Ising model

The LPA further simplifies the DE ansatz (56) by setting $Z_k \equiv 1$ and $Y_k \equiv 0$:

$$\Gamma_k = \int d^d x \left(U_k(\phi(x)) + \frac{1}{2}(\partial\phi)^2\right) \tag{71}$$

Then (61) is the only RG flow equation that remains in the hierarchy. In the case of the Ising model ($N = 1$), it simplifies further to

$$\partial_k U_k = \frac{1}{2}\int_q \partial_k R_k \left(q^2 + R_k + U'_k + 2\rho U''_k\right)^{-1} \tag{72}$$

Its initial condition $U_{k=\Lambda}$ is the potential term in the Hamiltonian of the Ising model. As discussed in a previous section (see (14) in II.B.2), a sensible choice can be

$$U_\Lambda(\rho) = \frac{\lambda_\Lambda}{2}(\rho - \rho_{min})^2. \tag{73}$$

It is now a good time to discuss the choice of regulator $R_k(q)$. In the *exact* NPRG equation (40), the choice of $R_k(q)$ only affects the path of the flow $\Gamma_\Lambda \rightsquigarrow \Gamma_0$ but not its endpoints. This will no longer be the case when approximations are made, and the dependence of Γ_0 on R_k becomes in turn a good indicator of the error provoked by the approximation being made. We draw this idea in figure 8. More generally speaking, for any physical

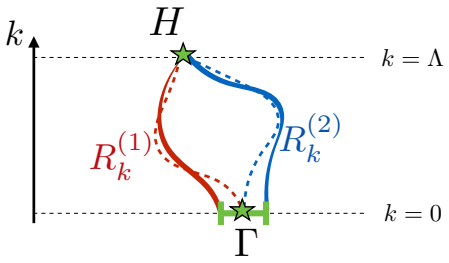


FIG. 8 Dependence of the RG flow on the regulator choice, with or without (dotted lines) approximation.

quantity Q, we should have

$$\frac{\delta Q}{\delta R_k} = 0$$

in the exact case. So as a thumb rule, we should *minimise* such dependence when making approximations. This is called the *principle of minimal sensitivity* (PMS). Another guiding rule is the *principle of fastest apparent convergence* (PFAC), which states that approximation schemes should be optimised so that higher order corrections are the smallest possible (compared to previous

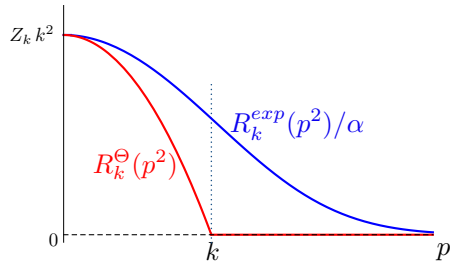


FIG. 9 Optimised and exponential regulators, (74) and (75).

terms). In general, the two rules are expected to be consistent with each other: The contrary indicates usually that the approximation is problematic.

Back to (72), we propose two choices of R_k (see figure 9 for plots), which are both known to yield quantitatively accurate critical exponents and are, furthermore, convenient for analytic and numerical applications:

$$R_k^{\Theta}(q) := Z_k(k^2-q^2)\theta(k^2-q^2)\,, \tag{74}$$

$$R_k^{exp}(q) := \alpha Z_k \frac{q^2}{e^{q^2/k^2}-1}\,. \tag{75}$$

Here θ denotes the Heaviside step function, α in (75) is a free parameter to optimise the regulator (according to the thumb rules discussed above). Note that, although $Z_k = 1$ in the LPA, we have kept Z_k explicit on the above expressions to use them in the following section, where the LPA' will be discussed.

The regulator R_k^{Θ} is particularly convenient because it allows evaluating the momentum integral in (72) easily:

$$k\partial_k U_k = \frac{4v_d}{d}\,\frac{k^d}{1+(U'+2\rho U'')k^{-2}}\,, \tag{76}$$

with a dimension-dependent constant[31] Let us now show and comment snapshots of the evolution flow of U_k that can be obtained by integrating the PDE (76), and possible scenarios that can be obtained by varying parameters in the initial condition (73). Note that the temperature enters via the value of the bare parameters.

We have already mentioned the assumption that we start with a "Mexican hat"-potential as initial condition, see figure 10. Note that we will tipically refer to one of its minima, as they are all equivalent. At the beginning of the flow, $k = \Lambda - \epsilon$, the integration of fluctuations puts more disorder to the system and thus makes the minima of U_k decrease towards $\phi = 0$ (we are considering a zero average magnetization as the measure of 'total disorder').

[31] The volume factor v_d stems from the variable transformation $x = q^2$ in loop-integrals,

$$\int d^dq\, f(q) = 2v_d \int_0^\infty dx\, x^{\frac{d}{2}-1} f(x)\,. \tag{77}$$

Obviously, this relation simplifies integrations where rotational symmetry only allows for quadratic dependence on momentum. v_d.

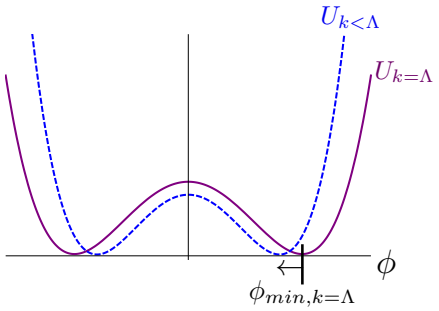


FIG. 10 Initial condition for the flow of the potential $U_k(\phi)$. Fluctuations put more disorder in the system and, hence, the minima tend to smaller (absolute) values in the field.

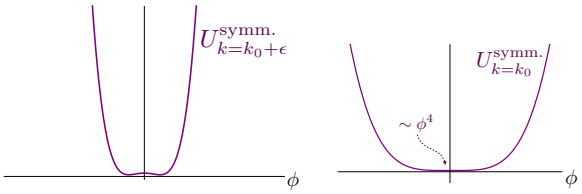


FIG. 11 Degeneration of the minima of the effective potential at a non-vanishing scale $k_0 > 0$ in the case of symmetry restoration.

We now sketch the three distinct, possible scenarios corresponding to the temperature being above, below or precisely at its critical value:

- Figure 11 illustrates the case, where the minima of U_k become degenerate at $\phi = 0$ for a non-vanishing scale, $k_0 > 0$. This relates to the transition from an ordered to a disordered regime. Fluctuations from lower scales $k < k_0$ increase the disorder in the system further. As a result, at the end of the flow the minimum of the potential remains at zero field expectation value, see figure 12. This means that we are in the high temperature phase $T > T_c$. The scale k_0 is related to the correlation scale $k_0 \sim \xi^{-1}$.

- Another possibility is that the minima of U_k are again driven towards smaller field expectation values, however, they never reach $\phi = 0$. That means, that below a scale k_0 the location of the minima remain fixed. From the scale k_0 and below, it is the value of the potential for $|\phi| < \phi_{min}$ that decreases, so that U_0 becomes convex in the limit $k \to 0$, i.e. $U_{k=0}$ is completely flat for $|\phi| < \phi_{min}$. This shape of the potential signals the low temperature phase $T < T_c$. Here, ϕ_{min} is the spontaneous magnetisation. This situation is depicted in figure 13.

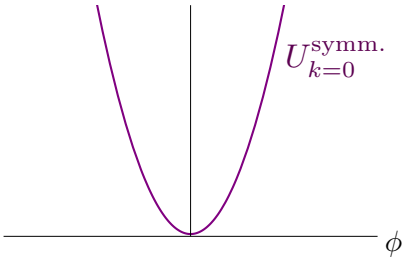


FIG. 12 Shape of the effective potential $U_{k=0}$ in the symmetric phase.

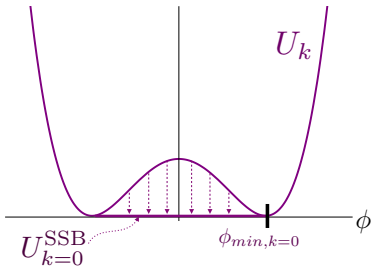


FIG. 13 Flow of the effective potential with $\phi_{min} > 0$ in the case of spontaneous symmetry breaking.

- The third distinct case is the point for which the temperature is precisely at the critical value, $T = T_c$. In this case the minima of the potential vanish for $k = 0$, $\phi_{min,k=0} = 0$. However, for every non-vanishing scale the potential has a non-vanishing expectation value, i.e. $\forall \epsilon > 0 : \phi_{min,k=\epsilon} > 0$. The fact that criticality can be seen only at $k = 0$ makes practical applications intricate. We discuss these issues in subsection VII.D.

C. LPA' and anomalous dimension

As was already mentioned, *LPA*' is a refinement of LPA. It allows the field normalisation Z_k to flow, but independently of the local field value

$$\Gamma_k = \int d^d x \left(U_k(\phi(x)) + \frac{1}{2} Z_k \left(\partial\phi\right)^2 \right) \tag{78}$$

It is also called the $O(\partial^2)$ approximation or the leading approximation. Recall that in the DE ansatz (56), the expression (60) is used to extract $Z_k(\rho)$ from Γ_k. Now in LPA', $Z_k(\rho)$ is taken to be independent of ρ and there is a choice to make. We shall adopt the following choice:

$$\begin{aligned} &Z_k(\rho) \rightsquigarrow Z_k(\rho_{min}) \\ &\left.\frac{\partial U_k(\rho)}{\partial \rho}\right|_{\rho_{min}} = 0. \end{aligned} \tag{79}$$

where the second equation determines implicitly ρ_{min}.

Remark. As a general rule, when we Taylor expand a function of ρ, we do it around ρ_{min} where $U_k(\rho)$ attains its minimum. The reason is that we are interested in the critical regime, where the external source necessarily vanishes: $h = 0$. Now with the DE Ansatz for Γ_k it is not hard to see that $h = \delta\Gamma_k/\delta\phi$ becomes proportional to $U_k'(\rho)$ when evaluated in a uniform field configuration [32]. Hence $h = 0 \Leftrightarrow \rho = \rho_{min}$. For LPA' in particular, we Taylor expand to order 0, *i.e.*, replace Z_k by its value at the minimum of the effective potential.

The advantage of LPA' (over LPA) is the possibility to access the anomalous dimension η. Recall that η can be read from the p dependence of the two-point function $\Gamma^{(2)}_{k=0}(p) \sim p^{2-\eta}$ (see (25) and (28)). Based on the scaling hypothesis, we show below VII.C.1 that

$$\Gamma^{(2)}_k(p, u_0) - \Gamma^{(2)}_k(0, u_0) \sim p^2 k^{-\eta}. \tag{80}$$

(The order of limits being taken here is: first send $p \to 0$ keeping k finite, and then look at the k-dependence). Now LPA' (78) implies $\Gamma^{(2)}_k(p, u_0) - \Gamma^{(2)}_k(0, u_0) \sim p^2 Z_k$. Compared to (80), we see that $Z_k \sim k^{-\eta}$. Therefore, LPA ($Z_k = 1$) will give the mean field value $\eta = 0$ while with LPA', one can get non-trivial η values.

1. Scaling of 1PI propagator (Derivation of 80)

At criticality ($\phi = 0$ and $h = 0$), we consider the following ratio

$$\frac{\Gamma^{(2)}_k(p, u_0)}{\Gamma^{(2)}_k(0, u_0)} = \tilde{f}\left(\frac{p}{k}, \frac{k}{u_0^{1/(4-d)}}\right) \sim f\left(\frac{p}{k}\right) \tag{81}$$

where we have used that the ratio needs to be a dimensionless quantity and finite in the limit $0 < k \ll u_0^{1/(4-d)}$.

By dimensional analysis, we find

$$\Gamma^{(2)}_k(0, u_0) = k^2 \, g\left(\frac{k}{u_0^{1/(4-d)}}\right) \sim k^2 \left(\frac{k}{u_0^{1/(4-d)}}\right)^{-x} \tag{82}$$

where, in the last step, we have introduced the *scaling hypothesis*, that is, the fact that at criticality we expect scale invariance which implies that quantities depending on scales are power laws.

Then, using (81), we find

$$\Gamma^{(2)}_k(p, u_0)) \sim k^2 \left(\frac{k}{u_0^{1/(4-d)}}\right)^{-x} f\left(\frac{p}{k}\right) \tag{83}$$

which, in the limit $k \to 0$ with $p > 0$, must be finite. We conclude that for $p/k \gg 1$

$$f\left(\frac{p}{k}\right) \sim \left(\frac{p}{k}\right)^{2-x} \tag{84}$$

[32] A reader finding it hard to see may want to learn taking functional derivatives in the beginning of subsection VII.A.1

which implies for $k \ll p \ll u_0^{1/(4-d)}$

$$\Gamma_k^{(2)}(p, u_0) \sim p^{2-x} \tag{85}$$

and, thus, x is the anomalous dimension: $x = \eta$.

In the other limit where $p \ll k$ (with $k \ll u_0^{1/(4-d)}$) the theory is regularized and the function $f(p/k)$ is the function f can therefore be expanded

$$f\left(\frac{p}{k}\right) = 1 + c\left(\frac{p}{k}\right)^2 + ... \tag{86}$$

Using (82), we can now compute the behavior of $\Gamma_k^{(2)}(p, u_0)$, in the opposite limit of (85) where $p \ll k$ (and $k \ll u_0^{1/(4-d)}$) which is the domain of validity of the DE:

$$\begin{aligned}\Gamma_k^{(2)}(p, u_0) - \Gamma_k^{(2)}(0, u_0) &\sim \Gamma_k^{(2)}(0, u_0)\left(f\left(\frac{p}{k}\right) - 1\right) \\ &\underset{p \ll k}{\sim} p^2 k^{-\eta}.\end{aligned} \tag{87}$$

Since, within the DE,

$$\Gamma_k^{(2)}(p, u_0, \boldsymbol{\phi} = 0) = U''(\boldsymbol{\phi} = 0) + Z_k p^2 + O(p^4), \tag{88}$$

we find that at criticality and in the limit $k \ll u_0^{1/(4-d)}$:

$$Z_k \sim k^{-\eta} \tag{89}$$

which shows that η can be computed either from the p-dependence of $\Gamma_k^{(2)}(p, u_0)$ when $u_0^{1/(4-d)} \gg p \gg k$ (which is unreachable within the DE but that can be obtained in the BMW scheme) or from the k-dependence of Z_k which is computable from the DE.

D. The co-moving frame, self-similarity and fixed-points

In this section we address and circumvent problems which arise in practical studies of the flow of the potential in LPA close to criticality. At first, we review why the naive approach necessarily fails. In a second step, we introduce the co-moving frame and the (closely related) notions of self-similarity and fixed-points. We'll be again using the $O(N)$-models in LPA and LPA' as templates.

The straightforward way to study the flow of the potential is to discretise the space and compute the values of the potential at fixed grid points, hence, $\phi(x) \to \phi_i := \phi(x_i)$, where the index i refers to the lattice sites. Those need to be chosen such that the form of the potential, cf. section VII.B, is approximated accurately. For the system far away from criticality and a sufficiently narrow grid this can be done easily: up

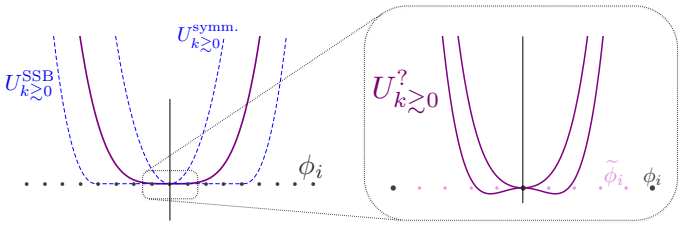


FIG. 14 Rescaling of the grid-resolution $\phi_i \to \widetilde{\phi}_j$ to study the effective potential of a system close to criticality.

to the accuracy of the grid-spacing we can determine the position of the minima. For the system right at criticality the minima must reach $\phi = 0$ only in the strict limit $k = 0$. That means, that for arbitrarily small k the grid must be fine enough to resolve if the minima are degenerate or distinct. At some point, however, the distance of the minima from the origin inevitably becomes smaller than the grid-spacing and, hence, we can not tell if we are right at the phase transition. At this point we must redefine the grid to smaller spacing. This is sketched in figure 14. In order to change the grid we have two possibilities, which correspond to active or passive transformations of the reference frame: we can either zoom the fields or shrink the lattice spacing.

This paragraph may not read well for some readers. The co-moving frame is a clever idea to zoom the fields ϕ as the minima approach the origin with decreasing scale k, simultaneously. Remember that in the flow at given k we are sensitive to fluctuations $p \approx k$. The idea of the co-moving frame is to rescale this region to its original size, so instead of looking on the flow with k from the outside we move together with the scale to smaller regions. In other words, instead of measuring the momenta p and k as dimensionful quantities, we introduce *dimensionless variables*, $\tilde{p} = p/k$. In that sense, we want to measure all quantities in units of k. From the scaling hypothesis we know that for a system close to criticality all quantities scale with a power-law. By dimensional analysis we can find and divide by the corresponding power in order to express all quantities in dimensionless variables.
As a consequence of the fact that there is only one scale k (with which we move), we have lost the explicit dependence on k in the flow equation. In other words, the flow does not know at which scale it is. Therefore, we can re-interpret the flow of the potential as its evolution in a dimensionless "RG-time"[33], $t := \log k/\Lambda$. When the shape of the potential does not change anymore in the evolution it has reached its *fixed-point*: $U^* : \partial_t U_k^* = 0$.

[33] Note that the system evolves from $t = 0 \to -\infty$.

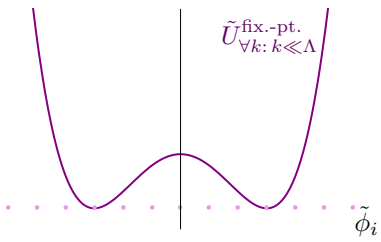


FIG. 15 Self-similar fixed-point potential $\tilde{U}_k$ at criticality in the co-moving frame. The zoom in the fields is done such that $\tilde{U}_k$ does not change its shape for smaller k. In particular the minima are fixed at non-zero postitions, even though the system is critical.

It is now *self-similar*, i.e. it is the same for all scales if we move with the scale. Figure 15 shows the self-similar fixed-point potential in the co-moving frame, where the grid is defined in the new variables $\tilde{\phi}$. To emphasise the connection to figure 14 we stress that the fixed-point potential is unchanged for arbitrarily small k.

We now turn to the co-moving frame for the $O(N)$-models in the LPA and advance to the LPA' below.
For the LPA it turns out to be sufficient to go to dimensionless variables to get rid of the explicit k-dependence in the flow equation. Here, on the left hand side of the flow of the potential[34] for an optimised regulator, (74),

$$k\partial_k U_k(\rho) = \frac{4v_d k^d}{d}\left\{\frac{1}{1+(U_k'+\rho U_k'')\,k^{-2}}+\frac{N-1}{1+U_k' k^{-2}}\right\}. \quad (90)$$

Note that we have also turned the derivative ∂_k into a dimensionless derivative, $k\partial_k = \partial \log k$, as we transform the equation to dimensionless quantities. By dimensional analysis in units of k, we find

$$[x]=k^{-1}\,,\ [p]=k\,,\ [\phi]=k^{\frac{d-2}{2}}\,,\ [U_k]=k^d\,. \quad (91)$$

The transformation to dimensionless quantities (tilded symbols) is given by

$$\begin{aligned} &p=\tilde{p}\,k\,,\ x=\tilde{x}/k\,,\\ &\phi(x)=k^{\frac{d-2}{2}}\,\tilde{\phi}(\tilde{x})\,,\\ &U_k(\phi(x))=k^d\,\tilde{U}_k\left(\tilde{\phi}(\tilde{x})\right). \end{aligned} \quad (92)$$

As a result, the flow equation of the potential is given by

$$\partial_t\tilde{U}_k = -d\,\tilde{U}_k' + (d-2)\,\tilde{\rho}\tilde{U}_k' + \frac{4v_d}{d}\left\{\frac{1}{1+\tilde{U}_k'(\tilde{\rho})+2\tilde{\rho}\tilde{U}_k''(\tilde{\rho})}+\frac{N-1}{1+\tilde{U}_k'(\tilde{\rho})}\right\}. \quad (93)$$

[34] We refrain from deriving (90) explicitly here as it can be easily inferred from (61) and (76).

The first two terms encode the dimensions of the potential and the fields, whereas the last term contains the dynamics from the radial and Goldstone modes.

By transforming to dimensionless variables we have succeeded in dropping the k-dependence in LPA. In the LPA' (and any higher order), however, we have to also consider the scaling of the wave-function renormalisation Z_k. The flow equation for the potential in LPA' is simply given by (61) with the condition $Y_k = 0$, hence,

$$\partial_t U_k(\rho) = \frac{1}{2}\int_q \partial_t R_k(q)\left(\frac{N-1}{Z_k q^2+R_k(q)+U_k'(\rho)}+\frac{1}{Z_k q^2+R_k(q)+U_k'(\rho)+2\rho U_k''(\rho)}\right) \quad (94)$$

Indeed, for a system near criticality, we have the scaling $Z_k \overset{k\to 0}{\sim} k^{-\eta}$ (where η is the anomalous dimension) so we will take this into account in our dimensional counting in k. It is convenient to define the running anomalous dimension η_k by

$$\eta_k = -\frac{k\partial_k Z_k}{Z_k}\,. \quad (95)$$

Now, the scaling behavior of Z_k implies $\eta_k \overset{k\to 0}{\longrightarrow} \eta$, *i.e.*, the fixed point of η_k is the anomalous dimension[35].
Expressed in the new variable, η_k, we find for the optimised regulator, (74),

$$k\partial_k R_k^{\Theta}(q) = -\eta_k R_k^{\Theta}(q) + 2Z_k k^2\Theta(k^2-p^2)\,. \quad (96)$$

Nevertheless, k will appear explicitly in the flow equation due again to the non-trivial running of Z_k, even in dimensionless variables. We cannot factor Z_k in both the numerator and the denominator of the flow equation, because the terms $\sim U_k'$, $\rho U_k''$ scale differently than the term quadratic in momentum. However, by a change of variables, we can eliminate the dependence on Z_k: we modify the field renormalisation in a k-dependent way to express the flow in renormalised (and dimensionless) quantities. The transformation of the field is given by

$$\phi(x) = Z_k^{-1/2} k^{\frac{d-2}{2}}\tilde{\phi}(\tilde{x})\,. \quad (97)$$

This ensures that for (92) with the modified definition of $\tilde{\phi}$, (97), the terms $\sim \tilde{U}_k'$, $\tilde{\rho}\tilde{U}_k''$ scale in the same way[36] as $Z_k q^2$. Therefore, the factor of Z_k cancels between the numerator and the denominator in the flow equation, leaving only the dimensionless η_k. As a result, for

[35] The initial condition for η_k at $k=\Lambda$ is given uniquely by the running. Indeed, as the anomalous dimension only changes in the critical region of small k, usually the initial condition $\eta_\Lambda = 0$ is applicable.

[36] The derivative is now taken with respect to the renormalised

a general regulator $R_k(q) = Z_k q^2 r(q^2/k^2)$, we obtain the flow of the potential

$$k\partial_k \tilde{U}_k = -d\,\tilde{U}_k + (d-2+\eta_k)\,\tilde{\rho}\tilde{U}'_k - v_d \int_0^\infty dy\, y^{d/2}\,(\eta_k r(y) + 2y r'(y)) \times \left\{ \frac{1}{y\,(1+r(y)) + \tilde{U}'_k + 2\tilde{\rho}\tilde{U}''_k} + \frac{N-1}{y\,(1+r(y)) + \tilde{U}'_k} \right\}, \tag{99}$$

with $y := q^2/k^2$ and $\tilde{U}'_k = \tilde{U}'_k(\tilde{\rho})$ and similarly for $U''(k)$. The structure of the equation is similar to the one in LPA, (93): On the right hand side, the first term encodes the dimension of the potential while the second term considers the dimension of the field, which is now taking into account the non-trivial scaling of the propagator, η_k. The dynamical part of the flow is given by the loop-integral.

The question is left, whether such a fixed-point potential exists. Indeed, for $d = 3$ (and arbitrary N) we find exactly one solution for $\tilde{U}^*$ (aside from the trivial solution $\tilde{U}^*_k =$ constant), which is the famous *Wilson–Fisher* fixed-point. The existance of only one solution is highly-nontrivial, because, in principle, an infinite number of fixed-point potentials mathematically could exist. In the present case, however, only a finite number is globally defined: except for a discrete number[37] all other solutions have singularities for a finite ρ.

Two comments address the question of convexity and dimensional dependence of the fixed-point potential:

- The fixed-point potential does not need to be convex (unlike the normal, dimensionful potential), because for a divergent correlation length the comoving frame is an *infinite* zoom in the fields,

$$\phi(x) = Z_k^{-1/2} k^{\frac{d-2}{2}} \tilde{\phi}(\tilde{x}) \overset{k\to 0}{\sim} \underbrace{k^{\frac{\eta+1}{2}}}_{\overset{k\to 0}{\longrightarrow} 0} \tilde{\phi}(\tilde{x})\,. \tag{100}$$

  for $d = 3$ and small η. Therefore, a finite range in $\tilde{\phi}$ corresponds to zero range in ϕ.

- The number of solutions depends on the dimension, even though the equation is the same. For $O(N)$-models in LPA', in $d = 2.9$ there are two non-trivial fixed-point potentials. For $d = 2$ there is an infinite number of possible fixed-point potentials, which constitutes a so-called *multi-critical fixed-point*.This result has been known from conformal field theories.

	MC	7-loop PT	$O(6)$	$O(4)$	$O(2)$	LPA
ν	0.63001(10)	0.6304(13)	0.6303(1)	0.632	0.628	0.6506
η	0.0362(2)	0.0335(25)	0.0357(2)	0.033	0.044	0
ω	0.832		0.827			

TABLE I Comparison of critical exponents for $N = 1$, $d = 3$ from Monte-Carlo lattice simulations, 7-loop PT and different orders of the DE in NPRG.

E. Quantitative results of the NPRG

In this subsection we compare quantitative results obtained in the NPRG with other methods. For $N = 1$, $d = 3$ we focus on the critical exponents ν, (6), and η, (80), and the *correction-to-scaling* exponent, ω, defined as to correction to the dominant scaling of the correlation length, viz.

$$\xi \sim |T - T_c|^{-\nu} \left(1 + A\,(T - T_c)^{-\omega}\right), \tag{101}$$

where A is a constant.

In principle, the results should depend on three approximations that have been made: the initial condition, the order of the DE and the choice of the regulator function. However, as we consider the self-similar fixed-point potential, the dependence on the initial condition is lost. The dependence on the choice of the regulator is mild, even for very different classes of regulators as given in (74), (75), thus, we shall not comment on it further. As a consequence, the only significant errors stem from truncational artifacts.

Monte-Carlo lattice computations serve as the benchmark, hence, we measure the accuracy with respect to results obtained there. In addition, we state the 7-loop PT predictions, in comparison with which we see the strength of the NPRG, as even comparatively simple truncations[38] yield good accuracy. The results are given in table I. [@ Bertrand: I took the numbers from your review, Lect.Notes Phys. 852 (2012) 49-132. Are those up-to-date? You said in the lecture that the PT results

field, hence, in a slight abuse of notation, we find

$$\rho(x) = Z_k^{-1} k^{d-2} \tilde{\rho}(\tilde{x})\,,$$
$$k^d U'_k(\rho) = \tilde{U}'_k(\tilde{\rho})\,\frac{\delta\tilde{\rho}}{\delta\rho} = Z_k k^2 \tilde{U}'_k(\tilde{\rho})\,,$$
$$k^d \rho U''_k(\rho) = k^d Z_k^{-1} k^{d-2} \tilde{\rho}\tilde{U}''_k(\tilde{\rho}) \left(\frac{\delta\tilde{\rho}}{\delta\rho}\right)^2 = Z_k k^2 \tilde{\rho}\tilde{U}''_k(\tilde{\rho})\,, \tag{98}$$

i.e. it scales like the quadratic part in the fields.

[37] Note that the discrete number of possible solutions is not necessarily finite.

[38] Note that in $d = 2$ the results are much worse. [Is there a (simple) reason for this? I assume that going to slightly higher order in the DE is not sufficient to improve?] For example, the exact result of Onsager $\eta^{\text{Onsager}} = 1/4$ is approximated by the $O(6)$ value by $\eta^{\text{O(6)}} = 0.237$ only.

were 6-loop, in your review it says 7-loop? Are they the same, or was that the comment you made that 'parts of the 7-loop contributions are taken into account there'?]

F. Field expansions

On top of the DE, we can make a field expansion of $U_k, Z_k, \ldots$. The simplest such truncation consists in using the LPA and in keeping only the first two terms of the expanion of U_k in powers of ϕ:

$$U_k(\rho) = g_{2,k}\rho + g_{4,k}\,\rho^2 + \ldots. \tag{102}$$

But as is explained above (see remark after (79)), it is better to expand U_k around its minimum instead

$$U_k(\rho) = u_{2,k}(\rho - \rho_{min})^2 + u_{3,k}(\rho - \rho_{min})^3 + \ldots$$

With these kinds of ansatz, the RG equation on Γ_k becomes a set a ordinary differential equations for the couplings retained in the *ansatz*:

$$\partial_t g_{n,k} = \beta_n\left(\{g_{p,k}\}\right). \tag{103}$$

If $U_k, Z_k, \ldots$ are not truncated in a field expansion, the RG equation on Γ_k becomes a set of coupled partial differential equations for these functions (as we have seen in the Ising model example VII.B). The initial condition at scale Λ is, as always, given by the Hamiltonian of the model.

An important remark follows. Let us notice that if the k-dependence of the couplings was neglected, the *ansatz* of (102) would exactly coincide with the ansatz chosen by Landau to study second order phase transitions. We know that it would lead to the mean field approximation which overall fails to predict the critical exponents. This is still the case considering a more 'refined' ansatz of the kind of the LPA', provided that we have removed the k-dependence. Hence, let us stress the remarkable fact that much of the physics is actually contained in the k-dependence.

VIII. THE BMW EXPANSION

So far, only $\Gamma_k^{(2)}(\mathbf{p})$ has been computed for the $O(N)$ models.

Appendix A: Gaussian free theory, perturbation theory, Mean Field and the classical approximation

We want to shed some light on these concepts, which are interrelated. Let's consider a theory with just one coupling constant like $g\varphi^4$. *Perturbation theory* stands for a perturbative expansion in the coupling constant or, equivalently, in $\hbar$:

$$\Gamma^{(2)} = p^2 + m^2 + \sum_i \left((g\hbar)^i \ldots\right) \tag{A1}$$

$$\Gamma^{(4)} = g + \sum_i \left((g\hbar)^i \ldots\right) \tag{A2}$$

The *gaussian free theory* is the limit $g \to 0$, so no interactions and only $\Gamma^{(2)}$ is non zero. On its part, *Mean Field = the classical approximation* stand for the limit $\hbar \to 0$, which is also the *tree-level* of perturbation theory.

Quantum criticality with two length scales

Hui Shao,[1,2,4] Wenan Guo,[1,3,*] Anders W. Sandvik[4,*]

[1]Department of Physics, Beijing Normal University, Beijing 100875, China
[2]Beijing Computational Science Research Center, Beijing 100084, China
[3]State Key Laboratory of Theoretical Physics, Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China
[4]Department of Physics, Boston University, Boston, Massachusetts 02215, USA

[*]To whom correspondence should be addressed;
E-mail: waguo@bnu.edu.cn, sandvik@bu.edu

The theory of deconfined quantum critical points describes phase transitions at temperature $T = 0$ outside the standard paradigm, predicting continuous transitions between certain ordered states where conventional theory requires discontinuities. Numerous computer simulations have offered no proof of such transitions, however, instead finding scaling violations which are neither predicted by the new theory nor conform with standard scenarios. Here we resolve this enigma by introducing a critical scaling form with two divergent length scales. Simulations of a quantum magnet with antiferromagnetic and dimerized ground states confirm the form, proving a continuous transition with deconfined excitations and also explaining anomalous scaling at $T > 0$. Our findings revise prevailing paradigms for quantum criticality, with potentially far-reaching implications for many strongly-correlated materials.

Introduction In analogy with classical phase transitions driven by thermal fluctuations, condensed matter systems can also undergo drastic changes as parameters regulating quantum fluctuations are tuned at low temperatures. Some of these *quantum phase transitions* can be theoretically understood as rather straight-forward generalizations of thermal phase transitions (*1, 2*), where, in the conventional Landau-Ginzburg-Wilson (LGW) paradigm, states of matter are characterized by order parameters. Many strongly-correlated quantum materials seem to defy such a description, however, and call for new ideas.

A promising proposal is the theory of deconfined quantum critical (DQC) points in two-dimensional (2D) quantum magnets (*3, 4*), where the order parameters of the antiferromagnetic (Néel) state and the competing dimerized state (the valence-bond-solid, VBS) are not fundamental variables but composites of fractional degrees of freedom carrying spin $S = 1/2$. These *spinons* are condensed and confined, respectively, in the Néel and VBS state, and become deconfined at the DQC point separating the two states. Establishing the applicability of the still controversial DQC scenario would be of great interest in condensed matter physics, where it may play an important role in strongly-correlated systems such as the cuprate superconductors (*5*). There are also intriguing DQC analogues to quark confinement and other aspects of high-energy physics, e.g., an emergent gauge field and the Higgs mechanism and boson (*6*).

The DQC theory represents the culmination of a large body of field-theoretic works on VBS states and quantum phase transitions out of the Néel state (*2, 7–10*). The postulated SU(N) symmetric non-compact (NC) CP^{N-1} action can be solved when $N \to \infty$ (*5, 11, 12*) but non-perturbative numerical simulations are required to study small N. The most natural physical realizations of the Néel–VBS transition for electronic SU(2) spins are frustrated quantum magnets (*9*), which, however, are notoriously difficult to study numerically (*13, 14*). Other models were therefore pursued. In the J-Q model (*15*), the Heisenberg exchange J between $S = 1/2$ spins is supplemented by a VBS-inducing four-spin term Q which is amenable to efficient quan-

tum Monte Carlo (QMC) simulations (*15–23*). While many properties of the J-Q model support the DQC scenario, it has not been possible to draw definite conclusions because of scaling violations affecting many properties. Similar anomalies were later observed in three-dimensional loop (*24*) and dimer (*25*) models, which are also potential realizations of the DQC point. Simulations of the NCCP[1] action as well have been hard to reconcile with the theory (*21, 26, 27*).

One interpretation of the unusual scaling behaviors is that the transitions are first-order, as generally required within the LGW framework for order–order transitions where unrelated symmetries are broken. The DQC theory would then not apply to any of the systems studied so far, thus casting doubts on the entire concept (*17,21,26*). In other interpretations the transition is continuous but unknown mechanisms cause strong corrections to scaling (*18, 27, 28*) or modify the scaling more fundamentally in some yet unexplained way (*19, 24*). The enigmatic current state of affairs is well summed up in the recent Ref. (*24*).

Here we present a resolution of the DQC puzzle based on a finite-size scaling Ansatz including the two divergent length scales of the theory—the standard correlation length ξ (which is the same for both order parameters, $\xi_{\rm spin} \propto \xi_{\rm dimer}$, and we use ξ for either of them) and a length $\xi' > \xi$ associated with the thickness of VBS domain walls and spinon confinement (the size of a spinon bound state). We show that, contrary to past assumptions, ξ' can govern the finite-size scaling even of quantities which are sensitive only to ξ in the thermodynamic limit. We carry out simulations of the J-Q model at low temperatures as well as in the lowest $S = 1$ (two-spinon) excited state, and demonstrate complete agreement with the two-length scaling hypothesis with no anomalous scaling corrections remaining.

Finite-size scaling forms Consider first a system with a single divergent correlation length $\xi \propto |\delta|^{-\nu}$, where $\delta = g - g_c$ is the distance to a phase transition driven by quantum fluctuations arising from non-commuting interactions controlled by g at $T = 0$. According to standard

finite-size scaling theory (*29*), close to $\delta = g - g_c = 0$ a singular quantity A takes the form

$$A(g, L) = L^{-\kappa/\nu} f(\delta L^{1/\nu}, L^{-\omega}), \tag{1}$$

where the exponents κ, ν, ω are tied to the universality class, κ also depends on A, and the scaling function f approaches a constant when $\delta \to 0$. We assume $\beta = 1/T \propto L^z$ (or, alternatively, $T = 0$) so that scaling arguments depending on β have been eliminated.

The form (1) fails for some properties of the J-Q model (*18,19,22*) and other DQC candidate systems (*24–26*). A prominent example is the spin stiffness, which for an infinite 2D system in the Néel phase should scale as $\rho_s \propto \delta^{z\nu}$ with $z = 1$ (*1, 3, 4*). To eliminate the size dependence when $\delta \neq 0$ and $L \to \infty$ in Eq. (1), we must have $\kappa = z\nu$ and $f(x, L^{-\omega}) \propto x^{z\nu}$ for large $x = \delta L^{1/\nu}$. Thus, $\rho_s(\delta = 0, L) \propto L^{-z}$ and $L\rho_s$ should be constant when $L \to \infty$ if $z = 1$. However, $L\rho_s(L)$ at criticality instead appears to diverge slowly (*17, 18, 21*). At first sight this might suggest $z < 1$, but other quantities, e.g., the magnetic susceptibility, instead behaves as if $z > 1$ (*30*). Strong scaling corrections have been suggested as a way out of this dilemma (*18, 19, 28*). Claims of a weak first-order transition have also persisted (*21, 26, 27*), though the continuous DQC scenario is supported by the absence of any of the usual first-order signals, e.g., the Binder cumulant does not exhibit any negative peak (*18, 24*).

To explain the scaling anomalies phenomenologically, in the presence of a second length $\xi' \propto \delta^{-\nu'}$ in the VBS, we propose that Eq (1) should be replaced by the form

$$A(g, L) = L^{-\kappa/\tilde{\nu}} f(\delta L^{1/\nu}, \delta L^{1/\nu'}, L^{-\omega}), \tag{2}$$

where, unlike what was assumed in the past, *$\tilde{\nu}$ is not necessarily the same as the exponent ν which governs the behavior of most observables in the thermodynamic limit*. Instead, we show that the criticality in the J-Q model generically has $\tilde{\nu} = \nu'$.

First assume $\tilde{\nu} = \nu$. The correct thermodynamic limit with $\kappa = z\nu$ for ρ_s can then be obtained from Eq. (2) if $f(x, y, L^{-\omega}) \propto x^{z\nu}$ for large x, y ($y = \delta L^{1/\nu'}$, $x = \delta L^{1/\nu}$) and,

ss before, $\rho_s(\delta = 0, L) \propto L^{-z}$. This can also be expressed using a scaling function where the second argument is the ratio of the two lengths; $\tilde{f}(\delta L^{1/\nu}, L^{1/\nu'-1/\nu}, L^{-\omega})$. If $\tilde{f}(\delta = 0)$ is constant when $L \to \infty$, then $L^{1/\nu'-1/\nu}$ acts like just another irrelevant field, similar to the standard scenario for dangerously irrelevant perturbations in classical clock models (*31*). Our proposal instead corresponds to the function $\tilde{f}$ behaving as a power of L for $\delta \to 0$, which was suggested in Ref. (*19*) though without any detailed forms or predictions. Using the form (2) and considering a different large-L limit, we can make new concrete predictions for the nature of the scaling anomalies. In the case of the stiffness, the correct thermodynamic limit is also obtained with $\tilde{\nu} = \nu'$ and $\kappa = z\nu$ if $f(x, y, L^{-\omega}) \propto y^{z\nu}$ for large L. Then $\rho_s(\delta = 0) \propto L^{-z\nu/\nu'}$.

This new scaling behavior corresponds to $\xi \propto (\xi')^{\nu/\nu'}$ saturating at $\xi \propto L^{\nu/\nu'}$ when $\xi' \to L$ upon approaching the critical point, in contrast to the standard scenario where ξ grows until it also reaches L. The criticality at distances $r < L^{\nu/\nu'}$ is still conventional, while $r > L^{\nu/\nu'}$ is governed by the unconventional power laws. Different behaviors for $r \ll L$ and $r \approx L$ were actually observed in the recent loop-model study (*24*) and a dangerously irrelevant field was proposed as a possible explanation, but with no quantitative predictions of the kind offered by our approach. The anomalous scaling law controlled by ν/ν', which we will confirm numerically, not only represents a novel property of the DQC point but also demonstrates an unexpected general richness of quantum criticality with two divergent lengths.

Quantum Monte Carlo Results The J-Q model (*15*) for $S = 1/2$ spins is defined using singlet projectors $P_{ij} = 1/4 - \mathbf{S}_i \cdot \mathbf{S}_j$ as

$$H = -J \sum_{\langle ij \rangle} P_{ij} - Q \sum_{\langle ijkl \rangle} P_{ij} P_{kl}, \tag{3}$$

where $\langle ij \rangle$ denotes nearest-neighbor sites on a periodic square lattice with L^2 sites and the pairs ij and kl in $\langle ijkl \rangle$ form horizontal and vertical edges of 2×2 plaquettes. This Hamiltonian has all symmetries of the square lattice and the VBS ground state existing for $g = J/Q <$

g_c (with $g_c \approx 0.045$) is columnar, breaking the translational and 90° rotational symmetries spontaneously. The Néel state for $g > g_c$ breaks the spin-rotation symmetry.

We will study several quantities in the neighborhood of g_c. Although we have argued that the asymptotic $L \to \infty$ behavior when $\delta \neq 0$ in Eq. (2) is controlled by the second argument of f, the critical scaling close to $\delta = 0$ can still be governed by the first argument. We will demonstrate that, depending on the quantity, either $\delta L^{1/\nu}$ or $\delta L^{1/\nu'}$ is the relevant argument, and, therefore, ν and ν' can be extracted using finite-size scaling with effectively one-parameter forms. We will do this for manifestly dimensionless quantities, $\kappa = 0$ in Eq. (2), before testing the anomalous powers of L in other quantities.

If the effective one-parameter scaling holds close to g_c, then Eq. (2) implies that $A(g, L_1) = A(g, L_2)$ at some $g = g^*(L_1, L_2)$ and a crossing-point analysis (Fisher's phenomenological renormalization) can be performed (*29*). For a $\kappa = 0$ quantity, if $L_1 = L$ and $L_2 = rL$ with r constant, a Taylor expansion of f shows that the crossing points $g^*(L)$ approach g_c as $g^*(L) - g_c \propto L^{-(1/\nu+\omega)}$ if ν is the relevant exponent (which we assume here for definiteness). $A^* = A(g^*)$ approaches its limit A_c as $A^*(L) - A_c \propto L^{-\omega}$, and one can also show that

$$\frac{1}{\nu^*(L)} = \frac{1}{\ln(r)} \ln \left(\frac{dA(g, rL)/dg}{dA(g, L)/dg} \right)_{g=g^*} \tag{4}$$

converges to $1/\nu$ at the rate $L^{-\omega}$. In practice, simulation data can be generated on a grid of points close to the crossing values, with polynomials used for interpolation and derivatives. We present details and tests of such a scheme for the Ising model in Supplementary Material.

