

**HOW TO MODEL MACROSCOPIC WORLDS:
TOWARDS THE PHILOSOPHY OF CONDENSED MATTER PHYSICS**

Sang Wook Yi

Department of Philosophy, Logic and Scientific Method
London School of Economics and Political Science

Thesis submitted for the Degree of Doctor of Philosophy
University of London
October 2000

UMI Number: U163447

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U163447

Published by ProQuest LLC 2014. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

THESES

F

7826



746253

Abstract

Condensed matter physics studies many-body phenomena, the phenomena involving a huge number of constituents interacting with each other strongly. My theme is modelling in condensed matter physics: the construction of mathematical/physical structures in order to understand many-body phenomena in the world. I study how condensed matter physicists learn about many-body phenomena from the successful employment of models. My proposal is to construe condensed matter physics as engaged essentially with the three activities: model-building, model-exploring and model-based understanding. General theories such as statistical mechanics guide the process of model-building as model-building methodology. I discuss the multiple layers of interactions among general theories (particularly among thermodynamics and statistical mechanics) and show that complementation and cooperation rather than reduction are better concepts for understanding their relation. In model-exploring stage, we probe a model in order to determine what exactly the model implies in itself and what it could with additional constraints. I investigate a number of epistemic roles of approximations in this stage. I also discuss what consists of model-based understanding. With the help of appropriate interpretative models, an well-understood model can provide us with our best representation of the phenomena (substantial theories). Finally, I investigate how physicists successfully deal with some crucial features of critical phenomena using renormalization group methods. I compare the renormalization group methods of condensed matter physics with those of quantum field theory, and argue. I claim that the mean field methods and renormalization group methods in condensed matter physics complement to each other.

To my parents

Table of Contents

<i>ACKNOWLEDGMENTS</i>	5
<i>INTRODUCTION</i>	
CONDENSED MATTER PHYSICS: MODEL-BUILDING, MODEL-EXPLORING AND MODEL-BASED UNDERSTANDING	6
CHAPTER 1	
GENERAL THEORIES AND SUBSTANTIAL THEORIES: MODEL-BUILDING METHODOLOGY AND MODEL-BASED UNDERSTANDING	9
CHAPTER 2	
THERMODYNAMICS AND STATISTICAL MECHANICS: ALTERNATIVE MODEL- BUILDING METHODOLOGIES OF CONDENSED MATTER PHYSICS	31
CHAPTER 3	
MODELS OF CONDENSED MATTER PHYSICS	47
CHAPTER 4	
APPROXIMATIONS IN CONDENSED MATTER PHYSICS	65
CHAPTER 5	
THE RENORMALIZATION GROUP METHOD IN CONDENSED MATTER PHYSICS	90
Figures	121
<i>BIBLIOGRAPHY</i>	125

Acknowledgments

There are lots of people I would like to thank, but I am afraid I shall name here only a very few of them. First of all, I wish to thank my supervisor, Professor Nancy Cartwright. It was reading of her book, *How the Laws of Physics Lie*, when I was busy in preparing for the entrance exam of the graduate school of Seoul National University in 1990, that prompted me to study philosophy of physics. For five years of my study at LSE, she has been always ready to help me through all the usual and unusual difficulties of my PhD study. I am extremely grateful for what she has taught me about philosophy and for what she has showed me about doing philosophy. Hasok Chang has always been a good friend to me, and a very good adviser on many issues. He is a great help to me intellectually as well as emotionally, and I really appreciate it. I have been really lucky to work for the *Measurement in Physics and Economics* research project housed at *the Centre for Philosophy of Natural and Social Science* at LSE. I learned a lot from my fellow researchers in the project, and greatly enjoyed the interdisciplinary research that helped me appreciate the importance of modelling both in natural and in social science. Craig Callendar, Jordi Cat, Hasok Chang, Arthur Fine, Stephan Hartmann, Carl Hofer, RIG Hughes, Fred Kronz, Chiung Liu, Michael Redhead, Paul Teller and Jan von Plato read some of earlier drafts of the chapters and gave me lots of illuminating comments. I wish to thank all of them. I have been extremely lucky during my education to have a number of really excellent teachers to whom I owe a lot. They are In Rae Cho, Moo Young Choi, Mi Kyung Kim, Young Sik Kim, Jan von Plato and Haw Ik Zhang. I want to pay my special tribute to them. Finally, I would like to thank my parents who have made every effort to give me all the things that were not allowed to them, my brothers who have always been honourable and good friends to me and my wife who has been the most wonderful thing to happen to my life.

INTRODUCTION

CONDENSED MATTER PHYSICS: MODEL-BUILDING, MODEL-EXPLORING AND MODEL-BASED UNDERSTANDING

My thesis is about condensed matter physics. Condensed matter physics (CMP) studies macroscopic objects and many-body phenomena, the phenomena involving a huge number of entities. A typical many-body phenomenon involves strong interactions among an enormous number of constituents. Many-body phenomena are hard to deal with by standard theoretical tools such as perturbation methods, which have been developed mainly for the study of few-body phenomena, such as the motion of planets. Consequently, condensed matter physicists have worked to devise a lot of methods and research strategies especially to overcome this difficulty and to gain an understanding of a number of fascinating many-body phenomena.¹

Important issues in contemporary philosophy of science are discussed largely within the context of few-body phenomena. For instance, the problem of reduction has been investigated in the context of the two-body problem in Newtonian mechanics and its relativistic or quantum counterparts.² The nature of idealisations and the role of models have been widely discussed with regard to the approximate nature of the idealisation involved in the simple (either classical or quantum) harmonic oscillator model.³ I claim that we can get new insights into many of these issues from studying them in the context of many-body phenomena. I try to show the fruitfulness of this kind of study here.

The main theme of my thesis is *modelling*, which is one of main research activities of physicists: the construction of mathematical/physical structures in order to understand target phenomena in the world. My interests in models are less in questions about their competition with statements as candidates for the true components of scientific theories,

¹) Leggett 1985

²) Cf. Sklar 1974, Maudlin 1994

but with what we can learn about the world (especially many-body phenomena in the world) from the successful employment of models. My proposal is to construe condensed matter physics as engaged essentially with the following three activities: model-building, model-exploring and model-based understanding.

A model should be carefully built out of a number of resources so that physicists can employ it later to understand its target phenomena. In the model-building stage, physicists may be helped by lots of different resources: by empirical data, by well-established heuristics, by a fresh look at an well-understood model and so on. Still, the most crucial factor in model-building in physics is the role of a few (general) theories such as statistical mechanics. These provide the core of our model-building methodology. They guide the entire process of a model-building in the sense that they usually provide the essential components of the model, and they tell us which way to go when we have to choose among different theoretical constraints on the model. Consequently, which general theory you pick as your model-building methodology will partially determine the character of the model (or how the model sees the world) and the explanatory strategy of the model (or how the model tackles a particular problem).

Also it is important to see how these general theories interact with each other. They certainly compete with each other, but in many cases they also complement and cooperate with each other to give us better understanding of their shared target phenomena. I study these multiple layers of interactions among general theories in order to show that not only fierce competition, but also complementation and co-operation among scientific theories are sometimes very fruitful. I also argue in particular that at least for the case of thermodynamics and statistical mechanics, the proper way of looking at the relationship is not through the concept of reduction but rather through the concepts of complementation and of co-operation.

Before we try to employ a model to understand its target phenomena, we need to understand the model itself first. This consists of the model-exploring stage. For instance

³) Cf. Giere 1988, pp. 62-91, Laymon 1985

we have to know what exactly the model implies in itself and what it could with additional constraints. We need to see whether a particular theoretical result of a model is a genuine feature of that model, or an artefact from the approximations involved in the derivation of the result. This point is seldom discussed by philosophers, although it is by physicists. A model in condensed matter physics is usually a highly mathematical and abstract object whose physical contents are not at all clear even to an expert in its first arrival. You have to develop the ‘canonical’ physical intuition about the model through the detailed investigation of it. I discuss how physicists do this job in order to separate the theoretical consequences of the basic architecture of a given model from those of the additional assumptions made for expediency.

After we understand a model itself, we are now ready to understand its target phenomena. With the help of appropriate interpretative models, an well-understood (abstract) model may provide us with our best representation of the phenomena under consideration. I discuss the nature of model-based understanding, and how we manage to understand so large range of phenomena with various types of models constructed under the guide of a few model-building methodologies.

The structure of the chapters is the following: I offer my proposal to see general theories as model-building methodologies in chapter 1. I defend this view in chapter 2 in the particular context of thermodynamics and statistical mechanics. I shall also criticise Sklar’s view seeking the reduction of thermodynamics to statistical mechanics there. In chapter 3, I study various model types in condensed matter physics to show how physicists use different types of models to achieve different goals in their pursuit of a model-based understanding of the world. In chapter 4, I study the model-building and the model-exploring stages in detail, highlighting the distinctive roles of two different types of approximations/idealizations in each stage. The renormalization group is the theme of chapter 5. There I explain how physicists successfully deal with some crucial features of critical phenomena using the renormalization group methods, and argue that here again what we see is not one theory (renormalization group) making another theory (mean field) redundant, but rather a happy complementation to each other.

CHAPTER 1

GENERAL THEORIES AND SUBSTANTIAL THEORIES:

MODEL-BUILDING METHODOLOGY AND MODEL-BASED UNDERSTANDING

1.1 Introduction

We want *understanding* (among other things) from our scientific theories. Condensed matter physics (CMP) studies the macroscopic, many-body phenomena that result from strong interactions¹ among huge number of entities (particles or fields) in a relatively dense state. Naturally we expect CMP to provide us with some understanding of these phenomena: how water changes its *phase* from liquid to gas apparently instantly at the exactly specifiable physical conditions, or what makes metal an excellent conductor of heat and electricity. This thesis is an attempt to investigate philosophically how condensed matter physicists are doing their research in order to get this understanding. This chapter sets the basic scheme: the general strategies or methodologies of model-building and the nature of model-based understanding.

As we will see in chapter 3, CMP is full of all sorts of models employed for various purposes. So it is *prima facie* natural to start with models of CMP in order to study the nature of ‘understanding’ we can get from CMP. Now a model has to be constructed before being put into use. Recently, several authors have taught us that a typical model in science is not just an ‘impoverished’ theory that can, if one wishes, always be derived from a relevant theory by deliberately omitting certain features of the full theory.² We need far more than a theory to build a successful model. On the other hand, a model is not always a product of the premature stage in scientific research before it grows up into its

¹) Strong interaction in this context has little to do with the strong force among nucleons. Usually it means a long-range interaction such as Coulomb interaction.

²) See for instance, the collection of papers in Morgan and Morrison 1999, especially Cartwright 1999a for the case of the BCS model, Hartmann 1996 for the case of the models in hadron physics, Morgan 1999 for the case of Fisher’s models of the monetary system and Morrison 1999 for the case of Prandtl’s model of a fluid.

full-fledged form, a mature theory.³ Physicists would build a model even when they had a ‘complete’ theory, sometimes even a *calculable* complete theory. I will argue in chapter 3 and in chapter 4 that this is because the model would be built to serve special epistemic purposes that a ‘complete’ theory could not. Understanding of many-body phenomena would be one of them.

Given the non-triviality of the model-building procedure, we need to study carefully how a model is built out of various resources before turning our discussion to the nature of model-based understanding. Moreover, as I will show later, the better grip of the model-building procedure will contribute to tackling several other important problems in philosophy of science such as how to understand scientific theories and what mathematics can do for an empirical science. In this chapter I shall focus on the role of theories in the model-building procedure, while in chapter 4 the role of approximations (and idealisations) will be discussed.

There have been lots of discussions in philosophy of science about the *nature* of scientific theories: what scientific theories are and are not. The syntactic view and the semantic view in their various forms are the two most influential views among philosophers of science at the moment. Recently the semantic view seems to have gained more support over the syntactic view, especially in the form of its more liberal variety such as Giere’s or Cartwright’s.⁴ But the die-hard syntactists never surrender and the battle between two camps goes on.⁵ I am going neither to defend one of these two competing views, nor to offer yet another alternative account of what scientific theories *are*. My focus is different. I am interested in what scientific theories can *do*, especially through and as models. I am also interested in what we can *learn* from the successful employment of scientific theories, again through and as models. To put it differently, my primary interest in scientific theories lies in their *functions* and *cognitive utility*. In order to address these issues, I will introduce in the next section a distinction among theories used in CMP,

³) Though some of the models can be construed as such. Lakatos 1970 provides us with an argument *tour de force* for this ‘heuristic’ role of scientific models.

⁴) See Giere 1984 and Cartwright 1983.

⁵) Cf. Savage 1990

which many physicists, I claim, implicitly make in their research. Then I will discuss the significance of the distinction as regards model-building and model-based understanding in sections 3 and 4. But before that a short preliminary remark is in order.

Throughout this chapter, I will assume as little as possible about the nature of scientific theories so that my conclusions about the function and the cognitive utility of theories should be as independent as possible of any particular view on what scientific theories are. For instance I do not exclude the possibility that the scientific theories we have here and now cannot be understood by one single account. I believe that some liberal kind of the semantic view such as Hughes' or Giere's is most satisfactory for understanding the theories of physical science.⁶ But it may well turn out that theories in so-called soft-science should be understood quite differently from physical theories. Still I am going to assume in this chapter the following: that working with models is central to the research of condensed matter physicists and that theories are one of major resources that scientists make use of when building a model. I take both claims to be uncontroversial.

1.2 General Theories and Substantial Theories in CMP

I claim that there is significant relationship between the *generality* of a theory in physics and its *substantiality*: the more general a theory in physics is, the less substantial it tends to be. Based on this relationship, I propose a categorisation of theories in physics into two groups: *general theories* and *substantial theories*. I will argue then that general theories function as a model-building methodology, while we expect substantial theories to give us a model-based understanding of the world. In order to explain what I mean by 'generality' and 'substantiality', let us start with the following observation.

Physicists tend to group their diverse research subjects in two different ways. Mostly they would say their research interests lie in elementary particle physics, in atomic physics or in solid state physics. These are clearly the classification based on the kind of phenomena they are investigating; so elementary particle physics studies elementary particle

⁶) See Hughes 1996, 1997 and Giere 1988

phenomena and solid state physics, solid state phenomena. But then there is another way of classifying theories in physics, which might be more familiar to philosophers; that is to say, there are classical mechanics, quantum mechanics, statistical mechanics, thermodynamics and so on. Notice that the majority of physicists would not describe their research area as classical mechanics or quantum mechanics, i.e., one from the second grouping.⁷ Exceptions are usually found among mathematical physicists who might happily confess researching on statistical mechanics for instance. As we will see later, there is a good reason for that.

More specifically in CMP, we can discern a difference in physicists' attitude towards two groups of 'theories': between theories like quantum mechanics or statistical mechanics and theories like the Landau theory of continuous phase transition or the Bloch theory of electrons in metal. They tend to think of a theory from the first group as the source of versatile modelling methods, general ideas, mathematical techniques and a particular way of looking at the world. So they would say, 'if we look at this phenomenon in the statistical mechanics formalism, we can model it by the two dimensional Heisenberg model'. Or they would say, 'if we want to treat it quantum mechanically, we have to consider the zero-temperature fluctuations explicitly'. On the other hand, physicists tend to regard a theory from the second group as their, however incomplete and tentative, intended representations of its target phenomena and as a source of their understanding of the world. So they would say, 'the Landau theory/model doesn't quite adequately deal with the order parameter fluctuation very near the critical point'. Or they would say, 'Although we know that superfluidity is clearly a quantum mechanical phenomenon, we

⁷) I cannot resist the temptation of telling my small episode relating to this point. Once I was solemnly corrected by my supervisor on the matter of describing my major research field. At that time I was working for my master thesis on some interesting properties of a model called the generalised Harper's equation. I thought I was doing statistical mechanics because all the mathematical techniques I was using and all the important theoretical ideas I was entertaining for the model basically came from statistical mechanics. But my supervisor, professor Moo Young Choi, persuaded me to identify myself as a condensed matter physicist, pointing out to me that the ultimate reason I was trying to understand the generalised Harper's equation is to understand the kind of phenomena the model might successfully describe, for instance the two dimensional Josephson junction array under incommensurate magnetic field. In short, I was less interested in the development of general-purpose tools such as statistical mechanics, than using those tools to study many-body phenomena. If I had been a mathematical physicist, my interests would have been exactly reversed.

do not understand it because we have not yet figured out the right theory/model that can cover all the aspects of the phenomenon’.

So why is there this difference? A quick and easy answer would be in terms of the intended range of target phenomena of a theory. A theory from the first group is supposed to apply to a wider range of phenomena, ideally perhaps to all the phenomena in the world. On the other hand, a theory from the second group is regarded as targeting more specific and restricted phenomena. So, according to this quick and easy answer all the differences in attitude towards the two groups of theories among condensed matter physicists boil down to the difference in the scope of the potential validity of a theory. For instance, quantum mechanics is expected to be valid, if ever, everywhere in the entire world,⁸ but the Landau theory of continuous phase transition is, obviously, valid only in the continuous phase transition phenomena.

One consequence of this answer is uncomfortable to swallow for those who believe that the aim of science is to pursue the most simple, unified, truth about the world. If this answer is correct, most physicists turn out to be spending most of their time pursuing limited and disparate truths about some specific phenomena rather than the unlimited and grand truth.

Certainly there is no doubt that a theory from the first group is more generally employed to explain a number of phenomena than a theory from the second group. But we cannot draw from this the conclusion that therefore theories in the first group are *universally* applicable while theories in the second group are not. This is because there are important ambiguities in the meaning of ‘generally applicable’. When physicists say that quantum mechanics for instance is generally applicable for the entire world, it could mean either that they would regard it to be true of the entire world or that they intend it to be employed for modelling as many phenomena as possible. The first meaning of ‘general applicability’ is closely connected with the idea of the so-called ‘fundamental’

representation of the world; that is, a theory is generally applicable only if it gives us the complete, true representation of the world. But the second meaning does not necessarily imply this; we can aim to employ a theory in modelling as often as possible whenever it is appropriate even if the theory doesn't even try to give us the 'fundamental' representation of the world.

I claim that in most cases (at least in CMP) what physicists mean is closer to the second; they generally do not confuse the intended scope of application with the assumed scope of truth. They rarely talk about a theory from the first group in terms of truth but rather in terms of effective applicability. So they would say, 'we could *treat* this phenomena quantum mechanically, but then we wouldn't understand some important features of it. On the other hand, we can successfully *model* most of its features classically'.⁹ Physicist, when pressed, will usually pay lip service to the fundamental truth of QM, saying, 'Of course, quantum mechanics is generally true of all the phenomena of the world'. But when asked for their reason, they would tell us, 'You see, we could treat all these phenomena quantum mechanically, if we wanted'. But that is the point: we do not always want to treat a phenomenon quantum mechanically. We are *selective*: we would construct a quantum mechanical model of a phenomenon only if the model could allow us some understanding of the phenomenon. For the same reason, we would construct a classical (or statistical mechanical or thermodynamic) model of the phenomenon only if we can gain some extra understanding of it by that particular way of modelling.¹⁰

For the entire thesis, I argue that we have to take seriously modelling and understanding based on the successful use of models to be central to research activities in CMP. From this modelling point of view then, the most important question we have to ask about a theory from the first group is how many resources we can draw from it when we model a particular phenomenon. If we look at the other side of the coin, given a phenomenon we are more interested in how many features of it can be understood as, say quantum

⁸) One could add 'everywhen' and say, 'everywhere and everywhen in the world'. But then even the theories from the second group are expected to be valid, if ever, everywhen. Modern physics is generally not a 'historical' science.

⁹) Cf. Goldenfeld 1992, p. 81 and Fisher 1988

mechanical or classical, rather than in whether quantum mechanics (or classical mechanics) is simply true of it.

So we conclude that the target phenomena of a theory from the first group are ‘general’ in the sense that its intended scope of application is potentially all the phenomena in the world. We try our best to use the various resources the theory offer and model as many phenomena as possible. On the other hand, a theory from the second group has a different relationship with its target phenomenon: we intend it to be our representation of that phenomenon, or a model from which we get our understanding of the phenomenon. Here the ‘truthfulness’ of a representation (or model) does matter, even if truth itself does not. For instance physicists do care and, if needed, are very much willing to debate on fiercely whether the Bloch model rather than the Sommerfeld model correctly captures the ‘essential physics’ of electron conduction phenomena in a metal.¹¹ The relationship between these two models is clearly that of competition: we have to settle down with one or the other as our best representation of the phenomena at the current stage of physics. Consider a similar situation with theories from the first group and appreciate the difference. Physicists are usually happy to ‘hybridise’ different modelling strategies of different theories in the first group if they can understand the phenomena under consideration better by doing so.¹² There, the relationship between two theories is not a competition, but rather a co-operation. I shall investigate this point more closely in chapter 2.

Notice that a theory from the second group can have a substantial degree of ‘generality’ depending on the nature of the phenomena it models. If the phenomenon is a very general phenomenon such as the conductivity of a metal, a theory of it, if successful, should be applied to a number of different metals. Also a theory is usually regarded to be better if it is applicable to a wider range of phenomena. The theory may not apply for all the metals in the world, but then we expect the theory to tell us in a physically significant way why it can explain these kinds of metals, but not the others. So in this sense, ‘general

¹⁰) More on this point in chapters 2 and 5

¹¹) See the discussion in Ashcroft and Mermin 1976, chapters 2 and 8

applicability' is a virtue for a theory from the second group as well as for a theory from the first. But we expect more from a theory from the second group than a set of versatile model-building tools.¹³ We want it to provide us with an acceptable theoretical explanation of the phenomenon under consideration and possibly with a reasonable representation of the phenomenon.¹⁴ This explains a subtle difference in the way general applicability is a virtue for theories from each group. While we try to employ a theory from the first group and build models of as many phenomena as possible, we do not always regard the versatility of a theory from the second group to be desirable. Instead we want 'substantial' physical differences between different phenomena to be reflected in theories from the second group. In this way, the theories will be less versatile and more specific. But there is no contradiction here. Unlike the theories from the first group, we want the theories from the second group to present us our best substantial knowledge of the world. So we must prioritise 'substantiality' over 'general applicability'. On the other hand, we could imagine that the theories from the first group present us our best knowledge about how to build a model of any phenomenon. Here the goal is to develop a successful methodology of model-building for as many phenomena as possible. As a result, we value most a versatile theory.

So the difference between the two groups runs deeper than it first looks. What I shall from now on call *general theories* (GT) such as quantum mechanics aims to provide a set of versatile, necessary tools for model-building, and thereby function as the background framework of what I will call *model-based understanding*. On the other hand, theories like the Bloch model aim to provide our best (however incomplete) representations of the world. I shall call them *substantial theories* (ST). A GT has no representational value in itself; all of its representation value comes through the models whose construction it guides, that is through its associated STs. Otherwise a GT would be an empty mathematical (or metaphysical) framework of how we think about the world.

¹²) Cf. Wise and Brock 1998

¹³) Cf. Cartwright, Shomar and Suarez 1995

One often talks about the ‘Quantum World’ or the ‘Classical World’. The Quantum World is the world of which quantum mechanics is universally true. Consequently, quantum mechanics as a GT is universally applicable for modelling every single feature of that world. In other words, a set of quantum mechanics-associated STs can give us the complete understanding of that world. Obviously if you believe in the modelling power of quantum mechanics, you will try to envisage as many features of the real world as possible in terms of their corresponding features of the Quantum World. But you can succeed in doing so only by constructing quantum mechanics-associated STs for one feature after another. And each time you succeed, you will identify that particular feature as quantum mechanical. So the identification of a feature in the world is done by a ST step by step, and never by a GT in a single stroke.

Here is a rough criterion for distinguishing a ST from a GT. A ST, unlike a GT, is usually taken to be a theory of something; that is to say, a ST specifies its target object/phenomena more or less clearly. The target phenomenon itself could be a complex one such as all known properties of metals. However, although the final dream theory of metals could be postulated as a motivational ideal for the research, most of CMP models have only a few target features that physicists want to understand. If a particular model turns out to be so good that it can explain other features that were not originally targeted in the model-building stage, then that is very well. But it is often the case that the very ‘targeting’ of particular phenomena in the model-building stage restricts the validity of the model in other phenomena, as I will argue in chapter 3.

