

Originally prepared for *ATITPhysics 2021 Summer School...*

Conformal Symmetry in Physics

Soner Albayrak

MAY 29, 2023

Summary

These are the notes of the lectures prepared for the “Advanced Topics in Theoretical Physics” summer school at *Feza Gürsey Merkezi*, September 2021. These notes are mostly based on other discussions and I provided the sources in relevant places. I will keep updating the notes after the summer school to keep it up-to-date and self-contained; there are also reminders [in blue](#) for me to add further discussion/comments.

For contact: contact@soneralbayrak.com

Preface

Remarks to the reader

I would like to make a few points crystal clear:

1. I do not update these notes on a regular basis.
2. The contents are correct to the best of my knowledge, but I have not put the extra effort to make sure that everything is book-level correct; nevertheless, I try to put as many references as possible when relevant, so please make the most of it!
3. I prefer a casual tone in my notes: if you don't like it, just deal with it! I believe this makes it easier for students to connect with written content, but it can also be a symptom for a midlife crisis of a theoretical physicist (I should be young enough for this to be correct though).
4. I believe that the level of this book is appropriate for an average senior undergraduate student, but I would dare say that it should be quite useful even for graduate students of theoretical physics.
5. Lastly, I provide several links as references here and there: I agree that this is a bad practice in academia and one should instead convert them to proper references in the bibliography. Nevertheless, it is faster for me to write this way and faster for the reader to just click on them where they appear, so I'll keep this practice. My upfront apologies if the links get broken in time: hopefully a snapshot will have been available here <https://web.archive.org/> (it would be rather amusing if this site itself becomes unavailable).

^aFor instance, I provide all links in their explicit form (not like [google](#) but as <https://www.google.com/>) so that the book on an e-reader without a browser or its printed version is as useful as its digital version.

About the format of the book

This book is publicly available online, and anyone can simply download and read it on their computer. Nevertheless, it is a lot easier on eyes to read a book on paper (or epaper), so I suspect many readers will simply print this book. To make this book convenient for all types of readers, I need to make a few design choices^a and the most important of it is the format of this book.

The most convenient paper that most students have easy access to is A4 (or letterpaper which is very similar in size): so if they choose to print this book, their easiest and cheapest option would be to use A4 paper. However that paper has been historically designed for typewriters which use large monospaced fonts, hence is not really appropriate for a digitally prepared book. Indeed, if you check your favorite-to-read book, you will most likely see that it uses a smaller paper size, with proportionally spaced small fonts.

What is wrong with using a large paper? It is empirically known that a document is properly legible if there are around 60-75 characters on a line: if there are more characters, it becomes harder to read and the reader may end up re-reading same line over and over again (doubling). When used with typewriter, A4 paper indeed has appropriate number of characters one a line, but as stated previously, this is not the optimal setup with the digital fonts, hence making A4 paper *too large for digital books*.

How to solve the problem that A4 is too large for a digital book? Obviously, we can use a typewriter font for which A4 is historically intended in the first place! However, such monospaced fonts (Courier being another example) are not aesthetically pleasing and do not belong to modern texts!

Of course, we can go with a modern proportionally spaced fonts but make the font size large enough such that a line has few enough characters for it to be easily legible. Although a better one than using monospaced fonts, this is still a suboptimal solution to the problem at hand...

Another solution which is somehow popular around the institutions is to use double-spacing among the lines. Indeed, regulations for master and doctorate thesis of various universities include compulsory large spacing among the lines, such as one and a half spacing or double spacing: you can also see this in my master and doctorate thesis: <https://arxiv.org/abs/1602.07676>, and <https://arxiv.org/abs/2107.13601>. Although this method can indeed prevent doubling to a degree, it is neither an aesthetic nor an efficient solution.

A somehow better solution than those listed above is to use multiple columns in the document. Indeed, this is the traditional approach in magazines and newspapers, and is immediately applicable in academic papers and manuscripts as well. However, A4 paper is not really big enough to have two columns of

text with around 60-75 characters (let alone three or more columns), and although one can go with smaller font sizes to make it more legible digitally, the printed version would still be hard to read either way (proper number of small font characters, or few number of normal font characters).

Common text editors such as Microsoft Word either go with one of the solutions above or do not solve the problem at all. On the contrary, L^AT_EX templates default to choose another approach: they stick to a modern font with an appropriate spacing and a single column, but they also increase the margins such that a line has proper number of characters. This is an ideal solution if the resultant text will be read digitally, however it leads to a waste of paper when printed.

In this book, we will not follow any of these design choices. Instead, we will go with the rather unorthodox *Tufte style*,^b an asymmetric allocation of the text in the paper. Indeed, the main text will be in the left of the paper, whereas we have another block of text on the right dedicated to the *sidenotes*,^c margin figures, and margin tables. We choose a rather narrow font family (*libertine*) and arrange

^bSee <https://www.ctan.org/tex-archive/macros/latex/contrib/tufte-latex/>

^cIn the traditional layout, one usually uses endnotes, margin notes, or footnotes; in this paper, most “notes” will be sidenotes with occasional footnotes.^d

^dSuch as this one.

the margins such that the main text is of 26 pica width and side text is of 14 pica width: for the 11pt and 9pt font sizes, this corresponds to roughly 66 and 44 characters of the libertine font for main and side text blocks respectively.^e Thus we have an ideally-sized main text block and acceptably-sized side text block for a modern proportionally spaced font in an A4 paper, and we do this without unnecessarily wasting the paper.^f

Literature recommendations

One can find many books, reviews, lecture notes, and articles about *conformal symmetry*, *conformal bootstrap program*, *quantum theory of fields*, *renormalization group*, *critical phenomena*, *AdS/CFT correspondence*, *cosmological bootstrap*, *flat space holography*, *string theory* and many other related concepts to conformal symmetry and its usage in theoretical physics. You can go ahead and search these keywords and see how large a literature there exists for each and every one of these topics.

So finding a source for conformal symmetry or its usage is no problem, it is the other way around: there are simply too many resources and one can feel overwhelmed and find it hard to navigate through those articles to learn the subject. So below, I'll list some resources that I would recommend both to naïve younglings who know nothing about these subjects and to those physics-hardened veterans who have heard of/worked on some of these subjects but haven't really got the chance to look at conformal symmetry directly. Please note that the list is *faaaaar* from being complete and is by no means supposed to include good resources and to exclude bad ones: I simply listed what I know (and would recommend) and it is extremely likely that there are reviews out there that would be far more suited yet are unbeknownst to me.

1. The traditional referencing source for conformal symmetry is the so-called *yellow book*, i.e. *Conformal Field Theory* by Francesco, Mathieu, and Sénéchal. It is a rather complete book, i.e. it covers all the basics, however it is published in 1997 and hence does not include recent developments captured by *conformal bootstrap program*. In short, conformal symmetry is a lot easier to solve in two spacetime dimensions and up until the last two decades all the major progress has been in two dimensions, hence naturally most of the book's content is for two dimensional conformal models.
2. The de facto referencing source for the conformal bootstrap program (which aims to constraint or solve conformal models in higher dimensions, i.e. $d \geq 3$) is the review article *The Conformal Bootstrap: Theory, Numerical Techniques, and Applications* by Poland (who happens to be my PhD advisor^g), Rychkov, and Vichi. The article really does not get into details (and not really

^eFor a nice discussion of these points along with the tools to compute approximate expected number of characters per line, see <https://ftp.cc.uoc.gr/mirrors/CTAN/macros/latex/contrib/memoir/memman.pdf>

^fI would like to acknowledge the following nice discussion with which I started to learn more about these *typographical* issues: <https://tex.stackexchange.com/questions/71172/why-are-default-latex-margins-so-big>.

^gThus I may be biased towards his work due to more exposure.

pedagogical), so it is not the best source to *learn* stuff. However, it is an excellent collection of sources, so one can use it to find other sources for individual topics.

3. The de facto lecture notes to learn the basics of the conformal symmetry is *TASI Lectures on the Conformal Bootstrap* by Simmons-Duffin (who happens to be my host during my 1-year visit to Caltech^h). Around similar times, Qualls shared his lecture notes *Lectures on Conformal Field Theory* as well (which is less focused on bootstrap but more on 2d CFTs).ⁱ A more recent review was given by Osborn, i.e. *Lectures on Conformal Field Theories in more than two dimensions*, however they are somehow more technical and less pedagogical.
4. Slava Rychkov maintains an active blog where he writes about his papers or his talks among other useful information, see <https://sites.google.com/site/slavyrychkov>. He also wrote lecture notes titled *EPFL Lectures on Conformal Field Theory in D= 3 Dimensions*, it is a bit older than Simmons-Duffin's lecture notes, but some may prefer the style. He also talked about the *philosophy* of the bootstrap approach as the 27th Ockham Lecture, see <https://www.merton.ox.ac.uk/event/27th-ockham-lecture-reductionism-vs-bootstrap-are-things-big-always-made-things-elementary>.
5. If you would like a *hardcore* resource, then there is the *holy book*.^j It is written by Dobrev, Mack, Petkova's, and Todorov in 1977 and is still extremely relevant today. The only problem is that it is rather mathematical (and technical), and you may find it hard – honestly I did not fully read the book myself, I only have used it as a reference to check stuff here and there.
6. If you would like a *layman review* of the conformal bootstrap, then I'd suggest the nature article of Poland and Simmons-Duffin, i.e. <https://doi.org/10.1038/nphys3761>.
7. If you prefer *watching* to *reading*, then there are recordings of excellent lectures in the summer schools of the conformal bootstrap collaboration.^k You can also view all videos of the collaboration at their Youtube channel:^l I personally advice 2017 & 2018 Bootstrap School videos as they were rather introductory and pedagogical!
8. These lectures will be mostly about conformal symmetry in $d \geq 3$ spacetime dimensions, hence $d = 2$ conformal symmetry (and related Virasoro symmetry) are not our focus. For those interested in $d = 2$ conformal symmetry (and its utility in String theory), I already named a few sources, but there are two books I would like to refer for sentimental reasons: *String and Symmetries*, edited by Gülen Aktaş, Cihan Saçlıoğlu, and Meral Serdaroglu; and *Conformal Field Theory: New Non-perturbative Methods*

^hThus I may be biased towards his work due to more exposure.

ⁱThere is also the lecture notes *Applied Conformal Field Theory* by Ginsparg. Honestly, I did not fully read the notes as they are rather focused on 2d CFTs which are not really my personal focus; however, some may find them quite useful.

^jSee <http://doi.org/10.1007/BFb0009678>.

^kSee
<http://bootstrapcollaboration.com/activities>

^lSee <https://www.youtube.com/channel/UCgWLG2q2275RuUJ5eNSCCFA/videos>.

ods In String And Field Theory, edited by Yavuz Nutku, Cihan Saçlıoğlu, and Teoman Turgut. The former book consists of the proceedings of the Feza Gürsey Memorial Conference 1 (1994); and the latter book consists of lecture notes of 1998 Summer Research Semester on Conformal Field Theories, *M*(atrix) Models and Dualities; both events held at Bogaziçi Üniversitesi-Tübitak Feza Gürsey Institute.

9. Lastly, I can recommend the first two chapters of my thesis as *an earnest attempt at a very pedagogical introduction* to conformal symmetry and the conformal bootstrap program. I tried to give a historical account of the subject, its basics, and recent developments. Even if you don't like it, you may find the references useful as I tried my best to organize and include as many references as possible. You can find it via the publisher's website^m or via arXiv.ⁿ

^mSee
<https://www.proquest.com/docview/2557212384/2C5A80A2ADB447E7PQ>.

ⁿSee <https://arxiv.org/abs/2107.13601>.

Contents

Preface	i
Remarks to the reader	i
About the format of the book	i
Literature recommendations	iii
Contents	vi
List of Tables	vii
List of Figures	viii
1 Why conformal symmetry?	1
1.1 Invitation: Critical phenomena	1
1.2 Connecting critical phenomena to conformal field theories	10
1.3 Using CFTs beyond statistical physics	23
2 What is conformal symmetry?	29
2.1 Invitation: how Galileon symmetry helps us derive the form for kinetic energy	29
2.2 Galilean symmetry and scale-invariance	31
2.3 Conformal algebra	37
2.4 Conformal algebra in the framework of QFTs	45
3 How can the conformal symmetry be utilized?	46
3.1 Review: how the Lorentz symmetry helps in quantum field theory	46
3.2 Extension to conformal symmetry	46
A Mathematical Preliminaries	47
A.1 The path from groups to algebras	47
A.2 Subgroups and subalgebras	50
A.3 Homomorphisms and automorphisms	53
B A brief review of distributions	57
B.1 Basics	57
B.2 More on intrinsic products of distributions	58
B.3 Subtleties with tensor products of distributions	60
Bibliography	63

List of Tables

1.1	Critical data for various fluids	9
1.2	Classification of fixed points	17
1.3	Ising Universality Class as a function of dimension of the space	19
1.4	Comparison of spaces with a maximal amount of symmetries	27

List of Figures

1.1	Phase diagram for fluids and uniaxial magnets	1
1.3	Uniaxial magnet as a set of one-dimensional spins on a lattice	2
1.2	Phase diagrams in three dimensions	2
1.4	Phase transitions in the Landau formalism	5
1.5	Critical opalescence	7
1.6	Uniaxial magnets near their Curie temperature	10
1.7	Example of a critical line: λ -line of ${}^4\text{He}$	11
1.8	RG flows for the $3d$ scalar field and the uniaxial magnet	15
1.9	IR equivalence	23
1.10	End points of renormalization flow and conformal field theories	25
1.11	Conformal symmetry in gauge theories	26
2.1	Constraint of Galilean symmetry on kinetic energy	30
2.2	Schematic relation between groups, fields, and algebras	31
2.3	Conformal transformations on S^n	39

1

Why conformal symmetry?

1.1 Invitation: Critical phenomena

A brief review of phase diagrams

Let us start with something we observe in everyday life and have been taught in kindergarten: matter exists in different phases. These phases can be differentiated by macroscopic parameters such as viscosity and compressibility; as we are all well aware by our everyday experiences, this translates into fluids having three different phases: *solid* (high viscosity & low compressibility), *liquid* (low viscosity & low compressibility), and *gaseous* (low viscosity & high compressibility). By changing the pressure and the temperature, we can change the phase of a fluid; the landscape of all phases of a given substance is shown in a *phase diagram*, e.g. Fig. (1.1).

The phase diagram of a system can be n -dimensional: the dimension of the diagram reflects the number of macroscopic parameters to tune with as we move between various phases. In the case of fluids, we have two parameters: *pressure P* and *temperature T*. Similarly, if we consider the phase diagram of a uniaxial magnet, it also has two parameters to tune: temperature and *external magnetic field h*.^o Other systems may have phase diagrams of different dimensions; for instance, [1] models biological structures as thermodynamic systems and classifies morphological patterns arising on the surface of tissues as various phases; in its application, the resultant phase diagram has three parameters, see Fig. (1.2).

One can naïvely think that we can always change the dimension of the phase diagram by considering other parameters; for instance, for fluids, we can also consider *molar volume V* as a parameter: the resultant graphical depiction of the phase diagram, along with its projection to various planes, can be seen in Fig. (1.2). It is true that such depictions can be quite useful, and one might conclude then that the phase diagram is in general of arbitrary dimension, as we can add more and more directions of arbitrary parameters. However, *what we mean by the dimension of the phase diagram* is the number of parameters necessary to tune to sweep all different phases. In the case of fluids, two parameters are sufficient to go over all phases, so we say that the phase diagram of fluids is two-dimensional, so is that of uniaxial magnets.

The dimension of the phase diagram, as a global property, may not be as so useful as *the dimension of the fixed points*, i.e. the number of parameters one needs to tune to reach a fixed point in the RG (renormalization group) framework. This is a local property, because it may differ from one fixed point to the other in the same phase diagram. In fact, the number of parameters necessary to reach a *critical fixed point* is same as the number of *relevant operators* (operators whose scaling dimension is less than the spacetime dimension) in the *conformal field theory* which describes that critical fixed point – more on that below! Note that we do not require knowledge of RG framework

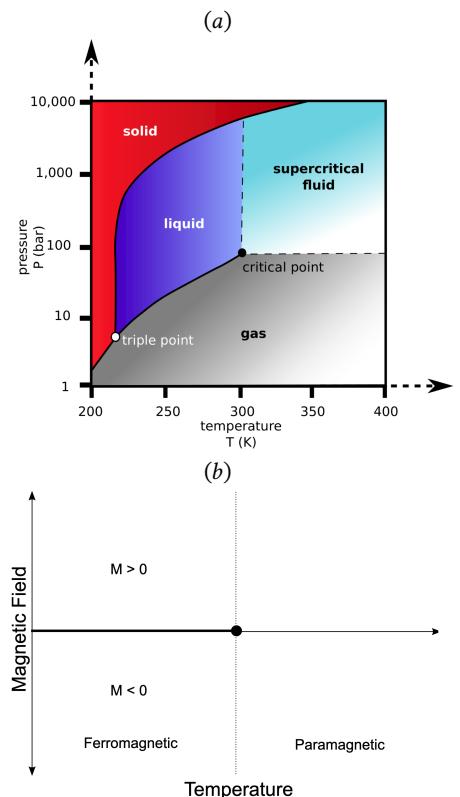


Figure 1.1: (a) Phase diagram for fluids (b) Phase diagram for uniaxial magnets

^oA uniaxial magnet is a magnet whose magnetization can point in only one dimension, see Fig. (1.3).

^pSee also <http://biomodel.uah.es/Jmol/plots/phase-diagrams/inicio.htm>.

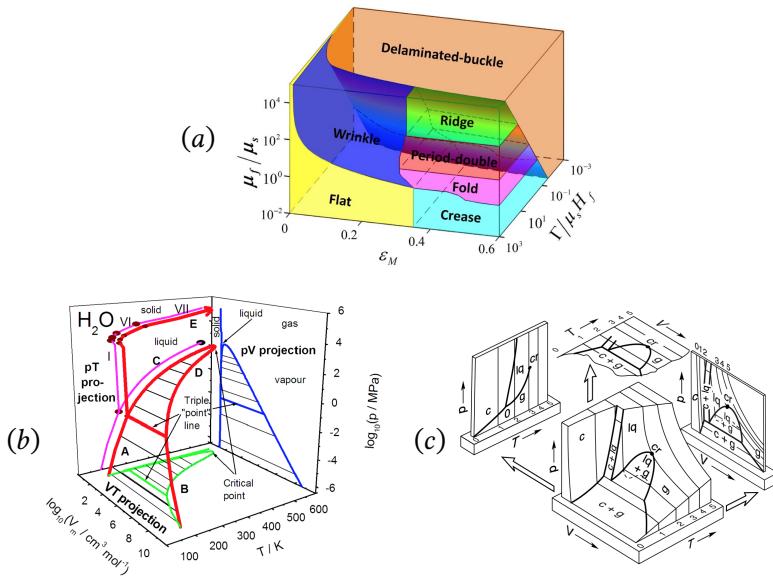


Figure 1.2: (a) A genuine three dimensional phase diagram example: various surface instability patterns induced on biological tissues [1]. One needs to tune three parameter to sweep all phases. (b) Three dimensional depiction of the phase diagram of water and its various projections [2]: one can still reach all phases by tuning only two parameters.^p(c) An illustrative orthographic (isometric) three-dimensional pVT diagram of water [2].

in these notes, so please remain comfortable if you did not understand what this paragraph is about.

Before we dive into critical fixed points or conformal field theories, let's take a step back and try to understand the phase diagrams in Fig. (1.1): they consist of some regions (bulk phases), lines dividing those regions (phase transitions), and endpoints of some lines (critical points). From statistical mechanics point of view, the phase diagram consists of *non-analyticities* of the free energy^q in thermodynamics limit,^r but actual computation of the free energy is almost always impossible so we do not have a clear handle on the derivation of the phases from the microscopic theory: in fact, the full phase diagram of a substance *can not be computed in general* starting from the microscopic theory! This is not because of our lack of knowledge or of a small technical problem, but of *the undecidable nature of the mathematical problem itself!*^s In short, we have the general formula but it will not do any good for us if we insist on computing the phases from the microscopic theory.

Let us instead construct a *phenomenological* model for the free energy.^t For that, we consider the free energy as a function of various phenomenological variables \$K_i\$ and write

$$e^{-\beta F(K_i)} = \int D\phi e^{-\beta L(K_i, \phi)} \quad (1.2)$$

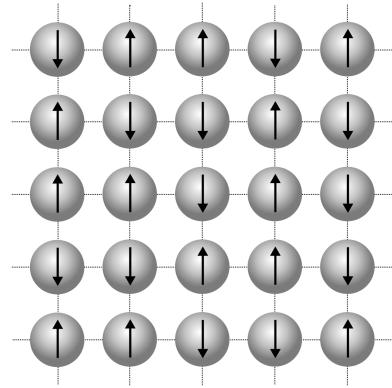


Figure 1.3: Uniaxial magnet as a set of one-dimensional spins on a lattice. If we examine this system at macroscopic scales, we do not observe individual spins but rather an average magnetization \$M(x)\$ which is a continuous field, hence we can describe it with a field theory. On microscopic scales, this theory is described by *the Ising model*.

^qFor those who have difficulty remembering their stat-mech courses, the free energy is the ultimate function from which all thermodynamic quantities can be derived via differentiation. I'll be rather ambiguous and do not specify *which* free energy I refer to (Helmholtz, Gibbs, Landau, etc.), because it will not matter for the qualitative discussion.

^rRemember that the free energy \$F\$ is given as

$$e^{-\beta F} = \text{Tr } e^{-\beta H} \quad (1.1)$$

for \$\beta \equiv (k_B T)^{-1}\$ where \$k_B\$ is the Boltzmann's constant and \$T\$ is the temperature. \$H\$ is the Hamiltonian and the trace denotes a sum over all degrees of freedom mentioned in \$H\$. The partition function \$Z\$ is analytic as the Hamiltonian is usually an analytic function and a *finite* sum of analytic functions is itself analytic, as long as the trace operation is a summation of *finitely many* degrees of freedom. Hence, the phases and phase transitions only appear in the thermodynamic limit, where the degree of freedom becomes infinite and the trace usually turns into a *path integral*.

^sHere by undecidable we mean that a general algorithm which is guaranteed to solve this problem in finite time cannot be found; see [3] for the case of phase diagrams. I highly recommend "Math Has a Fatal Flaw" video of the Youtube channel Veritasium for a broader

where $\int D\phi$ denotes a path integral over the field $\phi(x)$ which represents the degrees of freedom in the theory (Knowledge of path integrals is not essential for the discussion so just think of it as an ordinary integral if you are unfamiliar with the concept.). If we started with the microscopic theory, we would instead have trace over degrees of freedom, which would potentially include discrete sets (such as spins on lattice sites), but as we are constructing a *phenomenological model*, we are coarse-graining the microscopic details and representing the degrees of freedom by a continuous field $\phi(x)$.

The $L(K_i, \phi)$ is an arbitrary function for now, which would be the Hamiltonian in the microscopic theory. To leading order, we can assume that the dominant contribution to eqn. (1.2) comes from the form of $\phi(x)$ which minimizes $L(K_i, \phi)$, hence we have

$$e^{-\beta F(K_i)} = e^{-\beta L(K_i, \bar{\phi})} \quad (1.3)$$

where $\bar{\phi}(x) = \bar{\phi}$ is a uniform field that minimizes $L(K_i, \phi)$.

This phenomenological approach can then be summarized as follows:

- Determine a measurable macroscopic variable ϕ which changes over different phases: ϕ is called the *order parameter*.
- Write down an analytic function L of ϕ in terms of the phenomenological parameters K_i . L as a function of ϕ should have the symmetries of the Hamiltonian.
- Minimize L with respect to the variable ϕ : the minimum gives the free energy!
- If there is a non-analyticity in taking the minimum, that should give us the phase transition!

This approach is called *Landau formalism*, named after Lev Davidovich Landau, and $L(K_i, \phi)$ is usually called the *Landau potential*.^u Physically, Landau formalism is simply an application of the *mean field theory*: we approximate whatever it is of physical interest (such as magnetization field $M(x)$ of a magnet or density field $\rho(x)$ of the fluid) by its mean value and ignore the fluctuations!

Let us illustrate this approach for uniaxial magnets. The relevant parameter here is the magnetization $\vec{M}(x) = M\hat{z}$ of the magnet, which we take to be in the z -direction – note also that it takes its mean value M . Then the simplest form of Landau potential we can write down reads as

$$L(K_i, M) = K_0 + K_1 M^2 + K_2 M^4 - hM \quad (1.4)$$

which is analytic in M and contains the ferromagnetic coupling of the magnetic field M to the external magnetic field h .^v We note that the truncation of higher order terms (such as M^6) is not warranted in general, however it is valid as long as Landau formalism itself is applicable to the problem at hand.^w

view of the mathematical issue undecidability, along with incompleteness and inconsistency problems.

^uThe term “phenomenological” needs a little bit explanation. I use it the same way Goldenfeld uses in [4]. He has a beautiful explanation there and I would like to copy/paraphrase/edit some part of it here.

In constructing physical theories, we always use a certain level of description. Thus, in describing the long wavelength behavior of a magnet, we write down equations for the coarse-grained magnetization (here, coarse-grained means that the parameter is smoothed out below a certain detail). In describing the motion of a fluid, we write down equations of motion for the velocity field $v(r, t)$; however, the velocity field is actually a coarse-grained velocity of many particles in a small volume element in the neighborhood of the point r . In these and other examples, the variables of interest are always defined with respect to some coarse-graining process so that phenomena below some scale are subsumed somehow into the equations for the coarse-grained variables of interest.

In a phenomenological description of a system, we try to describe the behavior solely in terms of the coarse-grained variables, without reference to the microscopic physics on scales shorter than the coarse-graining length. Inevitably, the description of the coarse-grained variables introduces other parameters; for example, in a magnet, the parameters of the Landau free energy, or in a fluid, the coefficient of viscosity. These phenomenological parameters are determined by the microscopic physics: for example, the viscosity of a fluid may be calculated, with some approximation, from kinetic theory. A successful phenomenological theory contains only a finite number of such parameters (the smaller the better), and does not attempt to calculate the phenomenological parameters. These are taken to be inputs to the theory.

How does a phenomenological theory change when the microscopic physics is altered? From the above, the only possible change can be in the values of the phenomenological parameters. If new phenomenological parameters have to be introduced whenever the microscopic physics is changed, then the theory is not, by definition, phenomenological! In this sense, a successful phenomenological theory should be insensitive to the changes in the microscopic physics, although the phenomenological parameters may change.

An interesting corollary of this discussion is that all our physical theories are phenomeno-

As K_0 simply shifts the free energy, we can discard it.^x Likewise we can rescale M to absorb K_2 (which needs to be a positive number for L to have a minimum), hence we have

$$L(a, M) = aM^2 + M^4 - hM \quad (1.5)$$

which gives the free energy by minimization:

$$\{\bar{M}, F(a, h)\} = \begin{cases} \left\{-\frac{h}{2a}, -\frac{h^2}{4a}\right\} + H.O.T. & a > 0 \\ \left\{\pm\sqrt{\frac{-a}{2}}, -\frac{a^2}{4}\right\} + H.O.T. & a < 0 \end{cases} \quad (1.6)$$

where we wrote down the result for weak external magnetic field h for simplicity, and $H.O.T.$ refers to terms higher order in h .

It can be checked by looking at the full form of the result that the function is analytic unless $h = 0$. If $h = 0$, then there is a cusp at $a = 0$, i.e. second derivative of $F(a)$ has a discontinuity at $a = 0$. And for $a < 0, h = 0$, there are in fact two minima of the Landau potential, hence two possible choices for the magnetization. We can summarize these findings as follows:

1. There is no phase transition unless $h = 0$ ($F(a, h)$ is analytic there)
2. There are two minima of the Landau potential for $a < 0$ if $h = 0$: we have a *spontaneous symmetry breaking*. Physically, this means for $h = 0$ that there is no magnetization if a is positive and there is a magnetization $M \neq 0$ if a is negative but the sign of M can be either \pm . So we have the *disordered phase* for $a > 0$ and *ordered phase* $a < 0$, and there is a phase transition at $a = 0$ as a is varied at $h = 0$.
3. For $a < 0$, we have a degeneracy among the minimum at $h = 0$, but if we change h the degeneracy is lifted. Hence, if we keep $a < 0$ fixed and vary h from $h > 0$ to $h < 0$, the minimum jumps from $M > 0$ to $M < 0$: there is a non-analyticity for the whole line $a < 0$ at $h = 0$!
4. The first derivative of the free energy is discontinuous in h at $h = 0$ for any value of $a < 0$: the phase transition associated with this whole line is called *first order phase transition*!
5. The free energy and its first derivative are continuous at $a = 0, h = 0$: it is the second derivative that brings the non-analyticity when we change a . Hence we call the associated phase transition a *continuous phase transition*!^y

The situation is best understood with the aid of visuals, so please refer to Fig. (1.4).

We would like to note that Landau formalism is actually incorrect for $d \leq 4$ — see footnote w — and we should not use it for anything quantitative. However, the picture it depicts is actually qualitatively

logical, because in principle there may always be more degrees of freedom in finer and finer details!

^xHere I simplified the phenomenological Landau theory, interested reader may consult [4] for an excellent review.

^yThe sign needs to be negative for a ferromagnet so that parallel M and h decreases the energy, as opposed to antiferromagnets coupling $+hM$.

^wIn short, the Landau formalism is a sufficient description only above an *upper critical dimension*; physically, ignored fluctuations become important if the spacetime dimension is not high enough: for uniaxial magnets, the upper critical dimension is 4. If we are above the upper critical dimension, then the truncation of higher order terms, i.e. M^6 , is appropriate because their contribution diminishes as we near the critical point. This is actually closely linked to the *renormalization* of quantum fields and suppression of irrelevant terms as the energy scale tends to infinity. [Maybe I can put more explanation here or even show how this is true, i.e. \$t\$ versus \$\Lambda^{-1}\$.](#)

^xObservable quantities such as susceptibility are derivatives of the free energy.

^yIn the literature, such a transition is also called a *second-order phase transition*. This follows from Paul Ehrenfest's classification of phase transitions: he proposed that *phase transitions could be classified as n^{th} order if any n^{th} derivative of the free energy with respect to any of its arguments yields a discontinuity at the phase transition*. This classification is incomplete because there are thermodynamic quantities such as the specific heat that actually diverge at some phase transitions rather than exhibiting a simple discontinuity as the Ehrenfest classification implies. Thus, we will take a more modern approach and classify phase transitions as *first-order* or *continuous* as we did above.

