FRG lectures 2025

B. Delamotte, LPTMC, Sorbonne Université

1 Introduction

"*J'écrivais des silences, des nuits, je notais l'inexprimable. Je fixais des vertiges.* "
Arthur Rimbaud

"*Il y a un autre monde mais il est dans celui-ci.*"
Paul Eluard

1.1 What we shall talk about and what we shall leave out

In these notes, we shall be talking about the modern version of the renormalization group (RG) *à la* Wilson, now known as the functional renormalization group (FRG), the non-perturbative renormalization group (NPRG) or the non-perturbative functional renormalization group (NPFRG). The first part of the course will focus on methodology, using second-order phase transitions in simple models such as Ising or O(N) to show how the calculations are carried out and the fundamental concepts underlying the whole approach. Some structural aspects of the functional renormalization group will be demonstrated: RG flow in an infinite-dimensional space ("large river effect"), notion of continuous limit, convergence of the derivative expansion, ultraviolet (UV) and infrared (IR) asymptotic freedom, and so on. The relationship between perturbative and non-perturbative renormalization and some artifacts of perturbation theory will also be shown (alas, briefly). And finally, for this first part, the advantages and disadvantages of Wilsonian renormalization will be shown.

Although the first part of the course will present some physical results relating to phase transitions, it is in the second part of the course that non-trivial applications to physical systems, in this case quantum condensed matter, will be presented. Applications to gauge theories, cosmology and quantum gravity, for example, will not be covered. It will therefore be useful to consult the review:

"*The nonperturbative functional renormalization group and its applications*", N. Dupuis, L. Canet, A. Eichhorn, W. Metzner, J. M. Pawlowski, M. Tissier, N. Wschebor, Physics Reports 910, 1 (2021) and arXiv:2006.04853
which describes a fairly large number of physical applications of the method and provides an extensive bibliography in the field.

An introduction to functional renormalization is given in:
An Introduction to the Nonperturbative Renormalization Group, B. Delamotte, in: A. Schwenk, J. Polonyi (Eds.), Renormalization Group and Effective Field Theory Approaches to Many-Body Systems, Vol. 852 of Lecture Notes in Physics, Springer Berlin Heidelberg, 2012, pp. 49-132. doi:10.1007/978-3-642-27320-9_2 and arXiv:cond-mat/0702365
and in the notes made available by N. Dupuis on his homepage.

The curious can also consult the pedagogical (?) article introducing perturbative renormalization:
A hint of renormalization, B. Delamotte, Am. J. Phys. 72, 170-184 (2004) https://doi.org/10.1119/1.1624112 and Arxiv hep-th/0212049v3
as well as the book:
Quantum and Statistical Field Theory, M. Le Bellac, Oxford Science Publications, Clarendon Press, 1991.

1.2 Renormalization, a little history

"*Truth never triumphs, but its opponents eventually die*"
M. Planck

The history of renormalization and the renormalization group remains largely unwritten. While it is perfectly possible to understand Maxwell's equations without delving into the history of electromagnetism, this is not entirely the case for renormalization, and even less so for the renormalization group, as the weight of history continues to shape minds in 2025. What follows is by no means a faithful account of the historical development of renormalization, as gaps and overly approximate statements abound, and is intended only to suggest the reason for this strange state of affairs: we now have a renormalization method, admittedly incomplete but as powerful from a practical point of view as it is conceptually clear, which nevertheless remains largely ignored, if not condemned, by most of the physics community directly interested in the problems it could help to solve. This is why it is so important to pass on this method to the younger generation in the form of the course that is the subject of these notes.

Let's start with a bit of chronology.

- 1947: First Lamb shift calculation in quantum electrodynamics by H. Bethe; renormalization is truly born here,
- 1953 and 1954: E. Stueckelberg and A. Petermann, then M. Gell-Mann and F. E. Low: group law structuring perturbative renormalization and first notion of energy-scale-dependent coupling constant,
- 1957: First systematic procedure for eliminating UV divergences in quantum field theories by N. N. Bogoliubov and O. Parasiuk,
- 1959: Book by N. N. Bogoliubov and D. Shirkov entitled "The Theory of Quantized Fields", claims to explain the conceptual background of the renormalization group: largely misunderstood,
- 1966 - 1969: K. Heppe and then W. Zimmermann complete the construction initiated by Bogoliubov and Parasiuk; perturbative renormalization is fully under control (BPHZ program),
- 1960s: quantum field theory (QFT) was increasingly sidelined in particle physics, except for QED, which many saw as a fortunate exception that could not be generalized to other interactions (weak, strong and gravitational); statistical field theory (SFT) was not yet born,
- 1970 - 1972: proof of the renormalizability of gauge theories (Yang-Mills theories) by G. 't Hooft and M. Veltman: renewed interest in field theory in high-energy physics; the Glashow-Weinberg-Salam model developed in the 1960s was soon recognized as a serious candidate for the description of electromagnetic and weak interactions,
- 1970: formulation of the renormalization group equations by Callan and Symanzik,
- 1971: formulation by K. Wilson of his vision of the renormalization group and its applications to critical phenomena; SFT, although still in its infancy before this date, is truly born here,
- 1972: elaboration of the $\epsilon = 4 - d$ expansion with applications to statistical mechanics by K. Wilson and M. Fisher,
- 1973: proof of asymptotic freedom in Yang-Mills theories (one-loop calculation) by D.J. Gross and F. Wilczek and H.D. Politzer (for which they won the Nobel Prize),

- 70's and 80's: explosion of studies using renormalization and the perturbative renormalization group in high-energy physics, statistical physics and quantum condensed matter physics; very few studies were undertaken at this time following Wilson's non-perturbative ideas: "we think *à la* Wilson and calculate *à la* Feynman",

- 1984: J. Polchinski derives an exact Wilsonian RG equation and applies it to perturbative quantum field theory,

- 1994: U. Ellwanger, T. Morris and C. Wetterich independently derive an exact RG equation for the functional renormalization group different from Polchinski's (and related by Legendre transformation). This equation and the approximations it allows mark the revival of functional and non-perturbative renormalization *à la* Wilson.[1]

On the history of this field, we recommend reading K. Wilson's Nobel lecture: *The renormalization group and critical phenomena*, December 8, 1982.

Since 1995, FRG methods have been used in a large number of papers in many fields (see the bibliography of the review by N. Dupuis et al. cited above). Somewhat strangely, and despite the undeniable progress this method has made in many fields, it remains the prerogative of a fairly small community and is still regarded by many as an oddity leading to uncontrolled calculations that are best ignored (implying that perturbative calculations are, for their part, completely controlled!). Who in the string theory community, for example, has heard of the so-called "asymptotic safety" scenario for quantum gravity, showing that Hilbert-Einstein gravity could lead to a renormalizable quantum theory in a non-perturbative sense? Essentially nobody, although this scenario was initiated by S. Weinberg in the late 70s...

Historically, renormalization was developed in quantum field theory (in fact, essentially in QED) to eliminate the divergences that appear at second order in perturbation theory and at all subsequent orders.[2] These divergences are UV divergences, i.e. divergences originating in the theory's high-momentum regime (high-energy) or, to put it another way, short-distance behavior. The renormalization has therefore been constructed as a recursive algorithm, i.e. operating order by order of the perturbative series, for eliminating these divergences. The algorithm seemed a little magical and ill-defined to many physicists from the '50s to the '80s, but was justified *a posteriori* by the construction of a (renormalized) perturbation theory which, order by order, is well defined, at the sole cost of fixing a finite (and small) number of external parameters, such as masses and coupling constants. The method constructs a predictive theory, since an infinite number of independent quantities can be calculated on condition that a finite number of parameters not predicted by the theory are fixed by experiment.

The origin of the UV divergences is well known. In the transition from classical to quantum theory, fields must be considered not as functions, but as operator-valued distributions.[3] Since interactions are local, they translate into products of operators at the same space-time point (e.g. $\phi^4(x)$ or $\bar{\psi}(x)\gamma^\mu\psi(x)A_\mu(x)$), i.e. products at the same point of distributions, which is generally illegal. From classical theories that are local, the canonical quantization procedure generically produces distributional continuations that are sick at short range, and the perturbative renormalization constructs the right continuation order by order, modifying (we say, regularizing) the theory's short-range structure while preserving the long-range one.[4] Whether this problem reflects a real physical

[1] Note: this equation had been derived before them on several occasions, including in the 70s, but had either not been exploited, or had only been exploited in the perturbative framework.

[2] Perturbation theory (Feynman graphs) in QFT is a semi-classical expansion. For a given process that exists in classical electrodynamics, such as the scattering of an electron on a nucleus, the first order of perturbation theory gives the classical result and the successive orders a series in powers of $\hbar$. This is true for all quantum field theory

[3] To be precise if not pedantic, quantum fields must be considered as operator-valued tempered distributions.

[4] Of course, the notions of short and long distance should be made precise by specifying in relation to which distance scale they are considered short or long. At this stage, it is enough to think in terms of asymptotically short and asymptotically long distances, even if we can make all this much more quantitative

problem of quantum field theories or is merely an artefact of perturbation theory is not easily decided in the perturbative framework, and the reality is that the UV divergences encountered in perturbation theory can reflect real behavior of short-distance quantum "fluctuations" and can also be a pathology of the perturbative expansion itself. It is fair to say that, beyond the algorithm for eliminating UV divergences, it was not easy to find one's way through this conceptual imbroglio.

As for the renormalization group , it appeared at the time, i.e. before the 70s, as a condition for the coherence of the perturbative renormalization which, in certain cases, enabled it to be improved by resumming some of the terms of the perturbative expansion.[5] Logically, then, renormalization group comes after perturbative renormalization, which, from a practical point of view, can do without it. It seems to be no more than a welcome by-product of perturbative renormalization. Even with Callan and Symanzik's definitive formalization of the renormalization group equations used in perturbation theory, the latter does not constitute a computational method in itself, but is rather a post-processing of perturbative results. For example, from a perturbative result such as $X(t) = X_0(1 + \alpha\epsilon \log t + O(\epsilon^2))$, the Callan-Symanzik renormalization group can, if this is the case, show that this is the expansion of a power law which is therefore inevitably $X(t) = X_0\, t^{\alpha\epsilon+O(\epsilon^2)}$. This is by no means trivial, as reconstructing the power law means resumming powers of arbitrarily high order in ϵ, even though only the first order is initially known. However, the Callan-Symanzik renormalization group cannot be used to calculate the perturbative expansion, and Feynman-style techniques are required for this.

On the other hand, the renormalization group explains that coupling constants become functions of the scale (distance, energy, etc.) on which they are measured. Coupling constants are indeed constants in classical theories, as is the case for electric charge in classical electrodynamics: an experiment (scattering of two electrons, for example) can measure this charge, which is the same coupling constant involved in all classical electrodynamic phenomena involving only electrons. This is not the case in quantum electrodynamics and, more generally, in quantum field theory. A coupling constant is a measurable quantity (just like a mass) and, as such, must be determined from Green functions (or correlation functions in SFT), functions which are the physical quantities of the theory, unlike a Hamiltonian (or a Lagrangian).[6] For example, in ϕ^4 theory, the coupling constant of the quartic term can be measured, at least in principle, from the four-field correlation function $\Gamma^{(4)}(p_1, p_2, p_3)$, which is a physical quantity. These functions, which are actually numbers in classical field theory, become momentum-dependent functions in QFT, i.e. as soon as quantum "fluctuations" are taken into account (or statistical fluctuations in SFT). As a result, we can no longer speak of a single coupling constant, but of a scale-dependent coupling constant.[7] The Callan-Symanzik equations describe how the same field theory can be parameterized by coupling constants that differ according to the scale on which they are measured.

The ambition of the Wilsonian RG is quite different, as its creator claimed. It is, on the one hand, a genuine method for calculating correlation functions and, on the other, a method that is not a priori linked to perturbation theory. A word of warning is in order here. If the FRG could be used to calculate exactly all the correlation functions of a non-trivial statistical (or quantum) field theory model, it would have been solved. This is almost never the case in reality, and approximations have to be used to make use of the FRG as a computational method. Over and above its conceptual relevance, which is very high indeed, its ability to enable non-trivial calculations therefore rests

[5] It should be noted that the resummation of some of the perturbative diagrams, the so-called parquet diagrams for example, was practiced in the 60s independently of any reference to the renormalization group. The Russian school, following Landau, counted specialists in this technique among its ranks.

[6] Carlo Becchi, a great expert in field theory, used to say that a Lagrangian is an opinion. The a sorrowful spirit might object that Green functions in QFT are not physical either, only "on-shell" functions are. This is not the case in SFT, and this assertion about QFT is only valid because only S matrix elements are considered physical.

[7] In general, Green functions do not depend solely on an energy scale, but on all the kinematic configurations of the momenta on which they depend. If we speak of coupling constants depending on the energy scale only, it is because we are reasoning in terms of a given kinematic configuration (the angles between momenta) and we are studying the variation of Green function when we vary only the scale of its momenta.

on the possibility of implementing controlled approximations that go beyond perturbation theory.[8] But having controlled approximate methods is a subtle art. The physics community still believes far too often that perturbative methods are the only ones that can be controlled.[9] But this is not the case. The perturbative series encountered in physics are in the vast majority of cases non-convergent series, and writing a $O(\epsilon^{n+1})$ after a series of order n in ϵ makes no difference if the series is not convergent. Quantum electrodynamics is an exception in this respect, because it allows very precise calculations, which become more and more precise the higher the order of perturbation.[10] In many cases, renormalized series in field theory are at best asymptotic, Borel summable series. But even in these happy cases, the physical results depend on the resummation procedures of the renormalized series, which are largely arbitrary and for which it is very difficult to quantify the error bars.

FRG seeks to go beyond perturbative methods, but cannot escape the criticism of having difficulty quantifying the errors made. This was the fundamental criticism levelled at the FRG community by the perturbative RG community for several decades: due to approximations, even physical quantities seemed to depend catastrophically on non-physical and arbitrary parameters (the parameters used to perform coarse-graining, see the following). For a long time, these criticisms were justified. However, after the mid-90s, when a new, exact RG equation was developed by C. Wetterich (and also T. Morris and U. Ellwanger), the accumulation of completely non-trivial and accurate results left little doubt in the minds of specialists that the results obtained by this method were indeed inaccurate up to a certain point, but not catastrophically so, as was the case in earlier studies that did not rely on the same approximation schemes. More recently, great progress has been made in this direction, by showing that the most commonly used approximation scheme, the derivative expansion, leads, at least in certain cases and for certain quantities, to series that are convergent and therefore under control.[11] It is worth noting, however, that most of the perturbative RG community (and not just them) has not yet realized this change and remains on the idea that calculations based on the Wilsonian approach are out of control. Patience, it will change...

2 Phase transitions and functional methods

"*Vivre, c'est naître lentement*"
Antoine de Saint-Exupéry

In what follows, we will be using the Ising model (and its O(N) symmetric generalization) and the ϕ^4 model (known as the Ginzburg-Landau model) as physical supports to illustrate the concepts involved, and as frameworks for concrete calculations. They will be both simple enough not to add

[8] Some perturbative calculations can be performed very efficiently using the FRG, more efficiently in fact than using *à la* Feynman graphical methods. It remains true, however, that to this day and for many others, the usual graphical methods remain irreplaceable for perturbative calculations.

[9] By definition, exact methods for solving N-body systems - integrable or conformally invariant systems in $d = 2$ - provide controlled results. As for approximate methods, the conformal bootstrap is one of the very few that provide exact error bars. Unfortunately, very few physical quantities can be calculated using this method.