Now I reply to some possible doubts about the usefulness of my distinction. One might think that my distinction makes difference in ‘degree’ into difference in ‘kind’. Surely, if someone believes that, say, quantum mechanics gives us the complete representation of the world, then quantum mechanics becomes just a ST, but a very general one that has its target phenomena as all the phenomena of the world. I disagree. The main reason comes from the following two considerations. The first comes from looking closely at how GTs

¹⁴) According to Hughes, a theoretical explanation of a phenomenon is obtained (1) if it is given by an ‘acceptable’ theory and (2) if the representation offered by the models of the theory is ‘adequate’, which

are used in CMP. A GT provides a framework to set up the initial description of target phenomena, heuristics for picking up the set of candidate models, worked-out examples of successful applications of its models and standard way-outs when the researcher meets difficulties in working with the chosen model. In other words, a GT functions very much as *methodology*, in its original sense of ‘the system of methods and principles used in a particular discipline’.¹⁵ As a methodology, the semantic content of a GT is intrinsically indeterminate. Figuratively speaking, a GT rarely says anything definite about the world by itself. Of course it says various things through its ST on various aspects of the world, but only through them.

My second consideration is about how exactly physicists *employ* a GT in order to produce a ST. They do not necessarily use one GT in one ST: mixing GTs and hybridising different principles from different methodologies are common features of modelling in CMP.¹⁶ On the other hand, it is rare that they mix or hybridise different STs. Physicists may be instrumentalists in the sense that they want above all things models that work. But, unlike engineers, they are rarely *mere* instrumentalists about their models. If they believe for instance that the Bloch model of electrons in metal somehow captures the essence of electron physics in metallic phenomena, they do not hybridise the Bloch model with another model. Rather, they usually try to improve it by *refining* it, adding more realistic interactions or loosening unrealistic assumptions. This practice is easily understandable when we take GTs as model-building methodologies, and ST as our intended representations of target phenomena.

1.3 Model-Building Methodologies: Clarifications and Philosophical Implications

The essence of depicting a GT as a model-building methodology is to highlight how a GT guides the model-building process. Especially, how each GT provides us with a set of ‘autonomous’ principles with which we can look at the world in one particular way and build a model for certain features of it. Autonomous principles of a GT are not strictly

again means the models offer us explanation and understanding of the phenomenon. See Hughes 1993.
¹⁵) Collins Concise Dictionary, Fourth Edition 1999

reducible to another GT, and are also usually independent of each other within the GT. Taking the thermodynamic limit procedure is one notable example of these autonomous principles of thermodynamics. As we will see in chapter 2, you cannot justify this procedure by any other GT including an obvious candidate, statistical mechanics. You can give certain explanations why this procedure is ‘reasonable’ to follow only if you accept the framework of thermodynamics as your model-building guide. Still, this procedure is an essential element of every thermodynamic modelling of many-body phenomena in the world. Virtually every quantity in a typical thermodynamic model is obtained through this procedure. The procedure is in a way a method of turning what you can measure in real systems into the kind of quantities thermodynamics can deal with.

There is a long tradition in philosophy of science of talking about methodology in relation to theory change or choice, especially rational theory change or choice. All the important methodologists including Lakatos and Kuhn start their discussion of methodology considering a *series* of theories in the same research program (paradigm), or the temporal development of a theory through cumulative change or ‘revolutionary’ change. Then they typically ask questions like how we can decide one series of theories is better than another series of theories or how we can make sure the change a theory might experience through its development is rational or justifiable.¹⁷ In short they are interested in how to evaluate rationally a given theory with respect to its competing theories, that is in the methodology of (rational) theory appraisal.

My interests in methodology start from modelling: how to build a model using a set of resources. That is to say, my methodology is the methodology of model-building. I am particularly interested in how condensed matter physicists employ various resources in order to build a model of many-body phenomena and try to understand the phenomena through the model. A GT is one of the prime components among these resources, and functions as a methodology of model-building.¹⁸

¹⁶) Wise and Brock 1998

¹⁷) Kuhn 1970, Lakatos 1978, and Lakatos and Musgrave 1970.

Putting it this way, my sense of ‘methodology’ would seem to have little to do with the more traditional sense of ‘methodology’. But in fact there are some significant interconnections between two. First of all, many of the basic objects of traditional discussions in methodology are more or less same as mine. I say ‘more or less’ because the basic objects of the traditional discussion of methodology are often only loosely characterised. For instance, in Lakatos’ research program, ‘theory’ can mean a very general one such as quantum mechanics or a more specific one such as Bohr’s theory of the atom. It is understandable why Lakatos opted for this loose usage because he was mainly interested in the long-term, temporal change of a theory. As long as he can talk about the rationality behind these changes, the entity of the changes can be anything. In other words, for Lakatos’ methodological discussions, the difference between a general theory and a more specific theory is insignificant and even irrelevant.

On the other hand, I need the distinction between GT and ST because I am mainly interested in the relatively short-term process of model-building. Here the different role of each type of theory is crucial. Still, although both a GT and a ST could be discussed in terms of Lakatosian research program, GT is more susceptible to that sort of discussion because it is more likely to develop more dynamically as new methods are appended to the existing set of methods.

The similarity between GT and Lakatos’ research program (or Kuhn’s paradigm for that matter) goes further. In Lakatos’ research program, the positive heuristics help practitioners to develop the research program in an empirically successful way. A GT does the same thing for model-building, providing helpful suggestions and wisdom. But I am also interested in the models themselves and they constitute STs. While the traditional discussion of methodology in philosophy of science tried to give us a large-scale dynamic picture of scientific theory change, my discussion of methodology is an attempt to offer a small-scale dynamic picture of model-building. A ‘static’ counterpart would be the picture of how physicists study their models and get what they want from it. That is the topic of the next section.

¹⁸) For other resources than GTs in model-building, see chapters 3 and 4.

But before that, there are several philosophical implications of depicting GT as a methodology. As we regard GTs like thermodynamics or statistical mechanics as alternative methodologies, it becomes difficult to talk about reductions among them. After all, the empirical content of a methodology is given by its associated STs, and thus not fixed at any given time. So if someone wants to claim that a general theory **A** reduces another general theory **B**, he has to show a sort of meta-theorem to the effect that anything which can be successfully dealt with by **A** can be also dealt with by **B**; or in other words, all potential **A**-associated STs can be deduced (or explained) by the set of all potential **B**-associated STs. This would be indeed a truly impressive task if only some could do it. But even without the proof of the impossibility of this meta-theorem, we can appreciate how difficult the task would be.

Or, one should be satisfied with a more modest reduction such that all *currently available* **A**-associated STs can be deduced (or explained) by the set of all *currently available* **B**-associated STs. This is certainly an easier task. But I will show in chapter 2 that for the case of thermodynamics and statistical mechanics, this modest task cannot be achieved. There are substantial parts of many-body phenomena that can be dealt with successfully by some TD-associated SP's, but not by any of *purely* SM-associated SP's.

Treating a GT as a methodology might be thought of as just one way of looking at scientific theories. In a sense that is true, but there is one prominent virtue in emphasising the methodological nature of a GT. We can discern more clearly the autonomous and independent motivations and research histories of different disciplines of physics. This emphasis prevents us from drawing wrong conclusions about theories in physics, especially the ones misled by the similarity of formal/mathematical structures from two intrinsically different research traditions. I will argue in chapter 5 that despite the important mathematical connections in their formalism, the renormalization group methods in CMP are motivationally and conceptually quite different from those in quantum field theory.

But this point is more general: the mathematical structure of a GT is an important component, but it tends to obscure the crucial methodological differences between different GTs. Still methodological differences between GTs makes a big difference in their associated STs, and consequently our representation of the world. For instance, even though you could succeed in accommodating classical mechanics within an elaborated Hilbert space formalism (or quantum mechanics into a generalised Hamiltonian formalism), that would not mean that you have found a way of converting all classical models into quantum models (or vice versa). The know-how of model-building is unique to each GT, so you would need more than a transformation (from one mathematical formalism to another) in order to do the trick.

1.4 The Nature of Model-Based Understanding

I argued in section 1.2 that STs have ‘representational’ value in contrast to GTs. Exactly what kind of representational value a ST can have however is not an easy question to answer. Many authors have pointed out again and again the ‘unrealistic’ character of models: simplified structures, omission of significant causal mechanisms and so forth. Then what can we expect to learn from highly idealised theories? I claim that what we can get from substantial theories is *model-based understanding*. Before I explain what constitutes model-based understanding, we need a short preliminary discussion.

Understanding an object by the help of another object is an epistemic-subject relative notion. For instance Finns can understand their history from their history books, but I cannot as I do not have relevant epistemic capacities, in this case the familiarity with Finnish. Likewise, model-based understanding should be epistemic-subject relative. So we may not get understanding of, say, a phenomenon of boiling water when we have the ultimate (presumably messy and detailed) complete description of the phenomenon. Some form of artificial intelligence, if we are willing to accept that they have their sort of ‘understanding’, may understand boiling water out of this description. But who cares? What is important to us is some theory or models by which we as educated people can understand a given phenomenon.

So when we want some criteria that tell us when we have model-based understanding, the criteria should take into account certain built-in human capacities as well as a reasonable training in contemporary science way of looking at the world. Now let us consider possible candidates for the criteria. One obvious choice would be ‘empirical adequacy’.¹⁹ So one might say we get model-based understanding if the model of interest is empirically adequate. There is no doubt that empirically adequate models are highly regarded among condensed matter physicists. But empirical adequacy is usually taken to be an extra virtue of a model, which is not considered a necessary feature of model-based understanding. For instance, it is quite common in a theory of phase transitions that one feels legitimate to say that she understands a phenomenon by a particular model when the model gets the qualitative features of the phenomenon right, which are regarded as the essential physics of the phenomenon. On the other hand, even when we have a fantastic predictive success by a model, physicists usually hesitate to claim they understand the phenomenon without understanding the model itself.²⁰ When the model under consideration is too artificial, until someone figures out an ingenious interpretation of it that makes the physical content of physics more or less ‘transparent’ to the physicists community, physicists will not say they have model-based understanding with respect to that model. I will talk about more about the significance of interpretation in characterising models in chapter 3, but it is safe to say that model-based understanding does not necessarily presuppose empirical adequacy, at least in Van Fraassen’s rather strong sense of ‘empirical adequacy’.

How about instrumental utility? Hacking’s epistemologically demanding version of instrumental utility relates the reality of an entity to our ability to produce a new effect using that entity.²¹ Or more modestly, we generally value novel predictions as a virtue of a theory. Should we perhaps require a condition for model-based understanding in a similar spirit? So we might suggest that we get model-based understanding if we can produce a new, otherwise unexpected, effect thanks to the model of interest. Certainly

¹⁹) Van Fraassen 1980

²⁰) Physicists usually refuse to take seriously the Gaussian model despite its excellent empirical adequacy in its high temperature limit because they think that the interactions of the model is ‘unphysical’ and that the model becomes ‘meaningless’ (not just empirically inadequate) in the low temperatures.

producing a new effect using a disputed model helps a relevant physicists community to accept the model. For instance, there are several models in the theory of phase transitions that have risen up from a mere candidate mechanism of possible phase transitions to our best representation of a particular phase transition, mainly due to their prediction of certain unexpected effects. The growing interests in the Kosterlitz-Thouless transition model has been driven by its novel predictions of non-universal critical phenomena.²² But, physicists would not claim model-based understanding, despite the impressive effects the model predicted, if they had not thought they understood the Kosterlitz-Thouless mechanism on the physical grounds first. So here again, model-based understanding presupposes our understanding of the model itself before evaluating how well it performs in explaining the phenomena in the world.

Hasok Chang's suggestion is interesting in this respect. His topic is different from mine: it is about realism. He proposes the reformulation of realism as the pursuit of *ontological plausibility* in our system of knowledge. Ontological plausibility is conceived as a precondition for the intelligibility of a scientific theory under consideration. As a result, the kind of realism Chang defends facilitates the understanding of the world that goes beyond mere description or prediction. He acknowledges the inherent difficulties in making objective judgements of ontological plausibility, but he suggests that they become manageable if we adhere to the most basic ontological principles: those principles that are regarded as essential features of reality in the relevant epistemic community.²³

Here are a few examples of ontological principles that Chang discusses in the paper. (1) Principle of single value: a real physical property can have no more than one definite value in a given situation.²⁴ (2) Principle of 'no miracle': if there are regularities in nature, they cannot be suspended on isolated occasions to allow inexplicable happenings. (3) Principle of continuity: if the set-up (or cause) varies continuously, the outcome (or

²¹) Hacking 1983, pp. 262-275

²²) Goldenfeld 1992, pp. 345-350

²³) Chang forthcoming; also see Chang 1999 for related issues.

effect) cannot change abruptly and discontinuously.²⁵ Chang claims that the satisfaction of ontological principles is the basis of intelligibility in any account of reality; the denial of an ontological principle would strike us as nonsensical rather than false.

Translated into my discussion of models, Chang's suggestion boils down to a methodological dictum: if you want to be this kind of a realist, you should try to make your models not only empirically successful but also consistent with basic ontological principles we value most in our science. Chang's suggestion is illuminating not only because he forcefully claims that there should be more features for a theory/model other than empirical adequacy but also because he relates this extra virtue to the *intelligibility* of the model. For him an acceptable model should be intelligible to us by satisfying various ontological principles. If a model is empirically adequate, but not intelligible in this sense, according to Chang, we do not have a realistic belief about the model.

Something parallel goes on with model-based understanding. Throughout my discussions in this section, it becomes clear that we require the model under consideration to be understood (or intelligible) before we claim model-based understanding using the model. Here the understanding of a model typically consists of the following: figuring out at least one plausible (preferably realistic) physical mechanism for the model, determining the theoretical consequences of the model about its target phenomena and developing our physical intuitions about the model. A plausible physical mechanism for a model is given by what I call an 'interpretative' model: a concrete (physically possible) implementation of the abstract/mathematical structure of a given model, of which the physical constitutions are transparent enough to educated physicists.²⁶

In order to determine what the model implies about its target phenomena, we 'explore' the model using various mathematical techniques. One of the major purposes of this 'exploration' is to identify what the true features of the model are; in other words, what

²⁴) Chang discusses elsewhere how a French experimentalist Victor Regnault in the 1840s applied this principle to achieve an ontologically satisfactory way of determining temperatures. See Chang 2000.

²⁵) Leibniz composed a devastating attack on Descartes' physics relying on this principle. See Leibniz 1985 [1692].

the model can do with and without additional assumptions that are not a part of the original structure of the model.²⁷ Identifying these certainly helps us shape our physical intuitions about the model. We need however at least one successful application of the model in explaining a phenomenon in order to establish a set of ‘canonical’ physical intuitions about the model. Physicists then use their physical intuitions when they try to devise alternative plausible physical mechanisms (interpretative models) for the model and explain different kinds of phenomena. Once physicists come up with a shared set of physical intuitions about a particular model, they could treat the model as if it were a real physical object (with some non-trivial features) like an automobile or a fax machine. They have stabilised expectations about the possible reactions of the model under certain parameter change, just as we know that we have to push certain buttons in the unique order in order to fax documents.

In order to illustrate these points, consider the q -state Potts model. The Potts model is a *lattice* model; that is to say, it consists of an abstract (infinite) n -dimensional lattice with its only dynamic variables, ‘spins’, at each lattice point. Each spin, S_i , can be in one of q different states ($S_i = 1, 2, 3, \dots, q$). The interactions of the model are only between spins next to each other (nearest neighbourhood interactions) and given by the Kronecker delta function; that is, $\delta(S_i, S_j) = 1$ (if $S_i = S_j$), and 0 (otherwise).²⁸

If we are given this mathematical structure only, we cannot claim that we ‘understand’ the Potts model. Of course there is one obvious sense in which we do understand the Potts model; that is we can imagine an abstract structure of an infinite lattice with certain entities and certain interactions among them. But that is rather uninteresting sort of understanding, at least to physicists. For instance, we have no idea, at this stage, what ‘spins’ are intended to target, and we do not know what sort of physical interactions can implement this rather peculiar on-and-off type interactions between the spins. Moreover, it is not clear up to this point how we cash out the infinite size of the lattice, as we all know all physical objects CMP deals with are finite.

²⁶) More discussions about interpretative models, see chapter 3.

²⁷) More on the ‘exploration’ of the models, see chapter 4.

A number of interpretative models have been proposed for the Potts model, and some of them are quite successful. Here is one successful example. We can consider an interpretative model (with $q = 2$) where we model the magnetic moment of atoms in magnetic materials as ‘spin’, and their complicated electromagnetic interactions, as on-and-off interaction. This may look like a rather ‘unphysical’ interpretative model since we know that the interactions between magnetic materials are very messy. But it turns out that in many cases of ferromagnet, the crystal structure of the magnetic solid produces additional forces on the magnetic moments of the atoms which may produce alignment in a definite direction in space so that we can successfully model the interactions in the on-and-off style.²⁹

Here we have a plausible physical mechanism of ferromagnet that provides a successful interpretative model for the Potts model. The combination of the interactions among atomic magnetic moment and of the effects of crystal field offers one physically plausible way of getting ferromagnetic phenomena. Equipped with this interpretative model, the Potts model becomes intelligible to us. We understand for instance that the on-and-off interactions of the Potts model can represent complex interactions under certain physical constraints. Also we learn from how temperature and the ground state is related in ferromagnet, the ‘canonical’ way of co-ordinating temperature and interactions in other interpretative models for the Potts model.³⁰ That is to say, we develop a set of physical intuitions about the Potts model with which we can search for alternative physical mechanisms for the model in the future.

Here is another interpretative model for the Potts model ($q \rightarrow 1$ limit), which depicts an alternative physical mechanism. We take locations in space individuated by lattice points to be ‘spin’ and whether two neighbouring locations are connected or not, to be ‘interactions’. Then the Potts model equipped with this interpretative model becomes a

²⁸) Wu 1982

²⁹) Thouless 1989

³⁰) That is, you should model the physical interactions into the Potts model in such a way that all q -states are equally probable at the high temperature limit.

model of so-called *percolation* phenomena such as the spreading of a ground water mound in a porous rock. The nature of the physical mechanism of percolation is probabilistic ‘transmission’ of a certain property (say, being electrically connected) across spatial regions (say, into four different directions in the two dimensional lattice case) under given physical constraints (say, the transmission probability is inversely proportional to the distance from a ‘seed’). Now we learn from our experience of the ferromagnetic interpretative model of the Potts model how to calculate the average strength of (spontaneous) magnetisation in a ferromagnet. We can rely on that experience and on the physical intuition developed from that experience to calculate some important quantities in this alternative interpretative model of the Potts model. For instance, the rough counterpart quantity of magnetisation in this model is the number of locally connected regions, separated from each other (*‘clusters’*). So we can calculate how many clusters emerge on average at a given temperature and what the average size of those clusters is.

There are many cases in CMP where physicists hesitate to claim ‘understanding’ of a controversial phenomenon by a certain model despite the impressive empirical success of the model. Usually they are hesitant because they are not sure of the model itself.

Sometimes the model is too outrageously unrealistic under the conventional interpretation; or sometimes they simply do not know what kind of physical mechanism a given highly abstract model represents. Unlike mathematical physicists, condensed matter physicists are not willing to say they understand the phenomena in these cases. In other words, in order to achieve model-based understanding, the successful employment of a model for explaining/predicting certain phenomena are not enough. We need the understanding of the model as well.

After we understand a model, we may employ the model to understand its target phenomena in the world. This is done by ‘matching’ one of the interpretative models of the model with the central features of the phenomena. The matching should be motivated (ideally both theoretically and empirically) in the sense that we have good reason to believe that the central features of the phenomena can be thought of as having more or

less the same structure as postulated by the interpretative model. Citing the crystal field in a ferromagnet as a reason why the ‘spins’ of the Potts model can be a plausible representation of magnetic moment is a good example of this motivation.

When we see the ferromagnetic phenomena through a well-motivated interpretative model of the Potts model, we find that the central features of the ferromagnetic phenomena become ‘intelligible’. It is highly plausible to us that those phenomena can be produced by the physical mechanism proposed by the interpretative model because (1) we understand how we can obtain the counterpart of those phenomena from the Potts model with the appropriate interpretative model. (2) We know that the interpretative model is well motivated. In short the phenomena (through the Potts model) look physically sustainable in our world. Also when the matching is successful, we may be able to deduce various quantitative results out of the Potts model and see whether the model is empirically adequate as well. The positive answer from this empirical test will certainly raise our confidence that we genuinely have a model-based understanding of the phenomena, but the essence of our model-based understanding lies in the (qualitative) ‘matching’ the well-motivated interpretative model with the phenomena.

Therefore I propose the two-stage process of model-based understanding. model-based understanding consists of two stages: (1) understanding of the model under consideration, and this involves, among other things, exploring its potential explanatory power using various mathematical techniques, figuring out various plausible physical mechanisms for it and cultivating our physical intuition about the model; (2) matching the phenomenon with a well-motivated interpretative model of the model. The empirical success and instrumental utility all play their roles in the evaluation of how successful the model is, but not the essential part of model-based understanding.

The significance of including stage (1) into our voyage towards model-based understanding is that the physical content of a typical physics model, even though it is painstakingly built with the guidance of GTs, is not transparent. We need to develop the ‘canonical’ physical intuition about the model through, among other things, a number of

well-motivated interpretative models. Until the model is understood, we cannot use it to represent the phenomena in the world. In this sense, a model that has not been fully explored or understood still remains in the realm of GT. It has no concrete representational message. Only if we understand the model and employ the model to understand (and preferably deduce quantitatively correct results about) its target phenomena with the help of its interpretative models, we can grant it representational value. But then the representational value is almost entirely with the successful interpretative models, rather than with the original model. So an abstract physics model can transform itself into a ST only in its concretised form, qualified by one of its interpretative models.

My aim of this chapter is to put model-building and model-based understanding as the central part of CMP research of macroscopic phenomena. The model-building process is mainly done under the methodological guidance of GPs with the help of all sorts of things including Chang's ontological principles and empirical facts. And we now know that the main source of model-based understanding is a ST or a interpreted, successfully applied model with the help of its interpretative models. But then we realise that the model-exploration stage is an essential part of model-based understanding before turning them into a proper ST. So this is the full picture of how I see CMP research: model-building, model-exploring and model-based understanding. I'll discuss more closely each issue in the following chapters.

CHAPTER 2

THERMODYNAMICS AND STATISTICAL MECHANICS: ALTERNATIVE MODEL-BUILDING METHODOLOGIES OF CONDENSED MATTER PHYSICS

2.1 Introduction

Condensed matter physicists employ a number of resources in order to build models of many-body phenomena in the world and thereby to understand them. For instance, they may get an inspiration for a new model from newly available experimental data, or from a fresh look at an well-understood model. The original development of the Ising model as a simple model of ferromagnetic phenomena by Ising and Lenz can be regarded as an example of the first case.¹ And employing the Ising model to represent the social segregation among different ethnic groups in big cities is a good example of the second case.²

Still, the role of general theories (GTs) in the model-building procedure is most prominent. This is not at all surprising given that GTs usually *guide* the entire procedure of model-construction. When we *decide* to model a certain phenomenon classically for instance (even when we know that the phenomenon is due ultimately to quantum mechanical effects), we are thereby committed to deploying a number of the well-established model-building methods and strategies of classical physics. To take a concrete example, suppose we want to model classically the metallic phenomena of a solid. We would follow what one might call the (classical) ‘analytic’ method, ‘Treat the motion of electrons as that of a classical particle influenced by different causal factors (external force given by any externally applied field, an internal force exerted by the other electrons, an internal force exerted by the ions, the contribution of collisions manifested as a ‘friction-like’ form and etc.)’. On the other hand, if we decide to model the same phenomena quantum mechanically, the model-building strategies we may follow will be different. I will talk more about the details of the different model-building strategies in the theory of metals in chapter 4. In this chapter

¹) Brush 1983, pp. 233-243

²) Domb 1985

I will concentrate on the general theories themselves: what the essence of a general theory is and how general theories work together or relate to each other.