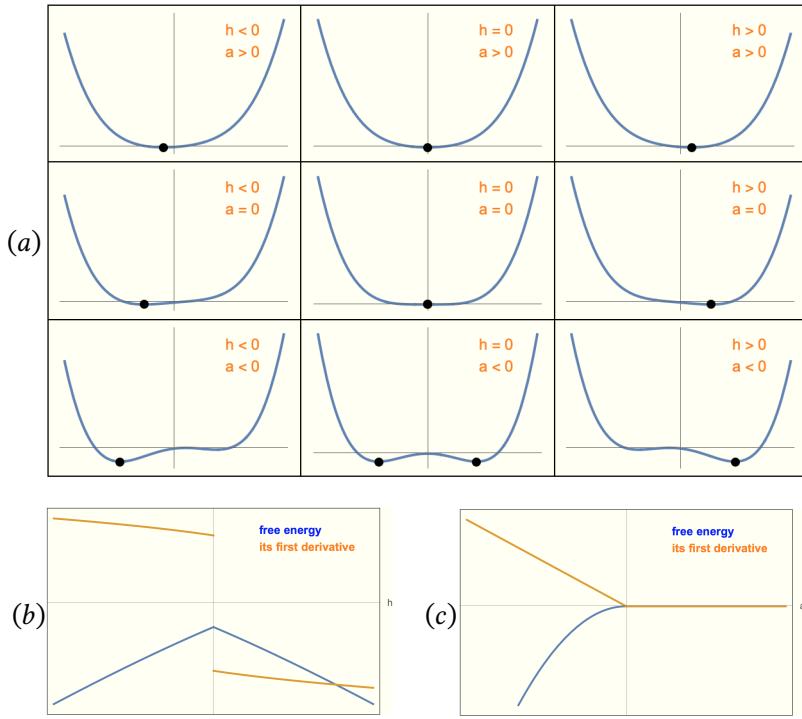


Figure 1.4: (a) The Landau potential $aM^2 + M^4 - hM$ and its minima with respect to the magnetization M for various values of the phenomenological variable a and the external magnetic field h . We see that there is no ordered phase at $h = 0$ if $a \geq 0$, i.e. the minimum is at $M = 0$. If there is external magnetic field, i.e. $h \neq 0$, then there is a net magnetization M which has the same sign as h as we have the ferromagnetic coupling $-hM$. On the other hand, we still have a net magnetization M even if $h = 0$ as long as $a < 0$; however, the minima are degenerate so we get a *spontaneous symmetry breaking*, i.e. M can take either positive or negative sign. However, even an infinitesimal h lifts the degeneracy so the minima *jumps* from one value to the other if h is changed continuously from $h < 0$ to $h > 0$ at $a < 0$ (this corresponds to the movement over last column): as the minimum moves discontinuously, we have an associated *first order phase transition*. On the contrary, if we move from $a > 0$ to $a < 0$ at $h = 0$ (this corresponds to the movement over middle row), the minimum changes from zero to a nonzero value continuously: there is a *continuous phase transition* at $a = h = 0$. (b) The landau potential at constant $a < 0$ as a function of h . We see that its first derivative has a discontinuity, showing that the phase transition is indeed *first-order*. (c) The landau potential at constant $h = 0$ as a function of a . We see that both the function and its first derivatives are continuous, hence the phase transition is indeed *continuous*.

correct: if we take that the phenomenological constant a to have a dependence $a \sim T - T_c$, the phases and the phase transitions described in Fig. (1.4) indeed matches the empirical phase diagram in Fig. (1.1).

Universality, critical opalescence, and critical exponents

Why did we review the phase diagrams in the previous section: what are we trying to achieve? Well, as the author, I'm trying to motivate conformal field theories but we are still not there yet! But if we go back to our original discussion, our motivation was to understand the nature of the phase diagrams, such as those in Fig. (1.1). We found out that in the case of uniaxial magnets, there is a continuous phase transition at the so-called *critical point* ($h = 0, T = T_c$) and there is a first order phase transition along the line which connects the critical point to the origin ($h = 0, 0 \leq T < T_c$). We might now ask how important our findings are: after all, uniaxial magnets (magnets whose magnetization can point only in 1 dimension) are highly-idealized and we may naïvely complain that these results do not have a wide range of applicability! However, this is simply not true! The description of and around the critical point is actually *universal*, and whatever information we can acquire regarding the critical point of the uniaxial magnet immediately applies to the critical point of most fluids! This concept is called *universality* and is related to the fact that *same conformal field theory* describes the critical point of both the uniaxial magnets and the most fluids!

The concept of universality can be best understood in the framework of renormalization group, and we'll review that in the next section. For now, let us focus on the implications of universality instead of deriving it:

- Critical phenomena, i.e. behavior observed at and around the critical points (i.e. continuous phase transitions) are *independent* of the microscopic theory! In other words, *local* properties of different systems around their critical points are same! This is in contrast to the *global* properties, i.e. the whole phase diagram, which varies from system to system!
- Critical phenomena depends on the *dimensionality* of the space and the *symmetries* of the order parameter! In fact, we can define *universality classes* based on these! For example, if the order parameter has \mathbb{Z}_2 symmetry,^z then we call that *Ising universality class* (why Ising, more below). For instance, both the density fluctuation ρ of a fluid and the magnetization M of a magnet has \mathbb{Z}_2 symmetry and fall under Ising universality class: they have identical critical behavior!

To summarize, phase diagrams of numerous materials include critical points which has continuous phase transitions, and the behavior of the system at and around the critical points can be categorized into a handful of different universality classes, which means understanding the critical behavior of one system helps us understand the critical behavior of bunch of irrelevant-looking other systems! *But what exactly are we referring to by critical behavior?*

Let us investigate what happens to a fluid near its critical point: if we start with a constant volume of ethane and heat it, it will become

^z \mathbb{Z}_2 symmetry refers to *evenness* of a system; for instance, the function x^2 has \mathbb{Z}_2 symmetry because it is invariant under $x \rightarrow -x$.

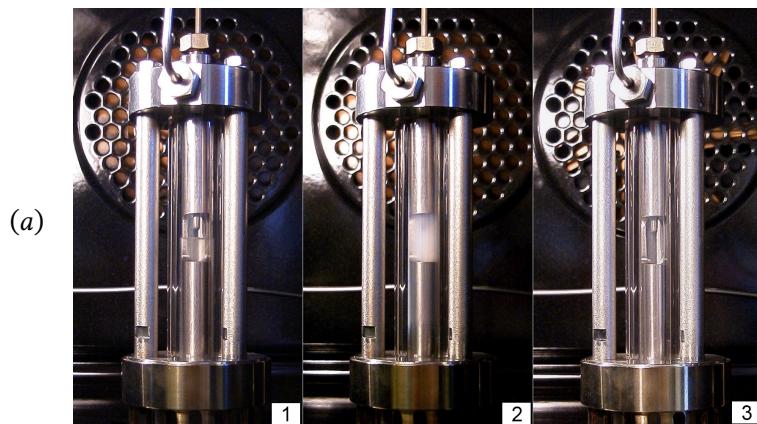


Table 1. Critical properties of various solvents (Reid et al., 1987)

Solvent	Molecular mass	Critical temperature	Critical pressure	Critical density
	g/mol	K	MPa (atm)	g/cm ³
Carbon dioxide (CO ₂)	44.01	304.1	7.38 (72.8)	0.469
Water (H ₂ O) [†]	18.015	647.096	22.064 (217.755)	0.322
Methane (CH ₄)	16.04	190.4	4.60 (45.4)	0.162
Ethane (C ₂ H ₆)	30.07	305.3	4.87 (48.1)	0.203
Propane (C ₃ H ₈)	44.09	369.8	4.25 (41.9)	0.217
Ethylene (C ₂ H ₄)	28.05	282.4	5.04 (49.7)	0.215
Propylene (C ₃ H ₆)	42.08	364.9	4.60 (45.4)	0.232
Methanol (CH ₃ OH)	32.04	512.6	8.09 (79.8)	0.272
Ethanol (C ₂ H ₅ OH)	46.07	513.9	6.14 (60.6)	0.276
Acetone (C ₃ H ₆ O)	58.08	508.1	4.70 (46.4)	0.278
Nitrous oxide (N ₂ O)	44.013	306.57	7.35 (72.5)	0.452

(b)

Figure 1.5: (a) Ethane of constant volume being heated: when it passes the critical point, the critical opalescence is observed, i.e. it becomes *milky*! (b) The critical point of various solvents. Despite the difference in critical temperature and critical pressure, the same phenomenon of critical opalescence is observed in all fluids: see Figure 1.1 for the critical point of fluids at the end of the evaporation line between liquid and gaseous phases.

opaque at the critical point! If we keep heating it so that it passes the critical point, it will become transparent again, see Fig. (1.5). This phenomenon is called *critical opalescence*. This becoming-opaque-at-critical-point feature is by no means special to ethane: we experimentally observe that fluids at their critical point become opaque in general!

Why is that? At the simplest level, we can understand it intuitively by a phenomenological model of fluids. Let's assume that a scalar field $\rho(x)$ denotes the density of the fluid at position x .^{aa} By continuity of this field, we expect that the densities $\rho(x)$ and $\rho(y)$ are correlated if y is around x : if the fluid is dense at a point x , it should be dense in its neighborhood as well.

As y gets further away from x , we expect the correlation to die down if the interactions in the fluid are short-ranged, hence we expect

^{aa}As this is a phenomenological model, we are coarse-graining the microscopic details, so $\rho(x)$ can be thought of as an average density, i.e. number of molecules in the sphere of radius Λ^{-1} centered at x divided by the volume of the sphere. If we take $\Lambda \rightarrow 0$, $\rho(x)$ becomes a constant and takes its mean value, which gets us back to the mean field theory (i.e. Landau framework). On the other hand if we take $\Lambda \rightarrow \infty$, we get rid of coarse-graining and get back to the microscopic theory (such as a lattice model). The analysis of the effect of the change in Λ is precisely the RG framework which we will review in section 1.2.

the correlation function^{ab} to have the form

$$\langle \phi(x)\phi(y) \rangle \sim (x-y)^p e^{-(x-y)/\xi} \quad (1.7)$$

where $\phi(x) \equiv \rho(x) - \bar{\rho}$ is the *fluctuation* in the density field and ξ is called the *correlation length*^{ad}

The correlation length is in general a function of the thermodynamic variables, such as the temperature. As we near a critical point by changing the temperature (or whatever the relevant variable is), the correlation length increases: in fact, at the critical point itself, the correlation length becomes infinity. Hence, the interactions in the fluid becomes long-ranged, and the fluctuations become rather strong, scattering various wavelengths of light and turning the substance into an opaque form!

To summarize, while investigating the critical behavior, we found out that the correlation length diverges at the critical point. In fact, we have a whole family of divergent quantities at the critical point, and it would be insightful for us to go over them briefly. However, to understand them better, let us first list what we have defined so far and define further quantities:^{af}

- t – reduced temperature^{ag}
- J – source field (e.g. reduced pressure for fluids and the external magnetic field h for magnets)
- f – free energy
- C – specific heat
- ϕ – order parameter (e.g. reduced density for fluids and magnetization M for magnets)
- χ – response function (e.g. compressibility for fluids and susceptibility for magnets)
- ξ – correlation length

Observable parameters such as specific heat C or compressibility χ of a fluid are in general complicated functions of thermodynamic variables; however, *they show a universal behavior and diverge near a critical point as a power law*:

$$C \propto |t|^{-\alpha}, \quad \chi \propto |t|^{-\gamma}, \quad \xi \propto |t|^{-\nu} \quad (1.10)$$

where the exponents α, γ, ν are called *critical exponents*^{ah,ai}. These exponents form the set of parameters that describe the critical behavior.

Earlier, when we looked at the implications of universality in page 6, we stated that the critical phenomena are independent of the microscopic theory and that it depends only on the *universality class* of the substance. What we meant there is precisely the values of these critical exponents: they *are* universal! For instance, we list examples of several fluids whose critical point lie in *Ising Universality Class* in Table (1.1). Despite the wild differences in critical temperature, they all exhibit the

^{ab}In these notes, we are using Dirac's bra/ket notation in extremely abusive manner! In eqn. (1.7), it refers to the expectation value with respect to Boltzman weight $e^{-\beta H}/Z$ as the probability; hence, $\langle A \rangle = Z^{-1} \int D\phi A e^{-\beta H[\phi]}$ for the partition function $Z = \int D\phi e^{-\beta H[\phi]}$. Thus, if $A = \phi\phi$, it becomes a correlation function of ϕ . Note that in the Landau framework we have the Landau potential L instead of the microscopic Hamiltonian H . This path integral form is extremely similar to what we use in high energy physics for quantum fields: there, we use the bra/ket notation for the similar quantity $\langle A \rangle = Z^{-1} \int D\phi A e^{-S[\phi]}$ for the vacuum generating function^{ac} $Z = \int D\phi e^{-S[\phi]}$ and the Euclidean action S — note that we are setting $\hbar = 1$ in these notes. For Lorentzian signature, we define the path integral through a Wick rotation.

We will use Dirac's bra/ket notation to denote conformally invariant structures as well, but there will not be any ambiguity: more on this later!

^{ac}If you haven't studied quantum field theory yet, the vacuum generating function Z is the function which can be used to compute any correlation function we like (we do this by taking functional derivatives of Z). By LSZ formalism, one can then extract scattering amplitudes or decay rates from the correlation functions; in short, full knowledge of Z is sufficient to extract most of what we are interested in computing in high energy physics.

^{ad}It will not be important for the present discussion but let us give some further information on the correlation function $\langle \phi(x)\phi(0) \rangle$ — the correlation function $\langle \phi(x)\phi(y) \rangle$ can be obtained from $\langle \phi(x)\phi(0) \rangle$ through the translation invariance of it, i.e. $\langle \phi(x)\phi(y) \rangle = \langle \phi(x-y)\phi(0) \rangle$. The exact analytic form of the correlations functions are not known in general (well, blame mathematicians not physicists) and are not really computable unless the model is exactly soluble (which is extremely rare). However, we can compute their asymptotic forms (either in weak coupling or in long distance separation by perturbation theory, or at the critical point by symmetry arguments). For fluids, we have

$$\begin{aligned} \langle \phi(\mathbf{x})\phi(0) \rangle &\sim \frac{1}{\xi^{d-2}} \frac{e^{-|\mathbf{x}|/\xi}}{(|\mathbf{x}|/\xi)^{(d-1)/2}} \left(1 \right. \\ &\left. + \frac{(d-1)(d-3)}{8(|\mathbf{x}|/\xi)} + \mathcal{O}\left(\frac{1}{(|\mathbf{x}|/\xi)^2}\right) \right) \end{aligned} \quad (1.8a)$$

as $|\mathbf{x}| \rightarrow \infty$ and ξ finite, and

$$\langle \phi(\mathbf{x})\phi(0) \rangle \sim \frac{a^\eta}{|\mathbf{x}|^{d-2+\eta}} \quad (1.8b)$$

as $\xi \rightarrow \infty$. Here d is the dimension of the space and a is a microscopic length scale (such as lattice spacing). See [5] for details on the first

Table 1.1: Data of various fluids for their critical points at the end of liquid-vapor transition

Substance	T_c (K)	ρ_c (kg/m^3)	Source
Deuterium (D_2O)	644	356	[7]
Sulfur hexafluoride (SF_6)	319	739	[8]
Carbon dioxide (CO_2)	304	468	[9]
Difluoromethane (HFC-32)	351	424	[10]
Pentafluoroethane (HFC-125)	339	570	[10]
Dinitrogen (N_2)	126	314	[11]
Neon (Ne)	44	484	[11]

same exact behavior around the critical point with the critical exponents

$$\alpha \approx 0.11, \quad \beta \approx 0.33, \quad \gamma \approx 1.24, \quad \delta \approx 4.79, \quad \eta \approx 0.04, \quad \nu \approx 0.63 \quad (1.11)$$

Note that these are the critical exponents in three dimensional space, hence are for *3d Ising Universality Class*: universality class does depend on the dimension of space hence these numbers change if we change the dimension of the space.

As we stated earlier, the uniaxial magnet at its Curie temperature with zero external magnetic field is also at the critical point described by Ising Universality Class, hence it exhibits exactly the same critical behavior with the same critical exponent. As we show in Fig. (1.6), various magnets have different Curie temperatures, but the divergence they exhibit is same.

Summary

We are now at page 9 but haven't really motivated why conformal field theories are worth studying in statistical physics, apart from mentioning its name here and there. We will get to that in the next section, but before that, let's finish setting the scene by summarizing our discussion so far.

- Matter has different phases (duh!) and all phases can be shown in an n -dimensional diagram called *phase diagram* where axes are thermodynamic variables such as temperature, pressure, magnetization, etc.
- Phase diagrams have *non-analyticities* in them, which divide them into different regions (phases). These non-analyticities may (but usually do) include a phase-transition which ends at a critical point.
- Phase transition lines are usually associated with a *first-order phase transition*, i.e. at least one of the first derivatives of the free

formula. The second formula is the *definition* of the critical exponent η : more on this below.

The Fourier transform of the two point function gives the *structure factor* $S(\mathbf{k})$ which is proportional to the intensity of the light scattered through an angle θ off the fluid where $|\mathbf{k}| = \frac{4\pi}{\lambda} \sin(\theta/2)$ and where λ is the wavelength of the incident light. At the critical point, this then means intensity $\sim |\mathbf{k}|^{\eta-2}$ ^{ae} indicating it diverges as $|\mathbf{k}| \rightarrow 0$ for $\eta < 2$, explaining the critical opalescence [4]. For a detailed discussion, see [6].

^{ae}This follows from dimensional analysis:

$$\begin{aligned} \int d^d x e^{-i\mathbf{k}\cdot\mathbf{x}} |\mathbf{x}|^{-d+2-\eta} &\propto \int_0^\infty |x|^{1-\eta} e^{-i|\mathbf{k}||x|\cos\theta} d|x| \\ &\propto |\mathbf{k}|^{\eta-2} \int_0^\infty y^{1-\eta} e^{-iy\cos\theta} dy \quad (1.9) \end{aligned}$$

^{af}Wikipedia does a good job at defining and listing them, see https://en.wikipedia.org/wiki/Critical_exponent.

^{ag}When we say reduced for a thermodynamic quantity x , we mean $\frac{x-x_c}{x_c}$ where x_c is its value at the critical point, e.g. reduced temperature $t \equiv \frac{T-T_c}{T_c}$.

^{ah}We have actually more critical exponents. If $t < 0$, we can also define $\phi \propto (-t)^\beta$. Likewise, if $t = 0$, we define $J \propto \phi^\delta$ and $\langle \phi(\mathbf{x})\phi(0) \rangle \propto |\mathbf{x}|^{2-d-\eta}$ (see eqn. (1.8b)).

^{ai}The critical exponents α, γ, ν have different values depending on $t > 0$ or $t < 0$.

^{ak}See Wikipedia and the references therein: https://en.wikipedia.org/wiki/Curie_temperature.

energy is discontinuous! Such phase transitions can be explained by spontaneous symmetry breaking, but this *does not have to* be the case! For instance, in magnets, this is the case: above and below the transition line are two different ground states of the potential due to breaking of the \mathbb{Z}_2 symmetry in the order parameter (magnetization M). In contrast, two sides of the evaporation line in fluids (liquids and gaseous) have the same Euclidean symmetry.^{al}

- Critical points at the end of phase transition lines have *continuous phase transition*:^{am} the correlation length at such points diverge! The critical surfaces does not have to be points (despite we have been talking about critical *points* so far), for instance, in ${}^4\text{He}$, there is a superfluid transition along the so-called λ -line and the phase transition along the whole line is a continuous phase transition, see Fig. (1.7).
- The behavior of systems around their critical points are independent of the microscopic details of the system but rather on the symmetries and the dimensionality of the space. This in effect divides critical phenomena into *universality classes*: understanding the critical behavior of one system is sufficient to understand the critical behavior of all systems in the same universality class (e.g. understanding the behavior of a uniaxial Nickel magnet at its Curie temperature is sufficient to understand the behavior of compressibility of water or heat capacity of carbon dioxide at their critical points)!

In the next section, we will see how these results relate to conformal field theories!

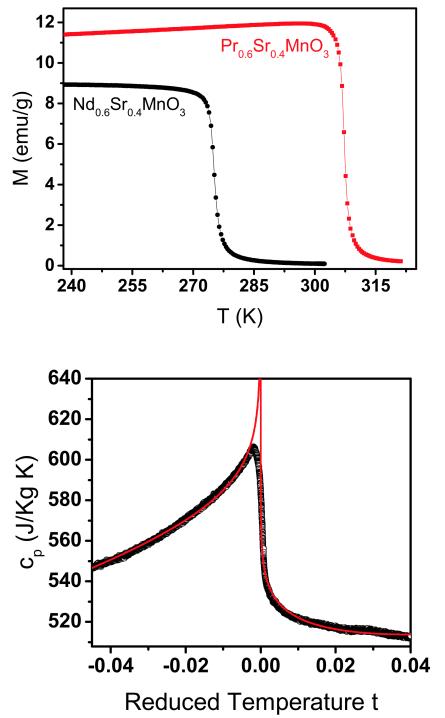
1.2 Connecting critical phenomena to conformal field theories

A brief review: renormalization group formalism

In the previous section, we concluded with the universality of the critical phenomena and how the correlation length diverges at the critical point. In this section, we will try to see why this is so and how this relates to conformally invariant theories.

^{al}It is interesting to note that all phases of matter are described by the same microscopic Hamiltonian (and the same partition function), but they may have different symmetries. For instance, in solids, the translation symmetry is broken (atoms are locked in specified positions in a lattice), so the symmetries of the fluid changes over the melting line. As a side note, a corollary of this is that we expect the melting line to extend indefinitely (unlike the evaporation line) because solids have a different symmetry group than liquids and that cannot change analytically! In contrast, liquids and gaseous already have the same symmetry group and it is perfectly fine for the evaporation line to end at a point, i.e. critical point.

^{am}I'm not sure if there are exceptions, i.e. if there are endpoints of some transition lines where no continuous phase transitions are associated with those points.



Material	Curie temperature (K)
Iron (Fe)	1043
Cobalt (Co)	1400
Nickel (Ni)	627
Gadolinium (Gd)	292
Dysprosium (Dy)	88
Manganese bismuthide (MnBi)	630
Manganese antimonide (MnSb)	587
Chromium(IV) oxide (CrO_2)	386
Manganese arsenide (MnAs)	318
Europium oxide (EuO)	69
Iron(III) oxide (Fe_2O_3)	948
Iron(II,III) oxide (FeOFe_2O_3)	858
$\text{NiO}-\text{Fe}_2\text{O}_3$	858
$\text{CuO}-\text{Fe}_2\text{O}_3$	728
$\text{MgO}-\text{Fe}_2\text{O}_3$	713
$\text{MnO}-\text{Fe}_2\text{O}_3$	573
Yttrium iron garnet ($\text{Y}_3\text{Fe}_5\text{O}_{12}$)	560
Neodymium magnets	583–673
Alnico	973–1133
Samarium–cobalt magnets	993–1073
Strontrium ferrite	723

Figure 1.6: The behavior of magnets near their critical point for magnets in 3d Ising Universality Class [12]: transition from ordered to disordered phase, the divergence of the specific heat, and Curie temperatures for various magnetic substances (top to bottom).^{ak}

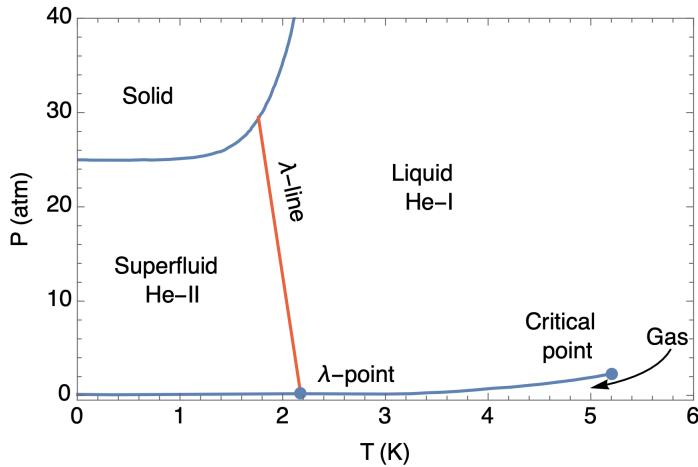


Figure 1.7: An example of critical line: the phase transition for ^4He along its whole λ -line is a continuous phase transition. Figure taken from [13].

As we stated earlier, the concept of universality can be best understood in the framework of renormalization group flows but I'll assume that the reader may not be familiar with it so I'll briefly summarize the main points of it qualitatively.

Let's remember our phenomenological model for the free energy in eqn. (1.2)

$$e^{-\beta F(K_i)} = \int D\phi e^{-\beta L[K_i, \phi]} \quad (1.12)$$

This is a phenomenological model, so the field $\phi(x)$ is coarse-grained below a certain length scale, say 1 micrometer. This means we cannot resolve anything below 1 micrometer but do so above it.

We have seen that the critical phenomena (such as critical opalescence in fluids) happen at a rather macroscopic scale, so let's coarse-grain all the fluctuations of the order parameter $\phi(x)$ between 1 micrometer and 1 millimeter. To do so, we separate *all functions* $\phi(x)$ as those which only fluctuate above 1 millimeter (call them $\phi_\ell(x)$) and those which fluctuate between 1 millimeter and 1 micrometer (call them $\phi_h(x)$).^{an} As $\int D\phi$ is a path-integral, it becomes $\int D\phi_h \int D\phi_\ell$,^{ao} hence we now have

$$e^{-\beta F(K_i)} = \int D\phi_\ell e^{-\beta L'[K_i, \phi_\ell]} \quad (1.13)$$

for

$$e^{-\beta L'[K_i, \phi_\ell]} = \int D\phi_h e^{-\beta L[K_i, \phi]} \quad (1.14)$$

If all possible terms allowed by the symmetries of the Hamiltonian are included in the effective potential L , then L' should not have any new terms; hence we can write $L'[K_i, \phi_\ell] = L[K'_i, \phi_\ell]$, thus we have

$$e^{-\beta F(K_i)} = \int D\phi e^{-\beta L[K_i, \phi]} = \int D\phi_\ell e^{-\beta L[K'_i, \phi_\ell]} \quad (1.15)$$

^{an}Here ℓ and h refer to low and high frequencies respectively. As a side note, it is actually a lot more straightforward to do this separation in Fourier space

^{ao}For anyone not familiar with the path-integrals, this may seem odd: we divided the whole integration to two parts, so they may expect $\int D\phi \rightarrow \int D\phi_h + \int D\phi_\ell$. One intuitive (yet not necessarily rigorous) way to see that this isn't the case is to consider the fact that the general function ϕ has both low and high frequency parts so it is not like $\int_a^b dx = \int_a^0 dx + \int_0^b dx$ but rather like $\int dxdy = \int dx \int dy$.

where the first integration is coarse-grained over 1 micrometer and second over 1 millimeter. Of course, there is nothing special about these values and we can generally write

$$e^{-\beta F(K_i)} = \int_{\Lambda} D\phi e^{-\beta L[K_i(\Lambda), \phi_{\Lambda}]} \quad (1.16)$$

where the integration is coarse-grained over distances Λ^{-1} and the phenomenological constants in the Landau potential are now functions of the parameter Λ .

What can we do with this parameter Λ ? On one hand, we can try to take it to infinity (effectively removing any coarse-graining process): this indicates we are trying to probe the microscopic theory and our phenomenological model will eventually break down; indeed, in statistical physics, we cannot take Λ bigger than the inverse lattice spacing, i.e. a^{-1} . In high energy physics, we do not have a natural microscopic length scale (assuming we are focusing on standard model without gravity) hence we may naively try to take Λ to infinity. In this limit, our theory may or may not fail; it is the experiments that ultimately would invalidate (or not) our models.^{ap} On the other hand, we can take Λ to zero: this means we are *integrating-out* all degrees of freedom, leaving us with the ultimate IR behavior of the theory!^{aq}

The change of Λ by *integrating-out* high frequency modes^{ar} as we have demonstrated above is called *renormalization group flow*, i.e. RG flow.^{as} The main take-away of this argument is that under the change of the *regularization parameter* Λ , the phenomenological parameters K_i change, i.e. $K_i = K_i(\Lambda)$.

The RG flow as we have briefly introduced above does nothing by itself: it does not compute anything, nor does it give any obvious simplification. In practice, one *combines* a computational technique with RG flow to *improve* that technique; for instance, one can combine RG flow with the perturbation theory to yield *RG improved perturbation*. We do not need to know anything about these in these notes.^{at} For our purposes, all that matters is that the phenomenological constants K_i in our Landau model changes with the regularization scale Λ .

Let's try to extend this formalism from statistical mechanics to high energy theory: in a d -dimensional Euclidean QFT,^{au} the vacuum generating function Z is given as^{av}

$$Z[K_i] = \int D\phi e^{-\int d^d x \mathcal{L}[K_i, \phi(x)]} \quad (1.17)$$

where $\phi(x)$ collectively denotes all fields, and K_i are usually called *coupling constants*.^{aw} We do not have a natural regularization parameter here as we did in the Landau theory (or inverse lattice spacing a^{-1} as in statistical mechanics), so the integration is naively over all length scales. However, this is too ambitious: as we stated in footnote ap, we can actually never write down a theory which is *guaranteed* to work at all length scales (according to our current understanding). Thus we have two choices: if our Lagrangian is *nonnormalizable*, we already

^{ap} If our model is so-called *nonnormalizable*, it will surely fail as we keep increasing Λ ; in fact, we may be able to predict beyond which point our model is unreliable. An example of this is Fermi's four-fermion theory of the weak interaction which breaks down around $\Lambda \sim 300$ GeV. On contrary, if our model is so-called *renormalizable*, then it is self-consistent in the phenomenological sense and hence is insensitive to microscopic theory: we can safely take Λ to ∞ . However, the resultant theory is *not necessarily the correct model for the phenomena it is written to describe*. In other words, despite being self-consistent at any scale Λ , *the predictions* of a renormalizable theory may diverge from the experimental values after a certain Λ , which indicates a modification of our model is necessary beyond a certain Λ . The fact that renormalizable theories cannot tell their regime of validity, along with the other fact that nonnormalizable theories necessarily break down beyond an energy scale, leave us with one firm conclusion: *we can never write down a physical theory which is guaranteed to work at all length scales!*

^{aq} As is commonplace in high energy and statistical physics, we will refer the behavior of a system at very low and very high energies as its IR and UV behavior, where they stand for infrared and ultraviolet (i.e. low and high frequency light).

^{ar}Equivalently coarse-graining low distance details.

^{as}In the full RG flow, we actually also rescale after coarse-graining, and do these infinitesimally so that there is *indeed* a continuous flow of coupling constants under RG.

^{at}How RG improves perturbation theory is a standard topic, and one can consult their favorite QFT book. In short, instead of having a perturbation expansion with the small parameter $g \log \left(\frac{p^2}{m^2} \right)$, we do a perturbation expansion with the small parameter g_{Λ} which is computed by RG flow of the parameter g . The second expansion is improved because log terms can get large for large momenta p^2 .

^{au}The argument is same for Lorentzian QFTs as well.

^{av}See footnote ab for a reminder on the vacuum generating function Z .

need a momentum cut-off Λ because our model is guaranteed to fail above a certain threshold. If our Lagrangian is *renormalizable*, then in principle we do not need a cut-off for consistency,^{ay} though the model may still break down (i.e. cannot explain experiments) after a certain energy threshold Λ . In short, for both cases, it is logical to keep an arbitrary regularization scale Λ in the integration, reminding us that we are not integrating over modes with higher frequency than arbitrary Λ :

$$Z[K_i] = \int_{\Lambda} D\phi e^{-\int d^d x \mathcal{L}[K_i(\Lambda), \phi_{\Lambda}(x)]} \quad (1.19)$$

Such theories are called *effective field theories*, and we just argued that all quantum field theories are in a sense effective field theories.^{az}

Let's get back to our main take-away both in statistical mechanics and high energy physics: the coupling constants in the Lagrangian (or Hamiltonian in statistical mechanics) change as we change the regularization scale Λ . But what exactly does that mean?