In the $S = 1$ sector, properties directly related to spinons can be studied with projector QMC simulations in a basis of valence bonds (singlet pairs) and two unpaired spins (*32, 34*). Previously the size of the spinon bound state in the J-Q model was extrapolated to the thermodynamic limit (*33*), but the results were inconclusive as to the rate of divergence upon approaching the critical point. Here we study the critical finite-size behavior. We define the size

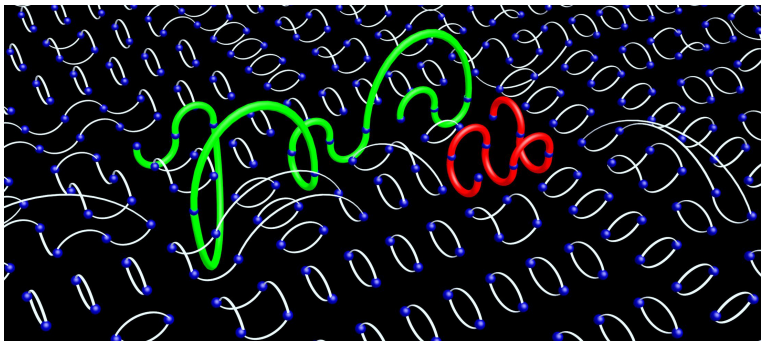


Figure 1: A QMC transition graph representing (*32, 34*) a sampled overlap $\langle\psi_L|\psi_R\rangle$ of $S = 1$ states, with two strings (spinons) in a background of loops formed by valence bonds. Arches above and below the plane represent $|\psi_R\rangle$ and $\langle\psi_L|$, respectively.

Λ of the spinon pair using the strings connecting the unpaired spins in valence-bond QMC simulations (*32, 34*), as illustrated in Fig. 1 and explained further in Supplementary Material.

If $\Lambda(g) \propto \xi'(g)$ when $L \to \infty$, then $\Lambda(g_c) \propto L$ follows from our proposed limit of Eq. (2). If Λ manifestly probes only the longer length scale also in a finite system, which we will test, then ν' is the exponent controlling the crossing points of Λ/L. Data and fits are presented in Fig. 2 (left). Unlike other quantities used previously to extract the critical point (*18*), the drift of $g*$ with L is monotonic in this case and the convergence is rapid. All $L \geq 16$ points are consistent with the expected power-law correction, with $1/\nu' + \omega \approx 3.0$ and $g_c = 0.04468(4)$, where the number in parenthesis indicates the statistical uncertainty (one standard deviation) in the preceding digit. The critical point agrees well with earlier estimates (*18*). The crossing value Λ^*/L also clearly converges and a slope analysis according to Eq. (4) gives $\nu' = 0.585(18)$.

Next, in Fig. 2 (right) we analyze a Binder ratio, defined with the z-component of the sublattice magnetization m_{sz} as $R_1 = \langle m_{sz}^2\rangle/\langle|m_{sz}|\rangle^2$ and computed in $T > 0$ simulations at $\beta = 1/T = L$ as in Ref. (*18*). Here the non-monotonic behavior of the crossing points necessitates several scaling corrections, unless only the largest sizes are used. In either case, the $L \to \infty$ behavior of g^* is fully consistent with g_c obtained from Λ/L. $R_1(g_c)$ has an uncertainty of over 1% because of the small value of the subleading exponent; $\omega \approx 0.4 \sim 0.5$. The slope

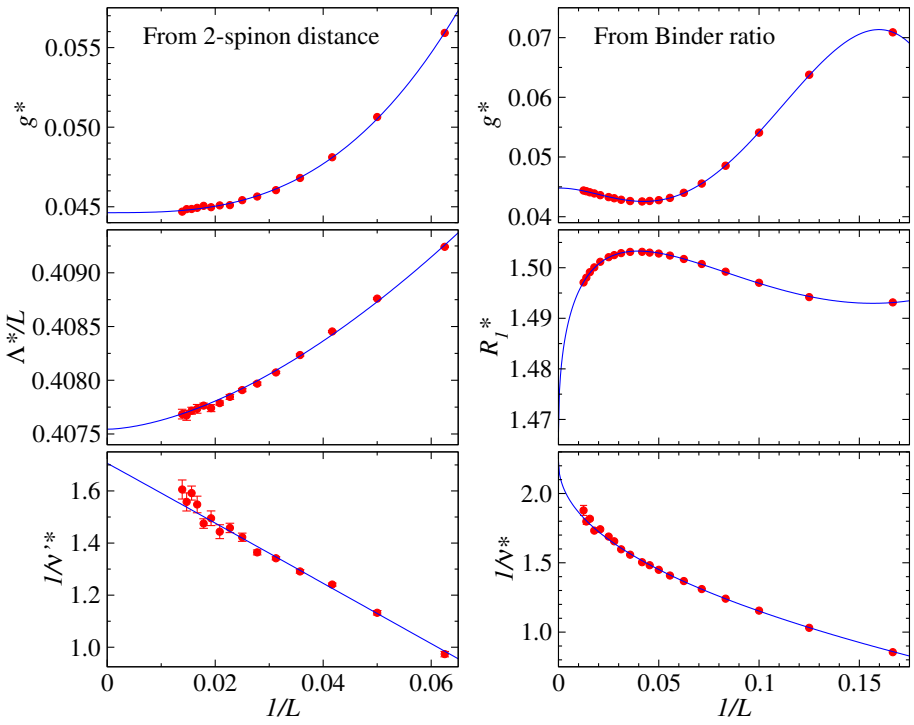


Figure 2: Crossing-point analysis of $(L, 2L)$ pairs for the size of the spinon bound state (left) and the Binder ratio (right). The monotonic quantities are fitted with simple power-law corrections, while two subleading corrections were included in the fits of the non-monotonic quantities.

estimator (4) of the exponent $1/\nu$ is monotonic and requires only a single $L^{-\omega}$ correction, also with a small exponent $\omega \approx 0.45$. The extrapolated exponent $\nu = 0.446(8)$ is close to the value obtained recently for the loop model (*24*).

The above results support a non-trivial deconfinement process where the size of the bound state diverges faster than the conventional correlation length. However, in the DQC theory the fundamental longer length scale ξ' is the thickness of a VBS domain wall. It can be extracted from the domain wall energy per unit length κ, which in the thermodynamic limit should scale as $\kappa \propto (\xi\xi')^{-1}$ (*4*). In Supplementary Material we re-derive this form using a two-length scaling Ansatz and discuss simulations of domain walls in a 3D clock model and the J-Q model. At criticality, in the conventional scenario (exemplified by the clock model) both ξ and ξ' saturate at L and $\kappa \propto L^{-2}$. For the J-Q model we instead find $\kappa \propto L^{-a}$ with $a = 1.715(15)$ for large L,

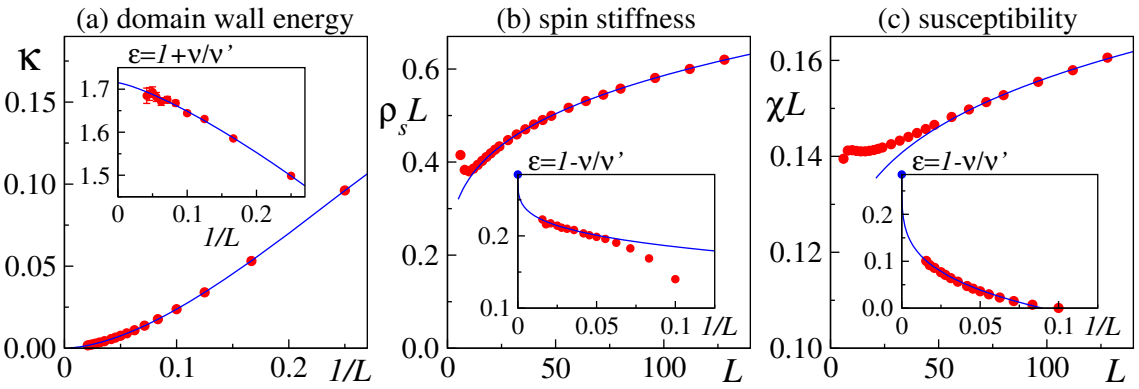


Figure 3: Consistent anomalous critical scaling of different quantities y at $J/Q = 0.0447$. The insets show running exponents $\epsilon(L) = \ln(y_L/y_{2L})/\ln(2)$ based on $(L, 2L)$ data. In (a), $1+\nu/\nu' = 1.715$ for $L \to \infty$ and a correction $\propto L^{-1.21}$ were determined by a fit to $\epsilon(L)$, while in the insets of (b) and (c) $1-\nu/\nu'$ was fixed at the corresponding value 0.285 and a correction $\propto L^{-\omega}$ with $\omega \approx 0.3$ was used to match the data for large systems. The same values of ν/ν' and ω were used in curves of the form $L^{1-\nu/\nu'}(a + bL^{-\omega})$ in the main (b,c) graphs.

as illustrated in Fig. 3(a). The natural interpretation of this unconventional scaling is that, when ξ' saturates at L, then ξ also can no longer grow but remains at $\xi \propto L^{\nu/\nu'}$. Thus $\kappa \propto L^{-(1+\nu/\nu')}$ with $\nu/\nu' = a - 1 = 0.715(15)$, which agrees reasonably well with the value $\nu/\nu' = 0.76(3)$ obtained from the quantities in Fig. 2. The results still leave open the possibility that the spinon confinement is governed by an exponent between ν and the domain-wall exponent ν' (*4*).

The critical spin stiffness ρ_s and long-wavelength susceptibility $\chi(k = 2\pi/L)$ were calculated at $\beta = L$. Conventional quantum-critical scaling (*2*) dictates that both these quantities should decay as $1/L$ when $z = 1$. Instead, Figs. 3(b,c) demonstrate clearly slower decays, $L\rho_s$ and $L\chi$ being slowly divergent, as had been found in earlier works as well (*17–19, 21, 30*). The new limit of the scaling function (2) requires $L\rho_s$ and $L\chi$ to diverge as $L^{1-\nu/\nu'}$. The behaviors are indeed consistent with $\nu/\nu' = 0.715$ extracted from κ and a correction $\propto L^{-\omega}$ with a small ω (close to the correction for R_1 in Fig. 2). The scaling of the three quantities is remarkably self-consistent, thus lending strong support to the new type of criticality where the magnetic properties are not decoupled from the longer VBS length scale ξ' for finite L. The results are incompatible with a first-order transition, where, $\kappa \to$ constant, $L\rho_s \to L$, $L\chi \to L$.

Discussion We have shown that the effects of the larger divergent length scale ξ' at the Néel–VBS transition are more dramatic than those caused by standard (*31*) dangerously-irrelevant perturbations, and we therefore propose the term *super-dangerous* for this case. The universality class, in the sense of the normal critical exponents in the thermodynamic limit at $T = 0$, are not affected by such perturbations, but anomalous power laws of the system size appear generically in finite-size scaling. We have determined the value $\nu/\nu' \approx 0.72$ for the exponent ratio governing the anomalous scaling in the J-Q spin model.

Loop and dimer models exhibit similar scaling anomalies (*24, 25*) and it would be interesting to test the consistency between different quantities as we have done here. In simulations of the NCCP^1 action (*21, 26, 27*), one would at first sight not expect any effects related to the longer DQC length scale, because the monopoles responsible for the VBS condensation are not present in the continuum theory (*3*). However, there could still be some other super dangerous operator present [see also Ref. (*24*)], perhaps related to lattice regularization.

The consequences of our findings extend also to $T > 0$ quantum criticality in the thermodynamic limit, because $1/T$ is the thickness of an equivalent system in the path integral formulation (*1, 2*). Anomalous finite-T behaviors of the J-Q model have already been observed (*18, 30*). For instance, the spin correlation length at $T > 0$, which should be affected by deconfined spinons, grows faster than the normally expected form $\propto T^{-1}$ and the susceptibility vanishes slower than T. Remarkably, the forms $\propto T^{-\nu'/\nu}$ and $T^{\nu/\nu'}$ can account for the respective behaviors. Thus, we find a strong rationale to revise the experimentally most important tenet of quantum criticality—the way $T = 0$ scaling is related to power laws in T at $T > 0$.

We conclude that quantum criticality with two divergent length scales is much richer than previously anticipated. Our findings may apply to a wide range of strongly-correlated quantum systems with more than one length scale and may help to resolve the mysteries still surrounding scaling behaviors in materials such as the high-T_c cuprate superconductors.

References and Notes

1. M. P. A. Fisher, P. B. Weichman, G. Grinstein, and D. S. Fisher, Phys. Rev. B **40**, 546 (1989).

2. A. V. Chubukov, S. Sachdev, and J. Ye, Phys. Rev. B **49**, 11919 (1994).

3. T. Senthil, A. Vishwanath, L. Balents, S. Sachdev, and M. P. A. Fisher, Science **303**, 1490 (2004).

4. T. Senthil, L. Balents, S. Sachdev, A. Vishwanath, and M. P. A. Fisher, Phys. Rev. B **70**, 144407 (2004).

5. R. K. Kaul and S. Sachdev, Phys. Rev. B **77**, 155105 (2008).

6. Y. Huh, P. Strack, and S. Sachdev, Phys. Rev. Lett. **111**, 166401 (2013).

7. N. Read and S. Sachdev, Phys. Rev. B **42**, 4568 (1990).

8. G. Murthy and S. Sachdev, Nucl. Phys. B **344**, 557 (1990).

9. N. Read and S. Sachdev, Phys. Rev. Lett. **66**, 1773 (1991).

10. O. I. Motrunich and A. Vishwanath, Phys. Rev. B **70**, 075104 (2004).

11. M. A. Metlitski, M. Hermele, T. Senthil and M. P. A. Fisher, Phys. Rev. B **78**, 214418 (2008).

12. E. Dyer, M. Mezei, S. S. Pufu, and S. Sachdev, J. High Energy Phys. **2015**, 37 (2015).

13. T. Li, F. Becca, W. Hu, and S. Sorella, Phys. Rev. B **86**, 075111 (2012).

14. S.-S. Gong, W. Zhu, D. N. Sheng, O. I. Motrunich, and M. P. A. Fisher, Phys. Rev. Lett. **113**. 027201 (2014).

15. A. W. Sandvik, Phys. Rev. Lett. **98**, 227202 (2007).

16. R. G. Melko and R. K. Kaul, Phys. Rev. Lett. **100**, 017203 (2008).

17. F. J. Jiang, M. Nyfeler, S. Chandrasekharan, and U. J. Wiese, J. Stat. Mech. (2008), P02009.

18. A. W. Sandvik, Phys. Rev. Lett. **104**, 177201 (2010).

19. R. K. Kaul, Phys. Rev. B **84**, 054407 (2011).

20. K. Harada, T. Suzuki, T. Okubo, H. Matsuo, J. Lou, H. Watanabe, S. Todo, and N. Kawashima, Phys. Rev. B. **88**, 220408 (2013).

21. K. Chen. Y. Huang, Y. Deng, A. B. Kuklov, N. V. Prokofev, and B. V. Svistunov, Phys. Rev. Lett. **110**, 185701 (2013).

22. M. S. Block, R. G. Melko, and R. K. Kaul, Phys. Rev. Lett. **111**, 137202 (2013).

23. S. Pujari, F. Alet, and K. Damle, Phys. Rev. B **91**, 104411 (2015).

24. A. Nahum, J. T. Chalker, P. Serna, M. Ortunõ, and A. M. Somoza, Phys. Rev. X **5**, 041048 (2015).

25. G. J. Sreejith and S. Powell, Phys. Rev. B **89**, 014404 (2014).

26. A. B. Kuklov, M. Matsumoto, N. V. Prokofev, B. V. Svistunov, and M. Troyer, Phys. Rev. Lett. **101**, 050405 (2008).

27. O. I. Motrunich and A. Vishwanath, arXiv:0805.1494 (2008).

28. L. Bartosch, Phys. Rev. B **88**, 195140 (2013).

29. J. M. Luck, Phys. Rev. B **31**, 3069 (1985).

30. A. W. Sandvik, V. N. Kotov, and O. P. Sushkov, Phys. Rev. Let. **106**, 207203 (2011).

31. F. Léonard and B. Delamotte, Phys. Rev. Lett **115**, 200601 (2015).

32. Y. Tang and A. W. Sandvik, Phys. Rev. Lett. **107**, 157201 (2011).

33. Y. Tang and A. W. Sandvik, Phys. Rev. Lett. **110**, 217213 (2013).

34. A. Banerjee and K. Damle, J. Stat. Mech. (2010) P08017.

35. G. Kamieniarzt and H. W. J. Blöte, J. Phys. A: Math. Gen. **26**, 201 (1993).

36. E. Marinari, *Optimized Monte Carlo Methods*, Advances in Computer Simulation, edited by J. Kertesz and I. Kondor (Springer, Berlin, 1997); arXiv:cond-mat/9612010.

37. A. W. Sandvik and H. G. Evertz, Phys. Rev. B. **82**, 024407 (2010).

38. H. Shao, W. Guo, and A. W. Sandvik, Phys. Rev. B **91**, 094426 (2015).

Supplementary Materials

www.sciencemag.org

Supporting on-line text

Figures S1 - S9

References (35-38)

Acknowledgments We would like to thank Kedar Damle for suggesting the definition of spinons using the entire strings created by unpaired spins in the valence-bond basis (*32, 34*). The research was supported by NSFC Grant No. 11175018 (WG), the Fundamental Research Funds for the Central Universities (WG), NSF Grant No. DMR-1410126 (AWS), and by the Simons Foundation (AWS). HS and WG gratefully acknowledge support from the Condensed

Matter Theory Visitors Program at Boston University and AWS is grateful to the Institute of Physics, Chinese Academy of Sciences, and Beijing Normal University for their support. Some of the computations were carried out using Boston University's Shared Computing Cluster.

Supplementary Materials

Supporting on-line text for **Quantum criticality with two length scales**

Hui Shao, Wenan Guo, and Anders W. Sandvik

Here we provide technical details of the methods and calculations reported in the main paper.

In **Sec. 1** we discuss the crossing-point analysis (phenomenological renormalization) underlying the finite-size scaling studies summarized in Fig. 2. We derive the scaling properties of the crossing points as a function of the system size and use the 2D Ising model as a bench-mark case to demonstrate the unbiased nature of the method when all sources of statistical errors and scaling corrections are considered.

In **Sec. 2** we present the scaling arguments underlying the analysis of the domain-wall energy κ of the critical J-Q model in Fig. 3(a). We begin by the simpler case of a generic system with discrete symmetry-breaking at a critical point with a single divergent length scale, deriving the scaling form of κ in the thermodynamic limit and for finite size. We then generalize to the case when the thickness of the domain wall diverges faster than the correlation length and discuss the different scenarios for finite-size scaling at criticality. We use the 2D Ising model as an example to illustrate Monte Carlo (MC) procedures we have developed for computing the free-energy differences needed for κ at thermal phase transitions. We also demonstrate conventional finite-size scaling in a classical system with a dangerously irrelevant perturbation; the 3D six-state clock model. For the J-Q model at $T = 0$, we supplement the results shown in Fig. 3(a) with other calculations, where domain walls are introduced with different types of boundary conditions, demonstrating anomalous scaling with the same exponent ratio ν/ν' in all cases.

In **Sec. 3**, to further motivate the two-length scaling form (2) and its unconventional limiting behavior [demonstrated numerically in Figs. 3(b,c)], we present derivations of the quantum-critical scaling forms of the spin stiffness and magnetic susceptibility by generalizing the approach of Fisher *et al.* (*1*) to the case of two divergent length scales.

In **Sec. 4** we provide some more details of the $T > 0$ and $T = 0$ (ground-state projector) QMC methods used in the studies of the J-Q model.

1 Crossing-point analysis

The crossing-point analysis employed in Fig. 2 is an extension of Fisher's "phenomenological renormalization", following essentially the formalism developed and tested in Ref. (*29*), but applying it to MC data instead of numerically exact transfer-matrix results. In Sec. 1.1 we

discuss formalities and derivations of the exponents governing the drifts of crossing points in the standard case, when there is a single divergent length scale. In Sec. 1.2 we discuss why the single-length scaling form (1) can still be used to analyze crossing points and extract the exponent ν controlling the shorter length scale, even in the case when the criticality is described by the two-length ansatz (2) with the anomalous limit controlled by the longer length scale. In Sec. 1.2 we discuss several practical issues and potential error sources (statistical as well as systematical) that should be properly taken into account when analyzing crossing points. We illustrate the procedure with data for the 2D Ising model, demonstrating the unbiased nature of the approach by reproducing its exactly known critical temperature and critical exponents to within small statistical errors.

1.1 Scaling corrections and crossing points

Consider first the standard case of a single divergent length scale (correlation length) $\xi \propto |\delta|^{-\nu}$ as a function of the distance $\delta = g - g_c$ to a critical point (classical, driven by thermal fluctuations at temperature $T > 0$ or a quantum phase transition at $T = 0$). For some other singular quantity A with the behavior $A \propto |\delta|^{\kappa}$ in the thermodynamic limit (valid for $g < g_c$, $g > g_c$, or both, depending on the quantity) the finite size scaling is governed by the form

$$A(\delta, L) = L^{-\kappa/\nu} f(\delta L^{1/\nu}, \lambda_1 L^{-\omega_1}, \lambda_2 L^{-\omega_2}, \cdots), \tag{5}$$

where $0 < \omega_i < \omega_{i+1}$ and the variables λ_i are irrelevant fields which in principle can be tuned by introducing some other interactions in the Hamiltonian.[1] Keeping only the most important irrelevant field, using the notation $\omega \equiv \omega_1$ for convenience, and suppressing the dependence of the unknown value of λ_1, we have Eq. (1) in the main text. The scaling function is non-singular and we can Taylor expand it in the neighborhood of the critical point;

$$A(\delta, L) = L^{-\kappa/\nu}(a_0 + a_1 \delta L^{1/\nu} + b_1 L^{-\omega} + \ldots). \tag{6}$$

For two system sizes $L_1 = L$ and $L_2 = rL$, the two curves $A(\delta, L_1)$ and $A(\delta, L_2)$ take the same value (cross each other) at the point

$$\delta^* = \frac{a_0}{a_1} \frac{1 - r^{-\kappa/\nu}}{r^{(1-\kappa)/\nu} - 1} L^{-1/\nu} + \frac{b_1}{a_1} \frac{1 - r^{-\kappa/\nu-\omega}}{r^{(1-\kappa)/\nu} - 1} L^{-1/\nu-\omega}. \tag{7}$$

Thus, in general the finite-size value $g^*(L)$ of the critical point defined using such curve-crossing points shifts with the system size as $g^*(L) - g_c \equiv \delta^* \propto L^{-1/\nu}$. However, if the

[1] Contributions originating from the regular part of the free energy are also of the same form and in some cases, where the exponents of the irrelevant fields are large, they can cause the leading scaling corrections.

quantity A is size-independent at the critical point, $\kappa = 0$, the first term in Eq. (7) vanishes and the shift is faster;

$$g^*(L) - g_c \propto L^{-(1/\nu+\omega)}, \tag{8}$$

where the constant of proportionality depends on the chosen aspect ratio r and the generally unknown coefficients of the Taylor expansion (6). The value of the quantity A at the crossing point is obtained by inserting δ^* into Eq. (6), which for both the general case $\kappa \neq 0$ and the special case $\kappa = 0$ can be written as

$$A^*(L) = A(\delta^*, L) = L^{-\kappa/\nu}(a + bL^{-\omega} + \ldots), \tag{9}$$

with some constants a and b. Thus, in principle a crossing point analysis can be used to obtain the leading critical exponents κ and ν, as well as the subleading exponent ω. However, it should be noted that the higher-order terms in Eq. (6) can play a significant role for system sizes attainable in practice, and often $1/\nu + \omega$ and ω extracted from fitting to power laws according to Eqs. (8) and (9) should be considered only as "effective" exponents which change with the range of system sizes considered (with the correct exponents obtained only for very large system sizes where the subleading corrections become negligible). To extract the critical point, a dimensionless quantity ($\kappa = 0$) should be chosen as the convergence then is the the most rapid, given by Eq. (8). The value of the critical point g_c obtained from fitting to this functional form is normally not very sensitive to the imperfection of the power-law correction with the effective value of the exponent, as long as the fit is statistically sound.

There are many other ways of analyzing crossing points. For instance, the exponent ν can be obtained more directly than its difficult extraction based on the correction terms in the shift analysis above. Consider a dimensionless quantity Q, such as the Binder ratio (or the corresponding cumulant). We then have, including also some terms of higher order in Eq. (6),

$$Q(\delta, L) = a_0 + a_1\delta L^{1/\nu} + a_2\delta^2 L^{2/\nu} + b_1 L^{-\omega} + c_1\delta L^{1/\nu-\omega} + \ldots \tag{10}$$

and from the derivative $s(\delta)$ with respect to δ or $g = g_c + \delta$ we have

$$s(\delta) = \frac{dQ(\delta, L)}{d\delta} = \frac{dQ(g, L)}{dg} = a_1 L^{1/\nu} + c_1 L^{1/\nu-\omega} + a_2\delta L^{2/\nu} + \ldots. \tag{11}$$

We will now assume that $s(\delta)$ is positive in the region of interest, and if not we redefine it with a minus sign. At $\delta = 0$ we then have

$$\ln[s(0)] = c + \frac{1}{\nu}\ln(L) + dL^{-\omega} + \ldots, \tag{12}$$

with some constants c and d. Thus, for large L, $\ln(s)$ at the critical point depends linearly on $\ln(L)$ and the slope is the exponent $1/\nu$. A drawback of this method for extracting ν is that the

critical point has to be determined first, and a careful analysis should also take into account the uncertainties in the estimated value of g_c.

To circumvent the requirement of having to determine g_c first, we observe that, instead of evaluating the derivative (11) exactly at the critical point we can use the crossing point of the quantity Q for two system sizes $(L_1, L_2) = (L, rL)$ [or, as in Ref. (*29*), one can use $L_2 = L_1 + \Delta L$ with a constant ΔL, which only modifies some unimportant prefactors of the results derived below]. Inserting the crossing value (8) of δ into (11) we obtain

$$\begin{aligned} s(\delta^*, L_n) &= a_1 L_n^{1/\nu} + c_1 L_n^{1/\nu-\omega} + a_2 d L_n^{1/\nu-\omega} + \ldots \\ &= a_1 L_n^{1/\nu}(1 + \tilde{b}_1 L_n^{-\omega} + \ldots), \quad n = 1, 2. \end{aligned} \tag{13}$$

Having access to two different slopes at the crossing point, we can take the difference of the logarithms of these and obtain

$$\ln[s(\delta^*, rL)] - \ln[s(\delta^*, L)] = \frac{1}{\nu}\ln(r) + aL^{-\omega} + \ldots, \tag{14}$$

with some constant a. We can therefore define an exponent estimate $\nu^*(L)$ corresponding to the crossing point,

$$\frac{1}{\nu^*(L)} = \frac{1}{\ln(r)} \ln\left(\frac{s(\delta^*, rL)}{s(\delta^*, L)}\right), \tag{15}$$

and this estimate approaches the correct exponent at the rate $L^{-\omega}$ for large L;

$$\frac{1}{\nu^*(L)} = \frac{1}{\nu} + eL^{-\omega} + \ldots, \tag{16}$$

with some constant e and various higher-order terms again left out.

With all the crossing-point quantities discussed above, the infinite-size values g_c, Q_c, and $1/\nu$ can be obtained by fitting data for several system-size pairs (L, rL), using Eqs. (8), (9), and (16). One can either use the leading form as written with only the leading correction $L^{-(1/\nu+\omega)}$ (in the case of g_c) or $L^{-\omega}$ (for Q_1 and $1/\nu$) if the system sizes are large enough for the higher-order terms to be safely neglected, or one can include higher-order terms explicitly and fit to a larger range of system sizes. The former method has the advantage of the optimum fit being easier to find, while fits with multiple power laws are some times challenging or affected by large fluctuations in the parameters unless the statistical errors are very small.

1.2 The case of two length scales

We now turn to systems with two divergent lengths, where the critical scaling is governed by Eq. (2). When the thermodynamic limit corresponds to the scaling function $f(x, y)$ being a

power of the first argument $x = \delta L^{1/\nu'}$ for large x and y, the effect of the second argument $y = \delta L^{1/\nu'}$ is the same as in the standard case of a dangerously irrelevant field scaling as $L^{-\omega}$. The crossing-point analysis then remains the same as in the previous section. In the anomalous case, which we have termed the *super dangerous* perturbation, the second scaling argument (the longer length scale) generically controls the $L \to \infty$ behavior and demands the modified powers of L in front of the scaling function. This case requires some additional discussion.

In general, the scaling in this case is much more complex. In the main paper we have discussed how the correct thermodynamic limit is obtained when the scaling function is controlled by $y = \delta L^{1/\nu'}$. This limit corresponds directly to the intuitive physical picture of the shorter length ξ saturating at $L^{\nu/\nu'}$ when the longer length ξ' reaches L, and, therefore, ξ should not be replaced by L at criticality but instead by $L^{\nu/\nu'}$. This change imposes an anomalous power law at criticality for any observable which can be written as some nonzero power of the correlation length close to the critical point. It should be noted that, there are special non-generic observables, such as the Binder ratio, which by construction neither have any L-dependent prefactors of the finite-size scaling function f nor any dependence on ξ in the thermodynamic limit (e.g., the Binder ratio takes constant values in the phases and a different value at the critical point). In such non-generic cases there are also no modified power laws, since there are no powers to be modified by the ratio ν/ν' in the first place. All other generic observables are expected to develop anomalous power laws.

We next note that, in the above large L limit of $f(x, y)$, both the arguments $x = \delta L^{1/\nu}$ and $y = \delta L^{1/\nu'}$ become large. When we are interested in crossing points close to $\delta = 0$, we are far from this limit, however. We can anticipate crossing points as in the single-length case when the first argument x is of order one, and then the second argument is very small, $y \sim L^{1/\nu'-1/\nu} \ll 1$. There is no *a priori* reason to expect that this limit is also controlled by y. The most natural assumption, which can be tested, is that y is irrelevant in this regime. Then we are back at a situation where the standard crossing-point analysis can be performed and the exponent delivered by such an analysis should generically be ν, not ν'. An exception is an observable which is manifestly dependent only on the longer length scale, in which case the shorter length scale will play the role of an irrelevant correction. The simplest quantity of this kind is a length scale which is proportional to the longer length ξ' itself. In the main text we have analyzed the size Λ of the spinon bound state and found its crossing points to be controlled by an exponent exponent ν' which is indeed significantly larger than ν, and also $\Lambda \sim L$ holds in the neighborhood of the critical point, as expected from the scaling function controlled by y when $\Lambda \sim \xi'$ in the thermodynamic limit.

We now have concluded that the limits $y \to \infty$ and $y \to 0$ are controlled by different exponents in the generic case; by ν' in the former case and ν in the latter case. This implies an interesting cross-over behavior between these limits. In principle, such a cross-over can also

be tested explicitly by numerical data, by graphing results for a wide range of system sizes and couplings (in the case of the J-Q model inside the VBS phase) against both $\delta L^{1/\nu}$ and $\delta L^{1/\nu'}$. One should see scaling collapse into common scaling functions in both cases, but only in their corresponding regimes controlled by their respective exponents ν and ν'. It would clearly be desirable to carry out such an analysis for the J-Q model, which we have not yet done due to the large computational resources required to do this properly for sufficiently large system sizes. We anticipate the analysis of the cross-over to be complicated also by the a small exponent ω of the leading scaling corrections, as demonstrated in Fig. 2 in the main paper.

Even if no tests of the cross-overs are available currently, the two limits $y \to 0$ and $y \to \infty$ have already been confirmed in this work; the former by the scaling of the Binder cumulant with the exponent ν (the shorter length scale) and the latter more indirectly by the presence of anomalous powers of L. An anomalous exponent which is very well converged as a function of the system size and completely inconsistent with any other previous scenario (neither large scaling corrections nor a first-order transition) is best provided by the domain-wall energy κ, which is analyzed in Fig. 3 of the main paper and also further below in Sec. 2.4.

1.3 Tests on the 2D Ising model

In order to demonstrate the reliability of the method of obtaining the critical point and exponents from crossing points, and to discuss practical issues in implementing it, we here present results based on the Binder cumulant U of the standard 2D Ising model;

$$U = \frac{1}{2}\left(3 - \frac{\langle m^4 \rangle}{\langle m^2 \rangle^2}\right), \tag{17}$$

where m is the magnetization

$$m = \frac{1}{N}\sum_{i=1}^{N} \sigma_i, \quad \sigma_i \in \{-1, +1\}. \tag{18}$$

MC simulations were carried out on lattices of size $L \times L$ with periodic boundary conditions, using a mix of Wulff and Swendsen-Wang (SW) cluster updates, with each sweep of Wulff updates (where on average $\approx N$ spins are flipped) followed by an SW update where the system is decomposed into clusters, each of which is flipped with probability $1/2$. The SW clusters are also used to measure $\langle m^2 \rangle$ and $\langle m^4 \rangle$ with improved estimators (after each SW update). We carried out simulations of sizes $L = 6, 7, \ldots, 20, 22, \ldots, 36, 40, \ldots, 64, 72, \ldots, 128$, at $20 - 30$ temperatures in the neighborhood of the relevant crossing points of the Binder cumulant for system-size pairs $(L, 2L)$, i.e., using aspect ratio $r = 2$ in the expressions of Sec. 1.1. Up to 5×10^9 measurements were collected for the smaller sizes and 10^8 for the largest sizes.

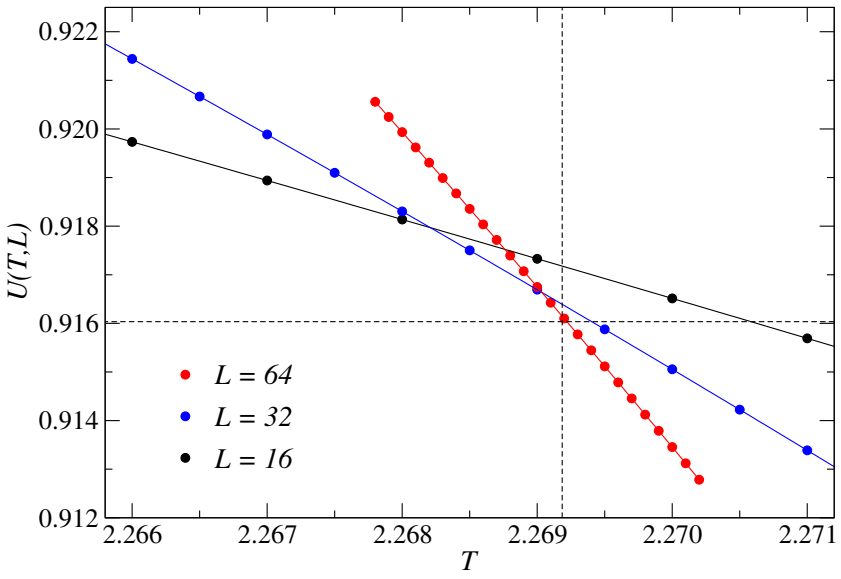


Figure S1: Binder cumulant of the 2D Ising model with $L = 16, 32, 64$ in the neighborhood of the points at which the curves cross each other. The vertical and horizontal dashed lines indicate the critical temperature T_c and the value of the cumulant at T_c, respectively. The solid curves are cubic polynomial fits to the data sets. Error bars are much smaller than the plot symbols.

Figure S1 shows examples of data for three different system sizes, where cubic polynomials have been fitted to the data. The crossing points are extracted numerically to machine precision using bisection. In order to analyze T_c and U_c in the thermodynamic limit, it suffices to consider a small number of points very close to each crossing point to be analyzed. To obtain ν from the slopes according to Eq. (15), where the derivative in Eq. (11) is taken of the fitted polynomials, it is better to have a more extended range of points. However, for a very large range a high order of the polynomial has to be used in order to obtain a good fit, and it is then better in practice to adapt the window size so that a relatively low order polynomial can be used. In the tests reported here, cubic polynomials were used and all fits were statistically sound.

In order to compute the statistical errors (error bars) a bootstrap method can be used, i.e., by generating a large number of random samples of the binned MC data. Each bootstrap sample is computed using $B(L,T)$ randomly chosen bins for each system size and temperature, where $B(L,T)$ is the total number of data bins available from simulations at (L,T). The standard deviations of the values (the horizontal and vertical crossing points and the slope-estimator for $1/\nu^*$) computed for these bootstrap samples correspond to the error bars, which later will be used in the fits to extrapolate to infinite size. In evaluating the cumulant (17), for the full data set or a bootstrap sample, the individual expectation values $\langle m_i^2 \rangle$ and $\langle m_i^4 \rangle$ should be computed first based on all the bins included in the sample, after which the ratio is evaluated. If one

instead uses ratios computed for each bin separately, a statistically significant systematical error can be introduced in the ratio, due to nonlinear contributions to the statistical error which do not vanish as the number of bins is increased (for fixed bin size) but do decrease properly in the bootstrap method when the sample size is increased.

We next fit crossing points for a series of system pairs to the expected forms, Eqs. (8), (9) with $\kappa = 0$, and (16), and compare with exact and previous numerical results for the 2D Ising model. Onsager's rigorous analytical solution gives $T_c = 2\ln^{-1}(\sqrt{2}+1) \approx 2.269185314$ and $\nu = 1$. The value of U at T_c is not known exactly, but Blöte obtained $U_c \approx 0.916035$ by extrapolating exact numerical finite-size transfer-matrix data to infinite size (*35*). For the Binder cumulant the dominant subleading correction has the exponent $\omega = 7/4$ (*35*). These results should all be obtained within statistical errors from the crossing point analysis of the MC data if sufficiently large systems are used and the data are analyzed using appropriate statistical methods. For small sizes the expected higher-order corrections will cause deviations beyond the statistical errors from the leading-order forms, which can be detected in the goodness of the fits to the leading forms (8),(9),(16). Our strategy is to remove small system sizes until a statistically sound fit is obtained for a given quantity.

The crossing points for the set of sizes $(L_i, 2L_i)$ are not all statistically independent, because the same system size can appear in two different pairs. One should therefore define the goodness of the fit, χ^2 per degree of freedom $N_{\rm dof}$ (the number of data points minus the number of parameters of the fit), with the full covariance matrix instead of just its diagonal elements (the standard error bars). Using V_i to denote some quantity defined based on the $(L_i, 2L_i)$ crossing point (the crossing temperature T^*, the value of U^* of U at the crossing point, or $1/\nu^*$ obtained from the slopes evaluated using the fitted polynomial), we thus use

$$\chi^2 = \sum_{i=1}^{M}\sum_{i=1}^{M}(\langle V_i\rangle - V_i^{\rm fit})[C^{-1}]_{ij}^2(\langle V_j\rangle - V_j^{\rm fit}), \tag{19}$$

where $\langle V_i\rangle$ is either the mean value obtained from all available bins or an average obtained from the bootstrap procedure discussed above (they should differ by an amount much smaller than the standard deviation based on the bootstrap analysis), $V_i^{\rm fit}$ is the value of the quantity evaluated using the fitted function (here a power-law correction to the infinite-size value), and M is the total number of system-size pairs used. The covariance matrix is defined as

$$C_{ij} = \Big\langle (V_i - \langle V_i\rangle)(V_j - \langle V_j\rangle)\Big\rangle, \tag{20}$$

where the expectation value for each pair i, j for which $C_{ij} \neq 0$ is again evaluated using bootstrap sampling (as explained above for the error bars, which correspond to the square-roots of the diagonal elements C_{ii}). We use of the order $100 - 1000$ bins and generate several thousand bootstrap samples to obtain accurate estimates of the covariance matrix.

To compute error bars on the extracted quantities, we repeat the fits to Eqs. (8), (9), and (16) several hundred times using the bootstrap method and define the final means and statistical errors (one standard deviation) using these bootstrap samples. For a statistically sound fit, $\langle\chi^2\rangle/N_{\rm dof} \approx 1$ is expected. To be more precise, we consider the standard deviation of the χ^2 distribution. For $N_{\rm dof}$ degrees of freedom, the standard deviation of $\chi^2/N_{\rm dof}$ is $(2/N_{\rm dof})^{1/2}$. We systematically eliminate the small sizes until $\langle\chi^2\rangle/N_{\rm dof}$ falls roughly within two standard deviations of its expected mean;

$$\frac{\langle\chi^2\rangle}{N_{\rm dof}} - 1 < 2\sqrt{\frac{2}{N_{\rm dof}}}. \tag{21}$$

Clearly this criterion is sensitive to the quality of the data—if the elements of the covariance matrix are very small, even fits including only relatively large system sizes can detect the presence of higher-order corrections and not pass our test, while with noisy data also small system sizes can be included (but the error bar on the final extrapolated value will be larger).

If a fit satisfies the goodness-of-fit criterion (21) it can still not be completely guaranteed that no effects of the higher-order corrections are present in the final result, but in general one would expect any remaining systematical errors to be small relative to the statistical error. In principle one can estimate the magnitude of the systematical error using the parameters obtained from the fit and some knowledge or estimate of the nature of the higher-order corrections. We will not attempt to do that here, because in general such knowledge will be very limited. To minimize possibly remaining systematical errors one can continue to exclude more system sizes even after the soundness criterion (21) is satisfied, at the price of increasing the statistical errors of the parameters extracted from the fits.

The above method implies a 'curse of good data', as less data points are actually included in the final fit when longer simulations are carried out for a fixed set of system sizes. However, the discarded data still contain valuable information on the convergence properties and in principle can be used to analyze higher-order scaling corrections (which we do not pursue here).

Results for the horizontal (temperature) and vertical (cumulant) crossing values of the 2D Ising model are shown in Fig. S2. For the horizontal points in (a), our fits start to satisfy the criterion (21) when including sizes $L \geq 12$ (the average goodness of the fit is then $\langle\chi^2\rangle/N_{\rm dof} \approx 1.6$ with $N_{\rm dof} = 20$) and we show that case in the figure. The fit gives $T_c = 2.2691855(5)$ and the exponent combination $1/\nu + \omega = 2.674(4)$. Thus, the critical temperature comes out correct within the remarkably small error bar, while $1/\nu + \omega$ is about twenty of its error bars outside the true value $1/\nu + \omega = 2.75$. As discussed above, it is typical in finite-size scaling that corrections-to-scaling exponents do not come out reliably until very large systems are used, and we therefore do not consider the mismatch as a failure here, rather as a confirmation of the known fact that the exponent should be considered as an "effective exponent" which slowly changes as larger system sizes are included.