Understandably, physicists are often quite opportunistic when they construct models. More specifically, it is not uncommon to employ more than one general theory in the construction of a model. The successful hybridisation of classical physics and quantum physics in the semi-classical models of various phenomena from electric conduction in a metal to quantum chaos is not an exception but rather a rule. What I hope to show in this chapter is that this opportunistic attitude is not just a temporary makeshift resulting from practical difficulties in real-time situations, but rather a proper way of doing research in CMP. Considering what physicists can gain from the cooperation between general theories in understanding many-body phenomena, they have a good reason to practise hybridisation in modelling.

More specifically I shall look at how condensed matter physicists use two of major general theories in CMP: thermodynamics (TD) and statistical mechanics (SM). With classical mechanics and quantum mechanics, TD and SM are employed most frequently in CMP models. Pick up any model from a standard textbook of CMP, or from research journals, and you are likely to find, either among the basic components of the model or during the theoretical investigation of it, one or another employment of well-established methods or strategies from TD or from SM.

For instance, almost always physicists start their theoretical description of a given phenomenon by defining it as either an 'equilibrium' or a 'non-equilibrium' phenomenon. So when they investigate the crystal structure of a metal as an equilibrium phenomenon, they are not denying that the metal will sooner or later undergo a certain non-equilibrium process and lose the rigid structure, in the short time scale due to microscopic fluctuations and in a longer time scale due to dissipation. When they talk about the crystal structure of a metal and picture it as a fixed geometric shape, they are rather saying that we are looking at a particular state (or time-scale) of a metal where the crystal structure is in 'equilibrium' enough. In order to investigate (or model) this particular aspect of a metal, they average out the small-time scale fluctuations and disregard the long-term dissipation effect. But things should change drastically, if they are interested in 'non-equilibrium' properties of a

metal, say how dissipation occurs in a metal. Then they no longer treat the crystal structure as a fixed shape, but rather model it as for instance an elastic medium.

One of the lessons we may learn from the above example is a familiar one: we can model the same phenomenon in different ways (sometimes incompatibly different ways) depending on the target features of the phenomenon we particularly want to understand. I will discuss this point in more detail in chapter 3. Here I used the example to illustrate how ubiquitous the employment of TD (or SM) is in CMP modelling and how the choice of a general theory affects the model-construction procedure significantly.

Given the importance of a general theory as a model-building methodology in CMP and physicists' tendency to employ more than one general theory in a single model the following question arises: how should we understand the relationship between more than one general theory employed in a single model? The traditional answer to this question is by the concept of reduction: one general theory (say, SM) reduces another general theory (say, TD). Then the use of two general theories in a single model may be taken to be innocuous because there is, in a certain fundamental sense, only one general theory used after all.

My thesis is that this way of understanding the relationship between GTs is unsatisfactory. In order to argue for my thesis in a more specific way, I shall discuss Lawrence Sklar's sophisticated view on reduction of TD to SM. Although my interests in the relationship between different GTs are partially motivated by the 'hybridisation' practise of physicists in modelling, my criticism of Sklar's reduction will not be restricted to modelling, but general. My aim is to show why the concept of reduction is generally unfruitful in understanding the relationship between TD and SM. I will also argue that a better way of understanding the relationship is through alternative (modelling) methodologies. As TD and SM are two major GTs of CMP, by showing the fruitfulness in thinking of them as (model-building) methodologies, I hope to provide an additional supporting argument for my claims in chapter 1.

2.2 Sklar's Reduction of Thermodynamics to Statistical Mechanics

Lawrence Sklar in his recent book, *Physics and Chance*³, proposes a sophisticated account of reduction of thermodynamics (TD) to statistical mechanics (SM). He admits that it is hard to capture the complicated relationship between TD and SM by a simplistic general model of reduction. So he says, 'One thing we shall not do is that many important cases of alleged reduction seem to have such special features that none of the simpler general structural models can do them justice. ... the alleged reduction of thermodynamics to statistical mechanics is ... one of those cases where the more you explore the details of what actually goes on, the more convinced you become that no simple, general account of reduction can do justice to all the special cases in mind.'⁴

He discusses what he thinks is wrong with several traditional accounts of reduction in science including Nagel's classical account of the reduction of TD to SM.⁵ He also acknowledges that in order to 'derive' TD's theoretical results from SM, we need extra resources *outside* of SM, sometimes from observations and sometimes from the 'derived' theory, TD, itself.⁶ Again he accepts that the reduced theory usually suffers changes through the reduction procedure. It is because 'the very acting of reducing one theory to another usually leads us to find flaws in the reduced theory as it was originally formulated, and to look for an alternative to it better suited to the reduction procedure'.⁷ Despite all these concessions, he concludes that broadly reductive relations hold between TD and SM, although we need to be careful about exactly what reductive relations they are.

I shall argue that Sklar's analysis of the alleged reduction of TD by SM is problematic in several respects. More specifically I will consider a few counterexamples to show that none of what Sklar takes to be the central features of progressive reduction in science holds in the case of thermodynamics and statistical mechanics. Then I will suggest the broader conclusion that a more useful way of understanding the relationship between TD and SM is as collaboration and competition among alternative methodologies rather than reduction of one theory to another.

³) Sklar 1993

⁴) Sklar 1993, p. 334

⁵) Sklar 1993, pp. 334-338

⁶) See his discussion of Krylov. Sklar 1993, pp. 369-73 and compare this with what Krylov himself claims in Krylov 1979.

Three themes are prominent in Sklar's view on reduction: unification, identification, and the importance of background reduction.

Unification. Sklar takes the progress of science to be 'marked by the continual success of attempts to unify a greater and greater range of phenomena in more and more comprehensive theoretical schemes'⁸. One way of achieving this unification is reduction, with 'the full range of phenomena handled by the reduced theory now being handled by the reducing theory'⁹.

Identification. Sklar contrasts two kinds of concepts-bridges between different theories: mere correlation and identification. For example, in the reduction of optics by classical electrodynamics, we can say either that each electromagnetic wave is accompanied by a light wave, or that light waves are nothing but electromagnetic waves. According to Sklar, in the first case we will wonder about the reason for these associations and consequently seek some explanation of them. On the other hand, the second case does not call for explanation. Sklar sees reduction by identification as one of the ways of unifying our theories of the world. Naturally, he argues, theoretical identification should be preferred to identification in reduction.¹⁰

Finally, this identificatory reduction takes place in a background programme of reduction that aims for 'the reduction of the theory of macroscopic matter to its micro-constituents by the identification of the macroscopic entities as structured out of microscopic entities'¹¹.

2.3 Unificatory Reduction of TD to SM?

One should be puzzled about the claim that SM unifies TD for it looks the other way around if we focus not on the principles of the two but rather on their concepts. What identificatory reduction boils down to in the case of TD and SM is finding out case-by-case SM's equivalents of TD's general concepts in every specific kind of

⁷) Sklar 1993, p. 339

⁸) Sklar 1993, p. 333

⁹) *Ibid.*

¹⁰) Sklar 1993, p. 340

¹¹) Sklar 1993, p. 341

thermodynamic system. For instance, TD's temperature is identified with the average kinetic energy of molecules in a particular ideal-gas in a box, and with the some function of the energy density of photons for radiation in a particular cavity.¹²

It seems more faithful to the spirit of theoretical unification, then, to say that TD's general concepts, such as temperature, 'unify' diverse SM manifestations across radically different systems. TD's concepts have every characteristic that advocates for unification would long for: truly general concepts applicable for a whole range of phenomena, simple and clear relations among themselves, etc...

In response one might admit that there are ambiguities in the concept of theoretical unification. The kind of unification Sklar takes as progress in science is not just the minimisation of theoretical concepts while maximising the coverage of phenomena under those concepts. So rather the desirable unification should give us, roughly speaking, the more correct ontology of the world with a more accurate description of it.¹³ After all Sklar believes that in unificatory reduction we often find the reduced theory flawed in one way or another, and in need of being corrected by the reduced theory.¹⁴ SM as a more correct theory should explain why TD is wrong in some of its results as well as why it is right in others.

So one way of cashing out the claim that SM unificatorily reduces TD is the following:

- (1) SM gives us, if not ultimately accurate, more accurate descriptions than TD.
- (2) SM can *correct* shortcomings of TD, but not vice versa.

¹²) Notice here that we are not making just token-token identifications, but rather context-conditioned type-type identifications.

¹³) In fact Sklar clearly shows his preference for this kind of unification when he criticises Kemeny-Oppenheim's *systematicity* approach in reduction which illustrates nicely the alternative concept of unificatory reduction as 'less theoretical vocabulary with more observation data'. Cf., Kemeny and Oppenheim 1956.

¹⁴) Sklar 1993, p. 339

But, I will argue, SM is not always more accurate than TD as we will see in the following examples. Moreover it is in fact corrected by TD to describe certain phenomena as accurately as TD does.

In TD, pressure and temperature are *intensive* variables, not depending on the size of the system. This implies that they are independent of the shape of the system. On the other hand, internal energy and entropy are *extensive* variables, proportional to the size of the system. This implies that if we conceptually split a system up into two parts, total internal energy and entropy are just the sum of those parts. The truth of this distinction can be tested experimentally. As long as it is experimentally valid, SM should be able to reproduce the distinction if it is to be as accurate in this respect as TD.¹⁵

But SM's treatment of *finite* systems cannot exactly reproduce the distinction. Roughly speaking, TD's distinction of intensive/extensive variables is tantamount to ignoring the surface effects; TD deals with only the bulk properties of its systems. But surface effects do contribute to the partition function of a finite system, and consequently to all SM quantities.

The way for SM to evade this difficulty is to consider *infinite* systems, to take the thermodynamic limit. Basically, the infinite size makes shape irrelevant for SM's pressure. Similarly the infinite system with realistic¹⁶ interactions among its constituents makes SM's entropy and internal energy truly extensive because the ratio of the summation of bulk interactions to that of surface interactions goes to zero in the thermodynamic limit. But surely real systems, however large they may be, are still finite. So here we have a clear case where TD 'corrects' SM so that SM can accommodate certain experimentally verified phenomena.

But one might argue that TD's distinction of intensive/extensive variables is only approximately valid rather than exactly valid even in the systems which TD can successfully deal with. SM's treatment of finite systems as finite is actually correct, but owing to the residual surface effects, too cumbersome to use in practice. That's

¹⁵) See Griffiths 1972.

¹⁶) Mostly it means short-ranged.

why we usually use the infinite system formalism of SM after taking the thermodynamic limit. The analogy here might be the case of Newtonian mechanics and the special theory of relativity. The mass of a moving object isn't actually constant, but can be regarded as constant for most practical purposes provided the speed of the object is much slower than that of light. Likewise TD's distinction of intensive/extensive variables is not exactly correct; nevertheless it is simple, versatile, and therefore worth keeping in SM.

Personally I do not think TD is approximately accurate in the sense that Newtonian mechanics is.¹⁷ But I shall not pursue this point for there is a case that looks hopeless to understand in terms of 'being approximately accurate'.

There are phase transitions around us: water boils to vapour, ferromagnets spontaneously align themselves along a certain axis, etc. TD has no intrinsic problem in dealing with these phenomena, but Yang and Lee showed in 1952 that no SM treatment of *finite* systems can have phase transitions. The introduction of infinite systems into SM is a necessary (but not sufficient) condition for describing phase transitions within the SM formalism.¹⁸ This has to do with the analyticity of SM's partition function for finite systems; the abrupt phase-changes of a system are represented by SM as singularities of its partition function. But no partition function of a finite system can have these singularities; only infinite systems can.

The important point is that the difference here between infinite systems and finite systems is not just quantitative but qualitative: the (non)existence of phase transitions. Unlike the intensive/extensive variables case, SM of infinite systems cannot be regarded as an 'approximately accurate' substitute for the 'accurate' SM of finite systems. Moreover here 'taking the thermodynamic limit', despite its clear unrealistic posits, actually improves the accuracy of SM in describing phase transitions. And the addition of this procedure to SM is motivated and justified by TD. Clearly TD corrects SM again in the region where it is more accurate, not vice versa.

¹⁷) Unlike the Newtonian mechanics case, there is no well-defined approximation procedure taking us from statistical mechanics to thermodynamics. The surface effects generally do not smoothly vanish as the size of a system goes to infinity. Moreover the justification for ignoring surface effects in a given finite system is sensitive to the specifics of the system.

¹⁸) Yang and Lee 1952

Consequently the (desirable kind of) unificatory reduction of TD by SM is not attained. What we find in actual practice is mutual correction rather than unidirectional unification.

2.4 The Case of Temperature: Reduction or Alternative Methodologies?

I shall use the case of temperature to illustrate the following theses (of course much more needs to be said to defend them in detail):

(TH1) Sklar's analysis of temperature as an example of identificatory reduction is problematic.

(TH2) TD and SM are not theories at some particular level, but general methodologies.

(TH3) A better way to look at the relationship between TD and SM is to take them as alternative methodologies.

Sklar divides the temperature of SM into two kinds: first there are SM concepts which are correlated with TD temperature in each specific kind of individual system, such as temperature as the mean kinetic energy of molecules of a particular ideal-gas at a particular time. Then there are SM concepts applicable to ensembles such as a defining parameter of a canonical ensemble.¹⁹

Sklar admits that there are subtleties in relating these two kinds of SM temperature concepts, but I won't discuss them. Nevertheless he does think that 'insofar as it is temperature as the instanced property of a particular system ... there seems nothing to block a *strict assertion of the identity* for that particular system with the appropriate microscopically characterised feature of it instanced at the same time'²⁰. And he suggests the determinable-determinate relationship (such as colour-red) to be the right sort of relation for TD's temperature and its corresponding microfeatures.

But it seems odd to say that the temperature of a certain gas-system, which is a property of the system, is 'identical' to the mean kinetic energy of gas molecules in the same way that light is identical to electromagnetic waves. After all temperature,

¹⁹) Sklar 1993, p. 351

²⁰) Sklar 1993, p. 353, my italics.

whether it is used as a general theoretical term or as a more concrete one represented by individual thermometer readings, is not material stuff. And the literal identification of properties represented by concepts with different histories, different usages, and different measurement techniques takes a lot of metaphysical back-up before it becomes clearly intelligible, let alone plausible. So why give up on the metaphysically less cumbersome claim that temperature is ‘correlated’ with the mean kinetic energy of gas molecules?

As we saw, Sklar emphasises that when we have correlation rather than identification we need to explain the correlation. Interestingly the above case of temperature is exactly like correlation in this respect. In order to make sense of what temperature means in a given concrete system, we have to find out the suitable ‘correlates’ of temperature in the system. This job is not done by just identifying certain plausible quantities in the system, but rather by setting a model of the correlation and explaining (or making intelligible) the correlation.

Let us consider the case of an ideal gas as an illustration. In order to find out the suitable ‘correlates’ of temperature, we set up a concrete model where one mole of a gas is confined in a rigid rectangular container. We consider the collisions between gas molecules and the container wall, and thereby the resulting pressure upon the wall. From this we can derive the explicit formula for the pressure,

$$P_x = \sum_{\dot{x}>0} 2p_x \dot{x} n(\dot{x}) \quad (2.1)$$

where P_x is the pressure exerted on the wall in the x -direction, p_x , the x -component of the momentum of a gas molecule, n , the number of such molecules per unit volume. Then we can employ the standard statistical mechanics formalism to ‘correlate’ the quantities in the above summation with their ensemble averages. Employing these correlations, we get,

$$PV = \frac{1}{3} \left\langle \sum_{j=1}^{3N} p_j \dot{x}_j \right\rangle \propto E_k \quad (2.2)$$

where V is the volume of the system, $\langle \dots \rangle$, the ensemble average, and E_k , the mean kinetic energy of the system. Then, here comes the final step. We recall another well-known correlation in thermodynamics, namely, Boyle-Charles' law ($PV = kT$), which relates the pressure and volume of a system with its average temperature. Combining (2.2) and this correlation, we see that the temperature of a gas can be correlated with the mean kinetic energy of the gas system. When we want to take into account the interactions among gas molecules, we have to generalise (2.2). The result is the Virial theorem, relating the mean kinetic energy and the mean potential energy with temperature in a gas with interactions.²¹ As we have seen, we *explain* why temperature is correlated with the mean kinetic energy of the molecules in the gas-system by setting up a concrete model. We don't just identify them with each other.

Consider another case: the temperature of radiation in a cavity. Here, temperature is correlated with the pressure in a different way.²² Then, should we say, following Sklar, that the temperature is 'identical' to the fourth order root of the energy density over a constant? Certainly that would sound absurd. On the other hand, as soon as we start to think in terms of 'correlation', this absurdity disappears. In fact, the correlation between the temperature of radiation in a cavity and the pressure of radiation can be explained again by setting up a concrete model of a radiation system. And here the famous Stefan-Boltzmann law plays an important role in correlating these two quantities.²³ The absurdity of simple identification emerges again more clearly if we remember that in this system pressure is a third of energy density²⁴. Then should we say that temperature in the radiation-system is somehow 'identical' to suitably modified pressure?²⁵ Certainly temperature being functionally 'correlated' with pressure is more natural.

Still it is true that we have an intuition that molecular kinetic energy has more than mere correlation to do with temperature in a gas-system. I guess the idea behind this intuition is that gas-molecules in an ideal gas are somehow 'responsible' for its temperature reading as well as for its pressure reading. Similarly we can say photons

²¹) Toda, Kubo and Saitô 1983, pp. 12-17

²²) $T = (u/\beta)^{1/4}$, where T is temperature, u , the energy density of radiation, and β , a constant.

²³) Toda, Kubo and Saitô 1983, pp. 74-78

²⁴) $P = u/3$, where P is pressure.

²⁵) $T = (3P/\beta)^{1/4}$

in a radiation system are responsible for its temperature and pressure readings. It will be interesting to clarify what this ‘responsible’ means. But identification seems the wrong approach. Rather they are more or less (causal) correlations.

Even Sklar himself is not very consistent about this point. In one place he talks about ‘the microscopic feature of the system *associated* with its temperature’²⁶. Yet in another, he says ‘the appropriate microscopic feature of the system with which to *identify* temperature’²⁷.

In order to understand the complicated relations between various concepts of temperature, I suggest we distinguish four different concepts of temperature.

TD-temperature₁: TD’s theoretical concept, absolute temperature related to the efficiency of an ideal heat engine²⁸

TD-temperature₂: TD’s phenomenological concept, thermometer reading of an equilibrium system

SM-temperature₁: SM’s theoretical concept, the derivative of internal energy by entropy, or physically speaking, the tendency-parameter of change of internal energy with respect to the number of accessible states

SM-temperature₂: manifestations of SM-temperature₁ in an individual system

The determinable/determinate relationship might hold between TD-temperature₁ and TD-temperature₂, and/or SM-temperature₁ and SM-temperature₂. For a sort of *generic/specific* relation holds in each pair. Historically, TD-temperature₁ was chosen among infinitely many absolute temperatures in such a way that TD-temperature₂ as measured on the Celsius scale could approximate it.²⁹ And, when we try to find SM-temperature₂ in a SM system, we use our physical intuition based on SM-temperature₁ to think hard to find what the manifestation of this abstract concept in this particular

²⁶) Sklar, *Op. cit.*, p. 352, my italics.

²⁷) *Ibid.*, p. 353, my italics.

²⁸) $\epsilon = 1 - T_1/T_2$, where ϵ is the maximum efficiency of an ideal heat engine operating between absolute temperatures T_1 and T_2 .

²⁹) Ter Harr and Wergeland 1966, pp. 20-3.

system could be. Still it is not similar to colour; for instance, a single system can manifest several kinds of SM-temperature₂ simultaneously (consider a system of a gas and electromagnetic radiation being in equilibrium with each other), while this is not possible for colours.

But the relation never holds between TD-temperature₁ and SM-temperature₂. The generic/specific relations do not hold for them because we do not use our intuition of the efficiency of an ideal heat engine to find the SM-temperature₂ in a concrete system. The correct relation between TD-temperature₁ and SM-temperature₂ is correlation rather than identification. This correlation is obtained in more than one ways. Either we correlate TD-temperature₁ with SM-temperature₁ on the theoretical level, and using the correlation between SM-temperature₁ and SM-temperature₂ in order to get the correlation between TD-temperature₁ and SM-temperature₂. Or you may try to correlate directly SM-temperature₂ with TD-temperature₂ in a given concrete system first, and thereby establish indirect correlation between TD-temperature₁ and SM-temperature₂ in that system. In fact this direct correlation between two concrete temperature concepts can be used to calibrate SM-temperature₂ in a familiar system using pre-established TD-temperature₂. This calibration is instrumental for SM to be operationalised.

Whatever path we may take, the correlation may not be found for some cases because it is not always possible to define the two theoretical temperatures in a given physical system. There are cases such as near-equilibrium chemical interactions where you can define thermodynamic temperatures but not statistical mechanical temperatures. We can define TD-temperature₁ in those situations as long as we can make sense of local equilibrium and of local entropy production. On the other hand, the number of accessible states is not well defined in this collection of open, localised sub-systems.³⁰ Consequently, SM-temperature₁ is not well defined. There are cases such as (highly) non-equilibrium phenomena where the number of accessible states is unambiguously defined. In these cases, SM-temperature₁ is well defined and clearly intelligible, but not TD-temperature₁.

³⁰) Kondepudi and Prigogine 1998

These observations point to a more general fact: TD-temperature₁ and SM-temperature₁ belong to two conceptually alternative theoretical schemes. In each scheme, they guide us to find the concrete counterparts of themselves in individual systems, namely TD-temperature₂ and SM-temperature₂. This point is illustrated by another example Sklar calls ‘concept-extension’: the negative temperature. TD-temperature₁ cannot be negative for efficiency cannot be over 100%. But SM-temperature₁ can be negative since in some, (highly) *non-equilibrium stationary*, two-level spin systems, you can make more ordered states by putting more energy into the system.³¹ There is nothing fancy about SM’s negative temperature because it naturally follows the definition of SM-temperature₁. There is no shame about TD’s non-negative temperature because as an equilibrium concept, TD-temperature₁ cannot be applied to highly non-equilibrium situations.³² Each concept is valid under its own prescribed conditions. SM-temperature₁ is an alternative to, rather than an extension of, TD-temperature₁.

2.5 TD and SM: Alternative, Competing Methodologies of CMP

In fact physicists use TD and SM as two distinct, not necessarily incompatible, ways of doing physics: two alternative methodologies. This is surely a part of the account of why TD and SM do not fit into the usual object-specific classification of physics, including atomic physics, condensed matter physics, laser physics.

As methodologies, models of TD and SM are widely used across these fields provided the interpretation of the models is satisfactory. What TD and SM amount to in specific applications critically depends on which interpretation the model under consideration assumes. Many philosophers tend to think of TD as a macroscopic theory and of SM as a microscopic theory. But this is not correct. Many models in SM are interpreted in dramatically diverse ways. For instance, the ‘spin’ of the Ising-Lenz model is interpreted as a microscopic portion of a ferromagnet, or as a macromolecule of haemoglobin, or as lots of other things. Moreover all of these interpretations, as long as they are successful, are equally legitimate. There is no *single correct interpretation* of ‘spin’. Consequently, models of SM are not microscopic but usually abstract: their contents are determined only through their applications (or

³¹) Toda, Kubo and Saito 1983, p. 40

³²) Although it can be applied to near equilibrium phenomena when generalised.

interpretative models).³³ There is no reason why SM should always have microscopic interpretative models. Rather SM is a general methodology with lots of abstract models that can be applied to explain various levels of phenomena.