To understand the situation better, let us consider a single scalar field $\phi(x)$ in three space dimensions with the Lagrangian:

$$\mathcal{L} = \frac{1}{2}(\partial\phi)^2 - \frac{m^2}{2}\phi(x)^2 - \frac{\lambda}{4!}\phi(x)^4 - h\phi(x) \quad (1.20)$$

We are describing a *massive* scalar field which self-interacts with the interaction strength λ . We observe three things:

1. We could have written $h(x)$ instead of h and treat the magnetic field $h(x)$ as a classical background field (i.e. it does not have a kinetic term).^{ba} Rather, h is simply a coupling constant in this setting, such as m^2 and λ .
2. The Lagrangian has \mathbb{Z}_2 symmetry (same symmetry with Ising Universality Class) if $h = 0$, i.e. it is invariant under $\phi(x) \rightarrow -\phi(x)$.
3. We could add other terms consistent with the \mathbb{Z}_2 symmetry to the Lagrangian, i.e. $\phi(x)^6$. The absence of such terms are yet not warranted but we'll argue later that it is a consistent and reasonable omission.

Let us now apply an RG transformation: as we discussed above, this will change the coupling constants (in this case m^2 , λ , and h) as we iterate the RG flow. If we start with *small* m^2 , λ , and h for $d > 4$,^{bb} we can show that λ gets smaller with RG: we call such coupling constants *irrelevant*.^{bc} On the other hand, m^2 and h are relevant coupling constants, i.e. they get larger under iterations of RG flow.^{bd} In the coordinate system where the coupling constants are axes, which is actually precisely the phase diagram we have been analyzing in the previous sections, relevant coupling constants form the directions that need to be tuned to reach the critical point; for instance, for $d > 4$, $(h, m^2, \lambda) = (0, 0, 0)$ describes the critical Ising model (m^2 plays the role of the critical temperature, much like a did in eqn. (1.5) for the Landau description of the

^{aw}In general, the Lagrangian^{ax} has the form

$$\mathcal{L}[K_i, \phi(x)] = \sum_i K_i \mathcal{O}_i(x) \quad (1.18)$$

where $\mathcal{O}_i(x)$ are *operators* in the theory and K_i are coupling constants, i.e. terms that couple the operators to the Laplacian. The operators can be of various forms, simple examples are $\mathcal{O}_1(x) = \phi(x)$ and $\mathcal{O}_2(x) = \phi(x)^2$. If the theory is free or we are perturbing around a free theory, such operators are not really independent (I can imagine some readers already shouting “one is the square of the other!”). However, if the theory is strongly coupled, it may not be possible to write down $\mathcal{O}_i(x)$ simply as $\phi(x)$ or $\phi(x)^2$; the very same operators that were $\phi(x)$ and $\phi(x)^2$ in the weak coupling regime can get radically different in the strong coupling regime: we will see an example of that in the *3d Ising model* below.

^{ax}In these notes we are abusing the language and refer both to the Lagrangian and the Lagrangian density as Lagrangian for short, which is a fairly common practice in the field.

^{ay}There are infinities if we don't regularize our integral, but in a normalizable theory these infinities can be removed. So, in effect, even if we put a cut-off Λ , we can consistently take $\Lambda \rightarrow \infty$; but we do not need to put a cut-off in the first place (we can regularize the infinities in other ways, such as dimensional regularization).

^{az}A related concept here is the *decoupling principle*. It basically states that if we have a very heavy particle in our model, its existence in low energies is only reflected through the modification of the coupling constants of lighter particles. In other words, in low energies, the heavy particles *decouple* from the effective description of the phenomena (in the jargon, their propagators are said to be *frozen*, or that they can be *integrated out*). As a corollary of the decoupling principle, we can never know certainly if there are heavier particles which have not been yet discovered by doing experiments at low energies; because if there are such particles, their contribution to the observables are indistinguishable from their absence with a different set of coupling constants of lighter particles.

^{ba}From QFT point of view, that makes much more sense as we can use $h(x)$ as a source term to take functional derivatives of the vacuum generating function Z to con-

Ising model) and we need to tune h and m^2 to reach this point ($\lambda \rightarrow 0$ by itself under RG). On the other hand, if $d < 4$, things get a little bit complicated: λ is also relevant now hence $(h, m^2, \lambda) = (0, 0, 0)$ no longer describes the Ising universality class: it now has 3 relevant directions hence 3 parameters need to be tuned (unlike the fluids or uniaxial magnet).^{be} However, if d is *very close to 4* (say $d = 3.99$), then λ is barely relevant and we can expect a new critical point close to $(0, 0, 0)$, which corresponds to Ising Universality class.^{bf} In fact, in their 1971 paper titled *Critical exponents in 3.99 dimensions* [15], Wilson and Fisher found such a point and computed the critical exponents as expansion of the small parameter $\epsilon = 4 - d$. One can reach this critical point (1) by tuning m^2/λ ^{bg} so that one is on the *critical surface* and then (2) by applying the RG flow which will take the system away from the Gaussian point and towards this new point: see Fig. (1.8) for the diagrammatic illustration of this procedure along with the RG flows on the phase diagram of the $3d$ scalar.

Thus, in summary, the description of a system (the values of coupling constants) change by at which scale we look at the system (different regularization Λ), and the modification in the description of the system as we change our perspective can be shown as a flow in the phase diagram for the system: this is called *RG flow* and Fig. (1.8) illustrate this for a scalar field and uniaxial magnet.

Let's comment more on Fig. (1.8a). We observe that there are four fixed points for the RG flow; *the Gaussian fixed point*, i.e. $\lambda = h = m^2 = 0$,^{bh} two sinks^{bi} at $m^2 = \pm\infty$, and *the Wilson-Fisher fixed point* at $(m^2, \lambda) = (m_c^2, \lambda_c)$. In comparison, Fig. (1.8b) has five fixed points: one Gaussian, two sinks, one critical fixed point (at Curie temperature), and one so-called discontinuity fixed point at $T = \infty$. We review types of different fixed points in Table (1.2) at the end of this section.

Let's take a step back and review our situation: we stated above that the $3d$ scalar theory at Wilson-Fisher fixed point is in $3d$ Ising Universality Class, but what exactly is the Lagrangian for this theory? Can we compute the critical exponents analytically and exactly? Even before all these, what exactly do we know about this point?

Our questions are rather valid and to the point (pun intended), but the answers will be disappointing. We remind the reader that we started *around* the Gaussian fixed point $m^2 = \lambda = h = 0$, and said that Wilson and Fisher found a new fixed point nearby. This procedure is in fact no coincidence: we have no way to map out the RG flow exactly and completely, and we have no way of doing actual computations (like literal computation of the path integrals and such) unless we are nearby the Gaussian fixed point. Thus, necessarily, we start around the Gaussian fixed point and look for new fixed points as Wilson and Fisher did.

The previous paragraph is rather disturbing: it basically means we somehow stumbled upon the Wilson-Fisher fixed point, and maybe there are a lot more fixed points that are unbeknownst to us! Indeed, by the conventional way of doing things (which I'll list below), there is no way we can guarantee that we do not have new fixed points, and

on the chosen UV regularization of the theory [16].

struct correlation functions. Our main point here is not that so we are simplifying the situation for a clearer discussion.

^{bb}In these notes, d always denotes the dimensionality of the space unless stated otherwise.

^{bc}More correctly, we call a coupling constant (or the direction it forms on the phase diagram) *irrelevant for the point p* if RG flow takes the coupling constant *towards p*. For instance, for the point $\lambda = 0$, the coupling constant λ is irrelevant because RG takes λ closer and closer to 0. On the contrary, if the RG transformation takes the value *away* from point p , then that direction is said to be *relevant for the point p*. For instance, for the point $h = 0$, the coupling constant h is relevant because RG takes any nonzero h further and further away from 0. If RG moves a variable x neither towards to nor away from a point p , then x is said to be a *marginal direction* for point p .

For the concept of being relevant and irrelevant to be mutually exclusive for a point p , the RG transform should not *reach* to that point in finitely many iterations; otherwise, it can move towards the point p in the first n iterations and move away from it in the remaining iterations, making the direction both irrelevant and relevant! Thus, the points p that we referred above should be *fixed-points*: points that RG transformation reaches in *infinitely* many iterations. Also, if we are *exactly* at those fixed points, then we stay at them under RG, i.e. they are *invariant points of RG flow*!

In summary, relevant and irrelevant directions are respectively the unstable and stable directions of the fixed points under perturbation.

^{bd}See footnote bc.

^{be}Remember that we should tune temperature T and external magnetization h for magnets and temperature T and pressure p for fluids.

^{bf}One motivation for this is called *cross-over* phenomena. Basically it says that despite λ is relevant, we still spend a lot of time around the critical point during RG (as λ is barely relevant), hence we can effectively treat λ as an irrelevant parameter. The end of section 5.4

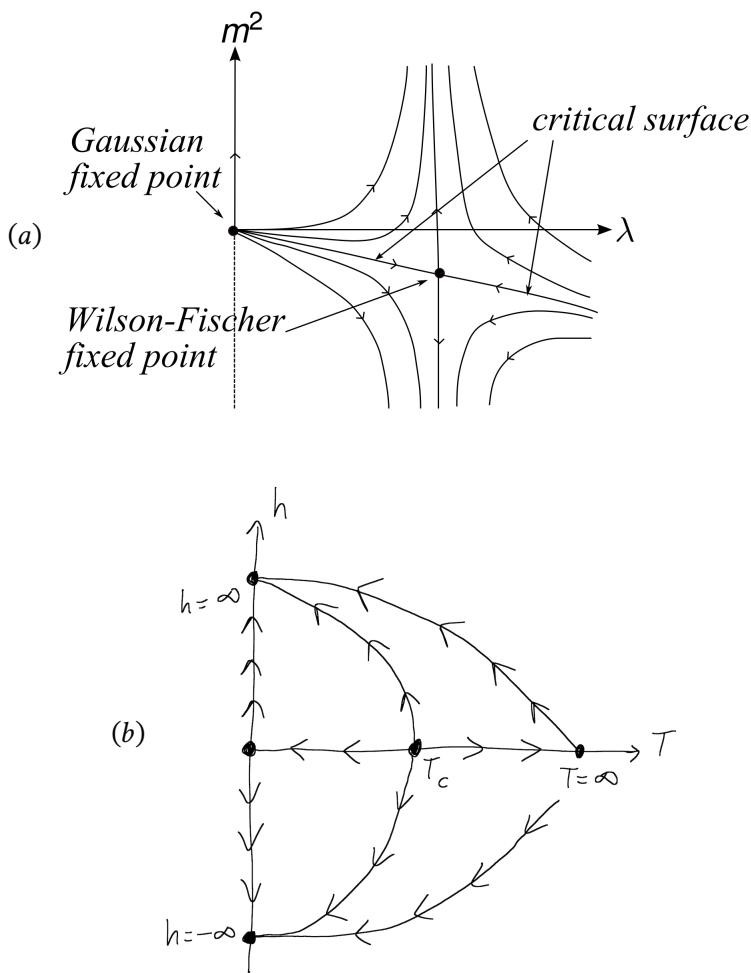


Figure 1.8: (a) Phase diagram of the 3d scalar field with \mathbb{Z}_2 symmetry, with the RG trajectories shown. We see that the coupling constants move towards four possible fixed points under RG transformation: two sinks at $m^2 = \pm\infty$, one Gaussian fixed point, and one interacting fixed point called *Wilson-Fisher fixed point*. (b) Phase diagram for uniaxial magnet, with the RG trajectories shown (see also Fig. (1.1)). There are five fixed points: one Gaussian fixed point at $h = T = 0$, two sinks at $h = \pm\infty$, one interacting fixed point at $\{T, h\} = \{T_c, 0\}$, and one non-interacting fixed point at $T = \infty$.

hence there is no guarantee that Wilson-Fisher fixed point is the correct one to describe the 3d Ising Universality Class. The more modern^{bj} way of mapping the fixed points of a phase diagram is in fact to *look for* conformal field theories with correct number of relevant operators and the correct symmetry group (more on this below). For instance, it is shown in [17] by conformal bootstrap techniques that “...3D Ising CFT is the only \mathbb{Z}_2 -symmetric CFT in 3 dimensions with exactly two relevant operators...”, which indicates that if the Wilson-Fisher fixes point is describing the 3d Ising CFT, then there can be no other fixed points

of [14] touches on this, in fact simply see the whole chapter for a thorough discussion.

^{bg}The precise value of this ratio to reach the critical point is not unique and depends
^{bh}It is called so because the Lagrangian is then Gaussian.

^{bi}These fixed points are called sinks because they do not have any relevant direction: a system never moves away from sink under RG flow!

^{bj}This adjective may be unfair to other methods, but with my limited knowledge this is how I see it.

with two relevant operators which describe a \mathbb{Z}_2 -symmetric CFT in the whole phase diagram!^{bk}

Before detailing this conformal bootstrap approach, we need to fully appreciate why this approach is powerful *and necessary* in the first place! To do so, let's give some general remarks and comment on a few points:

- Strongly coupled quantum field theories (or statistical models) are not soluble!^{bl} Not only do we not have the required mathematical tools to solve them, but also do we not have the mathematical framework to *hope* to solve them in future!^{bm}
- If we cannot analytically solve the problem, we can try numerically solving it on a computer! For that, we put the theory on a lattice (we need a discretization to employ the computer) and simulate its behavior. The branch of the physics that adopts this approach is called *lattice simulations*.^{bn} Despite its relative success, this approach is limited because (a) the computing power is limited and (b) the complexity of the problem can be rather high to utilize this approach!^{bo}
- We can solve free theories.^{bp} As we cannot solve strongly coupled theories, we can assume weak-coupling and do perturbation theory: this is basically what we do for most of the standard model!^{bq}
- If we cannot take the coupling constants to be small, we can try to vary other parameters and try to find a free theory this way. For instance, instead of considering a scalar field $\phi(x)$, we can consider N copies of it: most of the time the theory is free if $N \rightarrow \infty$; we can then (1) do an expansion around that free theory, (2) compute corrections order by order in $1/N$, (3) assume that truncation at some order is fine, (4) take $N = 1$ (or whatever the original number of copies) at the end. This procedure is called *large N expansion*, see [19] for a 200-page review of this concept!
- Another parameter that we can vary to approach a free theory is the dimension of the space itself! As we noted above, there exists a fixed point (Wilson-Fisher fixed point) if $d < 4$ and this fixed point merges with the Gaussian fixed point as d approaches 4. As we know how to deal with the Gaussian fixed point, we can take $\epsilon = d - 4$ to be small so that the Wilson-Fisher fixed point is close enough for us to use our Gaussian fixed point knowledge to analyze that critical point (and this is how that critical point was discovered in the first place)! This approach is called *epsilon expansion*, and in a sense similar to large N expansion: we expand in ϵ , compute first few terms, truncate the expansion, and take $\epsilon \rightarrow 1$ to get the critical point of 3d Ising model.^{br}

The items above show how desperate our situation actually is as theoretical physicists! Nevertheless, do not despair! We can still make

^{bk}There is always a chance that there is a fixed point with scale symmetry without conformal symmetry, but this is quite unlikely: we'll elaborate on this later.

^{bl}Of course, this is a generic statement and there are exceptions. However, to my knowledge, they are either in low dimensions (such as 1d or 2d models) or have additional symmetries (such as supersymmetry of conformal symmetry). [Put references here](#)

^{bm}We are talking about the impossibility of finding an algorithm to solve them *generically* because of the undecidable nature of the mathematical problem (see footnote ^s).

^{bn}[Add some sources here](#)

^{bo}For instance, the problem of solving 3d Ising model on the lattice is shown to be an *NP-complete problem* in 2000 [18]; basically meaning that it is not feasible in practice to try to solve this model on a lattice!

^{bp}Gaussian path-integrals are the only path-integrals we can compute exactly. [This is certainly true in standard statistical mechanics and QFT context but I wonder if there are exotic functions which can be analytically computed as well, maybe I can look into that!](#)

^{bq}We are rather lucky that the couplings of the electroweak theory (describing, electric, magnetic, and weak interactions) are rather small at the typical energies we observe the universe, so that we can use perturbation theory to describe the world successfully!

^{br}There is a whole bunch of literature on the validity of this approach and how to improve it. [Add references](#)

Table 1.2: Classification of fixed points

# of relevant directions	ξ	Type of fixed point	Physical Domain	Explanation
0	0	Sink	Bulk phase	\ddagger
1	0	Discontinuity FP	Plane of coexistence	\ddagger
1	0	Continuity FP	Bulk phase	$\ddot{\delta}$
2	0	Triple point	Triple point	$\ddot{\eta}$
2	∞	Critical FP	Critical manifold	σ^*
> 2	∞	Multicritical point	Multicritical point	\ddagger
> 2	0	Multiple coexistence FP	Multiple coexistence	$\ddot{\eta}$

progress in the realm of strongly coupled theories by realizing that fixed points are actually scale-invariant, hence they enjoy a higher symmetry than the rest of the phase diagram!

Let me expand on the scale invariance. Under RG transformation, the correlation length ξ scales linearly, i.e. it behaves as $\xi \rightarrow \ell\xi$. Thus, at the fixed points of the RG transformation, ξ is either 0 or ∞ . As a corollary, *any point at the phase diagram where the correlation length diverges is a fixed point of the RG flow!* Hence, by our discussion in the previous section we can conclude that the critical phenomena happen at some of the fixed points of the RG flow! *In short, understanding the nature of fixed points and the emergent scale invariance is invaluable to understanding critical phenomena!*

Let us end our scandalously lightening review of RG flow and turn to the relation to the conformal field theories in the next section; but before that, we present Table (1.2) as a general reference: it contains a brief categorization of fixed points according to their dimensionality and the value of correlation length, see Table 9.1 of [4] for the original/better version of it.

Dimensional counting and scale invariance

Let us consider the most generic form of the Lagrangian we gave in eqn. (1.18), i.e. $\mathcal{L}[K_i(\Lambda), \phi(\Lambda, x)] = \sum_i K_i(\Lambda) \mathcal{O}_i(\Lambda, x)$. The action then reads

$$S = \sum_i \int_{\Lambda} d^d x K_i(\Lambda) \mathcal{O}_i(\Lambda, x) \quad (1.21)$$

where Λ is the *momentum cut-off*, i.e. the regularization parameter.^{bt} If we scale $x \rightarrow \ell^{-1}x$, then the lower limit Λ^{-1} gets scaled as well, hence we get

$$\int_{\Lambda} d^d x K_i(\Lambda) \mathcal{O}_i(x) = \int_{\ell^{-1}\Lambda} d^d (\ell^{-1}x) K_i(\Lambda) \mathcal{O}_i(\Lambda, \ell^{-1}x) \quad (1.22)$$

If we further scale $\Lambda \rightarrow \ell\Lambda$, we then obtain the relation

$$K_i(\Lambda) \mathcal{O}_i(\Lambda, x) = \ell^{-d} K_i(\ell\Lambda) \mathcal{O}_i(\ell\Lambda, \ell^{-1}x) \quad (1.23)$$

In general, the coupling constants' functional form can be rather complicated; after all, this is what determines the flows in Fig. (1.8). However, near critical point, we can *linearize* the action of RG flow

\ddagger : No relevant directions, hence trajectories only flow to them: they correspond to *stable bulk phases*, and the nature of the coupling constant characterize the phase; e.g., in 3d Ising model with nearest neighbor ferromagnet interaction in external field h , there are two sinks at $h = \pm\infty \& T = 0$, indicating the stable bulk phase is a net magnetization in all temperatures, whose sign is determined by sign of h .

$\ddot{\delta}$: Correspond to points on a phase boundary and describe a *first order phase transition* where an order parameters exhibits discontinuous behavior (*We note that the interpretation of RG transformations near first order phase transformations is delicate* and this explanation is rather simplified, see [20] for a careful discussion). E.g., all points on the line $h = 0, T < T_c$ in a ferromagnet flow to a discontinuity FP at $h = 0, T = 0$. This FP is unstable towards the sinks. Lastly, this is the only fixed point where one of the eigenvalues of the real space RG flow is ℓ^d .^{bs}

$\ddot{\sigma}^*$: Represents a *phase of the system*, e.g. $h = 0, T = \infty$ which attracts all points on the line $h = 0, T > T_c$ in a ferromagnet. Phases described by this FP are unstable towards a sink.

$\ddot{\sigma}$: We explained this whole section.

$\ddot{\eta}$: Higher-dimensional generalization of what we have explained this whole section.

$\ddot{\eta}$: Nothing interesting.

^{bt}Let ℓ be a coarse-graining factor and d is the dimension of the system. If one of the eigenvalues of the RG flow near a critical point is ℓ^d , then that critical point is either a first-order or a discontinuity fixed point! All other fixed points have $y < d$ for the eigenvalue ℓ^y , see Table 1.2. This is called *Nienhuis-Nauenberg criterion*, see [21] for further details.

^{bs}As we stated earlier, we need this parameter in statistical mechanics and in non-normalizable quantum field theories. For renormalizable qft models, we can take it to infinity self-consistently but let's keep it for the sake of argument.

to leading order, hence the coupling constants takes a homogeneous form:^{bu}

$$K_i(\ell\Lambda) = \ell^{[K_i]} K_i(\Lambda) \quad (1.24)$$

where $[K_i]$ is called the *mass dimension* of K_i .^{bu} This then indicates $\mathcal{O}_i(\Lambda, x) = \ell^{[K_i]-d} O_i(\ell\Lambda, \ell^{-1}x)$. This equation suggests that \mathcal{O}_i also transforms homogeneously if we write it in terms of right variables. Instead of writing it as $\mathcal{O}(\Lambda, x)$, let us write it as $\mathcal{O}(\Lambda, \Lambda x)$ for the energy scale Λ and the dimensionless parameter Λx . Thus

$$\mathcal{O}_i(\ell\Lambda, \Lambda x) = \ell^{d-[K_i]} \mathcal{O}_i(\Lambda, \Lambda x) \quad (1.26)$$

Clearly $\mathcal{O}_i(\Lambda, \Lambda x)$ is a homogeneous function in its first variable:

$\mathcal{O}_i(\ell\Lambda, \Lambda x) = \ell^{[\mathcal{O}_i]} \mathcal{O}_i(\Lambda, \Lambda x)$ for the *mass dimension* of the operator \mathcal{O}_i . We then conclude

$$[K_i] + [\mathcal{O}_i] = d \quad (1.27)$$

In footnote bc, we categorized coupling constants by their relevance around the critical point, i.e. whether they need to be tuned or not. There, we said that an irrelevant coupling constant vanishes by RG flow around the Gaussian fixed point. We see from eqn. (1.24) and eqn. (1.27) that we can determine the relevance of coupling constants by the scaling dimension of the operator that accompanies them:

$$\mathcal{O} \text{ is called } \begin{cases} \text{a relevant} \\ \text{an irrelevant} \\ \text{a marginal} \end{cases} \text{ operator if } \begin{cases} [\mathcal{O}_i] < d \\ [\mathcal{O}_i] > d \\ [\mathcal{O}_i] = d \end{cases}. \quad (1.28)$$

Therefore, we can ignore all irrelevant operators \mathcal{O} in the Lagrangian if we are interested in studying the critical behavior, because their contribution goes to zero as we near the critical point. This is precisely why we truncated higher order terms, i.e. ϕ^6 , in eqn. (1.20).^{bw}

Around the Gaussian fixed point (i.e. the free theory), we can write down operators in terms of the fundamental field in the Lagrangian; for instance, we can take^{bx}

$$\mathcal{O}_1 = \phi, \mathcal{O}_2 = \phi^2, \mathcal{O}_3 = \phi^4, \mathcal{O}_4 = \phi^6, \mathcal{O}_5 = T_{\mu\nu} = g_{\mu\nu}\mathcal{L} + \partial_\mu\phi\partial_\nu\phi \quad (1.30)$$

where $T_{\mu\nu}$ is the stress tensor of the theory and \mathcal{L} refers to the Lagrangian in eqn. (1.20).^{by} As the Lagrangian includes a term $(\partial\phi)^2$ without a coupling constant in the front,^{bz} eqn. (1.27) indicates $[\partial\phi] = d/2$, hence $[\phi] = \frac{d-2}{2}$, thus this gives us

$$[\mathcal{O}_1] = \frac{d-2}{2}, [\mathcal{O}_2] = d-2, [\mathcal{O}_3] = 2d-4, [\mathcal{O}_4] = 3d-6, [\mathcal{O}_5] = d \quad (1.31)$$

We see that \mathcal{O}_1 and \mathcal{O}_2 are always relevant and that the stress tensor is always marginal. On the contrary, \mathcal{O}_3 (\mathcal{O}_4) is relevant only if $d < 4$ ($d < 3$), indicating that we were right to drop \mathcal{O}_4 around $d = 4$ to get the Wilson-Fisher fixed point at page 15.

^{bu}We call a function homogeneous in its n^{th} argument if it satisfies $f(\dots, ax_n, \dots) = a^k f(\dots, x_n, \dots)$.

^{bw}Actually, we are glossing over an important detail here. When we linearize RG flow, it takes the form of a matrix equation, i.e.

$$\begin{pmatrix} K_1 \\ K_2 \\ \vdots \end{pmatrix}' = M \begin{pmatrix} K_1 \\ K_2 \\ \vdots \end{pmatrix} \quad (1.25)$$

The coupling constants themselves do not simply scale as in eqn. (1.24) individually, but rather their combinations as eigenvectors of M does so. In fact eigenvalues of M determines how they'll scale as well, i.e. their mass-dimensions! However, as this detail is not really necessary for the main discussion, I'll assume in the text that M is diagonal hence K_i can satisfy eqn. (1.24).

^{bw}There is an important detail that we are glossing over: *dangerously irrelevant operators*. Some operators cannot be just dropped even though their contribution vanishes as we approach the critical point. This is simply because the partition function Z (or vacuum generating function in high energy physics) may not be analytic in that limit. To understand this, let's consider $Z = Z[K_r, K_i, K_d]$ for relevant K_r , irrelevant K_i , and dangerously irrelevant K_d coupling constants. Under RG, K_r gets bigger so we keep it there. On the contrary, $K_i \rightarrow 0$, hence we can replace it with 0: $Z = Z[K_r, 0, K_d]$. Similarly, K_d goes to zero as well, but we cannot replace it because the function Z is *not analytic* around $K_d = 0$. Instead, it behaves like this:

$$\lim_{K_d \rightarrow 0} Z[K_r, 0, K_d] = \lim_{K_d \rightarrow 0} \frac{1}{K_d^a} Z'[K_r] \quad (1.29)$$

for some parameter $a > 0$. In short, an irrelevant coupling constant is called *dangerously irrelevant* if the partition function is singular as that term vanishes.

This detail is not really important for our current discussion so we will omit any subtlety related to this.

^{bx}In QFT, we should actually define our operators after normal ordering, i.e. $\mathcal{O}_2(x) = : \phi(x)\phi(x) :$, $\mathcal{O}_3 = : \phi(x)\phi(x)\phi(x) :$, etc. This is to ensure that we can put multiple operators at the same spacetime point without any divergence (by normal ordering, we are actually removing the divergence). For instance, the correlator $\langle \phi(x)\phi(x) \rangle = \infty$ whereas $\langle : \phi(x)\phi(x) : \rangle = 0$. This detail is not important for the main discussion in the text.

The mass dimensions of the operators play critical role (pun intended) in the explanation of the critical phenomena. The critical exponents given in eqn. (1.11) can in fact be extracted from the mass dimension of the two coupling constants, hence from the mass dimensions of the accompanying operators by eqn. (1.27). Indeed, for Ising Universality Class, we can write down^{ca}

$$\begin{aligned}\alpha &= \frac{d - 2[\epsilon]}{d - [\epsilon]}, & \beta &= \frac{[\sigma]}{d - [\epsilon]}, & \gamma &= \frac{d - 2[\sigma]}{d - [\epsilon]}, & \delta &= \frac{d - [\sigma]}{[\sigma]}, \\ \nu &= \frac{1}{d - [\epsilon]}, & \eta &= 2[\sigma] - d + 2\end{aligned}\tag{1.32}$$

where σ and ϵ are standard symbols used for the relevant operators in the conformal field theory that describes the critical point of Ising Universality Class.

A careful reader may have noticed a tension between the two statements we have made:

1. The critical exponents listed in eqn. (1.11) ($\alpha \approx 0.11, \delta \approx 4.79, \dots$) can be extracted from the scaling dimensions of the operators via eqn. (1.32), i.e. $\alpha = \frac{d-2[\epsilon]}{d-[\epsilon]}$, $\delta = \frac{d-[\sigma]}{[\sigma]}$, etc.
2. The mass dimensions of the operators are given by eqn. (1.31), i.e. $[\mathcal{O}_1] = \frac{d-2}{2}$ and $[\mathcal{O}_2] = d-2$.

Clearly, the second point ensures that the critical exponents computed from them are *rational* numbers; for instance, if we take $\sigma = \mathcal{O}_1$ and $\epsilon = \mathcal{O}_2$, we compute $\alpha = \frac{4-d}{2}$ and $\delta = \frac{d+2}{d-2}$, which do not match the experimental values!

Well, we can resolve the problem by noting that Gaussian fixed point describes Ising universality class only if $d > 4$ and the experimental values are computed at $d = 3$. Indeed, if we were to do experiments at higher dimensions, we would confirm that the critical exponents are rational values! However, for $d = 3$, the correct critical point is Wilson-Fisher critical point, and the experiments indicate that the mass dimensions of the operator should become irrational: *we say that the operator gets an anomalous dimension if its mass dimension do not match its value at free theory*, see Table (1.3).

Let's try to understand *anomalous dimensions*. To do that, we need to go back to eqn. (1.26) and try to understand it physically. We have

$$\{x, \Lambda\} \rightarrow \{\ell^{-1}x, \ell\Lambda\} \Rightarrow \mathcal{O}_i(\Lambda, \Lambda x) \rightarrow \ell^{d-[K_i]} \mathcal{O}_i(\Lambda, \Lambda x)\tag{1.33}$$

Let's understand this for $\ell = 10$, $x = 10\text{mm}$, and $\Lambda^{-1} = 1\mu\text{m}$: we are redefining our rulers such that we now call a millimeter what we used to call 10 millimeters, and we are also dropping all the details between $1\mu\text{m}$ and $10\mu\text{m}$ in the original scale. With the new rulers, our coarse-graining cut-off is exactly same $\Lambda^{-1} = 1\mu\text{m}$; so effectively, we zoomed-out from the system (lost some fine-details) but rescaled our rulers such that the system measures same in the new setting. So we should *see* the system same, albeit a little-bit blurred.

^{by}The computation of the stress tensor is most conveniently done by changing the Euclidean/Lorentzian metric to an arbitrary one $g_{\mu\nu}$ and then variate the Lagrangian with respect to it, i.e. $T^{\mu\nu} \equiv \frac{2}{\sqrt{g}} \frac{\delta(\sqrt{g}\mathcal{L})}{\delta g_{\mu\nu}}$ for $g = |\det(g_{\mu\nu})|$.

^{bz}This is how we normalized our field ϕ in the first place, which is rather conventional.

^{ca}Add references for scaling relations.

Table 1.3: The behavior of the Ising Universality Class as we change the dimension of the space from 4 to 3. As we described in section 1.2, the Gaussian fixed point describes the Ising Universality Class for $d \geq 4$, and Wilson Fisher fixed point emerges as we continuously decrease d , which takes over the description of the Ising critical point. As this fixed point is strongly coupled (unlike Gaussian), the relevant operators σ and ϵ get anomalous dimensions γ_σ and γ_ϵ , and can no longer be related as ϕ^2 and ϕ were, i.e. $[\epsilon] \neq 2[\sigma]$.