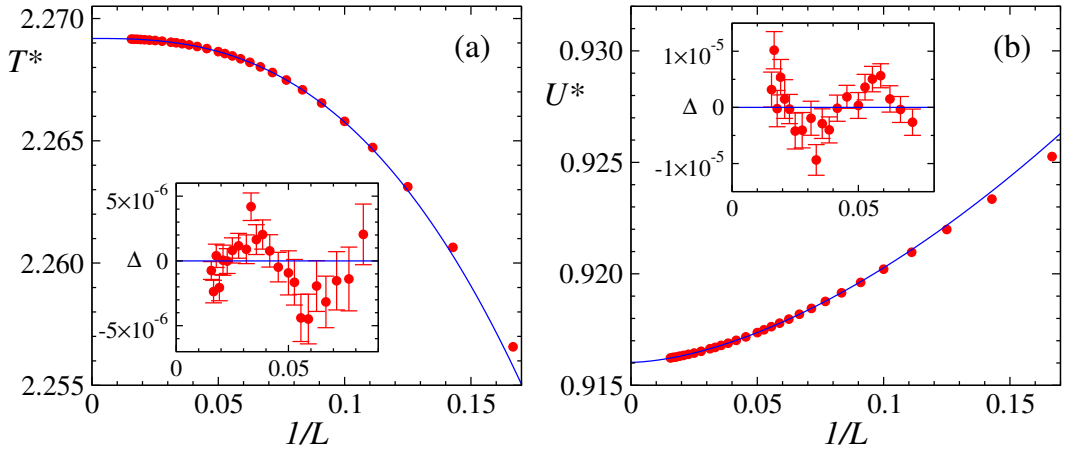


Figure S2: Results for the 2D Ising model. (a) Crossing temperature of the Binder cumulant for size pairs $(L, 2L)$ versus $1/L$, along with a fit of the $L \geq 12$ data to the form (8). (b) The cumulant at the crossing points, along with a fit to the form (9) for $L \geq 14$. In both (a) and (b), error bars are much too small to be visible. The insets shows the difference Δ between the data and the fitted functions including the error bars (for only the sizes included in the fits).

For the crossing value of the cumulant we find a similar trend. In this case a good fit requires that only the $L \geq 14$ points are used, giving the $U_c = 0.916031(3)$ and $\omega = 1.667(6)$, again with $\langle \chi^2 \rangle / N_{\rm dof} \approx 1.6$ ($N_{\rm dof} = 18$). The U_c value deviates by about an error bar from Blöte's result quoted above, while the correction exponent again is relatively far (considering the size of the error bar) from its asymptotic value $\omega = 1.75$. Interestingly, $1/\nu$ extracted as the difference of the two exponents comes out close to the correct value $1/\nu = 1$, within the statistical error.

The insets of Fig. S2 show the differences between the data points and the fitted curves. Here it can be seen that the points are not quite randomly distributed around 0, as they should be if the fitted functions are of the correct form. The overall shape with noisy minimums and maximums suggests the presence of a correction which is barely detectable for the range of system sizes at this level of statistics. One can then conclude that the deviations of $\langle \chi^2 \rangle / N_{\rm dof}$ by two standard deviations from 1 in these fits are not purely statistical fluctuations (which is not clear from the $\langle \chi^2 \rangle / N_{\rm dof}$ values alone), but due to the neglected higher-order corrections. Nevertheless, the most important extrapolated values T_c and U_c were not adversely affected statistically, thus demonstrating the ability of the effective exponent and the prefactor of the correction term in Eqs. (8) and (9) to reproduce the overall trend of the data sufficiently well for extrapolating to infinite size.

To illustrate the effect of excluding even more system sizes, with the minimum size $L = 28$ we obtain $T_c = 2.2691831(11)$, two error bars away from the correct value (still a statistically acceptable match), and $U_c = 0.916054(11)$, also about two error bars from the previous

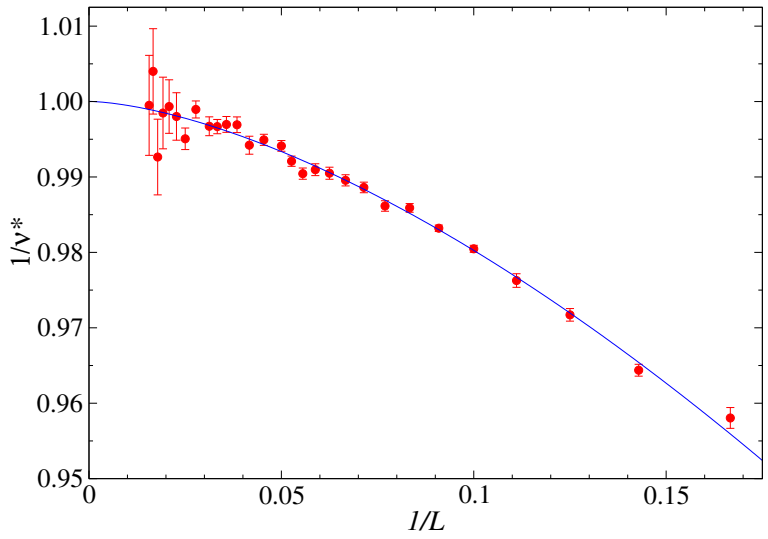


Figure S3: Estimates of the inverse of the correlation-length exponent ν of the 2D Ising model based on the slope expression (15) applied to the Binder cumulant. The curve is a fit to the form (8) including all points ($L \geq 6$).

(Blöte's) value. From the T_c fit we obtain $1/\nu + \omega = 2.70(4)$ in this case and from the U fit $\omega = 1.73(5)$. These exponents are now correct to within statistical errors, but the error bars are about 10 times as large as before, while the error bars on T_c and U only doubled. The average value of $\langle \chi^2 \rangle / N_{\rm dof}$ is very close to 1 for both these fits and the deviations from the fitted function look completely random. Upon excluding even more points, the error bars increase rapidly but the extracted parameters remain statistically in good agreement with their correct values.

Next, we extract the exponent ν using the log-slope formula (15). Fig. S3 shows the results along with a fit including all the system sizes ($L \geq 6$). Remarkably, the fit is statistically perfect, with $\langle \chi^2 \rangle / N_{\rm dof} \approx 1.0$, already at this small minimum size and the inverse exponent extrapolates to $1/\nu = 1.0001(7)$, in excellent agreement with the exact result 1. The slope data are much more noisy than the underlying U values and the error bars grow very rapidly with L for the largest sizes. The fit is therefore dominated by the smaller sizes. Naturally, the large error bars mask the effects of higher-order corrections, as discussed above. It is nevertheless remarkable that the extracted exponent $1/\nu$ does not show any effects of the neglected corrections at all, even though, again, the leading correction exponent, which comes out to $\omega = 1.57(7)$, is not very close to the correct value 1.75 and its error bar is large. Again, the flexibility of the leading finite-size term allows it to mimic the effects of the correction terms without significant effects in the extrapolation of the fit.

It should be noted that the 2D Ising model has logarithmic corrections in addition to the higher-order scaling corrections that we have neglected here (*35*), which is not a generic feature

of critical points (except for systems at their upper critical dimension). The logarithms of L multiply powers of L higher than those of the leading corrections and we therefore do not expect them to affect the procedures used above.

These results demonstrate the unbiased nature of the crossing-point analysis when it is carried out properly. We have used the same scheme to analyze the results for the J-Q model in Fig. 2 of the main text. In the left column, the behavior of Λ/L is similar to that of U of the Ising model in Fig. S2, with a relatively large correction exponent ω which makes the fits and extrapolations to $L \to \infty$ stable and and visually convincing. In the right column, it is clear that the leading correction exponent ω for R_1 is small, $\omega < 0.5$, and that there are other significant corrections present in the top two panels. The fact that the critical point nevertheless agrees perfectly to within small error bars with that extracted from the spinon bound state is very reassuring. As in the Ising model, the fit to $1/\nu^*$ only requires a single scaling correction, though it can not be excluded that this correction is an effective one, mimicking the collective effects of several corrections with the same sign. In any case, the extrapolations are stable, e.g., excluding some of the small-L points does not dramatically change the extrapolation, though of course the error bar grows.

We advocate the systematic curve-crossing method as outlined above to determine the critical temperature (or critical coupling of a quantum phase transition) and the critical exponents, instead of often used [also in DQC studies (*15, 20, 22*)] data-collapse techniques where many choices have to be made of the range of data included, use of corrections, etc. Although trends when increasing the system size can also be studied with data collapse [as done in Ref. (*20*))], the solid grounding of the present scheme directly to the finite-size scaling form (5) makes it the preferred method.

2 Domain-wall energy

As we discussed in the main text, the fundamental longer length scale ξ' in the DQC theory is the thickness of a domain wall in the VBS. In Fig. S4 we illustrate a generic domain wall in a 2D system in which a discrete symmetry is broken. In the case of a broken continuous symmetry, e.g., the magnetization vector in the XY spin model, there is no domain wall but the order parameter (its direction) gradually twists uniformly over the entire width L of the system. This case will be discussed in Sec. 3 in the context of a twist of the Néel order parameter of the J-Q model. For a discrete broken symmetry it is energetically favorable for the system to instead restrict the size of the region (the domain wall) over which the order parameter deviates significantly from the values imposed at the boundaries. Note, however, that the domain wall is not strictly fixed at some location, and, e.g., in an MC simulation the local order parameter will not detect the intrinsic width of a domain wall, because averaging is performed over all locations

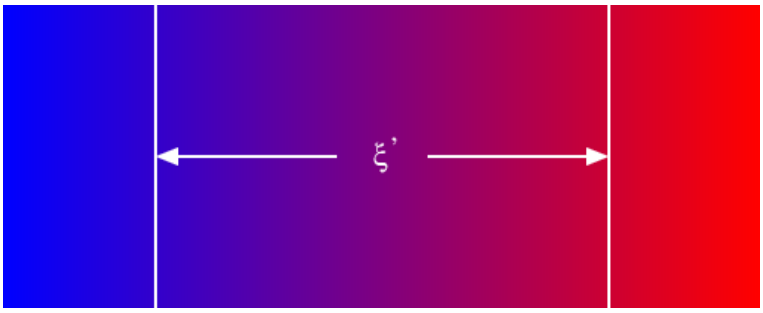


Figure S4: A domain wall in a generic 2D system where a discrete order parameter is locked at different values (directions) to the left and right and the twist between the two directions takes place over a region (domain-wall) of thickness ξ'.

of the wall. Therefore, other means have to be employed to detect the intrinsic domain-wall thickness, e.g., using suitably defined correlation functions.

As we showed in the main text, the length scale ξ' is conveniently present in the J-Q model in the finite-size scaling of the energy density κ of a VBS domain wall. Here, in Sec. 2.1 we derive the scaling form of κ, in the thermodynamic limit and for finite system size, using a simple ansatz generalizing the treatment by Fisher *et al.* (*1*) in a different context (considered further in Sec. 3) to the case of discrete symmetry breaking with two divergent length scales. The formalism applies both to classical and quantum systems. We present our MC procedures to compute κ at classical (thermal) phase transitions, using the 2D Ising model as a concrete example in Sec. 2.2. We also present results for the 3D classical six-state clock model in Sec. 2.3 before describing the details of the QMC calculations of κ for the J-Q model at $T = 0$ in Sec. 2.4.

2.1 Scaling forms

Let us first consider the case of a d-dimensional system with single divergent length scale $\xi \propto \delta^{-\nu}$. Following Fisher *et al.* (*1*), we consider the singular part of the free-energy density, which we can write for a classical system at finite temperature or a quantum system at $T = 0$ (in which case the free energy is just the ground state energy) as

$$f_s(\delta, L) \propto \delta^{\nu(d+z)} Y(\xi/L) \propto \xi^{-(d+z)} Y(\xi/L), \tag{22}$$

where the dynamic exponent $z = 0$ for a classical system. Introducing a domain wall, the free-energy difference with respect to the system without domain wall should scale in a similar way but with a different size-dependent function (*1*);

$$\Delta f_s(\delta, L) \propto \xi^{-(d+z)} \tilde{Y}(\xi/L). \tag{23}$$

This density should be understood as a quantity averaged over the inhomogeneous system (or, equivalently, in a finite system the domain wall location is not fixed and all properties are averages over all locations of the domain wall), and the total free-energy difference is

$$\Delta F_s(\delta, L) \propto \xi^{-(d+z)} \tilde{Y}(\xi/L) L^d, \tag{24}$$

where L^d is the volume of the system.

We can also write down a different expression for the free-energy difference, by explicitly considering the cost of twisting the order parameter. If the domain wall has width ξ and the total twist of the order parameter across the wall is $\Delta\phi$, then the cost per lattice link inside the wall is $\rho(\Delta\phi/\xi)^2$, which also defines the stiffness constant ρ. Outside the wall region the local energy cost vanishes, and, since the total volume occupied by the domain wall is ξL^{d-1} we have

$$\Delta F_s(\delta, L) = \rho(\Delta\phi)^2 \xi^{-1} L^{d-1}. \tag{25}$$

Consistency in the L dependence between this expression and Eq. (24) requires that the scaling function has the form $\tilde{Y} \propto \xi/L$, and therefore

$$\Delta F_s(\delta, L) \propto \xi^{-(d+z-1)} L^{d-1}. \tag{26}$$

The domain wall energy per unit volume of the wall is then

$$\kappa = \frac{\Delta F_s}{L^{d-1}} \propto \frac{1}{\xi^{d+z-1}}, \tag{27}$$

which no longer has any L dependence and, thus, represents the behavior in the thermodynamic limit. We can also read off the scaling of the stiffness constant,

$$\rho \propto \xi^{-(d+z-2)} \propto \delta^{\nu(d+z-2)}, \tag{28}$$

by comparing Eqs. (25). and (26).

Since we have written all expressions in terms of the correlation length, we can now switch to finite-size scaling at a critical point by simply making the substitution $\xi \to L$. For the domain wall energy (27) of interest here we obtain

$$\kappa(L) \propto L^{-(d+z-1)}. \tag{29}$$

Now consider a system with two length scales, with a conventional correlation length $\xi \sim \delta^{-\nu}$ and a domain wall thickness $\xi' \sim \delta^{-\nu'}$, with $\nu' > \nu$. A simple generalization of Eq. (24) suggests that

$$\Delta F_s(\delta, L) \propto \xi^{-(d+z)} \tilde{Y}(\xi/L, \xi'/L) L^d. \tag{30}$$

Note that only the shorter length scale should appear in front of the size-dependent scaling function $\tilde{Y}$ because the free energy in the thermodynamic limit should only depend on the two lengths in an additive way, $f_s = a\xi^{-(d+z)} + b\xi'^{-(d+z)}$, in order for the specific-heat exponent (α) relation $2 - \alpha = \nu(d + z)$ to hold, i.e., for hyper-scaling to apply (which we thus assume). Since ξ diverges slower than ξ', f_s is asymptotically dominated by the ξ term, and (30) should then describe the leading singular behavior.

We can also easily generalize Eq. (25) to a domain wall of thickness ξ';

$$\Delta F_s(\delta, L) = \rho(\Delta\phi)^2 \xi'^{-1} L^{d-1}. \tag{31}$$

Now consistency between Eqs. (30) and (31) for both the L dependence and the ξ' dependence requires that $\tilde{Y} \propto (L/\xi')(\xi^2/L^2)$, and we arrive at

$$\kappa \propto \frac{1}{\xi^{d+z-2}\xi'} \tag{32}$$

for the scaling of κ in the thermodynamic limit. Note the consistency of this form and the single-length form (27) when $\xi' \to \xi$. In the particular case of a DQC point ($d = 2$, $z = 1$), Eq. (32) reduces to $\kappa \propto (\xi\xi')^{-1}$, which was derived in a different way by Senthil *et al.* (*4*).

To convert Eq. (32) to finite-size scaling, in the standard treatment of two length scales arising from a dangerously irrelevant perturbation (*31*), the longer scale is not present in the leading finite-size scaling behavior. This can be understood physically as follows: Upon approaching the critical point from the ordered phase, when ξ' reaches L we simply replace ξ' by L. However, ξ continues to grow and controls the scaling behavior until it reaches L. At the critical point, also ξ is replaced by L and the critical finite-size scaling of κ obtained from (32) is, thus, identical to the single-length form (29). Since neither ν nor ν' appear here, there is no information on these exponents in the finite-size scaling of κ in the standard scenario.

As we have argued in the main text, there is also another possibility, namely, the growth of ξ is arrested at its value $\xi \propto L^{\nu/\nu'}$ when ξ' reaches L, leading to the finite-size scaling form

$$\kappa(L) \propto L^{-1-(d+z-2)\nu/\nu'}. \tag{33}$$

In the case of DQC, this reduces to $\kappa \propto L^{-(1+\nu/\nu')}$. It is very interesting that the ratio ν/ν' appears here in a simple way and can be extracted using critical finite-size scaling. The result in Fig. 3(a) leaves little doubt that $\kappa < 2$, which represents unambiguous evidence for anomalous scaling in the J-Q model. Below, in Sec. 2.4, we will present several different ways of computing the domain wall energy which all give the same value $\nu/\nu' \approx 0.72$.

2.2 2D Ising model

It is instructive to first test the domain-wall scaling using a simple system such as the 2D Ising model. A domain wall in the ferromagnet can be enforced in different ways using suitable

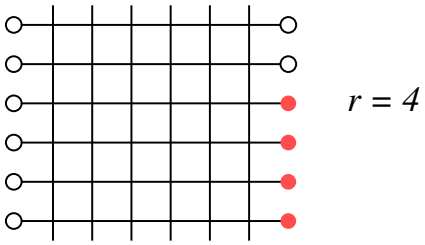


Figure S5: Boundary conditions used to induce a domain wall in the 2D Ising ferromagnet. The black open circles and red filled circles indicate down and up boundary spins, respectively. The vertical location r denotes the point at which the domain-wall inducing boundary is terminated. This location is updated, $r \to r \pm 1$, in MC updates in addition to the updates of the bulk spins. A full vertical domain wall is present when $r = L$.

boundary conditions. Here we use $L \times L$ systems with periodic boundaries in the y-direction and compare two different x boundaries, as illustrated in Fig. S5. The boundaries are open, with the edge columns coupled with the same strength J as the bulk coupling to fixed spins $\sigma_i = +1$ and $\sigma_i = -1$, equivalent to boundary fields of strength $\pm J$. Here the domain-wall imposing column of spins to the right extends only partially through the system, to illustrate the mechanism we use for computing the required free-energy difference.

It is not easy to compute the free energy in MC simulations, but it is relatively easy to compute a free-energy *difference*, if the two systems of interest, let us call them "1" and "2", can be simulated collectively as a partition function $Z_{12} = Z_1 + Z_2$. If there are updates switching the simulation between system states 1 and 2 with detailed balance satisfied, then the free-energy difference $\Delta F_{21} = F_2 - F_1 = \ln(Z_2/Z_1) = \ln(P_2/P_1)$, where P_1, P_2 are the probabilities of the simulation "visiting" the respective states. Such *multi-canonical* simulations (*36*) can be extended to an arbitrary number of systems $s = 1, \ldots, n$, and any F_{ij} can then be accessed, provided that the simulation can easily transition between the different states s.

In the studies of domain walls considered here, the different systems correspond to boundary conditions fluctuating between the normal periodic boundaries and the domain-wall boundaries. To enhance the ability of the system to fluctuate between these boundary conditions of interest, the whole boundary is not changed at once, but in small steps where the right boundary has a change from $\sigma_i = -1$ to $\sigma_i = +1$ at some vertical location $y = r$, as illustrated in Fig. S5. Thus, $r = 0$ corresponds to the normal periodic boundaries (no domain wall) and $r = L$ corresponds to the boundary enforcing a full vertical domain wall. For $0 < r < L$ the domain wall does not extend vertically through the whole system and instead has a horizontal part connecting to the location $y = r$ where the boundary changes. MC updates are used to move this location, $r \to r \pm 1$, using heat-bath acceptance probabilities.

We find that the probability $P(r)$ of the boundary conditions generated is the highest, as

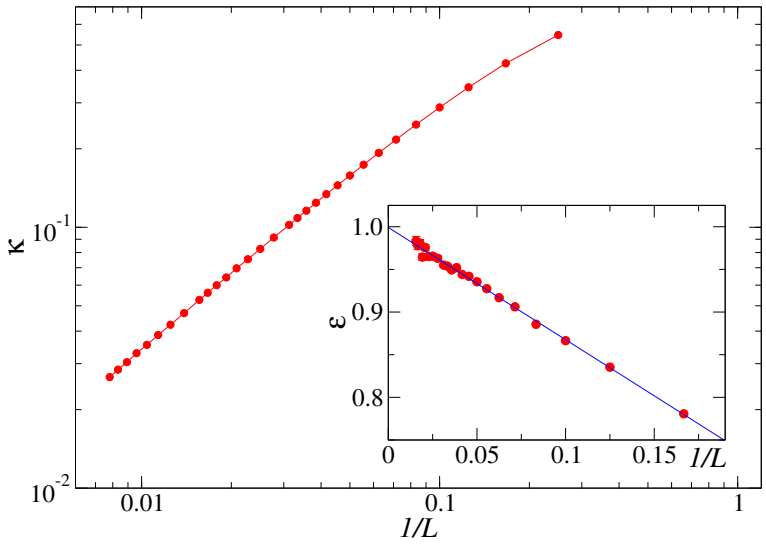


Figure S6: Scaling of the domain-wall energy per unit length in the 2D Ising model at the critical temperature. The inset shows the running decay exponent obtained from data pairs $\kappa(L)$ and $\kappa(2L)$ as $\epsilon(L) = \ln[\kappa(L)/\kappa(2L)]/\ln(2)$. The results have been fitted to a straight line, which extrapolates to the expected value, $\epsilon \to d - 1 = 1$, for $L \to \infty$.

expected, for $r = 0$. There is also a local maximum at $r = L$, and a minimum around $r = L/2$. To further increase the efficiency of the boundary moves, a weight factor $V(r)$ is multiplied with the Boltzmann probability for the spins and gradually adjusted such that the histogram $H(r)$ until the relative number of times the boundary is at r becomes almost flat. Then, the actual probability without the re-weighting factor is $P(r) = H(r)/V(r)$, and the free-energy difference between the systems with and without domain wall is (leaving out the unimportant temperature factor),

$$\Delta F = \ln\left(\frac{P(L)}{P(0)}\right). \tag{34}$$

MC results for κ are shown in Fig. S6. The inset shows the running exponent $\epsilon(L)$ extracted on the basis of size pairs $(L, 2L)$ by postulating $\kappa(L) = aL^{-\epsilon(L)}$ and $\kappa(2L) = a(2L)^{-\epsilon(L)}$, whence $\epsilon(L) = \ln[\kappa(L)/\kappa(2L)]/\ln(2)$. The results are fully compatible with $\epsilon(L) \to 1$ when $L \to \infty$, as predicted by Eq. (29) when $d = 2, z = 0$, with a correction $\propto L^{-1}$. We have also carried out simulations of the 3D Ising model and confirmed that $\epsilon(L) \to 2$.

2.3 3D clock model

The existence of two length scales in the DQC theory relies heavily (*3,4*) on an analogy with the classical 3D clock model, where the standard XY model is deformed by an external potential

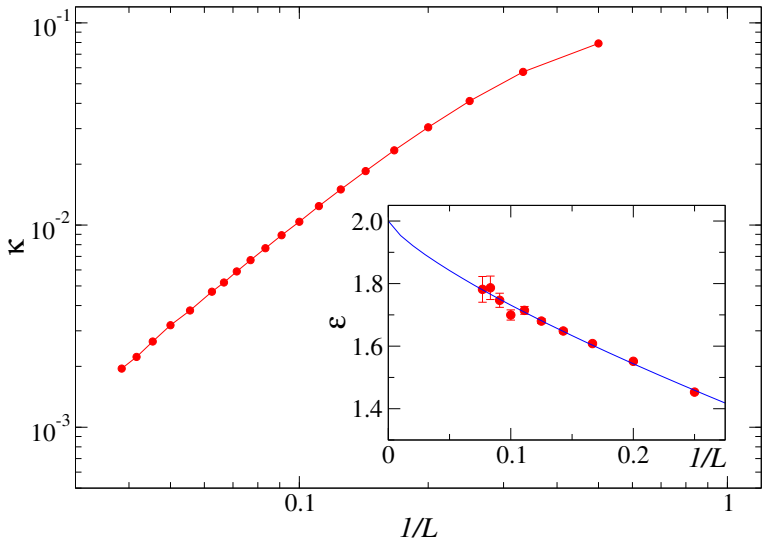


Figure S7: Scaling of the domain-wall energy per unit length in the 3D classical $q = 6$ clock model at its critical point ($T_c/J \approx 2.202$). The inset shows the running exponent obtained from data pairs $\kappa(L)$ and $\kappa(2L)$ as $\epsilon(L) = \ln[\kappa(L)/\kappa(2L)]/\ln(2)$ and a fit to the form $\epsilon(L) = 2 - aL^{-\omega}$ with $\omega \approx 0.74$.

$h\cos(q\Theta_i)$ for all the angles Θ_i. This term is known to act as a dangerously-irrelevant perturbation, leading to a domain-wall thickness $\xi' > \xi$. It is therefore natural to also test the scaling of the domain-wall energy in this case. Here we use the standard XY interaction between nearest neighbors on the 3D simple cubic lattice

$$H_{\rm XY} = -J\sum_{\langle ij\rangle}\cos(\Theta_i - \Theta_j), \tag{35}$$

where the angles are constrained to the q clock angles, $\Theta_i = n2\pi/q$, $n = 0, 1, \ldots, q-1$. The hard constraint is equivalent to the limit $h/J \to \infty$ with the cosine perturbation.

The exponent ν' should be independent of h/J (including the fully-constrained limit considered here) but depends on q, diverging as $q \to \infty$. There has been some controversy regarding methods to compute the exponent in MC simulations, as summarized in the recent Ref. (*31*), but for small q several calculations are nevertheless in good agreement with each other and we can use them as reference points.

In order for the exponent ratio ν/ν' to be significantly different from one we here use $q = 6$, in which case $\nu' \approx 1.44$ and, since the 3D XY exponent $\nu \approx 0.67$, the ratio $\nu/\nu' \approx 0.47$. Results for the domain-wall energy scaling at the critical point are shown in Fig. S7. The results are completely consistent with the form (29) with $d = 3, z = 0$, corresponding to the expected standard scenario where finite-size scaling is obtained from the thermodynamic-limit form by replacing both divergent length scales by L. The results are completely inconsistent with the

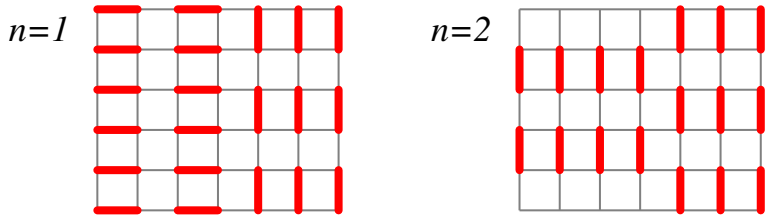


Figure S8: Simplified pictures of VBS domain walls with total twist angle $n\pi/2$, $n = 1, 2$, of the order parameter between the left and right boundaries. In the notation introduced in the text, the boundary conditions of these two cases are denoted as (h, v_1) and (v_2, v_1). In QMC simulations the dimerization at the open x boundaries is induced by weakening some of the interactions, thus explicitly breaking the symmetry between the possible VBS patterns. Periodic boundary conditions are employed in the y direction.

alternative scenario (33), where the decay exponent should approach $1 + \nu/\nu' \approx 1.47$. This result reinforces the unusual scaling of κ in the J-Q model, Fig. 3(a) of the main text, which we will discuss in more detail in the next section.

We also comment on the applicability of the generic two-length scaling form (2) in the main paper to κ in the clock model. Using the finite-size scaling we found above, we should have

$$\kappa(\delta, L) = L^{-2} f(\delta L^{1/\nu}, \delta L^{1/\nu'}). \tag{36}$$

To obtain the correct thermodynamic limit, $\kappa \to (\xi\xi')^{-1}$ when $L \to \infty$, we must have $f(x, y) \to x^\nu y^{\nu'}$, which is also natural because, given the form in the thermodynamic limit, f should be separable, $f(x, y) = f_x(x) f_y(y)$, where the two factors just correspond to the expected scaling forms for the length scales ξ and ξ' themselves. In contrast, in the J-Q model we have argued for an anomalous form which corresponds to a generally non-separable scaling function with the thermodynamic limit controlled only by the second argument.

2.4 J-Q model

In the J-Q model we are interested in ground state energies of systems with and without domain walls and these can be computed in standard QMC simulations. The multi-canonical approach employed in the previous section, developed to circumvent the difficulties of MC calculations of individual free energies at $T > 0$, are therefore neither useful nor needed. We use the projector QMC approach with $\mathrm{e}^{-\beta H}$ applied to a valence-bond trial state of the amplitude-product type (*37, 38*), choosing the "projection time" β sufficiently large, up to $\beta = 4L$, to converge the ground-state energy. Domain walls are introduced by boundary conditions in two different ways, schematically illustrated in Fig. S8.

The VBS order parameter is a vector $\mathbf{D} = (D_x, D_y)$, where the operators corresponding to the two components can be defined as

$$\hat{D}_x = \frac{1}{N}\sum_{i=1}^{N}(-1)^{x_i}\mathbf{S}_{x_i,y_i}\cdot\mathbf{S}_{x_i+1,y_i}, \quad \hat{D}_y = \frac{1}{N}\sum_{i=1}^{N}(-1)^{y_i}\mathbf{S}_{x_i,y_i}\cdot\mathbf{S}_{x_i,y_i+1}, \tag{37}$$

where (x_i, y_i) are the integer lattice coordinates of site i. Inside a columnar VBS phase of a large system, a histogram of the order parameter generated from the estimators of $\hat{D}_x$ and $\hat{D}_y$ in QMC simulations exhibits sharp peaks at the points $(1,0)$, $(0,1)$, $(-1,0)$, $(0,-1)$ times the magnitude D of the order parameter. These peaks are correspond to angles $m\pi/2$, $m = 0, 1, 2, 3$. As the critical point is approached, in simulations of the J-Q model the histograms develop an $U(1)$ symmetry, becoming completely circular symmetric at the DQC point (*15, 17*). The length scale ξ' controls this emergent $U(1)$ symmetry (*3*); upon course-graining the order parameter on length scales larger than ξ' the discrete Z_4 symmetry of the VBS is apparent, while on shorter length-scales $U(1)$ symmetry develops. The thickness of a domain wall forced by suitable boundary conditions is controlled by this same length scale.

The four-fold symmetry of the VBS on the square lattice allows for two different types of boundary conditions, as illustrated in Fig. S8. In the case labeled $n = 1$, the left and right side of the lattice is forced to have VBS order with horizontal and vertical dimers, respectively, which corresponds to an angular difference of the order parameter $\Delta\phi = \pi/2$. In the $n = 2$ graph, there is vertical dimer order at both edges, but with a relative shift of one lattice spacing, corresponding to an angular mismatch of $\Delta\phi = \pi$. In a large system, the elementary domain wall corresponds to $\Delta\phi = \pi/2$ and a π wall splits into two such elementary walls.

To compare the two cases and check for possible effects of interactions between two domain walls on the scaling of the energy, we have carried out projector QMC simulations with domain walls induced with total twist angles $\Delta\phi = \pi/2$ and π. Simulations without domain walls were carried out with similar boundary conditions, but with both the left and right walls at the same VBS angle ϕ. The energy differences can then be computed without any remaining effects of edge contributions to the total energy, which for a given type of edge is the same with and without domain walls present in the bulk. Denoting boundary conditions enforcing horizontal dimerization at one of the edges (as in the left edge of the $n = 1$ graph in Fig. S8) by h and vertical order with the two different phases (as shown in the $n = 2$ graph) by v_1 and v_2, the systems we study with different combinations of left and right boundaries are (h, h), (v_1, v_1), (h, v_1), and (v_2, v_1). The v_1 and v_2 boundaries are related by just a translation and therefore the edge contribution to the energy from these are the same. The domain wall contributions to the energy with the edge effects eliminated are then

$$\Delta E(\pi/2) = E(h, v_1) - [E(h, h) + E(v_1, v_1)]/2, \tag{38}$$

$$\Delta E(\pi) = E(v_2, v_1) - E(v_1, v_1), \tag{39}$$

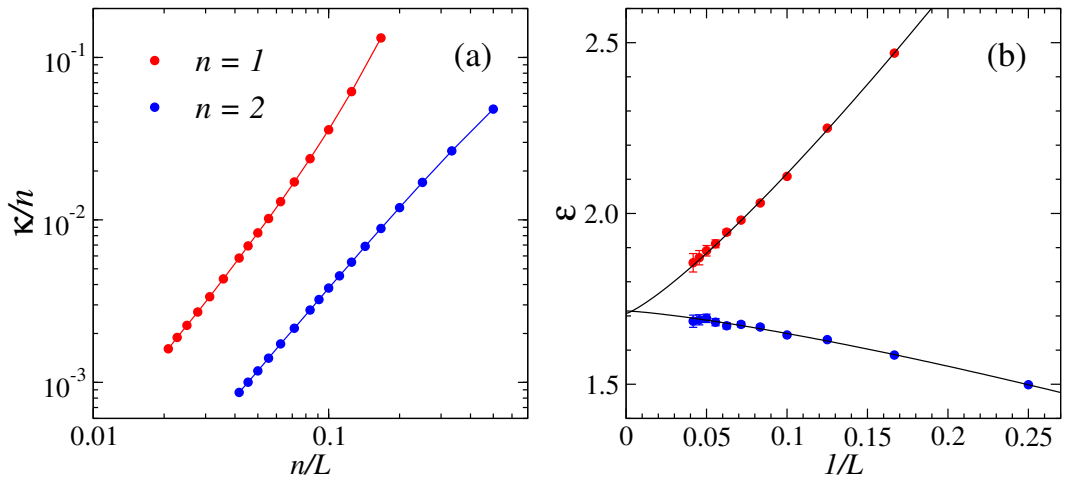


Figure S9: (a) Domain-wall energy in the critical J-Q model. (b) The exponent $1 + \nu/\nu'$ extracted from the data in (a) as a running exponent from system-size pairs $(L, 2L)$. Fits with power-law corrections including all data points are shown.

and the corresponding size-normalized energy density is $\kappa(\Delta\phi) = \Delta E(\Delta\phi)/L$.

QMC results for κ computed at the estimated critical point $J/Q = 0.0447$ are shown in Fig. S9(a) [where the the $\kappa = \pi$ results are the same as those already presented in Fig. 3(a)]. Here, to compare the energies on an equal footing, we divide κ by the number $n = 1, 2$ of domain walls induced when the VBS twist angle is $\Delta\phi = n\pi/2$ and plot the results against $(L/n)^{-1}$, L/n being the width over which a single domain wall is (on average) distributed. It is interesting, and at first sight surprising, that the $\pi/2$ domain wall is energetically much more expensive, since one would not expect any significant attractive interactions between the two domain walls in the $\Delta\phi = \pi$ case. We find that the lowering of the energy is due to enhanced fluctuations in the system with two domain walls. Recalling the emergent $U(1)$ symmetry discussed above and considering a $\pi/2$ domain wall between, say, boundaries at $\phi = 0$ and $\phi = \pi/2$, we expect the VBS angle in the center of the system to fluctuate mainly between these angles. In the case of the $\Delta\phi = \pi$ twist, there are similarly fluctuations between the angles at the edge, say $\phi = 0$ and $\phi = \pi$, but here the system has two possible paths to go between the edges, passing either through $\phi = \pi/2$ or $\phi = -\pi/2$. Since the system is critical, there is no reason to expect any breaking of this symmetry,

By constructing histograms of the order parameter we have confirmed these behaviors for moderate system sizes, while for larger systems the amplitude of the order parameter is reduced due to the critical nature of the domain walls and the histograms in both cases develop $U(1)$ symmetry. These results confirm that the system with π twist is "softer" than that with $\Delta\phi = \pi/2$, explaining the large overall differences between the $n = 1$ and $n = 2$ results in Fig. S9(a).

Apart from the different overall magnitudes, the power-law decay of κ with L for the largest systems is similar for $n = 1$ and 2. Fig. S9(b) shows the running exponents $\epsilon(L)$ extracted from system sizes $(L, 2L)$ in the same way as discussed in Sec. 2.2 for the Ising model. The two data sets asymptotically extrapolate to the same exponent, which we have argued is $1 + \nu'/\nu$, with $\nu/\nu' = 0.715(15)$. The corrections are perfectly captured by a power-law correction $\propto L^{-\omega}$, with the same exponent $\omega \approx 1.2 - 1.3$ but different signs of the prefactor. We have also carried out calculations slightly away from the estimated critical coupling, at $J/Q = 0.0450$ and 0.0445, and there are no significant differences in the exponent ratio extracted at these points.

These results are key to our claims of anomalous finite-size scaling in the J-Q model, as it is not possible to explain a non-integer decay exponent $\epsilon < 2$ for the domain walls within the conventional quantum-criticality scenario (as discussed above in Sec. 2.1), and the results also are completely inconsistent wit a first-order transition. In the latter case, VBS and Néel order would coexist at the transition point and a domain wall induced in the way explained above could possibly also be affected by coexistence inside the domain wall. However, regardless of the nature of the domain wall, the energy cost of the interface must scale linearly with the length of the domain wall, giving a finite κ and a vanishing exponent $\epsilon(L)$ when $L \to \infty$. Thus seems extremely unlikely given our data in Fig. S9(b).

In Ref. (*38*) we employed a different approach to studying domain walls in *periodic* systems, by restricting the trial state used in projector QMC simulations in the valence-bond basis to a topological (winding number) sector corresponding to the presence of a number of domain walls. There we also found anomalous scaling for κ, but with a somewhat larger exponent ratio $\nu/\nu' = 0.80(1)$, for a different variant of the J-Q model with products of three singlet projectors (the J-Q_3 model) instead of the two projectors used in the model (3) (the J-Q_2 model). We have also repeated this kind of calculation for the J-Q_2 model and again found $\nu/\nu' \approx 0.80$ for systems of small and moderate size. However, when larger systems are considered and the statistical accuracy is sufficiently high, drifts in the exponent toward smaller values become apparent. The asymptotic behavior is consistent with $\nu/\nu' \approx 0.72$ obtained above with the symmetry-breaking boundaries. The previous results in Ref. (*38*) were likely affected in the same way by remaining scaling corrections, and $\nu/\nu' \approx 0.72$ should hold universally for different variants of the J-Q model and for different ways of generating domain walls.

3 Finite-size scaling of the spin stiffness and susceptibility

In the main text we discussed the generic two-length finite-size scaling form (2) and its different limiting behaviors compatible with the the correct scaling of physical quantities in the thermodynamic limit. Here we discuss the behavior in the thermodynamic limit further, deriving the standard forms assumed in the main text for the spin stiffness ρ_s and the susceptibility χ in the

presence of two divergent length scales. We then argue for the unconventional size scaling. The scaling arguments generalize similar treatments by Fisher *et al.* (*1*) for a system with a single relevant length scale to a quantum phase transition with two divergent length scales, in a way analogous to the treatment of the domain-wall energy in the previous section.

The standard scenario of Fisher *et al.* (*1*) was formulated for interacting bosons and gives the scaling behaviors of the superfluid stiffness and the compressibility. The same formalism applies to a spin system as well (*2*), where the corresponding quantities are the spin stiffness ρ_s and uniform magnetic susceptibility χ, which we will use in the notation here. As in Sec. 2, we again start from the singular part of the free-energy density,

$$f_s(\delta, L, \beta) \propto \delta^{\nu(d+z)} Y(\xi/L, \xi^z/\beta), \tag{40}$$

where we now explicitly include the dependence on the inverse temperature β, which was assumed to be zero in the case of the quantum system ($z > 0$) in Sec. 2. In the end we will consider $\beta \to \infty$ but we will need finite β in the derivation of the susceptibility.

Upon imposing, by suitable boundary conditions, a total spatial phase twist $\Delta\phi$ of the continuous Néel order parameter uniformly distributed over the system, the increase in free energy is given by

$$\Delta f_s(\delta, L, \beta) = \rho_s \frac{(\Delta\phi)^2}{L^2} \propto \delta^{\nu(d+z)} \tilde{Y}_r(\xi/L, \xi^z/\beta). \tag{41}$$

Internal consistency of this scaling form demands that $\tilde{Y}_r$ behaves as $(\xi/L)^2$, thus,

$$\rho_s \propto \xi^2 \delta^{\nu(d+z)} \propto \delta^{\nu(d+z-2)}. \tag{42}$$

Similarly, $\chi(\Delta\phi)^2/\beta^2$ is the excess energy density needed to enforce a twist between $\tau = 0$ and $\tau = \beta$ in the imaginary-time direction;

$$\Delta f_s(\delta, L, \beta) = \chi \frac{(\Delta\phi)^2}{\beta^2} \propto \delta^{\nu(d+z)} \tilde{Y}_\tau(\xi/L, \xi^z/\beta), \tag{43}$$

where $\tilde{Y}_\tau$ has to behave as $(\xi^z/\beta)^2$. Thus, the susceptibility scales as

$$\chi \propto \xi^{2z} \delta^{\nu(d+z)} \propto \delta^{\nu(d-z)}. \tag{44}$$

The finite-size scaling properties at the critical point are simply obtained from Eqs. (42) and (44) by replacing $\delta^{-\nu} \propto \xi$ by the system length L, leading to

$$\rho_s \propto L^{-(d+z-2)}, \quad \chi \propto L^{-(d-z)}. \tag{45}$$

In the case of $z = 1$ (as in the DQC theory) both quantities scale as $1/L$ but note that the dependence on z is opposite for the two, which implies that the behavior seen in Fig. 3 in the main text can not be explained simply by $z \neq 1$.

We now generalize the above derivations to the case of two divergent length-scales, ξ and ξ', writing the free energy density as

$$f_s(\delta, L, \beta) \propto \delta^{\nu(d+z)} Y(\xi/L, \xi^z/\beta, \xi'/L, \xi'^z/\beta), \tag{46}$$

where we have made the assumption that the same dynamic exponent governs the two time scales associated with ξ and ξ' (and in principle we can generalize to two different exponents z and z'). The excess energy due to a spatial twist is

$$\Delta f_s(\delta, L, \beta) = \rho_s \frac{(\Delta\phi)^2}{L^2} \propto \delta^{\nu(d+z)} \tilde{Y}_r(\xi/L, \xi^z/\beta, \xi'/L, \xi'^z/\beta). \tag{47}$$

Here, at first sight, there are many ways in which $\tilde{Y}_r$ can depend on its arguments in order to contain the correct L dependence;

$$\tilde{Y}_r \propto \left(\frac{\xi}{L}\right)^a \left(\frac{\xi'}{L}\right)^{2-a}, \tag{48}$$

with arbitrary exponent a. However, upon approaching the critical point, when the longer length reaches L, we have $\xi'/L \approx 1$ and the only dependence on L at that point is in the factor $(\xi/L)^a$. Thus, we can argue that $a = 2$. For the thermodynamic limit we therefore reproduce the standard results, Eq. (42). In a similar way we also reproduce Eq. (44) for the susceptibility.