[10] It should be noted, however, that this does not mean that QED perturbative series are convergent series. They are in fact divergent, but the small coupling constant used for development ($\alpha/2\pi \simeq 1/860$ where $\alpha = 1/137$ is the fine structure constant) is responsible for the apparent convergence of the first orders of perturbation.

[11] Note now that the practical interest of the FRG is not just to provide accurate results in situations of strong (renormalized) couplings where perturbation theory is in trouble. As these notes will show, the FRG can also be used in certain cases to calculate quantities that are beyond the reach of any perturbative approach. The example of non-universal quantities such as a phase diagram or even a critical temperature is emblematic in this respect: this calculation is essentially impossible in the perturbative framework, but quite accessible, in some cases at least, to the FRG, including lattice models. But there are also cases where universal quantities cannot be calculated perturbatively, because they cannot be developed in series with a coupling constant. The Ising random-field model or the Kardar-Parisi-Zhang equation in dimensions greater than one, or certain multi-critical phenomena in O(N) models, provide such examples

to the complexity of the field, and rich enough not to give an ideal vision that is inapplicable to any more complicated system.

2.1 O(N) models, Hamiltonians, free energies and correlation functions

The Ising model is a model of classical spins on a d-dimensional lattice (assumed to be hypercubic), meaning it consists of random variables $\sigma_i = \pm 1$ located at the nodes of a lattice indexed by i. We will consider in the following hypercubic lattices and the lattice spacing will be denoted by a, and its size by L. Periodic boundary conditions are assumed, and we will primarily focus on the thermodynamic limit where $L \to \infty$. The interaction between spins is assumed to occur only between nearest neighbors and to be ferromagnetic. The interaction Hamiltonian in the presence of an external magnetic field h_i is therefore:

$$H = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - \sum_i h_i \sigma_i \tag{1}$$

where $\langle ij \rangle$ means the sum is taken over pairs of sites i and j that are nearest neighbors, and where $J > 0$ is the ferromagnetic exchange constant. In the presence of a constant magnetic field, the magnetic term becomes $-h \sum_i \sigma_i$, and the field couples to the magnetization mode, i.e., the sum of the spins whose average value is the total magnetization of the system. When $h_i = 0$, the system exhibits a $\mathbb{Z}_2$ symmetry of simultaneous inversion of all spins: $\sigma_i' = -\sigma_i$.

The model can be generalized to spins with N components of unit magnitude: $|\boldsymbol{\sigma}_i| = 1$, with a Hamiltonian exhibiting O(N) symmetry when the magnetic field is zero:

$$H = -J \sum_{\langle ij \rangle} \boldsymbol{\sigma}_i \cdot \boldsymbol{\sigma}_j - \sum_i \boldsymbol{h}_i \cdot \boldsymbol{\sigma}_i. \tag{2}$$

If we agree that $\mathrm{O}(1) = \mathbb{Z}_2$, the Ising model becomes a special case of the family of O(N) models.[12]

We will also be interested in "$\boldsymbol{\phi}^4$ models," which are models where $\boldsymbol{\phi}(x)$ is an N-component field varying between $-\infty$ and $+\infty$, defined on d-dimensional Euclidean space and featuring an internal O(N) symmetry.[13] The Hamiltonian we will consider is:

$$H = \int_x \left(\frac{1}{2} (\nabla \boldsymbol{\phi}(x))^2 + \frac{1}{2} r_0 \boldsymbol{\phi}^2(x) + \frac{1}{4!} g_0 (\boldsymbol{\phi}^2(x))^2 \right) - \int_x \boldsymbol{h}(x) \cdot \boldsymbol{\phi}(x) \tag{3}$$

where $\int_x = \int d^d x$. The partition functions of these models are

$$\mathcal{Z}[h] = \sum_{\boldsymbol{\sigma}_i} e^{-\beta H[\boldsymbol{\sigma}, \boldsymbol{h}]} \qquad \text{or} \qquad \mathcal{Z}[h] = \int D\boldsymbol{\phi}\, e^{-\beta H[\boldsymbol{\phi}, \boldsymbol{h}]} \tag{4}$$

depending on the type of model considered. In the previous equation, $\beta = 1/kT$, where T is the temperature, and the sum or integral represents a summation over all configurations of the $\boldsymbol{\sigma}_i$ or $\boldsymbol{\phi}(x)$ fields, so the integral is a functional integral. In all that follows, the β factor will be absorbed into H via a redefinition of the field and the constants r_0 and g_0, with the result that H will depend on

[12] Note: The inversion operation, which involves changing a spin to its opposite: $\boldsymbol{\sigma}_i' = -\boldsymbol{\sigma}_i$, is or is not a rotation depending on whether N is even or odd, respectively. For example, for $N = 3$, this inversion, represented by the matrix $\mathrm{diag}(-1, -1, -1)$, has determinant -1 and does not belong to SO(3), while for $N = 2$ or 4, it is indeed a rotation. On the other hand, mirror symmetries that change only one component of the vectors $\boldsymbol{\sigma}_i$ are never rotations; they are elements of O(N) but not of SO(N).

[13] The models studied exhibit two types of symmetry: symmetries originating from the isotropy group of the d-dimensional space on which they are defined, in our case the symmetries of the lattice or Euclidean space, and so-called internal symmetries, here O(N). The fields $\boldsymbol{\phi}(x)$ we consider are scalars for the isotropy group of Euclidean space and vectors for O(N) (more precisely, true scalars and true vectors for O(N)). Note that we could also consider a third type of model, where the random variables $\boldsymbol{\phi}_i$ vary between $-\infty$ and $+\infty$ and are defined on a lattice.

temperature, a forbidden characteristic for a true Hamiltonian. By abuse of notation and language, we will continue to denote it H and call it the Hamiltonian of the system. The partition functions will therefore be written as:

$$\mathcal{Z}[\boldsymbol{h}] = \sum_{\boldsymbol{\sigma}_i} e^{-H[\boldsymbol{\sigma},\boldsymbol{h}]} \quad \text{or} \quad \mathcal{Z}[\boldsymbol{h}] = \int D\boldsymbol{\phi}\, e^{-H[\boldsymbol{\phi},\boldsymbol{h}]} \tag{5}$$

By the same abuse of language, we will call

$$W[\boldsymbol{h}] = \log \mathcal{Z}[\boldsymbol{h}] \tag{6}$$

the (Helmholtz) free energy instead of $-kT \log \mathcal{Z}[h]$, as the $-kT$ factor plays no interesting role in what follows.

The correlation functions of the components of the field are obtained by differentiating $\mathcal{Z}[\boldsymbol{h}]$:

$$G^{(n)}_{\alpha_1,\cdots,\alpha_n}[i_1,\cdots,i_n;h] = \langle \sigma_{i_1}^{\alpha_1} \cdots \sigma_{i_n}^{\alpha_n} \rangle = \frac{1}{\mathcal{Z}[h]} \frac{\partial^n \mathcal{Z}[\boldsymbol{h}]}{\partial h {i_1}^{\alpha_n} \cdots \partial h {i_n}^{\alpha_n}} \tag{7}$$

or, in the continuous case:

$$G^{(n)}\alpha_1,\cdots,\alpha_n[x_1,\cdots,x_n;h] = \langle \phi\alpha_1(x_1)\cdots\phi_{\alpha_n}(x_n)\rangle = \frac{1}{\mathcal{Z}[h]} \frac{\delta^n \mathcal{Z}[\boldsymbol{h}]}{\delta h_{\alpha_1}(x_1)\cdots\delta h_{\alpha_n}(x_n)} \tag{8}$$

where these are now functional derivatives.

Of course, one can calculate these correlation functions in a zero external field, which means that after calculating the derivatives of $\mathcal{Z}[\boldsymbol{h}]$, we evaluate them at $\boldsymbol{h} = 0$. However, note that even in this case, we must first consider the functional dependence of $\mathcal{Z}$ on $\boldsymbol{h}$ in order to compute its derivatives with respect to $\boldsymbol{h}$. This is why $\mathcal{Z}[\boldsymbol{h}]$ is called the generating functional of correlation functions, and the dependence on $\boldsymbol{h}$ is indicated with square brackets rather than parentheses to follow the universally applied rule for denoting functionals. Similarly, $W[\boldsymbol{h}]$ generates the so-called connected correlation functions:

$$G^{(n)}_{c,\alpha_1,\cdots,\alpha_n}[i_1,\cdots,i_n;h] = \langle \sigma_{i_1}^{\alpha_1} \cdots \sigma_{i_n}^{\alpha_n} \rangle_c = \frac{\partial^n W[\boldsymbol{h}]}{\partial h {i_1}^{\alpha_n} \cdots \partial h {i_n}^{\alpha_n}} \tag{9}$$

and

$$G^{(n)}_{c,\alpha_1,\cdots,\alpha_n}[x_1,\cdots,x_n;h] = \langle \phi_{\alpha_1}(x_1)\cdots\phi_{\alpha_n}(x_n)\rangle_c = \frac{\delta^n W[\boldsymbol{h}]}{\delta h_{\alpha_1}(x_1)\cdots\delta h_{\alpha_n}(x_n)}. \tag{10}$$

For the two-point function (this would not be true for the four-point function or higher), the connected function is the correlation function of the field fluctuations around its mean value, as an elementary calculation shows:

$$G^{(2)}_{c,\alpha_1,\alpha_2}[x_1,x_2;h] = \langle \Big(\phi_{\alpha_1}(x_1) - \langle\phi_{\alpha_1}(x_1)\rangle\Big)\Big(\phi_{\alpha_2}(x_2) - \langle\phi_{\alpha_2}(x_2)\rangle\Big)\rangle. \tag{11}$$

In the following, when we deal with correlation functions, whether connected or not, in a constant field, we will write them as functions of $\boldsymbol{h}$: $G^{(2)}_{\alpha_1,\cdots,\alpha_n}(x_1,\cdots,x_n,\boldsymbol{h})$, and when they are in a zero field, we will omit the last argument and write them simply as $G^{(2)}_{\alpha_1,\cdots,\alpha_n}(x_1,\cdots,x_n)$. In both cases, the system being translation invariant, these functions depend on one less argument because they only depend on the differences $x_i - x_j$, or in other words, one of the x_i can be chosen as the origin.

The Fourier transform is defined by :

$$\tilde{f}(q) = \int_x e^{-iqx} f(x) \tag{12}$$

and its inverse by

$$f(x) = \int_q e^{iqx} \tilde{f}(q) \tag{13}$$

where $\int_q = \int d^d q/(2\pi)^d$. In cases where there is translation invariance, i.e. when the $\boldsymbol{h}$ field is constant, it is convenient to define new correlation functions in Fourier space by recognizing that the $\tilde{G}^{(n)}$ are proportional to $(2\pi)^d \delta(q_1 + \cdots + q_n)$ and thus defining:

$$\tilde{G}^{(n)}_{\alpha_1,\cdots,\alpha_n}(q_1, \cdots, q_n, h) = (2\pi)^d \delta\left(\sum_{i=1}^{n} q_i\right) G^{(n)}_{\alpha_1,\cdots,\alpha_n}(q_1, \cdots, q_{n-1}, h) \tag{14}$$

where the function on the right-hand side has no tilde. There should be no confusion with a function in direct space, as the context always tells us what kind of function we are talking about.

The average value is called the total magnetization:

$$\boldsymbol{M} = \langle \sum_i \boldsymbol{\sigma}_i \rangle. \tag{15}$$

We often prefer to work with the local magnetization

$$\boldsymbol{m}_i = \langle \boldsymbol{\sigma}_i \rangle = \frac{\partial W[\boldsymbol{h}]}{\partial \boldsymbol{h}_i} \tag{16}$$

or, in the continuum

$$\boldsymbol{m}(x) = \langle \boldsymbol{\phi}(x) \rangle = \frac{\delta W[\boldsymbol{h}]}{\delta \boldsymbol{h}(x)}. \tag{17}$$

Of course, if $\boldsymbol{h}$ is constant, $\boldsymbol{m}$ is also constant and is an intensive quantity, while $\boldsymbol{M}$ is extensive. Note that the function $G_c^{(2)}(x, y)$ can be written as

$$G^{(2)}_{c,\alpha\beta}(x, y) = \frac{\delta m_\alpha(x)}{\delta h_\beta(y)} \tag{18}$$

and is therefore nothing other than the response of the local magnetization at x to a variation of the magnetic field at y. It is thus a local generalization of the magnetic susceptibility, which represents the response of the magnetization to a variation of a constant magnetic field.

For what follows, it is important to define another free energy by changing the state variable from magnetic field to magnetization. This is done, as usual, by Legendre transformation, and so we define $\Gamma[\boldsymbol{m}]$, the Gibbs free energy, by (from now on, to establish the definitions of the quantities we are interested in, we will only give their definitions in the continuum):

$$\Gamma[\boldsymbol{m}] + W[\boldsymbol{h}] = \int_x \boldsymbol{h}(x) \cdot \boldsymbol{m}(x). \tag{19}$$

This free energy exists because W can be shown to be a convex function of $\boldsymbol{h}$, so Eq. (17) can be inverted and $\boldsymbol{h}$ can be calculated as a function of $\boldsymbol{m}$ and inserted into equation (19) to give Γ. Note that Eq. (19) is symmetrical in the simultaneous exchange $W \to \Gamma$ and $\boldsymbol{h} \to \boldsymbol{m}$, so from Eq. (17) we necessarily have the reciprocal relationship:

$$\boldsymbol{h}(x) = \frac{\delta \Gamma[\boldsymbol{m}]}{\delta \boldsymbol{m}(x)}, \tag{20}$$

a relation that can obviously be established directly from Eq.(19) by differentiating with respect to $\boldsymbol{m}(x)$.[14] It should be noted that in the absence of a magnetic field, $\boldsymbol{h} = 0$, the previous relation implies that the magnetization of the system is the one that minimizes the Gibbs free energy (a maximum is excluded because Γ is a convex function, as it is the Legendre transform of a convex function). In the presence of a constant $\boldsymbol{h}$ field, Eq.(20) is the equation of state of the system, which relates the magnetic field, magnetization, and temperature.

Just as we defined correlation functions by functionally differentiating $Z[\boldsymbol{h}]$ and $W[\boldsymbol{h}]$, we can define new functions by differentiating Γ. The functions obtained in this way are called the 1-Particle-Irreducible (1PI) correlation functions or vertex functions.[15] They are defined by:

$$\Gamma^{(n)}_{\alpha_1,\cdots,\alpha_n}[x_1,\cdots,x_n;m] = \frac{\delta^n \Gamma[\boldsymbol{m}]}{\delta m_{\alpha_1}(x_1)\cdots\delta m_{\alpha_n}(x_n)}. \tag{21}$$

Here too, we can evaluate Γ in a configuration where $\boldsymbol{m}(x)$ is constant, which implies that the external magnetic field is also constant. In this case, the system is invariant under translation and the $\Gamma^{(n)}$ depend only on $n-1$ space variables, $x_i - x_j$. Then, the Fourier transform of these functions depends only on $n-1$ momenta q_i and just like for the $G^{(n)}$, Eq.(14), we define functions $\Gamma^{(n)}_{\alpha_1,\cdots,\alpha_n}(q_1,\cdots,q_{n-1},m)$ without a tilde, once a factor of $(2\pi)^d$ and a Dirac delta for the conservation of the total momentum are extracted.

Of course, the set of 1PI functions contains the same information as the set of connected correlation functions, since one can go from W to Γ via a Legendre transform. One can thus establish, function by function, how to reconstruct the $G_c^{(n)}$ from the $\Gamma^{(p\leq n)}$.