On the other hand, TD is not macroscopic, but rather phenomenological in the sense that it starts with certain fundamental postulates the validity of which can be ascertained only by examining the phenomenon under consideration. The distinction of intensive/extensive variables in equilibrium TD is a good example. So is the extension of this distinction in non-equilibrium TD. The starting point of non-equilibrium TD is to *assume* that ‘local equilibrium’ is established in the system, which requires intensive variables of equilibrium TD to be variables of positions and time, and extensive variables, density-like quantities.³⁴ Roughly speaking, this assumption is tantamount to postulating the validity of a hydrodynamic treatment, the existence of smoothly-varying quantities averaged over *microscopically large but macroscopically small* regions of the system.

These observations point to the conclusion that TD and SM are not theories at some fixed level, but general methodologies. They often collaborate, as in the proof of the existence of phase transitions in the two-dimensional Ising-Lenz model³⁵, or compete, as in the early theory of superfluidity.³⁶ The relationship between TD and SM is more usefully understood as fruitful interactions between alternative methodologies rather than of one reducing the other.

Perhaps that is why Sklar emphasises the importance of background reduction: it can make the otherwise separate TD and SM talk about the same objects, ‘macroscopic, but microscopically structured’. Still his reduction fails because TD and SM *per se* are neither microscopic nor macroscopic. But, one might hope to ‘reduce’ all models of TD macroscopically interpreted by certain models of SM microscopically interpreted. I want to point out though that this ‘stretched’ sense of reduction cannot be

³³) I discuss this point at length in my “Models in Condensed Matter Physics”, MS.

³⁴) See Kondepudi and Prigogine 1998, chapters 15 and 16.

³⁵) Peierls in 1936, based on thermodynamic stability considerations, showed that there exists a phase transition at $T \neq 0$ in the two-dimensional Ising-Lenz model. Then Griffiths in 1964 proved it in a SM rigorous way.

³⁶) The competition between Tisza’s two-fluid model (TD approach) and Landau’s elementary excitation model (SM approach) is documented in Hoddson et al 1992.

guaranteed merely by hopes for the background reduction programme. It is a substantial question, and therefore should be carefully examined with respect to concrete examples.³⁷ No matter what the verdict is, however, it is a separate issue from my point that TD and SM as such are alternative methodologies.

³⁷ For instance, with respect to the examples in condensed matter physics, or theory of chemical reactions.

CHAPTER 3

MODELS IN CONDENSED MATTER PHYSICS

3.1 Introduction

In chapter 1, I argue that model-based understanding should be understood as a two-stage process: understanding of a model and understanding of its target phenomenon through an interpretative model of the model. In this chapter we will see how physicists attach a number of interpretative models to a given abstract model in order to understand a wide range of phenomena. Or they might opt for a single interpretative model for a given CMP model, and take the model-plus-interpretative model as our canonical representation of a particular phenomenon; to put it using a terminology of chapter 1, the model-plus-interpretative model complex becomes one of our substantial theories. I discuss in chapter 2 the competition and the collaboration between two different methodologies in CMP. In this chapter we will see with more concrete examples how physicists combine different methodologies, or more generally incongruent model-building components from different sources, to achieve a consistent and unified model-based understanding of the world.

I shall introduce several types of models in CMP in order to investigate the following two issues: (1) how model-based understanding is actually obtained. (2) how model-building is governed by methodologies or by the cooperation of several methodologies at the same time.

The essence of my answer to (1) in chapter 1 was that model-based understanding is obtained through the understanding of the model itself and through its interpretative models. In order to see more clearly the implications of this simple answer, I shall introduce two model-types: *free* models and *entrenched* models. The distinction between these two model-types is based on the relationship between the intrinsically abstract/formal structure of a typical CMP model on the one hand and its interpretative

models on the other. We shall see that the difference in this relationship brings about the difference in the kind of model-based understanding we can get from each model-type.

As I briefly mention in chapter 2, physicists often feel free to combine different resources in a single model. They may use more than one general theory or more than one level of descriptions. In order to see the fruitfulness of this ‘hybridisation’ in the CMP modelling, I shall introduce another set of model-types: *pure* models and *hybrid* models. This distinction is based on the constitution of the building-blocks of a model. We will see in my discussion of these model-types that, despite widespread belief in the virtue of universal explanation, the ‘hybridisation’ modelling strategy is essential for the consistent and complete understanding of the many-body phenomena.

I need to make some preliminary remarks about models of CMP in general before I get into the discussion of my classification. First, I should mention that the following classification is not meant to be exhaustive. After all, CMP is a quite large research field in both its topics and its methods. It comprises many types of theoretical modelling about all sorts of phenomena, materials, and interactions. We should not expect that a few kinds of model-types cover all these variegated activities. Nevertheless I claim that the substantial part of condensed matter physics modelling can be characterized by aforementioned two sets of model-types: free versus entrenched models and pure versus hybrid models.

In the following discussion most of examples of each model-type will be chosen from a particular research field of CMP, namely the theory of phase transitions. This choice is not accidental: to compare the theoretical virtue of pure models to that of hybrid models in the same field will help us clearly see the distinct features of each model-type. Moreover the spectacular critical behavior which many-body systems show during phase transitions are arguably *the* representative phenomena of CMP. CMP can be characterized as ‘many-body physics with strong interactions among a huge number of entities’ as many condensed matter physicists themselves often put it.¹ Phase transitions

¹) For instance, see Anderson 1983, or Leggett 1985.

are possible only from the complicated and coordinated interactions among a really huge number of entities in (relatively) condensed states; they are results of unique ‘many-body’ effects in CMP. If we have some specific types of models that turn out to be successful in dealing with phase transitions, we have good reason to expect that the model-types will cover a lot of modellings in other areas of CMP. That explains why I choose most of my examples from the theory of phase transitions.

Even when I illustrate my points about models in CMP using examples from the theory of phase transitions, this should not be taken to indicate that these model-types are applicable only to this particular research area. It does apply to many other areas in CMP although you can find a kind of correlation between research topics and types of model. For example, pure models are dominant in theories of phase transition, while hybrid models, more common in kinetic theory. There are some good reasons for these correlations and I will come back to it later.

3.2 The Role of the Interpretative Model: Free and Entrenched Models

I now propose the following distinction among CMP models based on the relationship of a given model with its interpretative models.

- A *free* model is associated with an *abstract* Hamiltonian that has the representations of the basic *uninterpreted* entities of the model and of the *primitive* interactions between them. Or, to put it differently, the relationship between a free model and its interpretative models is rather loose; there are usually many legitimate (and often successful) interpretative models for a single free model. In this sense, a free model is *free* in its relationship with interpretative models.
- An *entrenched* model is associated with a *relatively intuitive* Hamiltonian that has the representations of the basic *canonically-interpreted* entities of the model and of the *canonically-understood* interactions between them. Or again, to put it differently, the relationship between an entrenched model and its interpretative model is relatively tight; there is a historically well-rooted (and proved to be successful) interpretative model to a given entrenched model. In this sense, an entrenched model is *entrenched* in its relationship with its single interpretative model.

To say Hamiltonians of free models are abstract is not just to point out that they are written in a highly mathematised form and as a result we have no clear concrete intuition about them. In that sense, every Hamiltonian, arguably, turns out to be abstract. The additional abstract aspect of the Hamiltonians of free models is that they have *no uniquely legitimate* interpretative model. In other words, there is no point in talking about *the* canonical ‘picture’ that a given free model tells about the world. Usually a free model is associated with many different (even incompatible) but equally ‘good’ concrete representations of various phenomena.

On the other hand, the ‘relative intuitiveness’ of the Hamiltonian of an entrenched model does not mean that anybody can understand it without any difficulty. Rather, it means that each term of the Hamiltonian of an entrenched model is usually associated with well-known canonical interpretations so that physicists share more-or-less fixed physical intuitions about the model. Consequently, an entrenched model tends to have only one legitimate interpretative model and concerns a single phenomenon in the world.

The basic entities of a free model are utterly uninterpreted. For instance, a widely used term, ‘spin’ as a name for the basic entities of many free models does not have any intrinsic connection with well-known quantum mechanical spins. (That is why I am using scare quotes around it.) To take a specific example, ‘spins’ in the Heisenberg model can be either classical or quantum depending on how many degrees of freedom we may set for them in a particular interpretation.²

The ‘spins’ in the free models of CMP are generally interpreted as a kind of coarse-grained variable that is obtained from some averaging procedure over macroscopically few but microscopically many degrees of freedom in many-body systems. This is correct as long as such an averaging procedure is relevant and makes sense in that particular physical system. But, still, how to understand ‘spin’ in each concrete physical system is a non-trivial task and is not specified by giving just a free model. It is the role of

²) Rushbrooke, Baker Jr. and Wood 1974.

satisfactory context-specific interpretative models to provide for its free model the recipe for correctly identifying ‘spins’ in each physical system.

I am not denying that in some cases the inventor of a certain free model can have in her mind some particular microscopic picture to represent by the model. Lenz and Ising seemed to have a sort of quantum mechanical picture of ‘spins’ in their Ising-Lenz model.³ One could say that the Ising model has experienced a transition from being an entrenched model to a free model. But this fact, although historically significant, cannot be an intrinsic feature of the Ising model⁴ for we can see nothing in the model (at least in its present rendition), apart from *by fiat*, which prohibits us from having alternative readings of ‘spins’. In fact, as we will see in chapter 5, the real power of free models comes from the fact that the same abstract model can represent more than one physical system having drastically different microscopic pictures as long as we can equip it with corresponding, ‘successful’ interpretative models. That is why I claim that there is no unique ‘correct’ interpretative model for a free model.

Primitive interactions in a free model are again uninterpreted. Interaction-terms in the Hamiltonian of a free model are fully specified by three factors: entities involved in interactions, range of interactions, form of interactions or interaction strength. Now entities are uninterpreted in the way I explained before, and range and form of interactions are usually ‘artificial’ enough to allow no handy identification with either so-called fundamental interactions or familiar forms of phenomenological interactions such as the harmonic potential.

In many free models, the most commonly adopted interpretation of interactions is electromagnetic interactions. Nevertheless this does not mean that electromagnetic interactions offer the most natural interpretation of any given free model. Often we can represent by the same model Pauli’s exclusion principle, gravitational force, osmotic

³) See Brush 1967.

⁴) I will follow the physicists’ convention to call this model just the Ising model rather than the Lenz-Ising model, although the latter is historically more appropriate.

pressure, etc. if we have suitable interpretative models.⁵ If we recall the characteristics of the Potts model discussed in chapter 1, we can see now that it is a free model.

Moreover even in the case where some ‘rough’ identification of primitive interactions with electromagnetic interactions is acceptable, the match is indeed rough, actually too rough to be regarded as the original ‘electromagnetic interactions between electrons and atoms’ in the form of a Coulomb force. Usually the range of interactions is delimited, such as ‘nearest-neighbour interactions’ or ‘next nearest-neighbour interactions’, to make interactions short-ranged, while it is a well-known fact that the long-range nature of the Coulomb force renders its theoretical treatment much more difficult.

On the other hand, the canonical interpretations of the entities and of the interactions of an entrenched model are usually agreed on among physicists. Unlike free models, an entrenched model starts from a more or less familiar Hamiltonian that seems to give us a reasonable microscopic description of the system. It is usually written in terms of so-called fundamental interactions (in other words, using some familiar interpretative models) among less abstract entities than those of free models.

For instance, when you write the Hamiltonian for the Bloch model⁶, each term of the model represents four main contributions to the metallic phenomena in a solid: free-electron part, free-ion part, electron-electron interactions part, electron-ions interaction part. Although we need lots of idealisations in order to build this model and to get useful results out of it, the identification of each term in the Hamiltonian is taken to be non-problematic. Also physicists have developed a set of well-established physical intuitions about the Bloch model as a model of metallic phenomena. Consequently the tie between the Bloch model and its interpretative model for metallic phenomena has been so tightly made now that it is difficult to distinguish one from the other in the standard presentation of the Bloch model. So the Bloch model is an entrenched model.

⁵) See Fisher 1981

⁶) For more on this model, see chapter 4.

In this sense, each entrenched model has a ‘rough’ identification of its entities and interactions built into it. Why do I call it a ‘rough’ identification? It is because the intended interpretations of entrenched models are not aimed at picking up a unique entity or the exact form of interactions, but rather at restricting more tightly than the implicit conditions of free models do what could be the proper target systems of a given entrenched model.

For instance, ‘particles’ in the classical fluid model, a very popular form of entrenched model dealing with liquid-vapor phase transitions, are typically idealised, employing the classical image of non-permeable point-particles. Nevertheless, this feature is not very significant after all, for what we are actually interested in the model is not particles themselves but some coarse-grained density function-like quantities averaged over certain regions of intermediate size, neither microscopic nor macroscopic. So there is a two-step idealisation here: first from complex real particles to idealised point-particles, then second from these point-particles to a fluid-description of them. This second step is a non-trivial idealisation because it requires the emergence of a well-defined density-like function out of point-particles that cannot be true except when we take the thermodynamic limit for the system.

Interactions in entrenched models look quite familiar and promising for an obvious interpretation such as electron-electron interactions. But still we need to fill in several restrictions to these interactions to get desirable results. These restrictions range from fairly reasonable ones, considering their physical implications, to highly ad-hoc ones chosen for mathematical convenience or formal simplicity. After imposing these restrictions, we have ‘stylised’ interactions with canonical interpretations.

Whether a free model or an entrenched model, there is an important aspect of CMP models that is seldom explicitly stated when physicists present them. These are implicit general descriptions about target systems. These descriptions are *implicit* because many physicists think that presenting a Hamiltonian is enough to fix a model. But they are integral parts of models because they *constrain* which physical system could be modelled

by a given model. For instance, many free models in the theory of phase transition are supposed to be on some sort of underlying lattice with ‘spins’ fixed on each lattice point; the dynamics of the model is exhausted by changes of ‘spin’-values. This fact cannot be read off from the Hamiltonian of a free model.

Now all quantities in which we are interested from this dynamics are equilibrium quantities and this fact also assumes a lot of things: for instance, the existence of equilibrium quantities in the free model, the plausibility of the application of the free model in this physical system. Moreover how to characterise certain features of a free model is not determined by just presenting the Hamiltonian for it. For instance, in order to characterise how ‘ordered’ a system is, we have to find out a way of defining the order parameter in a given free model. But the formal structure of the model is rarely enough to determine the order parameter, and it should be supplied by the implicit descriptions of the free model. But order parameters may be the most important quantities in the theory of phase transition. Related to this, we need to have some general prescription for what would be the nature of ordered phase of a given free model. Usually to find out the nature of ordered phase is one of the most difficult problems of CMP except in a few trivial cases.⁷

The role of implicit description in an entrenched model is no less important than it is in a free model, if less conspicuous. For instance, many entrenched models in metallic phenomena implicitly require that the basic entities of their models (electrons and ions) will be confined within a metal, despite the collisions. Usually the implicit descriptions of an entrenched model are incorporated into physicists’ shared physical intuition about the model. More generally, the implicit descriptions of any model in CMP are incorporated into its context-specific interpretative models.

Now I shall discuss an example of free models, a general form of ‘spin’-lattice models: the ‘spin- s Heisenberg model’. Before presenting its Hamiltonian, I wish to make one important point clear: this model is a sort of post-hoc construction from various pre-

⁷) Anderson 1983

existing ‘spin’-models. In other words, this model itself with its full generality has never been proposed in order to explain particular phenomena; rather it’s main purpose is to offer a general framework where already known free models can be discussed in a more systematic way.

$$\text{“Spin-}s \text{ Heisenberg model”}: H = \sum_{i < j} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j + \mathbf{H} \cdot \sum_i \mathbf{S}_i \quad (3.1)$$

Here the Hamiltonian, H , contains two terms. The first term represent spin-spin interactions with \mathbf{S}_i (\mathbf{S}_j) being the i -th (j -th) vector-‘spin’ \mathbf{S} in the m -dimensional lattice. J_{ij} is the interaction strength between the i -th ‘spin’ and the j -th ‘spin’ and the summation over all possible pairs avoids double counting by constraining the sum only for $i < j$ cases. The second terms represent spin-external field interactions with \mathbf{H} being the external vector potential. Despite its similarity to the magnetic vector potential, the meaning of \mathbf{H} is not fixed. For example, if we adopt the lattice gas interpretative model of a particular form of the above model where ‘spins’ represent the occupancy (or vacancy) of a particle on the underlying lattice, \mathbf{H} becomes chemical potential.

The implicit description of target-systems of this model is the following: There is an *infinite* array of lattice points having spatial dimension, m , and on each lattice point, there are ‘spins’, the only dynamic variable of this model. The existence of a thermodynamic limit, or the existence of well-behaving thermodynamic quantities, is assumed, as well as the adequacy of an equilibrium treatment. What these assumptions amount to is that we are not interested in the actual dynamic evolution of each ‘spin’; instead we assume that sooner or later⁸ thermodynamic equilibrium will be achieved in the model so that we can use the standard thermal average in the model.⁹

There are several more subtle points in this model. The interaction strength, J_{ij} , is conventionally taken as positive. Then the Hamiltonian would have a lower energy if the

⁸) Again how soon it will be is a sort of questions free models are supposed to answer.

⁹) Not all free models have this assumption. Physicists actually proposed several free models which do not have well-behaving thermodynamic limits or the thermodynamic equilibrium in a standard sense. But physicists then have a burden of showing why those free models are *useful* enough to study.

inner-product of a pair of spins is positive owing to the minus sign in front of the spin-spin interaction term. This means, again implicitly, that the ordered phase of this model is something like all spins lining-up in the same direction. From this we can guess that the relevant order parameter should be some kind of average over spin-values. All this implicit information constrains our decision on what could be 'good' interpretative models for the spin- s Heisenberg model. For instance, we would require that the interpretative model should describe two-phase transitions such as the order-disorder transition or the metal-insulator transition, not multiple-phase transitions such as the Kosterlitz-Thouless transition in thin films.

Several examples of entrenched models will be discussed in chapter 4. It is noteworthy that most of models in the theory of phase transition are free models, while most of models in theory of metals are entrenched models. In order to appreciate the meaning of this correlation, let us recall the most important difference between a free model and an entrenched model. It has to do with whether there is any canonical interpretative model for a given abstract model.

If there is a canonical interpretative model, then the model is entrenched; if not, the model is free. When an entrenched model is entrenched enough to make it difficult to distinguish the model itself from its canonical interpretative model, we tend to forget the abstract nature of the model and treat it as an concrete 'picture' of the particular phenomenon it describes. When the model is also successful, we take it (in fact, the complex model-plus-its interpretative model) to be our best interpretation of the phenomenon under consideration. The model comes to belong to our stock of substantial theories. So the fact that most of the models in the theory of metals (and solid-state physics in general) are entrenched models shows that physicists tend to develop specific representations of individual metallic phenomena. This does not mean that an entrenched model developed as a representation of, say, the electric conduction in a metal cannot be employed to represent a totally unrelated phenomenon, say the vaporisation of mercury with the help of a different successful interpretative model. But then the virtue of the model would change from being our best representation of a particular phenomenon to

giving us a physical insight into the unifying connections between various different phenomena, an important virtue of a free model.

Similarly, there is some important reason why there are so many free models in the theory of phase transition. As we will see in more detail in chapter 5, one of the major task for the theory of phase transitions is to explain why almost identical patterns of phase transitions occur across radically different materials, apparently without good *prima facie* reasons. (The Phenomena of Universality) For instance, striking asymptotically identical pattern has been empirically found in the water-vapour transition and the order-disorder transition in ferromagnets, despite the fact that the microscopic composition and the nature of the interactions of each case are very different.

Now a free model may have many possibly *mutually incompatible*, but still successful, interpretative models, even for the same phenomena in principle.¹⁰ Each interpretative model gives us a coherent ‘picture’ of the basic entities and interactions among them for the phenomenon of interest. Still, there can be drastic incompatibility between two, otherwise satisfactory interpretative models of a single free model. It is easy to see then why free models are ideal for describing the phenomena of universality: the freedom of choice in its interpretative models enables a free model to accommodate a number of radically different phenomena within itself.

CMP studies a huge number of phenomena. We expect CMP to provide us with substantial theories of these phenomena that reflect the specificity and the diversity of many-body phenomena. These substantial theories are provided by entrenched models. On the other hand, we expect CMP to give, not just an unmanageable collection of numerous (entrenched) models for each minutely classified phenomenon, but also a unified understanding of certain, otherwise perplexing, patterns in many-body phenomena that appear across lots of different phenomena. This part is done by free models. Here we can see that two different types of CMP models serve different

theoretical purpose; representation of specific physical mechanism and unifying understanding of the diverse phenomena.

3.3 Patchwork Model-Building: Pure and Hybrid Models

There is another fruitful way of distinguishing models in CMP based on different modelling strategies. So I now propose the following model-types:

- A *pure* model is homogeneous in its constitution; either the building-blocks of the model all come from a single general theory such as classical mechanics, or a single level of description, such as macroscopic, is consistently adopted within the model.
- A *hybrid* model is heterogeneous in its constitution; either the building-blocks of the model come from more than one general theory as in the case of semi-classical models, or more than one level of description (for instance, both macroscopic and microscopic) are mixed within the model.

Pure models are pure in their model-building strategies. When we build a pure model, we generally try to *exhaust* the model-building potential of a given general theory. We use all known techniques and heuristics of a given general theory and try to make a successful model of its target phenomena. We may succeed or we may not. When we do not succeed, we need to look at other general theories for help; we turn our attention to hybrid models. Being pure about the level of description has a further reason. There is no doubt that we have to have consistency among various levels of descriptions, such as descriptions based on so-called fundamental interactions among fundamental particles and descriptions based on phenomenological interactions among macroscopically identified objects. Pure models do the foundational job for this consistency. They provide us our best descriptions on each level, which can then be put into the work of coordination. The second part of this job is done by hybrid models.

Hybrid models are heterogeneous in their composition in the sense that they combine in themselves several theories having distinct levels of description of a physical system, or,

¹⁰) It is rare that physicists entertain more than one interpretative model of a free model given a phenomenon at the same time. It is mainly because finding out a single interpretative model is difficult

sometimes, having different degrees of commitment from physicists. The range of theories that could be incorporated into a single hybrid model is fairly diverse: it could be a very phenomenological scattering formula, a more stable theory from other subfields of physics, and so forth. More specifically there is a very special way for a hybrid model to be heterogeneous, namely ‘semi-classical’. Often, but not always, a certain part of the microscopic description of the system is dealt with in a quantum mechanical way, while the other parts, such as the interaction process between the system and our measurable variables, is described in a classical way.

Let us see those points through some examples. The exchange interaction model of ‘spins’ is a hybrid models. ‘Spins’ in this model are ‘spins’ in the typical lattice pure-model’s sense. This hybrid model tries to provide one plausible link between abstract ‘spins’ and Coulomb interactions among electrons. It is well known that the wavefunction of two *free* electrons is degenerate: singlet and triplet states. But electrons are not free in real systems but strongly interact with each other in the form of Coulomb interactions. Now there is a finite energy gap between spin-singlet and spin-triplet states due to Coulomb interactions if we calculate the turn-on effect up to the first-order perturbation of it. Then we *define* this finite energy gap, J , as something resulting from spin-spin interactions and call it the ‘exchange interaction strength’ in the standard lattice pure model.¹¹

There are several important things to say about this model. First, the origin of these exchange interactions in this hybrid model is Coulomb interactions between electrons, not spin-spin interactions; real spin-spin interactions are too small to be influential in the first-order perturbation calculation. Second, this hybrid model seems to provide a genuine bridge principle between fundamental interactions between electrons and more phenomenological, exchange interactions in certain pure models. One might take this hybrid model to unravel some solid microscopic foundations behind the use of such a useful concept as ‘spin’ in pure models. These suppositions, however, are ungrounded:

enough, let alone two. But how to interpret a given free model *better* is always an important issue.
¹¹) Chaikin and Lubensky 1995, pp. 21-5

even in the ferromagnetic case, where this kind of hopeful interpretation is fairly plausible, spin-spin interaction terms do not represent electron *spin-spin* interactions, as I emphasized earlier. They describe Coulomb interactions among electrons. So the exchange interaction model of ‘spins’ does not provide a reductive explanation of the ‘spins’ in spin-lattice models in terms of quantum spins of electrons. ‘Spins’ in pure models cannot be regarded as streamlined quantum spins. If this hybrid model aims at showing some direct connections between quantum spins and ‘spins’ in pure models, it simply fails to achieve its aim. But, as I shall argue later, that was not the aim of hybrid models in general and certainly not in this particular case. Moreover, what we ultimately want from this hybrid model is not a model of two electrons but a generalization of this approach into the many-‘spins’ case, that is a spin-lattice model. After this kind of generalization is taken, as we saw in our exposition of free models, even the partial link between Coulomb interactions among electrons and exchange interactions of electrons will be lost, because now the model cannot be taken to apply exclusively to ferromagnet. Whatever justification this hybrid model might give to the use of the corresponding free model for ferromagnetic phenomena, it won’t give any comfort to other applications of the free model for different phenomena.