For the finite-size scaling there are two physically natural options, following from two possible behaviors of the shorter length scale ξ upon further approaching the critical point when ξ' has already reached L: (i) ξ continues to increase and eventually reaches L. The standard finite-size scaling forms (45) are then again obtained by replacing $\xi \propto \delta^{-\nu}$ by L. (ii) The two length scales are fundamentally tied together, and once ξ' has saturated ξ is locked into its corresponding value; $\xi \propto (\xi')^{\nu/\nu'} \propto L^{\nu/\nu'}$. Making this replacement in the thermodynamic-limit forms (42) and (44) leads to

$$\rho_s \propto L^{-(d+z-2)\nu/\nu'}, \quad \chi \propto L^{-(d-z)\nu/\nu'}, \tag{49}$$

exactly as we argued in the main text based on a direct finite-size scaling ansatz with an appropriate limit of the scaling function. As was shown in Fig. 3 in the main text, the forms (49) are in excellent qualitative agreement with data for the J-Q model, with both ρ_s and χ decreasing slower with L than in the standard forms (45). Quantitative agreement is observed when using the exponent ratio ν/ν' extracted from the scaling of the domain-wall energy.

4 Quantum Monte Carlo simulations

The QMC calculations of the spin stiffness and susceptibility were carried out with the standard Stochastic Series Expansion algorithm, using the same program as in Ref. (*18*), to which we

refer for technical references. For a given system size, the method produces unbiased results only affected by well-characterized statistical errors of the MC sampling.

Ground-state calculations in both the $S = 0$ and $S = 1$ sector were carried out with projector QMC simulations in the basis of valence bonds (singlet pairs) and unpaired spins, following Refs. (*18, 32*) and references cited there [see also Ref. (*38*)]. For system size N and total spin S, there are $(N - 2S)/2$ valence bonds and $2S$ unpaired spins with the same z-spin projection, i.e., the total spin-z projection of the state $S^z = S$. The degrees of freedom of a bra and ket state are importance-sampled, using the overlap of the two states as the sampling weight. This overlap is represented by a transition graph, where, in the case of the ground state with $S = 0$, the bonds form closed loops. For $S > 0$ there are $2S$ "open loops", or strings, where for a given string the end points fall on two unpaired spins, one in the bra and one in the ket. Such a configuration for $S = 1$ is illustrated in Fig. 1 of the main text.

This string connecting an unpaired spin in the bra and ket states is a representation of a spinon. The statistics of the individual strings and their cross-correlations provide information on the nature of the spinons and their collective states. In particular, the size of the lowest-energy $S = 1$ spinon bound state in a VBS can be defined in simulations with two unpaired spins. In Ref. (*33*) the distance between the unpaired spins (the end points of the strings) were used for this purpose. Here we use a slightly different measure, inspired by the arguments of Ref. (*34*) for a different problem,[2] using the entire strings in the following way: Each of the two spinons, 1, and 2, in the $S = 1$ state is associated with a string covering lattices sites located at $r_1(i)$, $r_2(j)$, with $i = 1, \ldots, n_1$ and $j = 1, \ldots, n_2$. We average the distance-squared $r_{ij}^2 = |\vec{r}_1(i) - \vec{r}_2(j)|^2$ between two points on the two strings over all the $n_1 n_2$ pairs of lattice sites and define the size as $\Lambda = \langle r_{ij}^2 \rangle^{1/2}$. We find that this definition provides a clearer signal of the spinon bound state diverging faster than the correlation length than definitions of Λ based on just the unpaired spins used in Refs. (*32, 33*).

[2]We thank Kedar Damle for suggesting a definition based on the entire spinon string.

PHYSICAL REVIEW A VOLUME 16, NUMBER 2 AUGUST 1977

Large-distance and long-time properties of a randomly stirred fluid

Dieter Forster*
Department of Physics, Temple University, Philadelphia, Pennsylvania 19122

David R. Nelson†
Department of Physics, Harvard University, Cambridge, Massachusetts 02138

Michael J. Stephen‡
Physics Department, Rutgers University, New Brunswick, New Jersey 08903
(Received 14 February 1977)

Dynamic renormalization-group methods are used to study the large-distance, long-time behavior of velocity correlations generated by the Navier-Stokes equations for a randomly stirred, incompressible fluid. Different models are defined, corresponding to a variety of Gaussian random forces. One of the models describes a fluid near thermal equilibrium, and gives rise to the usual long-time tail phenomena. Apart from simplifying the derivation of the latter, our methods clearly establish their universality, their connection with Galilean invariance, and their analytic form in two dimensions, $\sim(\log t)^{-1/2}/t$. Nontrivial behavior results when the model is formally continued below two dimensions. Although the physical interpretation of the Navier-Stokes equations below $d = 2$ is unclear, the results apply to a forced Burger's equation in one dimension. A large class of models produces a spectral function $E(k)$ which behaves as k^2 in three dimensions, as expected on the basis of equipartition. However, nonlinear effects (which become significant below *four* dimensions) control the infrared properties of models which force the Navier-Stokes equations at zero wave number.

I. INTRODUCTION

Renormalization group methods have enjoyed success in fields as disparate as high-energy physics,[1] critical phenomena,[2] and solid-state physics.[3] In particular there has been considerable progress in the application of renormalization group theory to study *dynamic* critical phenomena.[4] The Navier-Stokes equation for an incompressible fluid with a random forcing function bears a superficial resemblance to various models used in studies of nonlinear spin dynamics.[4,5] Here we exploit this similarity to give an analysis of the large-distance, long-time behavior of velocity correlations generated by the Navier-Stokes equations with a variety of different forcing functions. Although our analysis does *not* pertain to the properties of a fluid near its critical point,[6] renormalization group methods useful in the study of critical dynamics can be taken over directly.

In Sec. II we will discuss different possibilities for the behavior of the force-force correlations at small wave number. One of the models (model A), with the force regarded as a noise field simulating the effects of the molecular degrees of freedom, describes a fluid near thermal equilibrium. This model generates the familiar long-time tails in the renormalized viscosity, and produces new singularities at small wave numbers as well. Renormalization group theory leads naturally to a unified treatment of these anomalies and provides a scaling description of the breakdown of hydrodynamics which occurs for $d<2$.

Other models are also described in Sec. II. A kind of universality applies: large classes of models exhibit similar infrared, long-time properties. In particular, the more "realistic" models all exhibit a spectral density function $E(k)$ which goes as k^{d-1} at small wave number. This agrees with a result obtained by Saffman[7] for homogeneous isotropic turbulence. We should note, of course, that our considerations refer to the region of effectively small Reynolds' number, and $E(k)\sim k^{d-1}$ is simply a consequence of equipartition and phase-space considerations. It is all the more noteworthy, therefore, that one model (model B), which forces the Navier-Stokes equation even at $k=0$, leads to rather different results at small k. Here nonlinearities dominate the infrared behavior of $E(k)$ below four dimensions, and lead to logarithmic anomalies in $d=4$.

We have supplemented our analysis of the Navier-Stokes equation with a brief discussion of two additional equations appropriate to fluid behavior: A forced version of Burger's equation in one dimension, and a model of the diffusion of a passive scalar. The results for Burger's equation appear to be new, and it would be interesting to test them in a numerical simulation.

In Sec. II we define and discuss the model systems considered in this paper. We discuss the relationship of our analysis to previous work on

long-time tails, as well as the molecular basis for the continuum equations employed here. In Sec. III we present a recursion relation analysis of the large-distance, long-time properties of a variety of fluid systems. Although the calculations will be presented in some detail, reference will be made to the exposition of the dynamic renormalization group approach given for ferromagnetic systems by Ma and Mazenko.[8]

It will be shown in Sec. IV that many of the results of Sec. III can be derived quickly and efficiently by a direct graphical approach. The graphical analysis is simplified considerably by a Ward identity related to the Galilean invariance of the underlying dynamical equations. This same identity shows that certain results obtained in Sec. III are, in fact, valid to all orders in ϵ (here, ϵ is either $2-d$ or $4-d$). The graphical manipulations are conveniently performed using the formalism developed by Martin, Siggia, and Rose.[9] We will not derive this formalism here but note that it can be quickly developed from the equations of motion described in Sec. III using a path integral approach.[10] Although the graphical analysis quickly produces results derived somewhat more laboriously in Sec. III the recursion relation method appears to be more generally applicable, and should be of more utility when Ward identities do not produce such enormous simplifications.

Section V summarizes what has been accomplished.

II. DYNAMICAL EQUATIONS

A. The forced Navier-Stokes equation

Consider the Navier-Stokes equation describing an unbounded incompressible fluid subject to a random forcing function $\vec{f}(\vec{x},t)$, namely,

$$\partial_t\vec{v}+\lambda_0(\vec{v}\cdot\vec{\nabla})\vec{v}=-\lambda_0\vec{\nabla}p/\rho+\nu_0\nabla^2\vec{v}+\vec{f}\,, \tag{2.1}$$

where $\vec{v}=\vec{v}(\vec{x},t)$ is the fluid velocity, $p=p(\vec{x},t)$ is the pressure, ρ is the fluid mass density, ν_0 is the (unrenormalized) viscosity, and λ_0 is a perturbative parameter which will eventually be set to unity. The pressure term in (2.1) is used to enforce the condition of incompressibility,

$$\vec{\nabla}\cdot\vec{v}=0\,. \tag{2.2}$$

Equation (2.1) has often been considered as a model of homogeneous, isotropic turbulence.[11] The use of a statistically defined force sidesteps consideration of the onset of turbulence with increasing Reynolds number, and allows the generation of the statistically steady state. If we take the force to act only at long wavelengths, it is plausible that it sets up an inertial range cascade which, at short distances, is independent of the details of the force.[12]

We are concerned here with the *infrared, long-time* properties of correlations generated by $\vec{f}(\vec{x},t)$. Clearly, the long-wavelength fluid behavior will depend to some extent on the properties of the force. However, as mentioned in Sec. I, some degree of universality applies even at large distances. While we will not attempt to treat the formidable and probably more interesting problem of the ultraviolet (short-distance, short-time) correlations described by (2.1), i.e., of fully developed turbulence, the infrared properties are interesting in their own right. Batchelor and Proudman[13] originally considered the behavior of the large eddies in freely decaying (unforced) turbulence. Specifically, they studied the residual motion far downstream in a wind tunnel experiment, on a scale which is larger than the mesh which initially produced the turbulence. (In this region most of the energy in the fluid has been dissipated, and the Reynolds number is effectively small.) Saffman[7] has considered turbulence generated at an initial instant by a distribution of random impulsive forces, and finds that the spectral energy density behaves as k^{d-1} for small k. Although forced turbulence is a somewhat different problem than that treated by the above authors, we do find that most of the models considered in this section behave as predicted by Saffman. We consider, in addition, the low-frequency properties of the correlations, and corrections to the leading behavior of $E(k)$. Large-distance, long-time properties of freely decaying turbulence can be treated by the methods described in Sec. III, but will not be discussed further in this paper.

We now specify the statistical properties of the force entering Eq. (2.1). The force is taken to be purely solenoidal without loss of generality — any longitudinal component can be absorbed into the definition of the pressure. The problem is simplified further by assuming Gaussian "white noise" statistics for the force. Deviations from a strictly Gaussian force can be considered, but these do not alter the asymptotic infrared behavior. Thus, it is only necessary to specify the two-point force correlations which are of the form

$$\langle f_i(\vec{k},\omega)f_j(\vec{k}',\omega')\rangle = 2D(k)(2\pi)^{d+1}\delta(\vec{k}+\vec{k}')\delta(\omega+\omega')(\delta_{ij}-k_ik_j/k^2)\,, \tag{2.3}$$

where $\vec{f}(\vec{k},\omega)$ is the Fourier transform of $\vec{f}(\vec{x},t)$ in space and time,

$$\vec{f}(\vec{k},\omega)=\int dx\int dt\,\vec{f}(\vec{x},t)e^{i\omega t-i\vec{k}\cdot\vec{x}}\,. \tag{2.4}$$

Three prototype models will be discussed cor-

responding to different forcing functions $D(k)$:

$$\text{model A: } D(k) = D_0 k^2 \quad |\vec{k}| < \Lambda, \quad = 0, \text{ otherwise.} \tag{2.5}$$

$$\text{model B: } D(k) = D_0 \quad |\vec{k}| < \Lambda, \quad = 0, \text{ otherwise.} \tag{2.6}$$

$$\text{model C: } D(k) = D_0 \quad \tilde{\Lambda} < |\vec{k}| < \Lambda, \quad = 0, \text{ otherwise.} \tag{2.7}$$

Model A[14] can be considered simply as a Langevin model for a fluid near equilibrium. This connection will be discussed further in Sec. IIC. In this case the fluctuation-dissipation theorem requires $D_0 = \nu_0 k_B T/\rho$. It can also be thought of as representing some macroscopic stirring force whose spatial integral vanishes. Model B includes a statistically defined force which acts on the fluid even at $k = 0$. Heuristically, it corresponds to a macroscopic "shaking" of the fluid container.[15] While it is perhaps somewhat artificial to imagine exciting a fluid even at $k = 0$, model B does exhibit intriguing behavior below four dimensions. Model C, with $0 < \tilde{\Lambda} < \Lambda$, is perhaps the most realistic. The fluid is excited in a band in k space below $k = \Lambda$, and one is interested in the resulting correlations near $k = 0$. We shall show that the infrared behavior of model C is the same as that of model A, which is a further motivation for considering model A. Of course, we can consider variations, such as an $O(k^4)$ correction to $D(k)$ in model A. However, model A and model B will turn out to be representative of two broad universality classes; most variations turn out to be irrelevant variables in the sense of Wilson.[16]

It should be emphasized that the cutoff Λ occurring in Eqs. (5)–(7) has here the interpretation of an inverse stirring length (except in the case of a fluid near thermal equilibrium, see Sec. IIC). It will be in general much less than the inverse of any dissipation length scale (we will focus primarily on the limit of small viscosity) and of course much less than any molecular cutoff.

B. Burger's equation and the diffusion of a passive scalar

In Sec. III it will be shown that model A develops nontrivial behavior when formally continued below two dimensions. However, an incompressible fluid is not of much interest in, say, one dimension where the correlations vanish identically! Moreover, recent work by Frisch, Lesieur, and Sulem[17] suggests that Navier-Stokes turbulence may not be realizable for *any* dimension less than two.

Fortunately, the same renormalization group methods apply to a d-dimensional generalization of Burger's equation.[18] We consider a velocity field which, at $d = 3$, obeys

$$\partial_t \vec{v} + (\vec{v}\cdot\vec{\nabla})\vec{v} = \nu_0 \nabla^2 \vec{v} + \vec{f}, \tag{2.8}$$

with the restriction

$$\vec{\nabla} \times \vec{v} = 0. \tag{2.9}$$

Using the identity $(\vec{v}\cdot\vec{\nabla})\vec{v} = \frac{1}{2}\vec{\nabla} v^2 - \vec{v} \times (\vec{\nabla} \times \vec{v})$ and deleting the last term, this model is trivially continued to arbitrary d. In one dimension, the ultraviolet behavior is of interest in its own right, and is dominated by shock-wave excitations.[18] We will exhibit new singularities in the one-dimensional large-distance, long-time properties of this model for a Gaussian random stirring force with correlation

$$\langle f_i(\vec{k}, \omega) f_j(\vec{k}', \omega') \rangle = 2D_0 k_i k_j (2\pi)^{d+1} \delta(\vec{k} + \vec{k}') \delta(\omega + \omega'). \tag{2.10}$$

We have also considered a model of the diffusion of a passive scalar T which satisfies an equation of motion

$$\partial_t T + (\vec{v}\cdot\vec{\nabla})T = \kappa_0 \nabla^2 T. \tag{2.11}$$

Here $\vec{v}$ is the fluctuating velocity field appropriate to model A. This problem has also been treated in the long-time tail literature.[19]

C. Long-time tails and the d-dimensional Navier-Stokes equation

The problem of long-time tails in Green-Kubo functions is physically very different from those introduced above; it concerns the dynamical properties of a fluid near thermal equilibrium. More precisely, it concerns the question of whether the usual *linearized* Navier-Stokes equations correctly describe the space and time correlations of spontaneous velocity fluctuations (or, equivalently, the relaxation to equilibrium of sufficiently small externally induced fluctuations), at least for large distances and long times. It is, of course, always formally possible to reorganize the Liouville equation for the local microscopic velocity into the suggestive form[20]

$$\partial_t \vec{v}(\vec{x}, t) - \int_0^t dt' \int d\vec{x}'\, \nu_R(\vec{x} - \vec{x}', t - t') \nabla'^2 \vec{v}(\vec{x}', t') = \vec{f}_R(\vec{x}, t), \tag{2.12}$$

where the "random force" $\vec{f}_R$ vanishes on the average, is uncorrelated with the initial value $v(\vec{x}, 0)$, and where

$$\langle f_{Ri}(\vec{x}, t) f_{Rj}(\vec{x}', t') \rangle = -(k_B T/\rho) \nabla^2 \nu_R(x - x', t - t') \delta_{ij} \tag{2.13}$$

because of equipartition and the fluctuation-dissipation theorem. Here the angular brackets signify the usual thermal equilibrium average, of

course, and we denote here by $\vec{v}$ the transverse part of the unaveraged fluid velocity so that $\vec{\nabla}\cdot\vec{v}=0$. Equation (2.12) is then, of course, equivalent to the customary definition[21] of a renormalized viscosity in terms of the velocity-velocity correlation function. Conventional hydrodynamics results if one assumes that all of the many modes which contribute to $\vec{f}_R$ decay over microscopic distances and times. Then the spatial Fourier transform of ν_R is effectively of the form

$$\nu_R(k,t-t')=2\nu\delta(t-t') \tag{2.14}$$

for sufficiently small k and long times, where ν is the measured viscosity.

The long-time tails are corrections to (2.14) of the form $\sim(t-t')^{-d/2}$ in $d>2$ dimensions. They must result from microscopically slow modes which are still contained in $\vec{f}_R$. In accord with previous work[22] we make the plausible assumption that the only such terms are products of the conserved amplitudes which, for sufficiently small wave vectors, certainly decay over macroscopic times as well. (In order to exclude for simplicity terms which involve coupling of sound and heat diffusion modes we restrict consideration to an incompressible isothermal fluid.) A kinetic equation in which such mode coupling terms are still explicitly exhibited should be local and Markovian. However, it can have those properties only for wave numbers k below a cutoff Λ where Λ^{-1} is large on a microscopic scale but small on a macroscopic one. The former restriction is necessary, of course, for any continuum theory; small wavelength velocity fluctuations are, instead, treated as noise, along with all other molecular degrees of freedom. Note that for sufficiently large Λ^{-1} the relative importance of highly nonlinear terms is strongly reduced from phase-space considerations which our renormalization-group analysis renders more precise. The second restriction must be imposed, of course, to guarantee that *only* small wavelength fluctuations are treated as noise.

There is no difficulty in formally "deriving" the corresponding kinetic equation by way of projector techniques.[23] Using the standard Zwanzig-Mori procedure, one would project the Liouville dynamics onto that part of phase space which is spanned by, in principle, arbitrary powers of the velocity variable, with wave number k small than Λ, and treat the remainder as stochastic noise. In fact the resulting equation is strongly restricted by the requirements of momentum conservation, symmetry, incompressibility, and Galilean invariance. If we omit terms of higher than second order in v, and keep only the lowest-order terms in a gradient expansion, we obtain model A, i.e.,

$$(\partial_t-\nu_0\nabla^2)\vec{v}+(\vec{v}\cdot\vec{\nabla})\vec{v}=\vec{f}_0\,, \tag{2.15}$$

where because of equipartition and the fluctuation-dissipation theorem the two-point force correlations are given by (2.3), with

$$D(k)=\nu_0(k_BT/\rho)k^2\,. \tag{2.16}$$

Note that ν_0 is now a "bare" viscosity which does not as yet contain the contribution made by the mode coupling term. Since we further know, from statistical mechanics, that the equilibrium distribution of $\vec{v}$ is Gaussian, it is consistent to choose the noise force f_0 Gaussian as well.

This is the model which Zwanzig[22] has solved to second order in perturbation theory. For dimension $d\geq 3$, the result,

$$\nu_R(k=0,t)=(k_BT/\rho)\left[(d^2-2)/(d^2+2d)\right](8\pi\nu t)^{-d/2} \tag{2.17}$$

for large t, is fairly convincing since perturbation theory converges term by term. However, as Zwanzig noted, at $d=2$ perturbation theory is questionable since it diverges logarithmically, term by term, at low frequency and wave number. (The expansion parameter is $\bar{\lambda}^2=(k_BT/\rho\nu_0^2)$ whose "naive" dimension is $d-2$.) The renormalization group methods which we discuss below overcome this difficulty. Further, on the basis of Wilson's[16] ideas about irrelevant variables, they afford a convincing and simple demonstration that results obtained from model A will not be invalidated if one extends the model in several respects: to include higher derivative terms than those exhibited in (2.15), terms of higher order in $\vec{v}$, accompanying non-Gaussian contributions to the noise, or even velocity-dependent noise forces. For a near-equilibrium fluid, this is particularly important since model A can only be a crude approximation of the problem which one actually would like to solve. Of course, even renormalization-group methods do not allow one to escape entirely from the assumption, physically plausible as it may be, on which our considerations are based: That there is a coarse-grained level of description at which a kinetic equation, roughly of the structure studied here, describes the dynamics of an incompressible fluid properly.

We conclude this section with a brief note on Galilean invariance since it plays a major simplifying role in this work. Obviously, the transformation

$$\vec{v}(\vec{x},t)\to\vec{v}_0+\vec{v}'(\vec{x}-\vec{v}_0t,t)\,, \tag{2.18}$$

which is an exact symmetry of the microscopic dynamics, must also be a symmetry of semi-phenomenological models. It is important, in our case, because it *prescribes* the vertex of the nonlinearity in Eq. (2.15). [If one introduces a book-

keeping parameter λ_0 as in Eq. (2.1), it multiplies only the $\vec{v}_0$ in the argument of (2.18).] Of course, in formal "derivations" of kinetic models via projector methods, one must *require* that the "noise" term be separately invariant since otherwise an interpretation as a Langevin force would obviously be nonsensical. The standard Zwanzig-Mori technique is in accord with this requirement.

III. RECURSION RELATION ANALYSIS

A. Renormalization-group method

It is useful at this point to review the dynamic renormalization-group procedure.[4] Corresponding to the cutoff Λ in the definitions of the forcing function, the Fourier decomposition of the velocity field will be cut off for $|\vec{k}| > \Lambda$:

$$v_i(\vec{x}, t) = \int_{k<\Lambda} \frac{d^d k}{(2\pi)^d} \int \frac{d\omega}{2\pi} v_i(\vec{k}, \omega) e^{i\vec{k}\cdot\vec{x} - i\omega t} . \quad (3.1)$$

In reality, of course, the high-k velocity modes will eventually be excited[24] by the nonlinearities in the Navier-Stokes equation even for a cutoff force. The supposition here is that the ultraviolet excitations do not affect the infrared modes populating the steady state which develops for $k \ll \Lambda$. Indeed, in line with the usual arguments advanced for the universality of critical phenomena,[2] we expect that the short-distance properties are irrelevant to the large-distance, long-time behavior. Short-distance phenomena will, of course, affect the *amplitudes* of power-law singularities in the correlation functions. These amplitudes are nonuniversal, and cannot be determined by the renormalization techniques described here.

Substituting the decomposition (3.1) into (2.1), and using (2.2) to eliminate the pressure, we obtain the transformed Navier-Stokes equation,

$$v_l(\vec{k}, \omega) = G_0(k, \omega) f_l(\vec{k}, \omega) - \tfrac{1}{2} i \lambda_0 G_0(k, \omega) P_{lmn}(\vec{k}) \times \int_{q\Omega} v_m(\vec{q}, \Omega) v_n(\vec{k} - \vec{q}, \omega - \Omega) . \quad (3.2)$$

Here we have defined an unrenormalized propagator,

$$G_0(k, \omega) \equiv [-i\omega + \nu_0 k^2]^{-1} , \quad (3.3)$$

and the function

$$P_{lmn}(\vec{k}) \equiv P_{lm}(\vec{k}) k_n + P_{ln}(\vec{k}) k_m , \quad (3.4)$$

where $P_{ij}(\vec{k})$ is the transverse projection operator,

$$P_{ij}(\vec{k}) \equiv \delta_{ij} - k_i k_j / k^2 . \quad (3.5)$$

Furthermore, we have adopted a standard convention by defining

$$\int_{q\Omega} \equiv \int_{q<\Lambda} \frac{d^d q}{(2\pi)^d} \int_{-\infty}^{+\infty} \frac{d\Omega}{2\pi} . \quad (3.6)$$

A particular mode $v_l(\vec{k}, \omega)$ is coupled to the remaining degrees of freedom by the nonlinear term proportional to λ_0 on the right-hand side of (3.2). In principle Eq. (3.2) can be iterated in powers of this coupling.

In practice, however, there are difficulties. Consider model A for concreteness. Note that expressing time in units of $1/\nu_0$, velocity in units of $(D_0/\nu_0)^{1/2}$ and the force in units of $1/(D_0\nu_0)^{1/2}$ amounts to setting $D_0 = \nu_0 = 1$ and replacing the vertex λ_0 by

$$\bar{\lambda}_0 = \lambda_0 D_0^{1/2} / \nu_0^{3/2} \quad (3.7)$$

Thus while one naively calculates a perturbation expansion in powers of λ_0 the actual series involves powers of $\bar{\lambda}_0$. This is fortunate since λ_0 was only introduced to organize the expansion, and must eventually be set to unity. However, there are obvious difficulties with a naive perturbation theory for small viscosities ν_0. Note further that $\bar{\lambda}_0$ has the dimension of length to the power $\frac{1}{2}(d-2)$. Thus an expansion of the dimensionless ratio ν_R/ν_0, for example, where ν_R is the renormalized viscosity at zero frequency and wave number, must involve terms of the form

$$\nu_R/\nu_0 = 1 + \text{const}\ \bar{\lambda}_0^2 \int_q (1/q^2) + \cdots , \quad (3.8)$$

which, in two dimensions, diverge logarithmically due to small wave-vector fluctuations.

It should be unproblematic, however, to selectively assess the effect which modes in a shell $\Lambda e^{-l} < |\vec{k}| < \Lambda$ have on the dynamics of the remaining ones if l is not chosen too large. In effect this is what the renormalization group does. One projects the equations of motion onto the phase space spanned by modes with $0 < |\vec{k}| < \Lambda e^{-l}$, and pushes the remainder into the appropriately redefined, partially renormalized noise. This is clearly an iterative procedure from whose asymptotic behavior the properties of macroscopic modes can be extracted.

As explained by Ma and Mazenko,[8] the dynamic renormalization group procedure consists of two steps. First, we eliminate from (3.2) the modes $v_i^>(k, \omega)$ such that $\Lambda e^{-l} < |\vec{k}| < \Lambda$. This is done by formally solving the equations for $v_i^>(k, \omega)$ as a power series in λ_0. The solution, because of the nonlinearities, depends on the remaining modes $v_i^<(k, \omega)$. These formal solutions are then substituted into the equations for $v_i^<(\vec{k}, \omega)$ to eliminate their explicit dependence on $v_i^>(\vec{k}, \omega)$. Finally, the reduced set of equations is averaged over the part of the force $f_i^>(\vec{k}, \omega)$ that acts in the shell $\Lambda e^{-l} < |\vec{k}| < \Lambda$.[25] This redefines the coefficients which enter the reduced equations of motion. The fluctuating remainder is added to the noise force $f_i^<(\vec{k}, \omega)$,

and redefines the coefficients which characterize its spectrum. The parameter l is a measure of the fraction of the degrees of freedom which have been eliminated.

The second step consists of rescaling space, time, and the remaining velocities and forces in order to make the new set of equations look as much as possible like the original Navier-Stokes equation. For example, we would like to keep the coefficient of $\partial_t \vec{v}$ in Eq. (2.1) fixed at unity.

This procedure will in general produce complicated new couplings (higher powers of $\vec{v}$ and $\vec{\nabla}$), in addition to the original Navier-Stokes nonlinearities. However, these couplings turn out to be irrelevant variables, and can often be neglected to a leading approximation (for example, they can be neglected in model A above two dimensions, and also to first order in $\epsilon = 2 - d$ below two dimensions). The direct graphical approach discussed in Sec. IV bypasses the complications introduced by these new couplings.

The output of almost any renormalization group calculation can be expressed in terms of a homogeneity relation. Consider for concreteness the velocity-velocity correlation function,

$$G_{ij}(\vec{k}, \omega) = \frac{\langle v_i(\vec{k}, \omega) v_j(\vec{k}', \omega') \rangle}{(2\pi)^{d+1} \delta(\vec{k}+\vec{k}') \delta(\omega+\omega')} \tag{3.8}$$

for model A. We will bypass the problems which arise in a direct expansion in powers of $\bar{\lambda}_0$ by means of a scaling law (to be derived in Sec. III B), namely

$$G_{ij}(\vec{k}, \omega; \lambda_0) = e^{\alpha(l)} G_{ij}[e^l \vec{k}, e^{\alpha(l)}\omega; \bar{\lambda}(l)]. \tag{3.9}$$

This scaling or homogeneity relation holds for small k and ω. It expresses the fact that G_{ij} can be computed from both the original and the reduced set of equations. Here $\alpha(l)$ is a function characterizing the time rescaling necessary to preserve the form of the Navier-Stokes equation, and $\bar{\lambda}(l)$ is the effective coupling constant after a fraction $1 - e^{-dl}$ of the degrees of freedom have been eliminated.

The left-hand side of (3.9) is difficult to calculate if $\bar{\lambda}_0$ is large and/or, in two dimensions, k and ω are small. In fact we will discover that, for $d \geq 2$, $\bar{\lambda}(l)$ can be made as small as desired by choosing l sufficiently large. Since $\alpha(l)$ turns out to be a linearly increasing function of l, we can, for large l, compute the right-hand side of (3.9) by ordinary perturbation theory. The assertion that $\bar{\lambda}(l) \to 0$ as $l \to \infty$ might be called "infrared asymptotic freedom" in the language of quantum field theory.[1]

B. Model A

We now carry out the program outlined above for model A. The calculations are very similar to a calculation described in Sec. IV B of the paper by Ma and Mazenko.[8] These authors consider a different problem, but one which also involves a quadratic nonlinearity. The recursion relations are conveniently expressed in terms of Feynman graphs. The graphs which occur are the same as those appearing in the standard graphical expansions of the Navier-Stokes equation,[26] but with the interpretation that the internal lines carry momenta in the range $e^{-l}\Lambda < |\vec{k}| < \Lambda$ and the external lines carry momenta such that $|\vec{k}| < e^{-l}\Lambda$. Both internal and external frequencies are unbounded.

Intermediate recursion relations for the propagator, the force-force correlation, and for λ are shown schematically in Fig. 1 to leading order in $\bar{\lambda}$. The term "intermediate" is used because we have not yet rescaled space, time, etc. From Fig. 1, one can simply "read off" the intermediate recursion relations for ν, D, and λ. As an illustration, we evaluate the recursion formula implied by Fig. 1(a) in Appendix A. The reader is referred to Ref. 8 for more details.

The intermediate values of ν, D, and λ are

$$\nu_I = \nu_0[1 + A_d \bar{\lambda}_0^2 (1 - e^{-(d-2)l})/(d-2)], \tag{3.10}$$

$$D_I = D_0[1 + A_d \bar{\lambda}_0^2 (1 - e^{-(d-2)l})/(d-2)], \tag{3.11}$$

$$\lambda_I = \lambda_0. \tag{3.12}$$

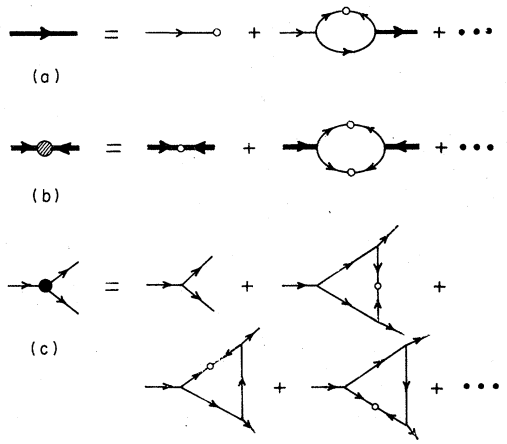


FIG. 1. Schematic representation of recursion formulas describing propagator, correlation function, and vertex renormalization to leading order for the Navier-Stokes equations. Intermediate frequencies are summed from $-\infty$ to $+\infty$, while the intermediate momenta $\vec{q}$ are integrated over the shell $e^{-l}\Lambda < |\vec{q}| < \Lambda$. A light line terminating in an open circle represents the bare propagator $G_0(k, \omega)$ times the random force. Heavy lines represent the full solution of the interacting Navier-Stokes equation, $v_I(\vec{k}, \omega)$. Figure 1(a) describes viscosity renormalization, Fig. 1(b) describes renormalization of the force-force correlations, and Fig. 1(c) gives the renormalization of the coupling λ_0. See Fig. 3 and Appendix A for more details.

Here A_d is a dimensionality-dependent constant, positive for $d \geq 2$ and equal to $1/16\pi$ in two dimensions, and we have, for convenience, set $\Lambda = 1$. As in the model considered in Ref. 8, the "vertex correction" contributions to λ_I shown in Fig. 1(c) vanish identically. These intermediate couplings enter an intermediate Navier-Stokes equation for $v_i^<(\vec{k}, \omega)$, which is of the form (3.2) but with ν_I replacing ν_0, λ_I replacing λ_0, and with $f_i^<(\vec{k}, \omega) \to f_{I,i}^<(\vec{k}, \omega)$ where

$$\langle f_{I,i}^<(\vec{k}, \omega) f_{I,i}^<(\vec{k}', \omega') \rangle = 2D_I k^2 (2\pi)^{d+1} \delta(k + k') \delta(\omega + \omega') P_{ij}(\vec{k}). \quad (3.13)$$

To implement the second step of the dynamic renormalization group procedure we set

$$\vec{v}^<(\vec{k}, \omega) = \zeta(l) \vec{v}'(\vec{k}', \omega'), \quad (3.14)$$

where

$$\vec{k}' = e^l \vec{k} \text{ and } \omega' = e^{\alpha(l)} \omega. \quad (3.15)$$

The rescaling of k compensates for the eliminated degrees of freedom, and $\zeta(l)$ and $\alpha(l)$ are to be determined. The force must be rescaled by

$$\vec{f}_I^<(\vec{k}, \omega) = e^{-\alpha(l)} \zeta(l) \vec{f}'(\vec{k}', \omega'). \quad (3.16)$$

The recursion relations for ν and D are then given by

$$\nu' \equiv \nu(l) = e^{\alpha(l) - 2l} \nu_I(l), \quad (3.17)$$

$$D' \equiv D(l) = e^{\alpha(l) - 2l} [\exp[\alpha(l) + \tfrac{1}{2} dl] / \zeta(l)]^2 D_I(l). \quad (3.18)$$

We note from (3.11) and (3.12) that ν and D are renormalized in the same way. This property, which persists to all orders in perturbation theory, is maintained by the rescaling if we choose

$$\zeta(l) = \exp[\alpha(l) + \tfrac{1}{2} dl]. \quad (3.19)$$

With this choice the recursion formula for λ is

$$\lambda' \equiv \lambda(l) = \exp[\alpha(l) - \tfrac{1}{2}(d+2) l] \lambda_I. \quad (3.20)$$

Because the renormalization group can obviously be iterated it is convenient to replace l by an infinitesimal parameter δ, which leads to *differential* recursion relations. More accurate values of the parameters $\nu(l)$, $D(l)$, and $\lambda(l)$ (which describe the system after a *finite* fraction $1 - e^{-dl}$ of the degrees of freedom have been eliminated) are then obtained by integrating these differential equations. On taking the limit $\delta \to 0$ we find[14]

$$d\nu(l)/dl = \nu(l)[-2 + z(l) + A_d \bar{\lambda}^2(l)], \quad (3.21)$$

$$dD(l)/dl = D(l)[-2 + z(l) + A_d \bar{\lambda}^2(l)], \quad (3.22)$$

$$d\lambda(l)/dl = \lambda(l)[-1 - \tfrac{1}{2} d + z(l)], \quad (3.23)$$

where

$$\alpha(l) = \int_0^l z(l') dl' \quad (3.24)$$

and

$$\bar{\lambda}(l) = \lambda(l)[D(l)/\nu^3(l)]^{1/2} \quad (3.25)$$

in accord with (3.7).

The function $z(l)$ is still arbitrary at this point. However, $z(l)$ drops out of the recursion relation for the reduced coupling $\bar{\lambda}$, namely,

$$d\bar{\lambda}(l)/dl = \tfrac{1}{2} \epsilon \bar{\lambda}(l) - A_d \bar{\lambda}^3(l), \quad (3.26)$$

where we have set

$$\epsilon \equiv 2 - d. \quad (3.27)$$

Above two dimensions ($\epsilon < 0$), the recursion formula (3.26) drives $\bar{\lambda}(l)$ to zero as $l \to \infty$. In exactly two dimensions, $\bar{\lambda}(l)$ still goes to zero, although much more slowly. Below $d = 2$, the fixed point at $\bar{\lambda} = 0$ is unstable to small perturbations, and $\bar{\lambda}(l)$ (for $\bar{\lambda}_0 > 0$) is driven to a stable fixed point,

$$\bar{\lambda}^* = (8\pi\epsilon)^{1/2}, \quad (3.28)$$

to order $\epsilon^{3/2}$. The solution of (3.25), which displays the behavior described above, is

$$\bar{\lambda}(l) = \bar{\lambda}_0 \exp \tfrac{1}{2} \epsilon l [1 + 2A_d \bar{\lambda}_0^2 (e^{\epsilon l} - 1)/\epsilon]^{1/2}. \quad (3.29)$$

The scaling relation to be derived below simplifies considerably if $z(l)$ is chosen to keep $\nu(l)$ and thus $D(l)$ fixed at their initial values. From (3.22) and (3.23) we see that the necessary, l-dependent choice is

$$z(l) = 2 - A_d \bar{\lambda}^2(l). \quad (3.30)$$

Thus, as $l \to \infty$, $z(l)$ approaches a fixed value,

$$\begin{aligned} z &\to 2, \quad d \geq 2, \\ z &\to 2 - \tfrac{1}{2}\epsilon, \quad d < 2. \end{aligned} \quad (3.31)$$

The homogeneity relation (3.9) follows from the fact that for $k < e^{-l}$, $G_{ij}(\vec{k}, \omega)$ can be computed both from the original and the reduced set of equations. Because of the velocity scaling (3.14) we have

$$G_{ij}(\vec{k}, \omega; \bar{\lambda}_0) = \frac{\langle v_i(\vec{k}, \omega) v_j(\vec{k}', \omega') \rangle}{(2\pi)^{d+1} \delta(\vec{k} + \vec{k}') \delta(\omega + \omega')} = \zeta^2(l) e^{-dl + \alpha(l)} \frac{\langle v_i^<(e^l \vec{k}, e^{\alpha(l)} \omega) v_j^<(e^l \vec{k}', e^{\alpha(l)} \omega') \rangle}{(2\pi)^{d+1} \delta(e^l \vec{k} + e^l \vec{k}') \delta(e^{\alpha(l)} \omega + e^{\alpha(l)} \omega')}$$

$$= \exp\left(\int_0^l z(l')\, dl' \right) G_{ij}\left[e^l \vec{k}, \exp\left(\int_0^l z(l')\, dl' \right) \omega; \bar{\lambda}(l) \right]. \quad (3.32)$$

Above two dimensions, $\int_0^l z(l')\,dl'$ approaches 2 for large l, and $\bar{\lambda}(l)$ tends to zero. In this case Eq. (3.32) implies that, for asymptotically small k, $G_{ij}(\vec{k},\omega)$ can be expressed in term of a scaling function,[2]

$$G_{ij}(\vec{k},\omega)=P_{ij}(\vec{k})k^{-2}\Phi(\omega/k^2). \tag{3.33}$$

This is just the prediction of conventional linearized hydrodynamics, provided we take

$$\Phi(x)=\chi\nu/[x^2+\nu^2], \tag{3.34}$$

where $\chi\equiv D_0/\nu_0=k_BT/\rho$, and ν is the measured viscosity.

Indeed, the renormalization (in the conventional sense) of this transport coefficient is easily extracted from (3.32) as well since for sufficiently large l [such that $\bar{\lambda}(l)\simeq 0$] we have more accurately

$$\int_0^l z(l')\,dl'-2l\equiv-\Delta=-A_d\int_0^\infty \bar{\lambda}(l')\,dl', \tag{3.35}$$

except for exponentially small terms. Since we know the correlation function for $\lambda\simeq 0$, we obtain (3.34), with

$$\nu/\nu_0=e^{\Delta}=[1+2A_d\bar{\lambda}_0^2\Lambda^{|\epsilon|}/|\epsilon|]^{1/2}, \tag{3.36}$$

where we have restored the wave-number cutoff Λ. In an equilibrium fluid above two dimensions, this formula gives the contribution of modes with wave number below Λ to the measured viscosity.

In order to discuss corrections to conventional hydrodynamics it is convenient to consider the renormalized viscosity $\nu_R(k,\omega)$, defined implicitly by representing $G_{ij}(\vec{k},\omega)$ in the form[21]

$$G_{ij}(\vec{k},\omega)=2P_{ij}(\vec{k})\,\mathrm{Re}\{\chi/[-i\omega+k^2\nu_R(k,\omega)]\}. \tag{3.37}$$

This definition agrees with that implied by Eq. (2.12). A scaling relation for $\nu_R(k,\omega)$ follows, namely

$$\nu_R(k,\omega;\bar{\lambda}_0)=e^{2l-\alpha(l)}\nu_R[e^lk,e^{\alpha(l)}\omega;\bar{\lambda}(l)]. \tag{3.38}$$

As an example consider the case of $k=0$. We evaluate the right-hand side of (3.38) at $l=l^*$ such that

$$e^{\alpha(l^*)}\omega=1. \tag{3.39}$$

By choosing ω small enough we can make the effective coupling, $\bar{\lambda}(l^*)\sim\bar{\lambda}_0\omega^{(1/4)|\epsilon|}$, as small as desired so that $\nu_R[0,1;\bar{\lambda}(l^*)]$ can be expanded. Noting that perturbation theory involves only even powers of $\bar{\lambda}$ it is clear that the lowest-order correction is of order $\omega^{(1/2)|\epsilon|}$. This argument, which is just a simple application of Wegner's theory[27] of the corrections to scaling, has thus produced the celebrated long-time tail correction to the renormalized viscosity. Explicitly in $d=3$ we find

$$\nu_R(0,\omega)=\nu+[7i\chi/120\pi\nu](i\omega/2\nu)^{1/2}+\mathcal{O}(\omega) \tag{3.40}$$

in agreement with the long-time tail literature.[19]

In exactly *two dimensions* the slow approach of $\bar{\lambda}(l)$ to zero gives logarithmic corrections to conventional hydrodynamics which, in contrast to the results for $d>2$, are not conveniently handled by ordinary perturbation theory but which the renormalization group method yields easily. Inserting the expressions (3.29) and (3.30) to compute $\alpha(l)$ near $d=2$ one obtains

$$\nu_R(k,\omega;\bar{\lambda}_0)=[1+\bar{\lambda}_0^2(e^{\epsilon l}-1)/8\pi\epsilon]^{1/2}\,\nu_R[e^lk,e^{z^*l}\omega;\bar{\lambda}(l)], \tag{3.41}$$

where z^* is the value given in (3.31). To find $\nu_R(k,\omega=0)$ we evaluate the right-hand side by choosing $l=l^*$ such that $e^{l^*}k=1$. We can take k small enough to make $\bar{\lambda}(l^*)$ as small as desired. Nevertheless, the wave-number argument on the right-hand side of (3.41), ke^{l^*}, remains firmly pinned at 1, and since even in two dimensions the perturbation expansion for ν_R is unproblematic for *finite* k (or ω), we can simply replace $\nu_R[1,0;\bar{\lambda}(l^*)]$ by its unrenormalized value, ν_0. The result,[14]

$$\nu_R(k,0)=\nu_0[1+\bar{\lambda}_0^2(k^{-\epsilon}-1)/8\pi\epsilon]^{1/2}, \tag{3.42}$$

is valid for $|\epsilon|\ll 1$, and gives a logarithmic correction to the hydrodynamic result in $d=2$, namely

$$\nu_R(k,0)=[(\chi/8\pi)\ln(1/k)]^{1/2} \tag{3.43}$$

for asymptotically small k. Similarly, (3.41) gives[22,14]

$$\nu_R(0,\omega)=[(\chi/16\pi)\ln(1/\omega)]^{1/2} \tag{3.44}$$

for asymptotically small frequency in two dimensions, and since $z^*=2-\frac{1}{2}\epsilon$ for $d<2$, one obtains $\nu_R\sim(1/\omega)^{\epsilon/(4-\epsilon)}$ for $d<2$.[19,14] [Equation (3.44) differs by a factor of $2^{1/2}$ from the result conventionally extracted from the mode coupling formula, see p. 117 in Pomeau and Résibois, Ref. 19. We believe that (3.44) is correct. The difference can be understood by arguing that the mode coupling formula computes not $\nu(\omega)$, but $\delta\nu(\omega)$.]