As an example, and because it will be useful later, let us establish the relation between $G_c^{(2)} = W^{(2)}$ and $\Gamma^{(2)}$ in the case where $N = 1$. We start with Eq. (19), which we differentiate with respect to $h(z)$:

$$\delta(x-z) = \frac{\delta h(x)}{\delta h(z)} = \int_y \frac{\delta^2\Gamma[m]}{\delta m(x)\delta m(y)}\frac{\delta m(y)}{\delta h(z)} = \int_y \Gamma^{(2)}[x,y;m]W^{(2)}[y,z;h]. \tag{22}$$

The previous relation, valid in the presence of any field $h(x)$, generalizes the matrix relation that connects a matrix to its inverse, $\sum_j A_{ij}\,A^{-1}_{jk} = \delta_{ik}$: $\Gamma^{(2)}$ and $W^{(2)}$ are therefore inverses of each other in a functional sense. The relation (22) simplifies when h is constant:

$$\int_y \Gamma^{(2)}(x-y,m)W^{(2)}(y-z,h) = \delta(x-z) \tag{23}$$

and even further when working in Fourier space:

$$\Gamma^{(2)}(p,m)\,W^{(2)}(p,h) = 1 \tag{24}$$

Thus, in a uniform field and in Fourier space, $\Gamma^{(2)}$ and $W^{(2)}$ are inverses in the ordinary sense of the term. Note that for O(N), the previous relation generalizes trivially to:

$$\Gamma^{(2)}_{\alpha\beta}(p,m)\,W^{(2)}_{\beta\gamma}(p,h) = \delta_{\alpha\gamma}. \tag{25}$$

[14] It is important to alert the beginner reader to a small subtlety when handling formulas involving the Legendre transform. Clearly, the definition of $\mathcal{Z}[\boldsymbol{h}]$ in Eq.(5) makes $\mathcal{Z}$ a functional of $\boldsymbol{h}(x)$, and the same obviously applies to $W[\boldsymbol{h}]$. The magnetization $\boldsymbol{m}(x)$ defined in Eq.(17) is therefore a derived quantity, obtained once $\boldsymbol{h}$ is specified: $\boldsymbol{m}(x)$ is a function of $\boldsymbol{h}(y)$. One can reverse the point of view and consider that it is $\boldsymbol{m}(x)$ that is fixed. The "correct" state function to consider is then Γ, and it is $\boldsymbol{h}(x)$ that becomes a function of $\boldsymbol{m}(y)$ via Eq.(20). Specifically, when taking a (functional) derivative with respect to either $\boldsymbol{h}(x)$ or $\boldsymbol{m}(x)$, one must consider that the conjugate variable, respectively $\boldsymbol{m}(x)$ and $\boldsymbol{h}(x)$, become functionals of the variable with respect to which we are differentiating. Thus, when differentiating Eq.(19) functionally with respect to $\boldsymbol{m}(x)$, one should not forget to differentiate $W[\boldsymbol{h}]$, on the grounds that it is a functional of $\boldsymbol{h}$ and not of $\boldsymbol{m}(x)$, and one must write: $\frac{\delta W[\boldsymbol{h}]}{\delta \boldsymbol{m}(x)} = \int_y \frac{\delta W[\boldsymbol{h}]}{\delta \boldsymbol{h}(y)}\frac{\delta \boldsymbol{h}(y)}{\delta \boldsymbol{m}(x)} = \int_y \boldsymbol{m}(y)\frac{\delta \boldsymbol{h}(y)}{\delta \boldsymbol{m}(x)}$.

[15] The terminology "connected" and "1PI" for the correlation functions $G_c^{(n)}$ and $\Gamma^{(n)}$ comes from their perturbative expansion in terms of Feynman diagrams. The connected functions only receive contributions from connected graphs, meaning graphs that are not made up of several disconnected subgraphs. The graphs contributing to $\Gamma^{(n)}$ are those that remain connected when any of their internal lines are cut, i.e., a line connecting two vertices.

where the Einstein summation convention over repeated indices has been used: $A_{\alpha\beta}B_{\beta\gamma} = \sum_\beta A_{\alpha\beta}B_{\beta\gamma}$.

Studying correlation functions consists in analyzing the information contained in the generating functionals (the free energies in particular) through Taylor expansions in powers of the fields. For example, for $N = 1$, if we perform an expansion around a constant field, h_0 for W and m_0 for Γ, we obtain:

$$W[h] = \sum_{n=0}^{\infty} \int_{x_1,\cdots,x_n} G_c^{(n)}(x_1,\cdots,x_n,h_0)\,(h(x_1)-h_0)\cdots(h(x_n)-h_0) \tag{26}$$

and

$$\Gamma[m] = \sum_{n=0}^{\infty} \int_{x_1,\cdots,x_n} \Gamma^{(n)}(x_1,\cdots,x_n,m_0)\,(m(x_1)-m_0)\cdots(m(x_n)-m_0)\,. \tag{27}$$

Another way to study the free energies is to perform an expansion in derivatives of the fields, that is, in powers of the moments in Fourier space. This will be very useful later. We will implement it on Γ, and this expansion reads for $N = 1$:

$$\Gamma[m] = \int_x \left\{ U(m(x)) + \frac{1}{2} Z\left(m(x)\right) \left(\nabla m(x)\right)^2 + O(\nabla^4) \right\} \tag{28}$$

where the function $Z\left(m(x)\right)$ must of course not be confused with the partition function $\mathcal{Z}[h]$. We note that for a constant field $m(x) = m_0$, $\Gamma(m_0) = \Omega U(m_0)$ where Ω is the volume of the system. The function U is called the effective potential in QFT and the thermodynamic potential (Gibbs potential) in statistical mechanics. It plays a crucial role in characterizing spontaneous symmetry breaking, for example. It is worth noting that by the isotropy of Euclidean space in d dimensions, only even powers of ∇ can appear in this derivative expansion, and that the functions U and Z are functions of m^2 only due to the $\mathbb{Z}_2$ symmetry. For a more complicated internal symmetry group, the derivative expansion must respect the fact that Γ is a scalar, and therefore must involve all scalars in the space derivatives and under the internal symmetry group. For example, for O(N) models, this expansion becomes:

$$\Gamma[\boldsymbol{m}] = \int_x \left\{ U(\boldsymbol{m}^2(x)) + \frac{1}{2} Z\left(\boldsymbol{m}^2(x)\right) \left(\nabla m_\alpha(x)\right)\left(\nabla m_\alpha(x)\right) + \frac{1}{4} Y\left(\boldsymbol{m}^2(x)\right) m_\alpha(x) m_\beta(x) \nabla m_\alpha(x) \nabla m_\beta(x) + O(\nabla^4) \right\} \tag{29}$$

and there are thus two independent functions appearing at the second order of the derivative expansion. Finally, note that $U^{(n)}(m_0) = \Gamma^{(n)}(p_1 = 0, \cdots, p_{n-1} = 0, m_0)$, so that the effective potential contains all the information about all the $\Gamma^{(n)}$ at zero momentum. Conversely, the $\Gamma^{(n)}$ contain all the information about the momentum dependence of a given order term in fields. Therefore, studying correlation functions or the functions $U, Z, \ldots$ are two different and complementary ways to analyze the information contained in the free energies of a system with n bodies.

2.2 Phase transitions

The O(N) models defined by the Hamiltonians of Eqs.(2) and (3) undergo a phase transition at $\boldsymbol{h} = 0$ and for all $d > 2$ when varying the temperature. For $N = 1$, there is a transition in $d = 2$ but not in $d = 1$. Strictly speaking, this occurs only in the thermodynamic limit: $L \to \infty$. These systems are in a disordered phase at high temperature where the magnetization is zero: $m_i = \langle \sigma_i \rangle = 0$ and in an ordered phase at low temperature where it is non-zero: $|m_i| = |\langle \sigma_i \rangle| > 0$. The magnetization is called the order parameter because its value, whether zero or non-zero, is characteristic of the phase in which the system resides. By definition, the temperature at which this transition occurs is called the critical temperature T_c. This temperature depends on d, N, the model considered, whether on

a lattice or in the continuum. If defined on a lattice, it depends on the structure of the lattice: in $d = 2$, for example, the critical temperature is different if we consider a square lattice or a triangular lattice. In the continuum, T_c depends on the coupling constant g_0 appearing in the Hamiltonian Eq.(3). For this reason, T_c is said to be non-universal: it depends on the microscopic details of the system.

For the models considered above, the transition is said to be continuous because the order parameter in the zero-field limit varies continuously as a function of temperature, although this variation is not analytic, see Fig. 2.2 where we can see how the non analyticity builds up as the systems size increases.

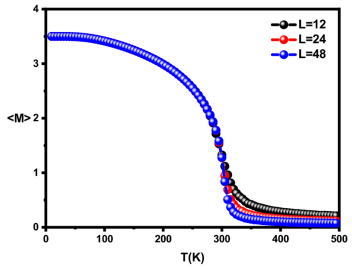

Figure 1: Magnetization of a magnetic system as a function of temperature for several lattice sizes.

In general, phase transitions are divided into two distinct types: continuous transitions as described above, and discontinuous transitions, or first-order transitions, where the order parameter undergoes a discontinuity at the transition. In fact, the important aspect is less whether the order parameter is continuous or not, but rather whether the correlation length diverges or remains finite at the transition. By definition, the correlation length ξ of a system is the typical length scale that gives the decay of the two-point correlation function. In $d = 3$ and at large distances compared to the lattice spacing, the two-point function behaves for $T \simeq T_c$ and $N = 1$ as:

$$G^{(2)}(x_i, x_j, h = 0) = \langle \sigma_i \sigma_j \rangle \underset{T \to T_c}{\sim} \frac{e^{-\frac{|x_i - x_j|}{\xi}}}{|x_i - x_j|^{d-2+\eta}} \tag{30}$$

where η is called the anomalous dimension of the field (we will discuss it later) and ξ is a function of temperature. For continuous transitions, ξ diverges at T_c and it is found that it diverges in a power-law manner. For example, approaching T_c in the high-temperature phase:

$$\xi(T) \underset{T \to T_c^+}{\simeq} \xi_0^+ (T - T_c)^{-\nu} \tag{31}$$

where ν is another critical exponent and ξ_0 is an amplitude that depends on the system considered. The divergence of the correlation length is the main feature of continuous transitions and will condition the scale invariance (and conformal invariance in many cases) that will appear at the transition.

Let us list some important properties of continuous transitions.

- **Scale invariance.** A statistical system such as the Ising model (and this is generically true) has two intrinsic scales: a microscopic scale, which in this case is the lattice spacing a, and a macroscopic length scale L which is its length, which in units of a, is an integer. The Hamiltonian in Eq. (1) does not contain any length scale other than a and L. In the thermodynamic limit, $L \to \infty$, only a remains. The dynamics of the system produces another scale, ξ, defined in Eq. (30). It is therefore expected generically that this scale is of the order of a, and indeed this is the case. But by finely tuning T so that $T \simeq T_c$, one finds that ξ can be arbitrarily

large compared to a, which is very unusual in physics. One can suspect that the physics of the system will be dominated by ξ, which already controls the behavior of the two-point function when the distances considered are large compared to a. In other words, close to the transition and at large distances, $G^{(2)}$ depends on ξ and no longer on a, which is still the natural scale of the problem: a no longer plays a role for long-distance correlations. Therefore, exactly at T_c, where $\xi = \infty$, and again for large distances compared to a, the system no longer presents any characteristic scale, and one expects the emergence of scale symmetry (invariance under dilation) at large length scales. In fact, the story is even more beautiful in many cases because the system becomes conformally invariant at T_c and at large scales, not just invariant under dilation.

- **Power-law behaviors.** Scale-invariant systems are generically characterized by power laws, the only laws that are consistent with this symmetry.[16] We indeed observe this phenomenon for phase transitions: near criticality, various interesting physical quantities obey power laws, as we have already seen for the correlation length in Eq.(31) and the two-point function when $\xi = \infty$, Eq. (30). The magnetization, for example, behaves when approaching T_c from below as:

$$m(T) \underset{T \to T_c^-}{\simeq} m_0 \, (T_c - T)^{\beta}. \tag{32}$$

The magnetic susceptibility at zero field, χ, which is the response of the magnetization to a variation in the magnetic field, behaves as:

$$\chi(T) = \frac{\partial M}{\partial h}_{|h=0} \underset{T \to T_c^+}{\simeq} \chi_0^+ \, (T - T_c)^{-\gamma}. \tag{33}$$

We also find that at $T = T_c$:

$$M(T) \underset{h \to 0}{\simeq} m_0' \, h^{1/\delta}. \tag{34}$$

The specific heat also diverges as a function of $T - T_c$ with an exponent called α, and we have already defined the anomalous dimension η in Eq.(30). In total, there are six critical exponents, called dominant. Other exponents, called subdominant or corrections to scaling laws, regulate the subdominant behaviors for $T - T_c$. For example, we can define the correction to scaling exponent ω by:

$$\xi(T) \underset{T \to T_c^+}{\simeq} \xi_0^+ (T - T_c)^{-\nu} \, (1 + A(T - T_c)^{\omega} + \cdots) \tag{35}$$

where A is an amplitude that depends on the system considered.

It should be noted that for certain quantities, critical exponents can be defined in both the high and low-temperature phases. For example, for the correlation length in the low-temperature phase:

$$\xi(T) \underset{T \to T_c^-}{\simeq} \xi_0^- \, (T_c - T)^{-\nu'}. \tag{36}$$

Exponents in the high and low-temperature phases, when they exist, are often the same (although the amplitudes are not). However, this is not always true (contrary to what is often stated).

[16] Here, I am cheating a little because the power laws shown below refer to the temperature dependence and not the distance dependence, which are the objects of dilations. But on further thought, we will see that we can substitute the deviation from the critical temperature, $T - T_c$, with the correlation length and thus transform the power laws in temperature into power laws in ξ. However, one must still remember that for $\xi < \infty$, there is not strictly speaking dilation invariance...

- **Universality.** We are very accustomed in physics to seeing very different systems at the microscopic level behave in the same way at large scales. The ideal gas law, Ohm's law, etc, are valid regardless of the nature of the substances considered as long as they satisfy certain prerequisites. For example, for the ideal gas law to apply, it is enough for the gas atoms to be considered point-like and interact only through perfectly elastic collisions. It doesn't matter whether it is oxygen or carbon dioxide, or whether they are made up of electrons, protons, and neutrons, which themselves are made up of quarks interacting via gluon exchange. This property, foundational to all of physics, comes from the fact that, most often, in terms of collective behaviors and large scales, the microscopic details have been "statistically averaged out" and only general characteristics of the systems, described by a small number of quantities, are relevant at these scales. The physics at short distances can thus be horribly complicated, even chaotic, but the stochastic behaviors that emerge are usually simple. Thus, from Avogadro's number of degrees of freedom, needed to describe a gas molecule by molecule, emerge a few quantities such as pressure, volume, and temperature in its large-scale description. From the unpredictable, chaotic microscopic level emerges the quasi-determinism of thermodynamic quantities.

  Of course, the central limit theorem plays a considerable role here by telling us that when summing independent random variables, we obtain Gaussian laws that have (almost) completely forgotten the underlying probability distribution. And the greater the number of summed variables, the more precise the prediction of the behavior of the sum becomes in relative value. Universality is the name of this phenomenon and has nothing to do with phase transitions (contrary to what is generally suggested). The surprise is therefore not that there are universal phenomena in physics: they are everywhere. The surprise is that even critical phenomena are subject to this.[17] The central limit theorem is indeed valid for independent random variables whose mean and variance exist. For Ising away from criticality, the spins are certainly correlated over typical distances of ξ, but since $\xi \simeq a$, their correlations are very weak, and the system appears as a collection of clusters of typical size ξ, largely independent of each other, even though the spins inside the clusters are not independent. For this reason, it is very easy to generalize the central limit theorem to statistical systems away from criticality, and universality works in this case in a "trivial" way.