Spacetime dimension d	$4 \rightarrow 3$
Fixed point	Gaussian \rightarrow Wilson-Fisher
\mathbb{Z}_2 -odd operator \mathcal{O}_1	$\phi(x) \rightarrow \sigma(x)$
\mathbb{Z}_2 -even operator \mathcal{O}_2	$\phi(x)^2 \rightarrow \epsilon(x)$
Mass dimension of \mathcal{O}_1 , i.e. $[\mathcal{O}_1]$	$\frac{d-2}{2} \rightarrow \frac{d-2}{2} + \gamma_\sigma$
Mass dimension of \mathcal{O}_2 , i.e. $[\mathcal{O}_2]$	$d-2 \rightarrow d-2 + \gamma_\epsilon$

There is no guarantee that the operators would look same under such a blurring change; in fact, we do not expect them to look so! However, at the critical point, we have *scale-invariance*, i.e. the system look *exactly same* if you zoom-out, i.e. we do not get blurring with the above prescription. And as we linearized the action of RG *around a fixed point*, it makes sense that the operator $\mathcal{O}(\Lambda, \Lambda x)$ remain same upto an overall factor in this prescription, at least in the leading order.

How about the fixed point itself, i.e. what happens exactly at the fixed point? Well, we stated that we reach the fix point only after *an infinite number of RG iterations*, so it should not depend on Λ . Another way to see this is that Λ describes the scale at which we look at the system, and the system is *scale-invariant* at the critical point, so it does not depend on Λ : the operator $\mathcal{O}(a, b)$ should become $\mathcal{O}(b/a)$ at the critical point for $\mathcal{O}(\Lambda, \Lambda x)$. But this means, we have

$$\mathcal{O}(\lambda x) = \frac{1}{\lambda^{\Delta_{\mathcal{O}}}} \mathcal{O}(x) \quad (1.34)$$

via eqn. (1.26). Here, Δ is called *the scaling dimension* of the field.

As is customary for scale-invariant and conformal invariant theories, we will use the Greek letter Δ to denote the scaling dimension. The free theory (which is itself scale-invariant) has the so-called *engineering scaling dimensions*, these were what we presented in eqn. (1.31). For a scalar field, as we saw there, we have

$$\Delta_{\text{free scalar}} = \frac{d-2}{2} \quad (1.35a)$$

which follows from setting the mass dimension of its kinematic term to d , i.e. $[(\partial\phi)^2] = d$. In contrast, the kinematic term for a *fermionic field* and gauge fields are $\bar{\psi}\not{\partial}\psi$ and $(\partial_{[\mu} A_{\nu]})^2$ hence^{cd}

$$\Delta_{\text{free fermion}} = \frac{d-1}{2}, \quad \Delta_{\text{gauge vector}} = \frac{d-2}{2} \quad (1.35b)$$

We can also define the antisymmetric tensor $F \equiv \partial_{[\mu} A_{\nu]}$, for which we have

$$\Delta_{\text{antisymmetric tensor}} = \frac{d}{2} \quad (1.35c)$$

These *engineering scaling dimensions* are valid at Gaussian fixed point, but they get modified when we move to another fixed point, especially if the theory at that fixed point is strongly coupled! Then we define *anomalous dimension* (denoted by γ) as

$$\Delta_{\mathcal{O}} = \Delta_{\mathcal{O}_{\text{free}}} + \gamma_{\mathcal{O}} \quad (1.36)$$

For instance, the operators σ and ϵ of 3d Ising model have the values^{cc}

$$\Delta_{\sigma} \sim \frac{1}{2} + 0.018, \quad \Delta_{\epsilon} = 1 + 0.413 \quad (1.37)$$

where the former two values are engineering scaling dimensions^{cd} and the latter two are the anomalous dimensions. As a corollary of the

^{cd}In these notes we follow the conventions $A_{[x}B_{y]} = \frac{1}{2}A_xB_y - \frac{1}{2}A_yB_x$ and $A_{\{x}B_{y\}} = \frac{1}{2}A_xB_y + \frac{1}{2}A_yB_x$.

^{cc}See [17] for the computations of these values by conformal bootstrap.

^{cd}As $\sigma \sim \phi$ and $\epsilon \sim \phi^2$ around the free theory, their engineering scaling dimensions are $[\sigma] = \frac{d-2}{2}$ and $[\epsilon] = d-2$.

above discussion, we can immediately say that *the fact that the anomalous dimension of the operators σ and ϵ are large indicates that the 3d Ising Model is strongly coupled and hence we cannot hope to describe it by traditional methods such as perturbation theory.*

How come do we get anomalous dimensions γ ? I am not asking the dynamics how γ emerges, but rather how come it is consistent that we can have irrational values for scaling dimensions of operators \mathcal{O} , yet we can have dimensionless action, i.e. $[\mathcal{L}] = d$? Indeed, if we consider the following relation at the critical point^{ce}

$$\begin{aligned} \int d^d x \left(\frac{\partial \mathcal{O}(x)}{\partial x} \right)^2 &= \int d^d(\lambda y) \left(\frac{\partial \mathcal{O}(\lambda y)}{\partial (\lambda y)} \right)^2 = \int d^d \lambda^{d-2} \left(\frac{\partial \mathcal{O}(\lambda y)}{\partial y} \right)^2 \\ &= \int d^d y \lambda^{d-2-2\Delta} \left(\frac{\partial \mathcal{O}(y)}{\partial y} \right)^2 \end{aligned} \quad (1.38)$$

which forces $\Delta = [\mathcal{O}]$, i.e. that the scaling dimension is equal to the mass dimension, which is always rational. What went wrong?

Scaling dimension should not be equal to mass dimension. To see that, consider the two-point correlation function $\langle \phi(x)\phi(0) \rangle$. By dimensional analysis, it has the mass dimension $[\langle \phi(x)\phi(0) \rangle] = d - 2$; at the critical point, we then expect the scaling behavior

$$\langle \phi(x)\phi(0) \rangle \sim \frac{1}{|x|^{d-2}} \quad (1.39)$$

But the definition of the critical exponent η follows from $\langle \phi(x)\phi(0) \rangle \sim |x|^{-d+2-\eta}$ hence this naively indicates that we always have $\eta = 0$, which clearly contradicts the empirical evidence!

The resolution to this problem is related to a common misconception. It is often stated that *there is no length scale at the fixed point of RG flow, hence there is no length scale at the critical point as well*. This statement is incorrect! The correct statement is this: *the physics near and at the critical point is governed by the long-range physics and is independent of the microscopic details; nevertheless, we still need the microscopic length scale for consistency*. Therefore, the operator $\mathcal{O}(x)$ actually has a dependence on the microscopic scale a : $\mathcal{O} = \mathcal{O}_a(x)$. If we scale all our rulers by λ , so that we call λ meters what we used to call 1 meter, then both a and x get scaled and we obtain $\mathcal{O}_{\lambda a}(\lambda x) = \lambda^{-[\mathcal{O}]} \mathcal{O}_a(x)$.^{cf} On the contrary, if we simply scale x as a parameter, then we get $\mathcal{O}_a(\lambda x) = \lambda^{-\Delta_\mathcal{O}} \mathcal{O}_a(x)$. Thus, we can consistently have^{cg}

$$\langle \phi_a(x)\phi_a(0) \rangle = \frac{a^\eta}{x^{2[\phi]+\eta}}, \quad \langle \phi(x)\phi(0) \rangle = \frac{1}{x^{2\Delta_\phi}} \quad (1.40)$$

for

$$\phi_a(x) = a^{\eta/2} \phi(x), \quad \eta = 2(\Delta_\phi - [\phi]) \quad (1.41)$$

With $\Delta_\mathcal{O} = \Delta_{\mathcal{O}_{\text{free}}} + \gamma_\mathcal{O}$ and $[\mathcal{O}] = \Delta_{\mathcal{O}_{\text{free}}}$, we see that η is simply twice the anomalous dimension γ .

What is this microscopic length scale a ? We really do not need to know! Indeed, a does not show up anywhere in our equations and

^{ce}It is important that we are precisely at the critical point so that neither the integral limits nor the operator $\mathcal{O}(x)$ does have any regulator Λ dependence.

^{cf}In eqn. (1.38), this means that we actually have the integral $\int_a^d d^d x \left(\frac{\partial \phi_a(x)}{\partial x} \right)^2$, and changing the integration variable $x \rightarrow \lambda x$ turns the integration limit a to λa .

^{cg}See eqn. (1.8).

affects nothing about the computations: this is why it is stated that the physics at critical point is independent of the microscopic details. We detailed this subtlety just to make sure that the concept of anomalous dimension makes sense to the reader: indeed, if we did not have a -dependence at the critical point, we could not have any anomalous dimension. Nevertheless, as a is only there to make sure dimension counting checks out, we'll forget about it in the rest of these notes: we will stick to $\phi(x)$ with its scaling dimension Δ_ϕ .^{ch}

This concludes our discussion of dimensional counting and scale-invariance. We'll summarize these findings in the next section.

Summary

- Renormalization group flow formalism allows us to look at a system from different scales, and compute how its behavior change as we go from UV scales to IR scales.
- With the RG, we observe that the behavior around fixed points are solely determined by the relevant directions, hence we expect a *universal scaling* around the critical point!^{ci}
- Fixed points of RG have scale-invariance, hence operators at that point satisfy the equation

$$\mathcal{O}(\lambda x) = \lambda^{-\Delta_{\mathcal{O}}} \mathcal{O}(x) \quad (1.42)$$

where $\Delta_{\mathcal{O}}$ is called *scaling dimension* of the operator \mathcal{O} .

- Scaling dimensions of *relevant operators*, whose scaling dimension is less than d , determine the critical exponents if that fixed point is associated with a critical phenomena. In Ising model, we have two relevant operators σ and ϵ and we can write critical exponents as $\alpha = \frac{d-2\Delta_\epsilon}{d-\Delta_\epsilon}$, $\beta = \frac{\Delta_\sigma}{d-\Delta_\epsilon}$, etc.
- We can compute scaling dimensions of the operators if we have a free theory, i.e. the scalar operator ϕ has the scaling dimension $\Delta_{\text{free scalar}} = \frac{d-2}{2}$.
- If the theory is strongly coupled, then scaling dimension of operators diverge from their values at free theory; the difference is called *anomalous dimension*: $\Delta_{\mathcal{O}} = \Delta_{\mathcal{O}_{\text{free}}} + \gamma_{\mathcal{O}}$.
- There is no analytic and exact way to solve strongly-coupled theories. We usually do several approximations or expansions: *discretization on a lattice*, *large N expansion*, ϵ -expansion, etc.^{cj} These methods may or may not yield reliable and robust results.
- The alternative (and relatively recently developed) approach is to use *conformal bootstrap program*. It uses rigorous self-consistency and symmetry conditions to solve the landscape of CFT's for a given universality class. We'll detail this program in the following chapters.

^{ch} It is worth stressing the importance of the remarkable conclusion of this discussion. The critical phenomena is observed at lengths of ξ , say of the order of microns or larger, and the microscopic length scale a is of the order of an atom. So naïvely, the ratio $\frac{a}{\xi}$ is sufficiently small that we can replace it by 0. However, if we do this, a does not appear in any final formula hence we cannot get any anomalous dimension: *so the critical phenomena is an area of physics where we cannot get rid of an extremely small parameter ($\frac{a}{\xi}$) if we would like to describe the phenomena correctly!* This is rather remarkable, and shows how incorrect the often-heard statement “*the only important length scale near the critical point is the correlation length*” really is!

There is no contradiction between *physics being independent of microscopic details* and *we cannot get rid of the microscopic length scale a* . Mathematically speaking, if we have a function $f(a, \xi)$ and we would like to take $a \rightarrow 0$, we need the function to be regular around $a \sim 0$. If it has the form $\lim_{a \rightarrow 0} f(a, \xi) \sim \lim_{a \rightarrow 0} a^{-\sigma} g(\xi)$, then we cannot take $a = 0$ in $f(a, \xi)$. The situation is simply this: we only need a as an overall factor to fix dimensions; it does not interfere with the rest of the expression!

The first explicit recognition of such a case (that a function has the behavior $\lim_{a \rightarrow 0} f(a, \xi) \sim \lim_{a \rightarrow 0} a^{-\sigma} g(\xi)$ so that one cannot discard a) is probably the calculation of how a converging shock wave is focused (see the footnote in page 197 of [4]). In [22], there is an extensive study of such problems and their solutions, accomplished by making the explicit hypothesis that anomalous dimensions exist combined with numerical methods. *Historically, however, the problem of anomalous dimensions in critical phenomena was discovered and solved apparently without knowledge or recognition of these other phenomena: this is rather interesting because Landau had the first-hand knowledge of anomalous dimensions in both the critical phenomena and the problem of the converging shock wave but apparently he made no reference to a connection between them!*

Finally, we would like to note that, rather ironically in a sense, we do not need the microscopic length scale a if we are away from the critical point: as the correlation length ξ is finite there, it does the job to fix the dimensions, see eqn. (1.8a) for an example!

^{ci}For a qualitative understanding, let us consider a system close to criticality. The trajectories of the systems on the critical manifold remain on the critical manifold and flow

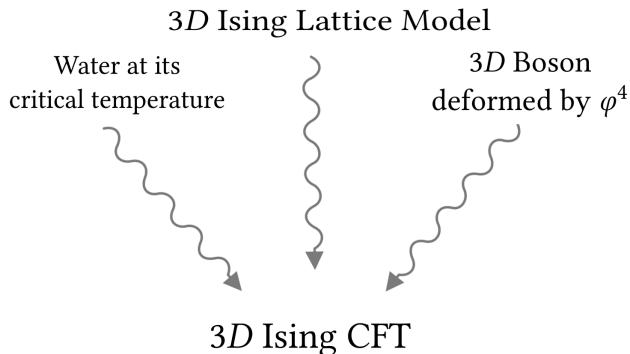


Figure 1.9: Many different phenomena at microscopic scales are described by the same conformal field theory at macroscopic scales; i.e. they fall in the same universality class! This is called *IR equivalence*.

- Why use conformal symmetry for critical phenomena? This follows from IR equivalence: many physically-relevant critical phenomena are in the same universality class with a known CFT; and, solving the CFT is sufficient to understand, say, the divergence of the heat capacity of water near its critical point —see Fig. (1.9) for illustration of this for Ising Universality Class.
- How do we know that we have *conformal symmetry* at fixed points? Technically, all we know is that we have *scale symmetry* at the fixed points and *not all scale invariant theories are conformally invariant!* There are strong motivations to anticipate conformal invariance (we'll discuss these), but for now, we can take a simpler approach: we *will assume* that the critical point is described by a conformal theory, compute the critical data, and check with the experiments to see if they agree!

1.3 Using CFTs beyond statistical physics

In the previous sections, we have used the problems in statistical mechanics to motivate the usage of *conformal field theories*; indeed, as we have seen there, it is essential to understand conformal field theories to describe and compute critical data for a huge collection of phenomena.^{ck} What made the study of conformal field theories all the more powerful was the concept of universality: we can divide critical phenomena into a handful of different universality classes and understanding one CFT is sufficient to understand all the problems in the same universality class. The most prominent example of this is the 3d Ising universality class; solving this CFT is sufficient to understand

1. Liquid-vapor transitions, e.g. water at its critical point
2. Uniaxial magnetic systems, e.g. an idealized Nickel magnet
3. Binary mixtures, e.g. polymer solutions
4. Coulombic systems, e.g. solvophobic electrolytes

to the fixed point. On the other hand, if the system is slightly off the critical manifold, it will approach to the critical point during part of the RG, but will ultimately repel from the critical manifold and flow away because of the unstable directions due to relevant operators.

Independent of the initial position off the critical manifold, the system will first flow toward the critical point because the only singularities of the flow field are the fixed points themselves! Likewise, independent of where they are off the critical manifold, they will be eventually repelled with the same eigenvalues because it is the same unstable directions that cause the behavior! *The fact that it is the same eigenvalues which drive all slightly off-critical systems away from the fixed point is the origin of universality!* [4]

^{cj}See the bullets listed in page 16 and the next one for further details. [Maybe put some resources here](#)

^{ck}To be fair, what is essential is to understand *scale invariant theories*, which as we have already stated are not necessarily conformally invariant. Nevertheless, this is almost always the case and we'll motivate this later.

5. Micellar systems, e.g. the process of aggregation of certain surfactant molecules in dilute aqueous solutions

Please see the section 3.1.1 of [23] for details on these cases alongside several resources.

Of course, there are several other universality classes and they can be used to explain yet other whole sets of phenomena in statistical mechanics, solid state physics, or chemistry; for instance

- The three-dimensional XY universality class (describes the superfluid transition of ${}^4\text{He}$ along the λ -line^{c_l} among others)
- The three-dimensional Heisenberg universality class (describes Curie transition in isotropic ferromagnets such as Ni and EuO, among others)

where further details can be found in the latter reference, i.e. > 100 page review [23].

For a practicality-minded reader, these should be more than enough as reasons to study conformal field theories. However, young researchers (especially in high energy physics) are often amazed by speculative theories or effective models in high energies, rather than low-energy phenomenological models with applications in *trivial* branches, such as chemistry.^{c_m} Therefore, I'd like to mention some other areas where conformal field theories are relevant.^{c_n}

- String theory, M-theory, or similar mathematically advanced models in general
- Flat space holography
- Quantum field theory
- AdS/CFT
- Cosmology

The first two items are really related to $2d$ CFTs: we know that string theory is a two dimensional conformal field theory on its worldsheet, and a scattering amplitude in four dimensional lorentzian space-time can be related to a $2d$ Euclidean conformal field theory which lives at the null infinity of the Lorentzian space.^{c_o} Our focus in these notes will be on higher dimensional $d \geq 3$ CFTs, so we can ignore these motivations.

I wrote *quantum field theory* as a motivation but it is rather ambiguous so let me unpack it. As we reviewed in section 1.2, any field theory can be subjected to RG flow, and its phase diagram shows the behavior of the model in general. Even if we are not interested in critical phenomena, the fixed points are still relevant because they are like landmarks on the space of all field theories; in fact, most theories flow from a CFT at UV to another CFT at IR, see Fig. (1.10).^{c_p}

^{c_l}See Fig. (1.7).

^{c_m} The tone in this sentence may be uncalled for, but as a young (?) researcher in high energy physics, I'd like to make this (self-)criticism: unfortunately, we can be too focused on high energy theories and a one-theory-to-rule-them-all model that would explain whole physics and that can be best understood only in high energies. Thus, we tend to feel that things at lower energies (and branches that explain them) are secondary to what we are studying; after all, we can (?) derive them from the high energy models. This is why (for someone who thinks that way) chemistry would be labeled trivial. Of course, I'm making these comments based on my own observations, which actually form a small subset of high energy physicists, making my generalization unwarranted.

For a second, let's assume that my comment has some truth in it, i.e. some young high energy physicists do actually overemphasize the roll of high energy physics and do see some of the other branches rather trivial. Within this assumption, it is then a nonzero chance that you, the reader, may feel the same way. So I'd like to address this and hope that it is beneficial to you in case you were unaware of it.

The point is that *high energy physics can be used to explain all low energy phenomena* is an assumption, and is not really backed-up. It is found on *constructionism*, that one can construct macroscopic behavior from the knowledge of microscopic model, but constructionism is actually not valid: there are *emergent phenomena*, concepts that only emerge at long distances and cannot be explained by short distance physics. Anderson famously discusses this in his seminal paper *More Is Different* in 1972, and points to the difference between *reductionism* and *constructionism*: we can indeed reduce low energy rules to some fundamental laws in high energy physics, but we cannot always start from the fundamental laws in high energy physics and construct low energy phenomena. I strongly urge the reader to check that article if they haven't already: see [24].

^{c_n}The list is by no means complete.

^{c_o} This is a relatively new field and the correspondence is not really that well-established, see [25] which roughly initiated the current research on this area. For an earnest attempt at an introductory review, see section 2 of my paper [26].

More importantly, as we do not know how to solve strongly coupled field theories in general, we may try to expand them around CFTs;^{cr} likewise, we can explain more than critical phenomena if we are close to a CFT. For instance, there is this concept called *walking*, which is relevant if the theory is not scale invariant but near a fixed point,^{cs} and walking behavior is related to many things including the electroweak phenomenology beyond the Standard Model in high energy physics and the so-called weakly first-order phase transitions in statistical mechanics.^{ct} Finally, a quantum theoretical model itself can be conformally invariant under various assumptions; for instance, SU(N) gauge theories with N_f flavors of fundamental fermions can flow at IR to an interacting CFT for particular values of N_f , i.e. *such gauge theories are expected to have a conformal window*, see Fig. (1.11).^{cu} To understand such concepts, we need to understand conformal field theories in higher dimensions.

Another motivation to study conformal field theories is the infamous conjecture that there is a duality between scattering amplitudes in curved spacetime and the conformal field theory at the boundary of that space.^{cw} As we do not have the UV completion of quantum gravity,^{cy} AdS/CFT correspondence is an attractive toy model to work with

^{cr}I need to put some source here

^{cs}The coupling constants are called *running* as they change with scale (such as fine structure constant of Quantum Electrodynamics), and are constant if the theory is scale invariant. If the theory is almost scale invariant, the coupling constants barely change hence are called *walking*!

^{ct}See [27, 28] as rather introductory reviews.

^{cu}For such gauge theories, Caswell [29] and Banks & Zaks [30] showed that there is a fixed point at the RG diagram (much like Wilson-Fisher fixed point in Fig. (1.8), but for gauge fields instead of scalars) and that an interacting CFT lives there. This fixed point is called Caswell-Banks-Zaks fixed point, and is relevant for asymptotically free gauge theories; in particular, for the quantum field theoretical description of strong force (i.e. QCD₄) in its conformal window.

So what is the conformal window? On one hand, N_f is bounded from above by the requirement that the theory should have asymptotic freedom, hence $N_f \leq 11N_c/2$. On the other hand, N_f should be higher than some critical N_c below which we get *chiral breaking* which dynamically generates a length scale. The value 11/2 above follows from perturbative computation of RG flow, and is sufficiently reliable. For N_c however, we do not know any rigorous method^{cv}, and there seems to be a disagreement to what that value should be; for instance, see Table 1 of [31] for disagreeing results for $N_f = 12$ QCD₄, and Table 5 of [32] for a similar situation in QED₃. For a conformal bootstrap approach to 4d CFTs at IR to which 4d gauge theories may flow, see [33].

One may think that such considerations may be of theoretical importance but of little practical value; after all, we know that QCD₄ has three colors and six flavors, so $N = 3$ and $N_f = 6$; thus we do not need to consider anything else. This naïve perspective is however incorrect: as we have stressed earlier (see footnote az for instance), all our theories are actually effective field theories and the effective number of flavors depend on which energy scales we are doing our experiments/computations (we can *integrate-out* heavier generations). In addition, that we have found six flavors so far does not guarantee that we will not find more at higher energies, so it is advisable to fully analyze the phase diagram of QCD for arbitrary values of flavor number (maybe I should put some references here for reasons for and against number of generations greater than 3). Furthermore, modifications of the standard model (such as with supersymmetry) can bring further flavors, increasing the importance of considerations with higher flavor numbers (such as $N_f = 12, 16$, see [34].)

detail later, the way conformal bootstrap tries to find N_c is simple: it starts with a value of N , and looks for existence of conformal field theories. If $N < N_c$, we can not find any self-consistent CFT so bootstrap method should disallow all possible CFT candidates. By varying N , we can determine N_c . [put some references here](#)

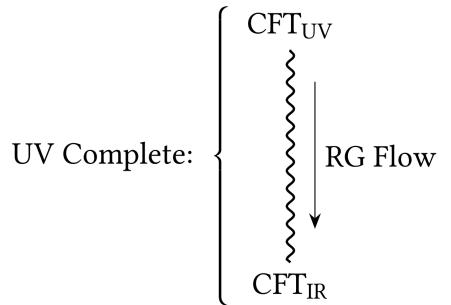


Figure 1.10: End points of renormalization flow and conformal field theories. All field theories flow from a fixed point at UV to another at IR, and fixed points are either described by a topological theory (i.e. $\xi = 0$), a scale-invariant but not conformally invariant theory ($\xi = \infty$), or a conformally invariant field theory ($\xi = \infty$). Optimistically, we can assume that many cases of interest have CFTs at the end points of the RG flow.

^{cp}This statement is probably an exaggeration. First of all, the end-points of *any* QFT has to be scale invariant, but some of them are simply trivial (i.e. the correlation length $\xi = 0$) and are called *topological* as there is no propagating degree of freedom left in them.^{cq}The rest have massless degrees of freedom, i.e. $\xi = \infty$, but only some of them are also conformally-invariant (there are examples of fixed points with scale invariance but not conformal invariance), and only a subset of them are interacting (hence interesting). As I do not know what are the respective ratios of these cases, I actually only made an optimistic assumption by saying *most theories* in the text.

^{cq}We should be very careful when we use the words trivial and topological at the same sentence! There can be many interesting non-trivial topologies, and that the theory is topological at the fixed point does not mean it doesn't have anything interesting in it. We

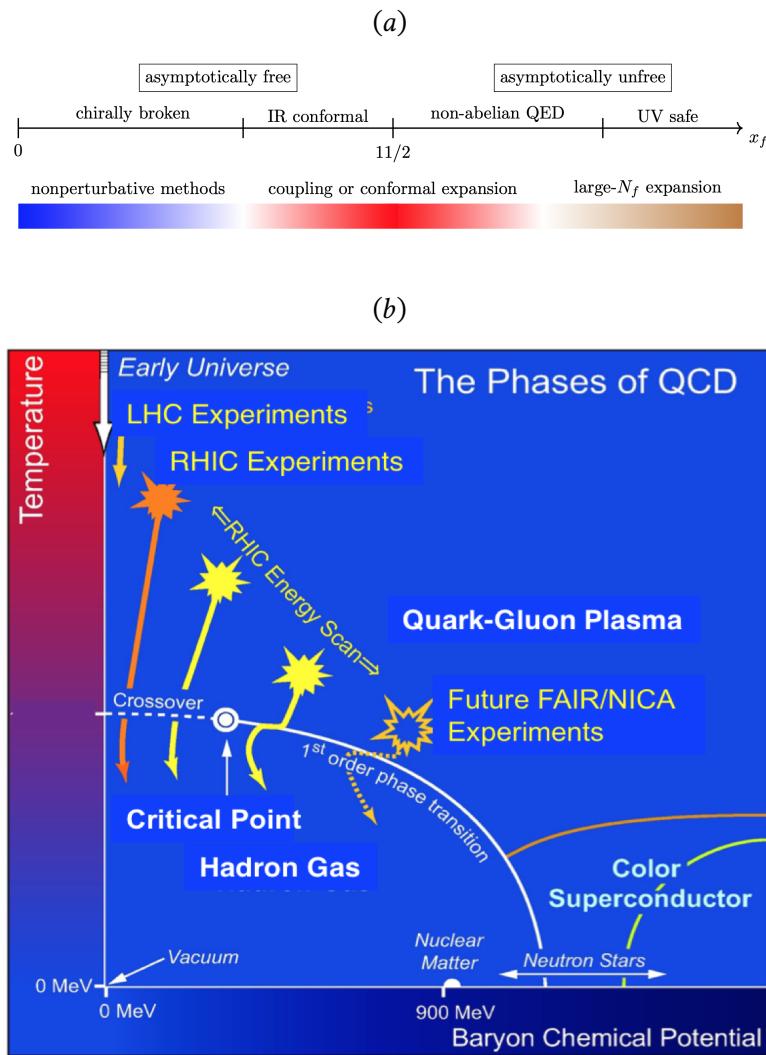


Figure 1.11: (a) Conjectured IR phase structure for $SU(N)$ gauge theories with N_f flavors of fundamental fermions as $N, N_f \rightarrow \infty$ for fixed ratio $x_f = N_f/N$, at zero temperature and zero chemical potential. Figure taken from [35]. (b) A schematic phase diagram for QCD_4 where the regions probed by different accelerator facilities are also indicated. We expect a CFT to describe the critical point and the critical manifold (the region that flows to critical point) at IR. Figure taken from [36].

the quantization of gravity in a controlled environment; indeed, AdS is simple enough to do actual computations of correlation functions of gravitons.

Lastly, we can mention cosmology as a motivation to study conformal field theories. By cosmological principle^{da}, we expect the universe to be governed by FLRW dynamics.^{db} Hence, it has the metric

$$ds^2 = -dt^2 + a(t)\eta_{ij}dx^i dx^j \quad (1.43)$$

just mean that we are not interested at those in these notes.

^{cv}To my knowledge, *numerical conformal bootstrap program* is the only non-perturbative and sufficiently rigorous method for limited computational power (after all, we can always do lattice computations as well if we have large computational powers), if we are not in $2d$ and if we are not assuming other simplifications (such as supersymmetry). As we will

^{cw}More specifically, there is a correspondence between correlators in Anti de Sitter (AdS)^{cx} in $d+1$ dimensions and the conformally invariant correlators at the boundary of that AdS [37–39]. This conjecture is often perceived as an outcome of string theory, but we can actually argue that we can arrive at this correspondence without using the string theory (at least intuitively). Penedones does a good job at pedagogically introducing this concept in [40]; if you would like a more complete (yet also denser) review, see [41].

^{cx}If you have not heard of de Sitter (dS) space before, it is basically a Lorentzian version of the standard sphere in d dimensions (which is Euclidean); so in a sense, it is the most symmetric space in the Lorentzian signature. AdS is almost same as dS, their only difference is the sign of the *curvature* of the space. In short, flat Lorentzian space, AdS and dS are three most symmetric spaces with curvature = 0, < 0, and > 0 respectively; see Table (1.4) for a comparison of these spaces.

^{cy}It is often stated that *we do not have a theory for quantum gravity*. This is misleading; at low energies, we know how to quantize gravity (and obtain the quanta graviton) and compute scattering amplitudes of these objects among themselves or with standard model particles. The problem is that this theory is nonnormalizable, hence predicts its own regime of validity (roughly, upto Planck scales). Thus, we *have* quantum gravity but we do not know its UV completion. If we remember from footnote ap, the situation is similar to Fermi's model of weak interactions: he had a theory for weak interactions, but it was not *UV-complete*, i.e. it had to break after going higher than some energy (some energy ~ 300 GeV in his case).

^{cz}As an abbreviation, we write down signatures as (p, q) . This basically means that the

Table 1.4: Comparison of spaces with a maximal amount of symmetries. The former two rows correspond to Euclidean spaces whereas the latter two do to Lorentzian spaces. It is known that not all curved spaces can be embedded in flat space; however, these spaces can actually be done so and the signature^{c_z} of the embedding space is given in the last column. See [41] for a brief discussion of these points.