A quantity of considerable interest in turbulence theory is the spectral energy density,[12]

$$E(k)=[\tfrac{1}{2}S_d/(2\pi)^{d+1}]k^{d-1}\int_{-\infty}^{+\infty}\mathrm{Tr}G_{ij}(\vec{k},\omega)\,d\omega, \tag{3.45}$$

where S_d is the surface area of a d-dimensional sphere, $S_d=2\pi^{(1/2)d}/\Gamma(\frac{1}{2}d)$. Because the prefactor $e^{\alpha(l)}$ in (3.9) is identical to the frequency rescaling. $\int \mathrm{Tr}G_{ij}(\vec{k},\omega)\,d\omega$ is a constant for small k in any dimension, and $E(k)\sim k^{d-1}$ in accord with the results of Saffman[7] for a somewhat different problem. In fact, as shown by Edwards and McComb,[28] all equal time correlations generated by a force whose autocorrelation goes as D_0k^2 are Gaussian distributed. Anomalies (controlled by the exponent z) only ap-

pear when the correlations are considered at unequal times.

C. A generalized Burger's equation

As explained in Sec. II, there are difficulties in the physical interpretation of results obtained for model A below two dimensions. For this reason, we consider the d-dimensional Burger's equation described in Sec. II B. Although it is interesting to survey its properties as a function of dimensionality, the model is rather unrealistic outside of one dimension since it does not conserve "energy" [i.e., the integral $\int d^dx\, v^2(\vec{x}, t)$] in the inviscid limit. We shall see that there are corresponding anomalies in the recursion relation analysis near $d=2$. In this subsection we merely display these peculiarities, and demonstrate that the small viscosity infrared properties are not readily susceptible to analysis near two dimensions. Results derived in Sec. IV will, in fact, allow explicit predictions to be made about this model in one dimension.

Upon Fourier transformation, Eq. (2.8) takes the form

$$v_i(\vec{k}, \omega) = G_0(k, \omega) f_i(\vec{k}, \omega) - \tfrac{1}{2} i\lambda_0 G_0(k, \omega) k_i \int_{q\Omega} v_j(\vec{q}, \Omega) v_j(\vec{k} - \vec{q}, \omega - \Omega)\,, \tag{3.46}$$

where $G_0(k, \omega)$ is again given by (3.3), and the parameter λ_0 has been inserted multiplying the nonlinearity. The recursion analysis proceed as sketched for model A; the graphs shown in Fig. 1 again appear, but with interpretations dictated by (3.46) and (2.10). Because of Galilean invariance the "vertex correction" graphs[29] shown in Fig. 1(c) again vanish when the external momenta and frequencies are taken to zero. However, the model develops nontrivial static properties, and therefore ν_0 and D_0 no longer renormalize in the same way. We therefore rescale the velocities and forces as in (3.14) and (3.16) but extend (3.19) by means of the parametrization

$$\zeta(l) = \exp\left(\int_0^l [z(l') + y(l') + \tfrac{1}{2}d]\, dl'\right). \tag{3.47}$$

In d dimensions the recursion relations for $\nu(l)$, $D(l)$, and $\lambda(l)$ are, to leading order in $\bar{\lambda} = \lambda D^{1/2}/\nu^{3/2}$,

$$d\nu(l)/dl = \nu(l)[-2 + z(l) + K_d \bar{\lambda}^2(l)(2-d)/4d], \tag{3.48}$$

$$dD(l)/dl = D(l)[-2 + z(l) - 2y(l) + K_d \bar{\lambda}^2(l)/4], \tag{3.49}$$

$$d\lambda(l)/dl = \lambda(l)[-1 - \tfrac{1}{2}d + z(l) + y(l)\,], \tag{3.50}$$

where $K_d = 1/[2^{d-1}\pi^{(1/2)d}\Gamma(\frac{1}{2}d)]$.

As in the discussion of model A, $y(l)$ and $z(l)$ can be chosen to be functions of $\bar{\lambda}(l)$ such that $\nu(l)$ and $D(l)$ remain at their initial values. However, difficulties arise when we consider the recursion formula for $\bar{\lambda}(l)$,

$$d\bar{\lambda}(l)/dl = \tfrac{1}{2}(2-d)\bar{\lambda}(l) + K_d[(2d-3)/4d]\bar{\lambda}^3(l)\,. \tag{3.51}$$

This equation differs qualitatively from the corresponding equation for model A [Eq. (3.26)] near $d=2$ because of the sign of the cubic term. A nontrivial fixed point exists *above* two dimensions, but it is unstable to small perturbations. If $\bar{\lambda}$ is initially larger than this critical coupling, $\bar{\lambda}(l)$ becomes intractably large for large l, and homogeneity relations such as (3.9) are of little use.

It is only below 1.5 dimensions, where the cubic term changes sign, that (3.51) has any similarity to (3.26). Of course, higher-order terms in $\bar{\lambda}(l)$ may be of importance in these dimensionalities. Note that the recursion relations for $\nu(l)$ and $D(l)$ can be made identical in one dimension, where energy becomes a conserved quantity. In Sec. IV we will show that ν and D are related by a fluctuation-dissipation theorem in one dimension. This relation, together with a Ward identity related to the Galilean invariance of the theory, will lead to a nontrivial prediction in one dimension. Specifically, we will demonstrate that the correlation function,

$$G(k, \omega) = \langle v(k, \omega) v(k', \omega')\rangle / (2\pi)^2 \delta(k + k')\delta(\omega + \omega') \tag{3.52}$$

scales with an exponent $z = \frac{3}{2}$, i.e.,

$$G(k, \omega) = k^{-3/2}\Phi(\omega/k^{3/2})\,. \tag{3.53}$$

D. Convection of a passive scalar

It is instructive to study the convection of a passive scalar by the velocity field appropriate to model A. We must consider the auxiliary equation of motion (2.11), where the variable T is intended to represent temperature, or the concentration of a labelled set of particles.[11,19]

Introducing the Fourier-Laplace transform of the scalar field $T(\vec{x}, t)$,

$$\hat{T}(\vec{k}, \omega) = \int_0^\infty dt\, e^{i\omega t} \tilde{T}(\vec{k}, t) = \int_0^\infty dt \int d^dx\, e^{i\omega t - i\vec{k}\cdot\vec{x}} T(\vec{x}, t)\,, \tag{3.54}$$

Eq. (2.11) takes the form

$$\hat{T}(\vec{k}, \omega) = \mathcal{G}_0(k, \omega)\tilde{T}(\vec{k}, t=0) - i\lambda_2 \mathcal{G}_0(k, \omega) k_i \int_{q\Omega} v_i(\vec{k} - \vec{q}, \omega - \Omega)\hat{T}(\vec{q}, \Omega)\,. \tag{3.55}$$

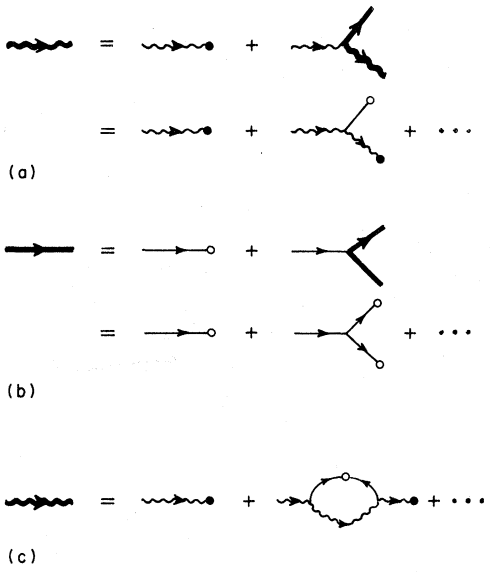


FIG. 2. Schematic representation of the coupled equations of motion describing the diffusion of a passive scalar. In Fig. 2(a), the light wavy line is the bare diffusion propagator, while the solid circle represents the initial conditions $\tilde{T}(\vec{k}, t=0)$. Heavy wavy lines indicate the solution of the diffusion equation before averaging over the random force. Figure 2(c) summarizes the recursion formula for the diffusion propagator obtained by averaging over components of the random force $f_i(\vec{k}, \omega)$ in the shell $e^{-l}\Lambda < |\vec{k}| < \Lambda$.

A parameter λ_2 has been inserted in the nonlinear term, and we have introduced the bare diffusion propagator

$$g_0(k, \omega) = [-i\omega + \kappa_0 k^2]^{-1} \ . \tag{3.56}$$

Equation (3.55), which is to be solved in conjunction with (3.2), is depicted schematically in Fig. 2(a) together with its iterated solution. Figure 2(b) shows a schematic representation of (3.2).

The relaxation from the initial condition $\tilde{T}(\vec{k}, t=0)$ results upon averaging over the fluctuating velocity field. As discussed earlier, this can be done gradually, by integrating out modes in successive shells of $\vec{k}$ space. The graphs which represent the recursion relation for the diffusion propagator are shown in Fig. 2(c). As usual, nonlinear corrections to the vertex vanish. Note that if $\lambda_2 \neq \lambda_0$, the theory is not formally Galilean invariant; even then vertex corrections vanish as a result of the transversality of the velocity field. However, since initially $\lambda_2 = \lambda_0 = 1$, Galilean invariance is assured at every recursive step if we choose the same rescaling for $\hat{T}$ as we did for the velocity field,

$$\hat{T}(\vec{k}, \omega) = \zeta(l)\hat{T}'\left\{e^{l}\vec{k}, \exp\left[\int_0^l z(l')\, dl'\right]\omega\right\}, \tag{3.57}$$

where as before

$$\zeta(l) = \exp\left[\tfrac{1}{2}dl + \int_0^l z(l')\, dl'\right]. \tag{3.58}$$

In this case $\lambda_2(l)$ and $\lambda(l)$ are identical for all l. The differential equation for $\kappa(l)$ is then

$$d\kappa(l)/dl = [-2 + z(l)]\kappa(l) + \frac{[(d-1)/d]K_d\bar{\lambda}^2(l)\nu^2(l)}{[\nu(l) + \kappa(l)]}, \tag{3.59}$$

where K_d was defined in (3.48). This equation is to be solved in conjunction with the equations (3.21)–(3.25) for model A.

After completely averaging over the fluctuating velocity field, the "temperature" relaxation is customarily described by

$$\langle \hat{T}(\vec{k}, \omega)\rangle = [-i\omega + k^2\kappa_R(k, \omega)]^{-1}\tilde{T}(\vec{k}, t=0). \tag{3.60}$$

It is then easy to show that the renormalized diffusion coefficient κ_R scales according to

$$\kappa_R(k, \omega; \bar{\kappa}_0, \bar{\lambda}_0) = e^{2l-\alpha(l)}\kappa_R[e^{l}k, e^{\alpha(l)}\omega; \bar{\kappa}(l), \bar{\lambda}(l)], \tag{3.61}$$

where the dimensionless ratio $\bar{\kappa}(l) = \kappa(l)/\nu(l)$ obeys the recursion relation

$$d\bar{\kappa}(l)/dl = -A_d\bar{\lambda}^2(l)(\bar{\kappa}(l) - 2(d-1)(d+2) \times \{(d^2-2)[\bar{\kappa}(l)+1]\}^{-1}). \tag{3.62}$$

The analysis is completed in the usual way. We can integrate the renormalization equations until $l = l^*$ such that, for example,

$$e^{2\alpha(l^*)}\omega^2 + \nu_0^2 e^{4l^*}k^4 = 1. \tag{3.64}$$

This ensures that diagrammatic corrections to (3.55) are cut off by either a finite frequency or a finite momentum so that no infrared divergences can complicate the analysis at $l = l^*$, even in two dimensions. By choosing k and ω sufficiently small, $\bar{\lambda}(l^*)$ can again be made as small as desired. Recalling that $z(l)$ was chosen to keep $\nu(l)$ fixed at ν_0, we find that $\kappa(l)$ is given by the equation

$$\frac{x^2(l)-1}{x_0^2-1}\left[\frac{[x(l)-1]/(x_0-1)}{[x(l)+1]/(x_0+1)}\right]^{\xi} = \left[1 + 2A_d\bar{\lambda}_0^2\frac{e^{\epsilon l}-1}{\epsilon}\right]^{-1}, \tag{3.65}$$

where

$$x(l) \equiv \xi[2\kappa(l)+1] \tag{3.66}$$

and

$$\xi^{-2} \equiv 1 + 8(d-1)(d+2)/(d^2-2). \tag{3.67}$$

This equation, together with (3.61) and (3.64), determines $\kappa_R(k, \omega)$ in implicit form. On setting $k=0$, it is straightforward to exhibit a long-time tail in

$\kappa_R(0,\omega)$[19] from the more general results derived here. The result in three dimensions is

$$\kappa_R(0,\omega)=\kappa+[i\chi/6\pi(\nu+\kappa)^{3/2}](i\omega)^{1/2}, \tag{3.68}$$

where κ is the measured diffusion coefficient which is given by $\kappa=\nu\kappa(l\to\infty)$. Similarly, the divergence of the super-Burnett coefficient in three dimensions is easily understood by a "corrections to scaling" analysis of $\kappa_R(k,\omega=0)$. In two dimensions the result, for asymptotically small k and ω, can be written in the intriguing form

$$\kappa_R(k,\omega)/\nu_R(k,\omega)=\bar{\kappa}(\infty)=\tfrac{1}{2}(1+\sqrt{17}). \tag{3.69}$$

This ratio remains universal if the theory is formally continued below $d=2$, except that $\bar{\kappa}(\infty)$ is replaced by the ϵ-dependent solution of $x(\infty)=1$.

We note, finally, that these results are unchanged if the scalar field $T(\vec{x},t)$ is itself driven by an independent Gaussian source of thermal noise, in addition to being convected by the velocity field. This assumption is a natural one for the case when $T(\vec{x},t)$ is to represent the density of a dilute solute.

E. Model B

The treatment of model B follows closely that sketched for model A, but with rather different results. Infrared singularities due to the nonlinearities modify the large distance properties of this model below four dimensions.

Construction of recursion relations for the equation of motion (3.2) [with a forcing function governed by (2.6)] proceeds as before. Nonlinear contributions to the renormalization of λ_0 again vanish, but D_0 and ν_0 are renormalized very differently. In particular the graph shown in Fig. 1(b) does not contribute to the constant part of the force autocorrelation, but generates instead a term proportional to k^2. This term is irrelevant when the force is rescaled to keep $D(l)$ fixed at D_0.

When velocities and forces are rescaled as indicated by Eqs. (3.14) and (3.16), we find [from (3.18) with the factor e^{-2l} removed] that the choice

$$\zeta(l)=\exp\left\{\int_0^l \tfrac{1}{2}[3z(l')+d]\,dl'\right. \tag{3.70}$$

is necessary to keep $D(l)$ fixed at D_0 and to preserve the form of the Navier-Stokes equations. The resulting recursion formulas for $\nu(l)$, $D(l)$, and $\lambda(l)$, accurate to $\Theta(\bar{\lambda}^2)$, are

$$d\nu(l)/dl=[-2+z(l)]\nu(l)+B_d\lambda^2(l)D(l)/\nu^2(l), \tag{3.71}$$

$$dD(l)/dl=0, \tag{3.72}$$

$$d\lambda(l)/dl=[-1-\tfrac{1}{2}d+\tfrac{3}{2}z(l)]\lambda(l), \tag{3.73}$$

where $B_d=K_d(d-2)/2d$ with K_d as defined in (3.48). The equation for the reduced coupling $\bar{\lambda}\equiv\lambda D^{1/2}/\nu^{3/2}$ is then

$$d\bar{\lambda}(l)/dl=\tfrac{1}{2}(4-d)\bar{\lambda}(l)-(\tfrac{3}{2})B_d\bar{\lambda}^3(l). \tag{3.74}$$

The choice of $z(l)$ which fixed $\nu(l)$ at ν_0 is

$$z(l)=2-B_d\bar{\lambda}^2(l). \tag{3.75}$$

Above four dimensions (3.64) exhibits a stable "hydrodynamic" fixed point at $\bar{\lambda}^*=0$, which leads to the conventional result $z\to 2$. Below $d=4$ however, a stable nontrivial fixed-point controls the infrared properties. Thus, the asymptotic behavior of the frequency rescaling exponent $z(l)$ is

$$\begin{aligned} &z(l)\to 2, \quad d\geq 4 \\ &z(l)\to 2-(4-d)/3, \quad d<4. \end{aligned} \tag{3.76}$$

The homogeneity law analogous to (3.32) for model B is

$$G_{ij}(\vec{k},\omega;\bar{\lambda}_0)=\exp\left[2\int_0^l z(l')\,dl'\right] \times G_{ij}\left\{e^l\vec{k},\ \exp\left[\int_0^l z(l')\,dl'\right]\omega;\bar{\lambda}(l)\right\}. \tag{3.77}$$

We will not pause to calculate the renormalized viscosity for this problem, although the methods of Secs. III B and III D can be extended straightforwardly. We note, however, that there are "long-time tails" in $\nu_R(k=0,\omega)$ above four dimensions, and logarithmic corrections in $d=4$. For example,

$$\nu_R(k,\omega=0)\sim\ln^{1/3}(1/k) \tag{3.78}$$

in four dimensions. Hydrodynamics breaks down for this model below four dimensions.

Because the prefactor in Eq. (3.77) differs from the frequency rescaling, $E(k)$ does *not* vary as k^{d-1} for this model. Indeed, using the results derived above, we readily deduce that

$$\begin{aligned} &E(k)\sim k^{d-3}, && d>4, \\ &E(k)\sim k/\ln^{1/3}(1/k), && d=4, \\ &E(k)\sim k^{1-(2/3)(4-d)}, && d<4. \end{aligned} \tag{3.79}$$

This anomalous behavior is a consequence of forcing the Navier-Stokes equation at macroscopically large wavelengths, i.e., of a nonvanishing function $D(k)$ at $k=0$.

F. Model C and universality

Model A and model B are representative of two broad universality classes; their infrared long-time properties characterize a large number of similar equations of motion. In particular, k^2 corrections to $D(k)$ for model B and k^4 corrections to $D(k)$ for model A[30] are irrelevant variables, and do not affect the asymptotic properties derived in this

section. The term irrelevant is used here in the same sense as in critical phenomena: the coefficients which parametrize such correction terms vanish at the fixed point, and their asymptotic approach is faster than that of the "leading irrelevant variable" $\bar{\lambda}$. For example, a term of the form $\mu\nabla^2 v^3(\vec{x}, t)$ in the equations of motion (because of momentum conservation and symmetry, there must be at least two gradient factors in this term) would lead to a recursion relation of the form

$$d\mu(l)/dl = [z(l) - 2 - d]\,\mu(l) + \text{diagrams}, \qquad (3.80)$$

and thus approach the fixed point as e^{-dl}, for $d \geq 2$. The exponent, $z - 2 - d$, simply characterizes the space and time dimensions of μ and is thus easily found. If, therefore, such a term were inserted in homogeneity relations like (3.38), it would only contribute a correction term of, at least, order $\omega^{(1/2)d}$ which is asymptotically negligible. In this fashion it is straightforward to demonstrate that terms of higher order in $\vec{\nabla}$, ∂_t, and $\vec{v}(\vec{x}, t)$ in the equations of motion are irrelevant, as are deviations from the Gaussian character of the noise or velocity-dependent noise forces. Of course, this analysis can only be carried out in a small neighborhood of the fixed point considered here. We have nothing to say about the mathematical possibility that a different fixed point might in fact be approached for which λ, μ etc. are finite.

Even barring this bizarre possibility one should note that the irrelevant variables, which for the equilibrium fluid are undoubtedly present in the initial model, renormalize the remaining parameters. That is, if the "time" $l = l_0$ is such that for $l > l_0$ all irrelevant parameters can be disregarded, the parameters ν_0, D_0, and λ_0 will have changed to $\nu(l_0)$, $D(l_0)$ and $\lambda(l_0)$. It is at this point, practically speaking, that our present explicit analysis starts.

Model C is more realistic than models A and B because the fluid is stirred only in a narrow band of wave numbers. Although somewhat idealized, it nevertheless represents a good example of the kind of "universality" discussed in the above paragraph. The large-distance and long-time properties of model C are just those of model A.

To demonstrate this, we use the renormalization group developed for model B to "integrate out" the pulse of force in $\vec{k}$ space. Repeated elimination of degrees of freedom occupying shells in $\vec{k}$ space gradually removes the constant part of the force autocorrelation. No new contribution to the constant part of the force-force correlations are generated at small k because, as discussed in Sec. III E, graphs like that in Fig. 1(b) contribute only k^2 corrections to the renormalized force autocorrelations. Recursion relations such as (3.71)–(3.73) need only be integrated until $l = l_0 = \ln(\Lambda/\tilde{\Lambda})$, when the lower edge of the pulse is reached. At this point, the system is described by renormalized equations of motion in which "partially dressed" couplings $\nu(l_0)$ and $\lambda(l_0)$ appear.

The crucial feature is that the force-force correlations no longer contain a constant part, and behave instead as k^2 for small k. Although this k^2 term was irrelevant in the analysis of model B, it now dominates the infrared behavior. No nonanalytic terms such as $|\vec{k}|$ or $k^{3/2}$ can appear at this point because the (analytic) recursion relations have only been integrated a finite amount of "time" $l_0 = \ln(\Lambda/\tilde{\Lambda})$. Model C now resembles model A, and can by analyzed by the methods developed in Sec. III B. In particular, the infrared properties of the two models should be identical.

These conclusions should not depend on the rather special rectangular shape of $D(k)$. We expect that any force whose autocorrelation is cut off both above and below will generate dynamics falling into the universality class exemplified by model A.

G. Results to all orders in ϵ

The reader may have noted that the recursion relations such as (3.23) and (3.73) derived for "convective" coupling constants in this Section are all extremely simple. No nonlinear terms $[\mathcal{O}(\lambda^3)]$ appear on the right-hand side of the recursion formulas for $\lambda(l)$. This feature is the reason, for example, for the simple canonical exponents $(t^{-d/2})$ which characterize long-time tail phenomena for model A above 2 dimensions, as well as the simple exponents $\frac{1}{2}$ of logarithmic corrections at $d = 2$. For model A this feature is quite general, a consequence of the Galilean invariance (2.18) of the underlying equations of motion which is also a symmetry of the shell integration. A graphical proof will be given in Sec. IV.

Accepting this multiplicative renormalization of $\lambda(l)$, results such as

$$z = 2 - \tfrac{1}{2}\epsilon \quad (\epsilon \equiv 2 - d > 0) \qquad (3.81)$$

for model A and

$$z = 2 - \tfrac{1}{3}\epsilon \quad (\epsilon \equiv 4 - d > 0) \qquad (3.82)$$

for model B appear to be correct to all orders in ϵ. The derivation of such results by means of recursion relations was sketched for model A in Ref. 14, and is given in detail for a problem in dynamic critical phenomena by Halperin, Hohenberg, and Siggia.[31] We will not pause to repeat such a demonstration here, but instead refer the reader to the graphical treatment in Sec. IV. As stated in that Section, we have been unsuccessful in producing a graphical proof that the result (3.82) for model B is correct to all orders.

IV. DIRECT GRAPHICAL ANALYSIS

At this point we demonstrate that results obtained in the previous Section can be quickly and efficiently derived using a direct graphical approach. Although the techniques employed here are perhaps more familiar than the recursion relation formalism of Sec. III, they do in fact rely on renormalization group ideas, such as Wilson's Feynman graph approach,[2] and the parquet graph resummation method.[32] They allow a convincing demonstration that results obtained for model A are in fact valid to all orders in $\epsilon \equiv 2-d$, and permit predictions to be made about the infrared properties of Burger's equation in one dimension. Part of the utility of this graphical approach rests on the Ward identity proved in Appendix B. As mentioned in the Introduction, we expect the recursion analysis to be of more utility in situations where Ward identities do not produce such enormous simplifications.

A. Model A

We first discuss the small k, small ω properties of the response function $G_R(k,\omega) \equiv \langle \delta v_i / \delta f_i \rangle (\vec{k},\omega)$ and the correlation function

$$G(k,\omega) = \mathrm{Tr} G_{ij}(\vec{k},\omega) \tag{4.1}$$

appropriate to model A. An analogy with critical phenomena suggests that these functions can be written in a scaled form,

$$\begin{aligned} G^{-1}(k,\omega) &= k^{2-\eta} g(\omega/k^z), \\ G_R^{-1}(k,\omega) &= k^{2-\eta} g_R(\omega/k^z), \end{aligned} \tag{4.2}$$

where the exponents η and z are to be determined. The fluctuation-dissipation theorem discussed in Appendix B requires that G^{-1} and G_R^{-1} scale with identical exponents η and z. In a linear treatment of the fluctuations, these two functions are simply given by

$$\begin{aligned} G_0(k,\omega) &= 2D_0 k^2/[\omega^2 + \nu_0^2 k^4], \\ G_{0R}(k,\omega) &= [-i\omega + \nu_0 k^2]^{-1}, \end{aligned} \tag{4.3}$$

which satisfy (4.2) with $\eta = 0$ and $z = 2$.

For model A at least, these exponents can be determined more generally by the following simple arguments:

(i) The *equal time* velocity fluctuations are given by

$$\frac{\langle v_i(\vec{k},t) v_j(\vec{k}',t)\rangle}{(2\pi)^d \delta(\vec{k}+\vec{k}')} = \int_{-\infty}^{+\infty} G(k,\omega)\, d\omega/2\pi = D_0/\nu_0, \tag{4.4}$$

where this last result follows from the Gaussian distribution of these fluctuations [Eq. (B1)]. This leads immediately to an exponent relation,

$$z = 2 - \eta. \tag{4.5}$$

(ii) A second relation between η and z follows from simple power counting arguments. In Appendix B it is shown that vertex corrections vanish for small k and ω. The requirement that perturbation theory be consistent with (4.2) and this result leads to an additional relation,

$$3\eta + z = \max(2, 4-d). \tag{4.6}$$

We thus have two relations for the exponents η and z. On substituting the values $\eta = 0, z = 2$ appropriate to linearized hydrodynamics into (4.6) we see that $d = 2$ is a special case. The following results, which agree with those of Sec. III, are then obtained:

(1) $d < 2$.

$$\eta = \tfrac{1}{2}(2-d), \quad z = 2 - \tfrac{1}{2}(2-d), \tag{4.7}$$

and the renormalized viscosity is singular,

$$\begin{aligned} \nu_R(k, \omega=0) &\sim k^{-(1/2)(2-d)}, \\ \nu_R(k=0,\omega) &\sim (1/\omega)^{(2-d)/(2+d)}. \end{aligned} \tag{4.8}$$

The correlation and response functions can be determined as expansions in powers of $\epsilon = 2-d$.[14]

(2) $d > 2$. To leading order the exponents and correlation and response functions in Eq. (4.2) are given by linearized hydrodynamics. There are, however, nontrivial corrections of the form discussed in Sec. III. These corrections may be determined by calculating the self-energy contributions to $G^{-1}(k,\omega)$ and $G_R^{-1}(k,\omega)$. In three dimensions, the renormalized viscosity appearing in (3.37) is then given by

$$\nu_R(0,\omega) = \nu + [7i\chi/120\pi\nu](i\omega/2\nu)^{1/2} + \Theta(\omega), \tag{4.9}$$

$$\nu_R(k,0) = \nu - (1+\tfrac{3}{4}\pi)(\chi/96\pi\nu)k + \Theta(k^2), \tag{4.10}$$

where as in Sec. III various constant terms have been absorbed into the measured viscosity ν, and $\chi = D_0/\nu_0 = k_B T/\rho$.

(3) $d = 2$. This case requires special consideration since logarithmic corrections arise. However, we can apply the parquet method used in critical phenomena[32] by writing a self-consistent equation for the renormalized viscosity. This equation can be determined as an expansion in the parameter $1/\nu_R$ which is logarithmically small:

$$\frac{\partial}{\partial k}\nu_R(k,0) = -\frac{\chi}{16\pi}\frac{1}{k\nu_R(k,0)} + \Theta(1/\nu_R^3). \tag{4.11}$$

Imposing the boundary condition $\nu_R(k=\Lambda,0) = \nu_0$, this equation integrates to

$$\nu_R(k,0) = \nu_0[1 + (\chi/8\pi\nu_0^2)\ln(\Lambda/k)]^{1/2}. \tag{4.12}$$

An analogous calculation gives

$$\nu_R(0,\omega)=\nu_0[1+(\chi/16\pi\nu_0^2)\ln(\nu_0\Lambda^2/\omega)]^{1/2}. \quad (4.13)$$

These results agree, of course, with those obtained in Sec. III by different methods.

B. Burger's equation in one dimension

The results described above can be extended to the Burger's equation model described in Sec. II B, but only in one dimension. It is only in one dimension that a fluctuation-dissipation theorem relates the response and correlation functions generated by Burger's equation in a simple way. Just as for model A, one can show that the vertex corrections to Burger's equation are negligible in the infrared limit. It follows, repeating the arguments of the previous subsection, that

$$\eta=\tfrac{1}{2}, \quad z=\tfrac{3}{2} \quad (4.14)$$

for Burger's equation in one dimension.

C. Model B

There is no obvious fluctuation-dissipation theorem for this model, which complicates the analysis. We might again expect that the response and correlation functions scale, and postulate the functional forms

$$G^{-1}(k,\omega)=k^z g(\omega/k^z), \quad G_R^{-1}(k,\omega)=k^{2-\eta}g_R(\omega/k^z). \quad (4.15)$$

Linearized hydrodynamics gives the results $\sigma=\eta=0$, $z=2$, and

$$G_0(k,\omega)=2D_0/[\omega^2+\nu_0^2k^4], \quad G_{0R}(k,\omega)=1/[-i\omega+\nu_0k^2]. \quad (4.16)$$

We determine the exponents to leading order in $\epsilon=4-d$ by the following arguments:

(i) In the limit $k\to 0$ we have $G_R^{-1}(0,\omega)=-i\omega$ which requires again that

$$z=2-\eta. \quad (4.17)$$

(ii) Arguments similar to those used in Appendix B show that the leading vertex corrections for this model vanish. We have not been able to extend this result to higher orders.

(iii) Neglecting vertex renormalization, power counting arguments relate the exponents η, z, and σ,

$$4\eta+z+\sigma=\max(2,6-d). \quad (4.18)$$

The critical dimension in this case is $d=4$. A detailed calculation of the self-energy to first order in $\epsilon=4-d$ shows that σ is at least of order ϵ^2. Thus the above relations give, to leading order in $\epsilon=4-d$,

$$\eta=\tfrac{1}{3}\epsilon+\mathcal{O}(\epsilon^2), \quad z=2-\tfrac{1}{3}\epsilon+\mathcal{O}(\epsilon^2). \quad (4.19)$$

V. SUMMARY

We have applied renormalization group methods useful in studies of dynamic critical phenomena to the large-distance, long-time properties of a randomly stirred fluid. Long-time tail phenomena and the large eddy properties of the forced Navier-Stokes equation are understood in terms of an attractive hydrodynamic fixed point above two dimensions. A stable, nontrivial fixed point appears below $d=2$. The slow approach to the hydrodynamic fixed point in exactly two dimensions leads to logarithmic corrections to conventional hydrodynamics. Although the physical significance of an incompressible fluid in less than two dimensions is unclear, we have produced a model, model B, for which hydrodynamics breaks down below *four* dimensions. This breakdown is described in terms of scaling laws and exponents, and is associated with a nontrivial fixed point.

Most results obtained for a fluid near thermal equilibrium were derived previously using the mode-coupling approximation.[19] The large eddy properties of a randomly stirred fluid presumably follow from the analogous approximation scheme in turbulence theory, Kraichnan's direct interaction approximation.[33] Many of the results obtained in renormalization studies of critical dynamics[4] were also anticipated by mode coupling theories.[34]

The advantage of the approach taken here is that it renders these essentially uncontrolled approximations systematic. Specifically, we find that simple low-order perturbation schemes (such as the mode-coupling or DIA integral equations) are adequate provided we are interested only in the small k, small ω properties of the correlations. Conversely, our results suggest that these approximations would be rather poorly suited to treat the *ultraviolet* behavior of a fluid. Heuristically at least, one can imagine running the recursion formulas presented here backwards in an attempt to study large wave-number properties. However, the recursion relations derived here [see, e.g., Eq. (3.26)] indicate that the effective coupling constant $\bar{\lambda}(l)$ will become *larger*, not smaller, at these wave numbers. Weak-coupling perturbation schemes simply will not work.

Accompanying the formal manipulations entering a renormalization group transformation are Wilson's ideas[16] about irrelevant variables. To test the importance of a particular term in a fluid equation, we simply study its recursion formula in the vicinity of the fixed point of interest. If this term decays rapidly to zero, it can be neglected in an analysis of the infrared long-time behavior. Similar reasoning leads us to neglect corrections to the Gaussian character of the random force in the lim-

it. To convincingly exclude non-Gaussian or velocity dependant contributions to the force using microscopic arguments alone appears to be rather difficult.

We would also like to point out the essential simplicity of the considerations presented here. In particular, the recursion relations (3.21)–(3.23) for Model A can be written down on general grounds, viz.: (i) $D_I > D_0$ is a stability requirement, (ii) $D_I/D_0 = \nu_I/\nu_0$ is a consequence of the fluctuation-dissipation theorem, and (iii) $\lambda_I = \lambda_0$ follows from Galilean invariance in conjunction with momentum conservation. Thus the only feature of Eqs. (3.21)–(3.23) which requires explicit calculation to order λ^2, is the magnitude of the coefficient A_d; in other words the amplitude of long-time tails, while their leading exponents are universally determined.

In conclusion, we believe there are advantages in describing fluctuating hydrodynamics in terms of fixed points and recursion flows. It is our hope that the techniques described here will be of utility in attacking other problems in fluid mechanics or in irreversible statistical mechanics.

APPENDIX A: EXPLICIT CALCULATION OF A RECURSION FORMULA

As an illustration of the analysis developed in Sec. III, we extract a recursion relation for model A from the graphs shown in Fig. 1(a). The procedure is very similar to that sketched by Ma and Mazenko,[8] but is complicated by a proliferation of indices appearing on transverse projection operators. The indices and momenta accompanying the Feynman graph of Fig. 1(a) are shown in Fig. 3, together with the meaning of its constituent elements.

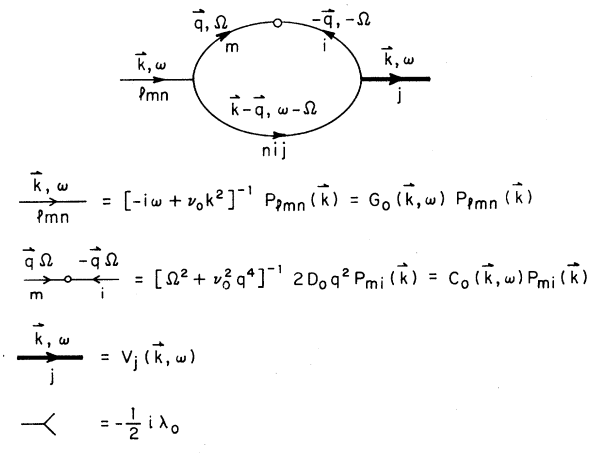


FIG. 3. The momenta, frequencies, and vector indices accompanying the Feynman graph shown in Fig. 1(a). The meanings of the various constituent elements are also indicated.

Writing out the algebraic equation associated with Fig. 1(a), we have

$$v_i^<(\vec{k},\omega) = G_0(k,\omega) f_i^<(\vec{k},\omega) + 4G_0(k,\omega)(\tfrac{1}{2}i\lambda_0)^2 \times P_{lmn}(\vec{k}) I_{mnj}(\vec{k},\omega) v_j^<(\vec{k},\omega) + \cdots , \quad \text{(A1)}$$

where

$$I_{mnj}(\vec{k},\omega) = \int_{q\Omega}^{>} P_{nij}(\vec{k}-\vec{q}) G_0(\vec{k}-\vec{q},\omega-\Omega) \times C_0(q,\Omega) P_{im}(\vec{q}) . \quad \text{(A2)}$$

A combinatorial factor 4 is associated with the graph of Fig. 3, and there is an implied summation over repeated indices in (A1) and (A2). The symbol $\int_{q\Omega}^{>}$ means $(2\pi)^{-(d+1)} \int_{-\infty}^{\infty} d\Omega \int d^d q$, where the momentum integrals are restricted to the domain $e^{-l} < |\vec{q}| < 1$, $e^{-l} < |\vec{k}-\vec{q}| < 1$. The upper momentum cutoff has been fixed, for convenience, at unity.

Equation (A1) can be rearranged to give an equation of motion of the form

$$[-i\omega + k^2 \nu_I(k,\omega)] v_i^<(\vec{k},\omega) = f_i^<(\vec{k},\omega) + \cdots , \quad \text{(A3)}$$

where $\nu_I(k,\omega) = \nu_0 + \Delta\nu_I(k,\omega)$, with

$$k^2 \Delta\nu_I(k,\omega) P_{ij}(\vec{k}) = \lambda_0^2 P_{lmn}(\vec{k}) I_{mnj}(\vec{k},\omega) . \quad \text{(A4)}$$

Thus the diagram in Fig. 3 renormalizes the viscosity. The appearance of the projection operator on the left-hand side of this equation is an obvious consequence of symmetry.

The frequency integral in (A2) is readily done, with the result

$$k^2 \Delta\nu_I(k,\omega) P_{ij}(\vec{k}) = \lambda_0^2 (D_0/\nu_0) P_{lmn}(\vec{k}) k_i \times \int^{>} \frac{d^d q}{(2\pi)^d} \frac{P_{nj}(\tfrac{1}{2}\vec{k}-\vec{q}) P_{mi}(\tfrac{1}{2}\vec{k}+\vec{q})}{-i\omega + 2\nu_0 q^2 + \tfrac{1}{2}\nu_0 k^2} , \quad \text{(A5)}$$

where we have made a convenient change of variables, namely $\vec{q} \to \vec{q} + \frac{1}{2}\vec{k}$, and used the properties of the projection operators. It is now apparent that, for small k and ω, we can evaluate the integral in (A5) at $k = 0$ and $\omega = 0$. The momentum integral is then restricted to the spherical shell $e^{-l} < |\vec{q}| < 1$, and the angular average of the projection operators is evaluated straightforwardly. We obtain finally

$$\Delta\nu_I(0,0) = \frac{\lambda_0^2}{2} \frac{D_0}{\nu_0^2} \frac{d^2-2}{d^2+2d} \frac{s_d}{(2\pi)^d} \int_{e^{-l}}^{1} \frac{q^{d-1} dq}{q^2} , \quad \text{(A6)}$$

where $s_d = 2\pi^{(1/2)d}/\Gamma(\frac{1}{2}d)$. Thus the "intermediate" viscosity $\nu_I = \nu_0 + \Delta\nu_I(0,0)$ is given by (3.10), with

$$A_d = \frac{1}{2}\left(\frac{d^2-2}{d^2+2d}\right)\frac{s_d}{(2\pi)^d} . \quad \text{(A7)}$$

It is a simple matter, in principle, to work out

graph rules that describe the extension of this second-order calculation to higher orders. We will not do so here but mention that (i) the sum of all $1PI$ diagrams with one incoming and one outgoing arrow renormalizes the viscosity ν_0, (ii) the sum of all $1PI$ diagrams with two incoming arrows renormalized the force strength D_0, and (iii) the sum of all $1PI$ diagrams with one incoming and two outgoing arrows renormalizes the coupling constant λ_0. $1PI$ diagrams are those which do not fall apart if a single line is cut, and they have no factors G_0 associated with the external lines. The momentum integrals over the internal lines are all restricted to the shell $e^{-l} < |\vec{q}| < 1$.

Finally, we note that if we integrate $\vec{q}$ in (A5) over the full momentum space, $0 < |\vec{q}| < 1 (= \Lambda)$, we obtain the quantity $\nu_R(k,\omega) - \nu_0$ defined by Eq. (3.37), to order $\bar{\lambda}_0^2$. The calculation of long-time tail corrections like (3.40) is based on (A5), in conjunction with the homogeneity relation (3.38) for ν_R. Equation (A5) is, of course, then just the standard-mode coupling formula.

APPENDIX B: FLUCTUATION-DISSIPATION THEOREM AND VERTEX CORRECTIONS IN MODEL A

A perturbation theory applicable to the Navier-Stokes equation has been developed by Martin, Siggia, and Rose (MSR)[9] and others. The theory involves correlation functions $[G(k,\omega)]$, response functions $[G_R(k,\omega)]$ and vertex functions (Γ), and in general all these must be worked out. In the case of model A a fluctuation-dissipation theorem (FDT) exists and the perturbation theory can be simplified considerably.

Several classes of classical processes for which FDT's exist have been discussed by Deker and Haake.[35] Model A exhibits detailed balance, and the irreversible terms in the equation of motion are linear in the velocity; it thus corresponds to the second class considered by Deker and Haake. It is then easy to show that the equal-time velocity fluctuations are determined by the Gaussian distribution

$$P_0 \sim \exp\left[-(\nu_0/2D_0)\sum_{\vec{k}} v_i(\vec{k})v_i^*(\vec{k})\right] \tag{B1}$$

and thus show no corrections in any dimension. In particular this result also applies to Burger's equation (2.8) with the forcing function (2.10).