  This is not the case for systems at criticality, where all the spins are strongly correlated because $\xi = \infty$. Therefore, it is difficult to know *a priori* whether any universality will survive this situation, and if so, what the "universality classes" will be, that is, the set of systems for which certain quantities have the same behavior. The surprise is therefore here: yes, despite the strong correlations, there is universality in critical phenomena in the sense that for certain classes of systems and for certain quantities, identical large-scale behaviors emerge in microscopically different systems. The most well-known universal quantities are the critical exponents, but there are many others, such as amplitude ratios between the high and low-temperature phases, such as ξ_0^+/ξ_0^-. The prediction of this universality is one of the great successes of RG (inaccessible, for example, to integrable models[18]), even though it is not easy to predict *a priori* the contours of a universality class, already because it is not easy to predict in general whether a given system undergoes a continuous phase transition or not.[19] The

[17] The surprise comes from the major characteristic of these phenomena: the divergence of the correlation length, which might (erroneously) suggest that since all degrees of freedom are coupled, no detail of the system can be neglected.

[18] Having an exact solution for a given model such as the ferromagnetic Ising model in $d = 2$ on a square lattice with nearest-neighbor interaction says nothing about another model, integrable or not, even if it is also invariant $\mathbb{Z}_2$: integrability says nothing about universality and a fortiori nothing about the contours of a universality class. One might think that conformal techniques would escape this criticism, since they are only concerned with symmetry properties of models, but it is particularly difficult with this technique to link a given model with a given set of conformal data.

[19] We often read that universality classes are determined by the space dimension d, the symmetry of the system

notion of RG fixed point is what best explains universality in critical phenomena, and it is not surprising that this concept is intimately linked to the generalization of the central limit theorem to strongly correlated systems.

- **Spontaneous symmetry breaking and ergodicity breaking.** Quite often, but not always[20], phase transitions are associated with spontaneous symmetry breaking. For the Ising model, this is the complete breaking of the $\mathbb{Z}_2$ symmetry, and for the $\mathrm{O}(N)$ model defined above, it is $\mathrm{O}(N)$ broken into $\mathrm{O}(N-1)$. This symmetry breaking can only occur in the thermodynamic limit because for any finite volume, even if the majority of spins on average point in the same direction, there is always a non-zero probability that a fluctuation will simultaneously change the direction of more than half of the spins. Of course, this probability vanishes as the number of spins tends to infinity, so that in the low-temperature phase, when the system has acquired a non-zero spontaneous magnetization, this magnetization can no longer change direction. Therefore, in these situations, there is also spontaneous ergodicity breaking because the system can no longer explore all the states that are, in principle, accessible to it.

  For systems exhibiting spontaneously broken continuous symmetry at the transition, Goldstone's theorem predicts the existence of modes whose susceptibility is infinite throughout the low-temperature phase (so-called massless modes). The theorem predicts that there will be as many as there are "broken generators," that is, as many as the difference between the number of generators of the group $\mathcal{G}$ of the Hamiltonian (symmetry of the high-temperature phase) and the number of generators of the group $\mathcal{H}$ of the system's ground state (symmetry of the low-temperature phase). In the case of $\mathrm{O}(N)$ broken into $\mathrm{O}(N-1)$, there are thus $N(N-1)/2-(N-1)(N-2)/2=N-1$ Goldstone modes in the low-temperature phase.[21]

- **Upper and lower critical dimensions.** It is observed, and this will be explained later, that for the Ising and $\mathrm{O}(N)$ models, the critical exponents, and more generally the universal quantities, are correctly predicted by mean-field theory. This dimension, called the upper critical dimension, d_c, is four for these models.[22] Note that the notion of upper critical dimension only makes sense for universal quantities. For non-universal quantities, nothing in particular happens when crossing dimension four, and they remain non-exactly determined in all dimensions

and whether or not the interactions are short-range. This implies that any system with the same symmetry, the same dimension and, for example, short-range interactions all share the same set of critical exponents. This is incorrect for several reasons. Firstly, and most fundamentally, two such systems can undergo phase transitions, one continuous and the other first-order: no universality here, of course. A simple example of this situation is given by a theory in the continuum, in $d=3$, with a field $\phi(x)$, of symmetry $\mathbb{Z}_2$ and whose Hamiltonian has a ϕ^4 term with a very negative coupling constant and a ϕ^6 term with a small positive coupling constant. Although this model has $\mathbb{Z}_2$ symmetry, it undergoes a first-order transition and is therefore obviously not in Ising universality class. Secondly, the symmetry of a Hamiltonian is not everything. A given Hamiltonian can have different types of order in its low-temperature phase (different spontaneous symmetry breaking schemes) depending on the value of its coupling constants. The Hamiltonian can therefore be invariant under a group $\mathcal{G}$, independent of the values of the couplings, and the low-temperature phases, i.e. the fundamental states of H (which depend on the couplings) can be invariant under various groups $\mathcal{H}_i$ with therefore different types of phase transition that are in general not in the same universality class. Instead of talking about the symmetry $\mathcal{G}$ of the model, it is therefore better to talk about the symmetry-breaking scheme: $\mathcal{G}\to\mathcal{H}_1$ or $\mathcal{G}\to\mathcal{H}_2$ or $\mathcal{G}\to\cdots$. Furthermore, the same system can undergo critical, tricritical and, more generally, multi-critical transitions. Here again, the critical exponents are not the same, depending on the degree of multi-criticality considered. Finally, it is not out of the question for the same system to have several RG critical fixed points, each with its own basin of attraction and therefore different universality classes. While this scenario is generally perturbatively inaccessible, it is certainly no proof of its impossibility.

[20] For example, in the liquid-vapor transition, which is generically first-order but whose first-order line ends with a second-order transition, there is no symmetry breaking.

[21] One must be careful not to over-interpret this result. It is possible that other modes are also massless in the low-temperature phase. This is the case for the $\mathrm{O}(N)$ model, where the last mode is also massless due to its interactions with the Goldstone modes. The mean-field analysis is wrong on this point because fluctuations must be taken into account for this effect to appear.

[22] In $d=4$, there are logarithmic corrections to mean-field theory, making it not entirely accurate in this dimension for universal quantities.

greater than four by the mean-field approximation. In fact, it is only in $d = \infty$ that mean-field theory becomes exact, including for non-universal quantities.

There also exists for the O(N) models a lower critical dimension, below which there is no longer a phase transition, or more precisely, where the phase transition temperature becomes zero. For the Ising model, this lower critical dimension is one, and for $N > 1$, it is two. The difference between these two cases arises from the fact that Ising symmetry is a discrete symmetry, while for $N > 1$, it is a continuous symmetry.[23] The presence of a lower critical dimension for O($N > 1$) is responsible for the existence of a perturbative theory in $d = 2 + \epsilon$ for these models, performed starting from the nonlinear sigma model, with the case $N = 2$ being once again very special.

2.3 Mean-field approximation

"*Ce qui est simple est faux, ce qui est compliqué est inutilisable.*"
Paul Valéry

The entire difficulty of studying statistical systems such as O(N) models lies in the presence of the very large number of degrees of freedom involved. This difficulty is even greater near a continuous transition because, in this case, all degrees of freedom are strongly correlated, and thus the system cannot *a priori* be considered as an assembly of small subsets essentially independent of each other.

The first approximation one might think of, the mean-field approximation in statistical mechanics or the classical approximation in QFT, consists of approximating the summation over the system's configurations in the partition function in a way that makes it calculable. There are many ways to achieve this. In all cases, the summation is truncated, which amounts to neglecting configurations of the system that are expected to play a subdominant role compared to the one(s) retained.

A possible formulation of the mean-field (MF) approximation for ϕ^4 theory consists of applying a saddle-point approximation to the functional integral in $\mathcal{Z}[h]$, keeping only the configuration that contributes the most to the integrand, i.e., the one that minimizes the Hamiltonian. In the remainder of this section, we will consider only the case where the external magnetic field is constant and, for simplicity of notation, restrict ourselves to $N = 1$, as the generalization to $N > 1$ is straightforward. The configuration ϕ_0 that minimizes the Hamiltonian at fixed h is assumed to be independent of x since h is constant and the gradient term in H is positive, thus only increasing the energy. We then obtain:

$$\mathcal{Z}(h) \to \mathcal{Z}^{\mathrm{MF}}(h) = e^{-H(\phi_0, h)} \tag{37}$$

where ϕ_0 is defined by

$$\frac{\partial H(\phi, h)}{\partial \phi}_{|\phi=\phi_0} = 0 \tag{38}$$

which defines ϕ_0 as a function of h.

Let us show that at this level of approximation, ϕ_0 is nothing more than the system's magnetization. To do this, let us calculate the magnetization from Eq. (37):

$$M = m\Omega = \frac{dW}{dh} = -\frac{\partial H(\phi, h)}{\partial \phi}_{|\phi_0} \frac{\partial \phi_0}{\partial h} - \frac{\partial H(\phi, h)}{\partial h}_{|\phi_0} = \Omega\phi_0 \tag{39}$$

[23] The Mermin-Wagner and Coleman theorem ensures that for systems in thermodynamic equilibrium with short-range interactions, a continuous symmetry cannot be broken in $d \leq 2$. Note that the case $N = 2$ and $d = 2$ is special because it features a very peculiar finite-temperature transition induced by the presence of topological defects, the famous Kosterlitz-Thouless transition.

where $\Omega = L^d$ is the system's volume. It follows that $m = \phi_0$, as expected, because the spin configuration corresponding to the magnetization is certainly the one that contributes the most to $\mathcal{Z}$.[24] From this, we deduce the expression for Γ at this level of approximation:

$$\mathcal{Z}^{\mathrm{MF}}(h) = e^{-H(\phi_0,h)} = e^{-\Gamma(m)+\Omega m h} \tag{40}$$

which implies the fundamental result:

$$\Gamma^{\mathrm{MF}}(m) = \Omega V(\phi = m) \tag{41}$$

where $V(\phi)$ is the potential appearing in H, also called the bare potential. It is the part independent of the field h and non-derivative, while the derivative part is referred to as the kinetic term. For the ϕ^4 theory, $V(\phi(x)) = r_0\phi^2(x)/2 + g_0\phi^4(x)/4!$.

The previous result generalizes very easily to the case where one considers a non-constant field $h(x)$, in which case we find:

$$\Gamma^{\mathrm{MF}}[m] = H[\phi(x) = m(x), h(x) = 0] \tag{42}$$

that is to say, Γ is nothing but the part of the Hamiltonian independent of the magnetic field. This result is crucial for several reasons:

- Discussions about a system's thermodynamics are often based on the system's Hamiltonian (or Lagrangian in QFT). This is *a priori* incorrect because the statistical physics of a system is determined by its free energies, not its Hamiltonian. However, the previous relation shows that if the mean-field approximation is a good approximation, i.e., if fluctuations around the mean field are small enough, then $\Gamma \simeq \Gamma^{\mathrm{MF}} = H$, and such discussions are valid. A striking example of this type of misuse is the discussion of symmetry breaking in a system based on the Hamiltonian, particularly on the double-well shape (or "Mexican hat") of the bare potential V. Such discussions are generally incorrect because what matters for symmetry breaking is not V but the effective potential U defined in Eqs.(28) and (29), as discussed later.[25]

- Equation (42) reveals a subtle issue: while H may very well be a non-convex function of ϕ, Γ must be convex with respect to m. For instance, if $r_0 < 0$ in H, Eq.(3), the potential for $N = 1$ takes the form of a double well with minima at $\pm\sqrt{-6r_0/g_0}$ and a local maximum at 0 (see Fig. 2). However, U can only equal V in its convex region, i.e., outside the interval $[-\sqrt{-6r_0/g_0}, +\sqrt{-6r_0/g_0}]$. Thus, at the mean-field level, the effective potential U is represented by taking the convex envelope of V, i.e., drawing a horizontal line between the two minima. Note that this part of the effective potential is physically inaccessible.

- The mean-field approximation predicts spontaneous magnetization for $r_0 < 0$ because, in this case, $h = U'(m) = V'(m)$ (outside the non-convex region) has two non-zero solutions $\pm\sqrt{-6r_0/g_0}$ as $h \to 0^{\pm}$. This corresponds to spontaneous symmetry breaking, correctly reproduced at the mean-field level for generic dimensions d. However, a transition is also predicted in $d = 1$, which is incorrect. For $\mathrm{O}(N > 1)$ models, the same would be true for $d \leq 2$, which is also incorrect, indicating that fluctuations around the mean field are large enough in low dimensions to destabilize its predictions, even qualitatively.

[24] The true magnetization of the system, not its mean-field approximation, is different from ϕ_0, as fluctuations around the mean-field configuration modify it. In ferromagnetic models, these fluctuations tend to reduce the magnetization compared to the mean-field prediction.

[25] There are cases where the bare and effective potentials differ significantly, making the mean-field analysis qualitatively incorrect. For example, mean-field theory may predict a second-order phase transition, whereas the actual transition is first-order, or vice versa. It may even predict no phase transition, while a second-order transition exists in reality. Examples include the 3-state Potts model in $d = 2$, frustrated magnetic systems in $d = 3$ on stacked triangular lattices, branching annihilating random walks, etc.

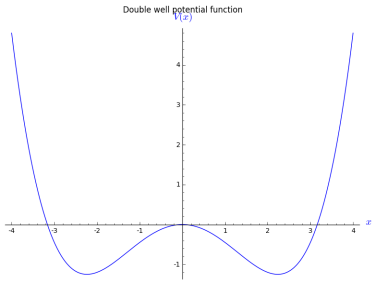

Figure 2: Double-well bare potential. At the mean-field level, this potential is identical to the effective potential in the region where it is convex, i.e., outside the region between its two minima.

- The mean-field approximation is exact in $d = \infty$. Note here that for dimensions above the upper critical dimension $d_c = 4$, universal quantities are correctly obtained using the mean-field approximation (see below). However, non-universal quantities are not: they generally continue to receive significant corrections from fluctuations, often making their mean-field approximation very poor. Only in infinite dimensions do these fluctuations disappear.[26]

One can go a bit further in the analysis of the mean-field approximation by calculating $G^{(2)}$, $\Gamma^{(2)}$, and $\Gamma^{(4)}$. The calculation of $\Gamma^{(2)}(p, m)$ in the presence of a constant magnetization m is immediate within the mean-field approximation using Eq.(42). We start by computing the first functional derivative of Γ:

$$\begin{aligned}\Gamma^{(1)}_{\alpha}[x_1, \boldsymbol{m}] &= \frac{\delta\Gamma}{\delta m_\alpha(x_1)} \\ &= \int_x \left\{\nabla m_\alpha(x)\cdot\nabla\delta(x-x_1) + r_0 m_\alpha(x)\delta(x-x_1) + \frac{g_0}{6}\boldsymbol{m}^2(x)m_\alpha(x)\delta(x-x_1)\right\} \\ &= \left(-\nabla^2 + r_0 + \frac{g_0}{6}\boldsymbol{m}^2(x_1)\right) m_\alpha(x_1).\end{aligned} \tag{43}$$

We then compute the second functional derivative and evaluate the result in a configuration with constant magnetization: $m_\alpha(x) = m_\alpha$:

$$\begin{aligned}\Gamma^{(2)}_{\alpha\beta}(x_1 - x_2, \boldsymbol{m}) &= \frac{\delta^2\Gamma}{\delta m_\alpha(x_1)\delta m_\alpha(x_2)}_{|\boldsymbol{m}(x)=\boldsymbol{m}} \\ &= \delta_{\alpha\beta}\left(-\nabla^2 + r_0 + \frac{g_0}{6}\boldsymbol{m}^2\right)\delta(x_1-x_2) + \frac{g_0}{3}m_\alpha m_\beta\delta(x_1-x_2).\end{aligned} \tag{44}$$

As expected, $\Gamma^{(2)}_{\alpha\beta}$ is a linear combination of the two (symmetric) rank-two tensors of $\mathrm{O}(N)$ that can be constructed with the vector $\boldsymbol{m}$, namely $\delta_{\alpha\beta}$ and $m_\alpha m_\beta$. In Fourier space, this becomes: [27]

$$\Gamma^{(2)}_{\alpha\beta}(p, \boldsymbol{m}) = \delta_{\alpha\beta}\left(p^2 + r_0 + \frac{g_0}{6}\boldsymbol{m}^2\right) + \frac{g_0}{3}m_\alpha m_\beta. \tag{45}$$

[26] From a technical standpoint (discussed later), it is the long-wavelength, or infrared, fluctuations that determine universal behaviors and, consequently, the value of d_c. Conversely, non-universal quantities depend on the microscopic details of the model and are thus affected by short-wavelength, or UV, fluctuations. These fluctuations are generally not suppressed as the dimension increases beyond d_c, so there is no reason for non-universal quantities to be well approximated in $d > d_c$.