One nice example of hybrid models from the theory of phase transitions is the Ornstein-Zernike (OZ) model of classical fluids¹². This model is especially clear about the bridging between microscopically defined parameters and experimentally obtainable quantities. The model starts by introducing the concept of the molecular pair-distribution function, which has figured centrally in all subsequent work on liquids:

$$\langle v(dr_1)v(dr_2) \rangle = n_2(r_1, r_2)dr_1dr_2 \quad (3.2),$$

where $v(dr)$ is a random variable representing the number of particles in a volume dr centred at r and $n_2(r_1, r_2)$ is the pair distribution function representing the correlation between particles at r_1 and r_2 . The bracket means taking the standard thermodynamic

¹²) Ornstein, L.S. and Zernike, F. (1914) *Proc. Akad. Sci. Amsterdam* 17: 783.

average. From this pair-correlation function, we can define quantities that represent deviations from randomness in a homogeneous fluid case, which is:

$$h(r) = n_2(r_1, r_2) / n_1(r_1)n_1(r_2) - 1 \quad (3.3),$$

where $r = r_1 - r_2$ and $n_1(r)$ is the local density centred at r . We call $h(r)$ the pair-correlation function.

Up to this point we have only dealt with the hydrodynamic treatment of many particle systems. To get the OZ model, we need help from a quite different theory, the theory of light scattering. Here the important quantity is the structure factor: the normalized intensity of the scattered light as a function of q , the difference in phase between incident wave (wave vector k_0) and scattered wave (wave vector k_s). That is,

$$S(q) = I(q) / I_0(q) \quad (3.4),$$

($S(q)$, structure factor; $I(q)$, the density of the scattered light; $I_0(q)$, $I(q)$ for the case of random scatterers).

Now we have two very standard theoretical treatments for distinct fields, a theory of liquids and a theory of light scattering respectively. Here comes the important link between the two: since $I(q)$ is obtained from some average process over the scatterer, $S(q)$ can be written in terms of the pair-distribution function of the scatterer and consequently, of the pair-correlation function. Moreover there is a theorem in thermodynamics, the fluctuation theorem, which connects the pair-correlation function with the isothermal compressibility, K_T in the following way:

$$\text{(Fluctuation Theorem): } kT\rho K_T = 1 + \rho \int h(r) dr \quad (3.5);$$

but,

$$S(k=0) / \rho = 1 + \rho \int h(r) dr \quad (3.6),$$

therefore,

$$S(k=0) / \rho = 1 + \rho \int h(r) dr = kT \rho K_T \quad (3.7),$$

where ρ is the (global) average density. What we have done up to this point is to combine three formulae from different fields and establishing a connection between experimentally meaningful quantities and a fairly abstract thermodynamic quantity, K_T . The OZ model shows a typical procedure in hybrid models to achieve ‘bridging’ between and incorporating several theories into a heterogeneous whole.

The real essence of the OZ model is to conceptually divide the total correlation into two parts: a direct part arising from the interactions of one-pair of molecules and an indirect part arising from mediating effect of intervention of other molecules. Nobody knows what exactly this division means in physical terms or whether this division is plausible. But you are free to rigorously define certain quantities out of known quantities; in this case, a new direct correlation function out of a (total) correlation function. Why do we do this? It is because we think we can ‘reasonably’ assume that the new function, unlike the old one, is sufficiently short ranged to have moments. Then we can have, for small k (that is, in the domain where the direct correlation is expected to be dominant), the approximate form of the structure factor in a Lorentzian form, which boils down to:

$$S(k=0) / \rho (= 1 + \rho \int h(r) dr = kT \rho K_T) \approx C \xi, \quad (3.8)$$

where C is a constant and ξ is the correlation length. This is a remarkable result for it means that in a critical region where the correlation length diverges, the forward scattering also diverges, which is so-called the ‘critical opalescence’ phenomena.

Despite its name, the direct-correlation function does not represent the ‘direct’ correlation between a molecule at r_1 and another molecule at r_2 , since it is ‘defined’ to be the remaining part of the total correlation function minus a certain integral which is supposed to capture only the two-step indirect process.¹³ The role of this new function is to guarantee the physical reasonableness of the short-range assumption that is crucial for the

whole derivation. Let me call this a ‘mediating definition’ strategy. It is not unreasonable and it is certainly an ingenious move to make, but the definition is clearly dictated by our hindsight about the range of interactions in many-body physics. For instance, Ornstein and Zernike assume that the direct correlation function would be of short range even if the total correlation function were long ranged.

Consequently, what the OZ model shows is not the reducibility of the scattering amplitude in terms of microscopic structure factor. Rather it shows the *cotenability* of our basic picture of fluids with an empirically well-known phenomena, ‘critical opalescence’. But to achieve this cotenability through ‘mediating definitions’ means that the role of hybrid models is not take by physicists something like ‘derivations’. The value of the OZ model is to provide a physically feasible ‘many-body’ structure for a well-known macroscopic phenomena.

The distinction of models into pure and hybrid shows what are central research strategy physicists adopt in order to understand as many many-body phenomena as possible. The first stage tries to exhaust as much as possible the capacity of a given general theory or model-building methodology to give coherent representations of those phenomena that that particular general theory is successful in dealing with. This stage produces pure models as its results. The second stage is to co-ordinate the successful results of the first stage into a consistent picture wherever different representations overlap in their target phenomena. This co-ordination is done by hybrid models, and it contributes to improve our model-based understanding.

The different functions of pure versus hybrid models can be understood using Nancy Cartwright’s concepts of ‘patchwork’ and of ‘the dappled world’.¹⁴ Pure models are intended to cover the dappled world as much as possible with the help of a single/pure model-building methodology. The range of phenomena in the world successfully understood by one particular kind of pure models can be thought of as one big patch with

¹³) Domb 1996, pp. 104-5

¹⁴) Cartwright 1999

an irregular shape. Then there are a number of patches corresponding to different general theories and model-building methodologies, which cover the world, and some of them may overlap. Then we need to have a consistent way of co-ordinating different patches, and the hybrid models do this job. They make sure that the overlapping parts of the different patches are either jointly acceptable or at least jointly intelligible, as we saw in the case of the OZ model.

3.4 Conclusion

We saw that CMP uses various types of models (and their associated interpretative models) to contribute our model-based understanding. In the next chapter we will look at the dynamics of model-building process more closely.

CHAPTER 4

APPROXIMATIONS IN CONDENSED MATTER PHYSICS

4.1. Introduction

In the last chapter, we looked at how physicists use different types of models for different purposes in order to understand condensed matter phenomena. For instance, we have learned that a free model, unlike an entrenched model, serves not so much as a representation of the phenomena but as a flexible structure with well-developed, but still rather abstract, physical intuitions. In this chapter we investigate a dynamical side of the model-based understanding in condensed matter physics; that is to say, we will ask (1) how models are *built* helped by various kinds of theoretical resources and (2) how these models are *explored* to reveal their explanatory and predictive potentials. More specifically I am going to highlight one important feature of these model-building and model-exploring stages: *approximations*.

For condensed matter physicists, performing approximations is a major part of their job when they are trying to understand certain phenomena employing models. Naturally, one of the virtues of a good physicist is to know how to employ approximations appropriately. What I am interested in is what role physicists intend approximations to play in both the model-building and the model-exploring stages. In order to investigate this, I shall ask first: what is the *nature* of approximations in condensed matter physics? Are they really 'approximate' to something as the conventional account has it? From the answer to the first question we can come back to our original question, namely for what purpose do we use approximations in our theoretical understanding of condensed matter phenomena? Do we use approximations only for pragmatic reasons? To put the question differently, are they just cleverly designed mathematical devices that allow physicists, for example, to evade difficulties with insoluble equations? But let us see first what the conventional account says about these issues.

4.2. The Passive vs. the Interactive Picture of Approximation

Traditionally, approximations have seldom been a focus of attention from philosophers of science. The reason is rather simple: approximations have been regarded as a drawback in our ability to understand the world theoretically; they are practically useful, but nevertheless expected to be minimised as our theory advances. From this comes the idea that we need to look only at how theories are ‘ideally’ employed (i.e., without approximations) in explanation or prediction in order to grasp the fundamental structure of how scientific theories are applied to the world. Then we could expect that the ‘actual’ practice of scientists (i.e., with approximations) follow essentially the same structure as the fundamental one, perhaps only ineptly.¹

Here is a caricature of what this conventional view of approximations would amount to in condensed matter physics. You start from a (ideally) complete description of a given *real* material, say a sample of copper. Then you connect this description, using various bridge principles, with (again ideally) true theories. So the theories tell you which equations you should write down for this sample of copper, and bridge principles will provide appropriate interpretations to each term in the equations with suitable boundary conditions. Now you *rigorously* derive some theoretical results from the equations and compare them with the experimental results obtained from the sample. If we assume (ideally) that our ability to get the relevant information from real physical systems by experiment is perfect, the two results should perfectly match with each other.

But obviously we do not expect that we can write down the complete description of a condensed-matter system full of complicated interactions among a huge number of constituents. Neither do we think that our current theories are the final, true theories about condensed-matter systems. Moreover, even granted the truth of our current theories, the intractability of its equations for a typical condensed-matter system (say, the

¹) *Cf.*, Laymon 1998, especially his discussion of ‘Idealizations as Approximations’. But notice that Laymon’s account is much more sophisticated than what I call the ‘conventional’ view. (More on his views see below) Perhaps Hempel 1966 is more appropriate as implicitly endorsing the conventional view, but then he rarely mentions approximation.

N-body Schrödinger equation for a given sample of copper with N enormously large) would prohibit us from calculating any interesting quantity of the system.

Here comes the approximation. We may try to gather as many relevant facts about the system as possible and put them into our description of it; then hope that our descriptions ‘approximate’ the complete description of the system ‘well-enough’, whatever that would mean. Similarly, we might believe that our current best theory is ‘approximately true’, if not literally true in every detail, although we may never know which part of it is true and which part, false. Third, we may use a number of ‘approximation methods’ specially developed in order to evade the intractability of the equations so that we can get some results for the comparison with the theoretical results.

I want to set aside the much-debated issue of ‘approximately true’ theories in order to concentrate on the other senses of ‘approximation’: the ‘approximate’ description of the physical system and the ‘approximation methods’ in our theoretical treatment of the system. It is unquestionably true that *some* approximations employed in physics have the kind of features the conventional view assumes. For instance, for a small angle, $\sin\theta \approx \theta$, so we perform a linear approximation in our classical description of a pendulum: we replace $\sin\theta$ with θ . Likewise we have a number of standard series expansion techniques such as Taylor expansion in theoretical physics by which we obtain the approximate value of an intractable mathematical function in terms of a sum of manageable terms. But clearly that is not the case for *all* approximations in physics. There are lots of approximations physicists help themselves with, which are neither an ‘approximately true’ description of the phenomena nor a mechanical result generated by a mathematical technique. For instance, condensed matter physicists often make the so-called ‘free’ electron approximation where electrons are assumed to be under no external potentials, which is absolutely false for all electrons in the world and not even approximately true for most of them. And there the conventional view falls wide off the mark.

Ronald Laymon's work on approximation and idealisation is probably the classic discussion nowadays, and he agrees.² He discusses this issue in the framework of idealisation. First he (as most of philosophers of science) understands approximations in a narrower sense than many physicists may opt for; approximation is something really being 'approximate' to the true description or value. After assuming approximation to be a relatively unproblematic concept, he focuses on idealisation, a false, simplified description that scientists use for various purposes. Roughly speaking, the concept of idealisation is a familiar analytical tool for philosophers of science in understanding the practice of scientists that scientists usually call 'approximation'. As I am going to investigate the details of how physicists use 'approximation' in order to build a model as well as to probe it, I shall respect their usage. Moreover, as we will see later, by not adopting philosopher's narrow conception of approximation, we can learn a lot about the practice of physicists. So unless there is no chance of confusion, whenever I say approximation in this chapter, I always mean it in its broad sense including both the cases of approximation (in the narrow sense) and of idealisation.

Laymon starts his most recent discussion of idealisation with a popular, but simplistic view that the idealisations of science are just approximations in the narrow sense.³ But he is well aware of the fact that many individual instances of idealisations in science cannot be properly dealt with by this simplistic view. For instance it is not approximately true that the universe contains only two bodies as assumed in Newton's derivation of Kepler's law. Nor is an assumption Samuelson made for his influential model, 'An Exact Consumption-Loan Model of Interest with and without the Social Contrivance of Money' that workers lives consist of three periods and they each produce one unit of a single completely perishable output (say, 'chocolate') in the first two periods.⁴ So he suggests that we should think of individual idealisations as elements of some larger (often not known in advance) structure that converges on the truth. He claims that scientists typically use a 'monotonically convergent' sequence of idealisations, each idealisation being more realistic than the earlier ones in the sequence, and that the use of idealisations

²) See Laymon 1985, 1989, 1990

³) Laymon 1998

can be justified only if idealisations are used in this way. In other words, a good idealisation should come from a sequence of idealisations in which each idealisation is *comparatively* truer than its earlier ones.

Laymon's account is true of many approximations in physics, taking care of several difficulties the conventional view faces. As Lakatos shows in his memorable defence of the methodology of research programmes, one major strategies of a research programme is to employ systematically less and less unrealistic situations (or models) in each stage of research in order to overcome almost inevitable falsifications and move to more and more satisfactory theories. My question is whether Laymon's elaborated account of approximation is entirely adequate for understanding approximations in condensed matter physics. My answer is no. I claim that we need more a sophisticated picture.

Notice that even according to Laymon's account we are bound to lose some information by performing approximations. The loss could be due to a 'not exact but only approximate' description, or to 'not rigorously valid but only approximately correct' mathematical procedures performed on those descriptions. In any case the loss is thought to be damaging. For instance, we may blame the loss resulting from the use of approximations for any unsatisfactory match between experimental and theoretical results when we have great confidence in the theory involved. In sum the point is simple: more approximations will weaken our ability to understand the world theoretically. Or, to put it as a slogan, "*the less approximation, the better*".

I would like to emphasise this point by characterising various forms of the conventional account of approximations, including Laymon's, as the *passive* picture of approximations.⁵ The picture sees the approximation in physics as more or less faithful to its literal meaning, 'an object, description, situation, etc. that is very like something else

⁴) Samuelson 1958

⁵) I think many philosophers of science accept this picture about approximations in science, if only implicitly. For instance, the picture fits well the views of those who see idealizations as approximations in the narrow sense. See Laymon 1998. See also Balzer et al. 1987 and Laymon 1989.

but which does not have all its features or qualities'⁶. The nature of approximations in this picture is then characterised by the (hopefully small) departure from a (ideal) target and by the (unavoidable) loss of our knowledge about the target. For our discussions, the target is either the complete description of a real physical system or the rigorous solutions of the relevant mathematical equations. Notice the negative implication the use of approximations has under the passive picture: our science as currently practised is defective due to our use of approximations.

Given this negative implication of approximations, the passive picture sees the only role of approximations as a pragmatic one: an approximation is the last resource we go to when we have no alternative. The use of approximations mirrors our epistemic limitation in applying our theories to the world. This makes clear why I call the picture *passive*; it does not recognise any active role for approximations in our theoretical work. More specifically, it does not allow any possibility that employing an appropriate approximation in fact complements our theories in order to improve our ability to understand the world theoretically.

In this chapter I offer an alternative account of approximations in physics: the *interactive* picture of approximations. The interactive picture starts from the recognition that approximations play quite diverse roles in theoretical physics. Some of them clearly serve pragmatic purposes as the passive picture says. For instance physicists, confronting an intractable integration, often impose certain constraints on the functional form of the integrand in order to obtain some manageable result from the integration.⁷ Many of them however play various epistemic roles in building models, extracting the correct physical content of a given model, clarifying the valid scope of a model, isolating relevant causal

⁶) Collins Cobuild English Language Dictionary 1992, p. 68

⁷) One example is the 'factoring-out' method. A typical scenario might be the following: you encounter a formidable-looking integral, $\int f(x,t)g(t)dt$, with an (in many cases) unknown function, $f(x,t)$. Let us suppose that you are interested in $f(x,t)$. As long as $f(x,t)$ can be factored out of the integral, we might be able to integrate the remaining $\int g(t)dt$, or if not, label it as a constant α , and proceed with our theoretical investigation of $f(x,t)$. If we can be pretty sure from other information that $f(x,t)$ varies very little with t during the interval we are interested in, we may feel justified in doing this factoring-out approximation. But even *without* any proper justification about the slow variation of $f(x,t)$, we may just decide to go ahead for that is the only way we know to proceed in a given investigation of $f(x,t)$.

mechanisms, etc. These roles of approximations are epistemic because they add something to our knowledge about the physical systems in the world, rather than lose something from it.

Moreover the interactive picture claims that approximations fulfill these epistemic roles by *co-operating* with general theories such as classical electromagnetism; they do not just *passively attach to* them. So the name, 'interactive' picture. In other words the interactive picture takes the use of approximations to be a significant part of our theoretical activities which complements general theories epistemically as well as pragmatically. Under this way of understanding approximations, the less use we make of approximations does not necessarily mean the better theoretical treatment of physical systems we achieve. We need to have some distinction between good, epistemically constructive approximations and bad, epistemically destructive ones. Perhaps the correct slogan should be: the more approximations, the better – but only if they help us learn more!

4.3. Setting the Stage: Representational vs. Explorative Approximation

Many philosophers of science now agree that models play a central place in our scientific understanding of the world.⁸ Of course they do not share the same views on every issue. For instance, the exact nature of the relationship between theories and models is still a controversial issue: are all models in physics 'driven' by theory just as the classical simple harmonic oscillator model might be by classical mechanics?⁹

Nevertheless, I will assume in this chapter the following theses: (T1) a typical way of getting theoretical understanding of a given system in physics consists of building a model for the system, working with the model in order to obtain certain theoretical results out of it, and interpreting the theoretical results with respect to the experimental results of the system; and (T2) one of the important roles of models is to allow us to represent certain parts of the world. I think that (T1) is at least descriptively correct for many

⁸) Cf. Gähde 1995, Giere 1988, Morgan and Morrison 1999

⁹) Cf. Cartwright et al. 1995

theoretical activities in physics. But still, as it seems, (T1) claims a too substantial point to be assumed without any argument. Perhaps it would be better to think of the following argument for the interactive picture as indirect support for the plausibility of (T1), if not a conclusive argument for it. For I shall make it clear that approximations play distinctive roles at each stage of our theoretical understanding set by (T1). I think it is fair to call (T2) uncontroversial.

Now I introduce two distinctive types of approximations in condensed matter physics: *representational approximation* and *exploratory approximation*. My claim is then that these two types are genuinely different kinds of approximations despite the fact that both are equally called ‘approximations’ by condensed matter physicists. A representational approximation is a *constitutive assumption* of a model in the sense that it represents a certain aspect of the real system in the model system and by doing so, becomes a building block of the model. On the other hand, an exploratory approximation is a *theoretical device* for probing a model in the sense that it helps us to extract the physical content from the model system. Notice that different types of approximation take a different place in the process of our theoretical understanding of condensed matter systems.

Representational approximations are employed when we *model* real condensed matter systems, while exploratory approximations are employed when we theoretically *probe* the model we have constructed of condensed matter systems.

Let me illustrate these definitions using a well-known model from the theory of metals: the Drude model.¹⁰ Three years after Thomson’s discovery of the electron in 1897, Drude constructed a model of electrical and thermal conduction in metals by applying the then highly successful kinetic theory of gases to metals. The Drude model was very successful in explaining a number of metallic phenomena and has served since as a basic model for further developments in the theory of metals.¹¹ Basically Drude modelled a metal as a gas of electrons in a sea of immobile positively charged particles. When Drude formulated

¹⁰) See Ashcroft and Mermin 1976, chapter 1.

¹¹) Despite the fact that the Drude model is a purely classical model, the phenomena that can be explained by the model are impressive: Ohm’s law, the Hall effect (partially), the AC electrical conductivity of metals, etc.

his model, the identity of these immobile positively charged particles was not known. But I will consider here the modern formulation of the Drude model where a metal is taken to be a gas of conduction electrons (originating from valence electrons of atoms) in a background of fixed, positively charged ions. Electrical conduction is modelled as the transport of the charges of conduction electrons, and heat conduction, as the transport of their energy.

Apart from a somewhat intuitive picture of moving electrons in a sea of fixed ions, the Drude model is defined by a set of three postulates about the ‘electrons’ and ‘ions’ of the model: the free-electron approximation, the independent-electron approximation, and the relaxation-time approximation. I use scare quotes for the two kinds of theoretical objects in the Drude model not because they are unobservable but because I want to emphasize the postulated nature of them. In other words, ‘electrons’ and ‘ions’ in the Drude model are little more than abstract objects which satisfy the given three postulates. This point is very important for my later argument because the ‘electrons’ and ‘ions’ of the Drude model are far from being ‘approximately’ similar to real electrons and ions in a metal. Since this feature is true of all objects in the models I will discuss later, I will drop the scare quotes from now on for simplicity.

What is the free-electron approximation? In the Drude model, there is only one kind of interaction: ‘collisions’ between ions and electrons. Except for collisions, an electron is assumed not to interact with ions. So the electrons in the Drude model are ‘free’ in that most of the time¹², they are not attracted by positive ions. It is therefore called the ‘free-electron’ approximation. Electrons in the Drude model are also assumed not to interact with each other at all. This implies that the Drude model depicts each electron as being causally separated from the other electrons despite the fact that there are presumably complicated interactions between enormous number of electrons in a metal sample. Physicists call this kind of model a ‘one-electron model’, and the approximation which

¹²) Strictly speaking, electrons of the Drude model are free for all time because the relaxation-time approximation requires the ‘collisions’ to be instantaneous.

ignores the electron-electron interactions, the ‘independent-electron’ approximation. Consequently, the Drude model is a one-electron model.

With these two approximations, the only causal mechanism left in the Drude model is collisions between electrons and ions. The collision in the Drude model is not so much a real time process of an electron’s bumping along from ion to ion. Rather it is a name for a particular *randomisation mechanism* in the following sense¹³: during ‘collisions’, an electron suffers an instantaneous, abrupt change in its velocity with a probability per unit time $1/\tau$ (τ , relaxation time), and immediately after each ‘collision’ an electron is to emerge with a velocity that is not related to its velocity just before the collision, but randomly directed and with a speed appropriate to the local temperature. The first feature is needed so that we can assign to an electron the definite probability for undergoing a collision in any infinitesimal time interval dt , dt/τ . The second feature is necessary for the instantaneous achievement of thermal equilibrium in a particularly simple way. Now it is clear why I call the ‘collisions’ in the Drude model a particular randomisation mechanism: they are *constructed* in such a way as to achieve certain desirable features in the model: a simple and definite collision probability and the effective establishment of local equilibrium. These specific assumptions for the collisions in the Drude model are together called the ‘relaxation-time’ approximation.