The FDT takes the form

$$G(k,\omega) = (D_0/\nu_0)[G_R(k,\omega) + G_R^*(k,\omega)] \tag{B2}$$

connecting the correlation and response functions. This relation can also be written as a relation connecting the self-energies. The self-energies are defined by

$$G(k,\omega) = G_R(k,\omega)[2D_0k^2 + \Sigma(k,\omega)]G_R^*(k,\omega)\,, \tag{B3}$$

$$G_R(k,\omega) = [-i\omega + \nu_0 k^2 - \Sigma_R(k,\omega)]^{-1}\,. \tag{B4}$$

Equation (B2) becomes

$$\Sigma(k,\omega) = -(D_0/\nu_0)[\Sigma_R(k,\omega) + \Sigma_R^*(k,\omega)]\,. \tag{B5}$$

The perturbation theory of MSR can be further simplified for model A by noting that vertex corrections vanish in the infrared limit. This result is essentially a consequence of Galilean invariance and is easily proved to all orders in perturbation theory. The theory of MSR involves three vertices. However, in the case when the steady-state distribution is Gaussian it has been shown by Kawasaki[36] that only one type of vertex need be considered. In the notation of MSR one then shows that the vertex $\Gamma^{(1)}_{ijm}(\vec{k}_1,\vec{k}_2;\, t_1,t_2,t_3)$ reduces to its bare value $\gamma_{ijm}(\vec{k}_1+\vec{k}_2)\delta(t_1-t_2)\delta(t_1-t_3)$ when the external momenta $\vec{k}_1$ and $\vec{k}_2$ are small. The bare interaction has the form

$$\gamma_{ijm}(\vec{k}) = -ik_m(\delta_{ij} - \hat{k}_i\hat{k}_j) - ik_j(\delta_{im} - \hat{k}_i\hat{k}_m)\,. \tag{B6}$$

Typical diagrams contributing to Γ are shown in Figs. 4(a) and 4(b). The method of proof employed here is similar to some methods used by Deker and Haake. We need the following:

(a) We consider each diagram of perturbation theory at fixed times of the vertices. We note that t_1 is always the latest time in a diagram because the t_1 vertex is connected to every other vertex through a series of retarded response functions.

(b) On the internal lines of a diagram we neglect the external momenta $\vec{k}_1$ and $\vec{k}_2$.

(c) The bare interaction has the form (B6) and one notices that the pressure terms will not contribute if the vertex is an internal one (i.e., has no external lines attached), because of the incompressibility condition. Thus, what enters is

$$\gamma'_{ijm}(\vec{k}) = -ik_m\delta_{ij} - ik_j\delta_{im}\,. \tag{B7}$$

(d) We use the FDT (B2) to replace all correlation functions by retarded or advanced response functions depending on the time labels attached to the

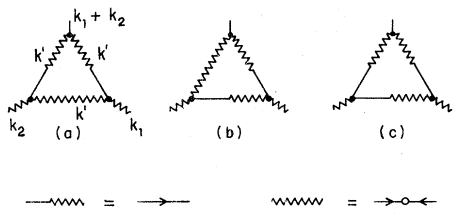


FIG. 4. The two third-order diagrams (a) and (b) in case $t_1 > t_2 > t_3$. Response functions G_R are denoted by ∿∿— and correlation functions G by ∿∿∿. In (a) we have neglected the external mometa on the internal lines of the diagram.

two vertices involved.

Making use of (a)–(d) alone it is possible to show that all the perturbation terms (except the bare interaction) for $\Gamma^{(1)}$ cancel. An example is given in Figs. 4(a) and 4(b). With the time order $t_1 > t_2 > t_3$ we use (d) to convert both these diagrams into the form shown in Fig. 4(c). The lower left vertex has the form $\gamma'_{ijm}(-\vec{k}')$ in 4(a) and $\gamma'_{ijm}(\vec{k}')$ in 4(b), the rest of the diagram being the same in both cases. As $\gamma'_{ijm}(-\vec{k}') + \gamma'_{ijm}(\vec{k}') = 0$, the sum of 4(a) and 4(b) vanishes.

In general we look at the vertex with the earliest time label. If this is an internal vertex there will be three diagrams which are distinct in that this vertex appears in the three different orientations, the rest of the diagrams being the same. When we use the FDT the three vertex ends become equivalent, and the sum of the three contributions will thus be proportional to (omitting the pressure terms) $\gamma'_{ijm}(\vec{k}) + \gamma'_{ijm}(\vec{k}_1) + \gamma'_{ijm}(\vec{k}_2) = 0$. If the vertex with the earliest time is an external one (either t_2 or t_3) it can appear as in Figs. 4(a) and 4(b), the rest of the diagram being the same. The sum of the two contributions will cancel as in the example of Fig. 4.

Thus in the limit of small external momenta the vertex $\Gamma^{(1)}$ reduces to its bare value (B6), with no corrections due to renormalization. This result also applies to Burger's equation in one dimension.

ACKNOWLEDGMENTS

This work was carried out independently at Temple University and at Harvard University. We would like to thank P. C. Martin and E. Siggia for helpful discussions. Dieter Forster would also like to acknowledge stimulating and useful conversations with M. Droz, M. Green, and J. D. Gunton. David R. Nelson is particularly indebted to U. Frisch for helpful comments.

---

*Work supported in part by the National Science Foundation under Grant No. DMR 76-11974.

†Work supported by a Junior Fellowship, Harvard Society of Fellows, and in part by the National Science Foundation under Grant No. DMR 72-02977 A03.

‡Work supported in part by the Rutgers University Research Council and the National Science Foundation under Contract No. GH 38458. Part of this work was done at Harvard University whose hospitality is gratefully acknowledged.

[1]See, e.g., S. Coleman, in *Properties of the Fundamental Interactions* 1971, edited by A. Zichichi (Editrice Compositori, Bologna, Italy, 1972); H. D. Politzer, Phys. Rep. 14C, 131 (1974).

[2]K. G. Wilson and J. Kogut, Phys. Rep. 12C, 77 (1974); M. E. Fisher, Rev. Mod. Phys. 42, 597 (1974); S.-K. Ma, Rev. Mod. Phys. 45, 589 (1973). See also S.-K. Ma, *Modern Theory of Critical Phenomena* (Benjamin, Reading, Mass., 1976).

[3]For an application to the Kondo problem, see K. G. Wilson, Rev. Mod. Phys. 47, 773 (1975).

[4]See, e.g., B. I. Halperin, P. C. Hohenberg, and S. K. Ma, Phys. Rev. Lett. 29, 1548 (1972); Phys. Rev. B 10, 137 (1974); B. I. Halperin, P. C. Hohenberg, and E. Siggia, Phys. Rev. Lett. 28, 548 (1974); S. Ma and G. Mazenko, Phys. Rev. Lett. 33, 1384 (1974).

[5]K. Kawasaki, Ann. Phys. (N.Y.) 61, 1 (1970).

[6]In particular, we restrict ourselves here to *incomprehensible* fluids. The compressibility of a fluid diverges at the critical point.

[7]P. G. Saffman, J. Fluid Mech. 27, 581 (1967).

[8]S.-K. Ma and G. Mazenko, Phys. Rev. B 11, 4077 (1975).

[9]P. C. Martin, E. G. Siggia, and H. A. Rose, Phys. Rev. A 8, 423 (1973).

[10]This has been done for spin systems by R. Bausch, H. K. Jansen, and H. Wagner (Institut für Festkörperforschung, KFA Jülich preprint).

[11]A. S. Monin and A. M. Yaglom, *Statistical Fluid Mechanics*, Vol. II (M.I.T., Cambridge, 1975).

[12]An introduction to modern turbulence theory is given by S. A. Orszag, in *Hydrodynamics*, edited by R. Balian (Gordon and Breach, New York, 1974).

[13]G. K. Batchelor and I. Proudman, Philos. Trans. R. Soc. A 248, 369 (1956).

[14]Results for model A have appeared elsewhere, D. Forster, D. R. Nelson, and M. J. Stephen, Phys. Rev. Lett. 36, 867 (1976).

[15]This interpretation is due to P. C. Martin.

[16]K. G. Wilson, Phys. Rev. B 4, 3174 (1971).

[17]U. Frisch, M. Lesieur, and P. L. Sulem, Phys. Rev. Lett. 37, 895 (1976).

[18]A discussion of Burger's equation is given by P. G. Saffman, in *Topics in Nonlinear Physics*, edited by N. J. Zabusky (Springer-Verlag, New York, 1968).

[19]B. J. Alder and T. E. Wainwright, Phys. Rev. A 1, 18 (1970); M. H. Ernst, E. H. Hauge, and J. M. J. van Leeuwen, Phys. Rev. Lett. 25, 1254 (1970) and Phys. Rev. A 4, 2055 (1971); J. R. Dorfman and E. G. D. Cohen, Phys. Rev. Lett. 25, 1257 (1970) and Phys. Rev. A 6, 776 (1972) and Phys. Rev. A 12, 292 (1975). For two recent reviews and many additional references, see Y. Pomeau and P. Résibois, Phys. Rep. 19C, 64 (1975), and R. Dorfman, in *Fundamental Problems in Statistical Mechanics*, edited by E. G. D. Cohen (North-Holland/American Elsevier, Amsterdam, 1975).

[20]R. Zwanzig, in *Lectures in Theoretical Physics*, edited by W. R. Brittin (Interscience, New York, 1961) and Phys. Rev. 124, 983 (1961); H. Mori, Prog. Theor. Phys. (Kyoto) 33, 423 (1965).

[21]L. P. Kadanoff and P. C. Martin, Ann. Phys. (N.Y.) 24, 419 (1963).

[22]See, e.g., R. Zwanzig, in *Proceedings of the I.U.P.A.P. Conference on Statistical Mechanics*, edited by S. A. Rice, K. Freed, and J. Light (University of Chicago Press, Chicago, 1972).

[23]See, e.g., K. S. J. Nordholm and R. Zwanzig, J. Stat.

Phys. 11, 143 (1974); L. S. García-Colín and R. M. Velasco, Phys. Rev. A 12, 646 (1975).

[24]Except in an equilibrium fluid where, as outlined in Sec. IIC, the cutoff is considered part of the equations of motion themselves, with high-k modes treated as noise.

[25]Strictly speaking, the average over $f^>$ should be done at constant $v^<$, rather than $f^<$.

[26]See, e.g., R. H. Kraichnan, J. Fluid Mech. 5, 497 (1959); W. W. Wyld, Jr., Ann. Phys. (N.Y.) 14, 143 (1961), and Ref. 10.

[27]F. J. Wegner, Phys. Rev. B 5, 4529 (1972).

[28]S. F. Edwards and W. D. McComb, J. Phys. A 2, 157 (1969).

[29]As shown in Ref. 9, renormalized dynamical theories in general involve *three* distinct quadratic nonlinearities in the equations of motion. However, the nonlinearities considered in Sec. III dominate the infrared behavior. The remaining nonlinearities (which do not appear in the unrenormalized equations of motion) involve higher powers of k and are irrelevant variables.

[30]In precisely two dimensions the equilibrium distribution of a fluid is expected to correspond to constant energy $\sum_k \vec{v}_k \cdot \vec{v}_{-k}$ *and* constant vorticity $\sum_k k^2 \vec{v}_k \cdot \vec{v}_{-k}$. This would be engendered by a forcing function $D(k) = \nu k^2/[\beta + \gamma k^2]$ as was pointed out to us by Herbert Wagner. The difference to the forcing function $D_0 k^2$ considered for model A is irrelevant for the infrared properties.

[31]B. I. Halperin, P. C. Hohenberg, and E. D. Siggia, Phys. Rev. B 13, 1299 (1976).

[32]Larkin and Khmelnitski, Zh. Eksp. Teor. Fiz. 56, 2087 (1969) [Sov. Phys.-JETP 29, 1123 (1969)]; and T. Tsuneto and E. Abrahams, Phys. Rev. Lett. 30, 217 (1973).

[33]R. H. Kraichman (private communication); see Ref. 26 and R. H. Kraichnan, Adv. Math. 16, 305 (1975).

[34]The relation between renormalization group and mode-coupling theories of critical dynamics is discussed by K. Kawasaki and J. Gunton, Phys. Rev. B 13, 4658 (1976).

[35]U. Deker and F. Haake, Phys. Rev. A 11, 2043 (1975).

[36]K. Kawasaki, Prog. Theor. Phys. 52, 1527 (1971).

Critical Phenomenon of the Order-Disorder Transition in Incompressible Active Fluids

Leiming Chen

College of Science, China University of Mining and Technology, Xuzhou Jiangsu, 221116, P. R. China

E-mail: leimingyandongyu@gmail.com

John Toner

Department of Physics and Institute of Theoretical Science, University of Oregon, Eugene, OR 97403
Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Str. 38, 01187 Dresden, Germany

E-mail: jjt@uoregon.edu

Chiu Fan Lee

Department of Bioengineering, Imperial College London, South Kensington Campus, London SW7 2AZ, U.K.

E-mail: c.lee@imperial.ac.uk

Abstract. We study incompressible systems of motile particles with alignment interactions. Unlike their compressible counterparts, in which the order-disorder (i.e., moving to static) transition, tuned by either noise or number density, is discontinuous, in incompressible systems this transition can be continuous, and belongs to a *new* universality class. We calculate the critical exponents to $\mathcal{O}(\epsilon)$ in an $\epsilon = 4-d$ expansion, and derive two exact scaling relations. This is the first analytic treatment of a phase transition in a new universality class in an active system.

PACS numbers: 05.65.+b

PACS numbers: 64.60.Ht

PACS numbers: 87.18Gh

Submitted to: *New J. Phys.*

1. Introduction

Emergent properties of interacting non-equilibrium systems are of widespread and fundamental interest. One of the simplest, but most striking, of these is the self-organized phenomenon of "flocking"- that is, collective motion (CM) in large groups of motile organisms [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. This phemenon is fascinating in part because its occurrence in two spatial dimensions requires the spontaneous breaking of a continuous symmetry, which is forbidden in thermal equilibrium by the Mermin-Wagner theorem [11]. It was initially hoped [2] that the transition into this novel state could be a continuous one belonging to a new universality class. However, it was subsequently realized, from both simulations and theoretical analysis [12, 13, 14, 15, 16, 17, 18] of the hydrodynamic equations [3, 4, 5, 6], that as this putative continuous transition is approached from the ordered side, but before it can be reached, the homogeneous CM state becomes unstable to modulation of the density along the mean velocity. The transition from the homogeneous CM state (i.e., the ordered state) to the disordered state proceeds via two first order transitions: one from homogeneous to banded, the next from banded to disordered.

Since this instability requires density variations, we reason that the instability might be eliminated by making the system incompressible. In this paper, we show that, indeed, the order-disorder transition is continuous in an incompressible system, and belongs to a new universality class. We demonstrate this by finding, in a dynamical renormalization group (DRG) analysis of the hydrodynamic equations for an incompressible active fluid in d spatial dimensions, a novel stable fixed point that controls the transition. This calculation is done to order $\mathcal{O}(\epsilon)$ in an $\epsilon = 4 - d$ expansion; to the same order, we calculate the critical exponents of the transition. We also obtain two scaling laws relating these critical exponents which are valid to all orders in ϵ (i.e., exact).

Our results are testable in both experiments and simulations. Three potential realizations are:

(i) Systems with strong repulsive short-ranged interactions between the active particles. Incompressibility has, in fact, been assumed in, e.g., recent experimental studies on cell motility [19]. In such systems, the compressibility will be non-zero,but small. Hence, our incompressible results will apply out to very large length scales, or, equivalently, very close to the transition, but will ultimately crossover to the compressible behavior (i.e., a small first order transition driven by the banding instability).

(ii) Systems with long-ranged repulsive interactions; here, true incompressibility is possible. Long ranged interactions are quite reasonable in certain contexts: birds, for example, can often see all the way across a flock [20].

(iii) Motile colloidal systems in fluid-filled microfluidic channels.The forces exerted by the active particles are, of course, tiny compared to what would be needed to compress the background fluid, so that fluid is effectively incompressible. Since the active particles drag the background fluid with them, their motion is effectively

incompressible as well. Indeed, experiments [21] show these systems do not exhibit the banding instability [12, 13, 14, 15, 16, 17, 18] found in all compressible active systems. This also suggests a numerical approach: simulating active particles moving through an incompressible fluid [22].

2. Generic model of incompressible active fluids

We formulate the most general hydrodynamic model for systems lacking both momentum conservation, and Galilean invariance, consistent with the symmetries of rotation and translation invariance, and the assumption of incompressibility. As the number density cannot fluctuate (by the assumption of incompressibility), the velocity field is the only hydrodynamic variable in the problem, which becomes soft as the transition is approached. Since the velocity is small near the transition, we can expand the equation of motion (EOM) in powers of the velocity. The symmetry constraints of translation and rotation invariance force the EOM valid at long wavelengths and times to take the form: [23, 3, 4, 5, 6]

$$\partial_t \mathbf{v} + \lambda(\mathbf{v}\cdot\nabla)\mathbf{v} = -\nabla\mathcal{P} - (a + b|\mathbf{v}|^2)\mathbf{v} + \mu\nabla^2\mathbf{v} + \mathbf{f}. \tag{1}$$

where the pressure $\mathcal{P}$ enforces the incompressibility condition $\nabla\cdot\mathbf{v} = 0$, $\mathbf{f}$ is a "white noise" with spatio-temporally Fourier transformed statistics:

$$\langle f_m(\mathbf{k},\omega) f_n(\mathbf{k}',\omega')\rangle = 2DP_{mn}(\mathbf{k})\delta(\mathbf{k}+\mathbf{k}')\delta(\omega+\omega') \ , \tag{2}$$

and $P_{mn}(\mathbf{k}) \equiv \delta_{mn} - k_m k_n/k^2$ is the transverse projection operator. This EOM (Eq. (1)) reduces, when $a = 0 = b$, to the classic model of a fluid forced at zero wavenumber treated by [24] (their "model B"). With $\lambda = 0$, it reduces to a simple, time-dependent Ginzburg-Landau (TDGL) [25, 26] dynamical model for an isotropic ferromagnet with long ranged dipolar interactions [27, 28].

Because our system lacks Galilean invariance, λ need not (and in general will not) be one, and the terms $-(a + b|\mathbf{v}|^2)\mathbf{v}$ are allowed in the EOM. The latter is crucial as it explains why there can be a polar ordered phase in an active system, which is not possible in a normal fluid.

3. Novel universality class

At the mean field level, for $a < 0, b > 0$ the system is in the ordered phase with $|\mathbf{v}| = \sqrt{-a/b}$, and for $a > 0, b > 0$ it is in the disordered phase with $|\mathbf{v}| = \mathbf{0}$. To go beyond this mean field description, we employ the DRG method [24] near the order-disorder transition. To do so, we spatio-temporally Fourier transform Eq. (1), and project orthogonal to wavevector $\mathbf{k}$; obtaining

$$\begin{aligned} v_l(\tilde{\mathbf{k}}) = G(\tilde{\mathbf{k}})\Bigg[f_l(\tilde{\mathbf{k}}) - \frac{\mathrm{i}\lambda}{2} P_{lmn}(\mathbf{k}) \int_{\tilde{\mathbf{q}}} v_m(\tilde{\mathbf{q}}) v_n(\tilde{\mathbf{k}}-\tilde{\mathbf{q}}) \\ - \frac{b}{3} Q_{lmnp}(\mathbf{k}) \int_{\tilde{\mathbf{q}},\tilde{\mathbf{h}}} v_m(\tilde{\mathbf{k}}-\tilde{\mathbf{q}}-\tilde{\mathbf{h}}) v_n(\tilde{\mathbf{q}}) v_p(\tilde{\mathbf{h}}) \Bigg] \end{aligned} \tag{3}$$

where we have adopted the reduced notations $\tilde{\mathbf{k}} \equiv (\mathbf{k}, \omega)$ and $\int_{\tilde{\mathbf{q}}} = \int_{\mathbf{q},\Omega} \equiv \int \frac{d^d q}{(2\pi)^d} \frac{d\Omega}{2\pi}$, and we have defined $P_{lmn}(\mathbf{k}) \equiv P_{lm}(\mathbf{k}) k_n + P_{ln}(\mathbf{k}) k_m$, $Q_{lmnp}(\mathbf{k}) \equiv P_{lm}(\mathbf{k})\delta_{np} + P_{ln}(\mathbf{k})\delta_{mp} + P_{lp}(\mathbf{k})\delta_{mn}$, and the "propagator" $G(\tilde{\mathbf{k}}) \equiv (-\mathrm{i}\omega + \mu k^2 + a)^{-1}$. Graphical representations of the various terms in Eq. (3) are shown in Fig. 1.

We now perform the standard DRG procedure [24], averaging over short wavelength degrees of freedom, and rescaling: $\mathbf{r} \to \mathbf{r}e^{\ell}$, $t \to te^{z\ell}$ and $\mathbf{v} \to e^{\chi\ell}\mathbf{v}$. Our procedure is identical to the calculation for model B in [24], except for a modified propagator, and some additional Feynmann graphs due to the extra $b|\mathbf{v}|^2\mathbf{v}$ non-linearity in our problem. At the one loop level, the non-vanishing graphical contributions to the various coefficients in Eq. (3) are shown in Fig. 2. More details of the calculation are given in

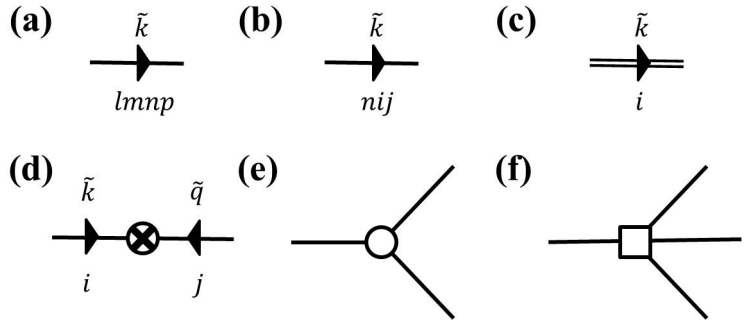


Figure 1. Graphical representations: (a)= $Q_{lmnp}(\mathbf{k})G(\tilde{\mathbf{k}})$; (b) = $P_{nij}(\mathbf{k})G(\tilde{\mathbf{k}})$; (c) = $v_i(\tilde{\mathbf{k}})$; (d) = $2DP_{ij}(\mathbf{k}) \mid G(\tilde{\mathbf{k}}) \mid^2$; (e) = $-\frac{\mathrm{i}}{2}\lambda$; (f) = $-\frac{b}{3}$.

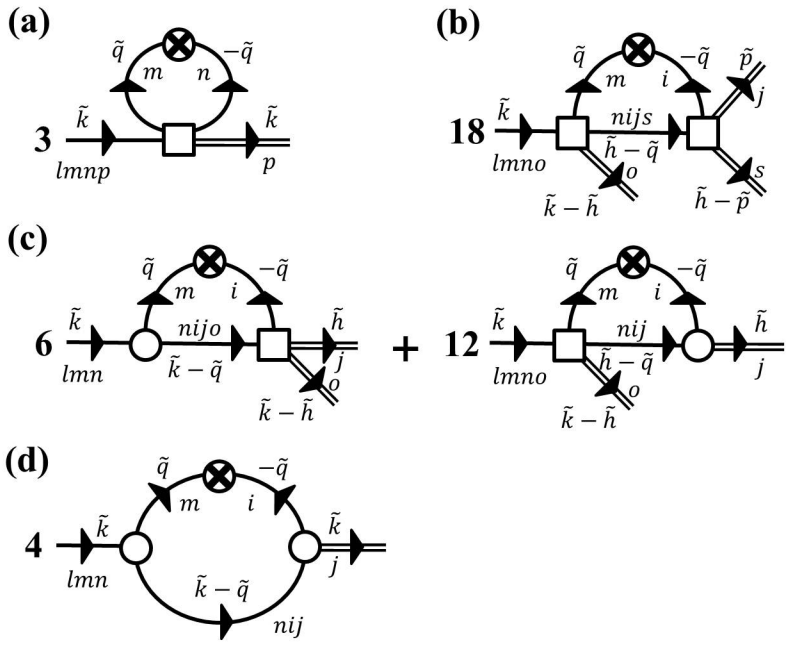


Figure 2. Non-vanishing diagrams at the one-loop level. Diagrams (a) to (d) contribute to a, b, λ and μ respectively.

the supplemental materials. We obtain, in d spatial dimensions, the following RG flow equations of the coefficients to one loop order and to linear order in $\epsilon \equiv 4-d$ [29]:

$$\frac{da}{d\ell} = za - \frac{9g_2}{2}(a-\mu\Lambda^2), \tag{4}$$

$$\frac{db}{d\ell} = \left(2\chi + z - \frac{17g_2}{2}\right) b, \tag{5}$$

$$\frac{d\lambda}{d\ell} = \left(\chi - 1 + z - \frac{5g_2}{3}\right) \lambda, \tag{6}$$

$$\frac{d\mu}{d\ell} = \left(-2 + z + \frac{g_1}{4}\right) \mu, \tag{7}$$

$$\frac{dD}{d\ell} = (-2\chi + z - d)\, D\ . \tag{8}$$

where we've defined dimensionless couplings:

$$g_1 \equiv \frac{S_d D\lambda^2}{(2\pi)^d \mu^3}\Lambda^{-\epsilon},\ \ g_2 \equiv \frac{S_d D b}{(2\pi)^d \mu^2}\Lambda^{-\epsilon}\,, \tag{9}$$

and where $S_d \equiv 2\pi^{d/2}/\Gamma(d/2)$ is the surface area of a unit sphere in d dimensions, $\epsilon \equiv 4-d$, and Λ is the ultraviolet wavevector cutoff. Since our interest is in the transition, we have, in the last four recursion relations (5-8), set $a = 0$, and have worked to linear order in a in (4). It is straightforward to verify that higher order terms in a affect none of our results up to and including linear order in $\epsilon = 4-d$.

From these RG flow equations, we can derive two closed flow equations for $g_{1,2}$ for arbitrary χ and z:

$$\frac{dg_1}{d\ell} = \epsilon g_1 - \frac{3}{4}g_1^2 - \frac{10}{3}g_1 g_2\ , \tag{10}$$

$$\frac{dg_2}{d\ell} = \epsilon g_2 - \frac{1}{2}g_1 g_2 - \frac{17}{2}g_2^2\ . \tag{11}$$

Although not necessary, it is convenient to make a special choice of z and χ such that μ and D are kept fixed at their bare values (i.e, μ_0 and D_0, respectively). We will hereafter adopt this choice of z and χ, which is

$$z = 2 - g_1/4 + \mathcal{O}(\epsilon^2) \quad , \quad \chi = \frac{z-d}{2} + \mathcal{O}(\epsilon^2)\,. \tag{12}$$

We will also hereafter use the subscript 0 to denote the bare (i.e., unrenormalized) values of the parameters.

Eq. (4) now becomes

$$\frac{da}{d\ell} = \left(2 - \frac{g_1}{4} - \frac{9}{2}g_2\right) a + \frac{9}{2}g_2\mu\Lambda^2\ . \tag{13}$$

Eqs. (10,11,13) have a non-Gaussian fixed point in $d < 4$:

$$g_1^* = \frac{124}{113}\epsilon + \mathcal{O}(\epsilon^2)\ ,\ g_2^* = \frac{6}{113}\epsilon + \mathcal{O}(\epsilon^2)\ ,\ a^* = \left[-\frac{27}{226}\epsilon + \mathcal{O}(\epsilon^2)\right]\mu\Lambda^2, \tag{14}$$

which can be shown by analyzing the three recursion relations to be a stable attractor of all points on a two-dimensional surface (the "critical surface") in the three-dimensional parameter space (g_1, g_2, a), but to be unstable with respect to displacements off this

critical surface. The flows on the critical surface are illustrated in Fig. 3. This is exactly the topology of renormalization group flows that corresponds to a continuous phase transition with universal exponents controlled by the fixed point that's stable within the critical surface. Hence, we conclude that the order-disorder *is* generically continuous in incomprssible active fluids. Furthermore, as far as we know, the fixed point we've obtained is novel, and thus the critical behaviour of incompressible active fluids belongs to a new universality class.

4. Critical exponents

The exponential runaway from the critical surface in the unstable direction near the stable fixed point (14) grows like $e^{y_a \ell}$, with the exponent

$$y_a = 2 - \frac{58}{113}\epsilon + \mathcal{O}(\epsilon^2). \tag{15}$$

This eigenvalue determines the critical exponent ν governing the critical behavior of the velocity correlation length ξ. In addition, the smallest (in magnitude) of the two negative eigenvalues gives the "correction to scaling exponent" y_2 [25]; we find $y_2 = -\frac{31}{113}\epsilon + \mathcal{O}(\epsilon^2)$.

A useful experimental probe of the transition is the velocity correlation function $\langle \mathbf{v}(\mathbf{r}+\mathbf{R}, t+T)\cdot\mathbf{v}(\mathbf{R},T)\rangle \equiv C(\mathbf{r},t)$, which depends on bare parameters b_0, μ_0, D_0, and λ_0 and, most importantly, the proximity to the phase transition $\delta a_0 \equiv a_0 - a_0^c$, where a_0^c is the value of a_0 at the transition. The RG connects the original $C(\mathbf{r},t)$ to that of the rescaled system:

$$C(\mathbf{r},t;\delta a_0, b_0, \lambda_0) = e^{2\chi\ell} C\Big(\mathbf{r}e^{-\ell}, te^{-z\ell}; \delta a(\ell), b(\ell), \lambda(\ell)\Big), \tag{16}$$

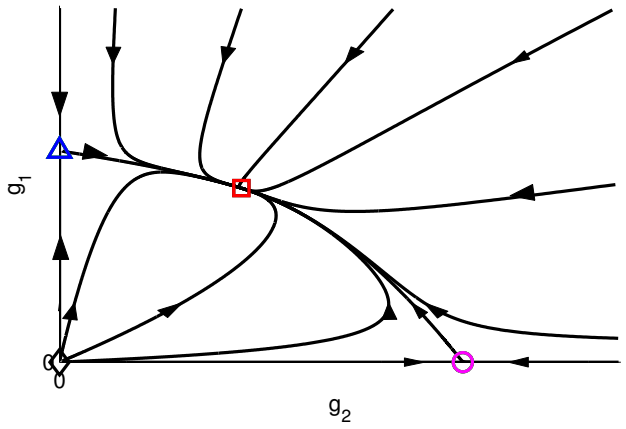


Figure 3. RG flows on the critical surface. Besides the unstable Gaussian fixed point (black diamond) and the stable fixed described in Eqs (14) (red square), there are two unstable fixed points: one at $g_1^* = 0$, $g_2^* = \frac{2\epsilon}{17}$, which is the fixed point of an isotropic ferromagnet with long-ranged dipolar interactions [27] (purple circle), and one at $g_2^* = 0$, $g_1^* = \frac{4\epsilon}{3}$, which is the fixed point of a fluid forced at zero wavevector (Model B of [24]) (blue triangle).

where we have not displayed μ_0 and D_0, since they are kept fixed in the RG. By choosing $\ell = \ln(\Lambda r)$, and using $\delta a(\ell) \approx \delta a_0 e^{y_a \ell}$, we can obtain from this a scaling form for large r:

$$C(\mathbf{r}, t) = r^{2-d-\eta} Y_{\pm}\left(\frac{r}{\xi}, \frac{t}{r^z}\right) \ . \tag{17}$$

In particular, the equal time correlation function scales as $r^{2-d-\eta}$. In Eq. 17, the exponent η is given by

$$\eta = 2 - d - 2\chi = 2 - z + \mathcal{O}(\epsilon^2) = \frac{31}{113}\epsilon + \mathcal{O}(\epsilon^2) \ , \tag{18}$$

the scaling functions $Y_{\pm}$ by

$$Y_{\pm}(x, y) = C(\Lambda^{-1}, y; \pm x^{y_a}, b^*, \lambda^*) \, , \tag{19}$$

and the diverging correlation length ξ by $\xi \equiv \Lambda^{-1} |\delta a_0|^{-\nu}$, where the correlation length exponent

$$\nu = \frac{1}{y_a} = \frac{1}{2} + \frac{29}{226}\epsilon + \mathcal{O}(\epsilon^2). \tag{20}$$

In our expression for the ϵ expansions for η, we have replaced χ and z by their values at the fixed point (14), which is valid given that r is large and the system is sufficiently close to the transition. We do so consistently when calculating other exponents as well. The first line of equation (18) is exact (i.e., independent of the ϵ-expansion), as it is simply the definition of η.

The order-disorder transition can be driven by tuning any one of many microscopic control parameters (e.g., density or noise strength).Whatever control parameter s is tuned, we expect $a_0 - a_0^c \propto (s - s_c)$ by analyticity near s_c, where s_c is the values of the control parameter s at the transition. As a result, the velocity correlation length ξ just defined diverges as $\xi \propto |s - s_c|^{-\nu}$ as *any* control parameter s is tuned.

The scaling functions $Y_{\pm}$ in equation (17) are different on the disordered (+) and ordered (−) sides of the transition, because the system is in different phases in the two cases. On the disordered side , we expect $Y_+(x, y)$ to decay exponentially with both x and y, while on the ordered side, $Y_-(x, y)$ has a more complicated scaling behavior that we'll discuss elsewhere [34].

Now we calculate the magnitude of the order parameter in the ordered state near the critical point. The RG connects the average velocity of the original system and that of the rescaled system with the relation

$$\langle \mathbf{v} \rangle (\delta a_0, b_0, \lambda_0) = e^{\chi \ell} \langle \mathbf{v} \rangle \left(\delta a_0 e^{y_a \ell}, b(\ell), \lambda(\ell) \right) \ . \tag{21}$$

We choose ℓ such that $\delta a_0 e^{y_a \ell}$ is of order 1. Therefore, ℓ is large since δa_0 is small near the critical point, and hence, both $b(\ell)$ and $\lambda(\ell)$ flow to their nonzero fixed values. Then all the singular dependence on $(s_c - s)$ on the RHS of the equality (21) is included in the exponential. This implies $|\langle \mathbf{v} \rangle| \sim |s_c - s|^{\beta}$ with

$$\beta = -\nu\chi = \frac{1}{2} - \frac{6}{113}\epsilon + \mathcal{O}(\epsilon^2) \ . \tag{22}$$

The first equality in this expression, which is *exact*, can be rewritten in terms of η using the definition of η embodied in the first equality of (18), giving the *exact* hyperscaling relation

$$\beta = \frac{\nu}{2}(d - 2 + \eta) \ . \tag{23}$$

In this respect, our system is similar to equilibrium systems, in which (23) also holds [26].

We study next the linear response of the system to a weak external field $\mathbf{H}$; that is, simply adding a small constant vector $\mathbf{H}$ to the RHS of (1). In this case, the RG leads to the scaling relation

$$\langle \mathbf{v} \rangle (\delta a_0, b_0, \lambda_0, H) = e^{\chi \ell} \langle \mathbf{v} \rangle \left(\delta a_0 e^{y_a \ell}, b(\ell), \lambda(\ell), H e^{y_H \ell} \right) \ , \tag{24}$$

where $H \equiv |\mathbf{H}|$ and y_H is the RG eigenvalue of the external field H at the fixed point (14). As there are no one loop graphical corrections to the external field, we can obtain y_H to $\mathcal{O}(\epsilon)$ by simple power counting, which gives

$$y_H = z - \chi + \mathcal{O}(\epsilon^2) \, . \tag{25}$$

Again choosing ℓ such that $\delta a_0 e^{y_a \ell}$ is of order 1, we obtain

$$\langle \mathbf{v} \rangle (\delta a_0, b_0, \lambda_0, H) = (\Lambda \xi)^{\chi} \langle \mathbf{v} \rangle \left(1, b^*, \lambda^*, H \left(\Lambda \xi \right)^{y_H} \right) \ , \tag{26}$$

where b^* and λ^* are the nonzero fixed values of $b(\ell)$ and $\lambda(\ell)$, respectively. Since the expectation value on the right hand side is evaluated in a system far from its critical region (since $\delta a = 1$), we expect linear response to the external field on that side with an order one susceptibility. Hence,

$$\langle \mathbf{v} \rangle (\delta a_0, b_0, \lambda_0, H) \sim (\Lambda \xi)^{\chi} H \left(\Lambda \xi \right)^{y_H \ell} \propto \xi^{\chi + y_H} H \, , \tag{27}$$

which implies a linear susceptibility χ_H which diverges as $|s - s_c|^{-\gamma}$ with

$$\gamma = \nu(\chi + y_H) = \nu z + \mathcal{O}(\epsilon^2) = 1 + \frac{27}{226}\epsilon + \mathcal{O}(\epsilon^2) \, . \tag{28}$$

Note that Eqs. (18,28) seem to suggest that η and γ satisfy Fisher's scaling law $\gamma = (2 - \eta)\nu$. However, since our system is out of equilibrium and thus the fluctuation dissipation theorem is not expected to hold, we do *not* expect Fisher's scaling law to hold; indeed, the $\mathcal{O}(\epsilon^2)$ terms probably violate it. The first line of equation (28) *is* exact, however, and can be used to derive another scaling law, as we'll now show.

Turning on a small field right at the transition, we can again relate then the average velocities of the original and the rescaled systems using Eq. (24). However,$\delta a(\ell)$ now flows to 0 for large ℓ since the system is right at the critical point. Therefore, by choosing $\ell = \ln \left(1/H \right) / y_H$, we obtain the H-dependence of $\mathbf{v}$:

$$\langle \mathbf{v} \rangle (\delta a_0, b_0, \lambda_0, H) = H^{-\frac{\chi}{y_H}} \langle \mathbf{v} \rangle \left(0, b^*, \lambda^*, 1 \right) \ . \tag{29}$$

This implies $\delta = -\frac{y_H}{\chi}$. Combining this with the exact first equalities in Eqs (22,28), we obtain Widom's scaling relation

$$\gamma = \beta(\delta - 1), \tag{30}$$

Exponents	Incomp. active fluids	Heisenberg model [27]	Heisenberg with dipolar interactions [30]
η	0.35 ± 0.08	0.033 ± 0.004	0.023 ± 0.015
β	0.43 ± 0.03	0.3645 ± 0.0025	0.38 ± 0.02
δ	3.48 ± 0.03	4.803 ± 0.037	4.45 ± 0.04
γ	1.11 ± 0.01	1.386 ± 0.004	1.37 ± 0.02
ν	0.67 ± 0.04	0.705 ± 0.003	0.69 ± 0.02

Table 1. Comparisons between the critical exponents obtained in this work and other models in spatial dimension $d = 3$.

which is *exact*. Plugging the ϵ-expansions of γ and β into this relation, we find

$$\delta = 3 + \frac{51}{113}\epsilon + \mathcal{O}(\epsilon^2)\,. \tag{31}$$

5. Numerical estimation

We can estimate the numerical values of the exponents in spatial dimension $d = 3$ as follows: We first choose a scaling relation satisfied by any three exponents (e.g., Eq. (23) in the main text for η, β, and ν). We then determine numerical values for two of them (e.g., ν and β) by simply setting $\epsilon = 1$ in the ϵ-expansion for them, and dropping the unknown $\mathcal{O}(\epsilon^2)$ terms. We now get the value of the third exponent (e.g., η) by requiring that the scaling law (i.e., Eq. (23) in the main text) hold exactly in $d = 3$. In this example, this gives $\beta = 1/2 - 6/113 \approx 0.447$, $\nu = 1/2 + 29/226 \approx 0.628$, and $\eta = 2\beta/\nu - 1 \approx 0.424$. Next, we take η and ν to be given by their respective ϵ-expansions with $\epsilon = 1$, and get β from the exact scaling relation. This gives: $\nu = 1/2 + 29/226 \approx 0.628$, $\eta = 31/113 \approx 0.274$, and $\beta = \nu(1 + \eta)/2 \approx 0.400$. Finally, we take β and η from their ϵ-expansions, and get ν from the exact scaling relation, obtaining $\eta = 31/113 \approx 0.274$, $\beta = 1/2 - 6/113 \approx 0.447$, and $\nu = 2\beta/(1 + \eta) = 0.702$.

Note that each exponent gets two possible values in this approach: one from directly setting $\epsilon = 1$ in the ϵ-expansion, and another by obtaining the exponent from the exact scaling relation in $d = 3$.

Applying the same approach to Widom's exact scaling relation (i.e., Eq. (30) in the main text) and the three associated exponents γ, β, and δ gives the possible values: $\beta \approx 0.447$ or $\beta \approx 0.457$, $\delta \approx 3.451$ or $\delta \approx 3.503$, and $\gamma \approx 1.119$ or $\gamma \approx 1.096$.

So if we look at the range of values we've found for each of the exponents, we have $0.274 \leq \eta \leq 0.424$, $0.628 \leq \nu \leq 0.702$, $0.400 \leq \beta \leq 0.457$, $3.451 \leq \delta \leq 3.503$, and $1.096 \leq \gamma \leq 1.119$. Assuming, as seems reasonable (and as is true for, e.g., the critical exponents for the equilibrium $\mathcal{O}(n)$ model [26]), that the correct values lie within the range spanned by the different approaches we've used here, we can conclude that, in spatial dimension $d = 3$, the critical exponents are as shown in the second column of Table 1.

Comparing the critical exponents with the known values for the two equilibrium

analogs of our system: the three-dimensional, three component Heisenberg model (i.e., the $\mathcal{O}(3)$ model) with and without dipolar interactions (third and fourth columns respectively in Table 1), we see that ν and β are very close in all three models. The situation is a little better for γ and δ. The biggest difference, however, is clearly in η, which is much larger in the incompressible active fluid. Thus experiments to determine this exponent, which, as can be seen from equation (17) in the main text, can be deduced from velocity correlations right at the critical point, will provide the clearest and most dramatic evidence for the non-equilibrium nature of this system, and the novelty of its universality class.

The values of the exponents in $d = 2$ obviously can *not* be reliably estimated quantitatively from the $4 - \epsilon$-expansion. We do note, however, that the ordered state is expected to exist and to have true long-ranged order. This is clear since true long-ranged order exists even in the *compressible* problem, which obviously has more fluctuations than the incompressible problem we've studied here. Hence, we do *not* expect this problem to be like the *equilibrium* $2d$ XY model, in which [31] the ordered state only has quasi-long-ranged order (i.e., algebraically decaying correlations). We therefore do not expect 2d incompressible active fluids to exhibit any of the singular behavior of exponents found in the 2d equilibrium XY model; in particular, there is no reason to expect $\nu = \infty$. Beyond this, there is little we can say quantitatively about $d = 2$ beyond the expectation that the critical exponents β, ν, η, δ, and γ should be further from their mean field values $\beta = \nu = 1/2$, $\eta = 0$, $\delta = 3$, and $\gamma = 1$ than they are in $d = 3$. This implies that in $d = 2$, β will be smaller, and the four other exponents will be bigger, than the values quoted above for $d = 3$. We also note that the exact scaling relations Eqs. (23,30) in the main text will hold in $d = 2$, and that all of the exponents will be universal (i.e., the same for all incompressible active fluids) in $d = 2$.