Name of the space	Signature	Curvature	Signature of the embedding space
Sphere	(d, 0)	Positive	(d + 1, 0)
Lobachevski	(d, 0)	Negative	(d, 1)
de Sitter	(d - 1, 1)	Positive	(d, 1)
Anti de Sitter	(d - 1, 1)	Negative	(d - 1, 2)

where $a(t)$ is called the expansion parameter. Note that η_{ij} here stands for the spatial metric which we are taking to be flat.^{d_c} We define the conformal time τ as

$$d\tau = \frac{dt}{a(t)} \quad (1.44)$$

hence the metric can be written as

$$ds^2 = a(\tau)^2 (-d\tau^2 + \eta_{ij}dx^i dx^j) \quad (1.45)$$

We define the Hubble parameter $H(t)$ as

$$H(t) = \frac{1}{a(t)} \frac{da(t)}{dt} \quad (1.46)$$

Let us now assume an inflationary cosmological model, which means that there was a time back in the early universe when there was a sustained period of accelerated expansion, i.e. $\frac{d^2 a(t)}{dt^2} > 0$. Let's assume that Hubble parameter had changed *very slowly* and hence can be taken as a constant. This then implies

$$a(t) = e^{H(t-t_0)} \quad (1.47)$$

which means $d\tau = e^{-H(t-t_0)} dt$, indicating that

$$\tau = -\frac{1}{H} \ln(a(\tau)) \quad (1.48)$$

If we now insert this back to eqn. (1.45), we obtain

$$ds^2 = \frac{-d\tau^2 + \eta_{ij}dx^i dx^j}{H^2 \tau^2} \quad (1.49)$$

But this is precisely the metric for de Sitter space in the so-called Poincaré frame!^{d_d} Thus, *if universe is homogeneous, isotropic, spatially almost flat, and had expanded during an inflationary period with an almost constant Hubble constant, then the spacetime of the universe during that period is approximately given as de Sitter space!*^{d_e}

metric of that space is given as $ds^2 = \sum_{i=1}^p dx_i^2 - \sum_{j=p+1}^{p+q} dx_j^2$.

^{d_a}The principle that states the universe is isotropic and homogeneous at sufficiently large scales. This is based on the standard model of cosmology, i.e. *concordance model*, and it is calculated in [42] based on this model that the inhomogeneities in the universe are below $< 1\%$ for distances of $\gtrsim 300$ megaparsec. The principle is widely accepted at the very least as a leading order approximation to the universe; however, this assumption is not necessarily warranted in short distances (see for instance [43] which claims that the principle is used incorrectly at too short distances in the analysis of supernovae data). In fact, even at large distances, there are certain examples that violate the principle, i.e. very large quasar groups that create inhomogeneities at cosmological scales (for instance *Clowes–Campusano large quasar group* which is longer than 600 megaparsec, see https://en.wikipedia.org/wiki/Large_quasar_group). In summary, the principle is now accepted in general but may be ruled out in future, see the video *New Evidence against the Standard Model of Cosmology* of Sabine Hossenfelder for a beautiful discussion of these points.

^{d_b}FLRW stands for the initials of the scientists Friedmann, Lemaître, Robertson, and Walker. Basically, this is the most symmetric universe which expands over time, an approximate model of our own universe.

^{d_c}In general, η_{ij} could be the metric of a maximally symmetric $d - 1$ space.

^{d_d}[add references here](#)

^{d_e}See [44] for a detailed review of cosmology and these arguments.

Well, how does this relate to conformal symmetries? As we review in Table (1.4), dS and AdS are almost same, and the boundary correlators of dS can also be explained by a CFT living there.^{df} You can ask, then, why do we hear too much about AdS/CFT but not that much about dS/CFT? I'm no expert in this field but I guess we can name two sound reasons for that:

- Just because the symmetries can be there does not mean they are not broken! For instance, conformal invariance is almost always broken in inflationary models^{dg} and only re-merges in slow-roll inflation^{dh} [45], and even then, it is exact only in the limit where dynamic gravity decouples (so-called *decoupling limit*). The interactions of slow-roll inflation beyond the decoupling limit are then controlled by the weakly broken conformal symmetry [46–48].
- Even if we had the symmetries intact, the relation between boundary CFT operators with the bulk fields^{di} in AdS and dS are somewhat different. To understand this, let's rewrite eqn. (1.49) again, but this time both for dS and AdS:

$$ds^2 = \frac{-d\tau^2 + \eta_{ij}^{(dS)} dx_i dx_j}{\tau^2} \text{ (dS)}, \quad ds^2 = \frac{dz^2 + \eta_{ij}^{(AdS)} dx_i dx_j}{z^2} \text{ (AdS)} \quad (1.50)$$

where $\eta^{(dS)}$ and $\eta^{(AdS)}$ are the boundary metrics respectively. We take $\eta^{(dS)}$ to be Euclidean for physical reasons (it is the metric for the spatial part of the universe), but we can take $\eta^{(AdS)}$ to be Lorentzian, meaning that a Lorentzian CFT lives at the boundary.

Now, we can see an intuitive difference between these two: In AdS, the CFT operators act as *boundary conditions* on bulk fields; they determine bulk fields' values as $z \rightarrow 0$.^{dj} On the other hand, the CFT operators act as *final value conditions* on dS bulk fields; they determine bulk fields' value as $\tau \rightarrow 0$. We say *final value condition* because, $-\infty < \tau < 0$ and the CFT basically lives at the infinite future of the bulk fields. Thus it somehow makes sense to use CFT knowledge to fix the behavior of AdS fields to determine their time evaluation, but in dS, knowledge of CFT is already equivalent to the future knowledge of bulk fields, so it does not make a similar sense to use them.

Despite these reasons, CFT knowledge is essential in cosmology and the modern approach is to use it as much as possible to constraint ambiguities in the dS correlators. This approach is similar to *conformal bootstrap program*,^{dk} and is in fact named *cosmological bootstrap program*, see [49–51].^{dl}

^{df}Unlike AdS, the CFT at the boundary of dS has to be a Euclidean CFT. One way to see this is to check the symmetries: the symmetry algebra of a conformal theory at a spacetime of signature (p, q) is given as $so(p+1, q+1)$, and this should be the isometry of the space, which is roughly the last column in Table (1.4).

^{dg}It is interesting to note that (contrary to conformal invariance) scale invariance remains an approximate symmetry in all inflationary models.

^{dh}Inflation models where the *potential energy* of the inflation dominates the *kinetic energy*!

^{di}The combination “bulk x ” refers to the quantity x within the curved spacetime, whereas “boundary x ” refers to the quantity x at the boundary of the curved spacetime.

^{dj}We get to the boundary as $z \rightarrow 0$.

^{dk}As we stated earlier, this program aims to classify and constraint conformal symmetries by non-perturbative tools such as analyticity or unitarity.

^{dl}I have been citing a lot of Daniel Baumann's papers in the last few pages. This is not just necessarily because of his papers' excellence (although they are really good): I really do not command the literature that well and am biased towards Baumann's work as I am working in his group.

2

What is conformal symmetry?

2.1 Invitation: how Galileon symmetry helps us derive the form for kinetic energy

In these notes, we will analyze the effect of conformal symmetry in field theory. In the previous section, we have motivated why conformal symmetry is a relevant symmetry in many areas of physics and that it should be considered in field theories for various applications, but we did not touch on *how* a symmetry can help us get any result.

The presence of a symmetry puts many constraints on the mathematical expressions for various quantities, and we call these *kinematic constraints*.^{dm} In this section we will introduce and illustrate the concept of kinematic constraint with the simplest example of Galilean symmetry!

We start by the concept of kinetic energy: we observe that moving objects can exert force on other objects upon collision, i.e. move them as well, and hence they should have some kind of energy due to their motion: their *kinetic energy*. We can reason that this should depend on the *amount of matter* (hence the mass m) and on the *velocity of matter* (hence \vec{v}), thus we say $E_k = f(m, \vec{v})$ for an arbitrary function f .

How do we determine the function f ? Well, we can reason kinetic energy is equal to work done on the object, and use work is the integration of force over distance, i.e. $E_k = -\Delta W = -\int \vec{F} \cdot d\vec{x}$. We can take $E_k = \Delta W$ as definition of work, but how about $\Delta W = \int \vec{F} \cdot d\vec{x}$? Where does this come from? Also, how do we solve this without $\vec{F} \cdot d\vec{x} = m\vec{v} \cdot d\vec{v}$ which follows from $\vec{F} = m\vec{a}$?

Clearly using kinetic energy-work definition really cannot fix the form of kinetic energy unless we take Newton's second law and assume that the work is integration of force over distance. Well, if we are accepting $\Delta W = \int \vec{F} \cdot d\vec{x}$ to derive $E_k = m\vec{v} \cdot \vec{v}$, we could have accepted the latter in the first place, so this is not really a useful derivation.

Let us instead use Galilean symmetry to show that the form of kinetic energy is actually a kinematic constraint that we can derive without any other unwarranted assumptions as above.^{dn} Instead, we will only assume

- Mass is invariant under Galilean transformations
- Temperature is invariant under Galilean transformations

With these assumptions, we *define* kinetic energy as the heat released once the object is brought to rest, thus it changes the temperature at a proportional rate to its kinetic energy, which we can reliably measure with a thermometer. By this definition, we can argue that the kinetic energy is *additive* in mass, which we can also easily verify empirically; thus, $E_k = mf(\vec{v})$.

Let's now consider a thought experiment. We throw two identical balls of clay at each other with identical speeds, say v , in a train

^{dm} To my knowledge, the difference between *kinematic constraints* and *dynamic constraints* is that the first set have to do with the motion without the source of the motion (at least in Newtonian mechanics), whereas the second set varies between different sources of motion. For instance, the relation $F = ma$ would be a kinematic equation whereas the equation $F \propto \frac{q_1 q_2}{r^2}$ is a dynamic one as the former has to do with motion and is valid for any moving object whereas the latter have to do with a particular interaction (electromagnetism) and is valid only for objects charged under this interaction. This definition generalizes beyond Newtonian physics as follows: *kinematic things* (equations, constraints, data, etc.) have to do with the symmetries in the theory (hence independent of its internal details) whereas *dynamic things* are model-dependent. For instance, the conservation of momentum is a kinematic constraint (i.e. it follows from the translation invariance of the theory), but conservation of parity is a dynamic constraint (it depends on which Lagrangian we choose for the model, hence model-dependent).

I'm not aware how conventional the above nomenclature really is; furthermore, the difference between them is rather blurry in the modern perspective: take masslessness of electron for instance, it follows from gauge symmetry so one may claim it is a kinematic constraint; however, one can also argue that gauge symmetry is a property of chosen model, hence masslessness of photon is dependent on model after all. Clearly the distinction between kinematic and dynamic is not really unambiguous in such situations.

In CFT realm, the distinction between what is kinematic and what is dynamic will be somewhat more clear: anything that is fixed by conformal symmetry will be kinematic and what is left will be dynamic; for instance, three point correlators of three identical scalar operators have to have the form $\langle \phi(x_1)\phi(x_2)\phi(x_3) \rangle = \lambda(x_1-x_2)^{-\Delta}(x_2-x_3)^{-\Delta}(x_3-x_1)^{-\Delta}$. This form is a *kinematic constraint* as it is fixed by the symmetry. In contrast, λ represents *dynamic data* as it is model-dependent and can't be fixed by symmetry alone.

In short, kinematic vs dynamic may/may not be a standard distinction so please take it as I have described in these notes (and blame me for any ambiguity). Nevertheless, it is still consistent with the Newtonian terminology as anything that has to do with the motion (orig-

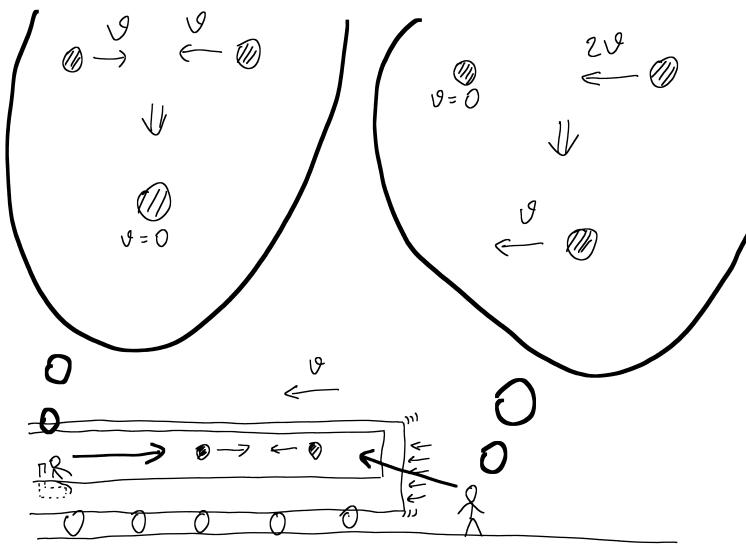


Figure 2.1: Constraint of Galilean symmetry on kinetic energy. Two identical clays of mass m and opposite velocities of speed v collide to full stop in a train. Let $E = mf(\vec{v})$ be the formula for kinetic energy; the observer in train then observes $E_{\text{initial total}} = 2mf(\vec{v})$ and $E_{\text{final total}} = 0$. In contrast, an outside observer observes $E_{\text{initial total}} = mf(2\vec{v})$ and $E_{\text{final total}} = 2mf(\vec{v})$. As the difference in energy is converted to heat (i.e. change in temperature) and as we assume that the temperature is Galilean invariant, both observers should observe same energy difference, leading to $f(2\vec{v}) = 4f(\vec{v})$, yielding that $f(\vec{v}) \propto \vec{v}^2$.

moving in the same dimension of the movement of clays with speed v (see Fig. (2.1) for illustration.). For an observer in the train, the clays have initial and final kinetic energies as $(mf(\vec{v}), mf(\vec{v}))$ and 0 respectively (note that clays merge at the collision). For an observer outside on the other hand, the initial and final energies read as $(mf(2\vec{v}), 0)$ and $(2mf(\vec{v}))$. As the difference in initial and final energies have turned into heat, and as we assume the temperature is a Galilean invariant, both observers should measure same energy difference:

$$mf(\vec{v}) + mf(\vec{v}) - 0 = mf(2\vec{v}) + 0 - 2mf(\vec{v}) \quad (2.1)$$

which leads to $f(2\vec{v}) = 4f(\vec{v})$, yielding

$$E_k \propto m\vec{v} \cdot \vec{v} \quad (2.2)$$

We can now consider the differential form of this equation, i.e. $dE \propto m\vec{v} \cdot d\vec{v}$ and use the relation $\vec{v} \cdot d\vec{v} = \frac{d\vec{x}}{dt} \cdot (\vec{a} dt) = \vec{a} \cdot d\vec{x}$, leading to $dE \propto m\vec{a} \cdot d\vec{x}$. We can then define the force as the gradient of kinetic energy, i.e. $\vec{F} = \vec{\nabla} E_k$, which gives us second law of Newton's law for free, i.e.

$$\vec{F} \propto m\vec{a} \quad (2.3)$$

inal meaning of kinematic) actually follows from the Galilean symmetry!

^{dn} See <https://physics.stackexchange.com/questions/535/why-does-kinetic-energy-increase-quadratically-not-linearly-with-speed> for a discussion of various derivations, including the one we'll detail below.

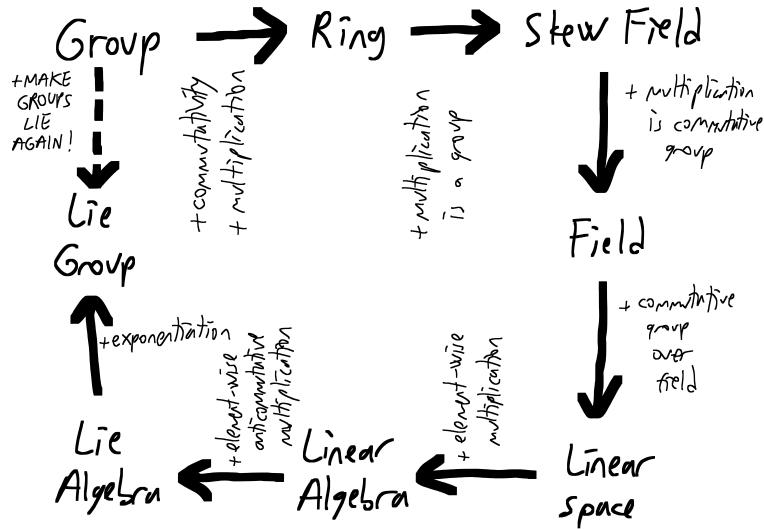


Figure 2.2: Schematic relation between groups, fields, and algebras.
Please see section 2.2 for details.

The proportionality constants depend on our conventions and we can of course always normalize mass m such that the above relation is an equality.

In this simple example, we see that assuming a symmetry brings actually a constraint to our equations: kinetic energy needs to be quadratic in velocity, which is true for any system with Galilean invariance! In the rest of the notes, we'll try to do a similar thing with conformal invariance: it will also bring various kinematic constraints which will in turn simplify the computation of numerous quantities in field theories!

2.2 Galilean symmetry and scale-invariance

As we have seen in the previous section, imposing invariance under a symmetry helps us constrain the physical formulae. However, to fully understand how a symmetry is embedded in mathematical equations, we need to understand the concept of group and algebra, which we will briefly review here.

Primer: groups, fields, and algebras

Before we start, we may want to set the conventions and review the basics of groups, fields, and algebras. This subsection is tangential to the main direction of the notes, and can be skipped if the reader is confident in their knowledge in this area of math. In particular, this subsection can be summarized with the schematic relations shown in Fig. (2.2).

Group: A set of elements $g_i \in G$ such that

1. there is a distinguished element e such that $g * e = e * g = g$ for all g
2. there is a mapping $g \rightarrow \hat{g}$ of G onto G such that $g * \hat{g} = \hat{g} * g = e$
3. operator $*$ is associative: $g_1 * (g_2 * g_3) = (g_1 * g_2) * g_3$

where e is usually called identity and \hat{g} is usually called the inverse element, denoted as g^{-1} , and $a * b$ is denoted as ab . If $*$ is commutative, $a * b = b * a$ is usually denoted as $a + b = b + a$, and \hat{g} denoted as $-g$. In this case e is denoted as 0. In this section, for brevity, we will use the terminology such that any commuting operation will be called *addition* and any not-necessarily-commuting operation will be called *multiplication*, i.e. $a + b = b + a$ and $ab \stackrel{?}{=} ba$.

$|G|$ is the order of the group and is equal to the number of elements of G . One of the most important groups is the symmetric group S_n , which is the set of all permutations of n elements. In fact, in general, the collection of all transformations of a set X (one to one mappings of X onto itself) is a group.

Some common examples of groups include

- \mathbb{C} under addition, $\mathbb{C} \setminus \{0\}$ under multiplication
- The additive group of the ring \mathbb{Z} (the set of integers) is called the infinite cyclic group.
- The $(\text{mod}-n)$ additive group of the ring $\mathbb{Z}_n = \{0, 1, \dots, n-1\}$ is the cyclic group of order n .

Ring (Commutative Group with multiplication): Any commutative group with additional multiplication operation is called a ring; more precisely, the group R is a ring if

1. R is commutative
2. group operation is addition
3. multiplication is defined in R with respect to addition:

$$a(b+c) = ab + ac, (b+c)a = ba + ca \quad (2.4)$$

The adjectives of the ring follows from the properties of the multiplication; for instance, if the multiplication is commutative (associative) the ring is called commutative ring (associative Ring).

The set \mathbb{Z} (and the set $n\mathbb{Z}$, integers divisible by n) are rings under arithmetical addition and multiplication. The set $\mathbb{Z}_n = \{0, 1, \dots, n-1\}$ is a ring under addition and multiplication module n . If n is a prime number, it is also a field!

Skew field (Ring whose multiplication is a group): A ring κ is a skew field if the set $\kappa \setminus \{0\}$ is a group under multiplication! They are also called division ring as division can be defined in these rings! For instance, quaternions form a noncommutative skew field.

Field (commutative skew field): A ring κ is a field if the set $\kappa \setminus \{0\}$ is a commutative group under multiplication! For instance, \mathbb{R} and \mathbb{C} are fields. $\mathbb{Z}_n = \{0, 1, \dots, n-1\}$ for n prime is also a field. The main difference between a field and skew field is that multiplication is commutative in one and not in the other, i.e. $a * b \neq b * a$ for two quaternions but is so for complex numbers.

Linear space (A commutative group over a field): \mathcal{L} is a linear space over κ if

- \mathcal{L} is a commutative group
- κ is a field
- For all $(\lambda, x) \in (\kappa, \mathcal{L})$, there is $\lambda x \in \mathcal{L}$
- $\lambda(x + y) = \lambda x + \lambda y$, $\lambda(\mu x) = \lambda \mu x$, $(\lambda + \mu)x = \lambda x + \mu x$
- $1x = x$ where 1 is identity of κ

Elements of linear spaces are usually called *vectors*.^{do} For instance $\mathfrak{M}(m, n, \kappa)$ of all $m \times n$ matrices from a field κ is linear space over κ , which is of dimension $m \times n$. We note that the relevant operations are element-wise addition (i.e. $M_1 + M_2$) and multiplication by field elements (i.e. cM). For this linear space, we can always choose the basis as the matrices $e_{ij} = \delta_{ij}$.

Linear algebra (A commutative ring over a field = A linear space with multiplication among its elements): A linear space \mathcal{L} over a field κ is a linear algebra over κ if

- There is element-wise multiplication of \mathcal{L} , hence \mathcal{L} is a ring
- $(\lambda x)y = x(\lambda y) = \lambda(xy)$ for λ element of the field and x, y element of the linear algebra

In order to define a multiplication in a linear algebra \mathcal{L} of finite dimension, it is sufficient to give the product of its basis vectors:

$$e_i e_j = \sum_k c_{ij}^k e_k \quad (2.5)$$

Here, c_{ij}^k are called *structure constants* of the algebra \mathcal{L} .

The most common example for a linear algebra would be the space (n, κ) of all $n \times n$ matrices over the field κ , which forms an associative and non-commutative linear algebra.^{dp}

Lie algebra (A linear algebra whose element-wise multiplication is anticommutative): A linear algebra \mathcal{I} over a field κ is a *Lie algebra* if

- Multiplication in \mathcal{I} is anticommutative.^{dq} They are denoted as $[X, Y] = -[Y, X]$ and are called commutator.

^{do} We can remind ourselves following properties of vectors:

- Vectors x_1, \dots, x_n are *linearly independent* if $\sum_{m=1}^n \lambda_m x_m = 0 \Rightarrow \lambda_{1,\dots,n} = 0$ for $\lambda_m \in \kappa$
- A set $\{e_\alpha\}$ of vectors forms a *basis* of \mathcal{L} if for any $x \in \mathcal{L}$ we have a unique linear combination $x = \lambda_\alpha e_\alpha$.
- Any set of n linearly independent vectors of n -dimensional linear space is a basis.
- **Dimension of \mathcal{L} :** The maximal number of linearly independent vectors of a linear space \mathcal{L} . If we can find arbitrary many independent vectors, \mathcal{L} is infinite dimensional.

^{dp} In this algebra, the element-wise multiplication is simply the matrix multiplication, hence identity element is the identity matrix δ_{ij} . As we can choose the basis $e_{ij} = \delta_{ij}$, the element multiplication reads $e_{ij}e_{km} = \delta_{jk}e_{im}$, which yields the structure constants $c_{ij,km}^{rs} = \delta_{ir}\delta_{jk}\delta_{ms}$.

^{dq} If we have $[X, Y] = 0$ for any elements X, Y of a Lie algebra \mathcal{I} , then \mathcal{I} is said to be *commutative*.

- Elements satisfy *the Jacobi identity* $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$

The structure constants of a Lie algebra satisfy the identities

$$c_{ij}^k = -c_{ji}^k, \quad \sum_j (c_{ij}^\ell c_{km}^j + c_{mj}^\ell c_{ik}^j + c_{kj}^\ell c_{mi}^j) = 0 \quad (2.6)$$

Conversely, any algebra whose structure constants satisfy these conditions is a Lie algebra.

In general, we can obtain a *Lie algebra* from any algebra \mathcal{L} over a field κ by choosing $[X, Y] = XY - YX$, if multiplication of \mathcal{L} is associative. With this construction, \mathcal{I} has same elements with \mathcal{L} .^{dr} For instance, the set of vectors in a three-dimensional space is a Lie algebra under vector multiplication.

Exponential map (from algebra back to the group): We have started with groups first and constructed lie algebras from them step by step, which is summarized in Fig. (2.2). We can close the loop now and get back to the groups from algebras via so-called *exponential map*!

We will not dwell on technical details or subtleties; instead, let's simply state blankly that the group space^{ds} can be a differentiable manifold for some groups and such groups are called Lie groups. For instance, the translations in two dimensions have group elements $g \in \mathbb{R}^2$, e.g. $(1, 3) + (2, -5) = (3, -2)$. Thus, the group space is \mathbb{R}^2 , a differentiable manifold, and hence the translations form a Lie group.

The group manifold for the Lie groups can be rather complicated; however, if we sufficiently zoom around the identity element of the group, the group space should become flat (i.e. Euclidean space) and we should be able to construct a linear algebra. In fact, we can go ahead and construct a Lie algebra (whose identity element matches identity element of the group). By exponentiating the action of the Lie algebra, we can zoom out of the infinitesimal neighborhood around the identity: the exponentiation of the action of the Lie algebra generates the action of the Lie group, *at least part of it which can be obtained continuously from the identity element*.

This explanation is rather abstract, so let's illustrate it with the simplest case: translations in d -dimensions. Clearly, the group space is \mathbb{R}^d and the elements of the group are usual d -dimensional vectors, i.e. $\vec{a} \in \mathbb{R}^d$. The action of the group *on the group space* is trivial, i.e.

$$a^i : x^i \rightarrow x^i + a^i \quad (2.8)$$

which is inhomogeneous but can be upgraded to a homogeneous transformation^{dt} as

$$\begin{pmatrix} a^1 \\ a^2 \\ \vdots \\ a^d \end{pmatrix} : \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^d \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & \dots & 0 & a^1 \\ 0 & 1 & \dots & 0 & a^2 \\ \ddots & & \ddots & & \vdots \\ 0 & 0 & \dots & 1 & a^d \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^d \\ 1 \end{pmatrix} \quad (2.9)$$

^{dr}A concrete example of this procedure is given by the set of all $n \times n$ matrices over the field κ , i.e. $\mathfrak{M}(n, \kappa)$, which is a Lie algebra over κ . The commutation relations for its basis matrices e_{ij} and the structure constants are given as

$$[e_{ij}, e_{km}] = \delta_{jk} e_{im} - \delta_{mi} e_{kj}, \quad c_{ij,km}^{rs} = \delta_{ir} \delta_{jk} \delta_{ms} - \delta_{kr} \delta_{im} \delta_{js} \quad (2.7)$$

^{ds}The abstract space filled with the elements of the group.

^{dt}In this context, we state that the transformation $x \rightarrow M \cdot x$ is homogeneous whereas $x \rightarrow M \cdot x + a$ is inhomogeneous.

which is actually a special case of *Affine group*.^{du} If we instead focus on the action of the group on *scalar functions defined on \mathbb{R}^d* , then we have

$$a^i : f(x^i) \rightarrow f(x^i + a^i) \quad (2.10)$$

If we denote by $T(a^i)$ the group action of the translation, then we have

$$T(a^i)f(x^i) = f(x^i + a^i) \quad (2.11)$$

To find T , we Taylor-expand for small a^i ,^{dv} which reads

$$T(0)f(x^i) + a^j \frac{\partial T(a_k)}{\partial a_j} \Big|_{a_k=0} f(x^i) + \dots = f(x^i) + a^j \partial_j f(x^i) + \dots \quad (2.12)$$

Thus, the solution reads

$$T(a_i) = \exp(a_i \partial_j) \quad (2.13)$$

which can be straightforwardly checked. Thus, the partial derivative *generates* the translation and its exponentiation gives the group element of translations. We then immediately see that

$$[\partial_i, \partial_j] = 0 \quad (2.14)$$

implies the translation algebra is abelian: we have d -copies of the algebra \mathbb{R} .

Translations, rotations, and dilations

In the previous subsection, we have given bunch of technical details regarding the definitions of mathematical concepts of algebras, groups, etc. Such definitions and technicalities will be barely needed, but are important for the sake of completeness.

With the technicalities dealt with, we can now focus on deriving the Galilean and dilation algebras. We will do this by analyzing the relevant transformations to extract the algebras. Earlier, we showed that the translation algebra is generated by ∂_i and this can be seen from the action of the translation group on scalar functions on the group space, i.e.

$$e^{a_i \partial_i} f(x_i) = f(x_i + a_i) \quad (2.15)$$

But this can be generalized further: if the action of the group changes a parameter κ_i to $\kappa_i + a_i$, the relevant element of the algebra is $\frac{\partial}{\partial \kappa_i}$, i.e. $\exp\left\{a_i \frac{\partial}{\partial \kappa_i}\right\} f(\kappa_i) = f(\kappa_i + a_i)$, for any $\kappa_i = \kappa_i(x_j)$.

Let's illustrate this with scaling transformations, i.e. $x^i \rightarrow e^\lambda x^i$ for $\lambda \in \mathbb{R}$.^{dw} This is a one dimensional group and is called the dilation group. Clearly, dilation does not translate x^i ; however, we see that it does translate the logarithm of the norm of x_i :

$$\lambda : \frac{1}{2} \log(x_i x_i) \rightarrow \frac{1}{2} \log(x_i x_i) + \lambda \quad (2.16)$$

^{du}We do not need to know anything about these in these notes. [Maybe I should mention physical uses of affine group](#)

^{dv}This is no coincidence: $a^i = 0$ is the identity element of the translation group, which is why we are expanding around it!

^{dw}Note that this transformation does not change the *direction* of x_i as $e^\lambda \geq 0$. This is intentional because the transformation $x_i \rightarrow -x_i$ is not part of the dilation group.

Therefore, we can immediately write down the differential representation of the element in the dilation algebra $d(1)$:

$$d(1) = \left\{ \frac{\partial}{\partial \left(\frac{1}{2} \log(x_i x_i) \right)} \right\} \quad (2.17)$$

If we define the norm $r = \sqrt{x_i x_i}$, then $\frac{1}{2} \log(x_i x_i) = \log(r)$ and $\partial/\partial \log(r) = r \frac{\partial}{\partial r}$. As $r \frac{\partial}{\partial r} = x_i \partial_i$, ^{dx} we conclude

$$d(1) = \{x_i \partial_i\} \quad (2.18)$$

which is the generator of scaling transformations, i.e.

$$e^{\lambda x^i \partial_i} f(x^i) = f(e^\lambda x^i) \quad (2.19)$$

A somehow more nontrivial example can be given by rotations. In \mathbb{R}^d , rotation transforms a vector x_i to x'_i by a matrix multiplication in Cartesian coordinates, i.e.

$$x'_i = M_{ij} x_j \quad (2.20)$$

Clearly we do not have a translation. However, we can switch to generalized spherical coordinates such that

$$\begin{aligned} x_1 &= r \sin \theta_1 \sin \theta_2 \dots \sin \theta_{d-3} \sin \theta_{d-2} \cos \phi \\ x_2 &= r \sin \theta_1 \sin \theta_2 \dots \sin \theta_{d-3} \sin \theta_{d-2} \sin \phi \\ x_3 &= r \sin \theta_1 \sin \theta_2 \dots \sin \theta_{d-3} \cos \theta_{d-2} \\ x_4 &= r \sin \theta_1 \sin \theta_2 \dots \cos \theta_{d-3} \\ &\vdots \\ x_d &= r \cos \theta_1 \end{aligned} \quad (2.21)$$

If we now do a rotation in $x_1 - x_2$ plane by an angle α , the transformation is simply a translation, i.e.

$$\alpha : (r, \theta_1, \dots, \theta_{d-3}, \phi) \rightarrow (r, \theta_1, \dots, \theta_{d-3}, \phi + \alpha) \quad (2.22)$$

hence we can immediately write down the generator of the rotation as $\frac{\partial}{\partial \phi}$. But by chain rule, it simply reads as

$$\frac{\partial}{\partial \phi} = \frac{\partial x_i}{\partial \phi} \frac{\partial}{\partial x_i} = x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1} \quad (2.23)$$

which indicates that a rotation in $x_1 - x_2$ plane is generated by $x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1}$. But there is nothing special about this plane and we can always choose other spherical coordinates where ϕ lies in $x_i - x_j$ plane, hence we conclude

$$m_{ij} = x_i \partial_j - x_j \partial_i \quad (2.24)$$

generates a rotation in $x_i - x_j$ plane. By exponentiation, we can then write

$$e^{\theta_{ij} m_{ij}} f(x_k) = f(M_{k\ell} x_\ell) \quad (2.25)$$

^{dx}One intuitive way to see this is to note that $r \partial_r$ is invariant under rotations, hence we need an $so(d)$ scalar (i.e. no open indices). The only such scalar with one partial derivative is $x_i \partial_i$.

where $M_{k\ell}$ is the rotation matrix that rotates \vec{x} in $x_i - x_j$ plane by an angle θ_{ij} .