The above three approximations are constitutive assumptions of the Drude model. In other words, the Drude model *is* the set of these three approximations with certain interpretative background¹⁴. They are bricks and beams of the Drude model, so to speak. Moreover these three approximations contribute to the building of the Drude model by

¹³) Actually for some time after the model was proposed, physicists were so taken by this intuitive picture of ‘collisions’ that they discussed problems such as the proper aiming of an electron at an ion in each collision at which they were not very successful. Notice that the Drude model does not force us to believe this ‘billiard-ball’ picture of collisions. The model needs a certain mechanism ‘defined’ by the relaxation-time approximation, but the mechanism is abstract enough to be compatible with many different pictures. The ‘billiard-ball’ picture was brought into the model as an analogy from the kinetic theory of gases, but not all features of the kinetic theory of gases need be fruitful in the Drude model of metals. See Ashcroft and Mermin 1976, p. 6

¹⁴) The interpretative background of the Drude model includes some definitional characteristics (electrons and ions are classical particles, but only electrons are mobile, etc.) and implicit postulates (despite collisions electrons are confined to the interior of metals, etc.).

representing *real* electrons and ions in metals in the model in a specific way. So as I define the term, they are *representational approximations*.

On the other hand *within* the Drude model, we may employ various approximation methods in order to calculate certain theoretical quantities. For instance, when analysing the current induced by a time-dependent electromagnetic field in the Drude model, we approximate the force contribution of the electromagnetic field towards the equation of motion for an electron by ignoring the force from the magnetic field and taking only the force due to the electric field.¹⁵ The result is the following:

$$d\mathbf{p}/dt = -\mathbf{p}/\tau - e\mathbf{E} \quad (4.1),$$

where \mathbf{p} is the (average) momentum per electron. The justification for this approximation is the familiar order-of-magnitude consideration. The contribution of the magnetic field to the equation of motion is smaller than that of the electric field by a factor v/c (v , the mean speed of an electron and c , the speed of light). This factor is typically 10^{-10} for metals¹⁶, and therefore the magnetic field contribution can be safely ignored. Similar considerations lead us to ignore the second-order time infinitesimals in the derivation of the equation of motion for an electron itself.¹⁷

The obvious reason for introducing this order-of-magnitude approximation is to make the investigation of the behaviour of electrons under the time-dependent electromagnetic field in the Drude model simpler and easier. In other words, the approximation is intended as a method which facilitates our theoretical study of what the Drude model implies about the behaviour of its electrons under that specific situation, which is in fact the famous phenomenon of *plasma oscillation*. So following my definition, the order-of-

¹⁵) Under the passive picture of approximations this approximation is very peculiar because it cannot be understood as something 'very similar but not exactly' to the electromagnetic field: the electromagnetic field cannot *exist* without a magnetic field however small its magnetic contribution to certain phenomena might be.

¹⁶) Ashcroft and Mermin 1976, p. 16

¹⁷) *Ibid.*, p. 11

magnitude approximation in the Drude model is an exploratory approximation; it is a theoretical device that helps us to extract the physical content of the Drude model.

I claim that representational and exploratory approximations are genuinely different from each other. They differ in their procedural path: one takes us from certain features of real condensed matter systems to our representation of them in a model, and the other, from the model to its theoretical implications. They also differ in our motivation for using them. We employ representational approximations in our model-building because we want our models to be as simple and clear in their causal mechanisms as possible so long as the models help us understand the condensed matter phenomena of interest. On the other hand we employ exploratory approximations in our investigation of a given model itself because without approximations it is often impossible to know what the physical content of the model is. To put it differently, the two types of approximations contribute at different stages of our theoretical activities described in (T1).

All of the three representational approximations in the Drude model are grossly false, let alone approximately true, to the real interactions among electrons and ions in metal.¹⁸ Nevertheless as far as the metallic phenomena that the Drude model successfully treats are concerned, the simple and rather artificially constructed ‘interactions’ prescribed by the three representational approximations are a legitimate part of our knowledge in modelling metals. For we learn from the success of the Drude model lots of valuable information about how to model metallic phenomena. For instance, we learn from it that electron-electron interactions need not be modelled in the treatment of metals in order to explain a large number of metallic phenomena.¹⁹ Also we learn that the specific details of

¹⁸) One might say that the *net effect* of the three representational approximations in the Drude model is approximately true to certain metallic phenomena. I fully agree on that point. In fact that is why the Drude model is so successful. But the point is irrelevant to the question of whether *the behaviour* of the electrons and ions in the Drude model is approximately true to that of the real electrons and ions in a metal.

¹⁹) This does not mean of course that it need not be modelled in order to explain *all* metallic phenomena. That is a different, empirical question which should be answered through the extensive modelling of the other metallic phenomena not dealt with by the Drude model. In fact we now know that electron-electron interactions should be taken into account in order to model superconductivity. We also know however that the independent-electron approximation is still valid in order to understand a lot more metallic phenomena which cannot be properly understood by the Drude model.

electron-ions interactions need not be modelled in order to construct a reasonably good model of metals. This kind of knowledge then provides useful heuristics when we build another model of metals, or of other condensed matter systems. Here we see one epistemic role of representational approximations.

Exploratory approximations allow us to have a different kind of knowledge: knowledge about what a given model can and cannot accommodate. Recall my example of an exploratory approximation in the Drude model: ignoring the magnetic field contribution in the equation of motion for an electron. One of the features of the Drude model resulting from this approximation is that the AC (i.e., time-dependent) induced current becomes linearly dependent on the external electric field just as the DC (time-independent) induced current is.²⁰ This linearity is a nice feature of the Drude model since it allows us to take many general results we know about the DC phenomena into the AC regime and to get a unified treatment of the electromagnetic phenomena in the model. The linearity is introduced by an exploratory approximation and the nature of the exploratory approximation (order-of-magnitude considerations) makes it sure that it is (approximately) a genuine feature of the Drude model. So the exploratory approximation tells us an important fact about the Drude model, namely that it can accommodate DC and AC electromagnetic phenomena in a unified way. Here we see one epistemic role of exploratory approximations. So representational and exploratory approximations are again different in their epistemic roles.

4.4. Is the Distinction Significant Enough to Matter?

Someone who might be sceptical about the significance of the distinction between these two kinds of approximations in condensed matter physics could ask the following question: why shouldn't we say that we *represent* in the Drude model the time-dependent electric field *without* its corresponding magnetic field? Let me first clarify what exactly this question is supposed to challenge. I do not think that the challenge is the rather obvious point that we could build a *new* model (different from the original Drude model)

which has an additional (apart from the original three) representational approximation of representing a time-dependent electric field without an accompanying magnetic field. Surely an electric field cannot be time-dependent without the accompanying magnetic field under Maxwell's law. Nevertheless you *could* think of whatever model you want. The important question is whether the new model is interesting enough to investigate, and it does not look very promising in this respect.²¹

Recall that I define my typology of approximations *functionally*, one as constitutive assumptions of a model and the other as a theoretical device for the process of model-probing. Consequently my typology does not deny the possibility that the addition of an appropriate exploratory approximation produces the same 'effect' in a given model to that of another representational approximation. For instance, we could either represent 'collisions' in a Drude model to be instantaneous, or build the model without this representational approximation and then introduce the exploratory approximation of ignoring the time needed for the 'collisions' afterwards. The important question we should ask is whether this possibility is a rule or an exception.

Therefore I think the real challenge raised by the question is this: Can we obtain the *same* (Drude) model by replacing the order-of-magnitude exploratory approximation with its corresponding representational approximation? If the answer is yes, then the exploratory approximations are after all not substantially different from representational approximations, and consequently my typology of approximations in condensed matter physics is not very significant.²² My general reply to this challenge will be given at the end of section 3 after I discuss the nature and roles of approximations in condensed matter physics. Here I will give my specific answer only to the case of the order-of-magnitude approximation in the Drude model.

²⁰) The current density \mathbf{j} in the AC case is still linear to the electric field \mathbf{E} just as it is in the DC case. This linearity is theoretically very useful. But the linearity would be unavailable if we included the magnetic field contribution in the Drude model.

²¹) This new model cannot describe any time dependent electric phenomena at all unless the inventor of the model could also invent a new electromagnetic theory suitable for the model.

²²) Michael Redhead in his discussions about models and approximations claims that seeking approximate solutions of the exact equations/models is equivalent to seeking exact solutions to approximate equations/models. So he would answer yes to the question. See Redhead 1980.

My specific answer is simply to point out that we *cannot* obtain the same Drude model by replacing the order-of-magnitude exploratory approximation of ignoring the magnetic contribution to the equation of motion of an electron, by a corresponding representational approximation representing no magnetic field in the Drude model. Recall that the magnetic field *itself* has the same magnitude as the electric field. The justification for ignoring the magnetic field contribution in formula (1) (i.e., the equation of motion of an electron) was a *comparative*, not an absolute, one: there are two components from the total force for an electron, one from the electric field and the other from the magnetic field, and the latter is negligible *with respect to* the latter.

The crucial point is that we do not ignore the magnetic contribution simply because the proportion factor of it, v/c , is very very small. If that very very small effect were the only contribution, we should not ignore it. This is in fact what happens when we consider the AC dielectric constant in the Drude model. Consider the following formulae in the Drude model:

$$\nabla \times (\nabla \times \mathbf{E}) = -\nabla^2 \mathbf{E} \quad (4.2)$$

$$\nabla \times (\nabla \times \mathbf{E}) = (i\omega/c) \nabla \times \mathbf{H} \quad (4.3)$$

$$(i\omega/c) \nabla \times \mathbf{H} = \omega^2 / c^2 (1 + 4\pi i \sigma / \omega) \mathbf{E} \quad (4.4)$$

$$-\nabla^2 \mathbf{E} = \omega^2 / c^2 (1 + 4\pi i \sigma / \omega) \mathbf{E} = \omega^2 / c^2 \epsilon(\omega) \mathbf{E} \quad (4.5)$$

where \mathbf{E} is an electric field, \mathbf{H} , a magnetic field, and $\epsilon(\omega)$, the complex AC dielectric constant. Formula (4.2) is a simple result of vector calculus, and Formula (4.3) is obtained by taking a curl of one of Maxwell's equation.²³ Formula (4.4) is also from one of Maxwell's equations using the linear relation between the (induced) current density, \mathbf{j} and the electric field, \mathbf{E} . Finally formula (4.5) follows from formulae (4.2), (4.3), and (4.4) altogether.

²³) In CGS units

Recall that we ‘ignore’ the magnetic contribution by the order-of-magnitude exploratory approximation and get the linear relation between \mathbf{j} and \mathbf{E} . Recall also however that the order-of-magnitude approximation is based on a comparative reasoning. As we can see in formula (4.3), we do not ignore the magnetic field, \mathbf{H} , *despite* its $1/c$ factor, when we calculate the AC dielectric constant. This is because the contribution of \mathbf{H} to the AC dielectric constant is the *only* contribution that reflects the time-dependent effects from \mathbf{E} . If we ignored \mathbf{H} in formula (4.3), we would not have the imaginary part of the dielectric constant in formula (4.5). Then the AC dielectric constant would collapse into the DC dielectric constant (which is simply 1 in CGS units) and the Drude model would become incapable of dealing with any time-dependent phenomena such as plasma oscillation. Understandably physicists do not use an exploratory approximation that ignores the magnetic contribution in this case.

The moral is this: representing no magnetic field instead of a particular exploratory approximation during the calculation of the AC conductivity does not give us the same Drude model. It can only offer us a much less attractive impoverished version of the Drude model that cannot accommodate time-dependent phenomena. In this particular case, we can clearly recognize the distinctive roles that two different kinds of approximations play.

4.5. Approximation and Model: the Interactive Picture of Approximation

As we can see from the above discussion, the kind of considerations involved in the choice of representational approximations are different from those for exploratory approximations. For instance, physicists (although they could if they wanted) do not choose to represent a time-dependent electric field without magnetic field in the Drude model. On the other hand, when physicists probe the capability of the Drude model to deal with various metallic phenomena, they are happy to ignore the magnetic contribution in one case (the AC conductivity) and not in the other case (the AC dielectric constant).

I claim that the interactive picture can explain this difference. Physicists would not use a representational approximation that entirely ignores the magnetic field in the Drude model because it would substantially diminish the value of the model as a model of metals. On the other hand they do use the exploratory approximation of ignoring the contribution due to the magnetic field based on order-of-magnitude considerations, because it allows them to write down theoretical results of the model in a more transparent and manageable form. More specifically, it allows them among other things to describe in the Drude model the electromagnetic phenomena in a unified way.

The interactive picture says that representational approximations are constitutive assumptions of a model. Given the nature of representational approximations, physicists wouldn't adopt one which builds a shortcoming into the model. The interactive picture says that exploratory approximations are theoretical devices for probing a model. Again, given the nature of exploratory approximations, physicists would adopt it in a model whenever they think the use of it can facilitate the probing of the physical content of the model.

Now I will consider two more approximations in condensed matter physics to argue for the interactive picture. Recall that the interactive picture, unlike the passive picture, argues that approximations play various different epistemic roles. Moreover it claims that approximations fulfil these roles by cooperating with general theories. The following two examples will fit nicely into this interactive picture.

Let me consider first the long-wavelength approximation in the Drude model. Recall formula (4.1). The formula is in fact a special case of the following more general equation:

$$d\mathbf{p}/dt = -\mathbf{p}/\tau + \mathbf{f}(t) \quad (4.6)$$

where \mathbf{p} is again the *average* momentum per electron and $\mathbf{f}(t)$ is the external force applied to the electrons of the Drude model. The first term on the right hand side of (4.6) (the

collision term) represents the net-effects of the electron's collisions with ions on an electron under the relaxation-time approximation.

Notice that (4.6) is dependent only on time with no spatial dependence. That is because \mathbf{p} is a variable averaged over space, and the relaxation-time approximation assigns the same collision probability ($1/\tau$) to an electron no matter where it is located, and we consider a spatially uniform force in (4.6). If we consider a special case of (6) where the external force is given by a time-dependent, *but* spatially uniform electric field, then we will get (4.1). Equipped with (4.1), the Drude model can cope with the time dependent, *but* spatially uniform electrical phenomena as well as time independent ones.

The problem arises when we want to consider in the Drude model time dependent *and* spatially varying electrical phenomena such as the propagation of electromagnetic waves. It is no good to put in a form of electric force that is both time and space dependent, $-e\mathbf{E}(\mathbf{x}, t)$, as $\mathbf{f}(t)$ in (4.6) since the other two terms in (4.6) implicitly require the spatial uniformity of the Drude model. Especially the relaxation approximation cannot be valid if the Drude model has any spatial dependence. For if the model had spatial dependence, the collision term would not be independent of the spatial location of an electron. Then, using the space-dependent electric force in (4.6) would make the equation inconsistent, not mathematically but physically with respect to one of constitutive assumptions of the Drude model.

Of course one can try to evade this difficulty by proposing a *new* equation of motion which has both time and space dependence and replacing (4.6) with it. Then the new equation (by construction) may accommodate the time dependent and spatially varying electrical phenomena, not in the Drude model but in a new model where the relaxation-time approximation is no longer valid. But it is not clear why we should make this move since we can come up with a way that allows us to clarify the exact physical implications of each constraint in the problem and to coordinate them in a certain way. And indeed that is what condensed matter physicists do in this case.

The reconciliation starts from looking at the relaxation-time approximation in the Drude model more carefully and thinking about which conditions assure its validity. The core idea is this: we may be able to rewrite (6) with spatial dependence even under the relaxation-time approximation in a way that preserves some desirable features of the Drude model, such as the linear dependence of the current density on an external electric field.

Recall that under the relaxation-time approximation, a series of collisions separated with an average time interval, τ , is the only ‘internal’ causal mechanism for an electron. Thus, between collisions the only force an electron would experience is an external force. To be more specific, let us consider an electric force due to $\mathbf{E}(\mathbf{x}, t)$ as an external force. Then the current density (the local average of electron density, proportional to \mathbf{p}) in the Drude model at a certain point and time will be entirely determined by what the electric field has done to each electron at that point since its last collision. Under the relaxation-time approximation, this last collision takes place with an overwhelming probability no more than a few mean free paths²⁴ away from the point.²⁵ So if the external electric field does not vary appreciably over distances comparable to the electronic mean free path, the field will be locally uniform (although globally not) enough to determine linearly the current density at the point. In other words, both (4.7) and (4.8) hold in the Drude model:

$$\mathbf{j}(\omega) = \sigma(\omega)\mathbf{E}(\omega), \quad (4.7)$$

$$\mathbf{j}(\mathbf{r}, \omega) = \sigma(\omega)\mathbf{E}(\mathbf{r}, \omega). \quad (4.8)$$

This is a nice result because (4.8) allows us to deal with a much larger range of AC phenomena in the Drude model than (4.7) alone does.

In sum, when we consider a spatially-varying electric field in the Drude model, we are under a threat of inconsistency. An electron in the Drude model is assumed to ‘feel’ the same force, independent of its location, and (4.7) results from that assumption. But the

²⁴) The mean-free path of an electron is the average length an electron can cover between one collision and the next one. This is of the order of $1/\tau$.

spatially-varying electric field requires that an electron should feel a different force depending on its location. The solution is this: as long as the wavelength of λ of the field is very large compared to the mean free path, there is no real inconsistency. We can apply the Drude model even to the case of a spatially-varying (but only very smoothly) electric field. This approximation is called the *long-wavelength approximation*.

The long-wavelength approximation is an exploratory approximation: it is used for probing the Drude model's ability to deal with AC phenomena. The diagnosis is that the Drude model is compatible only with those AC phenomena due to a 'comparatively long' wavelength electric field. The precise meaning of 'comparatively long' is given by one of the constitutive assumptions in the Drude model, that is the relaxation-time approximation.

One interesting point about this exploratory approximation is that it is not motivated by any kind of mathematical intractability whatsoever. This means that the role of this particular approximation is not pragmatic as the passive picture might suggest. Rather we try to *coordinate* a certain constitutive feature of the model with another constraint imposed by our cherished theory, in this case classical electromagnetism. By doing so, the long-wavelength approximation clarifies the valid scope of the Drude model: the Drude model is valid only for long-wavelength AC phenomena. This is a clearly epistemic role: the approximation helps us realise a new fact about the Drude model that was not clear from just looking at the constitutive assumptions of the model. As we can see in this example, doing approximations does not necessarily mean losing some valuable information. Sometimes we can gain from it.

Let me consider another example, the Bloch approximation in the Bloch model. The Bloch model is the fundamental model for the band structure of metals. Also it, unlike the (classical) Drude model, is one of the most successful *quantum* models in the theory of metals. The constitutive assumptions of the Bloch model are the following: (B1) the independent-electron approximation, (B2) ions form a perfectly periodic, fixed lattice,

²⁵) Ashcroft and Mermin 1976, p. 25

and (B3) every electron experiences the same effective potential which has the *same* periodicity as the underlying lattice. Let us call (B2) and (B3) together the *Bloch approximation*.

The Bloch approximation is a representational approximation: it is perhaps the most crucial building-block of the Bloch model although it is false in real metals. Almost every important result in the Bloch model follows from (B3)²⁶, while (B2) is necessary to get (B3). The motivation for the Bloch approximation is twofold: first, to reduce a many-electron problem to a one-electron problem with an effective potential, and the second, to investigate the effects of the periodic structure of the lattice on electronic motion. Let me give a rough sketch of how these motivations are embodied in order to get at the Bloch approximation.

Consider the motion of an electron in a model of metals consisting of only electrons and ions. In principle we need to investigate complicated interactions among ions and electrons in order to determine the motion of an electron. But suppose that ions form a perfect crystal, a fixed lattice structure with a certain periodicity, i.e. suppose (B2). Then the motion of an electron is given by its interaction with the other electrons and its interaction with the fixed lattice. But under the independent-electron approximation (B1), the only interaction the electron can have is its interaction with the lattice. Now let us express this interaction by an *effective* one-electron potential. Since electrons are identical to each other, each electron should feel the same effective one-electron potential. The exact form of this potential will depend on the exact lattice structure under study. But if ions are similar enough to each other in the lattice so that we can model them as identical,²⁷ the perfect lattice will produce the periodic potential that has the same periodicity as the lattice itself. Then since the lattice is the only thing that an electron can 'feel', the effective one-electron potential should have the same periodicity as the lattice. Thus (B3).

²⁶) For instance, a lot of properties of electronic levels (band structure) follow from (B3).

Notice that unlike the Drude model the Bloch model does not explicitly mention ‘collisions’ at all. Does this mean that there are no ‘collisions’ between electrons and ions in the Bloch model? Not really. By *construction*, the ‘correct’ one-electron effective potential is supposed to contain all interactions between electrons and the ionic lattice. So this ‘correct’ one-electron effective potential should contain some contributions from collision-like behaviour of electrons with ions. We should rather say that the Bloch model does not separately *recognise* ‘collisions’ as its causal mechanism to be responsible for the metallic phenomena explained by the model. In fact it is a theoretical result of the Bloch model that an electron has a non-vanishing velocity without any retardation, which means there are no ‘collisions’ in a classical sense in the model.

The Bloch approximation suggests to us a good heuristic for modelling metals. Its advice is to separate conceptually the problem of describing metallic phenomena into two parts and attack each part one by one:²⁸ (1) consider the motion of an electron moving through a (fictitious) perfect ionic lattice and experiencing a certain periodic potential due to the lattice, and (2) treat any deviations from the first part (quasi-periodic lattice, impurities, dislocations, lattice vibrations, etc.) by modifying the results from the first part. The first part consists of investigating the effect of periodicity of the ionic lattice on electronic motions. On the other hand, the second part concerns the ‘disturbance’ of this effect when we introduce other causal factors.

As a whole the heuristic aims at analysing the effects of various causal factors on electronic properties in metals separately from each other. The Bloch model with the Bloch approximation picks up a causal mechanism from the ionic lattice to electrons through the effective one-electron potential. The fact that this potential has the same periodicity as the lattice explains a lot of metallic phenomena. On the other hand, under the Bloch approximation the model does not identify ‘collisions’ as a separate causal mechanism. In this way the Bloch approximation allows physicists to study exclusively

²⁷) Although this assumption is essential for getting at (B3), it is rarely mentioned in the textbook. Cf., Ashcroft and Mermin 1976, pp. 132-3

²⁸) Daniela Bailer-Jones takes this as a general prescription for modelling. See Bailer-Jones 1999, 2000.

the effects of the periodicity of typical lattice structures of metals in a highly idealised Bloch model.

After finding out what kinds of phenomena the periodicity of the lattice can give rise to in the Bloch model, physicists then can consider a more complicated model by introducing more causal mechanisms one by one. What they usually do is to apply the same heuristic as the Bloch approximation. For instance if physicists want to introduce a new causal mechanism by allowing ions to ‘vibrate’ with respect to their equilibrium positions, they (1) consider a new model with an electron moving through a (fictitious) perfect ionic lattice *with* lattice vibrations, and (2) try to figure out the composite effects of both causal factors (the periodicity of the lattice and lattice vibration) in this model. In the first part of this theoretical investigation, physicists need another representational approximation in order to study *exclusively* the effects of lattice vibrations on metallic phenomena without paying too much attention to the motion of electrons. So they adopt the adiabatic approximation where electrons are assumed to adjust themselves very quickly to the (relatively slow) vibration of ions so that they will always be in their ground states.²⁹ In the second part, they try to understand how a new causal mechanism (lattice vibrations) modifies the effective one-electron potential given by the Bloch approximation.

Both in the Bloch model and in a new model with lattice vibrations, what we see is the powerful method of theoretical physics: the *analytical method*. And the two representational approximations (the Bloch approximation and the adiabatic approximation) are the core of the application of this method to our theoretical investigation of metals. They help us pick up various causal factors (the periodicity of the lattice, lattice vibrations, etc.) when we construct models of metals. Then they also help us study one particular channel of causal mechanism exclusively (the effective one-electron potential, normal modes, etc.) Here we can see how representational approximations can add something to our understanding of condensed matter systems. They teach us how to model condensed matter systems.

²⁹) Ashcroft and Mermin 1976, pp. 425-6

I have argued that the interactive picture accounts for many facts about the nature and roles of approximations in condensed matter physics far better than the passive picture. Now I can more fully answer the challenge about the significance of my typology of approximations. As I said before, although I do not deny the possibility that one representational approximation and another exploratory approximation have the same physical content, I think that such cases are exceptions, not the rule.