6. Conclusion & Outlook

We have studied the order-disorder transition in incompressible active fluids using a dynamical $\epsilon = 4 - d$ expansion. This is the first study of the static to moving phase transition in active matter to go beyond mean-field theory, and include the effects of fluctuations on the transition. We find a stable non-Gaussian fixed point, which implies a continuous transition, whose critical exponents were calculated to $\mathcal{O}(\epsilon)$. This fixed point is new, and all of the critical exponents differ from those found for any previously known phase transition. Therefore, the universality class of this transition is new. In addition, we found that among the five critical exponents we calculated, there are two *exact* scaling relations which are the same as those in equilibrium ferromagnetic transitions. This is despite the fact that our system is fundamentally nonequilibrium, and that the universality class of its transition is new. Furthermore, we connected our model with two classic universality classes discovered in the early days of the renormalization group, namely the randomly stirred fluid model [24], and the dipolar ferromagnet model [27] (see Fig. 3). Specifically, we now know that the two fixed points associated to the two

classic universality classes are unstable in the combined model we considered here.

Our predictions on the critical exponents can be tested in motile systems (both experimental and simulated) with strong repulsive interactions between the particles, e.g., in bacterial suspensions like those studied in [19], provided that the system can be tuned to reach the critical point. The exponent β can be determined by measuring the average velocity. The exponents η and the correlation length (which determines ν) can be obtained by the velocity correlation functions. The exponents γ and δ can be obtained by measuring the response of the average velocity to an external perturbation. In magnetotactic bacteria; the perturbation can be a magnetic field [32], in chemotactic bacteria, it can be a nutrient gradient [33].

Future theoretical work [34] on this problem will include working out in quantitative detail the cutoff of the continuous transition by the banding instability in systems with a small, but non-zero, compressibility.

Acknowledgments

J.T. thanks Nicholas Guttenberg for explaining how to simulate systems with long ranged interactions. He also thanks the Department of Physics, University of California, Berkeley, CA; the Max Planck Institute for the Physics of Complex Systems (MPI-PKS), Dresden, Germany; the Aspen Center for Physics, Aspen, CO; the Isaac Newton Institute, Cambridge, UK; the Kavli Institute for Theoretical Physics, Santa Barbara, CA; and the Department of Bioengineering, Imperial College, London, UK for their hospitality while this work was underway. He also thanks the US NSF for support by awards #EF-1137815 and #1006171; and the Simons Foundation for support by award #225579. L.C. acknowledges support by the National Science Foundation of China (under Grant No. 11474354) and the Fundamental Research Funds for the Central Universities (under Grant No. 2013XK04). CFL thanks the MPI-PKS where the early stage of this work was performed.

References

[1] Reynolds C 1987 Flocks, Herds, and Schools: A Distributed Behavioral Model *Computer Graphics* **21**, 25; Deneubourg J L and Goss S 1989 Collective patterns and decision-making *Ethology, Ecology, Evolution* **1**, 295; Huth A and Wissel C 1990 The movement of fish schools: a simulation model, in *Biological Motion*, eds. Alt W and Hoffmann E (Springer Verlag) p. 577-590; Partridge B L 1982 The Structure and Function of Fish Schools *Scientific American*, 114-123

[2] Vicsek T, Czirók A, Ben-Jacob E, Cohen I, Shochet O 1995 Novel Type of Phase Transition in a System of Self-Driven Particles *Phys. Rev. Lett.* **75**, 1226; Czirok A, Stanley H E, and Vicsek T 1997 Spontaneously ordered motion of self-propelled particles *J. Phys. A* **30**, 1375

[3] Toner J and Tu Y-h 1995 Long-Range Order in a Two-Dimensional Dynamical *XY* Model: How Birds Fly Together *Phys. Rev. Lett.* **75**, 4326

[4] Tu Y-h, Toner J and Ulm M 1998 Sound Waves and the Absence of Galilean Invariance in Flocks *Phys. Rev. Lett.* **80**, 4819

[5] Toner J and Tu Y-h 1998 Flocks, herds, and schools: A quantitative theory of flocking *Phys. Rev. E* **58**, 4828

[6] Toner J, Tu Y-h and Ramaswamy S 2005 Hydrodynamics and phases of flocks *Ann. Phys.* **318**, 170

[7] Loomis W 1982 *The Development of Dictyostelium discoideum* (Academic, New York); Bonner J T 1967 *The Cellular Slime Molds* (Princeton University Press, Princeton, NJ)

[8] Rappel W J, Nicol A, Sarkissian A, Levine H and Loomis W F 1999 Self-organized vortex state in two-dimensional Dictyostelium dynamics *Phys. Rev. Lett.* **83**, 1247

[9] Voituriez R, Joanny J F and Prost J 2005 Spontaneous flow transition in active polar gels *Europhysics Letters* **70**, 404

[10] Kruse K, Joanny J F, Jülicher F, Prost J and Sekimoto K 2005 Generic theory of active polar gels: a paradigm for cytoskeletal dynamics *European Physical Journal E* **16** 5

[11] Mermin N D and Wagner H 1966 Absence of Ferromagnetism or Antiferromagnetism in One- or Two-Dimensional Isotropic Heisenberg Models *Phys. Rev. Lett.* **17**, 1133; Hohenberg P C 1967 Existence of Long-Range Order in One and Two Dimensions *Phys. Rev.* **158**, 383

[12] Bertin E, Droz M and Gregoire G 2006 Boltzmann and hydrodynamic description for self-propelled particles *Phys. Rev. E* **74**, 022101

[13] Bertin E, Droz M and Gregoire G 2009 Hydrodynamic equations for self-propelled particles: microscopic derivation and stability analysis *J. Phys. A: Math. Theor.* **42**, 445001

[14] Mishra S, Baskaran A and Marchetti M C 2010 Fluctuations and pattern formation in self-propelled particles *Phys. Rev. E* **81**, 061916

[15] Farrell F D C, Marchetti M C, Marenduzzo D and Tailleur J 2012 Pattern Formation in Self-Propelled Particles with Density-Dependent Motility *Phys. Rev. Lett.* **108**, 248101

[16] Yamanaka S and Ohta T 2014 Formation and collision of traveling bands in interacting deformable self-propelled particles *Phys. Rev. E* **89**, 012918

[17] Bialké J, Löwen H and Speck T 2013 Microscopic theory for the phase separation of self-propelled repulsive disks *EPL* **103**, 30008

[18] Ihle T 2013 Invasion-wave-induced first-order phase transition in systems of active particles *Phys. Rev. E.* **88**, 040303

[19] Wensink H H *et al.* 2012 Meso-scale turbulence in living fluids *Proc. Nat. Acad. Sci.* **109**, 14308

[20] Pearce D J G, Miller A M, Rowlands G and Turner M S 2014 Role of projection in the control of bird flocks *Proc. Nat. Acad. Sci.* **111**, 10422

[21] Bricard A, Caussin J-B, Desreumaux N, Dauchot O and Bartolo D 2013 Emergence of macroscopic directed motion in populations of motile colloids Nature **503**, 95

[22] Zöttl A and Stark H 2014 Hydrodynamics Determines Collective Motion and Phase Behavior of Active Colloids in Quasi-Two-Dimensional Confinement *Phys. Rev. Lett.* **112**, 118101

[23] This equation of motion is simply what one would obtain from the equation of motion of references [3, 4, 5, 6] by expanding for small $\mathbf{v}$ and dropping all density-dependent terms.

[24] See, e.g., Forster D, Nelson D R and Stephen M J 1977 Large-distance and long-time properties of a randomly stirred fluid *Phys. Rev. A* **16**, 732

[25] Ma S-K 2000 *Modern Theory of Critical Phenomena* (Westview Press)

[26] Chaikin P M and Lubensky T C 1995 *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, U.K.)

[27] Fisher M E and Aharony A 1973 Dipolar Interactions at Ferromagnetic Critical Points *Phys. Rev. Lett.* **30**, 559; Aharony A and Fisher M E 1973 Critical Behavior of Magnets with Dipolar Interactions. I. Renormalization Group near Four Dimensions *Phys. Rev. B* **8**, 3323; Aharony A 1973 Critical Behavior of Magnets with Dipolar Interactions. II. Feynman-Graph Expansion for Ferromagnets near Four Dimensions *Phys. Rev. B* **8**, 3342; Bruce A D and Aharony A 1974 Critical exponents of ferromagnets with dipolar interactions: Second-order ϵ expansion *Phys. Rev. B* **10**, 2078

[28] The connection between our problem and the dipolar magnet is that the long-ranged dipolar interaction in magnetic systems couples to, and therefore suppresses, the longitudinal component of the magnetization. See [27] for more details.

[29] Supplemental Material.
[30] Guida R and Zinn-Justin J 1998 Critical exponents of the N-vector model *J. Phys. A* **31**, 8103 (1998); Zinn-Justin J 2001 Precise determination of critical exponents and equation of state by field theory methods Phys. Rep. **344**, 159
[31] Kosterlitz J M and Thouless D J 1973 Ordering, metastability and phase transitions in two-dimensional systems *J. Phys. C* **6**, 1181; Kosterlitz J M 1974 The critical properties of the two-dimensional xy model *J. Phys. C* **7**, 1046; Kosterlitz J M and Thouless D J 1978 Two-Dimensional Physics, in *Progress in Low Temperature Physics*, edited by D. F. Brewer (North-Holland, Amsterdam), Vol. VII-B
[32] Blakemore R 1975 Magnetotactic bacteria *Science* **190**, 377
[33] Adler J 1966 Chemotaxis in Bacteria Science **153**, 708
[34] Chen L, Toner J and Lee C F, unpublished.

Critical Phenomenon of the Order-Disorder Transition in Incompressible Flocks: Supplemental Material

Leiming Chen*
College of Science, China University of Mining and Technology, Xuzhou Jiangsu, 221116, P. R. China

John Toner†
Department of Physics and Institute of Theoretical Science, University of Oregon, Eugene, OR 97403 *and Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Str. 38, 01187 Dresden, Germany*

Chiu Fan Lee‡
Department of Bioengineering, Imperial College London, South Kensington Campus, London SW7 2AZ, U.K.
(Dated: February 2, 2015)

The various constituent elements of Feynman diagrams are illustrated in Fig. 1 in the main text. The one loop graphical corrections to various coefficients in the model Eq. (1) in the main text are illustrated in Fig. 2 in the main text. Since we are only interested in the RG flows near the critical point, we have evaluated all of these graphs at $a = 0$, except for the graph (a) for a, the a-dependence of which we need to determine the "thermal" eigenvalue y_a. The corrections from graphs (a), (b), (c), and (d) are given respectively by:

$$\delta a = \frac{(d-1)(d+2)}{d}\frac{S_d}{(2\pi)^d}\frac{Db}{\mu\Lambda^2+a}d\ell, \tag{1}$$

$$\delta b = -\left[d+1+\frac{6(d^2-2)}{d(d+2)}\right]\frac{S_d}{(2\pi)^d}\frac{b^2D}{\mu^2}\Lambda^{d-4}d\ell, \tag{2}$$

$$\delta\lambda = -\frac{2(d^2-2)}{d(d+2)}\frac{S_d}{(2\pi)^d}\frac{Db\lambda}{\mu^2}\Lambda^{d-4}d\ell - \frac{(d-2)}{d}\frac{S_d}{(2\pi)^d}\frac{Db\lambda}{\mu^2}\Lambda^{d-4}d\ell. \tag{3}$$

$$\delta\mu = \frac{(d-2)}{2d}\frac{S_d}{(2\pi)^d}\frac{D\lambda^2}{\mu^2}\Lambda^{d-4}d\ell, \tag{4}$$

Combining these corrections (including the zero correction to D) with the rescalings described in the main text, and setting dimension $d = 4$ in all of these expressions (which is sufficient to obtain results to first order in $\epsilon = 4-d$), leads to the recursion relations (4–8) in the main text, (although in (4) we have in addition expanded to linear order in a, which is sufficient to determine the exponents to $\mathcal{O}(\epsilon)$).

*Electronic address: leimingyandongyu@gmail.com
†Electronic address: jjt@uoregon.edu
‡Electronic address: c.lee@imperial.ac.uk

I. EVALUATION OF THE FEYNMAN DIAGRAMS

A. Graph (a)

This graph represents an additional contribution $\delta\left(\partial_t v_l\right)$ to $\partial_t v_l$ given by:

$$\begin{aligned}
\delta\left(\partial_t v_l\right) &= -bQ_{lmnp}(\mathbf{k})v_p(\tilde{\mathbf{k}})\int_{\mathbf{q},\Omega}\frac{2DP_{mn}(\mathbf{q})}{(\mu q^2+a)^2+\Omega^2} \\
&= -bDQ_{lmnp}(\mathbf{k})v_p(\tilde{\mathbf{k}})\int_{\mathbf{q}}\frac{P_{mn}(\mathbf{q})}{\mu q^2+a} \\
&= -bDQ_{lmnp}(\mathbf{k})v_p(\tilde{\mathbf{k}})\int_{\mathbf{q}}\frac{\delta_{mn}(1-\frac{1}{d})}{\mu q^2+a} \\
&= -(d+2)bDv_l(\tilde{\mathbf{k}})\left(1-\frac{1}{d}\right)\int_{\mathbf{q}}\frac{1}{\mu q^2+a} \\
&= -(d+2)bDv_l(\tilde{\mathbf{k}})\left(1-\frac{1}{d}\right)\frac{S_d}{(2\pi)^d}\frac{\Lambda^d d\ell}{\mu\Lambda^2+a}\,,
\end{aligned} \tag{5}$$

where $P_{mn}(\mathbf{q})=\delta_{mn}-q_m q_n/q^2$ is the transverse projection operator, and Q_{lmnp} is defined in the main text after Eq. (3). In going from the second to the third line above, we have used the well-known identity $\langle P_{mn}(\mathbf{q})\rangle_{\hat{\mathbf{q}}}=\left(1-\frac{1}{d}\right)\delta_{mn}$, where $\langle\rangle_{\hat{\mathbf{q}}}$ denotes the average over directions $\hat{\mathbf{q}}$ of $\mathbf{q}$ for fixed $|\mathbf{q}|$. This identity is derived in part (II) of these Supplemental Materials. Clearly this correction to $\partial_t v_l(\mathbf{k},\omega)$ is exactly what one would obtain by adding to the parameter a (hiding in the "propagator") in Eq. (3) in the main text a correction δa given by the coefficient of v_l in (5); i.e., Eq. (1).

B. Graph (b)

This graph represents an additional contribution $\delta\left(\partial_t v_l\right)$ to $\partial_t v_l$ given by:

$$\delta\left(\partial_t v_l\right)=18\left(\frac{b}{3}\right)^2 2DQ_{lmno}(\mathbf{k})\int_{\tilde{\mathbf{p}},\tilde{\mathbf{h}}}v_j(\tilde{\mathbf{p}})v_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})v_s(\tilde{\mathbf{h}}-\tilde{\mathbf{p}})\int_{\mathbf{q},\Omega}\frac{P_{mi}(\mathbf{q})Q_{nijs}(\mathbf{k}-\mathbf{q})}{(\mu^2q^4+\Omega^2)[\mu|\mathbf{h}-\mathbf{q}|^2-\mathrm{i}(\omega_h-\Omega)]}\,. \tag{6}$$

where the combinatoric prefactor of 18 arises because there are 3 ways to pick the leg with index o on the left. Then, once this choice has been made, there are 2 ways to pick the leg with index m on the left, and 3 ways to pick the one with index i on the right.

To lowest order in the external momenta ($\mathbf{k}$, $\mathbf{p}$ and $\mathbf{h}$) and frequencies ω, ω_p, and ω_h, we can set all of these external momenta and frequencies equal to zero in the integrand of the integral over $\tilde{\mathbf{q}}$ in Eq. (6). This proves, as we shall see, to make this graph into a renormalization of the cubic non-linearity b. Making this simplification, Eq. (6) becomes, after using

$$\frac{1}{\mu q^2+\mathrm{i}\Omega}=\frac{\mu q^2-\mathrm{i}\Omega}{(\mu^2q^4+\Omega^2)}\,, \tag{7}$$

and dropping an integral of an odd function of Ω,

$$\delta\left(\partial_t v_l\right)=4b^2DQ_{lmno}(\mathbf{k})\int_{\tilde{\mathbf{p}},\tilde{\mathbf{h}}}v_j(\tilde{\mathbf{p}})v_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})v_s(\tilde{\mathbf{h}}-\tilde{\mathbf{p}})\int_{\mathbf{q},\Omega}\frac{\mu q^2P_{mi}(\mathbf{q})Q_{nijs}(\mathbf{q})}{(\mu^2q^4+\Omega^2)^2}\,. \tag{8}$$

Using the definition of $Q_{nijs}(\mathbf{q})$, and contracting indices, we obtain

$$P_{mi}(\mathbf{q})Q_{nijs}(\mathbf{q})=P_{mn}(\mathbf{q})\delta_{js}+P_{ms}(\mathbf{q})P_{jn}(\mathbf{q})+P_{jm}(\mathbf{q})P_{ns}(\mathbf{q})\,. \tag{9}$$

As we've done repeatedly throughout these Supplemental Materials, we'll replace this tensor with its average over all directions of $\hat{\mathbf{q}}$. This is easily obtained from the known direction averages (39) and (53), and gives, after a little algebra,

$$\langle P_{mi}(\mathbf{q})Q_{nijs}(\mathbf{q})\rangle_{\hat{\mathbf{q}}}=\frac{d+1}{d+2}\delta_{mn}\delta_{js}+\frac{d^2-2}{d(d+2)}\left(\delta_{ms}\delta_{jn}+\delta_{jm}\delta_{ns}\right)\,. \tag{10}$$

Inserting this into our earlier expression for $\delta\left(\partial_t v_l\right)$ and performing a few tensor index contractions gives

$$\begin{aligned}\delta\left(\partial_t v_l\right) &= 4b^2 D\left(\frac{d+1}{d+2}\delta_{js}Q_{lmmo}(\mathbf{k}) + \frac{d^2-2}{d(d+2)}\left(Q_{lsjo}(\mathbf{k}) + Q_{ljso}(\mathbf{k})\right)\right) \\ &\times \int_{\tilde{\mathbf{p}},\tilde{\mathbf{h}}} v_j(\tilde{\mathbf{p}})v_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})v_s(\tilde{\mathbf{h}}-\tilde{\mathbf{p}}) \int_{\mathbf{q},\Omega} \frac{\mu q^2}{(\mu^2 q^4+\Omega^2)^2}\,. \end{aligned} \tag{11}$$

The trace in this expression can be evaluated as

$$Q_{lmmo}(\mathbf{k}) = P_{lm}(\mathbf{k})\delta_{mo} + P_{lm}(\mathbf{k})\delta_{mo} + P_{lo}(\mathbf{k})\delta_{mm} = (d+2)P_{lo}(\mathbf{k})\,, \tag{12}$$

and the integral over frequency Ω and $\mathbf{q}$ is readily evaluated, and is given by

$$\int_{\mathbf{q},\Omega} \frac{\mu q^2}{(\mu^2 q^4+\Omega^2)^2} = \frac{1}{4\mu^2}\frac{S_d}{(2\pi)^d}\Lambda^{d-4}d\ell\,. \tag{13}$$

Putting (12) and (13) into (11), and taking advantage of the complete symmetry of $\int_{\tilde{\mathbf{p}},\tilde{\mathbf{h}}} v_j(\tilde{\mathbf{p}})v_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})v_s(\tilde{\mathbf{h}}-\tilde{\mathbf{p}})$ under interchanges of the indices j, o, and s to symmetrize the tensor prefactor gives

$$\delta\left(\partial_t v_l\right) = \frac{b^2 D}{3\mu^2}\left(d+1+6\frac{(d^2-2)}{d(d+2)}\right)\frac{S_d}{(2\pi)^d}\Lambda^{d-4}d\ell Q_{ljso}(\mathbf{k})\int_{\tilde{\mathbf{p}},\tilde{\mathbf{h}}} v_j(\tilde{\mathbf{p}})v_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})v_s(\tilde{\mathbf{h}}-\tilde{\mathbf{p}})\,. \tag{14}$$

This is readily recognized as a contribution to $-\frac{b}{3}$ term in Eq. (3) in the main text; hence, the correction of b is given by (2).

C. The First Graph in (c)

This graph represents an additional contribution $\delta\left(\partial_t v_l\right)$ to $\partial_t v_l$ given by:

$$\delta\left(\partial_t v_l\right) = 6\left(-\frac{i\lambda}{2}\right)\left(-\frac{b}{3}\right)P_{lmn}(\mathbf{k})\int_{\tilde{\mathbf{h}}} v_j(\tilde{\mathbf{h}})v_s(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})\int_{\mathbf{q},\Omega} Q_{nijs}(\mathbf{k}-\mathbf{q})\frac{2DP_{mi}(\mathbf{q})}{(\mu^2 q^4+\Omega^2)[\mu|\mathbf{k}-\mathbf{q}|^2 - \mathrm{i}(\omega-\Omega)]}\,, \tag{15}$$

where $P_{\ell mn}(\mathbf{k})$ is defined in the main text after Eq. (3), and the combinatoric prefactor of 6 arises because there are 2 ways to pick the leg with index m on the left, and 3 ways to pick the one with index i on the right.

The piece of $\delta\left(\partial_t v_l\right)$ linear in the external momentum $\mathbf{k}$ is immediately recognized as a contribution to $-\frac{i}{2}\lambda$ term in Eq. (3) in the main text. Since there is already an implicit factor of $\mathbf{k}$ in the $P_{lmn}(\mathbf{k})$ in this expression, we can evaluate this graph to linear order in $\mathbf{k}$ by setting both $\mathbf{k}$ and the external frequency ω to zero in the integrand of the integral over $\tilde{\mathbf{q}}$. Doing so, and in addition using (7) gives, after dropping an integral of an odd function of Ω that vanishes,

$$\delta\left(\partial_t v_l\right) = 2i\lambda b D P_{lmn}(\mathbf{k})\int_{\tilde{\mathbf{h}}} v_j(\tilde{\mathbf{h}})v_s(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})\int_{\mathbf{q},\Omega}\frac{\mu q^2 P_{mi}(\mathbf{q})Q_{nijs}(\mathbf{q})}{(\mu^2 q^4+\Omega^2)^2}\,, \tag{16}$$

where we've used the fact that $Q_{nijs}(\mathbf{q})$ is an even function of $\mathbf{q}$. Using the definition of $Q_{nijs}(\mathbf{q})$, and performing the tensor index contractions, we can simplify the numerator of the integrand as follows:

$$P_{mi}(\mathbf{q})Q_{nijs}(\mathbf{q}) = P_{mn}\delta_{js} + P_{jn}P_{ms} + P_{ns}P_{mj}\,, \tag{17}$$

where we've also used the fact that, e.g., $P_{mi}(\mathbf{q})P_{in}(\mathbf{q}) = P_{mn}(\mathbf{q})$ (which is a consequence of the definition of $P_{mi}(\mathbf{q})$ as a projection operator). Using this result (17), and taking the angle average of the numerator of (16) (which is the only factor in the integral that depends on the direction of $\mathbf{q}$) gives,

$$\left\langle P_{mi}(\mathbf{q})Q_{nijs}(\mathbf{q})\right\rangle_{\hat{\mathbf{q}}} = \left(\frac{d+1}{d+2}\right)\delta_{mn}\delta_{js} + \left(\frac{d^2-2}{d(d+2)}\right)\left(\delta_{jn}\delta_{ms} + \delta_{ns}\delta_{mj}\right)\,. \tag{18}$$

In deriving this expression, we've made liberal use of the angle averages (53) and (39).

The first term on the right hand side of (18) contributes nothing, since it contracts two of the indices on the prefactor $P_{lmn}(\mathbf{k})$ together, which gives zero, as can be seen from the definition of $P_{lmn}(\mathbf{k})$:

$$P_{lmm}(\mathbf{k}) = P_{lm}(\mathbf{k})k_m + P_{lm}(\mathbf{k})k_m = 0\,. \tag{19}$$

Keeping only the second term gives

$$\begin{aligned}\delta\left(\partial_t v_l\right) &= 2i\lambda bD\left(\frac{d^2-2}{d(d+2)}\right)P_{lmn}(\mathbf{k})\left(\delta_{jn}\delta_{ms}+\delta_{ns}\delta_{mj}\right)v_j(\tilde{\mathbf{h}})v_s(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})\int_{\mathbf{q},\Omega}\frac{\mu q^2}{[\Omega^2+\mu^2q^4]^2}\\ &= \frac{i\lambda bDS_d\Lambda^{d-4}d\ell}{2\mu^2}\left(\frac{d^2-2}{d(d+2)}\right)P_{lmn}(\mathbf{k})\left[v_n(\tilde{\mathbf{h}})v_m(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})+v_m(\tilde{\mathbf{h}})v_n(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})\right]\\ &= \frac{i\lambda bDS_d\Lambda^{d-4}d\ell}{\mu^2}\left(\frac{d^2-2}{d(d+2)}\right)P_{lmn}(\mathbf{k})v_m(\tilde{\mathbf{h}})v_n(\tilde{\mathbf{k}}-\tilde{\mathbf{h}}),\end{aligned} \tag{20}$$

where in the last step we have used the symmetry of $P_{lmn}(\mathbf{k})$ under interchange of its last two indices.

This is immediately recognized as a contribution to $-\frac{i}{2}\lambda$ term in Eq. (3) in the main text, which implies a correction to λ given by

$$\delta\lambda = \frac{-2\lambda bDS_d\Lambda^{d-4}d\ell}{\mu^2}\left(\frac{d^2-2}{d(d+2)}\right), \tag{21}$$

which is just the first term on the RHS of Eq. (3).

D. The Second Graph in (c)

This graph represents an additional contribution $\delta\left(\partial_t v_l\right)$ to $\partial_t v_l$ given by

$$\delta\left(\partial_t v_l\right) = 12\left(\frac{-\mathrm{i}\lambda}{2}\right)\left(-\frac{b}{3}\right)2DQ_{lmno}(\mathbf{k})\int_{\tilde{\mathbf{h}}}v_j(\tilde{\mathbf{h}})v_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})\int_{\mathbf{q},\Omega}\frac{P_{mi}(\mathbf{q})P_{nij}(\mathbf{h}-\mathbf{q})}{(\mu^2q^4+\Omega^2)[\mu|\mathbf{h}-\mathbf{q}|^2-\mathrm{i}(\omega-\Omega)]}. \tag{22}$$

where the combinatoric prefactor of 12 arises because there are 3 ways to pick the leg with index o on the left. Then, once this choice has been made, there are 2 ways to pick the leg with index m on the left, and 2 ways to pick the one with index i on the right.

To extract from (22) the contribution to $-\frac{i}{2}\lambda$ term in Eq. (3) in the main text, we need the piece of $\delta\left(\partial_t v_l\right)$ that is linear in either the external momentum $\mathbf{k}$ or $\mathbf{h}$. We notice that the integral over $\tilde{\mathbf{q}}$ vanishes if we set $\mathbf{h}=0$ in its integrand. This implies the entire term is at least of order $\mathbf{h}$. Thus, to obtain a correction to λ we can simply set the external frequency $\omega=0$ in the integrand; doing so, and integrating over Ω, we obtain

$$\delta\left(\partial_t v_l\right) = 2\mathrm{i}\lambda bDQ_{lmno}(\mathbf{k})\int_{\tilde{\mathbf{h}}}v_j(\tilde{\mathbf{h}})v_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})\int_{\mathbf{q}}\frac{P_{mi}(\mathbf{q})}{\mu q^2[\mu q^2+\mu|\mathbf{h}-\mathbf{q}|^2]}\left[P_{ni}(\mathbf{h}-\mathbf{q})(h_j-q_j)+P_{nj}(\mathbf{h}-\mathbf{q})(h_i-q_i)\right]. \tag{23}$$

This can be rewritten as

$$\delta\left(\partial_t v_l\right) = 2\mathrm{i}\lambda bDQ_{lmno}(\mathbf{k})\int_{\tilde{\mathbf{h}}}v_j(\tilde{\mathbf{h}})v_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})\left(I^{(2)}_{jmn}(\mathbf{h})+I^{(3)}_{jmn}(\mathbf{h})+I^{(4)}_{jmn}(\mathbf{h})+I^{(5)}_{jmn}(\mathbf{h})\right) \tag{24}$$

where $I^{(2)}_{jmn}(\mathbf{h})$ is given by equation (48) (with the obvious substitution $\mathbf{k}\to\mathbf{h}$), and the other integrals are defined as:

$$I^{(3)}_{jmn}(\mathbf{h}) \equiv h_j\int_{\mathbf{q}}\frac{P_{mi}(\mathbf{q})P_{ni}(\mathbf{h}-\mathbf{q})}{\mu q^2[\mu q^2+\mu|\mathbf{h}-\mathbf{q}|^2]}, \tag{25}$$

$$I^{(4)}_{jmn}(\mathbf{h}) \equiv h_i\int_{\mathbf{q}}\frac{P_{mi}(\mathbf{q})P_{nj}(\mathbf{h}-\mathbf{q})}{\mu q^2[\mu q^2+\mu|\mathbf{h}-\mathbf{q}|^2]}, \tag{26}$$

$$I^{(5)}_{jmn}(\mathbf{h}) \equiv -\int_{\mathbf{q}}\frac{P_{mi}(\mathbf{q})P_{nj}(\mathbf{h}-\mathbf{q})q_i}{\mu q^2[\mu q^2+\mu|\mathbf{h}-\mathbf{q}|^2]}. \tag{27}$$

Immediately, by the properties of the projection operator we get

$$I^{(5)}_{jmn}(\mathbf{h}) = 0. \tag{28}$$

We notice that both $I^{(3)}_{jmn}(\mathbf{h})$ and $I^{(4)}_{jmn}(\mathbf{h})$ are already proportional to $\mathbf{h}$, so we can simply set $\mathbf{h}=\mathbf{0}$ inside the integral. Thus, we obtain

$$I^{(3)}_{jmn}(\mathbf{h}) = h_j\int_{\mathbf{q}}\frac{P_{mi}(\mathbf{q})P_{ni}(\mathbf{q})}{2\mu^2q^4} = h_j\int_{\mathbf{q}}\frac{P_{mn}(\mathbf{q})}{2\mu^2q^4} = h_j\int_{\mathbf{q}}\frac{\langle P_{mn}(\mathbf{q})\rangle_{\hat{\mathbf{q}}}}{2\mu^2q^4} = \frac{d-1}{2d}\frac{S_d}{(2\pi)^d}\mu^{-2}\Lambda^{d-4}d\ell\, h_j\delta_{mn}, \tag{29}$$

$$\begin{aligned} I^{(4)}_{jmn}(\mathbf{h}) &= h_i\int_{\mathbf{q}}\frac{P_{mi}(\mathbf{q})P_{nj}(\mathbf{q})}{2\mu^2q^4} = h_i\int_{\mathbf{q}}\frac{\langle P_{mi}(\mathbf{q})P_{nj}(\mathbf{q})\rangle_{\hat{\mathbf{q}}}}{2\mu^2q^4} \\ &= \frac{1}{2d(d+2)}\frac{S_d}{(2\pi)^d}\mu^{-2}\Lambda^{d-4}d\ell\,\left[\left(d^2-3\right)\delta_{nj}h_m+\delta_{mn}h_j+\delta_{mj}h_n\right]. \end{aligned} \tag{30}$$

Plugging the values (48), (29), (30), and (28) of the various integrals into Eq. (24), we obtain

$$\delta\left(\partial_t v_l\right) = \mathrm{i}\frac{\lambda bD}{2\mu^2}\frac{S_d}{(2\pi)^d}\Lambda^{d-4}d\ell Q_{lmno}(\mathbf{k})\int_{\tilde{\mathbf{h}}}v_j(\tilde{\mathbf{h}})v_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})\left[\frac{1}{d}\delta_{mj}h_n+\frac{2d^2-d-10}{d(d+2)}\delta_{nj}h_m+\frac{d+1}{d}\delta_{mn}h_j\right]. \tag{31}$$

The third piece vanishes due to the incompressibility condition $h_jv_j(\tilde{\mathbf{h}})=0$. The first and the second pieces can be grouped together since Q_{lmno} is invariant under the interchange of m and n. Therefore, the above expression can be simplified as

$$\begin{aligned} \delta\left(\partial_t v_l\right) &= \mathrm{i}\frac{(d-2)}{d}\frac{\lambda bD}{\mu^2}\frac{S_d}{(2\pi)^d}\Lambda^{d-4}d\ell Q_{lmno}(\mathbf{k})\int_{\tilde{\mathbf{h}}}v_j(\tilde{\mathbf{h}})v_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})\delta_{mj}h_n \\ &= \mathrm{i}\frac{(d-2)}{d}\frac{\lambda bD}{\mu^2}\frac{S_d}{(2\pi)^d}\Lambda^{d-4}d\ell\int_{\tilde{\mathbf{h}}}\left[P_{lj}(\mathbf{k})h_o+P_{\ell n}(\mathbf{k})h_n\delta_{jo}+P_{\ell o}(\mathbf{k})h_j\right]v_j(\tilde{\mathbf{h}})v_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}}) \\ &= \mathrm{i}\frac{(d-2)}{d}\frac{\lambda bD}{\mu^2}\frac{S_d}{(2\pi)^d}\Lambda^{d-4}d\ell\int_{\tilde{\mathbf{h}}}P_{lj}(\mathbf{k})h_ov_j(\tilde{\mathbf{h}})v_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}}) \\ &= \mathrm{i}\frac{(d-2)}{d}\frac{\lambda bD}{\mu^2}\frac{S_d}{(2\pi)^d}\Lambda^{d-4}d\ell\int_{\tilde{\mathbf{h}}}P_{lj}(\mathbf{k})k_ov_j(\tilde{\mathbf{h}})v_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}}) \\ &= \mathrm{i}\frac{d-2}{2d}\frac{\lambda bD}{\mu^2}\frac{S_d}{(2\pi)^d}\Lambda^{d-4}d\ell\int_{\tilde{\mathbf{h}}}P_{ljo}(\mathbf{k})v_j(\tilde{\mathbf{h}})v_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}}), \end{aligned} \tag{32}$$

where in the second equality we have dropped the second and third pieces. The second piece can be dropped because it vanishes, as can be seen by simply changing variables of integration from $\mathbf{h}$ to $\mathbf{k}-\mathbf{h}$; this gives

$$\int_{\tilde{\mathbf{h}}}P_{\ell n}(\mathbf{k})h_nv_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})v_o(\mathbf{h}) = \int_{\tilde{\mathbf{h}}}P_{\ell n}(\mathbf{k})(k_n-h_n)v_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})v_o(\tilde{\mathbf{h}})\,.$$

Adding the left and the right hand side of this equation, and dividing by 2, implies

$$\begin{aligned} \int_{\tilde{\mathbf{h}}}P_{\ell n}(\mathbf{k})h_nv_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})v_o(\tilde{\mathbf{h}}) &= \frac{1}{2}\left(\int_{\tilde{\mathbf{h}}}P_{\ell n}(\mathbf{k})h_nv_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})v_o(\tilde{\mathbf{h}})+\int_{\tilde{\mathbf{h}}}P_{\ell n}(\mathbf{k})(k_n-h_n)v_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})v_o(\tilde{\mathbf{h}})\right) \\ &= \frac{1}{2}\int_{\tilde{\mathbf{h}}}P_{\ell n}(\mathbf{k})k_nv_o(\tilde{\mathbf{k}}-\tilde{\mathbf{h}})v_o(\tilde{\mathbf{h}}) = 0\,, \end{aligned} \tag{33}$$

with the last equality following from $P_{\ell n}(\mathbf{k})k_n=0$. The third piece can be dropped due to the incompressibility condition $h_jv_j(\tilde{\mathbf{h}})=0$. The remaining piece of (32) is readily recognized as a contribution to $-\frac{i}{2}\lambda$ term in Eq. (3) in the main text. This implies a correction to λ given by the second piece on the RHS of Eq. (3).

E. Graph (d)

This graph represents an additional contribution $\delta\left(\partial_t v_l\right)$ to $\partial_t v_l$ given by:

$$\begin{aligned} \delta\left(\partial_t v_l\right) &= -\lambda^2P_{lmn}(\mathbf{k})v_j(\mathbf{k})\int_{\mathbf{q},\Omega}\frac{2DP_{mi}(\mathbf{q})P_{nij}(\mathbf{k}-\mathbf{q})}{(\mu^2q^4+\Omega^2)[\mu|\mathbf{k}-\mathbf{q}|^2-\mathrm{i}(\omega-\Omega)]} \\ &= -D\lambda^2P_{lmn}(\mathbf{k})v_j(\mathbf{k})\int_{\mathbf{q}}\frac{P_{mi}(\mathbf{q})P_{nij}(\mathbf{k}-\mathbf{q})}{\mu q^2[\mu|\mathbf{k}-\mathbf{q}|^2-\mathrm{i}\omega+\mu q^2]}\,. \end{aligned} \tag{34}$$

The integral in this expression is readily seen to vanish when $\mathbf{k} \to \mathbf{0}$, since the integrand then becomes odd in $\mathbf{q}$. Hence, the integral is at least of order $\mathbf{k}$, so the entire term (include the implicit first power of $\mathbf{k}$ coming from the $P_{lmn}(\mathbf{k})$ in front), is $\mathcal{O}(k^2)$. Since we do not need to keep any terms in the equations of motion higher order in $\mathbf{k}$ and ω than $\mathcal{O}(k^2)$, this means that we can safely set $\omega = 0$ inside the integral. Keeping just the $\mathcal{O}(k)$ piece of the integral then gives us a modification to the equation of motion of $\mathcal{O}(k^2\mathbf{v})$, which is clearly a renormalization of the diffusion constant μ. So setting $\omega = 0$ in the integrand for the reasons just discussed, and then writing $P_{nij}(\mathbf{k}-\mathbf{q})$ using its definition as given in the main text after Eq. (3); this gives for the integral in (34):

$$\int_{\mathbf{q}} \frac{P_{im}(\mathbf{q})P_{nij}(\mathbf{k}-\mathbf{q})}{\mu q^2[\mu|\mathbf{k}-\mathbf{q}|^2 - \mathrm{i}\omega + \mu q^2]} = \int_{\mathbf{q}} \frac{P_{mi}(\mathbf{q})\left[P_{jn}(\mathbf{q}-\mathbf{k})(k_i - q_i) + P_{in}(\mathbf{q}-\mathbf{k})(k_j - q_j)\right]}{\mu q^2[\mu|\mathbf{k}-\mathbf{q}|^2 + \mu q^2]}\,. \tag{35}$$

The term proportional to k_j in this expression can be dropped, since $k_j v_j = 0$ (this is just the incompressibility condition $\nabla \cdot \mathbf{v} = 0$ written in Fourier space). The term proportional to q_i can be dropped since $P_{mi}(\mathbf{q})q_i = 0$ by the properties of the transverse projection operator $P_{mi}(\mathbf{q})$. This leaves two terms in the integral, which can be written as

$$I^{(1)}_{jmn}(\mathbf{k}) \equiv k_i \int_{\mathbf{q}} \frac{P_{mi}(\mathbf{q})P_{jn}(\mathbf{q}-\mathbf{k})}{\mu q^2[\mu|\mathbf{k}-\mathbf{q}|^2 + \mu q^2]}\,, \tag{36}$$

and

$$I^{(2)}_{jmn}(\mathbf{k}) \equiv -\int_{\mathbf{q}} \frac{P_{mi}(\mathbf{q})P_{ni}(\mathbf{k}-\mathbf{q})q_j}{\mu q^2[\mu q^2 + \mu|\mathbf{k}-\mathbf{q}|^2]}\,. \tag{37}$$

Since $I^{(1)}_{jmn}(\mathbf{k})$ already has an explicit factor of k in front, we can evaluate it to linear order in k by setting $\mathbf{k} = \mathbf{0}$ inside the integral. Doing so gives

$$I^{(1)}_{jmn}(\mathbf{k}) = \frac{1}{2\mu^2} k_i \int_{\mathbf{q}} \frac{P_{mi}(\mathbf{q})P_{jn}(\mathbf{q})}{q^4}\,, \tag{38}$$

The integral in this expression can now be evaluated by replacing the only piece that depends on the direction $\hat{\mathbf{q}}$ of $\mathbf{q}$, namely, the factor $P_{mi}(\mathbf{q})P_{jn}(\mathbf{q})$, with its angle average. As shown in (II), this average is given by

$$\langle P_{mi}(\mathbf{q})P_{jn}(\mathbf{q})\rangle_{\hat{\mathbf{q}}} = \left(\frac{d^2-3}{d(d+2)}\right)\delta_{mi}\delta_{jn} + \frac{1}{d(d+2)}(\delta_{mn}\delta_{ij} + \delta_{mj}\delta_{ni})\,. \tag{39}$$

Inserting this into (38) gives

$$I^{(1)}_{jmn}(\mathbf{k}) = \frac{1}{2\mu^2}\left[\left(\frac{d^2-3}{d(d+2)}\right)k_m\delta_{jn} + \frac{1}{d(d+2)}\left(k_j\delta_{mn} + k_n\delta_{mj}\right)\right]\frac{S_d}{(2\pi)^d}\Lambda^{d-4}d\ell\,. \tag{40}$$

Now let us expand $I^{(2)}_{jmn}(\mathbf{k})$ to linear order in k. Changing variables of integration from $\mathbf{q}$ to a shifted variable $\mathbf{p}$ defined by:

$$\mathbf{q} = \mathbf{p} + \frac{\mathbf{k}}{2} \tag{41}$$

gives

$$\begin{aligned} I^{(2)}_{jmn}(\mathbf{k}) = \quad & -\int_{\mathbf{p}} \frac{P_{mi}(\mathbf{p}_+)P_{ni}(\mathbf{p}_-)p_j}{\Gamma(\mathbf{p}_+)[\Gamma(\mathbf{p}_+)+\Gamma(\mathbf{p}_-)]} - \frac{k_j}{2}\int_{\mathbf{p}} \frac{P_{mi}(\mathbf{p}_+)P_{ni}(\mathbf{p}_-)}{\Gamma(\mathbf{p}_+)[\Gamma(\mathbf{p}_+)+\Gamma(\mathbf{p}_-)]} \\ \equiv\ & I^{(2.1)}_{jmn}(\mathbf{k}) + I^{(2.2)}_{jmn}(\mathbf{k}), \end{aligned} \tag{42}$$

where we've defined

$$\mathbf{p}_+ \equiv \mathbf{p} + \frac{\mathbf{k}}{2}, \quad \mathbf{p}_- \equiv \mathbf{p} - \frac{\mathbf{k}}{2}, \tag{43}$$

and $I^{(2.1)}_{jmn}(\mathbf{k})$ and $I^{(2.2)}_{jmn}(\mathbf{k})$ to be the first and second terms in the expression for $I^{(2)}_{jmn}(\mathbf{k})$. Since $I^{(2.2)}_{jmn}(\mathbf{k})$ has an explicit factor of k in front, we can evaluate this term to linear order in $\mathbf{k}$ by setting $\mathbf{k} = 0$ inside the integral. This leads to

$$I^{(2.2)}_{jmn}(\mathbf{k}) = -\frac{k_j}{2}\int_{\mathbf{p}} \frac{P_{mi}(\mathbf{p})P_{ni}(\mathbf{p})}{\Gamma(\mathbf{p})[\Gamma(\mathbf{p})+\Gamma(\mathbf{p})]} \ = \ -\frac{d-1}{4d}\frac{S_d}{(2\pi)^d}\mu^{-2}\Lambda^{d-4}d\ell\, k_j\delta_{mn}. \tag{44}$$