Unlike translations or dilations, the algebra of rotations is nontrivial, i.e. we have *non-zero* structure constants! That's a fancy way of saying that rotations in different planes do not commute, e.g. rotating in $x-y$ plane by $\pi/3$ and then rotating in $y-z$ plane by $\pi/6$ yields a different vector than the one we obtain by the reverse order of rotations. We can see this explicitly as

$$m_{12}m_{23} - m_{23}m_{12} = x_1\partial_3 - x_3\partial_1 = m_{13} \quad (2.26)$$

which actually tells us that the difference of whether we apply $x-y$ rotation or $y-z$ rotation is itself equivalent to a rotation in $x-z$ plane! One can actually compute all possible cases and write down that

$$[m_{ij}, m_{k\ell}] = \eta_{i\ell}m_{jk} + \eta_{jk}m_{li} + \eta_{ik}m_{\ell j} + \eta_{j\ell}m_{ik} \quad (2.27)$$

Clearly, this contrasts with the commutators for the generators of translation and dilation algebras, which we will denote as p_i and d : $[p_i, p_j] = 0$, $[d, d] = 0$ (rather trivially).

Now that we constructed the individual algebras for translations, dilations, and rotation, we can try to combine them to generate bigger algebras. A pedestrian way to do this is to use their differential forms and compute the commutators directly; for instance,

$$[p_i, d] = \partial_i(x^j\partial_j) - x^j\partial_j\partial_i = \partial_i = p_i \quad (2.28)$$

With explicit computation, we can then write down

$$[p_i, m_{jk}] = \eta_{ij}p_k - \eta_{ik}p_j \quad (2.29a)$$

$$[p_i, d] = p_i \quad (2.29b)$$

$$[m_{ij}, d] = 0 \quad (2.29c)$$

^{dy}Our procedure is non-canonical: one usually starts with algebra abstractly, deals with its representations, and move on to groups with exponential map. We instead glossed over representations and started with the action of the symmetry to derive the algebra. I believe this ordering is more pedagogical for a physicist.

2.3 Conformal algebra

Special conformal transformations

In the previous section, we reviewed the mathematical background on groups and algebras, and showed that the algebra of translation, rotation, and dilation symmetries can be extracted from their action via exponential map. ^{dy} As algebras, they have the elements

$$p_i = \partial_i, \quad d = x_i\partial_i, \quad m_{ij} = x_i\partial_j - x_j\partial_i \quad (2.30)$$

with the commutation relations given in eqn. (2.29). As group elements, they act as

$$\begin{aligned} D(\lambda) : x_i &\rightarrow e^\lambda x_i \\ M(\alpha) : (r, \theta_i, \phi) &\rightarrow (r, \theta_i, \phi + \alpha) \\ P(a_i) : x_i &\rightarrow x_i + a_i \end{aligned} \quad (2.31)$$

There is an interesting observation: both dilations and rotations leave two points invariant: origin and the infinity.^{dz} In comparison, translations leave only infinity invariant, i.e. $0 + a_i \neq 0$ but $\infty + a_i = \infty$. Therefore, translations *break* the symmetry between origin and the infinity. It is thus tempting to consider another transformation, which moves the infinity but keeps the origin fixed. We can achieve this by exchanging the origin and the infinity before and after we do a usual translation!

One can in principle write down many transformations that exchange the origin with infinity but the simplest such transformation is *inversion* I , which acts in one dimension as $x \rightarrow \frac{1}{x}$. Indeed, $x = 0$ would go to ∞ and vice versa. In higher dimensions, we define it as^{ea}

$$I : x_i \rightarrow \frac{x_i}{x^2} \quad (2.32)$$

Now, if we consider a translation in inverted coordinates, i.e.

$$k_i : \frac{x_i}{x^2} \rightarrow \frac{x_i}{x^2} + a_i \quad (2.33)$$

then this transformation leaves the origin invariant and moves infinity. We call this transformation *special conformal transformations*!

It is straightforward to check that

$$x'_i = \frac{x_i + a_i x^2}{1 + 2a \cdot x + a^2 x^2} \quad (2.34)$$

satisfies $\frac{x'_i}{x'_k x'_k} = \frac{x_i}{x_k x_k} + a_i$, hence the transformation can be rewritten as^{eb}

$$k_i : x_i \rightarrow \frac{x_i + a_i x^2}{1 + 2a \cdot x + a^2 x^2} \quad (2.36)$$

There is an advantage to the form presented in eqn. (2.33): it is precisely a translation for the parameter $\kappa(x_i) = \frac{x_i}{x^2}$, hence as discussed in section [dv](#), we can immediately write down the algebra element as

$$k_i = \frac{\partial}{\partial \left(\frac{x_i}{x^2} \right)} \quad (2.37)$$

which then gives^{ec}

$$k_i = (x^2 \delta_{ik} - 2x_i x_k) \partial_k \quad (2.42)$$

The algebra generated by k_i is called *special conformal algebra*, and is denoted as $sc(d)$.

One can now use the explicit form above to work out the new commutation relations:

$$[k_i, k_j] = 0 \quad (2.43a)$$

$$[k_i, d] = -k_i \quad (2.43b)$$

$$[p_i, k_j] = 2m_{ij} - 2\eta_{ij}d \quad (2.43c)$$

$$[m_{ij}, k_\ell] = \eta_{j\ell} k_i - \eta_{i\ell} k_j \quad (2.43d)$$

^{dz}If we consider a limiting point to infinity, then rotations would change its direction. But if we consider infinity as a point and add it to \mathbb{R}^d , then the action of the rotation on the resultant space is indeed invariant if acted on the point ∞ , see Fig. (2.3).

^{ea}Note that this a discrete transformation: we only have two elements of the inversion group, the identity E and the inversion I with the relation $I * I = E$.

^{eb}It is also straightforward to check that this transformation leaves origin invariant and moves the infinity:

$$\lim_{x_i \rightarrow 0} x'_i = 0, \quad \lim_{x_i \rightarrow \infty} x'_i = \frac{a_i}{a^2} \quad (2.35)$$

^{ec}We use chain rule to derive this. First note that

$$\frac{\partial}{\partial x_i} = \frac{\partial \left(\frac{x_k}{x^2} \right)}{\partial x_i} \frac{\partial}{\partial \left(\frac{x_k}{x^2} \right)} \quad (2.38)$$

which gives

$$\frac{\partial}{\partial x_i} = (\delta_{ik} x^{-2} - 2x_i x_k x^{-4}) \frac{\partial}{\partial \left(\frac{x_k}{x^2} \right)} \quad (2.39)$$

Thus, we have

$$\frac{\partial}{\partial \left(\frac{x_i}{x^2} \right)} = M_{ik} \frac{\partial}{\partial x_k} \quad (2.40)$$

for

$$(\delta_{ik} x^{-2} - 2x_i x_k x^{-4}) M_{kl} = \delta_{il} \quad (2.41)$$

We can then make an ansatz for $M_{kl} = c_1 \delta_{kl} + c_2 x_k x_l$ and solve for c_i .

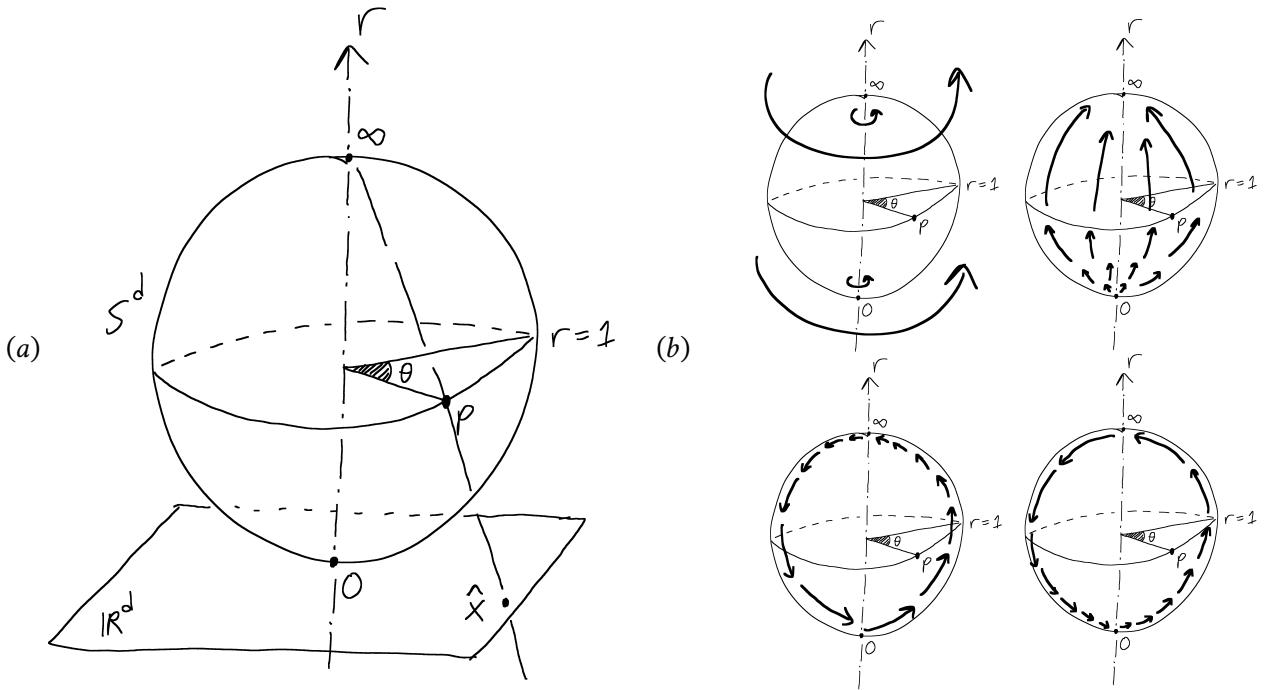


Figure 2.3: (a) The Euclidean space \mathbb{R}^d is topologically uncompact; one can construct a one-point compactification of it, which is d -sphere: $S^d = \mathbb{R}^d \cup \{\infty\}$. One way to see this correspondence is through a stereographic projection, which is shown in the image: we embed S^d and \mathbb{R}^d into \mathbb{R}^{d+1} such that the south pole of the sphere is at the origin of \mathbb{R}^d (their only common point in \mathbb{R}^{d+1}). We then consider a line in \mathbb{R}^{d+1} passing through the north pole of S^d , and identify the points that it intersects on S^d and on \mathbb{R}^d , e.g. \hat{x} and P . This procedure maps origin of \mathbb{R}^d to the south pole, inside the unit circle of \mathbb{R}^d to south hemisphere, and outside the unit circle of \mathbb{R}^d to the north hemisphere. The only point that is unmatched by this procedure is the north pole itself, which we identify with ∞ . Lastly, we see in the image that the axis that connects the poles is labeled as r because different altitudes on S^d correspond to different norms on the \mathbb{R}^d , whereas θ schematically refers to the angular part of a position in \mathbb{R}^d , i.e. (θ_i, ϕ) in eqn. (2.21). (b) Conformal transformations as movements of points on S^d . Under rotations (top-left), the d -sphere rotates such that altitudes remain unchanged. Under dilations (top-right), the sphere instead stretch on the south hemisphere and compress in the north hemisphere (or visa-versa) without rotation. Under translations (bottom-left), the sphere rotates about any particular axis (not necessarily around r as rotations do), but it also stretch and compress so that the north pole remains at the north pole. Special conformal transformations (bottom-right) has exactly the same effect, except the south pole remains unchanged instead of north pole. As expected, origin is unchanged by rotations, dilations, and special conformal transformations whereas infinity is unchanged by rotations, dilations, and translations.

We observe a similarity between translations and special conformal transformations in the commutation relations. As our way of deriving the special conformal algebra relied on the action of translations on

inverted coordinates, this similarity is no surprise to us. We can use this similarity to rewrite the full set of commutation relations as

$$[m_{ij}, m_{kl}] = \eta_{il}m_{jk} + \eta_{jk}m_{li} + \eta_{ik}m_{lj} + \eta_{jl}m_{ik} \quad (2.44a)$$

$$[m_{ij}, p_k^\pm] = \eta_{jk}p_i^\pm - \eta_{ik}p_j^\pm \quad (2.44b)$$

$$[p_i^+, p_j^-] = 2\eta_{ij}d - 2m_{ij} \quad (2.44c)$$

$$[p_i^\pm, d] = \pm p_i \quad (2.44d)$$

where we defined

$$p_i^+ := p_i, \quad p_i^- := -k_i. \quad (2.45)$$

The action of these transformations on \mathbb{R}^d can be visualized by *compactifying* the space to S^d , see Fig. (2.3) for details. In the next section, we will analyze how we can construct bigger algebras starting with these commutation relations.

(Semi)direct sums of algebras

In principle, there are three cases when we try to combine two algebras g_1 and g_2 with elements $e_{1i} \in g_1$ and $e_{2j} \in g_2$:

1. $[e_{1i}, e_{2j}] = 0 \forall i, j$. In this case we can combine g_1 and g_2 as a *direct sum*:

$$g = g_1 \oplus g_2 \quad (2.46)$$

2. $[e_{1i}, e_{2j}] \in g_1 \forall i, j$. In this case, we can combine g_1 and g_2 as a *semi-direct sum*:

$$g = g_1 \oplus_s g_2 \quad (2.47)$$

3. Otherwise. In this case, we cannot write down an algebra g by combining just g_1 and g_2 .

As an example for the last case, we can take $g_1 = t(d)$ and $g_2 = sc(d)$: as we can see from eqn. (2.43c), the commutator generates elements that belong to neither $t(d)$ nor $sc(d)$, so these algebras cannot be combined *without* including more elements.

As for the direct sum of two algebras, we can consider rotations and dilations, i.e. $g_1 = so(d)$ and $g_2 = d(1)$. As $[m_{ij}, d] = 0$, we can construct a new algebra $g = so(d) \oplus d(1)$.

Let us now focus on the second case: semi-direct sums of algebras. The first example we can construct is the *Euclidean algebra* $e(d)$, which consists of translations and rotations. As a commutation of these two actions is a translation itself, we can choose $g_1 = t(d)$ and $g_2 = so(d)$ and write

$$e(d) = t(d) \oplus_s so(d) \quad (2.48)$$

For notational simplicity we have been focusing on algebras on \mathbb{R}^d , but this is not a requirement. If we consider the space $\mathbb{R}^{p,q}$, we can also write down

$$g(p, q) = t(p, q) \oplus_s so(p, q) \quad (2.49)$$

In particular, for $q = 1$, this becomes the *Poincaré* algebra:

$$p(d-1, 1) = t(d-1, 1) \oplus_s so(d-1, 1) \quad (2.50)$$

The second example for the semi-direct sum would be between the Euclidean algebra $e(d)$ and the dilation algebra $d(1)$, i.e.

$$g(d) = e(d) \oplus_s d(1) \quad (2.51)$$

or, in Lorentzian signature,

$$g(d-1, 1) = p(d-1, 1) \oplus_s d(1) \quad (2.52)$$

which is *the relevant algebra for scale invariant but not conformally invariant field theories with the usual Poincaré invariance*. By using the fact that rotations and scaling commute, we can rewrite this algebra as^{ed}

$$g(d-1, 1) = t(d-1, 1) \oplus_s (so(d-1, 1) \oplus d(1)) \quad (2.53)$$

Realization of conformal algebra in the embedding space

The commutation relations of the generators of *general conformal algebra* $gc(d)$ was given in eqn. (2.44):

$$[m_{ij}, m_{kl}] = \eta_{il}m_{jk} + \eta_{jk}m_{li} + \eta_{ik}m_{lj} + \eta_{jl}m_{ik} \quad (2.54a)$$

$$[m_{ij}, p_k^\pm] = \eta_{jk}p_i^\pm - \eta_{ik}p_j^\pm \quad (2.54b)$$

$$[p_i^+, p_j^-] = 2\eta_{ij}d - 2m_{ij} \quad (2.54c)$$

$$[p_i^\pm, d] = \pm p_i \quad (2.54d)$$

The action of these operations can be best described in S^d , which is the compactification of \mathbb{R}^d : we illustrate this in Fig. (2.3). We see that d generates a uniform stretching/compression of S^d (1-generator), m generates a rotation in $i - j$ plane for $i \neq r$ ($\binom{d}{2}$ generators), and p^\pm generate a combination of stretching/compression and rotation in $r - i$ plane for $i \neq r$ (d generators each).

A somewhat nicer way to understand the conformal algebra is through the re-organization of its elements in terms of M_{IJ} as follows:

$$\begin{aligned} m_{ij} &= M_{ij} \\ d &= M_{0,-1} \\ p_i^+ &= M_{i,-1} - M_{i,0} \\ p_i^- &= M_{i,-1} + M_{i,0} \end{aligned} \quad (2.55)$$

where the index $I = -1, 0, 1, \dots, d$ whereas the index $i = 1, 2, \dots, d$. If we now work out the commutation relations for M_{IJ} , we find that

$$[M_{IJ}, M_{KL}] = \eta_{IL}M_{JK} + \eta_{JK}M_{LI} + \eta_{IK}M_{LJ} + \eta_{JL}M_{IK} \quad (2.56)$$

which is precisely the same algebra as eqn. (2.44a) with the metric $\eta_{IJ} = \text{diag}(-1, 1, \dots, 1)$. Therefore, we conclude that the Euclidean conformal

^{ed}This algebra is also called *Weyl algebra*, and can be denoted as $w(3, 1)$, see section 4.1.4 of [52]. But this does not seem to be a universal usage; for instance, see page 176 of [53] which denotes the relevant group as *homogeneous Weyl group* (denoted by HW) instead. It is unclear to me what terminology the mathematicians use for this algebra. [maybe I should actually ask a mathematician](#).

algebra is isomorphic to the Lorentzian rotation algebra in two higher dimensions:^{ee}

$$gc(d) \simeq so(d+1, 1) \quad (2.57)$$

But this statement can be straightforwardly generalized to other signatures:

$$gc(p, q) \simeq so(p+1, q+1) \quad (2.58)$$

The isomorphism between conformal algebra and rotation algebra is great news because we are very accustomed to working with rotation algebra and its representations: scalars, vectors, tensors, etc. However it is not immediately clear how we can make use of this isomorphism directly because the rotation algebra acts on $\mathbb{R}^{d+1,1}$ whereas we want the action of the conformal group on \mathbb{R}^d . In short, we need to get rid of two extra degrees of freedom in a rotation-invariant way.

Let us label the position in $\mathbb{R}^{d+1,1}$ with X_I and in \mathbb{R}^d with x_i . We are trying to find the relation $X_I = X_I(x_i)$. As we would like to embed x_i into X_I , we will take $X_i = x_i$. Furthermore, as $X_{-1,0}$ are scalar functions of x_i , they can only depend on the norm of the vector, hence we have

$$X_I = (X_{-1}(x^2), X_0(x^2), x_i) \quad (2.59)$$

We would like to find the functions $X_{-1,0}$.

To find the functional form of X_I , we use the fact that conformal transformations act as rotations on X_I . We first change our coordinate system so that x_i reads as $(x, 0, \dots, 0)$.^{ef} We then have $X_I = (X_{-1}(x^2), X_0(x^2), x, 0, \dots, 0)$.^{eg} Let's now consider an infinitesimal special conformal transformation; by eqn. (2.36) we get $x_i = (x, 0, \dots) \rightarrow x'_i = (x - \alpha_1 x^2, \alpha_{i \neq 1} x^2)$, hence

$$(X_{-1}(x^2), X_0(x^2), x_i) \rightarrow (X_{-1}(x^2 - 2\alpha_1 x^3), X_0(x^2 - 2\alpha_1 x^3), x - \alpha_1 x^2, \alpha_{i \neq 1} x^2) \quad (2.60)$$

Now, we need this transformation to be a rotation. If we focus on X_1 , an infinitesimal rotation would take it to the form

$$X_1 \rightarrow X_1 + c_{-1}(\alpha_i)X_{-1}(x^2) + c_0(\alpha_i)X_0(x^2) \quad (2.61)$$

for some functions c_i linear in α_i . With $X_1 = x$ and $X'_1 = x - \alpha_1 x^2$, this is only possible if $X_{-1,0}$ are linear in x^2 , thus we conclude

$$X_I = (a_{-1} + b_{-1}x^2, a_0 + b_0x^2, x_i) \quad (2.62)$$

for some unknowns a_i and b_i . Let us again go back to our earlier coordinate system where $X_I = (a_{-1} + b_{-1}x^2, a_0 + b_0x^2, x, 0, \dots)$, and this time consider a dilation $x \rightarrow \lambda x$. We then have

$$(a_{-1} + b_{-1}x^2, a_0 + b_0x^2, x, 0, \dots) \rightarrow (a_{-1} + b_{-1}\lambda^2 x^2, a_0 + b_0\lambda^2 x^2, \lambda x, 0, \dots) \quad (2.63)$$

It is impossible to get this transformation as a rotation; under a rotation, X_1 would mix with the other nonzero entries linearly, hence it would become $x \rightarrow x + \theta_1 \text{constant} + \theta_2 x^2$ which is not of the form λx .

^{ee}As a consistency check, we see that the number of generators match as they should: $\frac{d(d-1)}{2} + d + d + 1 = \frac{(d+2)(d+1)}{2}$.

^{ef}We can do this without losing generality by choosing our coordinate system such that the unit vector \hat{e}_1 points along \vec{x} .

^{eg}Note that we are abusing the notation here: $x \equiv \sqrt{x_i x_i}$.

How do we solve this problem? Well, we established algebraically that the conformal algebra $gc(d)$ is isomorphic to a rotation algebra $so(d+1, 1)$ so it *has to* act as a rotation in the embedding space! However, it does not necessarily act on the *points* of the embedding space. If we instead take the action of the rotation algebra on the *rays* of the embedding space, we can identify $X \sim \Lambda X$ for any $\Lambda > 0$, ^{eh} thus we can write

$$(a_{-1} + b_{-1}x^2, a_0 + b_0x^2, x, 0, \dots) \rightarrow (a_{-1} + b_{-1}\lambda^2x^2, a_0 + b_0\lambda^2x^2, \lambda x, 0, \dots) \sim (a_{-1}\lambda^{-1} + b_{-1}\lambda x^2, a_0\lambda^{-1} + b_0\lambda x^2, x, 0, \dots) \quad (2.64)$$

which can indeed be realized as a rotation among X_{-1} and X_0 : we conclude that dilations correspond to rotations between these two coordinates, which makes sense as dilation is a 1-parameter transformation and hence should correspond to an $so(1, 1)$ transformation!

This *acting-on-the-rays* business may look intimidating to some readers but it is actually quite expected: we derived it from the action of the dilation operation, but we can also see this from the counting of the degrees of freedom: x_i has d -degrees of freedom whereas X_I has $d+2$ degrees of freedom: we need two $so(d+1, 1)$ invariant constraints! One such constraint is $X_I X^I = \text{constant}$ and the other is $X_I \sim \lambda X_I$.

Let us now use these two constraints to find the expressions for a_i and b_i . $X_I X^I$ reads as

$$X^2 = (b_0^2 - b_{-1}^2)x^4 + (1 + 2a_0b_0 - 2a_{-1}b_{-1})x^2 + (a_0^2 - a_{-1}^2) \quad (2.65)$$

This should not change under an arbitrary conformal transformation $x \rightarrow x'$. This is only possible if it is independent of x , ^{ei} which relate a_i and b_i , allowing us to write

$$X_I = \left(c \left[\frac{1 + 2a_0b_0}{2b_0} + b_0x^2 \right], a_0 + b_0x^2, x_i \right) \quad (2.66)$$

for an unknown $c = \pm 1$. We can set $c = 1$ without losing any generality (we can at most reflect X_{-1} direction as our definition); furthermore we can set $b_0 = -\frac{1}{2}$ by using $X \sim \lambda X$. Thus ^{ej}

$$X^I = \left(1 - a_0 + \frac{1}{2}x^2, a_0 - \frac{1}{2}x^2, x^i \right) \quad (2.67)$$

We are almost there! It is very tempting to conjecture that $a_0 = \frac{1}{2}$, giving us a very nice final result of $X^I = (\frac{1}{2} + \frac{1}{2}x^2, \frac{1}{2} - \frac{1}{2}x^2, x^i)$, and this is indeed correct! But to see why, we need to fix an ambiguity in our $X_i \leftrightarrow x_i$ relation.

To see the problem, let's remind ourselves that $X^I \sim \lambda X^I$, hence a position x^i corresponds to any λX^i , which is another way to say that our point x^i in \mathbb{R}^d corresponds to a *ray* X^i . However, we would like to extract the point x^i out of the ray X^i unambiguously, and we do this by *fixing the gauge*, i.e. removing the redundancy. Looking at the form of (X^{-1}, X^0, X^i) in eqn. (2.67), we can declare that we have the relation

$$x^i = \frac{X^i}{X^+} \quad (2.68)$$

^{eh}mention why/when lambda needs to be positive, euclidean vs lorentzian cases, etc

^{ei} We can actually do more than this by noting that X^2 should not only be constant but also remain invariant under our identification $X \rightarrow \lambda X$. We can then choose $X^2 = 0$, which completely fixes the form of X_A . However, despite being a huge simplification, the choice of $X^2 = 0$ is not unique in that $X^2 = \infty$ is also invariant under $X \rightarrow \lambda X$. We can of course say that ∞ is not part of $\mathbb{R}^{d+1,1}$, hence X^2 can't be chosen that way, but this could be circumcised via the compactification of the space by including the infinity as a point. There are probably mathematical arguments against these so at the end $X^2 = 0$ can be the unique choice from this perspective (or even that $X^2 = \infty$ is equivalent to $X^2 = 0$), however we do not need any of these arguments. We will take the longer path, leave $X^2 = \text{constant}$ without specifying what it is and then determine it by matching the action of conformal transformation on \mathbb{R}^d and $\mathbb{R}^{d+1,1}$, which will uniquely fix that $X^2 = 0$.

^{ej}Note that we also use $X^{(-1)} = -X_{(-1)}$ when we go from X_I to X^I .

where we define the *lightcone coordinates*^{ek}

$$X^\pm = X^{(-1)} \pm X^{(0)} \quad (2.69)$$

We now have an unambiguous mapping from X_I to x_i !

With our mapping fix, let us now consider an infinitesimal conformal transformation. In the embedding space, the general rotation transformation reads as

$$X'_I = e^{\omega_{IJ}} X^J \quad (2.70)$$

where ω_{IJ} is the rotation matrix. If the transformation is infinitesimal, we then have

$$\delta X_I = \delta w_{IJ} X^J \quad (2.71)$$

The rotation matrices satisfy the algebra in eqn. (2.56). In particular, this means

$$\delta\omega_{ij} = \delta\theta_{ij}, \quad \delta\omega_{0,-1} = \delta\lambda, \quad \delta\omega_{i,-1} = \frac{\delta c_i + \delta a_i}{2}, \quad \delta\omega_{i,0} = \frac{\delta c_i - \delta a_i}{2} \quad (2.72)$$

which follows from eqn. (2.55); here, θ, λ, c, a denote the parameters for rotation, dilation, translation, and special conformal transformation respectively.^{el} The δ in front of them reminds us that we are working with infinitesimal transformations!

Under an infinitesimal transformation, we have

$$x^i \rightarrow \frac{X^i + \delta X^i}{X^+ + \delta X^+} = \frac{X^i}{X^+} + \frac{\delta X^i}{X^+} - \frac{X^i \delta X^+}{X_+^2} \quad (2.73)$$

thus $\delta x^i = x'_i - x_i$ reads as

$$\delta x^i = \delta X^i - x^i \delta X^+ \quad (2.74)$$

where we use the fact that $X^+ = 1$ in our choice of Poincaré section in eqn. (2.67). By using $X^{-1} = -X_{-1}$ and contracting both sides with η_{ij} ,^{em} we get

$$\delta x_i = \delta w_{iJ} X^J + x_i \delta w_{-1J} X^J - x_i \delta w_{0J} X^J \quad (2.75)$$

where we used eqn. (2.71).

Let us analyze the equation above. In components, it reads as

$$\delta x_i = \delta w_{i,-1} X^{-1} + \delta w_{i0} X^0 + \delta w_{ij} x^j + x_i \delta w_{-1,0} + x_i x^j (\delta w_{-1,j} - \delta w_{0,j}) \quad (2.76)$$

which with eqn. (2.72) becomes

$$\delta x_i = \frac{\delta c_i}{2} + \delta \theta_{ij} x^j - \delta \lambda x_i + \frac{\delta a^k}{2} (X^- \eta_{ik} - 2x_i x_k) \quad (2.77)$$

The first three terms are the expected infinitesimal translation, rotation, and dilation. The last term is equivalent to a special conformal transformation only if $X^- = x^2$,^{en} yielding the constraint

$$X^- = X^{-1} - X^0 = 1 - 2a_0 + x^2 = x^2 \quad (2.78)$$

^{ek}In Minkowski spacetime with the metric $\eta = \text{diag}(-1, 1, 1, 1)$, any coordinate $X^0 \pm X^i$ is called a *lightcone coordinate*, because light moves along these coordinates. Even though the embedding space $\mathbb{R}^{d+1,1}$ is not exactly the physical space, it is still endowed with a Lorentzian metric hence $X^{-1} \pm X^0$ are called lightcone coordinates.

^{el}It is straightforward to see θ_{ij} and λ generate rotations and dilations due to their index structures. To decide which one of $\{a_i, c_i\}$ generate translations (and the other special conformal transformations), we will take a pedestrian way and decide at the very end by checking their infinitesimal action, see below!

^{em}For our particular case here this is simply δ_{ij} as we are working with \mathbb{R}^d , i.e. Euclidean signature.

^{en}See eqn. (2.36).

thus we conclude $a_0 = \frac{1}{2}$. This completes the proof: we have a map between $x^i \in \mathbb{R}^d$ and $X^I \in \mathbb{R}^{d+1,1}$ via

$$X^I = X^+ \left(\frac{1+x^2}{2}, \frac{1-x^2}{2}, x^i \right) \quad (2.79)$$

for an arbitrary nonzero X^+ . One key observation is that this X^I not only satisfies $X_I X^I = \text{constant}$ but actually $X_I X^I = 0$,^{eo} therefore we say

There is a one to one map between points in \mathbb{R}^d and null rays in $\mathbb{R}^{d+1,1}$.