The reason comes from my discussion of this section. Representational and exploratory approximations serve different purpose: the former is to build models or to isolate certain causal mechanisms, and the latter is to extract the correct physical content of a model or to clarify the valid scope of a model. It is not unthinkable that some approximations are good at all these epistemic roles. But as we saw in the case of the magnetic field contribution in the Drude model, the considerations for each type of role are usually different. Ignoring the magnetic field entirely in the model does not give us a good representational approximation while it does give us a good exploratory approximation depending on the specific problems we address in the model. Similarly, representing a model to be suitable only to the long-wavelength electric phenomena without any independent reason is unlikely to be a good representational approximation, since the model, after the exploring process, might turn out to be applicable to all wavelengths.

4.6. Conclusion

Approximations actively play various epistemic roles in theoretical physics. They can add something to our knowledge of how to model physical systems in the world. The proper way of looking at the relationship between general theories and approximations is through coordination rather than passive appendix. This is the account of the interactive picture of approximations, and I have argued for it discussing it in condensed matter physics.

We can distinguish two types of approximations in condensed matter physics based on which roles they play in the process of our theoretical understanding assumed by (T1).

Representational approximations serve as constitutive assumptions of a model, and exploratory approximations, as model-probing devices. Condensed matter physicists use them for other purposes as well, such as isolating causal mechanisms, coordinating the model with general theories, and clarifying the valid scope of a model. These two types of approximations convincingly illustrate how inadequate the passive picture is. Both representational and exploratory approximations in condensed matter physics are best understood in terms of the interactive picture of approximations.

One of the motivations for me to look at approximations in this way comes from an observation that physicists show different attitudes towards different kinds of approximations. Sometimes they are fastidiously meticulous about the ‘correctness’ of certain approximations: whether a given series expansion does converge to the true value, and if it does, whether it converges rapidly enough so that it is safe to take just a few leading terms as an approximate. And sometimes they very much care whether a given idealised model can be improved by replacing an unrealistic approximation by more realistic one. These are the cases that the passive picture can deal with. But there are other cases where physicists are quite happy with the ‘unrealistic’ nature of certain approximations. Galilean idealisation is one of them.³⁰ Moreover this different attitude is closely related to the different points at which physicists employ approximations in their building and exploring of models. The different types of approximations are use at the different stage of their model-based research. Consequently, physicists’ attitudes towards these two types of approximations will depend on what considerations they think most important in each stage of their research. As the kind of considerations taken to be crucial at the model-building stage are generally different from that at the model-exploring stage. Physicists tend to have different attitudes towards different types of approximation.

³⁰) McMullin 1985

CHAPTER 5

THE RENORMALIZATION GROUP METHOD IN CONDENSED MATTER PHYSICS

5.1 Introduction

In this chapter I discuss the *renormalization group* (RG) method in condensed matter physics (CMP): how its main ideas are motivated by other theoretical developments in CMP, how it works and what its philosophical implications are. The RG method is one of the most successful theoretical frameworks employed by condensed matter physicists for modelling and understanding macroscopic phenomena. For instance it provides, among other things, the most impressive explanation of critical phenomena by present physics. By looking at how this method has been developed and how it works, we can learn a lot about the gist of successful modelling in CMP.

Except for recent notable exceptions,¹ the RG method has been discussed by philosophers usually in the context of quantum field theory (QFT).² In contrast, the rich and independent theoretical tradition of the RG method in CMP has been largely neglected in the literature. I will talk a little about the relationship between these two methods, more on their substantial differences than on their formal similarities. In fact I will claim that there are a number of theoretical, conceptual and motivational differences between these two mathematically closely related methods.

More specifically I will argue that the theoretical need for the RG method in CMP has been motivated by different considerations from the notorious problem of infinities in QFT, and that the RG in CMP has many theoretical implications not shared by the RG of QFT and vice versa. It is important to realise the significance of these issues. Some authors (*Cf.* Huggett and Weingard 1995) seem to assume that the RG of CMP is just another application of the RG ideas of QFT, and that any philosophical conclusions possibly correct for the RG of QFT

¹) See Batterman 1998, Batterman 2000 and Hughes 1999 for the insightful discussion on several philosophical implications of the RG in CMP. Also there are a few historical and conceptual works by physicists on the topic. Domb 1996 is an excellent history of the theory of critical phenomena, offering a balanced view on the role of the RG method in this history. Fisher 1999 is insightful on the differences of the RG in CMP and the RG in QFT.

²) For the historical and philosophical discussions of the RG method in QFT, see Brown 1993, Cao 1999, Cao and Schweber 1993, Huggett and Weingard 1995 and 1996, Schweber 1994 and Teller 1994, chapter 7.

are automatically valid about the RG of CMP. Understandably Huggett and Weingard (1995, 1996) illustrate their claims about the principles of the RG in QFT, often by examples of CMP. In section 5.4 I will challenge this view and investigate a number of other related issues.

I use the word ‘method’ in the expression ‘RG method’ on purpose, and thus avoid a probably more conventional name, the RG *theory*. Both, the ‘RG method’ and the ‘RG theory’, are used by physicists to characterise the RG ideas and applications in CMP. It is fair to say that physicists tend to use the ‘RG theory’ when describing various abstract formalisations of the RG ideas such as a general theory of ϵ -expansion in momentum space. On the other hand, they opt for the ‘RG method’ when discussing concrete implementations of the RG ideas such as a dynamic RG treatment of non-equilibrium phenomena.³

Still there is a tendency among some physicists (especially field theorists working on CMP problems) to talk freely about the RG theory *as if* RG were a universal algorithm, to which we may put any physical questions expecting beautiful answers in a row. I think that this tendency is at best misleading and at worst simply misrepresents how the RG method works in CMP. First of all even when it is appropriate to talk about a RG ‘theory’, the definite article in ‘*the* RG theory’ is unjustifiable for, if my claims in section 5.4 are correct, there are significant variations in the nature of RG from QFT to CMP.

More importantly the RG in CMP cannot be identified by a fixed formal theory one can uniquely designate. Each of the RG applications in CMP is an *implementation* of a set of loosely connected ideas, carefully *designed* for modelling condensed matter phenomena. In order to achieve a successful RG modelling of a given phenomenon, physicists need a lot of context-sensitive creativity and not just the formal theory of RG. Consequently a large degree of variations in formalism and in structural assumptions exist among the different RG implementations.

Consider the idea of *scaling* for instance. It is one of the most important ideas of the RG method in CMP. But there are many different ways of embodying this idea depending on the specifics of the problem one might want to study, as we can see in section 5.2. Still we can

³) Cf. Goldenfeld 1992, chapters 9-12

understand all of them under the banner of the RG method. To put it differently, there is no unique set of (mathematical or implicit) features shared by *all* RG implementations in CMP. One of the consequences of this feature of the RG in CMP is that you cannot write down an unambiguous procedure for applying the RG for a given physical problem. In this sense the RG is not a theory in contrast to the theory of graph expansion in QFT for instance.

Each implementation of RG picks up only some of the RG ideas and *adapts* them to the specific nature of a given problem. The unifying power of the RG method partly lies in its *plasticity*. All these considerations are reflected in my second claim that the RG in CMP is not a uniquely identifiable mathematical theory, but rather a collection of interrelated and mutually inspired ideas and theoretical techniques, that is to say, a *method*. I am going to argue for this in section 5.5.

The RG method in CMP has achieved its most notable success with *critical phenomena*, phenomena involving phase transitions of many-body systems. This does not mean however that before the arrival of the RG, critical phenomena were a sort of mystery for condensed matter physicists. An influential and versatile method called the *mean field* (MF) method was (and still is) used to explain a number of characteristics of phase transitions both qualitatively and quantitatively. The mean field method has been continuously bolstered by a number of sophisticated series expansion techniques in precisely calculating the values of *critical exponents*.

Whether physicists or philosophers, those who are overwhelmingly impressed by the power of the RG method in CMP tend to think of the relationship between the RG method and the MF method as *replacement*. That is, the RG theory (as they would prefer to call it), being more correct and offering us the more comprehensive explanation of critical phenomena, has made the mean field theory *effectively redundant*. I shall argue that talking of replacement in this case is not even factually correct let alone philosophically fruitful. What has happened in the development of the theory of phase transition is best understood as the constructive *complementation* between two methods. The RG method is good at describing the scaling behaviour of a many-body system near its critical points; but it is silent about other important issues such as the nature of phases far from critical region. In contrast the MF method is inherently unreliable very near critical points because its fundamental assumption about the analyticity of thermodynamic functions breaks down there; but it is excellent at answering

exactly those questions with which the RG method cannot help us. In sum we need both the MF method and the RG method in order to get a comprehensive explanation of critical phenomena. This is why complementation, rather than replacement, is a better concept to understand the relationship between these two methods, or so I will argue in section 5.6.

But first of all we will look at the theoretical background with respect to which RG has emerged as a new method in CMP modelling. Then we can discuss what the RG method in CMP actually consists of and how it works in section 5.3.

5.2 Correlation Length, Universality, Scaling and the Mean Field Method

Imagine a many-body system, say a gas in a box. We can measure a number of macroscopic properties of this gas such as its density, compressibility, specific heat, etc, while keeping its external parameters such as pressure and temperature fixed. Let us suppose that the system is in equilibrium, which means that these macroscopic properties are not changing over time. Notice that I am not talking about equilibrium in terms of its more theoretical definition in statistical mechanics, namely being characterised by the microcanonical measure. Rather I want to start my discussion with an empirical fact that the macroscopic properties of *some* physical systems are not changing appreciably over time. I do this because I want to make it clear the link between this empirical observation and some of important theoretical concepts in CMP. In fact you can take a many-body system being unchanging as regards its macroscopic properties as the *de facto* definition of it being a *thermodynamic system*.

Now divide the system into two roughly equal halves. If the system is big enough, some macroscopic properties of each sub-system, the *intensive* properties, will then remain the same as those of the original system.⁴ This fact generally holds true if we repeat this kind of division many times. It is obvious however that this cannot be true after too many repetitions. For instance, there is some point at which the compressibility of a sub-system fluctuates substantially. In fact the stability of (intensive) macroscopic properties is bound to fail at the level of very small sub-systems as the concept of compressibility is not even applicable to atoms.

⁴) The other kind of macroscopic properties, extensive properties are proportional to the volume of the system. So in this case the values of extensive parameters of the halved sub-system will be a half of those of the whole system.

The length scale at which the (intensive) macroscopic properties of a sub-system are appreciably different from those of the whole-system gives a measure of what is called the *correlation length* of the system.⁵ We could say that any two regions of a system would significantly influence each other only if they are apart less than the correlation length of the system. Consequently, the microscopic fluctuations in different regions of the system apart more than its correlation length could be regarded as being effectively independent of each other. In this case, two regions of the system are virtually so ‘disconnected’ from each other that when we separate the two, there is no appreciable change in their macroscopic properties.

This is a nice example of how elegantly a rather simple theoretical concept of CMP can capture the physical content of well-established empirical facts about many-body systems. A many-body system under a given set of external conditions typically exhibits more or less stable macroscopic properties in its entire region; but only up to a certain length scale. We understand this phenomenon employing the concept of correlation length. Two regions of the system, separated further than its correlation length, are not substantially ‘correlated’ with each other. As a result it will make no appreciable difference to their macroscopic properties if the connection is actually severed, for instance if the system is physically divided into two parts. On the other hand if the system is divided so finely that the dimension of its dividends is smaller than its correlation length, the physical properties of the dividends will be different from one to another significantly. This is because the physically significant interactions among the dividends are cut off in this case and therefore their contributions to the physical properties of the dividends are lost.

It is noteworthy that the correlation length of a many-body system generally depends on the external parameters as well as the specifics of the system. The physical implication of this fact can never be clearer than in the phenomenon of phase transition. Consider a ferromagnet that exhibits below its Curie temperature so-called *spontaneous magnetisation*, magnetisation without any applied magnetic field. The correlation length of a ferromagnet varies a lot depending on its temperature and applied magnetic field. Especially, as the system approaches its paramagnet-ferromagnet phase transition point (i.e., its Curie temperature with no applied magnetic field), it increases continuously and finally becomes infinite at the

⁵) A formal definition of correlation length is the following: ξ is the correlation length of a system if the correlation function of the system, $C(\mathbf{r}, \mathbf{r}')$, behaves like $\exp(-|\mathbf{r} - \mathbf{r}'|/\xi)$ in the thermodynamic limit. The

transition point. From our earlier discussion of correlation length, we can immediately conclude that at this particular external condition, every bit of the ferromagnet is strongly correlated with each other. This fact makes this type of phase transition, the so-called *continuous* phase transition, very difficult to study mathematically: many favourite analytic techniques of physicists such as perturbative expansions are not applicable.⁶ As we shall see in section 3.1, the RG method is a nice way of evading this difficulty.

Another empirical fact we know about the continuous phase transition is its *universality*. It turns out many (not all)⁷ properties of a class of systems *near*⁸ their critical points are largely independent of the microscopic details of the interactions between their constituents. For instance it has been recognised since the early twentieth century that the liquid-gas phase transition shares many features with that of the paramagnet-ferromagnet phase transition.⁹ So if we draw the graph of the *reduced* density, $\rho - \rho_C$, for water (where ρ is the density of water and ρ_C , the density at the transition point), it shows the striking resemblance with that of the magnetisation, m , in the paramagnet-ferromagnet transition. Similarly, the isothermal compressibility of water shows the almost same behaviour with the zero field susceptibility of a ferromagnet. In modern terminology, we say that both transitions belong to the same universality class, in this case to the Ising universality class. One of the pivotal achievements of the RG method in theory of phase transition is to explain these phenomena of universality.

The phenomena of universality can be vividly seen through another important empirical fact that has fascinated physicists for a long time in various fields: *scaling*. In its simplest form, scaling just means that two measurable quantities depend on each other in a power-law fashion. A familiar example is Kepler's law, relating the (mean) radius R of a planet's orbit to the period T of the orbit:

$$T \propto R^{3/2} \tag{5.1}$$

correlation function, $C(r, r')$, gives a measure of how much two regions r centred at \mathbf{r} and at \mathbf{r}' are correlated with each other. Cf. Goldenfeld 1992, pp. 95-98

⁶) Many of important thermodynamic functions such as free energy (or their derivatives such as specific heat) is *not analytic* (meaning mathematically singular) on the critical point. It is not due to the divergence of these quantities although they are inevitably divergent when we take the thermodynamic limit of them. Rather it results from the fact that the continuous change of a phase into another phase cannot be represented in a mathematically smooth way. Notice that this is already a quite different reason from the parallel difficulties of perturbative expansions in QFT, that is, the problem of undesirable infinities.

⁷) See section 5.3.3 for the discussion of this point.

⁸) See section 5.3.1 and 5.3.2 for the significance of this 'nearness' of universal critical phenomena.

The scaling exponent of Kepler's law, $3/2$ in (5.1) can be deduced from dimensional analysis once we assume the inverse square law $1/r^2$ for the gravitational force. It is known that the scaling exponents obtainable from simple dimensional analysis will be generally rational numbers.¹⁰

Several important scaling relations like (1) were experimentally found in critical phenomena. In ferromagnets for instance physicists have established the following power-law behaviours of important thermodynamic functions in the vicinity of its phase transition point:

$$C \sim A |t|^{-\alpha} \quad (5.2)$$

$$\lim_{H \rightarrow 0^+} M \propto (-t)^\beta \quad (5.3)$$

$$\chi \equiv (\partial M / \partial H) |_{H=0} \propto |t|^{-\gamma} \quad (5.4)$$

$$M \propto |h|^{1/\delta} \quad (5.5)$$

where C is the specific heat in zero field (apart from terms which are regular in t); t , the *reduced temperature* $t \equiv (T - T_c) / T_c$ with T_c , the critical temperature; M , the spontaneous magnetisation; χ , the zero field susceptibility; h , the reduced external magnetic field, $h \equiv H / k_B T_c$. Notice that the above scaling relations phenomenologically define *critical exponents* such as $\alpha, \beta, \gamma, \delta$.

Interestingly experimentalists found in early 60's that many physical systems with radically different microscopic compositions share the same critical exponents. For instance the same critical exponents of ferromagnets are shared by ordinary liquids such as water as regards their liquid-gas phase transition. Again, properly modelled, many percolation phenomena exhibit these same critical exponents. This is a clear case of universality of critical phenomena.

Experimentalists also found that the critical exponents are in general *irrational* numbers. It implies that simple dimensional analysis is inadequate for understanding critical phenomena. Simple dimensional analysis is known to be adequate if there is only one 'significant' length

⁹) Domb 1996, pp. 9-14

¹⁰) Cardy 1996, pp. 3-4

scale in the physical problem under consideration. For instance, there is a single length scale in Kepler's problem, namely the radius of a planet's orbit, which is responsible for the rational scaling exponent, $3/2$. Although there are many other length scales involved in the motion of the planets around the sun, for instance the radii of the planets, they can be ignored when we study the large-scale features of the planets' motion such as Kepler's law. The irrationality of critical exponents then implies that this convenient feature of many physical problems, 'the separation of significant length scale', is not valid in critical phenomena. In order to study critical phenomena properly, we have to take into account physics of all length scales, from the minute-scale random fluctuations of atoms to the extremely long-range correlations among the parts of a system. This fact gives a serious difficulty for theoretical physicists because many conventional techniques developed for more expedient problems are inapplicable.

The task for condensed matter physicists is then to explain how these scaling behaviours are brought about and hopefully obtain numerically accurate values of critical exponents. By doing so, they could say something about how the tantalising phenomena of universality should be understood. Before and after the arrival of the RG method in CMP, condensed matter physicists used a powerful and versatile theory of phase transitions, the *mean field* (MF) method¹¹, to do this job. The physical idea behind the MF method is rather simple: it amounts to ignoring individual correlations among constituents of a many-body system, and postulating instead that every constituent responds to the same field obtained by averaging the correlations over the entire system. Let me explain this point by a concrete example, the ferromagnetic Ising model.

The fundamental quantity we have to consider for the ferromagnet Ising model is its partition function, and it is given as follows:

$$Z = \text{Tr} \exp[(1/2)\beta \sum J(r, r')s(r)s(r') + \beta H \sum s(r)] \quad (5.6)$$

J is the interaction strength between a spin at r and a spin at r' , $\beta, 1/kT$ where k is Boltzmann constant and H is an external magnetic field; each spin can take either $+1$ or -1 as its value.

¹¹) There are many variations of the MF method in many branches of many-body physics from atomic physics to solid state physics, hence the name the mean field 'method' rather than the mean field 'theory'.

In short this model describes a ferromagnet as a many-body system of ‘spins’ interacting with each other as well as with the external magnetic field.

Now it is very difficult to calculate the partition function even for this simple model of ferromagnets. So one might try to replace (5.6) with a new partition function describing a non-interacting system under a uniform field. This must be done in such a way as to encapsulate the essential physics of the problem, which is a non-trivial task. In the present case we know that the applied field H will give rise to some non-zero magnetisation $M = \langle s \rangle$ where $\langle s \rangle$ is the thermodynamic average of all spins. So we start from the following identity,

$$s(r)s(r') = (M + (s(r) - M))(M + (s(r') - M)) \quad (5.7)$$

and then to expand this to the first order in the fluctuation, $\delta s(r) \equiv s(r) - M$:

$$s(r)s(r') = M^2 + M\delta s(r) + M\delta s(r') + O((\delta s(r))^2) \quad (5.8)$$

$$\therefore s(r)s(r') \approx -M^2 + M(s(r) + s(r')). \quad (5.9)$$

As we can see from (5.9), the MF treatment of the model given by (5.6) replaces spin-spin interactions by spin-magnetisation interactions plus a constant. To put it in a different way, the MF method deprives spins of any correlations with each other. This is rather drastic. From the earlier discussion, we know that correlations within a correlation length do matter. For instance two spins neighbouring each other are strongly correlated. On this physical ground we can expect that the MF method will give us quantitatively correct results only when the correlation length of a given system is very small.

In fact we can utilise this connection between correlation length and the MF method in order to formulate a (rough) criterion for the valid scope of the MF method, namely the Ginzburg criterion¹² given as follows:

$$(\text{Ginzburg criterion}): \text{the MF method is valid when } \xi^{4-d} \ll R^4 \quad (5.10),$$

¹²) Cardy 1996, pp. 25-6

where ξ is a correlation length of a given system, d , the spatial dimension of the system and R , the range of interaction. Roughly speaking then, the MF method is valid when a measure of the range of interactions among constituents (given by R) is much bigger than the correlation length. Or to use our division-analogy before, the MF method is valid when we can get the stability of the macroscopic properties even to very minute scales.

This criterion is very useful for understanding why the MF method is very successful in explaining some phenomena such as (Type-1) superconductivity by the BCS theory, but even qualitatively unsatisfactory for some other phenomena such as the near-critical behaviour of continuous phase transitions. The explanation goes like this. For most systems R is of the same order as a , the microscopic length scale (for instance lattice spacing). We therefore expect that the MF method becomes invalid by the time the correlation length becomes a few times a as the system approaches its critical point. This is why the MF method is quite good at explaining far-from-critical point properties of the system while unsatisfactory as regards near-critical behaviours.¹³ On the other hand, for some systems such as type-I superconductors, R is large, of the order of the size of a Cooper pair. For such a system, it is very difficult to see the deviation from the MF predictions, even when experimenters probe the system very close to its critical point.

The MF method offers its explanation of the phenomena of universal critical behaviour. In fact the MF method gives us a remarkable degree of universality, in many cases too remarkable to be consistent with experimental data. The critical behaviour, according to the predictions of the MF method, is independent of almost every detail of the system under consideration. Actually, there is only one single set of critical exponent values according to the MF method, often referred to as the *Landau values*.

In sum we see how correlation length, scaling, universality and the MF method are interrelated in CMP research. For instance the scaling relations in critical phenomena provide us with a remarkable example of universality, and again the correlation length shows a scaling. It is also important to notice that each concept (or method) constrains the others in such a way to delimit the usefulness or validity of them. For instance correlation length delimits the validity of the scaling of macroscopic properties. Universality can be understood

¹³) This point is crucial for my argument in section 5.3.4.

by the MF method, but then the validity of the MF method relies on the comparison of the range of interaction with respect to the size of correlation length.

5.3 The RG Method in CMP

5.3.1 Motivation

I shall first try to convey the main ideas of the RG method in CMP before going into the details of its mathematical formalism. Suppose we want to investigate some specific problem about a condensed matter system. Instead of solving the problem in its original form, we may first try to re-express the relevant parameters of the problem in terms of some others, hoping that this re-expression renders the problem easier to tackle. For instance in the problem of how long-range order is established, we may decide to ‘integrate out’ the short-distance degrees of freedom of a given model and represent the system with a new model. Or, we may modify the effects of large-scale disturbances in a model of fluid turbulence and arrive at another model. The point is that we start with a model of certain phenomena, and transform it into a *new* one, by which we wish to understand the phenomena better.

Another important point is that we never even try to preserve everything by this change of models. After all, by the change we move from one model to another as our intended representation of the phenomena. Unless the change is trivial, these two models do not share all of their features; loss is inevitable. Naturally, we want to keep the crucial features of the phenomena we are interested in as intact as possible while allowing the change of some other features. For instance in the earlier example of the establishment of long-range order, we may allow somewhat radical modifications in the short-distance features of an original model as long as it preserves the long-distance characteristics of the model.