[27] Note that we have given the same name to $\Gamma^{(2)}_{\alpha\beta}(x_1 - x_2, m)$ and $\Gamma^{(2)}_{\alpha\beta}(x_1, x_2, m)$ even though the latter depends on one additional argument. Its Fourier transform thus depends on two momenta and is proportional to a delta of total momentum conservation, as well as $(2\pi)^d$. When we consider a situation where the system is translationally invariant and in accordance with our convention set out in Eq.(14), we will only use the Fourier transform of the function $\Gamma^{(2)}_{\alpha\beta}(x_1 - x_2, m)$, which depends on a single momentum.

From Eq. (25), the function $G^{(2)}_{c,\alpha\beta}(p, \boldsymbol{h}=0)$ is the inverse of $\Gamma\alpha\beta^{(2)}(p, \boldsymbol{m})$ with $\boldsymbol{m}$ the magnetization at zero field, i.e., $\boldsymbol{m}=0$ in the high-temperature (symmetric) phase and $|\boldsymbol{m}| = \pm\sqrt{-6r_0/g_0}$ in the low-temperature (broken) phase:[28]

$$G^{(2)}c, \alpha\beta(p, h=0) = \frac{\delta\alpha\beta}{p^2 + r_0} \qquad \text{for} \quad T > T_c \tag{46}$$

and

$$G^{(2)}_{c,\alpha\beta}(p, h=0) = \frac{\delta_{\alpha\beta}}{p^2} - \frac{2|r_0|}{p^2(p^2+2|r_0|)}\frac{m_\alpha m_\beta}{\boldsymbol{m}^2} \quad \text{pour} \quad T < T_c\ . \tag{47}$$

Without loss of generality, we can take $\boldsymbol{m}$ in the direction 1, i.e., $m_\alpha = m\delta_{\alpha 1}$. In this case, in the broken phase ($T < T_c$):

$$\begin{aligned} G^{(2)}_{c,11}(p, h=0) &= \frac{1}{p^2 + 2|r_0|} \\ G^{(2)}_{c,22}(p, h=0) &= \frac{1}{p^2} \end{aligned} \tag{48}$$

These are called longitudinal susceptibility and transverse susceptibility, respectively (of course, the direction 22 is not special, and all directions 33, $\cdots$, NN play the same role). In the symmetric phase, these two susceptibilities are naturally identical.

Recall that a correlation function proportional to $1/(p^2 + \xi^{-2})$ implies an exponential decay at large distances as $\exp(-r/\xi)$, which identifies ξ as the correlation length. Thus, in the high-temperature phase, $\xi = r_0^{-1/2}$, while in the low-temperature phase, the transverse directions have an infinite correlation length (this is the Goldstone theorem), and the longitudinal direction allows the definition of a correlation length $\xi = (2|r_0|)^{-1/2}$. Assuming r_0 is a regular function of T and is thus expandable, it must be proportional to $T - T_c$ asymptotically close to the transition. This allows us to conclude that the critical exponent $\nu = 1/2$ within the mean-field approximation and that it takes the same value in the low-temperature phase.[29] We also note that the amplitudes $\xi_0^{\pm}$ of ξ are not identical and that their ratio is $1/\sqrt{2}$, a universal number, i.e., here independent of g_0.

2.4 Free Propagator and Perturbation Theory

Let us show in detail that the function $G^{(2)}$ is the Green function of the differential operator appearing in the quadratic part of H. To do so, let us rewrite this quadratic part:

$$\begin{aligned} H^{\text{quad.}} &= \int_x \phi(x)\left(-\nabla^2 + r_0\right)\phi(x) \\ &= \int_{xy} \phi(x)\left[\left(-\nabla_x^2 + r_0\right)\delta(x-y)\right]\phi(y) \\ &= \int_{xy} \phi(x) A(x,y) \phi(y) \end{aligned} \tag{49}$$

with A the differential operator: $A(x,y) = \left(-\nabla_x^2 + r_0\right)\delta(x-y)$. Equation (49) generalizes to the continuous case the discrete sum $\sum_{ij}\phi_i A_{ij}\phi_j$ where the ϕ_i would be the components of a vector in

[28] The inverse in O(N) space is straightforward to calculate if one recalls that $G^{(2)}c, \alpha\beta$ is also an O(N) tensor and can thus be decomposed on the tensors $m_\alpha m_\beta$ and $\delta_{\alpha\beta}$.

[29] A word of caution is necessary here. For $N = 1$, there are no transverse directions, and the calculation above, although quantitatively incorrect because ν receives corrections for $d < 4$, is qualitatively correct. For $N > 1$, however, it suggests that $N - 1$ modes have an infinite correlation length but not the last one. In fact, the coupling between the transverse modes and the longitudinal mode also makes the latter have an infinite correlation length, and the mean-field approach is qualitatively incorrect here. Nevertheless, one can still define for $T < T_c$ a length, called the Josephson length, that separates Goldstone-type behavior from critical behavior, and for O(N) models, this length diverges at the transition with an exponent equal to ν, the same as in the high-temperature phase.

an n-dimensional space and A_{ij} the matrix elements of a diagonal matrix. Provided that none of the A_{ii} are zero, this matrix is invertible, and we obviously have $\sum_j A_{ij}A^{-1}_{jk} = \delta_{ik}$. In the continuous case, the relation between the operator A and its inverse is written:

$$\int_y A(x,y)A^{-1}(y,z) = \delta(x-z). \tag{50}$$

If A is a translation-invariant operator and can thus be written (abusively keeping the same name): $A(x,y) = A(x-y)$, then in Fourier transform, $A^{-1}(q) = 1/A(q)$. This applies to the operator A in Eq. (49), and the inversion relation (50) takes the form:

$$\begin{aligned}\delta(x-z) &= \int_y \left[\left(-\nabla_x^2 + r_0\right)\delta(x-y)\right] A^{-1}(y-z) \\ &= \left(-\nabla_x^2 + r_0\right)\int_y \delta(x-y)A^{-1}(y-z) \\ &= \left(-\nabla_x^2 + r_0\right) A^{-1}(x-z)\end{aligned} \tag{51}$$

which, in Fourier space, becomes:

$$(p^2 + r_0)A^{-1}(p) = 1 \tag{52}$$

and we thus see that $A^{-1}(p) = G^{(2)\mathrm{MF}}(p)$. We conclude that in the mean-field approximation, the function $G^{(2)}$ is the inverse of the quadratic part of the Hamiltonian. It is called the free propagator of the theory (free because, in QFT, a Gaussian theory describes free particles). This function is one of the essential ingredients of perturbation theory, as it consists of an expansion of $\mathcal{Z}[h]$ in powers of g_0 around the Gaussian theory:

$$\mathcal{Z}[h] = \int D\phi \left(1 + g_0\int_{x_1}\phi^4(x_1) + \frac{1}{2}g_0^2\int_{x_1x_2}\phi^4(x_1)\phi^4(x_2) + \cdots\right) e^{-H_G[\phi]+\int h\phi} \tag{53}$$

where $H_G[\phi]$ is the Gaussian part of the Hamiltonian at zero field: $H_G[\phi] = 1/2\int_x[(\nabla\phi(x))^2 + r_0\phi^2(x)]$.

The Gaussian functional integral is the only one that can be computed in any dimension:

$$\int D\phi, e^{-\frac{1}{2}\int_x \phi(x)A(x)\phi(x)+\int_x h(x)\phi(x)} \propto \frac{1}{\sqrt{\det A}}\, e^{\frac{1}{2}\int_x h(x)A^{-1}(x)h(x)} \tag{54}$$

and, of course, from this expression, one can compute the expectation value of an arbitrary product of the field by functional differentiation with respect to $h(x)$. Thus, we can compute any term such as:

$$\int D\phi\phi(x_1)\cdots\phi(x_n)e^{-H_G[\phi]+\int h\phi} \tag{55}$$

and therefore all the terms of the perturbative expansion.[30] From the above expression, we see that it involves the inverse of the operator A which comes from the quadratic part of H and which is the free propagator, highlighting its crucial importance in perturbation theory.

[30]What remains difficult in perturbation theory is not the computation of the functional integral but rather the computation of "loop integrals," i.e., the integrals over the points $x_1, x_2, \cdots$. There are as many such integrals (each being in d dimensions) as there are independent loops in the Feynman diagrams representing a given contribution to the perturbative expansion.

3 The Functional Renormalization Group: Introduction and Definitions

"*L'ignorance est la nuit qui commence l'abîme*"
V. Hugo

What is systematically neglected in the mean-field approximation, within the functional integral, is the contribution of field configurations other than the one corresponding to its mean value. We will call fluctuations around the mean field the contribution of these configurations. Going beyond the mean field thus consists of reincorporating these fluctuations that were neglected at the mean-field level in order to eventually compute the free energies W or Γ.

The idea of Wilsonian renormalization in its modern version is therefore to construct a trajectory starting from the mean-field situation where fluctuations have been neglected, and reaching the free energy Γ.

Historically, it was rather W than Γ that was targeted as the final point of the trajectory after integrating out fluctuations. Of course, in principle, there is not much difference between obtaining Γ instead of W, since one can reconstruct one from the other via a Legendre transformation. However, in practice, the implementation of approximations does not lead to the same results at all, depending on which quantity they are applied to: in this case, the equivalence between the two formulations no longer holds. Moreover, as we now know—but which was not understood until quite recently— Γ is a much better starting point for implementing approximations that go beyond perturbation theory than W.[31]

The central idea of Wilsonian renormalization is to organize the summation over fluctuations scale by scale in momentum space. In real space, this corresponds to Kadanoff's block spin method: at the lattice scale, one starts with the initial Hamiltonian and the corresponding Boltzmann weight, then integrates out the fluctuations over a block of spins to obtain a new Hamiltonian, from which a new Boltzmann weight follows for the spin configurations that have not yet been integrated out. The process is then iterated until all configurations of the system have been integrated out. This generates a flow of Hamiltonians that depends on the scale via the size of the blocks, linking the initial Hamiltonian to W. Integrating this flow therefore means computing W as a function of H, which amounts to solving the model.

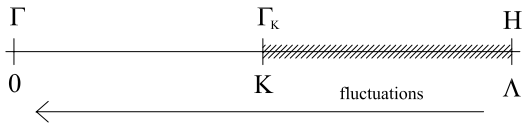

Figure 3: Integration of fluctuations in Γ_k. Fluctuations of large wavelength, i.e., small momentum compared to k, are frozen, and those of short wavelength are integrated without modification into Γ_k.

The idea with a flow of Γ (rather than H or W) is very similar: it involves starting from the Hamiltonian, which is defined at the lattice scale, and successively integrating, scale by scale in

[31] In fact, the relative neglect of Wilsonian methods stems from this: for a long time, the conceptual framework proposed by Wilson was very appealing, but the practical results obtained did not live up to expectations. That this was due to the practical implementation of the chosen approximation schemes, and more importantly, that it was possible to remedy this simply, was not at all clear and actually took decades to be understood and overcome through the use of Γ. In the meantime, proponents of perturbative methods had made their peace with it: one was thinking *à la* Wilson but was calculating *à la* Feynman.

momentum space, over the fluctuations. Since $H = \Gamma^{\mathrm{MF}}$ and the mean-field approximation consists of neglecting all configurations in the functional integral except for one, at the lattice scale, it involves freezing all fluctuations such that

$$H = \Gamma^{\mathrm{MF}} \qquad \text{is replaced by} \qquad H = \Gamma_\Lambda \tag{56}$$

where Λ is the inverse of a microscopic length scale used to define H, for example, a remnant of the lattice spacing for the Ising model. At scale Λ, the idea is to replace $H = \Gamma$, which is just an approximate relationship valid only at the mean-field approximation, by the relationship $H = \Gamma_\Lambda$, by cleverly freezing the fluctuations at this scale such that this relationship becomes exact. The idea is then to proceed like the block spins, progressively integrating the fluctuations in a Gibbs free energy Γ_k, depending on a scale k, until reaching Γ at $k = 0$ where, by definition, all fluctuations will have been integrated: $\Gamma_{k=0} = \Gamma$. Thus, we will have created a flow between $H = \Gamma_{k=\Lambda}$ and $\Gamma = \Gamma_{k=0}$, where, at an intermediate scale $\Lambda > k > 0$, only fluctuations with wavelengths between $a = \Lambda^{-1}$ and the running scale $l = k^{-1}$ have been integrated out, which is in the spirit of block spins.

To summarize, the idea is to construct a family of free energies Γ_k with a parameter interpolating between the Hamiltonian for $k = \Lambda$ and the free energy Γ that we aim to calculate at $k = 0$. To do this, the idea is to freeze at scale k the fluctuations with momenta smaller than k, i.e., the fluctuations with large wavelengths compared to k^{-1}: we integrate the "fast" modes and freeze the "slow" modes. The family of Γ_k must therefore satisfy:

- $\Gamma_{k=\Lambda} = H$
- $\Gamma_{k=0} = \Gamma$
- Γ_k is the free energy of the fast modes (relative to k) that have been integrated out.

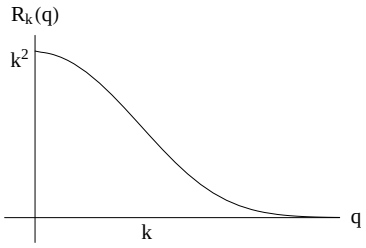

Figure 4: Typical form of the regulator $R_k(q)$.

We thus need to modify our model to freeze the slow modes at scale k. This is simple to do because, in a statistical sense, a mode is frozen when it no longer contributes to the partition function, implying that its correlation length is small. Indeed, consider a mode whose correlation length is zero. Its propagator also vanishes (this can be seen at the mean-field level in Eq.(46)) and it no longer participates in the system's dynamics. To modify its correlation length, we see from Eq.(46) that we must change the value of r_0 by making it very large. Physically, this is equivalent to placing the system at very high temperature. In this case, the system is completely disordered (the thermal energy kT becomes much larger than the magnetic energy J in lattice models like Ising and O(N)), and it behaves like a system of spins that do not interact with each other: it is no longer truly a many-body system but a collection of independent one-body systems for which the mean-field approximation becomes exact.