Our choice of Poincaré section in eqn. (2.79) works well everywhere except for a specific ray: $\{X^I | X^I \sim \lambda X^I, X^+ = 0\}$, i.e. the ray for which $X^+ = 0$. Our previous reasoning still applies here, so that it needs to be of the form $(\frac{1}{2} + bx^2, -\frac{1}{2} - bx^2, x^i)$ ($X^{0,-1}$ are linear in x^2) where we used $X \sim \lambda X$. But $X_I X^I = \text{constant}$ dictates $x^i = 0$, leaving us with one point: $(\frac{1}{2}, -\frac{1}{2}, 0)$.^{eq} In lightcone coordinates, this is $X^- = 1, X^{\text{rest}} = 0$. This is awfully similar to the origin, i.e. $X^+ = 1, X^{\text{rest}} = 0$. I'm sure some readers have already guessed by now; this ray corresponds to the infinity.^{er}

We conclude

$$X^I = \begin{cases} X^+ \left(\frac{1+x^2}{2}, \frac{1-x^2}{2}, x^i \right) & x \in \mathbb{R}^d \\ X^- \left(\frac{1}{2}, -\frac{1}{2}, 0 \right) & x = \infty \end{cases} \quad (2.80)$$

- asymmetry is because we choose X^I/X^+ ; we could have chosen X^I/X^-
- $X^+ \leftrightarrow X^-$ corresponds to inversion.
- Other Poincaré section possible, this one is standard

2.4 Conformal algebra in the framework of QFTs

- Killing equations and relevant discussion
- Relativity and Lorentzian transformations
- Poincare group
- generalization to conformal group

^{eo} As we discussed in the Footnote ei, we expected this result. Nevertheless, I am not aware if there is other arguments which require the matching to be such: I have checked several resources which use/discuss the embedding space formalism, [54–68],^{ep} but I could not find a satisfactory answer. I expect there exists studies where this is detailed (possibly in math literature), but the only reference I could find was [69] (see eqn b.5 there), where they derive the requirement of nullness through our way as well. Another reference where at least $X_A X^A \neq 0$ is considered is [70] (see eqn. (IV.1) there); however, they simply choose the null cone with the reasoning that it is also invariant under $X \rightarrow \lambda X$, which I find unsatisfactory.

^{ep}Most papers simply refer to earlier ones without a detailed explanation. I checked up to ~ 100 years old papers, but some of them are not in English (i.e. [71, 72]), so it is likely that I'm missing the old fundamental ones because of the language barrier. In addition, the terminology in old papers are rather different, making the task rather difficult. For instance, Kastrup derives the relation between the physical space and the embedding space by starting with the geometric interpretation of the special conformal group in [73], but one could easily miss that he is actually building the relation between x_i and X_A in our notation. For the record, he does not discuss $X_A X^A = 0$, he simply introduces X^+ (λ in his notation) such that $X^2 = 0$ without even mentioning it.

^{eq}It may seem interesting that $X^+ = 0$ ray is also a null ray! A priori it did not have to be so! But as we will see below, this ray actually corresponds to the point infinity in the compactification of \mathbb{R}^d , which can be reached by conformal transformations. Thus this ray should be reachable by rotation from other rays, hence it needs to be null as well!

^{er}Another way to obtain the $X^+ = 0$ ray is to start with eqn. (2.79) and take $X^+ \rightarrow 0$. To get a nonzero and finite point, we need to take $x^2 \rightarrow \infty$ with $X^+ x^2$ remaining constant. This takes us to $(\frac{1}{2}, -\frac{1}{2}, 0)$ as well. In this approach, it is clearer that $X^+ = 0$ null-ray corresponds to $x = \infty$.

3

How can the conformal symmetry be utilized?

3.1 Review: how the Lorentz symmetry helps in quantum field theory

Primer: basics of qft

- quantum mechanics, ward identities, path integral, quantization, LSZ formalism, correlation functions

Reminder: Wigner's classification of operators – little group formalism

3.2 Extension to conformal symmetry

A

Mathematical Preliminaries

This appendix is a quick review of some parts of first chapter of [74].

A.1 The path from groups to algebras

Below, we will construct our way starting from groups via the path

$$\begin{aligned} \text{group} &\rightarrow \text{ring} \rightarrow \text{skew field} \rightarrow \text{field} \rightarrow \text{linear space} \\ &\quad \rightarrow \text{linear algebra} \rightarrow \text{Lie algebra} \end{aligned} \quad (\text{A.1})$$

This perspective will help us realize how central groups play a role in mathematics as well as physics.

Group

Definition: A set of elements $g_i \in G$ such that

- \exists a distinguished element e such that $g * e = e * g = g, \forall g$
- \exists a mapping $g \rightarrow \hat{g}$ of G onto G such that $g * \hat{g} = \hat{g} * g = e$
- operator $*$ is associative: $g_1 * (g_2 * g_3) = (g_1 * g_2) * g_3$

e and \hat{g} are called identity and inverse respectively. In general, \hat{g} is denoted as g^{-1} , and $a * b$ is denoted as ab .^{es} If $*$ is commutative, $a * b = b * a$ is usually denoted as $a + b = b + a$, and \hat{g} denoted as $-g$. In this case e is denoted as 0.^{et}

A few useful information:

- $|G|$ is the order of the group, and is equal to the number of elements of G .^{eu}
- The set S_n of all permutations of n elements is called the symmetric group.^{ev}
- In general, the collection of all transformations of a set X (one to one mappings of X onto itself) is a group.
- The additive group of the ring $\mathbb{Z}_n = \{0, 1, \dots, n - 1\}$ is the cyclic group of order n .^{ew}

Ring = Commutative group with multiplication

Definition: The group R is a ring if

- R is commutative,
- group operation is addition,
- multiplication is defined in R with respect to addition:^{ex}

$$a(b + c) = ab + ac, (b + c)a = ba + ca \quad (\text{A.2})$$

^{es}For instance, $\mathbb{C} \setminus \{0\}$ (complex numbers except zero) form a group under multiplication with 1 being the identity element; and, one uses the notation ab and a^{-1} for multiplication and inverse. Note that the notation persists beyond this domain for multiplication, although the numbers no longer necessarily form a group.

^{et}For instance, \mathbb{C} form a commutative group under addition with 0 being the identity element; and, one uses the notation $a + b$ and $-a$ for addition and inverse. Like multiplication, this addition notation also persists beyond complex numbers even though they no longer necessarily form a group.

^{eu}For instance, the commutative group \mathbb{Z}_2 consists of two elements: e and $-e$, with the only nontrivial relation being $(-e) * (-e) = e$. We denote that as $|\mathbb{Z}_2| = 2$.

^{ev}For instance, S_3 has the elements (123), (132), (231), (213), (312), and (321) where (ijk) denotes the permutation $\{1 \rightarrow i, 2 \rightarrow j, 3 \rightarrow k\}$. Note that (123) is the identity element, and $|S_3| = 6$. One can actually show that $|S_n| = n!$.

^{ew}One calls the additive group of the ring \mathbb{Z} (the set of integers) *the infinite cyclic group*.

^{ex}In principle, the multiplication does not have to have any properties beyond its distributive property in eqn. (A.2). Its additional structure is used to label the ring, e.g. *Commutative Ring*, *Associative Ring*, and so on.

For instance, \mathbb{Z} (and the set $n\mathbb{Z}$, integers divisible by n) are rings under arithmetical addition and multiplication. The set $\mathbb{Z}_n = \{0, 1, \dots, n - 1\}$ is a ring under addition and multiplication modulo n .^{ey}

^{ey}We will see that it is also a field if n is a prime number.

Skew field = Ring whose multiplication is a group

A ring κ is a skew field if the set $\kappa \setminus \{0\}$ is a group under multiplication!

Field = Skew field whose multiplication is commutative

A ring κ is a field if the set $\kappa \setminus \{0\}$ is a commutative group under multiplication!^{ez}

Linear space = A commutative group over a field

Definition: \mathcal{L} is a linear space over κ if

- \mathcal{L} is a commutative group
- κ is a field
- For all $(\lambda, \vec{x}) \in (\kappa, \mathcal{L})$, there is $\lambda\vec{x} \in \mathcal{L}$
- $\lambda(\vec{x} + \vec{y}) = \lambda\vec{x} + \lambda\vec{y}$, $\lambda(\mu\vec{x}) = \lambda\mu\vec{x}$, $(\lambda + \mu)\vec{x} = \lambda\vec{x} + \mu\vec{x}$
- $1\vec{x} = \vec{x}$ where 1 is the identity of κ

^{ez} \mathbb{R} , \mathbb{C} , and \mathbb{Z}_n for n prime are examples to a field.

^{fa} The maximal number of linearly independent vectors of a linear space \mathcal{L} is called the *dimension* of \mathcal{L} . If we can find arbitrary many independent vectors, \mathcal{L} is infinite dimensional.

^{fb} Any set of n linearly independent vectors of n -dimensional linear space is a basis.

^{fc}The dimension of this linear space is $m \times n$, with the matrices $e_{ij} = \delta_{ij}$ being a possible basis.

^{fd} c_{ij}^k are called *structure constants* of the algebra \mathcal{L} .

Elements of linear spaces are usually called *vectors*. The set of vectors x_1, \dots, x_n are called are *linearly independent* if $\sum_{m=1}^n \lambda_m x_m = 0 \Rightarrow \lambda_{1,\dots,n} = 0$ for $\lambda_m \in \kappa$.^{fa}

A set $\{e_\alpha\}$ of vectors forms a *basis* of \mathcal{L} if for any $x \in \mathcal{L}$ we have a unique linear combination $x = \lambda_\alpha e_\alpha$.^{fb}

Example: The set $\mathfrak{M}(m, n, \kappa)$ of all $m \times n$ matrices from a field κ is a linear space over κ under element-wise addition and multiplication by field elements.^{fc}

Linear algebra = A commutative ring over a field = A linear space with multiplication among its elements

A linear space \mathcal{L} over a field κ is a linear algebra over κ if

- There is element-wise multiplication of \mathcal{L} , hence \mathcal{L} is a ring
- $(\lambda x)y = x(\lambda y) = \lambda(xy)$

If the linear algebra is finite dimensional, then it is sufficient to give a product of its basis vectors

$$e_i e_j = \sum_k c_{ij}^k e_k \quad (\text{A.3})$$

to define a multiplication in the algebra.^{fd}

Examples:

- The space $\mathfrak{M}(n, \kappa)$ of all $n \times n$ matrices over the field κ is a linear algebra under matrix multiplication.^{fe} The identity element is $I_n := \text{diag}(1, \dots, 1)$ and we can choose $e_{ij} = \delta_{ij}$ as the basis.^{ff}

We can define a map from the linear algebra \mathcal{L} over the field κ to the underlying field κ , i.e. $\mathcal{L} \rightarrow \kappa$. This map is called the determinant:

$$\det a = \sum (-1)^\alpha a_{1i_1} \cdots a_{ni_n} \quad (\text{A.4})$$

Here, the summation is over all permutations of $(1, \dots, n)$ and α is the inversions in a permutation.^{fg} This map preserves the group structure, i.e. $\det ab = \det a \times \det b$ for all $a, b \in \mathfrak{M}(n, \kappa)$.

- The group $\text{GL}(n, \kappa)$ is a special case of the linear algebra $\mathfrak{M}(n, \kappa)$: it is the set of all non-singular^{fh} matrices of degree n over a field κ .^{fi}
- The skew field of quaternions over the field κ , denoted by $\mathbf{H}(\kappa)$, is also a linear algebra.^{fj} The basis elements $\mathbf{e}_{0,1,2,3}$ satisfy:^{fk}

$$\mathbf{e}_0^2 = -\mathbf{e}_i^2 = \mathbf{e}_0, \quad \mathbf{e}_0 \mathbf{e}_i = \mathbf{e}_i, \quad \mathbf{e}_i \mathbf{e}_j = (-1)^\alpha \mathbf{e}_k \quad (\text{A.5})$$

where α is the number of inversions for (i, j, k) from $(1, 2, 3)$. Note that every non-zero element are invertible in this algebra.

Lie algebra = A linear algebra whose element-wise multiplication is anticommutative

- A linear algebra \mathcal{I} over a field κ is a *Lie algebra* if
 - Multiplication in \mathcal{I} is anticommutative. They are denoted as $[X, Y] = -[Y, X]$ and are called commutator.
 - Elements satisfy *Jacobi identity*:

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0 \quad (\text{A.6})$$

- The structure constants of a Lie algebra satisfy the identities

$$c_{ij}^k = -c_{ji}^k, \quad \sum_j (c_{ij}^\ell c_{km}^j + c_{mj}^\ell c_{ik}^j + c_{kj}^\ell c_{mi}^j) = 0 \quad (\text{A.7})$$

Conversely, any algebra whose structure constants satisfy these conditions is a Lie algebra.

- If we have $[X, Y] = 0$ for any elements X, Y of a Lie algebra \mathcal{I} , then \mathcal{I} is said to be *commutative*.
- Examples:
 - We can obtain a *Lie algebra* from any algebra \mathcal{L} over a field κ by choosing $[X, Y] = XY - YX$, if multiplication of \mathcal{L} is associative. With this construction, \mathcal{I} has same elements with \mathcal{L} .

^{fe}It is an associative and non-commutative algebra, e.g. $(a_{ij}b_{jk})c_{kl} = a_{ij}(b_{jk}c_{kl})$ and $a_{ij}b_{jk} \neq b_{ij}a_{jk}$ for matrices a, b, c .

^{ff}Note that element-wise multiplication becomes $e_{ij}e_{km} = \delta_{jk}e_{im}$ in this basis, hence the structure constants are $c_{ijkm}^{rs} = \delta_{ir}\delta_{jk}\delta_{ms}$.

^{fg}An inversion in a permutation (i_1, \dots, i_n) of the numbers $(1, \dots, n)$ is a pair (i_j, i_k) such that $j < k$ and $i_j > i_k$. If the number of inversions is even (odd), the permutation is said to be even (odd).

^{fh}Matrices with nonzero determinants are called non-singular.

^{fi}In general, the space $\mathfrak{M}(n, \kappa)$ is an associative linear algebra: $n \times n$ matrices form a commutative group under addition and they have an associative multiplication among themselves. For non-singular matrices, the elements form a group under multiplication as well, hence $\text{GL}(n, \kappa)$.

^{fj}It is a four dimensional linear space κ^4 with an associative and linear multiplication.

^{fk}One usually uses the shorthand notations $\mathbf{H} \equiv \mathbf{H}(\mathbb{R})$ and $\alpha \mathbf{e}_0 + \beta \mathbf{e}_1 + \gamma \mathbf{e}_2 + \delta \mathbf{e}_3 = \alpha + \beta \mathbf{i} + \gamma \mathbf{j} + \delta \mathbf{k}$

- A concrete example of above explanation: the set $\mathfrak{M}(n, \kappa)$ is a Lie algebra over κ . The commutation relations for its basis matrices e_{ij} and the structure constants are given as

$$[e_{ij}, e_{km}] = \delta_{jk}e_{im} - \delta_{mi}e_{kj}, \quad c_{ij,km}^{rs} = \delta_{ir}\delta_{jk}\delta_{ms} - \delta_{kr}\delta_{im}\delta_{js} \quad (\text{A.8})$$

^fSee section A.2.

- The set of vectors in a three-dimensional space is a Lie algebra under vector multiplication.

A.2 Subgroups and subalgebras

Subgroup

Basics

- H is subgroup of G if
 - H is a nonempty set
 - $g_1g_2^{-1} \in H$ for all $g_1, g_2 \in H$
- H inherits the group operation of G .
- A subgroup $N \subset G$ is said to be *invariant* (or normal) if $gN = Ng$ for any $g \in G$. Here, gN and Ng denote cosets.^f Another way to look at this is as follows: $gNg^{-1} = N$, hence a general conjugation of an element of N still leads to another element of N if N is a normal subgroup.
- Intersections of two subgroups of G is also a subgroup of G .
- To any subgroup $H, K \subset G$, there is their *commutator subgroup*, denoted by $[H, K]$, and generated by $hh^{-1}k^{-1}$ for $h \in H$ and $k \in K$. The subgroup $[G, G]$ is called *commutator subgroup* of G .
- The intersection of all subgroups containing a given set $A \subset G$ is a subgroup of G , and is called the *group generated by A*.

Cosets

- For the subgroup $H \subset G$ and the element $g_1 \in G$, the sets g_1H (Hg_1) are left (right) cosets.
- Left (right) cosets are either coincide or mutually distinct. Hence G can be decompose into different g_1H or Hg_1 .
- The set of left cosets is denoted as G/H and is called *left coset space* or *left quotient space*.
- The set of right cosets is denoted as $H\backslash G$ and is called *right coset space* or *right quotient space*.
- For the subgroups $H, K \subset G$, HgK is said to be a (H, K) two-sided coset. Set of these cosets and is denoted as $H\backslash G/K$

Centralizer & Normalizer

^{fm}See <http://mathworld.wolfram.com/GroupKernel.html>.

- The collection $Z(A)$ is said to be the centralizer of A if
 - A is a set of elements of a group G .
 - $Z(A)$ is the collection of all elements of G which commute with every element of A .
 - Centralizer of $A \subset G$, $Z(A)$, is a subgroup of G .
 - Centralizer of $A \subset G$, $Z(A)$, coincides with the centralizer of the subgroup generated by A .
 - The centralizer of a group G is said to be its *center*, and denoted by $Z \equiv Z(G)$.
- If H is a subgroup of G , then the set of elements of H , which commute with every element of A , is said to be the *centralizer* of A in H .
- The *normalizer* of a subgroup $H' \subset G$ is collection of $g \in G$ such that $gH' = H'g$. In other words, the normalizer of a subgroup $H' \subset G$ is the maximal subgroup of G in which H' is an invariant subgroup.
- If $H_{1,2} \subset G$ are two subgroups, one can define the normalizer of H_1 in H_2 .

Examples

- The additive groups \mathbf{R} of real numbers and \mathbf{Z} of integers are subgroups of the additive group \mathbf{C} of complex numbers.
- The multiplicative groups \mathbf{R}_+ of positive numbers and \mathbf{T} of numbers with absolute values equal to 1 are subgroups of the multiplicative group $\mathbf{C} \setminus \{0\}$ of complex numbers.
- The collection of matrices $g \in \mathrm{GL}(n, \kappa)$ with $\det g = 1$ forms an invariant subgroup of $\mathrm{GL}(n, \kappa)$, denoted as $\mathrm{SL}(n, \kappa)$:
 - **The proof:** Let \mathbb{R}^* be the multiplicative group of nonzero real numbers. Let

$$\phi : \mathrm{GL}(n, \mathbb{R}) \rightarrow \mathbb{R}^* \quad (\text{A.9})$$

be the map given by $\phi(X) = \det X$ for each $X \in \mathrm{GL}(n, \mathbb{R})$. This map is well defined as $\det X \neq 0$ for $X \in \mathrm{GL}(n, \mathbb{R})$. By the property of determinant, we know that $\det XY = \det X \det Y$ hence ϕ is a group homomorphism. The kernel of ϕ is

$$\ker(\phi) = \{X \in \mathrm{GL}(n, \mathbb{R}) | \phi(X) = \det X = 1\} = \mathrm{SL}(n, \mathbb{R}) \quad (\text{A.10})$$

and since kernels of group homomorphism's are always normal subgroups, we conclude $\mathrm{SL}(n, \mathbb{R})$ is a normal subgroup of $\mathrm{GL}(n, \mathbb{R})$.^{fm}

- Matrices g with $\det g = 1$ are called *unimodular*.
- The collection of parallel shifts of the plane is a subgroup of the group of the Euclidean motions of the plane.
- The scalar matrices, λI_n with $\lambda \in \kappa$, consist the center of the group $\mathrm{GL}(n, \kappa)$.
- The collection $\mathrm{SL}(n, \kappa)$ of unimodular matrices is the commutator subgroup of $\mathrm{GL}(n, \kappa)$.
- The commutator subgroup of the group $S_{\pm}(n, \kappa)$ of upper (lower) triangular $x \times n$ matrices coincides with the subgroup $N_{\pm}(n, \kappa)$ of special triangular matrices, i.e. matrices with 1's on the main diagonal.
- Let $N_+^{(k)}(n, \kappa)$ denote the subgroup of $N_+(n, \kappa)$, where it consists of upper triangular matrices $n = (n_{ij})$ such that $n_{ij} = 0$ for $0 < j - i \leq k$. Then, the commutator subgroup of $N_+(n, \kappa)$ and $N_+^{(k)}(n, \kappa)$ coincides with $N_+^{(k+1)}(n, \kappa)$.
- The center of the group $N_+(n, \kappa)$ consists of the matrices of the form $I_n + \lambda e_{1n}$, where $\lambda \in \kappa$.

Subring, subfield, linear subspace, linear subalgebra

Basics

- Subring, subfield, linear subspace, linear subalgebra are defined analogous to the subgroup.
- **Ideal:** A subring J of a ring R is said to be left (right) ideal of R if $rJ \subset J$ ($Jr \subset J$) for any $r \in R$.
- A subring which is simultaneously both left and right ideal is said to be an *ideal* (or two-sided ideal).
- One can analogously define left, right, and two-sided ideals of a linear algebra in the same way.
- In analogy with subgroups, we can define *subring*, *ideal*, *linear subspace*, etc. generated by A where A is a given set such that $A \subset S$ for S being the ring, linear space, etc. (whatever the case is).
- For linear subspaces \mathcal{L}_1 and \mathcal{L}_2 with $x \in \mathcal{L}_1$ and $y \in \mathcal{L}_2$, the *linear subspace generated by the linear subspaces \mathcal{L}_1 and \mathcal{L}_2* consists of all the elements $x + y$. It is called the *sum* of \mathcal{L}_1 and \mathcal{L}_2 .
- In Lie algebras, every left ideal is also a right ideal, and vice versa. So all Lie algebra ideals are two-sided ideals.
- In analogy with subgroups, we can define *normalizer*, *centralizer*, *commutator subalgebra* etc:

- The *centralizer* of a subset A of a ring R is the collection $Z(A)$ of elements of R which commute with every element of A .
- The *normalizer* of a subring S is the collection $N(S)$ of elements r such that $Sr \subset S$ and $rS \subset S$. It is the largest subring of R in which S is an ideal.
- If \mathcal{I} is a Lie algebra, then the centralizer of a subset A consists of all elements z such that $[z, a] = 0$ for all $a \in A$. The centralizer of a Lie algebra \mathcal{I} is said to be its center. The Lie subalgebra $[\mathcal{I}_1, \mathcal{I}_2]$ generated by all elements of the form $[x, y]$, where $x \in \mathcal{I}_1$ and $y \in \mathcal{I}_2$, is said to be the *commutator subalgebra* of Lie subalgebras \mathcal{I}_1 and \mathcal{I}_2 .

Examples

- \mathcal{L}_1 is a linear subspace of \mathcal{L} if
 - \mathcal{L} is a linear space over a field κ
 - \mathcal{L}_1 is also a linear space over the field κ
 - $\lambda x + \mu y \in \mathcal{L}_1$ for $x, y \in \mathcal{L}_1$ and $\lambda, \mu \in \kappa$
- A linear subspace \mathcal{L}_1 of a linear algebra \mathcal{L} is said to be a subalgebra in \mathcal{L} if the product of any elements of \mathcal{L}_1 is again an element of \mathcal{L}_1 .
- The field \mathbb{R} is a subfield of \mathbb{C} .
- The set $n\mathbb{Z}$ is an ideal of the ring \mathbb{Z} .
- The collection of elements of the form $\lambda e_0 + 0e_1 + 0e_2 + 0e_3$ is the center of the skew field $\mathbf{H}(\kappa)$ of quaternions.
- The Lie algebra $\mathfrak{n}_+(n, \kappa)$ of upper triangular matrices with 0's on the main diagonal is the commutator subalgebra of the Lie algebra $\mathfrak{s}_+(n, \kappa)$ of upper triangular matrices with commutation operation $[a, b] = ab - ba$.

A.3 Homomorphisms and automorphisms

In groups

Basics

- A mapping f of a group G_1 into a group G_2 is said to be a *homomorphism* if for any $g, h \in G_1$ we have $f(gh) = f(g)f(h)$.
- A one-to-one homomorphism of G_1 onto G_2 is called an *isomorphism*: we say that groups G_1 and G_2 are isomorphic.
- An isomorphism of a group G onto itself is called an *automorphism* of G .

- The image $f(H_1)$ of a subgroup $H_1 \subset G_1$ under a homomorphic mapping $f : G_1 \rightarrow G_2$ is a subgroup of G_2 , and the full inverse image $f^{-1}(H_2)$ of a subgroup $H_2 \subset G_2$ is a subgroup of G_1 .
- $N = f^{-1}(e')$, where e' is the identity element of G_2 , is an invariant subgroup of G_1 , called a *kernel* of the homomorphism f .
- The subgroup $f(G_1) \subset G_2$ is isomorphic to the quotient group G_1/N , i.e. to the coset space G_1/N with multiplication of cosets defined by the equality $g_1N * g_2N = g_1g_2N$.
- Inner automorphism:
 - The mapping $g \rightarrow g_0gg_0^{-1}$ is an automorphism of G . These automorphisms are called *inner*.
 - The element of G , which is the image of an element g under an inner automorphism, is called *conjugate* with g .
 - The group G decomposes into classes of mutually conjugate elements, and one of these classes is the set $\{e\}$.

Representation of a group

- A homomorphic mapping of a group G into the group of transformations of a set X is called a *representation* of the group by transformations of X . This mapping defines the *action* of the group G on X .
- The image of $x \in X$ under the transformation corresponding to $g \in G$ is denoted by $g \circ x$. So, for all $g_1, g_2 \in G$ and for all $x \in X$ one has the equality $(g_1g_2) \circ x = g_1 \circ (g_2 \circ x)$.
- The set X , on which the action of G is defined, is called the *G -set* or the *G -space*.
- Actions of a group G on sets X and Y are said to be *equivalent* if there exists a one-to-one correspondence $f : X \rightarrow Y$ such that $g \circ (f(x)) = f(g \circ x)$ for all $x \in X$, $g \in G$.
- The set N of elements of G , to which there corresponds the identity transformation in X , is an invariant subgroup of G , called the *kernel of ineffectiveness* of the action of G on X . If $N = \{e\}$, where e is the identity element of G , then G is said to act effectively on X .
- Let us denote by Y^X the space of mappings of a set X into a set Y .
 - If X is a G -set, then the action of G in Y^X is defined by the equality

$$(g \circ f)(x) = f(g^{-1} \circ x) \equiv f_g(x)$$

- If Y is a G -set, then the action of G in Y^X is given by the equality

$$(g \circ f)(x) = g \circ f(x)$$

Note that the equation in the first case is well-defined:

$$(g_1 g_2 \circ f)(x) = f((g_1 g_2)^{-1} \circ x) = f(g_2^{-1} g_1^{-1} \circ x) = f_{g_2}(g_1^{-1} \circ x) = (g_1 \circ f_{g_2})(x) = (g_1 \circ (g_2 \circ f))(x)$$

Examples

- Let \mathbf{R} be the additive group of real numbers, and \mathbf{R}^+ be the multiplicative group of all positive real numbers. Since $a^{x+t} = a^x a^t$, then for given $a \neq 1$, $a > 0$, the mapping $x \rightarrow a^x$ provides an isomorphism between \mathbf{R} and \mathbf{R}^+ .
- The mapping $x \rightarrow e^{2\pi i x}$ of the additive group \mathbf{R} onto the multiplicative group \mathbf{T} is homomorphic. Its kernel coincides with the additive group \mathbf{Z} of integers.
- The quotient group $\mathbf{Z}/n\mathbf{Z}$ is isomorphic to \mathbf{Z}_n .
- The mapping $g \rightarrow \det g$ is a homomorphism of $\mathrm{GL}(n, \kappa)$ onto $\kappa \setminus \{0\}$.
- The matrix $b = (b_{ij})$ with $b_{ij} = a_{ji}$ is said to be the transpose of the matrix $a = (a_{ij})$, it is denoted by a^t . It is evident that $(\lambda a + \mu b)^t = \lambda a^t + \mu b^t$, $(a^{-1})^t = (a^t)^{-1}$, and $(ab)^t = b^t a^t$. The mapping $a \rightarrow (a^{-1})^t$ is an automorphism of the group $\mathrm{GL}(n, \kappa)$.
- The left shifts $g \rightarrow g_0 g$ and the right shifts $g \rightarrow gg_0^{-1}$ define equivalent effective actions of G on G . The equivalence is given by the mapping $g \rightarrow g^{-1}$.
- The transformations $g \rightarrow g_0 g g_0^{-1}$ define the action of G on G with the kernel of ineffectiveness coinciding with the center Z of G .
- Any matrix from $\mathrm{GL}(n, \mathbb{C})$ with simple eigenvalues is conjugate to a diagonal matrix.