The calculation of $I^{(2.1)}_{jmn}(\mathbf{k})$ requires more effort. Expanding the numerator we get

$$I^{(2.1)}_{jmn}(\mathbf{k}) = -\int_{\mathbf{p}} \frac{p_j}{\Gamma(\mathbf{p}_+)} \frac{1}{[\Gamma(\mathbf{p}_+)+\Gamma(\mathbf{p}_-)]} \left(\delta_{mi} - \frac{p^+_m p^+_i}{p^2_+}\right)\left(\delta_{ni} - \frac{p^-_n p^-_i}{p^2_-}\right)$$
$$= -\int_{\mathbf{p}} \frac{p_j}{\Gamma(\mathbf{p}_+)} \frac{\delta_{mn} - \frac{p^-_m p^-_n}{p^2_-} - \frac{p^+_m p^+_n}{p^2_+}}{[\Gamma(\mathbf{p}_+)+\Gamma(\mathbf{p}_-)]} - \int_{\mathbf{p}} \frac{p_j p^+_m p^-_n}{\Gamma(\mathbf{p}_+)} \frac{\mathbf{p}_+\cdot\mathbf{p}_-}{[\Gamma(\mathbf{p}_+)+\Gamma(\mathbf{p}_-)]p^2_- p^2_+}. \quad (45)$$

Note that we have purposely written each term on the RHS of the second equality in Eq. (45) as an even function of $\mathbf{k}$ multiplied by a non-even function. We can simply set $\mathbf{k}=0$ inside the even part since it cannot be expanded to given a linear piece in k. Therefore, Eq. (45) can be simplified as

$$I^{(2.1)}_{jmn}(\mathbf{k}) = -\int_{\mathbf{p}} \frac{p_j}{\Gamma(\mathbf{p}_+)} \frac{p^2\delta_{mn} - 2p_m p_n}{2\Gamma(\mathbf{p})p^2} - \int_{\mathbf{p}} \frac{p_j p^+_m p^-_n}{\Gamma(\mathbf{p}_+)} \frac{1}{2\Gamma(\mathbf{p})p^2}.$$
$$= -\int_{\mathbf{p}} \frac{p_j}{\Gamma(\mathbf{p}_+)} \frac{p^2\delta_{mn} - 2p_m p_n}{2\Gamma(\mathbf{p})p^2} - \int_{\mathbf{p}} \frac{p_j}{\Gamma(\mathbf{p}_+)} \frac{1}{2\Gamma(\mathbf{p})p^2}\left(p_m p_n - \frac{p_m k_n}{2} + \frac{p_n k_m}{2} - \frac{k_m k_n}{4}\right)$$
$$= -\int_{\mathbf{p}} \frac{p_j}{\Gamma(\mathbf{p}_+)} \frac{1}{2\Gamma(\mathbf{p})p^2}\left(p^2\delta_{mn} - p_m p_n - \frac{p_m k_n}{2} + \frac{p_n k_m}{2} - \frac{k_m k_n}{4}\right)$$
$$= -\frac{1}{2\mu^2}\int_{\mathbf{p}} \frac{p_j}{p^6}\left(1 - \frac{p_s k_s}{p^2}\right)\left(p^2\delta_{mn} - p_m p_n - \frac{p_m k_n}{2} + \frac{p_n k_m}{2}\right) + O(k^2)$$
$$= -\frac{1}{2\mu^2}\int_{\mathbf{p}} \frac{1}{p^6}\left(-\frac{p_j p_m k_n}{2} + \frac{p_j p_n k_m}{2} - \delta_{mn} p_j p_s k_s + \frac{p_j p_s p_m p_n k_s}{p^2}\right) + O(k^2)\,. \quad (46)$$

The integral over $\mathbf{p}$ in this expression can now be evaluated by replacing $p_j p_m$, $p_j p_n$, $p_j p_s$, and $p_j p_s p_m p_n$ with their angular averages over all directions of $\mathbf{p}$ for fixed $|\mathbf{p}|$, as given by equations (52) and (60) of section (II). This gives

$$I^{(2.1)}_{jmn}(\mathbf{k}) = -\frac{1}{2\mu^2}\frac{S_d}{(2\pi)^d}\Lambda^{d-4} d\ell \left(-\frac{\delta_{mj}k_n}{2d} + \frac{\delta_{jn}k_m}{2d} - \frac{\delta_{mn}k_j}{d} + \frac{\delta_{mn}k_j + \delta_{nj}k_m + \delta_{mj}k_n}{d(d+2)}\right) + O(k^2)$$
$$= -\frac{1}{2\mu^2}\frac{S_d}{(2\pi)^d}\Lambda^{d-4} d\ell \left(-\frac{1}{2(d+2)}\delta_{mj}k_n - \frac{d+1}{d(d+2)}\delta_{mn}k_j + \frac{d+4}{2d(d+2)}\delta_{nj}k_m\right) + O(k^2)$$
$$= \frac{1}{4d(d+2)}\frac{S_d}{(2\pi)^d}\mu^{-2}\Lambda^{d-4} d\ell \left[d\delta_{mj}k_n - (d+4)\delta_{nj}k_m + 2(d+1)\delta_{mn}k_j\right] + O(k^2). \quad (47)$$

Plugging Eqs. (44,47) into Eq. (42) we get

$$I^{(2)}_{jmn}(\mathbf{k}) = \frac{1}{4d(d+2)}\frac{S_d}{(2\pi)^d}\mu^{-2}\Lambda^{d-4} d\ell \left[d\delta_{mj}k_n - (d+4)\delta_{nj}k_m + (-d^2+d+4)\delta_{mn}k_j\right]. \quad (48)$$

The terms in $I^{(1)}_{jmn}(\mathbf{k})$ (40) and $I^{(2)}_{jmn}(\mathbf{k})$ (48) that are proportional to k_j may be dropped, since they multiply $v_j(\tilde{\mathbf{k}})$, and, hence, vanish by the incompressibility condition $k_j v_j = 0$. Dropping them, and adding these two integrals $I^{(1)}_{jmn}(\mathbf{k})$ and $I^{(2)}_{jmn}(\mathbf{k})$ makes the entire correction to the equation of motion coming from graph II become:

$$\delta\left(\partial_t v_l\right) = -D\lambda^2 P_{lmn}(\mathbf{k}) v_j(\tilde{\mathbf{k}})\left(I^{(1)}_{jmn}(\mathbf{k}) + I^{(2)}_{jmn}(\mathbf{k})\right) = -\frac{D\lambda^2}{2\mu^2}\frac{S_d}{(2\pi)^d}\Lambda^{d-4} d\ell P_{lmn}(\mathbf{k}) v_j(\tilde{\mathbf{k}})$$
$$\times \left[\left(1 - \frac{2}{d} + \frac{1}{d(d+2)} - \frac{d+4}{2d(d+2)}\right)k_m\delta_{jn} + \left(\frac{1}{d(d+2)} + \frac{d}{2d(d+2)}\right)k_n\delta_{jm}\right]. \quad (49)$$

Simplifying, and performing the tensor index contractions, gives

$$\delta\left(\partial_t v_l\right) = -\frac{D\lambda^2}{2\mu^2}\frac{S_d}{(2\pi)^d}\Lambda^{d-4} d\ell \left[P_{lmj}(\mathbf{k})k_m\left(1 - \frac{5}{2d}\right) + P_{ljn}(\mathbf{k})k_n\frac{1}{2d}\right] v_j(\tilde{\mathbf{k}})$$
$$= -\frac{D\lambda^2}{2\mu^2}\frac{S_d}{(2\pi)^d}\Lambda^{d-4} d\ell \left[P_{lmj}(\mathbf{k})k_m\left(1 - \frac{2}{d}\right)\right] v_j(\tilde{\mathbf{k}})\,, \quad (50)$$

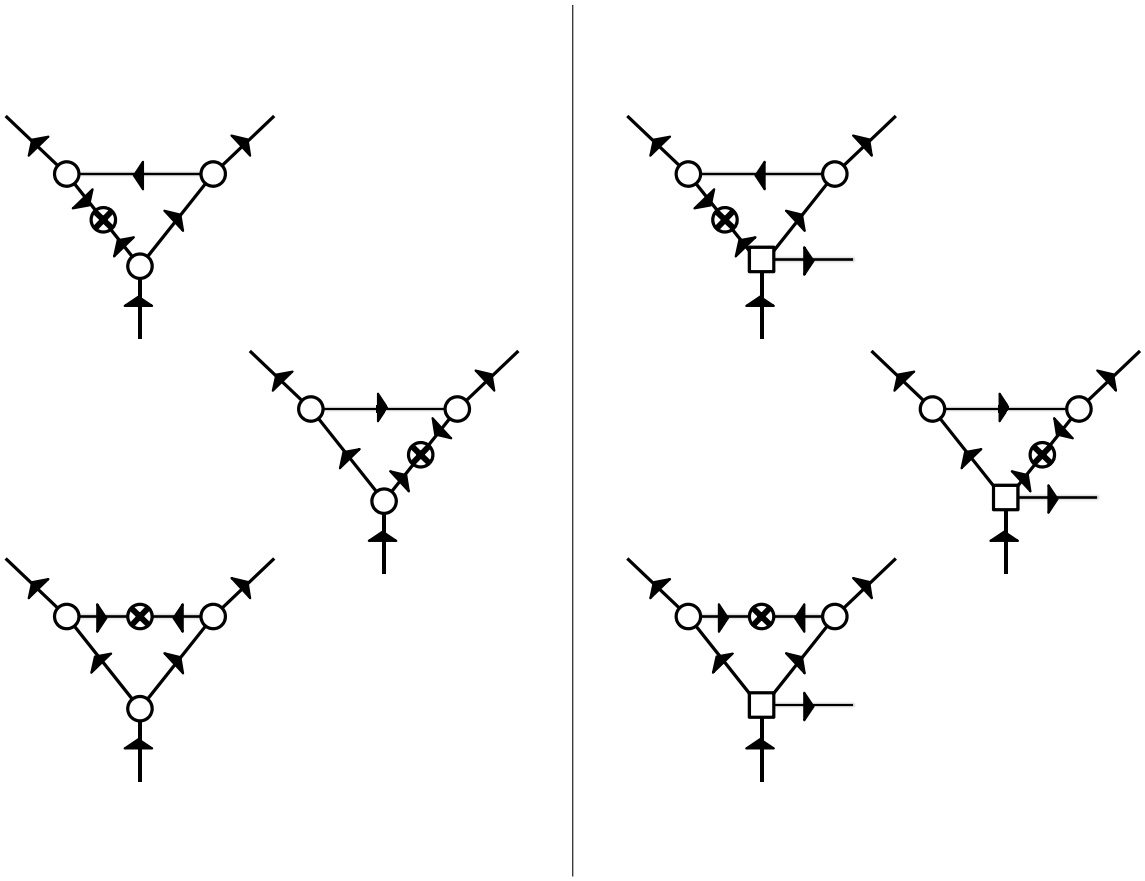


FIG. 1: Vanishing sets of diagrams. *Left*: The three diagrams consist of three three-point vertices. They cancel each other, leading to zero net contribution to λ. *Right*: The three diagrams consist of one four-point vertex and two three-point vertices. They again cancel each other, leading to zero net contribution to b.

where in the second step we have used the symmetry of $P_{ljn}(\mathbf{k})$ to write $P_{ljn}(\mathbf{k})k_n = P_{lmj}(\mathbf{k})k_m$. Now from the definition of $P_{lmj}(\mathbf{k})$, we have $P_{lmj}(\mathbf{k})k_m = P_{lm}(\mathbf{k})k_j k_m + P_{lj}(\mathbf{k})k_m k_m$. The first term in this expression vanishes by the fundamental property of the transverse projection operator, while the second is just $k^2 P_{lj}(\mathbf{k})$. Thus, we finally obtain

$$\delta\left(\partial_t v_l\right) \;=\; -\left(\left(\frac{d-2}{d}\right)\frac{D\lambda^2}{2\mu^2}\frac{S_d}{(2\pi)^d}\Lambda^{d-4}d\ell\right)k^2 P_{lj}(\mathbf{k})v_j(\tilde{\mathbf{k}})\,, \tag{51}$$

which is exactly what one would get by adding to μ (hiding in the "propagator") in Eq. (3) in the main text a correction $\delta\mu$ given by Eq. (4).

F. Vanishing diagrams

Besides the five non-vanishing one-loop diagrams, there are also two sets of one-loop diagrams that cancel exactly, giving zero contribution to the corrections (Fig. 1).

II. ANGULAR AVERAGES

In this section we derive the various angular averages used in the previous sections.
We begin by deriving the identity

$$\langle\frac{q_m q_n}{q^2}\rangle_{\hat{\mathbf{q}}} = \frac{1}{d}\delta_{mn}\,, \tag{52}$$

where $\langle\rangle_{\hat{\mathbf{q}}}$ denotes the average over directions $\hat{\mathbf{q}}$ of $\mathbf{q}$ for fixed $|\mathbf{q}|$. The identity (52) follows by symmetry: the average in question clearly must vanish when $m \neq n$, since then the quantity being averaged is odd in $\mathbf{q}$. Furthermore, when $m = n$, the average must be independent of the value that m and n both equal. Hence, this average must be proportional to δ_{mn}.

The constant of proportionality in (52) is easily determined by noting that the trace of this average over mn is $\langle\frac{q_m q_m}{q^2}\rangle_{\hat{\mathbf{q}}} = \langle\frac{q^2}{q^2}\rangle_{\hat{\mathbf{q}}} = \langle 1\rangle_{\hat{\mathbf{q}}} = 1$. This forces the prefactor of $\frac{1}{d}$ in (52).

From (52), it obviously follows that

$$\langle P_{mn}\rangle_{\hat{\mathbf{q}}} = \langle\delta_{mn} - \frac{q_m q_n}{q^2}\rangle_{\hat{\mathbf{q}}} = (1 - \frac{1}{d})\delta_{mn}\,, \tag{53}$$

which is the identity we used in equation (5).

We now consider the average of *two* projection operators $\langle P_{mi}(\mathbf{q})P_{jn}(\mathbf{q})\rangle_{\hat{\mathbf{q}}}$ that appears in (39). Using the definition of the projection operator, this can be written as follows:

$$\langle P_{mi}(\mathbf{q})P_{jn}(\mathbf{q})\rangle_{\hat{\mathbf{q}}} = \langle\left(\delta_{mi} - \frac{q_m q_i}{q^2}\right)\left(\delta_{jn} - \frac{q_j q_n}{q^2}\right)\rangle_{\hat{\mathbf{q}}} = \delta_{mi}\delta_{jn} - \delta_{mi}\langle\frac{q_j q_n}{q^2}\rangle_{\hat{\mathbf{q}}} - \delta_{jn}\langle\frac{q_m q_i}{q^2}\rangle_{\hat{\mathbf{q}}} + \langle\frac{q_i q_j q_m q_n}{q^4}\rangle_{\hat{\mathbf{q}}}\,. \tag{54}$$

The first two angular averages on the right hand side of this expression can be read off from (52). The last is new, and can be evaluated as follows:

First, note that by symmetry, this average vanishes unless the four indices i, j, k, l, are equal in pairs. Furthermore, if they are equal in pairs, but the pairs are different (e.g., if $i = j = x$ and $m = n = z$), then the average will have one value, independent of what the values of the two pairs of indices are (e.g., if $i = j = y$ and $m = n = x$, the average would be the same as in the example just cited. The only other non-zero possibility is that *all four* indices are equal, in which case the average is the same no matter which index all four are equal to (i.e., the average when $i = j = m = n = x$ is the same as that when $i = j = m = n = z$). Furthermore, this average must be completely symmetric under *any* interchange of its indices. This can be summarized by saying that the average must take the form:

$$\begin{aligned}&\langle\frac{q_i q_j q_m q_n}{q^4}\rangle_{\hat{\mathbf{q}}}\\ =\;& A\Upsilon_{ijmn} + B(\delta_{mi}\delta_{jn} + \delta_{ij}\delta_{mn} + \delta_{in}\delta_{jm})\,,\end{aligned} \tag{55}$$

where $\Upsilon_{ijmn} = 1$ if and only if $i = j = m = n$, and is zero otherwise, and A and B are unknown, dimension (d) dependent constants that we'll now determine.

We can derive one condition on A and B by taking the trace of (55) over any two indices (say, i and j). This gives

$$\langle\frac{q_m q_n}{q^2}\rangle_{\hat{\mathbf{q}}} = (A + (d+2)B)\delta_{mn}\,. \tag{56}$$

Comparing this with (52) gives

$$A + (d+2)B = \frac{1}{d}\,. \tag{57}$$

A second condition can be derived by explicitly evaluating the angle average when all four indices are equal. Since it doesn't matter what value they all equal, we'll chose it to be z. Defining θ to be the angle between the z-axis and $\mathbf{q}$, we can obtain the needed average in d-dimensions by integrating in hyperspherical coordinates:

$$\langle\frac{q_z^4}{q^4}\rangle_{\hat{\mathbf{q}}} = \frac{\int_0^\pi d\theta\cos^4\theta\sin^{d-2}\theta}{\int_0^\pi d\theta\sin^{d-2}\theta} = \frac{3}{d(d+2)}\,. \tag{58}$$

Comparing this with (55) evaluated for $i = j = m = m$ gives

$$A + 3B = \frac{3}{d(d+2)}\,. \tag{59}$$

The simultaneous solution for A and B of equations (57) and (59) is $A = 0$ and $B = \frac{1}{d(d+2)}$. Using these in (55) gives

$$\langle\frac{q_i q_j q_m q_n}{q^4}\rangle_{\hat{\mathbf{q}}} = \frac{1}{d(d+2)}(\delta_{mi}\delta_{jn} + \delta_{ij}\delta_{mn} + \delta_{in}\delta_{jm})\,. \tag{60}$$

Using this and (52) in (54) gives (39).

Renormalization Group Method for Differential Equations

Xiangjun Xing

Institute of Natural Sciences, and Department of Physics and Astronomy,

Shanghai Jiao Tong University, Shanghai, 200240 China[*]

(Dated: March 17, 2016)

Abstract

In this notes, we discuss the renormalization group method for solving differential equation.

PACS numbers:

[*]Electronic address: xxing@sjtu.edu.cn

I. DUFFY'S EQUATION

Consider Duffy's equation:

$$y''(t) + y(t) + \epsilon\, y(t)^3 = 0. \tag{1.1}$$

This is the equation satisfied by a nonlinear oscillator, whose total Hamiltonian is:

$$H = \frac{1}{2}p^2 + \frac{1}{2}y^2 + \frac{\epsilon}{4}y^4. \tag{1.2}$$

The Hamiltonian equations are:

$$p' = -\frac{\partial H}{\partial y} = -y - \epsilon y^2, \tag{1.3}$$

$$y' = \frac{\partial H}{\partial p} = p, \tag{1.4}$$

which is equivalent to Eq. (1.1). Three orbits in the phase space corresponding to different values of ϵ are illustrated in Fig. 1.

The total energy is a conserved quantity:

$$E = \frac{1}{2}y'(t)^2 + \frac{1}{2}y(t)^2 + \frac{\epsilon}{4}y(t)^4 \tag{1.5}$$

Therefore $y(t)$ remains bounded as a function of t.

The orbits are always periodic, hence $y(t)$ can be expanded as Fourier series. The period is however generically a function of amplitude. This is a generic property of nonlinear oscillator. Let ω be the base frequency, we have

$$\begin{aligned} y(t) &= A\cos\omega t + B\cos 3\omega t + C\cos 5\omega t + \ldots \\ &= \sum_{k=0} \Gamma_{2k+1}\cos(2k+1)\omega t. \end{aligned} \tag{1.6}$$

There are reasons why $2k\omega$ do not appear in the expansion. We can always make a time translation to Eq. (1.6) to obtain a more general class of solutions:

$$\begin{aligned} y(t) &= A\cos\omega(t+\tau) + B\cos 3\omega(t+\tau) + C\cos 5\omega(t+\tau) + \ldots. \\ &= \sum_{k=0} \Gamma_{2k+1}\cos(2k+1)\omega(t+\tau). \end{aligned} \tag{1.7}$$

In general, the solutions to a second order ODE form a two dimensional manifold. Since τ is independent of all other parameters, the other parameters in Eq. (1.7), namely $B, C, \ldots, \omega$,

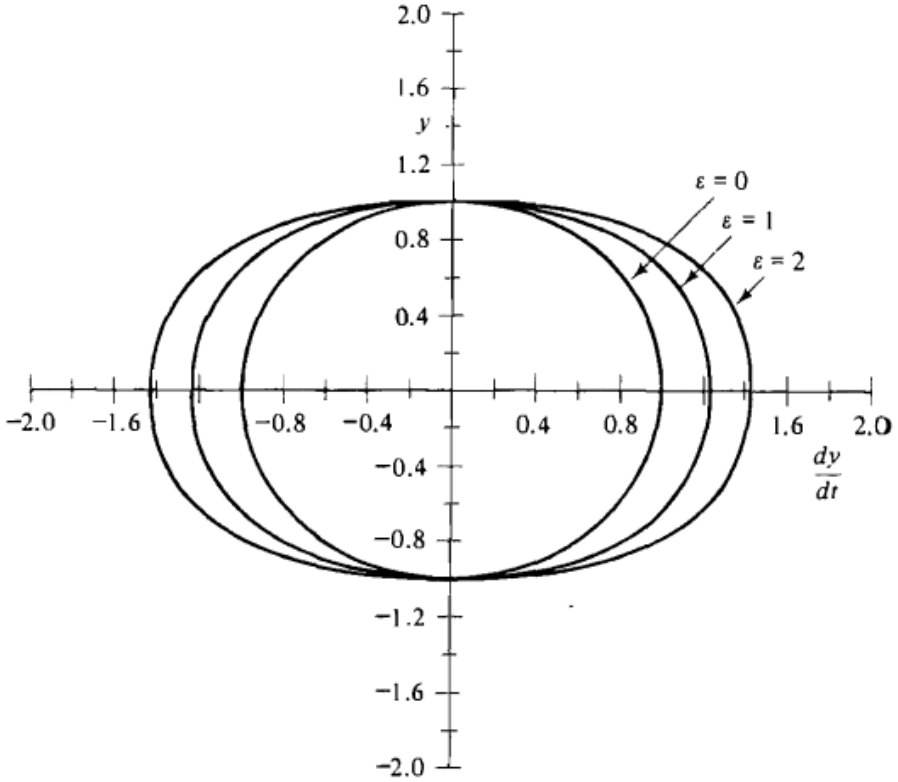


Figure 11.1 A phase-plane plot (y versus dy/dt) of solutions to Duffing's equation $d^2y/dt^2 + y + \varepsilon y^3 = 0$ $[y(0) = 1, y'(0) = 0]$ for $\varepsilon = 0$, 1, and 2. The orbits shown are constant-energy curves [see (11.1.8)] which satisfy $(dy/dt)^2 + y^2 + \varepsilon y^4/2 = 1 + \varepsilon/2$.

FIG. 1: Excerpt from the book by Bender and Orszag.

must be related to each other that only one of them is independent. Let it be A. Hence $B, C, \ldots, \omega$ can all be expressed as functions of A and ϵ. To find these coefficients, let us substitute the preceding expansion back into the Duffy's equation, and re-expand the equation as Fourier series. Each coefficient much vanish, hence we obtain:

$$3A^3\epsilon + 3A^2B\epsilon + A\left(6B^2\epsilon + 6BC\epsilon + 6C^2\epsilon - 4\omega^2 + 4\right) + 3B^2C\epsilon = 0, \tag{1.8a}$$

$$B\left(6A^2\epsilon + 6AC\epsilon + 6C^2\epsilon - 36\omega^2 + 4\right) + A^2\epsilon(A + 3C) + 3B^3\epsilon = 0, \tag{1.8b}$$

$$3A^2B\epsilon + 6A^2C\epsilon + 3AB^2\epsilon + 6B^2C\epsilon + 3C^3\epsilon - 100C\omega^2 + 4C = 0. \tag{1.8c}$$

To solve these equations for B, C, ω in a whole is clearly too complicated. Let us instead

solve them perturbatively. We write B, C, ω as perturbation series of ϵ:

$$B = 0 + B_1\,\epsilon + B_2\,\epsilon^2 + B_3\,\epsilon^3 + \cdots, \tag{1.9a}$$

$$C = 0 + C_1\,\epsilon + C_2\,\epsilon^2 + C_3\,\epsilon^3 + \cdots, \tag{1.9b}$$

$$\omega = 1 + \omega_1\,\epsilon + \omega_2\,\epsilon^2 + \omega_3\,\epsilon^3 + \cdots. \tag{1.9c}$$

There is no zero-th order term for B and C, because in the limit $\epsilon \to 0$, the solution must reduces to that of linear oscillator:

$$y(t) \to A\,\cos\omega t. \tag{1.10}$$

Substituting Eqs. (1.9) back into Eqs. (1.8), and re-expanding the latter in terms of ϵ, and comparing coefficients at each order, we find all the coefficients:

$$\begin{aligned} &B_1 = \tfrac{A^3}{32}, && C_1 = 0, && \omega_1 = \tfrac{3A^2}{8}, \\ &B_2 = -\tfrac{21A^5}{1024}, && C_2 = \tfrac{A^5}{1024}, && \omega_1 = -\tfrac{15A^4}{256}, \\ &B_3 = \tfrac{417A^5}{32768}, && C_3 = -\tfrac{43A^5}{32768}, && \omega_3 = \tfrac{123A^4}{8192} \end{aligned} \tag{1.11}$$

Substituting these back into Eqs. (1.9), we find

$$B = \frac{1}{32}A^3\epsilon - \frac{21}{1024}A^5\epsilon^2 + \frac{417}{32768}A^7\epsilon^3 + O(\epsilon^4), \tag{1.12a}$$

$$C = \frac{1}{1024}A^5\epsilon^2 - \frac{43}{32768}A^7\epsilon^3 + O(\epsilon^4), \tag{1.12b}$$

$$\omega = 1 + \frac{3}{8}A^2\epsilon - \frac{15}{256}A^4\epsilon^2 + \frac{123}{8192}A^6\epsilon^3 + O(\epsilon^4). \tag{1.12c}$$

Eqs. (1.7) and (1.12) form the solution to the order of ϵ^3. The solution contains two arbitrary parameters A, τ, which can be fixed by imposing initial conditions. Obviously, we can go to arbitrary higher (but finite) order if we need more accurate approximation.

II. NAIVE PERTURBATION METHOD

If we try to solve the problem using naive perturbation method, we would expand the solution in terms of ϵ:

$$y = y_0 + \epsilon\,y_1 + \epsilon^2\,y_2 + \epsilon^3\,y_3 + \ldots. \tag{2.1}$$

Substituting this back into Eq. (1.1), and expanding in terms of ϵ, and comparing coefficients order by order we find

$$y_0'' + y_0 = 0, \tag{2.2a}$$

$$y_1'' + y_1 = -y_0^3, \tag{2.2b}$$

$$y_2'' + y_2 = -3y_0^2 y_1, \tag{2.2c}$$

$$y_3'' + y_3 = -(3y_0 y_1^2 + 3y_0^2 y_2). \tag{2.2d}$$

We can solve all these equations order by order. That is, we first solve for zeroth order y_0, which contains two arbitrary constants. We shall however select a special solution in the form of

$$y_0(t) = A\,\cos(t). \tag{2.3a}$$

Substituting this back into Eq. (2.2b), we can solve for y_1:

$$y_1(t) = -\frac{3}{8}A^3 t\sin(t) + \frac{1}{32}A^3\cos(3t) + c_2\sin(t) + c_1\cos(t),$$

where c_1, c_2 are two arbitrary constants, which can be fixed only if we impose a specific initial conditions on $y_1(t)$. We shall however not impose any particular initial condition on $y_1(t)$ here. Rather, we shall impose relevant initial conditions only on the total solution. At this stage, we shall only note that we are free to choose any convenient values for c_1, c_2, since they correspond to addition of homogeneous solution to $y_1(t)$. This is amount to tuning the zero-th order solution $y_0(t)$. Hence from now on we have the following simple form for $y_1(t)$:

$$y_1(t) = -\frac{3}{8}A^3 t\sin(t) + \frac{1}{32}A^3\cos(3t). \tag{2.3b}$$

The first term grows with time t without bound, and therefore is a secular term. It makes the first order perturbation solution $y_0 + \epsilon y_1$ a bad approximation for large time. We will have to cure this problem, either using traditional multi scale method, or using renormalization group transformation.

Let us now substitute Eq. (2.2c), and solve for y_2. Again we choose two arbitrary constants such that y_2 does not contain any component of $\sin(t)$ and $\cos(t)$. We find

$$y_2(t) = \frac{A^5}{1024}\Big[-72t^2\cos(t) + 60t\sin(t) - 36t\sin(3t) - 21\cos(3t) + \cos(5t)\Big]. \tag{2.3c}$$

which contains second order secular terms that grow with t^2. We can keep going to work out higher order corrections, which contains higher and higher order secular terms. We can easily conclude that naive perturbation method becomes useless for this type of problems. We can actually expand the solution Eqs.(1.6) in terms of ϵ and verify explicitly that Eqs. (2.3) are all correct. The problem is how to recover the globally well behaved solution from a perturbation solution that contains secular terms at every order?

III. RENORMALIZATION GROUP METHOD

Let us define $\delta\omega = \omega - 1 = O(\epsilon)$, and write the argument of the trigonometric functions ωt in the following form:

$$\omega\, t = t + \delta\omega\, t = (t + t_0) + \delta\omega \left(t - \frac{1}{\delta\omega} t_0 \right), \tag{3.1}$$

where t_0 is an arbitrary constant. Defining

$$\mu = t_0/\delta\omega, \tag{3.2}$$

we can also write

$$\omega\, t = (t + \delta\omega\, \mu) + \delta\omega\, (t - \mu)\,. \tag{3.3}$$

Note that $t_0 = \delta\omega\, \mu$ is a slow variable comparing with μ. We are shifting t and $\delta\omega\, t$ in opposite direction such that their sum remains fixed. Such an operation is allowable even if t_0, μ are themselves functions of t!

Let us shift the time in the perturbative solution Eq. (1.6) (with $t_0 = t_0(\mu) = \delta\omega\, \mu$):

$$\begin{aligned} y(t) \;&=\; A\, \cos\left[(t + t_0) + \delta\omega\, (t - \mu)\right] \\ &+\; B\, \cos 3\left[(t + t_0) + \delta\omega\, (t - \mu)\right] \\ &+\; C\, \cos 5\left[(t + t_0) + \delta\omega\, (t - \mu)\right] \\ &+\; \ldots . \end{aligned} \tag{3.4}$$

Now we shall formally treat μ as an independent parameter. Parameters $B, C, \delta\omega$ of course still depend on ϵ. The basic idea is to expand the above expression in terms of ϵ while treating μ, t_0 as independent of ϵ. It is only after expansion that we impose the relation Eq. (3.2), then Eq. (3.4) (actually its series expansion in terms of ϵ) reduces to Eq. (1.6), which is the solution we are looking for.

Eq. (3.4) allows us to derive a useful relation. Recalling relation Eq. (3.2), and the fact that the function Eq. (3.4) is actually independent of μ, we can write

$$0 = \frac{d}{d\mu} y(t) = \frac{\partial y(t)}{\partial \mu} + \frac{\partial y(t)}{\partial t_0} \frac{dt_0}{d\mu}. \tag{3.5}$$

This is the renormalization group equation which states the invariance of the solution $y(t)$ with respect to the tuning of parameter μ.

Let us now expand Eq. (3.4) in terms of ϵ with t_0, μ treated as fixed constants. We find

$$\begin{aligned}
y(t) &= \tilde{y}_0(t) + \epsilon \tilde{y}_1(t) + \epsilon^2 \tilde{y}_2(t) + \epsilon^3 \tilde{y}_3(t) + \cdots, && \text{(3.6a)} \\
\tilde{y}_0(t) &= A\, \cos(t + t_0), && \text{(3.6b)} \\
\tilde{y}_1(t) &= -\frac{3}{8} A^3 (t - \mu) \sin(t + t_0) + \frac{1}{32} A^3 \cos 3(t + t_0), && \text{(3.6c)} \\
\tilde{y}_2(t) &= \frac{A^5}{1024} \Big[-72(t-\mu)^2 \cos(t + t_0) + 60(t - \mu) \sin(t + t_0) \\
&\quad - 36(t - \mu) \sin 3(t + t_0) - 21 \cos 3(t + t_0) + \cos 5(t + t_0) \Big]. && \text{(3.6d)}
\end{aligned}$$

Comparing these with Eqs. (2.3), we easily see the differences between y_k and $\tilde{y}_k$. To obtain $\tilde{y}_k$ from y_k, we only need to do two things: 1) replace all t inside trigonometric functions by $t + t_0$, and 2) replace all t outside trigonometric functions by $t - \mu$.

Let us now reverse the logic, and impose the invariance relation Eq. (3.8) on Eq. (3.6). This invariance relation establishes t_0 as a function of μ. We can expand $t_0(\mu)$ in Taylor series:

$$t_0(\mu) = \epsilon\, f_1(\mu) + \epsilon^2\, f_2(\mu) + \ldots. \tag{3.7}$$

There is no zeroth order term in this expansion, because we know that t_0 is a slow variable. It does not change with μ in the limit $\epsilon \to 0$.

Generally speaking, Eq. (3.6) (expanded up to infinite order, in principle,) contains three parameters A, t_0, μ. We know however all solutions to a second order ODE form a two dimensional manifold. This implies that only two of the three parameters are truly independent. That is, if we tune μ, there is a way to tune the other two parameters A, t_0 simultaneously such that the solution Eq. (y-sol-tilde) remains fixed. This invariance allows us to determine A, t_0 as functions of μ, and the invariance relation Eq. (3.8) becomes generalized:

$$0 = \frac{d}{d\mu} y(t) = \frac{\partial y(t)}{\partial \mu} + \frac{\partial y(t)}{\partial t_0} \frac{dt_0}{d\mu} + \frac{\partial y(t)}{\partial A} \frac{dA}{d\mu}. \tag{3.8}$$

This is the most general form of renormalization group equation for our case.

We can now substitute the expansion Eq. (3.7) back into Eq. (3.6) and expand $y(t)$ in terms of ϵ. This expression is independent of the parameter μ, and this independence must be guaranteed at every order of ϵ. That is, if we take the derivative of this series expansion with respect to μ, we must obtain zero, to the every order of ϵ. Using Wolfram Mathematica, we find that to the zeroth order, the invariance is automatically guaranteed. To the second order, we find:

$$\begin{aligned} 0 = \frac{d}{d\mu}y(t) \;=\; & \frac{1}{8}A\sin(t)\left(3A^2 - 8f_1'(\mu)\right)\epsilon \\ & + \; 4A\Big[-36A^4\mu\cos(t) - 15A^4\sin(t) + 9A^4\sin(3t) + 36A^4 t\cos(t) \\ & + \; 32f_1(\mu)\cos(t)\left(3A^2 - 8f_1'(\mu)\right) - 24A^2 f_1'(\mu)(4(t-\mu)\cos(t) \\ & + \; \sin(3t)) - 256\sin(t)f_2'(\mu)\Big]\epsilon^2 + O(\epsilon^3). \end{aligned} \tag{3.9}$$

To the first order, therefore, we obtain:

$$3A^2 - 8f_1'(\mu) = 0, \rightarrow f_1(\mu) = \frac{3}{8}A^2\mu. \tag{3.10}$$

We choose a particular value of integral constant to simplify the analyses. Substituting this back into the second order result, and setting it to zero, we find

$$15A^4 + 256f_2'(\mu) = 0, \tag{3.11}$$

$$\rightarrow \; f_2(\mu) = -\frac{15}{256}A^4\mu. \tag{3.12}$$

Substituting these results into Eq. (3.7), we obtain $t_0(\mu)$ up to the second order:

$$t_0(\mu) = \frac{3}{8}A^2\mu\epsilon - \frac{15}{256}A^4\mu\epsilon^2 + O(\epsilon^3). \tag{3.13}$$

Comparing this with Eq. (3.2), we find

$$\delta\omega = \omega - 1 = \frac{3}{8}A^2\epsilon - \frac{15}{256}A^4\epsilon^2 + O(\epsilon^3). \tag{3.14}$$

But this is identical to Eq. (1.9c) obtained previously. Recall that we have

$$t + t_0(t) = t + \delta\omega\, t = \omega t. \tag{3.15}$$

Finally we shall set $\mu = t$ in Eq. (3.6), so that all secular terms vanish. (Recall that we are allowed to select different value of μ for different t. This amounts to choose μ as a function of t.) Eq. (3.6) then reduces to

$$y(t) = A\cos(\omega t) + \left(\frac{1}{32}A^3\epsilon - \frac{21}{1024}A^5\epsilon^2\right)\cos(3\omega t) + \frac{1}{1024}A^5\epsilon^2\cos(5\omega t). \tag{3.16}$$

But this is just Eq. (1.7) truncated at the second order of ϵ. Hence we can obtain the globally uniform solution from the perturbative expansion, using the renormalization group transformation.

IV. RAYLEIGH'S EQUATION

Duffy's is special in the sense that as we tune the parameter μ, only the phase t_0 evolves, whilst the amplitude A remains fixed. In this section let us consider a different example, where both phase and amplitude evolve during the renormalization group flow. Consider the Rayleigh's equation:

$$\ddot{y} + y = \epsilon \left(\dot{y} - \frac{1}{3}\dot{y}^3 \right). \tag{4.1}$$

Multiplying both sides by $dy = \dot{y}dt$, we obtain:

$$d\left(\frac{1}{2}\dot{y}^2 + \frac{1}{2}y^2 \right) = \epsilon \left(\dot{y}^2 - \frac{1}{3}\dot{y}^4 \right) dt. \tag{4.2}$$

The quantity inside bracket in l.h.s. can be identified as the energy, where as the r.h.s. can be understood as energy dissipation/input in dt. If $\dot{y}^2 > 3$, l.h.s. is negative, and the total energy decreases. By contrast, if $\dot{y}^2 < 3$, the total energy increases. The total energy is not conserved. The system will settle down to a periodic orbit, where the total energy averaged over a period is conserved.

We carry out standard perturbation analyses. At zeroth order, we find:

$$y_0(t) = R\,\cos(t + t_0), \tag{4.3}$$

At first order, we have

$$y_1(t) = \frac{R^3}{96}\sin 3(t + t_0) + \frac{R}{2}\left(1 - \frac{R^2}{4}\right)(t - \mu)\cos(t + t_0). \tag{4.4}$$

The second term is secular. Note that we have chosen there is no component of homogeneous solution $\sin(t + t_0)$. At second order, we have:

$$\begin{aligned} y_2(t) = &- \frac{R^5}{3072}\cos 5(t + t_0) + \frac{R^3\,(3R^2 - 8)}{1024}\cos 3(t + t_0) \\ &- \frac{1}{256}R^3\left(R^2 - 4\right)(t - \mu)\sin 3(t + t_0) \\ &- \frac{1}{256}R\left(R^4 - 32\right)(t - \mu)\sin(t + t_0) \\ &+ \frac{1}{128}R\left(3R^4 - 16R^2 + 16\right)(t - \mu)^2\cos(t + t_0). \end{aligned} \tag{4.5}$$

The full solution is given by

$$y(t) = y_0(t) + \epsilon\, y_1(t) + \epsilon^2\, y_2(t) + \cdots . \tag{4.6}$$

which contains three parameters R, t_0, μ. According to our previous discussion, only two of these three are truly independent. Hence we can express R, t_0 as functions of μ such that the solution is independent of μ. The RG equation then reads:

$$0 = \frac{d}{d\mu} y(t) = \frac{\partial y(t)}{\partial \mu} + \frac{\partial y(t)}{\partial t_0}\frac{dt_0}{d\mu} + \frac{\partial y(t)}{\partial R}\frac{dR}{d\mu}. \tag{4.7}$$

We now insert Eq. (4.6) into Eq. (4.7), expand the r.h.s. into trigonometric series. Let us first consider the coefficient of $\cos(t + t_0)$. Anticipating that $R'(\mu) = O(\epsilon), t_0(\mu) = O(\epsilon^2)$, we can throw aways many terms are at least of order of ϵ^3, and find that the coefficient is given by

$$-\frac{1}{64}\Big[3\epsilon R(\mu)^2(t-\mu) - 4t\epsilon + 4\mu\epsilon - 8\Big]\Big[8R'(\mu) + \epsilon R(\mu)^3 - 4\epsilon R(\mu)\Big] + O(\epsilon^3). \tag{4.8}$$

But this must vanish at the order of ϵ^2. Hence we deduce:

$$R'(\mu) = -\frac{1}{8}\epsilon R(\mu)\left(R(\mu)^2 - 4\right) + O(\epsilon^3). \tag{4.9}$$

Similarly, we can extract the coefficient of $\sin(t + t_0)$, and set it to zero, and find

$$0 = -R(\mu)t_0'(\mu) + \frac{1}{256}\epsilon^2 R(\mu)^5 - \frac{1}{8}\epsilon^2 R(\mu) + O(\epsilon^3). \tag{4.10}$$

Solving this equation, we find

$$t_0'(\mu) = -\frac{1}{8}\epsilon^2 + \frac{1}{256}\epsilon^2 R(\mu)^4 + O(\epsilon^3). \tag{4.11}$$

We can also extract coefficients of $\cos 3(t + t_0), \sin 3(t + t_0), \cos 5(t + t_0)$, and after using Eqs. (4.9) and (4.11), find that they do not yield any extra information at order of ϵ^2.

What we should do now is to solve the flow equations (4.9) and (4.11) to find functions $R(\mu), t_0(\mu)$, substituting them back to Eq. (4.6), and finally set the arbitrary parameter μ to t. The final renormalized solution is given by

$$\begin{aligned} y_R(t) \;&=\; R(t)\,\cos(t + t_0(t)) + \epsilon\,\frac{R(t)^3}{96}\sin 3(t + t_0(t)) \\ &+\; \epsilon^2\,\frac{R(t)^3\,(3R(t)^2 - 8)}{1024}\cos 3(t + t_0(t)) - \epsilon^2\frac{R(t)^5}{3072}\cos 5(t + t_0(t)). \end{aligned} \tag{4.12}$$