The problem with the previous idea – changing the temperature of the entire system – is that it treats all fluctuation modes, whether fast or slow, equally. However, we want to freeze only the slow modes (relative to k). The remedy is to assign a large r_0 only to these modes while leaving the fast modes unchanged. Therefore, we modify the initial Hamiltonian by adding a term that only freezes the slow modes, i.e., giving them a "large mass" via a quadratic term in the field that is nearly zero for the fast modes:

$$H[\phi] \;\rightarrow\; H_k[\phi] = H[\phi] + \Delta H_k[\phi] = H[\phi] + \frac{1}{2}\int_q R_k(q)\phi(q)\phi(-q) \tag{57}$$

where we have written $H_k[\phi]$ in Fourier space for convenience, and where $R_k(q)$, called the regulator or cut-off function, has a typical form given in Fig. 4.[32]

We note the following two properties of R_k:

- $R_{k=0}(q) \equiv 0$ such that $H_{k=0}[\phi] = H[\phi]$
- $R_{k=\Lambda}(q) \sim \Lambda^2$ such that the term $\Delta H_{k=\Lambda}$ is very large compared to the other terms in H.

The partition function associated with this model is

$$\mathcal{Z}k[h] = \int D\phi, e^{-H_k[\phi]+\int_x h(x)\phi(x)} \tag{58}$$

Note the change in notation compared to what was previously used: now, the magnetic field term is explicitly separated from the Hamiltonian H_k, which no longer depends on it. The Legendre transform of $W_k[h] = \log \mathcal{Z}k[h]$, denoted $\Gamma_k^{\text{Leg}}[m]$, is

$$\Gamma_k^{\text{Leg}}[m] + W_k[h] = \int_x h(x)m(x). \tag{59}$$

It is convenient for the following to rewrite Γ_k^{Leg} in the form

$$\Gamma_k^{\text{Leg}}[m] = \Gamma_k[m] + \frac{1}{2}\int_q R_k(q)m(q)m(-q) \tag{60}$$

where $\Gamma_k[m]$, called the modified Legendre transform, is the object we will work with from now on. The difference between Γ_k^{Leg} and Γ_k is a trivial term in m since it is quadratic, but it is convenient to subtract it from Γ_k^{Leg} to have the property $\Gamma k = \Lambda = H$ as we will show later. We then obtain:

$$\Gamma_k[m] + W_k[h] = \int_x h(x)m(x) - \frac{1}{2}\int_{xy} R_k(x-y)m(x)m(y) \tag{61}$$

where we have written the term coming from the regulator in position space this time.

Of course, decreasing k in R_k corresponds to integrating more and more fluctuations, and this leads to an evolution of $\Gamma_k[m]$ with k, called the renormalization flow. Let us now derive the exact flow equation for $\Gamma_k[m]$.

3.1 An exact identity for $\Gamma_k[m]$ and the limit $k \to \Lambda$

Intuitively, one might expect that if the function $R_k(q)$ becomes very large for all fluctuation modes regardless of their wavelength, the fluctuations of the system regulated by R_k will become very small,

[32] It will be shown later that in fact the value of $R_k(q)$ at $q = 0$ is not exactly k^2, as it will be convenient to include a contribution from the renormalization of the field. For what concerns us at this stage, this is a detail, but we must remember that we will have to return to the precise form of the regulator.

and the mean-field approximation will cease to be an approximation—it will become exact. This is what we will now demonstrate, using an exact identity satisfied by Γ_k.

Let us begin with the example of a theory involving a scalar field. Starting from Eq. (61), we rewrite it as:

$$e^{-\Gamma_k[m]} e^{\int_x h(x)m(x) - \frac{1}{2}\int_{xy} R_k(x-y)m(x)m(y)} = \int D\phi\, e^{-H[\phi] - \frac{1}{2}\int_{xy} R_k(x-y)\phi(x)\phi(y) + \int_x h(x)\phi(x)}. \tag{62}$$

Furthermore, using Eqs. (20) and (60), we obtain:

$$\frac{\delta\Gamma_k[m]}{\delta m(x)} = h(x) - \int_y R_k(x-y)m(y). \tag{63}$$

Substituting Eq. (63) into Eq. (62), we finally obtain:

$$e^{-\Gamma_k[m]} = \int D\phi\, e^{-H[\phi] + \int_x \frac{\delta\Gamma_k[m]}{\delta m(x)}(\phi(x)-m(x)) - \frac{1}{2}\int_{xy} R_k(x-y)(\phi(x)-m(x))(\phi(y)-m(y))}. \tag{64}$$

This equation is very interesting because it trivially shows that if $\forall q$, $R_k(q) \to \infty$, then $\Gamma_k = H$, or in other words, the mean-field approximation becomes exact in this limit.

To show this, we use the functional analogue of the relation

$$\sqrt{\frac{2\pi}{a}}\, e^{-\frac{1}{2}ax^2} \underset{a\to\infty}{\sim} \delta(x), \tag{65}$$

which reads

$$e^{-\frac{1}{2}\int_{xy} R_k(x-y)(\phi(x)-m(x))(\phi(y)-m(y))} \underset{R_k\to\infty}{\sim} \delta(\phi - m), \tag{66}$$

where the proportionality factor analogous to $\sqrt{2\pi/a}$ has been omitted since it is irrelevant for the following. When this relation holds, the functional integral can be performed, and the previously announced result follows trivially.

In practice, there are two cases:

- Either $R_{k=\Lambda}(q) \equiv \infty$, in which case $\Gamma_\Lambda = H$,
- Or, as is most often the case, $R_{k=\Lambda}(q) \simeq \Lambda^2$, in which case $\Gamma_\Lambda \simeq H$, because although Λ is very large compared to most scales in the problem, it is not infinite.[33]

To conclude this section, we highlight several points:

- The regulator R_Λ indeed fulfills its intended role of freezing fluctuations of modes with $q < \Lambda$, since for $k = \Lambda$, all (or almost all) fluctuations are frozen and the mean-field becomes exact.
- It is Γ_Λ that becomes equal to H and not $\Gamma_\Lambda^{\text{Leg}}$ as defined in Eq. (59), which becomes $H + \Delta H_\Lambda$, as expected for the Legendre transform of W_Λ.
- The term $-\frac{1}{2}\int_q R_k(q)m(q)m(-q)$ used in Eq. (60) to define the modified Legendre transform $\Gamma_k[m]$ serves to compensate, on average, the term ΔH_k introduced in the regularized partition function, Eq. (58). When all fluctuations are frozen, that is, for $k = \Lambda$, these two terms coincide, ensuring that $\Gamma_\Lambda = H$.
- From the above, we deduce that the renormalization group flow of Γ_k will allow us to connect the Hamiltonian H to the free energy Γ, since at $k = \Lambda$ (the initial condition of the flow), $\Gamma_{k=\Lambda} = H$, and at $k = 0$ (the end point of the flow), $\Gamma_{k=0} = \Gamma$, because $R_{k=0} \equiv 0$.

[33] Note that this is irrelevant for the calculation of universal quantities, which are independent of short-distance (i.e., ultraviolet) details, but it is important for the calculation of non-universal quantities that require the most accurate possible description of the UV sector.

3.2 The exact flow equation for $\Gamma_k[m]$

Let us first calculate the flow equation for Γ_k in the case $N = 1$.
First, note that

$$\partial_{k|_h} \neq \partial_{k|_m} \tag{67}$$

because, at fixed h, m is a function of k, and conversely, at fixed m, it is h that becomes a function of k. The relationship between the two derivatives is

$$\partial_{k|_m} = \partial_{k|_h} + \int_x \partial_k h(x)_{|_m} \frac{\delta}{\delta h(x)} \tag{68}$$

which generalizes the usual derivative change when changing the set of variables held fixed.

We now apply this derivative to Eq.(61) and obtain:

$$\partial_{k|_m}\Gamma_k[m] + \left(\partial_{k|_h} + \int_x \partial_k h(x)_{|_m} \frac{\delta}{\delta h(x)}\right) W_k[h] = \int_x \partial_k h(x)_{|_m} m(x) - \frac{1}{2}\int_{xy} \partial_k R_k(x-y)m(x)m(y). \tag{69}$$

From the definition of $W_k[h]$ we get

$$\frac{\delta W_k[h]}{\delta h(x)} = \langle \phi(x) \rangle = m(x) \tag{70}$$

where the average is taken in the presence of the ΔH_k term. From Eq.(58) we also obtain:

$$\begin{aligned}
\partial_{k|_h} W_k[h] &= \frac{1}{Z_k[h]} \int D\phi \left(-\frac{1}{2}\int_{xy} \partial_k R_k(x-y)\phi(x)\phi(y)\right) e^{-H_k[\phi] + \int_x h\phi} \\
&= -\frac{1}{2}\int_{xy} \partial_k R_k(x-y)\langle \phi(x)\phi(y)\rangle \\
&= -\frac{1}{2}\int_{xy} \partial_k R_k(x-y)\left(G^{(2)}_{c,k}[x,y;h] + m(x)m(y)\right).
\end{aligned} \tag{71}$$

Substituting Eqs.(70) and (71) into Eq.(69), we get:

$$\partial_k \Gamma_k[m] = -\frac{1}{2}\int_{xy} \partial_k R_k(x-y) G^{(2)}_{c,k}[x,y;h]. \tag{72}$$

Since the functions $G^{(2)}_{c,k}[x,y;h]$ and $\Gamma^{(2),\mathrm{Leg}}_k[x,y;m]$ are inverses of each other, see Eq. (22), we finally deduce the flow equation:

$$\partial_k \Gamma_k[m] = -\frac{1}{2}\int_{xy} \partial_k R_k(x-y)\left(\Gamma^{(2)}_k[x,y;m] + R_k(x-y)\right)^{-1} \tag{73}$$

often called the Wetterich equation, this equation has the initial condition $\Gamma_\Lambda = H$ if $R_\Lambda(q) = \infty$.

The previous equation can be easily generalized to $\mathrm{O}(N)$ symmetric models. In this case, it is obviously convenient to take a term ΔH_k that respects the symmetry:

$$\Delta H_k[\boldsymbol{\phi}] = \frac{1}{2}\int_{xy} R_k(x-y)\boldsymbol{\phi}(x)\cdot\boldsymbol{\phi}(y) = \frac{1}{2}\int_{xy} \phi_\alpha(x)\Big(R_k(x-y)\delta_{\alpha\beta}\Big)\phi_\beta(y). \tag{74}$$

Just like the cutoff function: $R_{k,\alpha\beta}(x-y) = R_k(x-y)\delta_{\alpha\beta}$, the function $\Gamma^{(2)}$ is also an $\mathrm{O}(N)$ tensor, as seen in Eq. (21). By repeating the derivation steps above, we unsurprisingly find for the flow equation:

$$\partial_k \Gamma_k[m] = -\frac{1}{2}\mathrm{Tr}\int_{xy} \partial_k R_k(x-y)\left(\Gamma^{(2)}_k[x,y;m] + R_k(x-y)\right)^{-1} \tag{75}$$

where the trace is taken over the group indices, here $\mathrm{O}(N)$.

More generally, if H is symmetric under a group $\mathcal{G}$, we will take a term ΔH_k symmetric under $\mathcal{G}$ if it exists, and the flow equation will be formally identical to the one above with the trace taken over the indices of $\mathcal{G}$.

3.2.1 Some properties of the exact flow equation and the approximations to be implemented

- The Wetterich equation (75) is an exact equation that, if solved exactly, gives $\Gamma[m]$ at the end of the flow, starting from H as the initial condition, which constitutes an exact solution of the field theory defined from H.[34]

  This equation, which seems simple at first glance, is actually not: it is a functional, integral, nonlinear, and partial differential equation. It is generally not exactly solvable and therefore requires approximations to be usable. In the following, we will see two approximation schemes: the derivative expansion, which is by far the most popular, and the BMW (Blaizot-Méndez-Wschebor) method, which is more powerful but also more complicated to implement.

- It may seem surprising at first that an exact equation claiming to compute the free energy of a system depends crucially on a cutoff function $R_k(q)$, which is arbitrary except that it must vanish identically at $k = 0$ and tend to infinity as k approaches Λ. The solution to this paradox is, of course, that $\Gamma_{k=0} = \Gamma$ regardless of the choice of the cutoff function, as long as it satisfies $R_{k=0} \equiv 0$. In other words, the details of the trajectory connecting $\Gamma_\Lambda = H$ to $\Gamma_{k=0} = \Gamma$ do not matter as long as the starting and ending points are correct. See Fig. 5 for a schematic description of the flow.

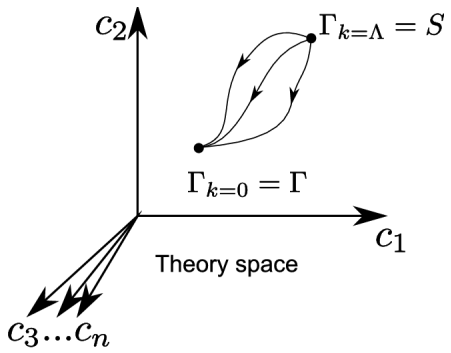

Figure 5: Different exact RG flows in the theory space corresponding to different choices of the function R_k. In QFT, the initial condition is the action of the model, denoted S, and in statistical mechanics, it is the Hamiltonian H. The axes schematically represent the different coupling constants. They are generally infinite in number because along the renormalization flow, all couplings (compatible with the symmetries) are generated.

But this statement is too brief. As mentioned above, RG is only of interest when approximations are used to perform actual calculations. And when this is done, the final result of the integration of the flow generally retains an artificial dependence on the choice of the function R_k. In other words, in the analogy to Fig. 5 in the presence of approximations, the flows do not all lead to the same point $\Gamma_{k=0}$, meaning they show a dispersion of $\Gamma_{k=0}$.

For a long time, part of the physics community active in this research field considered this dispersion to be a fatal flaw for the Wilsonian renormalization method, which implied two things: (i) a physical quantity cannot depend on a non-physical choice, and (ii) this dependence

[34] Note that an exact equation can also be formulated for lattice theories. Of course, two theories, one in the continuum and the other on a lattice, can belong to the same universality class. Their universal quantities will therefore be identical, but the non-universal ones will be different. The formulation of the flow on the lattice has been successfully implemented for calculating critical temperatures of ferromagnetic spin models with O(N) symmetry on a three-dimensional cubic lattice with nearest-neighbor spin interactions.

is specific to the Wilsonian method and discredits it for actual (i.e., approximate) calculations. However, we must be very cautious with assertions (i) and (ii) above.

First of all, one can only agree with (i) if no approximation is made. But generically, and in all areas of physics, the results of approximate calculations depend on the method of calculation. Take, for example, the simple perturbative calculation of the frequency $\Omega/2\pi$ of a classical anharmonic oscillator with Hamiltonian $H = p^2/2m + m\omega_0^2 x^2/2 + gx^4$. The standard method for perturbatively calculating Ω consists of considering the gx^4 term as a perturbation to the harmonic oscillator with frequency $\omega_0/2\pi$ and calculating Ω as a Taylor series in g. But nothing prevents us from first splitting the harmonic potential term: $m\omega_0^2 x^2/2 = m\omega^2 x^2/2 + m(\omega_0^2 - \omega^2)x^2/2$ and performing perturbation theory by considering $m(\omega_0^2 - \omega^2)x^2/2 + gx^4$ as the perturbation. One then finds, at all orders in perturbation theory, that Ω depends on ω, which is clearly an artifact of perturbation theory.[35] Thus, it is found that there is not just one perturbation theory, but, for a given finite order, an infinity of non-equivalent ways to construct perturbation theories. This is also true in field theory, and just because one is chosen does not mean that the others do not exist.