In rings, fields, linear spaces, and algebras

Basics

- Homomorphisms of linear spaces are called *linear mappings*. In other words, a mapping $A : \mathfrak{L}_1 \rightarrow \mathfrak{L}_2$ of a linear space \mathfrak{L}_1 over a field κ into a linear space \mathfrak{L}_2 over the same field is linear if for any $\lambda, \mu \in \kappa$ and for any $x, y \in \mathfrak{L}_1$ one has the equality $A(\lambda x + \mu y) = \lambda Ax + \mu Ay$.
- To every subspace $\mathfrak{M}_1 \subset \mathfrak{L}_1$ there corresponds the set of linear mappings $A : \mathfrak{L}_1 \rightarrow \mathfrak{L}_2$ such that $A(\mathfrak{M}_1) = 0$, and to every subspace $\mathfrak{M}_2 \subset \mathfrak{L}_2$ there corresponds the set of linear mappings $A : \mathfrak{L}_1 \rightarrow \mathfrak{L}_2$ such that $A(\mathfrak{L}_1) \subset \mathfrak{M}_2$

- The complete inverse image $A^{-1}(0)$ of the null vector from \mathfrak{L}_2 is called the *kernel* of the linear mapping $A : \mathfrak{L}_1 \rightarrow \mathfrak{L}_2$. It is a linear subspace in \mathfrak{L}_1 . The image $A(\mathfrak{L}_1)$ of the space \mathfrak{L}_1 under this mapping is a linear subspace in \mathfrak{L}_2 .
- The linear mapping $A : \mathfrak{L}_1 \rightarrow \mathfrak{L}_2$ for which $A^{-1}(0) = 0$ and $A(\mathfrak{L}_1) = \mathfrak{L}_2$ is invertible, i.e. there is the linear mapping $A^{-1} : \mathfrak{L}_2 \rightarrow \mathfrak{L}_1$ such that $A^{-1}A = E_1$ (E_k is the identity mapping in \mathfrak{L}_k , $k = 1, 2$). In this case we also have $AA^{-1} = E_2$.
- If linear mappings A and B of a space \mathfrak{L} into itself are invertible, then the composition AB is also invertible and $(AB)^{-1} = B^{-1}A^{-1}$.
- The set of invertible linear mappings of an n -dimensional linear space \mathfrak{L} over a field κ is a group which is isomorphic to the group $\mathrm{GL}(n, \kappa)$ of non-singular matrices of order n over the field κ , and so it is also denoted by $\mathrm{GL}(n, \kappa)$.
- The set of linear mappings of \mathfrak{L}_1 into \mathfrak{L}_2 is a linear space over κ : $(\lambda A + \mu B)\mathbf{x} = \lambda A\mathbf{x} + \mu B\mathbf{x}$, $\mathbf{x} \in \mathfrak{L}_1$. In particular, the space of linear mappings of a linear space \mathfrak{L} into κ (i.e. the space of functions f on \mathfrak{L} with values in κ satisfying the relation $f(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda f(\mathbf{x}) + \mu f(\mathbf{y})$) is linear. These functions are called *linear functionals* or *linear forms* on \mathfrak{L} .
- The linear space consisting of linear functionals on \mathfrak{L} is called *conjugate* (or *dual*) to \mathfrak{L} and is denoted by \mathfrak{L}' .
- To every linear operator $A : \mathfrak{L} \rightarrow \mathfrak{M}$ there corresponds the conjugate linear operator $A' : \mathfrak{M}' \rightarrow \mathfrak{L}'$, defined by the equality $(A'f)(\mathbf{x}) = f(A\mathbf{x})$, $\mathbf{x} \in \mathfrak{L}$, $f \in \mathfrak{M}'$.
- A finite dimensional linear space \mathfrak{L} is isomorphic to its conjugate space \mathfrak{L}' . The following equalities hold:

$$(\lambda A + \mu B)' = \lambda A' + \mu B', \quad (A^{-1})' = (A')^{-1}, \quad (AB)' = B'A'$$

B

A brief review of distributions

In this appendix, we briefly review the concept of distributions and present some of their properties.^{[fn](#)}

B.1 Basics

Let Ω be an open subset of \mathbb{R}^n . We then denote by $\mathcal{D}(\Omega)$ the space of functions that vanish outside the bounded subset of Ω and whose all order partial derivatives are continuous.^{[fo](#)} Such objects are also called *test functions*;^{[fp](#)} if $\phi(x)$ is such a function and if $\int f(x)\phi(x)dx$ exists, then we call $f(x)$ a *generalized function*. These integrals can also be viewed as an operation of a *linear functional* f on the test function ϕ and denoted as $\langle f, \phi \rangle$. We denote the space of all continuous linear functionals on $\mathcal{D}(\Omega)$ by $\mathcal{D}'(\Omega)$ and call these functionals *distributions*.^{[fq](#)}

We constructed a functional f from the function $f(x)$ by identifying $\langle f, \phi \rangle$ and $\int f(x)\phi(x)dx$ but not all functionals in $\mathcal{D}'(\Omega)$ can be constructed this way.^{[fr](#)} Furthermore, even if a functional is constructed this way, the properties of the function $f(x)$ do not carry over to the distribution f unless the function is *locally integrable*, i.e. unless $\int_{\Omega} f(x)\phi(x)dx$ is absolutely convergent for all $\phi \in \mathcal{D}(\Omega)$.^{[fs](#)} However, we actually do have another method to construct any distribution out of test functions:

Given any distribution $f \in \mathcal{D}'(\Omega)$, there exists a sequence $\{\phi_n\}$ of test functions such that $\phi_n \rightarrow f$ as distributions.

This definition of a distribution allows us to carry over usual operations on functions to distributions. Indeed, we can define $T \cdot f$ as $T \cdot f = \lim T \cdot \phi_n$ if $f = \lim_{n \rightarrow \infty} \phi_n$.^{[ft](#)}

The theorem above is quite useful to represent distributions in terms of test functions, however the test functions in $\mathcal{D}(\Omega)$ can be limited as they have to have bounded support. Indeed, a common function in Physics is the Gaussian $f(x) = e^{-|x|^2}$, which checks out to be a C^∞ function but fails to have a compact support (is nonzero for all $x \in \mathbb{R}^n$). Such functions belong to the *Schwartz class* $\mathcal{S}(\mathbb{R}^n)$, which is the algebra of smooth functions vanishing at infinity faster than any power of $|x|^{-1}$, together with all derivatives. These functions are particularly important for Fourier transform because if $f \in \mathcal{S}(\mathbb{R}^n)$ and g is the Fourier transform of f , then $g \in \mathcal{S}(\mathbb{R}^n)$ as well.

We denote the space of all continuous linear functionals on $\mathcal{S}(\mathbb{R}^n)$ by $\mathcal{S}'(\mathbb{R}^n)$ and call these functionals *tempered distributions*. As $\mathcal{D}(\mathbb{R}^n) \subset \mathcal{S}(\mathbb{R}^n)$, and as any f in $\mathcal{S}'(\mathbb{R}^n)$ is a functional $\langle f, \phi \rangle$ defined for all $\phi \in \mathcal{S}(\mathbb{R}^n)$, tempered distributions form a subclass of distributions. In particular, Fourier transform of a tempered distribution is also a tempered distribution.

The Dirac- δ distribution is a tempered distribution and hence can be constructed as a limiting sequence of Schwartz functions; indeed,

^{[fn](#)}For a rather pedagogical introduction, we refer the reader to [75].

^{[fp](#)}In other words, $\mathcal{D}(\Omega)$ is the algebra of C^∞ -functions with compact support on Ω .

^{[fq](#)}For example,

$$\phi(x) = \begin{cases} \exp(-x^{-2} - (1-x)^{-2}) & 1 > x > 0 \\ 0 & \text{otherwise} \end{cases} \quad (\text{B.1})$$

is a such a test function on $\Omega = \mathbb{R}$.

^{[fq](#)}As an example, the function $f(x) = |x|^{-n+k}$ for $k > 0$ gives rise to a distribution in $\mathcal{D}'(\Omega)$ for $\Omega \in \mathbb{R}^n$ by setting

$$\langle f, \phi \rangle = \int_{\Omega} \phi(x) |x|^{-n+k} d^n x \quad (\text{B.2})$$

because the integration is convergent for all test functions.

^{[fr](#)}For instance, there does not exist a mathematically rigorous Dirac-Delta function $\delta(x)$ to construct the Dirac-Delta distribution δ ; what physicists denote as $\delta(x)$ is a somewhat sloppy reference to the actual distribution δ . As we will see below, the correct way to define $\delta(x)$ is to take a limit of a sequence of test functions $\delta_n(x)$.

^{[fs](#)}For example, $f(x) = 1/|x|$ on \mathbb{R} is not locally integrable, hence its properties do not necessarily carry over to the distribution f defined as $\langle f, \phi \rangle = \int \phi(x)/|x| dx$. Indeed, even though $f(x)$ is a non-negative function, f is not a non-negative distribution, i.e. $\langle f, \phi \rangle$ can be negative even if the test function satisfies $\phi(x) \geq 0$.

^{[ft](#)}This requires the limit $\lim_{n \rightarrow \infty} T \cdot \phi_n$ to be independent of the choice of the sequence $\{\phi_n\}$ approximating f , which is true for most interesting operations such as translation or differentiation.

we can define it as

$$\langle \delta, \phi \rangle = \lim_{\epsilon \rightarrow 0} \frac{1}{\sqrt{\pi}|\epsilon|} \int_{-\infty}^{\infty} e^{-(x/\epsilon)^2} \phi(x) dx \quad (\text{B.3})$$

where $\delta_\epsilon(x) \equiv \frac{1}{\sqrt{\pi}|\epsilon|} e^{-(x/\epsilon)^2}$ is the Schwartz function whose sequence limits to the δ distribution. As an abusive notation, it is common to write $\delta(x)$ and call it a δ -function, despite such a function does not really exist. Indeed, it should be understood that $\delta(x)$ is only a shorthand notation such that

$$\int \delta(x) \phi(x) dx \equiv \lim_{\epsilon \rightarrow 0} \int \delta_\epsilon(x) \phi(x) dx \quad (\text{B.4})$$

It can be explicitly checked through its definition that Dirac- δ distribution has the action $\langle \delta, \phi \rangle = \phi(0)$. By using the adjoint identities we can derive related actions, e.g. $\langle \delta', \phi \rangle = -\phi'(0)$.^{fu} Likewise, by using its definition as a limit of $\delta_\epsilon(x)$, one can work out the action of tensor products of δ -distributions (i.e. $\delta(x)\delta(y)$), though ordinary products (i.e. $\delta^2(x)$) are not really defined in $\mathcal{D}'(\mathbb{R}^n)$.^{fv}

A related distribution to the Dirac- δ is $\delta_\pm \in \mathcal{D}'(\mathbb{R})$, defined as

$$\langle \delta_\pm, \phi \rangle = \lim_{\epsilon \rightarrow 0} \int \frac{1}{x \pm i\epsilon} \phi(x) dx \quad (\text{B.5})$$

which is related to δ -distribution by the relation

$$\lim_{\epsilon \rightarrow 0} \int \frac{1}{x \pm i\epsilon} \phi(x) dx = \int \left(\mathcal{P} \frac{1}{x} \right) \phi(x) dx \mp i\pi \lim_{\epsilon \rightarrow 0} \int \delta_\epsilon(x) \phi(x) dx \quad (\text{B.6})$$

where the *principle value* of $1/x$ can be defined as $\mathcal{P} \frac{1}{x} = \partial_x \log |x|$. With an abuse of notation, we can simply write

$$\delta_\pm(x) = \mathcal{P} \frac{1}{x} \mp i\pi \delta(x) \quad (\text{B.7})$$

as well. However, we should be careful while writing distributions as if they are functions as in (B.7): $\mathcal{D}'(\mathbb{R})$ cannot be imbedded in an associative differential algebra meaning that not only products of distributions can be ill-defined (see footnote fv) but also products of one distribution with ordinary function needs to be treated with care as well!^{fw}

B.2 More on intrinsic products of distributions

As we discussed above, it is not possible to define the multiplication of distributions while preserving generality; nevertheless, there are ways to do this for specific tempered distributions. An important category is when the distribution can be represented as the boundary value of a holomorphic function on $\mathbb{C} \setminus \mathbb{R}$ with support in the upper half plane: in this case, one just multiplies the holomorphic representatives and takes

^{fu} Adjoint identities in distribution theory are relations of the form $\int T\psi(x)\phi(x)dx = \int \psi(x)S\phi(x)dx$, where one tries to find the operation S for a given operation T . In the case of derivatives, it is straightforward to find that $S = -\partial$ for $T = \partial$ via integration by parts, leading to $\langle \delta', \phi \rangle = -\phi'(0)$.

^{fv} The limit of $[\delta_\epsilon(x)]^2$ does not converge in $\mathcal{D}'(\mathbb{R})$ for real-valued $\delta_\epsilon(x)$. For complex-valued sequences, the expression does converge however contains ambiguities; for instance, one can show that the action of the distributions δ^2 and $c_1\delta + c_2\delta'$ can be made equivalent for arbitrary complex coefficients c_i . This rather indicates that we should not attempt to define δ^2 as an element of the space of distributions [76].

^{fw} For instance, the product can be non-commutative. As an example, note that the relations

$$\mathcal{R} = \{\delta(x) \cdot 1 = \delta(x), \delta(x) \cdot x = 0, x \cdot \mathcal{P} \frac{1}{x} = 1\} \quad (\text{B.8})$$

indicate that

$$(\delta(x) \cdot x) \cdot \mathcal{P} \frac{1}{x} \neq \delta(x) \cdot \left(x \cdot \mathcal{P} \frac{1}{x} \right) \quad (\text{B.9})$$

Ultimately, \mathcal{R} needs to change if one imbeds $\mathcal{D}'(\mathbb{R})$ to an associative algebra \mathcal{A} ; in fact, in such an imbedding, either the derivatives or the product of continuous functions cannot be preserved due to the famous *impossibility result* of L. Schwartz [76].

the boundary value of the product. As an example, consider $\delta_+(x)$: this can be viewed as the boundary value of the function $\hat{\delta}_+(z)$ defined as

$$\hat{\delta}_+(z) = \begin{cases} 1/z & \text{Im } z > 0 \\ 0 & \text{Im } z < 0 \end{cases} \quad (\text{B.10})$$

Thus, we can write $\delta_+(x)\delta_+(x)$ as the boundary value of $\hat{\delta}_+(z)\hat{\delta}_+(z)$:

$$[\delta_+(x)]^2 = \lim_{\epsilon \rightarrow 0} \frac{1}{(x + i\epsilon)^2} \quad (\text{B.11})$$

where we realize that this indicates $[\delta_+(x)]^2 = -\partial_x \delta_+(x)$, meaning that its action can be written as $\langle \delta_+^2, \phi \rangle = \langle \delta_+, \phi' \rangle$.

This category of distributions is actually quite important for relativistic theories such as Lorentzian quantum field theory. In the rigorous way of deriving things, one starts in the Euclidean region (with operators taking functional values) and then does the Wick rotation, effectively moving the complex time x^0 from upper half-plane (it) to the real-line (t). This operation moves the operators to the boundary of their holomorphic domain, turning them into distributions. In particular, all Weightman functions become tempered distributions, as required by his temperedness axiom.^{fx}

This approach has been generalized to turn the span of some distributions into an associative commutative differential algebra. Importantly, define the distribution

$$\mathcal{P}x^{-m} \equiv \frac{(-1)^{m-1}}{(m-1)!} \frac{d^m}{dx^m} \log|x| \quad (\text{B.12})$$

which is the generalization of the principle value of $1/x$ to higher powers $-m \in \mathbb{N}$. It has been shown in [81] that we can consistently define an algebra with the multiplication table

$$\begin{aligned} \mathcal{P}x^{-m} \cdot \mathcal{P}x^{-n} &= \mathcal{P}x^{-m-n} \\ \delta^{(m)}(x) \cdot \delta^{(n)}(x) &= 0 \\ \mathcal{P}x^{-m} \cdot \delta^{(n-1)}(x) &= (-1)^m \frac{n!}{(m+n)!} \delta^{(m+n-1)}(x) \end{aligned} \quad (\text{B.13})$$

where $\delta^{(m)}(x)$ represents the m -th derivative of delta distribution. We can check the consistency of this multiplication with (B.11):

$$\begin{aligned} [\delta_+(x)]^2 &= \mathcal{P}x^{-1} \cdot \mathcal{P}x^{-1} - 2i\pi \mathcal{P}x^{-1} \cdot \delta(x) + (-i\pi)^2 \delta(x) \cdot \delta(x) \\ &= \mathcal{P}x^{-2} + i\pi \delta^{(1)}(x) + 0 \\ &= -\partial_x (\mathcal{P}x^{-1} - i\pi \delta(x)) \\ &= -\partial_x \delta_+(x) \end{aligned} \quad (\text{B.14})$$

which is indeed what we found by extending $\delta_+(x)$ to $\hat{\delta}_+(z)$.

^{fx} A precise discussion of this is given in [77] and [78]. A heuristic discussion of the argument can be found in [79], see also the relatively recent review [80] for a nice introduction.

B.3 Subtleties with tensor products of distributions

Compared to intrinsic products of distributions such as $\delta(x)^2$, tensor products of distributions such as $\delta(x)\delta(y)$ are milder: we can define them in various ways, such as *Fourier products*, *strict products*, and *model products* [76]. Nevertheless, they can also behave unexpectedly; indeed, we will reproduce below the discussion in §18 of the latter reference, where it is shown that direct use of such products with differential equations can give incorrect results.

We will illustrate this with a concrete example. Consider the system of partial differential equations:^{f_y}

$$\begin{aligned} (\partial_t + \partial_x)u_1 &= u_1 u_2 \\ (\partial_t - \partial_x)u_2 &= -u_1 u_2 \\ u_j(x, 0) &= a_j(x), \quad j = 1, 2 \end{aligned} \tag{B.15}$$

Given arbitrarily large & nonnegative initial data, the existence and uniqueness of a generalized solution to this system in the Colombeau algebra $G(\mathbb{R}^2)$ has been established. For $a_1(x) = \alpha_1\delta(x+1)$ and $a_2(x) = \alpha_2\delta(x-1)$, the solutions u_i converge to the distributions v_i that read as

$$\begin{aligned} v_1 &= \alpha_1\delta_1 + (\alpha_2 - \beta)\theta_1, \\ v_2 &= \alpha_2\delta_2 - (\alpha_2 - \beta)\theta_2 \end{aligned} \tag{B.16a}$$

with

$$\beta = -2 \log \left(1 - \exp\left(-\frac{\alpha_1}{2}\right) + \exp\left(-\frac{\alpha_1 + \alpha_2}{2}\right) \right) \tag{B.16b}$$

where we use the shorthand notation

$$\begin{aligned} \delta_1(x, t) &= \delta(x-t+1), & \delta_2(x, t) &= \delta(x+t-1) \\ \theta_1(x, t) &= \delta(x-t+1) \otimes H(t-1), & \theta_2(x, t) &= \delta(x+t-1) \otimes H(t-1) \end{aligned} \tag{B.17}$$

for the Heaviside distribution H .

This solution has been obtained by *regularizing the products of distributions* and then taking the appropriate limit. We will show below that a naive tensor multiplication of distributions actually gives an incorrect result.

From a physical standpoint, it is obvious that two delta waves propagate along lightcone coordinates, interact at $t = 1$, and then scatter with different amplitudes. Indeed, we can see that this is the case from the actual solution in (B.16). So even if we did not have the solution, it would have made sense to propose an ansatz of the form

$$\begin{aligned} u_1 &= \alpha_1(\delta_1 + c_1\theta_1) \\ u_2 &= \alpha_2(\delta_2 + c_2\theta_2) \end{aligned} \tag{B.18}$$

for coefficients c_i to be determined. Once we insert these into the system of equations, we are forced to evaluate tensor products of delta

^{f_y}This is an example of a semilinear hyperbolic system (which are used to model advection and nonlinear interaction). This particular model is also called *the migrating predator-prey model*: it models two species moving along the x -axis with speeds ± 1 and interacting upon collision. The system is interesting as it describes strong interaction, i.e. its solution does not split into a regular and a singular piece.

distributions: but these are well defined,^{fz} thus (B.15) become

$$\begin{aligned}\alpha_1 c_1 \delta(x) \delta(t-1) &= \alpha_1 \alpha_2 \left(\frac{1}{2} + \frac{c_1 + c_2}{4} + \frac{c_1 c_2}{8} \right) \delta(x) \delta(t-1) \\ \alpha_2 c_2 \delta(x) \delta(t-1) &= -\alpha_1 \alpha_2 \left(\frac{1}{2} + \frac{c_1 + c_2}{4} + \frac{c_1 c_2}{8} \right) \delta(x) \delta(t-1)\end{aligned}\quad (\text{B.20})$$

This gives us a system of algebraic equations, and with the condition $c_i \geq -1$ and $\alpha_i \geq 0$ (so that u_i have nonnegative coefficients), we obtain the unique solution

$$\begin{aligned}v_1 &= \alpha_1 \delta_1 + (\alpha_2 - \beta') \theta_1, \\ v_2 &= \alpha_2 \delta_2 - (\alpha_2 - \beta') \theta_2\end{aligned}\quad (\text{B.21a})$$

with

$$\beta' = \alpha_1 + 4 - \sqrt{(\alpha_1 + 4)^2 + 2\alpha_2(\alpha_1 - 4) + \alpha_2^2} \quad (\text{B.21b})$$

Clearly this does not match the correct answer in (B.16)!^{ga}

This example illustrates why one needs to exercise care even for tensor products of distributions, in the context of differential equations at the very least. The root of the problem here is that terms such as $\delta(x)\delta(y)$ represent limits of sequences, and that the derivatives do not necessarily commute with this limit: in the correct approach, one solves the system of equations with sequences, and take the limit *only after that*. Nevertheless, it is depressing to observe that the incorrect approach has yielded a self-consistent result, without any indication for a problem along the way, and that the incorrect result matches the correct one in the first few orders of α_i .

^{fz} $\delta_1 \delta_2$, $\delta_i \theta_j$, and $\theta_1 \theta_2$ exist as Fourier product, as strict product, and as model product with rotationally symmetric mollifiers employed, respectively. They read as

$$\delta_1 \delta_2 = \frac{1}{2} \delta(x) \delta(t-1), \quad (\text{B.19a})$$

$$\delta_1 \theta_2 = \delta_2 \theta_1 = \frac{1}{4} \delta(x) \delta(t-1), \quad (\text{B.19b})$$

$$\theta_1 \theta_2 = \frac{1}{8} \delta(x) \delta(t-1) \quad (\text{B.19c})$$

^{ga}It is interesting to note that $\beta' \approx \beta + \mathcal{O}(\mu^3)$ if $\alpha_1 \sim \alpha_2 \sim \mu$.

Bibliography

- [1] Q. Wang and X. Zhao, *Scientific reports* **5** no. 1, (2015) 1–10.
- [2] L. Glasser, *Journal of chemical education* **81** no. 3, (2004) 414.
- [3] J. Bausch, T. S. Cubitt, and J. D. Watson, *Nature Communications* **12** no. 1, (2021) 1–8.
- [4] N. Goldenfeld,, *Lectures on phase transitions and the renormalization group*. 1992.
- [5] M. E. Fisher, *Physica* **28** no. 2, (1962) 172–180.
- [6] J. Levy, J. Zinn-Justin, M. Levy, and J.-C. Le Guillou,, *Phase Transitions Cargèse 1980*, vol. 72. Springer Science & Business Media, 2012.
- [7] D. Sullivan, G. Neilson, H. Fischer, and A. Rennie, *Journal of Physics: Condensed Matter* **12** no. 15, (2000) 3531.
- [8] A. Haupt and J. Straub, *Phys. Rev. E* **59** (Feb, 1999) 1795–1802.
<https://link.aps.org/doi/10.1103/PhysRevE.59.1795>.
- [9] P. Damay, F. m. c. Leclercq, R. Magli, F. Formisano, and P. Lindner, *Phys. Rev. B* **58** (Nov, 1998) 12038–12043.
<https://link.aps.org/doi/10.1103/PhysRevB.58.12038>.
- [10] S. Kuwabara, H. Aoyama, H. Sato, and K. Watanabe, *Journal of Chemical and Engineering Data* **40** no. 1, (1995) 112–116.
- [11] M. W. Pestak and M. H. W. Chan, *Phys. Rev. B* **30** (Jul, 1984) 274–288.
<https://link.aps.org/doi/10.1103/PhysRevB.30.274>.
- [12] A. Oleaga, A. Salazar, M. C. Hatnean, and G. Balakrishnan, *Physical Review B* **92** no. 2, (2015) 024409.
- [13] S. M. Chester, W. Landry, J. Liu, D. Poland, D. Simmons-Duffin, N. Su, and A. Vichi, *JHEP* **06** (2020) 142.
- [14] J. L. Cardy,, *Scaling and renormalization in statistical physics*. 1996.
- [15] K. G. Wilson and M. E. Fisher, *Phys. Rev. Lett.* **28** (1972) 240–243.
- [16] D. Simmons-Duffin,, “The Conformal Bootstrap,” in *Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings*, pp. 1–74. 2017. [arXiv:1602.07982 \[hep-th\]](https://arxiv.org/abs/1602.07982).
- [17] F. Kos, D. Poland, and D. Simmons-Duffin, *JHEP* **11** (2014) 109.
- [18] S. Istrail,, “Statistical mechanics, three-dimensionality and np-completeness: I. universality of intracatability for the partition function of the ising model across non-planar surfaces,” in *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, pp. 87–96. 2000.
- [19] M. Moshe and J. Zinn-Justin, *Phys. Rept.* **385** (2003) 69–228.
- [20] A. D. Sokal, A. C. D. van Enter, and R. Fernandez, *Phys. Rev. Lett.* **66** (1991) 3253–3256.
- [21] B. Nienhuis and M. Nauenberg, *Phys. Rev. Lett.* **35** (1975) 477–479.
- [22] G. I. Barenblatt, G. I. Barenblatt, and B. G. Isaakovich,, *Scaling, self-similarity, and intermediate asymptotics: dimensional analysis and intermediate asymptotics*. No. 14. Cambridge University Press, 1996.

- [23] A. Pelissetto and E. Vicari, *Phys. Rept.* **368** (2002) 549–727.
- [24] P. W. Anderson, *Science* **177** no. 4047, (1972) 393–396.
- [25] S. Pasterski, S.-H. Shao, and A. Strominger, *Phys. Rev. D* **96** no. 6, (2017) 065026.
- [26] S. Albayrak, C. Chowdhury, and S. Kharel, *Phys. Rev. D* **102** (2020) 126020.
- [27] V. Gorbenko, S. Rychkov, and B. Zan, *JHEP* **10** (2018) 108.
- [28] V. Gorbenko, S. Rychkov, and B. Zan, *SciPost Phys.* **5** no. 5, (2018) 050.
- [29] W. E. Caswell, *Phys. Rev. Lett.* **33** (1974) 244.
- [30] T. Banks and A. Zaks, *Nucl. Phys. B* **196** (1982) 189–204.
- [31] T. DeGrand, *Rev. Mod. Phys.* **88** (2016) 015001.
- [32] S. Gukov, *Nucl. Phys. B* **919** (2017) 583–638.
- [33] D. Karateev, P. Kravchuk, M. Serone, and A. Vichi, *JHEP* **06** (2019) 088.
- [34] Y. Aoki, T. Aoyama, M. Kurachi, T. Maskawa, K.-i. Nagai, H. Ohki, A. Shibata, K. Yamawaki, and T. Yamazaki, , *PoS LATTICE2011* (2011) 053.
- [35] J.-W. Lee, *Phys. Rev. D* **103** no. 7, (2021) 076006.
- [36] T. K. Nayak, *J. Phys. Conf. Ser.* **1602** no. 1, (2020) 012003.
- [37] J. M. Maldacena, *Adv. Theor. Math. Phys.* **2** (1998) 231–252.
- [38] S. S. Gubser, I. R. Klebanov, and A. M. Polyakov, *Phys. Lett. B* **428** (1998) 105–114.
- [39] E. Witten, *Adv. Theor. Math. Phys.* **2** (1998) 253–291.
- [40] J. Penedones,, “TASI lectures on AdS/CFT,” in *Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings*, pp. 75–136. 2017. [arXiv:1608.04948 \[hep-th\]](https://arxiv.org/abs/1608.04948).
- [41] H. Nastase, [arXiv:0712.0689 \[hep-th\]](https://arxiv.org/abs/0712.0689).
- [42] P. Hunt and S. Sarkar, *Mon. Not. Roy. Astron. Soc.* **401** (2010) 547.
- [43] J. Colin, R. Mohayaee, M. Rameez, and S. Sarkar, *Astron. Astrophys.* **631** (2019) L13.
- [44] D. Baumann, *PoS TASI2017* (2018) 009.
- [45] D. Baumann, D. Green, and T. Hartman, *JHEP* **12** (2019) 134.
- [46] N. Arkani-Hamed and J. Maldacena, [arXiv:1503.08043 \[hep-th\]](https://arxiv.org/abs/1503.08043).
- [47] N. Arkani-Hamed, D. Baumann, H. Lee, and G. L. Pimentel, *JHEP* **04** (2020) 105.
- [48] I. Mata, S. Raju, and S. Trivedi, *JHEP* **07** (2013) 015.
- [49] D. Baumann, C. Duaso Pueyo, A. Joyce, H. Lee, and G. L. Pimentel, *JHEP* **12** (2020) 204.
- [50] D. Baumann, C. Duaso Pueyo, and A. Joyce, *AAPPS Bull.* **30** no. 6, (2020) 2–9.
- [51] D. Baumann, C. Duaso Pueyo, A. Joyce, H. Lee, and G. L. Pimentel, *SciPost Phys.* **11** (2021) 071.

- [52] T. Asselmeyer-Maluga, ed., *At the Frontier of Spacetime*. Springer, Cham, 2016.
- [53] M. Lachièze-Rey,, *Theoretical and observational cosmology*, vol. 541. Springer Science & Business Media, 2012.
- [54] L. Cornalba, M. S. Costa, and J. Penedones, *JHEP* **03** (2010) 133.
- [55] S. Weinberg, *Phys. Rev. D* **82** (2010) 045031.
- [56] W. A. Hepner, *Nuovo Cim.* **26** (1962) 351–368.
- [57] G. Mack and I. T. Todorov, *Phys. Rev. D* **8** (1973) 1764–1787.
- [58] M. Luscher and G. Mack, *Commun. Math. Phys.* **41** (1975) 203–234.
- [59] M. Kologlu, P. Kravchuk, D. Simmons-Duffin, and A. Zhiboedov, *JHEP* **01** (2021) 128.
- [60] S. Ferrara, A. F. Grillo, and R. Gatto, *Annals Phys.* **76** (1973) 161–188.
- [61] P. Di Francesco, P. Mathieu, and D. Senechal,, *Conformal Field Theory*. Graduate Texts in Contemporary Physics. Springer-Verlag, New York, 1997.
- [62] P. A. M. Dirac, *Annals Math.* **37** (1936) 429–442.
- [63] V. K. Dobrev, G. Mack, V. B. Petkova, S. G. Petrova, and I. T. Todorov,, *Harmonic Analysis on the n-Dimensional Lorentz Group and Its Application to Conformal Quantum Field Theory*, vol. 63. 1977.
- [64] D. Simmons-Duffin, *JHEP* **04** (2014) 146.
- [65] M. S. Costa, J. Penedones, D. Poland, and S. Rychkov, *JHEP* **11** (2011) 071.
- [66] M. Schottenloher, ed., *A mathematical introduction to conformal field theory*, vol. 759. 2008.
- [67] S. Ferrara, P. Gatto, and A. F. Grilla, *Springer Tracts Mod. Phys.* **67** (1973) 1–64.
- [68] P. Kravchuk and D. Simmons-Duffin, *JHEP* **11** (2018) 102.
- [69] D. G. Boulware, L. S. Brown, and R. D. Peccei, *Phys. Rev. D* **2** (1970) 293–298.
- [70] G. Mack and A. Salam, *Annals Phys.* **53** (1969) 174–202.
- [71] H. A. Kastrup, *Annalen Phys.* **464** (1962) 388–428.
- [72] H. A. Kastrup, *Nucl. Phys.* **58** (1964) 561–579.
- [73] H. A. Kastrup, *Phys. Rev.* **150** (1966) 1183–1193.
- [74] N. Vilenkin and A. Klimyk,, *Representation of Lie Groups and Special Functions: Volume 1: Simplest Lie Groups, Special Functions and Integral Transforms*. Mathematics and its Applications. Springer Netherlands, 2012. https://books.google.com.tw/books?id=_g7vCAAAQBAJ.
- [75] R. Strichartz,, *A Guide to Distribution Theory and Fourier Transforms*. Studies in advanced mathematics. World Scientific, 2003. <https://books.google.nl/books?id=T7vEOGGDCh4C>.
- [76] M. Oberguggenberger,, *Multiplication of Distributions and Applications to Partial Differential Equations*. Pitman research notes in mathematics series. Longman Scientific & Technical, 1992. <https://books.google.nl/books?id=7zPvAAAAMAAJ>.

- [77] R. F. Streater and A. S. Wightman,, *PCT, spin and statistics, and all that.* 1989.
- [78] R. Jost,, *The general theory of quantized fields*, vol. 4. American Mathematical Society, Providence, RI, 1965.
- [79] R. Haag,, *Local quantum physics: Fields, particles, algebras.* 1992.
- [80] T. Hartman, S. Jain, and S. Kundu, *JHEP* **05** (2016) 099.
- [81] V. K. Ivanov, *Siberian Mathematical Journal* **20** no. 4, (Jul, 1979) 509–516.
<https://doi.org/10.1007/BF00970364>.