Note that in field theory, there are intrinsically different perturbation theories: expansion in $\epsilon = 4 - d$ or $\epsilon = d - 2$, expansion at fixed d in coupling constant (for example, the "massive zero momentum scheme"), expansion in $1/N$, to name just a few of the most famous. It must be understood, however, that even for a given perturbative method, there are several non-equivalent ways to carry out the calculations that lead to different results at finite order. For this reason, it is said that the results obtained are dependent on the renormalization scheme. This answers point (ii) above: choosing a cutoff function R_k in FRG is analogous to choosing a renormalization scheme in perturbation theory, and in both cases, the physical results obtained at a given order of approximation depend on this choice. Therefore, FRG is not the only method to suffer from this arbitrariness, and the art, whether or not perturbative, is to optimize the choice of the scheme, see below for the optimal selection of R_k.

Finally, note that since perturbative series are generally not convergent, they require resummation (when possible, though not always) to be useful, which introduces its own arbitrariness because the resummed results depend on the resummation methods used.[36] FRG is in a better position in this regard, see below.

- If the microscopic Hamiltonian H (and the functional measure) are symmetric under a group G, and if there exists a cutoff function R_k such that the term ΔH_k respects this symmetry, then Γ_k is symmetric under G for all k, and thus $\Gamma = \Gamma_{k=0}$ is also symmetric. It is then said that the symmetry G has no anomaly: in quantum terms, the symmetry of the classical theory, i.e., the symmetry of H, is preserved at the quantum level, that is, on Γ and thus for all correlation functions.

  It may happen that no quadratic term ΔH_k in the fields respects the symmetry G, while the theory, i.e., Γ, remains invariant under G. This means that the symmetry is broken by ΔH_k

[35] It should be noted here that this dependence on ω, which may seem catastrophic, can actually be turned into an advantage. The idea is to search for "the best harmonic oscillator" around which to perform the perturbative calculation. Indeed, if g is large, Ω can be very different from ω_0 – it turns out that this is the case – which makes the perturbative series for large values of g very poor. The idea is to choose ω such that the difference $\Omega_n - \Omega_{n-1}$, where Ω_n is the value of Ω at the n-th order of perturbation, is as small as possible, ensuring the fastest possible convergence of the sequence of Ω_n. At any order n, this makes ω a function of g, which is entirely legitimate since ω is arbitrary. This procedure's result is that the convergence of these Ω_n is very rapid, unlike the perturbative series in g, which is not convergent and requires quite sophisticated resummations and large perturbative orders to achieve the same level of precision. This method of improving perturbation theory is sometimes called "improved perturbation theory". The optimization of the choice of non-physical parameters in FRG, the function R_k in this case, will be an important theme of these notes.

[36] Note that QED avoids this issue due to the smallness of the coupling constant that orders the perturbation theory, see footnote 11.

for all finite k and is only restored for $k = 0$. This is the case for gauge symmetries in QFT, as these symmetries impose the nullity of the mass of the gauge bosons and are therefore incompatible with a quadratic term in the fields such as ΔH_k. This symmetry-breaking term can be controlled by modified Ward (more precisely, Slavnov-Taylor) identities, which become, in the limit $k \to 0$, the true Ward identities of the gauge symmetry.

However, an additional difficulty arises once approximations are implemented, as in this case, the final result of the flow integration retains an artificial dependence on R_k as explained above, making it delicate to ensure the gauge invariance of the results of the flow integration, even at $k = 0$. This is true beyond the example of gauge invariance and is encountered in any situation where a quadratic cutoff term in the fields is incompatible with a symmetry or with any other property of the model.[37]

- One can derive an exact RG flow equation in the presence of fermions. The trace is then replaced by a supertrace (fermions contribute with an opposite sign to bosons).
- The equation (75) closely resembles the result at one loop, since at this order:

$$\Gamma_k = H + \frac{1}{2}\mathrm{Tr}\log\left(H^{(2)} + R_k\right). \tag{76}$$

Thus, by replacing $H^{(2)}$ – which is $\Gamma^{(2)}$ at the mean field approximation – with the functional $\Gamma_k^{(2)}[m]$ and differentiating with respect to k, we turn the one-loop result into an exact result! There is a diagrammatic representation of the Wetterich equation highlighting its one-loop structure, see Fig. 7.

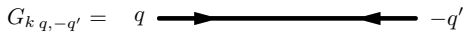

Figure 6: Diagrammatic representation of the full propagator $G^{(2)}_{c,k}[q, -q'; h]$ in Fourier space. It is a function of both x and y in direct space and of two arguments q and $-q'$ in Fourier space because there is no translation invariance when evaluating $G^{(2)}_{c,k}$ in a non-constant external field $h(x)$.

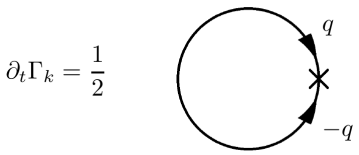

Figure 7: Diagrammatic representation of the Wetterich equation (75). In this equation, the integrals over x and y force, in Fourier space, the second argument of $G^{(2)}_{c,k}[q, -q'; h]$ to be $q' = q$. The remaining integral over q is represented by a loop. The cross corresponds to $\partial_k R_k(q)$.

- This one-loop structure has a very important practical consequence: only one integral needs to be computed, and when evaluated in a constant field, it is one-dimensional due to rotational invariance. This is very different from perturbation theory, where l-loop diagrams require integrals of dimension d times l. This represents a significant simplification compared to perturbation theory.

[37] In out-of-equilibrium statistical mechanics, the notion of causality often plays an important role and is not strictly speaking a symmetry. Here too, care must be taken to choose an R_k that does not violate this property.

- The perturbative expansion can be recovered from the Wetterich equation and, for example, the proof of the perturbative renormalizability of the ϕ^4 theory in $d = 4$ is much simpler in this formalism than in the diagrammatic formalism.

- Due to the term $\partial_k R_k$ in the Wetterich equation (75), only the q^2 moments of the order of k^2 or less contribute to the flow at the scale k (we will revisit this point in detail later). Thus, the RG flow is regular in the ultraviolet. Regarding the infrared, the term in R_k of the propagator: $(\Gamma^{(2)} + R_k)^{-1}$ regularizes its behavior at small momenta, which is normal since the "coarse-graining" procedure itself involves freezing the modes of long-wavelength fluctuations. All divergences in perturbation theory are thus avoided: the RG flow is computed directly and not first, as in perturbation theory, the relation between bare and renormalized quantities from which the RG flow is computed in a second step.

- From what has been said above, it follows that k acts as an infrared regulator (for $k \neq 0$), somewhat similar to a finite-size box $\sim k^{-1}$. Thus, for $k > 0$, there is no phase transition in the system regularized by R_k, and therefore no infrared singularity in Γ_k. At finite k, everything is regular and can be expanded in series of the field and/or its derivatives. We can therefore conclude that:

  (i) Singularities of Γ build up as k decreases and are thus smoothed by k in Γ_k;

  (ii) The precursor of the critical behavior should already appear at finite k for much larger momenta than k. This is indeed what is observed when calculating (using the BMW method, see below) the function $\Gamma_k^{(2)}(p, m = 0)$: for a critical model, i.e., one whose Hamiltonian H renders it critical, the behavior at small p: $p \ll k$ is regular and expandable in a series of p^2/k^2, while at $p \gg k$ it behaves as $p^{2-\eta}$, which is the expected behavior of $\Gamma^{(2)}_{k=0}(p)$ for any $|p| > 0$ at criticality. We thus see the critical behavior in p developing in the region where k is negligible compared to $|p|$, and as k decreases, this region extends more and more until it covers the entire p domain for $k = 0$. We will see that the region $p \lesssim k$ is the region where the derivative expansion is valid, beyond which it ceases to be valid (see below).

  (iii) There is a fundamental difference between the flow of Hamiltonians as historically formulated by Wilson (or the flow of $W_k[h]$) and the flow of Γ_k: in the first case, k is an ultraviolet cutoff because the effective Wilson Hamiltonian at the scale k, obtained after integrating over the fast modes (block-spin principle), is the Hamiltonian for the slow modes, i.e., those that have not yet been integrated. In the second case, Γ_k is the (modified) free energy for the fast modes, i.e., those that have already been integrated. The Wilsonian Hamiltonian at scale k has no simple physical interpretation because, to read the resulting physics, i.e., to obtain the free energy, one still needs to calculate a functional integral, that over the slow fluctuation modes. On the other hand, Γ_k is a kind of precursor to Γ, since it is a free energy, albeit strange because it corresponds to the incomplete integration over the fluctuations, but on which one can already read part of the physics, as explained above.

4 The Derivative Expansion

As explained earlier, the flow equation of Γ_k (75) cannot generally be solved, and recourse to approximations is inevitable. The most famous of these approximations is perturbation theory.

We will now see that the formulation of field theory in the form of the Wetterich equation (75) allows for approximations that are not based on the smallness of any parameter, whether it is the deviation from the upper critical dimension ($\epsilon = 4 - d$) or lower critical dimension ($\epsilon = d - 2$), the coupling constant g_0 (fixed-dimension expansion around the gaussian), or the inverse of the number

of components of the field, $1/N$ (expansion around the spherical model at $N = \infty$). The most famous of these approximations is the derivative expansion of $\Gamma_k[m]$, and we will now describe it in detail for $N = 1$.

As explained earlier, the fluctuations of slow modes responsible for infrared singularities at criticality are cut off in Γ_k by the regulator R_k. As a result, Γ_k is a regular functional of $m(x)$, even if the initial system, unregulated, is critical. We can therefore attempt a derivative expansion of this functional, which, for example, for $\Gamma_k^{(2)}(p, m = 0)$ will be an expansion in powers of p^2/k^2, regular as long as $k \neq 0$. For $N = 1$, this expansion is quite analogous to the one given in Eq. (28):

$$\Gamma_k[m] = \int_x \left\{ U_k(m(x)) + \frac{1}{2} Z_k\left(m(x)\right) \left(\nabla m(x)\right)^2 + O(\nabla^4) \right\}. \tag{77}$$

In the above equation, the k dependence is carried by the functions $U_k, Z_k, \cdots$.[38] It would be the same for $N > 1$ with the analogue of Eq. (29). Of course, the issue of the convergence of this expansion is crucial to determine whether we can derive non-trivial physics from it, and, as explained earlier, this is indeed the case for scales smaller than typically k. In the limit $k \to 0$, we can therefore hope to compute thermodynamic quantities reliably with this expansion, that is, those defined from correlation functions at zero moments, such as the correlation length, susceptibility, magnetization, specific heat, etc.

The flow equation for Γ_k (75), once the derivative expansion is inserted, becomes an infinite set of coupled flow equations for the functions $\{U_k(m), Z_k(m), \cdots\}$. To be useful, this infinite tower of equations must be truncated by keeping only a finite number of functions corresponding to a finite order in the derivative expansion. The simplest truncation is to keep only the potential U_k as a function of k. This is known as the local potential approximation (LPA), which we will now study in detail.

4.1 The Local Potential Approximation in the Ising case

In this approximation, we replace Γ_k by Γ_k^{LPA} given by

$$\Gamma_k[m] \to \Gamma_k^{\mathrm{LPA}}[m] = \int_x \left\{ U_k(m(x)) + \frac{1}{2} \left(\nabla m(x)\right)^2 \right\}. \tag{78}$$

Due to the $\mathbb{Z}_2$ symmetry, $U_k(m(x))$ is actually a function of $\rho(x) = m^2(x)/2$ only, and it may sometimes be useful to work with ρ rather than m to make the invariance under $\mathbb{Z}_2$ manifest in the equations we manipulate.

Note that in the LPA, the derivative term already present in H has been retained (but with $Z_k(m)$ replaced by 1) because otherwise the system would have no dynamics, as there would be no coupling between the $m(x)$ at different points.

The potential U_k is defined from Γ_k by:

$$\Omega U_k(m) = \Gamma_k[m]_{|_{m(x)=m}} \tag{79}$$

where m is constant and Ω is the volume of the system.[39] From this definition, we can see that there is no need to consider x-dependent fields $m(x)$ to obtain the functional dependence of U_k on m, which will greatly simplify our calculations.

The flow of U_k is obtained by applying the operator ∂_k to both sides of Eq. (79), replacing $\partial_k \Gamma_k$ with the right-hand side of the Wetterich equation, and calculating the right-hand side using the

[38] Recall that the function $Z_k(m(x))$ has nothing to do with the partition function $\mathcal{Z}_k[h]$.

[39] It is important to give a definition of U_k that is valid independently of the order of the derivative expansion being considered. Thus, the only source of error in its flow equation will come from truncating the derivative expansion, not from a definition that changes with the order considered.

LPA ansatz, Eq. (78):

$$\begin{aligned}\partial_k U_k(m) &\underset{\text{LPA}}{=} \frac{1}{2}\int_{xy} \partial_k R_k(x-y)\left(\Gamma_k^{(2),\text{LPA}}(x-y,m)+R_k(x-y)\right)^{-1} \\ &\underset{\text{LPA}}{=} \frac{1}{2}\int_{q} \partial_k R_k(q)\left(\Gamma_k^{(2),\text{LPA}}(q,m)+R_k(q)\right)^{-1}.\end{aligned} \tag{80}$$

The calculation of $\Gamma_k^{(2),\text{LPA}}(x-y,m)$ is straightforward:

$$\Gamma_k^{(2),\text{LPA}}(x-y,m) = \left(-\nabla_x^2 + U_k''(m)\right)\delta(x-y). \tag{81}$$

This gives the LPA flow for U_k:

$$\partial_k U_k(m) = \frac{1}{2}\int_q \frac{\partial_k R_k(q)}{q^2 + R_k(q) + U_k''(m)}. \tag{82}$$

It is now time to choose a cutoff function, and we will take, for example:

$$R_k(q) = (k^2 - q^2)\theta(k^2 - q^2) \tag{83}$$

where $\theta(x)$ is the Heaviside step function. This regulator has several advantages: it significantly simplifies the LPA flow and may be optimal at the LPA for calculating critical exponents, in the sense that it leads to results closest to the best-known results in $d = 3$ (though there is no rigorous proof of this, it is an empirical observation). Using

$$\int \frac{d^d q}{(2\pi)^d} f(q^2) = 2v_d \int_0^\infty dx\, x^{d/2-1} f(x) \tag{84}$$

with $v_d = (2^{d+1}\pi^{d/2}\Gamma(d/2))^{-1}$, we obtain:

$$\boxed{\partial_k U_k(m) \underset{\text{LPA}}{=} \frac{4v_d}{d}\frac{k^{d+1}}{k^2 + U_k''(m)}}. \tag{85}$$

This is a nonlinear partial differential equation, yet still very simple, and its numerical integration is straightforward and without surprises, at least for $d > 2$ (we will come back to the case $d = 2$).

To integrate this equation, we need an initial condition. In our case, it is given by the "bare" ϕ^4 potential, the one from the Hamiltonian of the system, Eq. (3). By fixing the value (positive) of g_0, there are several cases depending on the value of r_0.

To qualitatively understand what happens when integrating the flow of U_k, we recall that fluctuations tend to disorganize the system, so the value of the critical temperature for a model like Ising is smaller than its mean-field value: $T_c < T_c^{\text{MF}}$. For the ϕ^4 theory, the situation is the same: the value of r_0 for which the model is critical is smaller than the mean-field value (which is zero): $r_{0,c} < r_{0,c}^{\text{MF}} = 0$. Therefore, there are three possibilities, described qualitatively below and confirmed by the numerical integration of the flow given by Eq. (85).

(i) Let $r_{k=\Lambda} = r_0 > 0$. Even at the mean-field level, the system is in the disordered phase and the initial potential is convex, U-shaped. Since fluctuations can only accentuate this, the potential remains convex throughout the flow, deforming slightly due to fluctuations, mainly Gaussian, since the system is deeply in the high-temperature phase and its correlation length is consequently small.

(ii) Let $0 = r_{0,c}^{\text{MF}} > r_{k=\Lambda} = r_0 > r_{0,c}$. The system has a bare potential, i.e., a mean-field potential, with a double well since $r_0 < 0$. It therefore has two minima at $\pm\sqrt{-6r_0/g_0}$. However, it is in the high-temperature phase, so its spontaneous magnetization is zero, and the effective potential $U = U_{k=0}$ must have only one minimum at the origin. We deduce that the potential $U_k(m)$ must

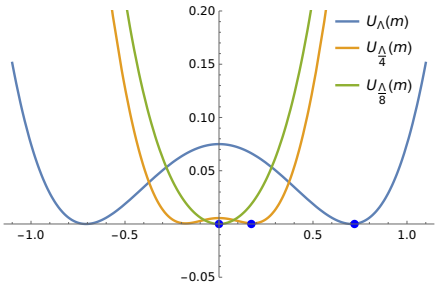

Figure 8: Typical flow of the potential U_k in the ordered phase ($r_0 > r_{0,c}$). The position of the potential minimum, marked by a dot, varies with the scale k and approaches the origin as k decreases, as might be expected since fluctuations tend to disorganize the system and thus reduce the spontaneous magnetization compared with the mean field. At a finite scale k_0, the minima reach the origin. This scale is typically the inverse of the correlation length ξ. For $k < k_0$, the minimum remains at the origin and the potential then deforms relatively little.

deform during the flow such that its minima (dependent on k) approach the origin until they meet at the origin for a value $k_0 > 0$, see Fig. 8. For the remainder of the flow, i.e., for $k < k_0$, the flow is quite similar to case (i): the potential continues to deform slightly but remains convex, U-shaped, until $k = 0$. We can easily guess the value of k_0: in the high-temperature phase, when the correlation length is large and finite, $\xi \gg a = \Lambda^{-1}$, we realize that $\xi < \infty$ only for length scales larger than or of the order of the correlation length: for much smaller length scales, it behaves as though the correlation length were infinite. Consequently, as long as $k \gg \xi^{-1}$, the potential evolves significantly with the flow and its minima approach the origin, and when k becomes of the order of ξ^{-1}, the system no longer appears critical, its high-temperature phase becomes manifest, and the only remaining minimum is at the origin. We conclude that $k_0 \simeq \xi^{-1}$.

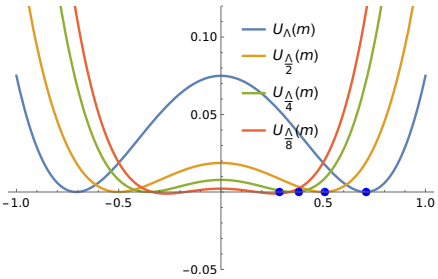

Figure 9: Typical flow of U_k in the broken phase ($r_0 < r_{0,c}$). The position of the potential minimum, marked by a dot, varies with scale k and decreases as k decreases. For $k \simeq \xi^{-1}$, the position of the minimum has almost reached its value at $k = 0$, which is that of spontaneous magnetization, since for $k < \xi^{-1}$ fluctuations are so small that the value of this minimum remains practically unchanged. On the other hand, the potential continues to flatten between minima since $R_{k=0}(q) \equiv 0$ so that $U = U_{k=0}$ is obtained by a true Legendre transform and is therefore convex.

(iii) Let $r_{k=\Lambda} = r_0 < r_{0,c}$. At the end of the flow, the potential must exhibit minima at the spontaneous magnetization $\pm m_{\text{sp.}}$. In other words, the minima of the potential $U_k(m)$ evolve with k until they reach a non-zero value at $k = 0$, which gives the spontaneous magnetization, see Fig. 9. Notice that U_k is not necessarily convex since it originates from a Legendre transform to which a quadratic term has been subtracted to give Γ_k. However, at $k = 0$, this quadratic term vanishes, and thus $U = U_{k=0}$ must be convex. Of course, the convexity property could be broken by the

LPA, and this is indeed what happens for certain choices of regulator. This is not the case for the regulator given in Eq. (83). In fact, by numerically integrating the flow, we find that the minima of U_k stabilize at (approximately) $\pm m_{\text{sp.}}$ for a value $k \simeq \xi^{-1}$, where U_k is still far from being convex, and during the remainder of the flow, the position of the minima changes very little, but the inner part of the potential, located between the two minima, flattens until it becomes entirely flat at $k = 0$. Notice that since $U'_{k=0}(m) = h$, the two minima at $\pm m_{\text{sp.}}$ are reached at the limits $h \to 0^{\pm}$, and there is thus a discontinuity in the magnetization at $h = 0$ in the broken symmetry phase. A consequence is that the flat part of the potential located between $\pm m_{\text{sp.}}$ is not thermodynamically accessible, i.e., not accessible by changing a parameter such as temperature or magnetic field.

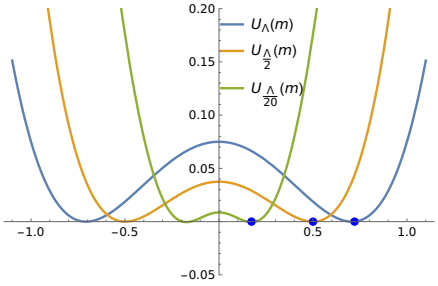

Figure 10: Typical flow of U_k at criticality ($r_0 = r_{0,c}$). The position of the minimum, marked by a dot, varies with the scale k and approaches the origin as k decreases, as in previous cases. However, being exactly at criticality, it is only at $k = 0$ that the two minima reach the origin. After a transient regime where $k \lesssim \Lambda$, the flow is self-similar, as no scale is relevant to the flow: Λ is much greater than k and therefore no longer plays a role, and the inverse of the correlation length is zero. We then enter the scale-invariant regime. In this $k \ll \Lambda$ regime, the potential deforms but remains identical to itself up to a rescaling of the two axes.

Criticality corresponds to the intermediate situation between cases (ii) and (iii), which is confirmed by the qualitative argument given in (ii) leading to $k_0 \simeq \xi^{-1}$. This qualitative argument also tells us that at criticality, the potential U_k has two minima throughout the flow, and it is only at the end of the flow, at $k = 0$, that these minima (and the local maximum located between the two minima) reach the origin, see Fig. 10. We must remember that at criticality, scale invariance is an emerging symmetry, at least for length scales much larger than the lattice spacing $a \sim \Lambda^{-1}$. This tells us that after a transient, non-scale-invariant regime where $k \lesssim \Lambda$, any scale has disappeared in the system, and therefore in the flow for $k \ll \Lambda$, except for k: the system being critical and, at large scales, appearing scale-invariant since $\xi = \infty$, only k remains as the relevant scale in the regularized system.

To understand what happens, let us ask the following question: what is a scale-invariant system in which we have introduced a reference length $l \sim k^{-1}$? This question, however trivial it may seem, has its subtleties. Let us imagine that we are in a spaceship far from any galaxy and that our world is scale-invariant (which it is not in reality). In this world, we can make a ruler that serves as our reference length to measure all the lengths in our world. Now let us imagine that we are communicating by radio with an extraterrestrial being at the other end of the universe. We could not tell them the direction in which our spaceship is pointing because all directions are equivalent in a universe devoid of matter (rotational invariance), nor could we tell them our position (translational invariance). We could not even tell them the size of our reference ruler, since by assumption our world is scale-invariant: for any experimental protocol A, aiming to communicate the size of our ruler, we could imagine another protocol B, where all lengths are scaled by a common factor, and it would lead to the same physical conclusions. Whether our size is one Angström or one light-

year is therefore an uncommunicable, and thus physically irrelevant, quantity, just like our absolute orientation or absolute position in space.[40]

What is important to understand is that from our explanations, our extraterrestrial could reproduce a spaceship B exactly like our spaceship A, except for its size, as it would be built using a length standard l_B, which he cannot compare absolutely with our length standard l_A. However, all proportions of spaceship B would be preserved, i.e., all ratios between lengths: if the length of the table in spaceship A is $2\,l_A$, then the table in spaceship B will be $2\,l_B$. This leads to the somewhat puzzling conclusion that the two tables in our previous example have both different absolute lengths and identical relative lengths.[41] More generally, any equation that is true in A and contains only lengths measured in A in units of l_A would also be valid in B, provided that the lengths there are measured in units of l_B. Put another way, equations written in A and B would not be identical in general if they involved lengths, but they would become so if they were written in terms of length ratios, i.e. in terms of lengths rendered dimensionless in A because measured in units of l_A and rendered dimensionless in B because measured in units of l_B.[42]

This lengthy preamble was meant to prepare the reader for what follows: the notion of dimensionless quantities in terms of k and the notion of a fixed point when scale invariance emerges.

First, note that unlike the discussion above, by changing H to $H_k = H + \Delta H_k$, we have not only introduced a length scale k^{-1} but also affected the scale invariance of H: if H is the Hamiltonian of a critical system, this is not the case for H_k when $k > 0$, and scale invariance is lost in the regularized theory. But just like in our example above, only the scale k has been introduced into H_k. Thus, while it is true that the flow equation (85) explicitly depends on k, and therefore it is not the same at different scales k and k', this should no longer be the case for the flow equation of quantities rendered dimensionless using scale k. We thus introduce the following dimensionless quantities:

$$\begin{aligned} \tilde{x} &= k\,x \\ \tilde{p} &= p/k \\ \tilde{m}(\tilde{x}) &= k^{-\frac{d-2}{2}} m(x) \\ \tilde{U}_k(\tilde{m}(\tilde{x})) &= k^{-d} U_k(m(x)) \end{aligned} \tag{86}$$

where the first two equations are trivial, and the next two follow from Eq. (78) and the fact that $\Gamma_k[m]$ is dimensionless. In fact, it is more coherent and convenient for the following to introduce a "renormalization time", also dimensionless, defined by $t = \log(k/\Lambda)$, rather than continuing to work with k, which has a dimension. We will therefore define:

$$\tilde{U}_t(\tilde{m}(\tilde{x})) = k^{-d} U_k(m(x)). \tag{87}$$

We can rewrite $\Gamma_k^{\mathrm{LPA}}[m]$ in terms of these variables:

$$\Gamma_t^{\mathrm{LPA}}[\tilde{m}] = \int_{\tilde{x}} \left\{ \tilde{U}_t(\tilde{m}(\tilde{x})) + \frac{1}{2} \left(\tilde{\nabla} \tilde{m}(\tilde{x}) \right)^2 \right\} \tag{88}$$

[40] Of course, our universe is not scale-invariant because we are made of atoms, which indeed have fixed sizes. As far as we know, we cannot invent an experiment B where, for example, the Bohr radius would be modified at will: the mass of the electron seems to be the same everywhere in the universe. It is the mass of elementary particles which is at the origin of the non-invariance by dilation of our universe because from this mass, from $\hbar$ and from the speed of light we can build a length scale, the Compton wavelength $\hbar/(mc)$ which is universal. However, the Ising model does not care about the mass of elementary particles, and at criticality, it indeed appears scale-invariant, at least at length scales much larger than the lattice spacing.

[41] This is puzzling but general to any symmetry. For rotational invariance for instance, the two spaceships would be oriented differently in space but the relative positions of all the internal elements of the spaceships would be identical.

[42] The above answers a potential question from the reader: l_A is a symbol representing a length; but what exactly is the value of this length if we wanted to replace the symbol with its numerical value? At first glance, the question seems unanswered, as to give a numerical value to l_A, we would need to measure the length of the reference ruler, but by assumption, there is no reference length other than the ruler itself. However, the solution to this paradox is simple: l_A is 1... in units of l_A.

and repeat the derivation of the flow equation for the potential at LPA. However, it is just as simple to perform the change of variables directly on Eq. (85) and compute $\partial_t \tilde{U}_t(\tilde{m}) = k\,\partial_k \tilde{U}_t(\tilde{m})$. As usual when performing such a change of variables, one must be careful that $\partial_{k|_m} \neq \partial_{k|_{\tilde{m}}}$. We have:

$$\partial_{t_{|_{\tilde{m}}}} = k\,\partial_{k_{|_{\tilde{m}}}} = k\,\partial_{k_{|_m}} + k\,\partial_k m_{|_{\tilde{m}}} \frac{\partial}{\partial m} = k\,\partial_{k_{|_m}} + \frac{d-2}{2}\tilde{m}\frac{\partial}{\partial \tilde{m}} \tag{89}$$

and thus,

$$\left(\partial_{t_{|_{\tilde{m}}}} - \frac{d-2}{2}\tilde{m}\frac{\partial}{\partial \tilde{m}}\right)\tilde{U}_t(\tilde{m}) = k\,\partial_{k_{|_m}}\left(k^{-d}U_k(m)\right) = -d\,\tilde{U}_t(\tilde{m}) + k^{1-d}\,\partial_{k_{|_m}} U_k(m) \tag{90}$$

or,

$$\left(\partial_t + d - \frac{d-2}{2}\tilde{m}\frac{\partial}{\partial \tilde{m}}\right)\tilde{U}_t(\tilde{m}) = k^{-d}k\,\partial_k U_k(m) \tag{91}$$

where it is understood that the operators ∂_t and ∂_k must be computed in each case while keeping $\tilde{m}$ and m fixed, respectively. The right-hand side must be rewritten in terms of $\tilde{U}_t(\tilde{m})$ so that the equation only involves this quantity on both sides. Using Eq. (85), we finally obtain the flow equation for the dimensionless potential at the LPA when the regulator in Eq. (83) is used:

$$\boxed{\partial_t \tilde{U}_t(\tilde{m}) = -d\,\tilde{U}_t(\tilde{m}) + \frac{d-2}{2}\tilde{m}\,\tilde{U}_t'(\tilde{m}) + \frac{4v_d}{d}\frac{1}{1+\tilde{U}_t''(\tilde{m})}}\,. \tag{92}$$

As announced in the long discussion above, any explicit dependence on k has disappeared from this equation. It has been transformed into terms that take into account the dimensions of the different quantities involved: d for the potential and $\frac{d-2}{2}$ for the field.[43]

In situations where the initial potential defines a critical theory and where, consequently, scale invariance will emerge at large distances, that is, as $t \to -\infty$, we expect that the evolution of the potential U_k becomes self-similar at small k, meaning it no longer distorts under the coarse-graining operation except in accordance with its dimension. In other words, once measured in units of k, the potential, $\tilde{U}_t$, should no longer evolve at all as a function of the scale: $\partial_t \tilde{U}_t(\tilde{m}) = 0$. The solution to this equation, denoted $\tilde{U}^*$, is called the fixed point potential of the renormalization group (at the LPA).

We have thus replaced the criticality condition with the scale invariance condition (at large distances), and this in turn with the fixed point condition of the theory regularized by R_k.[44] This is very clever from a practical point of view because, as we will see, it allows a very fine characterization of the critical point.

[43] Note that by redefining $\tilde{U}_t$ and $\tilde{m}$, we can make the $4v_d/d$ factor completely disappear from the previous equation. It is even advisable to do so in the search for a solution to this equation, especially when the dimension is varied between 4 and 2, as this avoids the appearance of large numbers that can disrupt numerical analysis.

[44] In technical terms, scale invariance is translated into the Ward identity for dilatation on Γ. The fixed point condition on Γ_k, the cancellation of the right-hand side of Eq. (92) in the case of the LPA, is the Ward identity of scale invariance on Γ_k, modified by the presence of the regulator, and which becomes the true Ward identity for dilatation when the regulator vanishes, i.e., for $k = 0$.