From our discussion in chapter 4 of the function of representational approximations in the model-building stage, we know that this distinction of important (thus preserved) and unimportant (thus modified) features is not absolute, but clearly depends on the nature of the problems we are studying. Indeed the RG applications in CMP provide us with plenty of examples illustrating this point. For instance, when we are more interested in how exactly the short-range fluctuations are organised down to a very small scale in the aforementioned fluid

model, we may be willing to bear with modifications of large-scale disturbances, but require the preservation of the short-range structure.¹⁴

Sometimes this transformation of one model into another is recursive in the sense that we can apply the same transformation again to the modified model, and get a further modified model. Furthermore this recursive transformation may have a group structure (actually semi-group structure). Particularly, the ‘nature’ of the transformation could have something to do with a renormalization, that is the rescaling of, say, the spatial dimension) of a model accompanied by appropriate changes in its interaction parameters. Combining these two features, we get the *renormalization group* of a series of models. The RG method in CMP is to use the RG transformation of a given CMP model in order to study some highlighted features of the model. Notice that even in this rough description of it there is no mention of the problem of infinity, one of the focal points in philosophical discussion of the RG in QFT.¹⁵

5.3.2 The RG Method and the Universality of Critical Phenomena

Now I will explain how the RG method works in a concrete case taking the two-dimensional Ising model (Ising-2D) as an illustration. The Ising-2D model is a free model, a model not fixed to any particular interpretation.¹⁶ It describes an abstract two-dimensional (infinite)¹⁷ lattice system of which primitive entities are ‘spins’ located on each lattice point. (Figure 5.1) Here ‘spin’ is just a name for the only dynamic variable of this system, not intrinsically related to quantum mechanical spins. Actually ‘spins’ in CMP models can be either classical or quantum, depending on the nature of the physical system the model is intended to describe. ‘Spins’ in the Ising-2D model are interacting with each other as well as with an external field (usually taken to be a magnetic field). The Ising-2D model describes quite diverse phenomena including the order-disorder transition of a ferromagnet and the liquid-gas transition of a simple fluid.

Recall our earlier discussion of correlation length. Figuratively speaking, the system ‘looks’ the same whether we look at the whole system or certain part of it provided the correlation length is large enough. To put it differently, a system with a large correlation length looks the same whether we see it closely or a bit from a distance. As the Ising-2D system approaches

¹⁴) Cardy 1996, p. 29

¹⁵) Cf. Cao and Schweber 1993 and Teller 1994

¹⁶) For the discussion about various model-types of CMP, see chapter 3.

its phase transition point, its correlation length gets larger and larger. So we can expect that in its critical region the Ising-2D system exhibits so-called *scale invariance*: the invariance of the critical properties of the system under the (length) scale change. We can formalise this point using Kadanoff's *block-spin* idea.

Kadanoff proposed that when the Ising-2D system is near its critical point, if we consider a new system consisting of *block-spins* obtained from a local average of several neighbouring spins of the original system, the new system should show the same critical behaviour as the original system.¹⁸ From this postulate and assumptions about the scaling of two important thermodynamic variables of the system, reduced temperature, t , and reduced field, h , he could explain several empirically known scaling relations of the system. There is more than one way of defining block-spins and I choose one of the simplest, shown in Figure 5.2. As we can see, each block spin comes from 9 original spins and the value of the block spin is determined by a simple majority rule, that is, the arithmetic mean of all the spins in a given block spin. As original spins have a value of either $\frac{1}{2}$ or $-\frac{1}{2}$, the block spins will get either $\frac{1}{2}$ or $-\frac{1}{2}$ as well.

Now we should decide how these block-spins interact with each other. A bold conjecture of Kadanoff was to require that the new system should be governed by the same form of *Hamiltonian* as the original system (Invariance of Hamiltonian), which means that the block-spins also should interact only with their nearest neighbour block-spins and with the external magnetic field. This proposal, although innovative and seemingly reasonable, was actually an unfortunate one since the new system is in general governed by a different Hamiltonian. We can easily see this point from Figure 5.3. Even when the original Hamiltonian has only nearest-neighbour interactions, the block-spins usually come to have next nearest-neighbour interactions.

This feature turns out to be not special to the 2D-Ising model, but a general characteristic of any block-spin transformation for models of critical phenomena. To illustrate this point, let us consider a rather unphysical model of Ising-spins where spins interact only with next nearest-neighbour, not the nearest ones. Still, after a block-spin transformation, the nearest-neighbour block-spins have to interact each other! In this case, the new system clearly does not share the

¹⁷) That is, after taking the thermodynamic limit.

Hamiltonian with the original system. The point we can learn from this consideration is this: the block-spin transformation and its generalised version, the RG transformation, will generally generate a *different* system from its starting system governed by a *different* Hamiltonian. The two model-systems share some properties just as two different physical systems could share certain characteristics. Still we need to keep in mind that the original Hamiltonian and its block-transformed Hamiltonian represent two different physical systems in order not to draw wrong conclusions out of the RG method in CMP.

Now then how can we determine the new (renormalized) Hamiltonian of block-spins in the 2D-Ising model? Recall that in statistical mechanics the partition function is the fundamental quantity in describing a physical system. Many of the important quantities of the statistical mechanics description of a system come from the derivatives of a partition function. Consequently as long as two systems with different Hamiltonians share the same partition function, the SM description of them will be the same in many respects. So it is reasonable to require that the renormalized system with block-spins should have the same partition function as the original system; that is,

$$Z = \text{Tr}_{s'} e^{-\mathcal{H}'(s')} = \text{Tr}_s e^{-\mathcal{H}(s)} = Z' \quad (5.11),$$

where \mathcal{H} is the (reduced) Hamiltonian of the original system and \mathcal{H}' , that of the renormalized system.¹⁹ We can implement our particular majority rule of the block-spin transformation by introducing the following projection operator,

$$\begin{aligned} T(s'; s_1, s_2, \dots, s_9) &= 1 \quad \text{if } s' \sum_i s_i > 0 \\ &= 0 \quad \text{otherwise.} \end{aligned} \quad (5.12).$$

Using the constraint (5.11) and the projection operator defined by (5.12), we can now formally define the renormalized Hamiltonina as follows:

$$e^{-\mathcal{H}'(s')} \equiv \text{Tr}_s \prod_{\text{blocks}} T(s'; s_1, s_2, \dots, s_9) e^{-\mathcal{H}(s)} \quad (5.13).$$

¹⁸) Kadanoff 1966

¹⁹) Following a commonly adopted notational convention among condensed matter physicists, I let $\beta(=kT)$ be absorbed into the redefined Hamiltonian. This newly defined Hamiltonina is called a *reduced* Hamiltonian.

The meaning of (5.13) is this: the new Hamiltonian is determined by ‘tracing-out’ all the contribution of the interactions within each block to the original partition function. The resultant, partially traced-out, partition function implicitly defines the renormalized Hamiltonian of block-spins.

The idea behind this ‘tracing-out’ is that we construct a new model and preserve in it all the ‘large distance’ physical features of the 2D-Ising model. The trick is that we ‘trace-out’ all the ‘short-distance’ physics of the 2D-Ising model and obtains its ‘renormalized’ model by requiring the partition function of the two models to be the same. This procedure can be applied to the new model again, and results in another new model and so on ... The hope is that by studying this chain of models and their RG transformation we might be able to understand better the critical behaviour of the 2D-Ising model.

Unfortunately, the partial-trace in (5.13) is hopeless to work out even for very simple projection operator like (5.12). Moreover, standard approximation methods are hard to apply, as there are often no small parameters to work with if we consider the RG transformation in real space (*real space RG*). This is why many condensed matter physicists find another kind of RG, the *momentum space RG*, attractive; there are often small parameters such as the famous ϵ in Wilson’s ϵ -expansion in the momentum space RG. I’ll talk more about this issue in section 5.3.3.

Now let us come back to the real space RG. Without doing the relevant calculation of the RG transformation explicitly, we can still learn a lot about the critical behaviour of the 2D-Ising model by studying the general feature of its so-called *RG-flow*. In order to discuss the RG-flow, we need to think of all the possible couplings of a given Hamiltonian as forming a vector $\{\mathbf{K}\} \equiv \{K_1, K_2, \dots\}$. In the original 2D-Ising model, there is only the nearest neighbour coupling, say K_1 , with all the other $K_i = 0$. But, as discussed before, the renormalization transformation will eventually generate all the other couplings. We may therefore picture the RG transformation as acting on the parameter space of all possible couplings $\{\mathbf{K}\}$:

$$\begin{aligned} \{\mathbf{K}'\} &= \mathcal{R}(\{\mathbf{K}\}) \\ \{\mathbf{K}''\} &= \mathcal{R}(\{\mathbf{K}'\}) = \mathcal{R}^2(\{\mathbf{K}\}) \\ &\vdots \end{aligned}$$

$$\{\mathbf{K}^n\} = \mathcal{R}^n(\{\mathbf{K}\}) \quad (5.14)$$

We may think of each $\{\mathbf{K}^i\}$ as representing a model (or its Hamiltonian) which results from i -times application of the RG transformation, \mathcal{R} , to the 2D-Ising model. If we follow these \mathbf{K} 's, we get the *RG-flow* of the 2D-Ising model. (See Figure 5.4)

The basic strategy of the RG method in the theory of critical phenomena is to *assume* the existence of a fixed point in the RG transformation under consideration (especially when we know that there is a phase transition in its corresponding system) and to study the consequences of the assumption. That is, we assume there exists in parameter space a fixed point, \mathbf{K}^* which satisfies $\mathcal{R}(\{\mathbf{K}^*\}) = \{\mathbf{K}^*\}$. Roughly speaking, as a RG transformation increases (in absolute terms) the canonical length scale of the models along a RG flow, the model picked up by \mathbf{K}^* represents the long-range behaviour (including its critical behaviour) that all the models in the RG flow share.

In order to illustrate this point, let us consider the correlation length ξ of the 2D-Ising model. From our discussion in section 5.2, we know that at the critical point, ξ should be infinite. Now as the block-spin transformation we have considered for the 2D-Ising model reduces the size of the system by 3 (scale factor), the following holds for the correlation length: $\xi[\mathbf{K}]/3 = \xi[\mathcal{R}(\{\mathbf{K}\})]$. But then for \mathbf{K}^* , $\xi[\mathbf{K}^*]/3 = \xi[\mathcal{R}(\{\mathbf{K}^*\})] = \xi[\mathbf{K}^*]$ must hold. This means that $\xi[\mathbf{K}^*]$ can only be zero or infinity. The (critical) fixed point corresponding to an infinite correlation length represents the singular critical phenomena, while the (trivial) fixed point corresponding to zero correlation length represents the bulk phases of the system. Here we can see how the concept of fixed points of the RG transformation captures one important feature of phase transitions, an infinite correlation length.

In fact many important critical phenomena including the universality of phase transitions are explained by looking at how the RG flow behaves locally near the critical fixed point, rather than at the critical fixed point itself. So we also assume that the RG transformation is differentiable at its fixed point so that we can linearise \mathcal{R} about \mathbf{K}^* as follows:

$$K_a' - K_a^* \sim \sum_b T_{ab}(K_b - K_b^*) \quad (5.15)$$

where $T_{ab} = (\partial K_a' / \partial K_b) |_{K=K^*}$. Call the (left) eigenvalues of the matrix \mathbf{T} , λ^i , and its eigenvectors $\{u_i\}$, so that

$$\sum_a u_a^i T_{ab} = \lambda^i u_b^i. \quad (5.16)$$

Basically we can extract from λ^i most of the important characteristics of the critical behaviour of the 2D-Ising model including its critical exponents. In order to do that, we define *scaling variables*, $v_i \equiv \sum_a u_a^i (K_a - K_a^*)$, which are linear combinations of the deviations $K_a - K_a^*$ from the fixed point. Scaling variables transform multiplicatively near the fixed point:

$$\begin{aligned} v_i' &= \sum_a u_a^i (K_a' - K_a^*) = \sum_{a,b} u_a^i T_{ab} (K_b - K_b^*) \\ &= \sum_b \lambda^i u_b^i (K_b - K_b^*) = \lambda^i v_i. \end{aligned} \quad (5.17)$$

It is convenient to define the quantities y_i , *renormalization group eigenvalues*, by $\lambda^i = b^{y_i}$: where b is a scale factor (3 for the particular block transformation we chose before). There are three cases to distinguish:

- $y_i > 0$: repeated renormalization group iterations drive v_i always from its fixed point value. In this case, v_i is said to be *relevant*.
- $y_i < 0$: repeated renormalization group iteration drive v_i away from its fixed point value. If we start sufficiently close to a fixed point, v_i will iterate towards zero. In this case, v_i is said to be *irrelevant*.
- If $y_i = 0$: v_i is said to be *marginal*. In this case, we cannot tell from the linearised equations whether v_i will move away from the fixed point or towards it. v_i is related with logarithmic corrections to scaling.

In order to explain the physical meaning of this classification let us consider a fixed point about which \mathcal{R} has n relevant eigenvalues and n' irrelevant eigenvalues²⁰, which makes the dimension of the parameter space, $n + n'$. Near the fixed point, consider a linear space spanned by the irrelevant eigenvectors, an n' -dimensional hypersurface. From the definition of irrelevant eigenvectors, we can see that successive renormalization will lead any point on this surface to the fixed point. The hypersurface is called the *critical surface*, as the critical phenomena of the model on this surface will be given by the fixed point (and the near-fixed

²⁰) Although strictly speaking the number of irrelevant eigenvalues are infinite.

point behaviour of the RG-flow). The irrelevancy of irrelevant eigenvectors comes from this insensitivity: couplings related to irrelevant eigenvectors are irrelevant as far as the description of critical phenomena is concerned.

On the other hand, the models not on the critical surface will not exhibit the same critical behaviour as the models on the critical surface. Successive RG transformations will lead them away from the critical surface. In order to 'bring' these models on to the critical surface, we need to control the couplings related to the relevant eigenvectors. From the viewpoint of experimentalists, this controlling amounts to controlling certain physical parameters of the experiment such as temperature, pressure or magnetic field. We may refer to these as 'knobs' in order to highlight its operational implications for an experimentalist.²¹ In order to end up on the n' -dimensional critical surface, and then to the fixed point, she must therefore adjust exactly n knobs.

In a sense, the distinction between the relevant/irrelevant 'knobs' explains the universality of critical phenomena. First of all, the existence of irrelevant 'knobs' explains why many systems, quite different in various respects, exhibit the same critical behaviour and thus belong to the same universality class. It is because the RG transformation is unresponsive to any change of irrelevant 'knobs' in the sense that the differences become insignificant through the successive application of the RG transformation. Although different systems are indeed different in their various characteristics, and their differences do matter in some other contexts,²² the RG transformation is designed in such a way as to disregard any difference that is not responsible for most of the critical phenomena.

On the other hand, the existence of relevant 'knobs' explains why there is more than one universality class as regards critical phenomena. After all, if all physical features of a system could be regarded as irrelevant, there should have been one universality class described by its unique fixed point. The successive application of the RG transformation along a relevant eigenvector leads a starting model away from a given fixed point and ultimately leads it to another fixed point. Consequently, we may say that the relevant 'knobs' for models will classify them into a number of different universality classes, each of which is described by different fixed points.

²¹) Cf. Cardy 1996, pp.41-43

So here is a picture of how these things work. (Figure 5.4) There can be a number of fixed points in the parameter space; each of them describes a different universality class. Any model represented in the parameter space would, following the RG-flow in the direction of relevant eigenvectors, ‘attract’ onto one of the critical surfaces. On a critical surface it would again ‘attract’ to its fixed point, and its critical behaviour will be described by the fixed point. Here we can see that the role of relevant eigenvectors is to ‘separate’ different regions in the parameter space for its corresponding universality class. On the other hand, the role of irrelevant eigenvectors is to ‘guide’ each model on a critical surface to its fixed point.

As an illustration consider the order-disorder phase transition of a ferromagnet. Here two ‘knobs’ must be adjusted to make a sample of a ferromagnet exhibit a phase transition: the magnetic field should be slowly lowered to zero while temperature is set to a specific value T_C (the critical temperature of the ferromagnet). Interestingly, the same is true of the liquid-gas transition of a simple fluid; only in this case the ‘knobs’ are temperature and pressure. The RG explanation of these facts roughly goes like the following: (1) As far as their critical phenomena are concerned, we may model both ferromagnet and a simple fluid by the 2D-Ising model.²³ (2) There are two relevant eigenvectors as regards the critical fixed point of the 2D-Ising model, and all the other eigenvectors are irrelevant. (3) Therefore, despite their differences in many other physical features, both a ferromagnet and a simple fluid belong to the same 2D-Ising universality class.

In fact, the 2D-Ising universality class includes a wider range of models than those models that can be modelled by the 2D-Ising model with nearest-neighbour interactions. Consider Figure 5.4 again. Here we see a two-dimensional hypersurface (spanned by two reduced coupling constants, K_1 and K_2) in the parameter space we considered before. K_1 equals J_1/kT , where J_1 is the strength of the nearest-neighbour interactions, and K_2 , J_2/kT where J_2 is the strength of the next nearest-neighbour interactions. As we can see from the RG flow, both the 2D-Ising model with the nearest neighbour interactions and a new model (2D-Ising* model) with only the next nearest-neighbour interactions are governed by the same critical fixed point. So their critical behaviour will be the same.

²²) Especially, for the so-called non-universal critical phenomena.

But notice that the 2D-Ising model and the 2D-Ising* model do *not* share every features in critical phenomena, although they belong to the same universality class. For instance, the critical temperature for each model is generally different as we can see from Figure 5.4. Critical temperature is an example of *non-universal* critical behaviour. The existence of the non-universal critical behaviour highlights the fact that the RG transformation will pick up only some of the critical behaviours which are universal such as critical exponents, the number of relevant 'knobs'.

Now consider a more general model that has both the nearest-neighbour interactions and the next nearest-neighbour interactions. As we change the temperature 'knob', the model follows the dotted line in Figure 5.4. Here we can see the 'relevance' of the temperature 'knob. When the temperature is above T_C , the successive application of the RG transformation will lead the model to the high temperature fixed point, and the universal critical behaviour of the model will be the same as a model with no interaction. On the other hand, when the temperature is below T_C , the successive application of the RG transformation will lead the model to the zero temperature fixed point, and the universal critical behaviour of the model will be the same as a model with infinite K_1 and K_2 interactions.

Our next question is how we can get the empirically known critical scaling relations of, say, a ferromagnet out of the RG transformation. As defined by (5.2) – (5.5), the scaling relations for the ferromagnet are related to the power-law behaviour of thermodynamic functions such as spontaneous magnetization M , or of response functions such as specific heat C .

Thermodynamic functions and response functions are first or second derivatives of the free energy of the 2D-Ising model respectively. So what we need to know is how the free energy of the 2D-Ising model is changed by the RG transformation. It turns out that we can write the singular part of the free energy in terms of the reduced temperature and the reduced magnetic field (two 'knobs') and their corresponding RG eigenvalues:

$$f_s(t, h) = |t/t_0|^{d/y_t} \Phi [(h/h_0)/(|t/t_0|^{y_h/y_t})] \quad (5.17)$$

²³) In fact the relation between a simple fluid and the 2D-Ising model is indirect: first, a simple fluid is modelled by the lattice-gas model, and then the lattice-gas model turns out to be mathematically equivalent to the 2D-Ising model.

where f_s is the part of free energy density that is not analytic at the critical point, Φ , a scaling function, y_t , the RG eigenvalue for temperature ‘knob’, and y_h , the RG eigenvalue for external magnetic field ‘knob’.

Now it is rather straightforward to get the critical exponents from (5.17). For instance, specific heat is $\partial^2 f / \partial t^2 |_{h=0} \propto |t|^{d/y_t - 2}$, so that $\alpha = 2 - d/y_t$. Likewise, spontaneous magnetization is $\partial f / \partial h |_{h=0} \propto (-t)^{(d - y_h)/y_t}$, so that $\beta = (d - y_h)/y_t$. Similarly we get $\gamma = (2y_h - d)/y_t$ and $\delta = y_h / (d - y_h)$. We see that the four principal critical exponents are given in terms of the two RG eigenvalues. This implies that there must be two so-called *hyper-scaling* relations between α , β , γ and δ such that $\alpha + 2\beta + \gamma = 2$ and $\alpha + \beta(1 + \delta) = 2$. These relations were postulated from experimental results, before the introduction of the RG method into the theory of phase transition. The RG explanation of the hyperscaling relation illustrates how simple and clear the RG explanation of critical phenomena can be. There are only two ‘relevant’ RG eigenvalues for the 2D-Ising model, and all four critical exponents are expressible in terms of these eigenvalues, so two relation should hold between four critical exponents.

Notice that no irrelevant scaling variable nor eigenvalues appear in (5.17). In general irrelevant scaling variables do affect the scaling function, Φ , in (5.17), and their contribution is manifested as a correction to the scaling function. The correction is usually small, but not always. When it is large, we need to probe the system very near to its critical point in order to see the scaling relations predicted by (5.17).²⁴ This fact shows the ‘approximate’ nature of the RG transformation in the context of CMP. The ‘equivalent’ models in a same universality class are equivalent only with respect to some critical behaviours. Moreover, their equivalence will emerge when we probe them very close to the critical fixed point. Sufficiently far from the fixed point where the linear approximation of \mathcal{R} is not valid, they are different even in these critical behaviours.

5.4 A Comparison: the RG in CMP and the RG in QFT

Several issues about renormalization in QFT have been discussed by philosophers as well as historians.²⁵ Two issues seem most prominent and actively debated. The first issue concerns

²⁴) Cardy 1996, pp. 48-9

²⁵) Cf. Brown 1993; Cao 1998, 1999; Cao and Schweber 1993; Huggett and Weingard 1995, 1996; Schweber 1994; Teller 1994, 1998

the nature of the ‘infinities’ that are frequently popping up in QFT calculations. Do they somehow correspond to genuine infinities in nature? Or are they just artefacts of our clumsy current theories? Fortunately physicists had managed to develop various procedures to evade this problem and extract meaningful answers from the relevant QFT calculations. The procedures range from a rather simplistic cutoff renormalization to a more sophisticated regularization program, and more recently to the RG method. The second issue is whether these procedures are justifiable. At a first glance, the introduction of the procedures could look *ad hoc* since the development of the procedures was mainly motivated by physicists’ urge to eliminate infinities and get the final results to be finite.

Paul Teller suggests three ways of looking at these issues.²⁶ First, the *cutoffs* approach. Here the order of doing two mathematical procedures is crucial. When we calculate in QFT a potentially diverging quantity such as mass, we make it finite by introducing a cutoff. Then we absorb a finite correction to bare mass into our definition of ‘observed’ mass. The rationale here is that we can observe only ‘dressed’ mass, not bare mass. The next step is to take the limit of the cutoff and make our definition of observed mass cutoff-independent. This approach emphasizes that we only deal with finite quantities at each step of calculation, so the problem of infinity is illusory. The justification of renormalization in this approach is given by the fact that each step is mathematically consistent and respectable.

But some theorists feel uncomfortable about this approach. Their complaints are directed to the conceptual problems of the cutoff-dependent intermediate quantities. These quantities do not satisfy various theoretical constraints physicists value highly, such as gauge invariance, unitarity, etc., and are therefore not satisfactory. So they argue that we have to conceptualize QFT calculations entirely in terms of cutoff-independent quantities. Teller calls this way of thinking about the renormalization in QFT the *real-infinities* approach. According to this approach, what we do during renormalization is literally to throw away genuine infinities. It is clear that the advocates of this approach take very seriously the infinities in QFT calculations. Still it is hard to see, according to Teller, how this discarding of genuine infinities can be justified.

²⁶) Teller 1994, pp.159-69

The third approach is named by Teller the *mask-of-ignorance* approach. This approach points out that in order to take seriously the infinite quantities in QFT calculations we have to be sure of the correctness of our current theories at all energies. But clearly we know that our current theories for instance QED break down at high enough energy. So we should think (according to this approach) that our current, infinity-plagued theory must be, at best, an approximation to some divergence-free ‘correct theory’. Notice that this approach implicitly assumes that there is no genuine infinity in nature so that the ‘correct theory’ should be divergence-free. As we do not have this ‘correct theory’ yet, we do not know from which energy level our approximate theory becomes reliable. In the same vein we do not know which regularization scheme we should choose in order to arrive from our approximate theory to the ‘correct theory’. Here comes the renormalization procedure as a rescue: it circumvents the problem and extracts cutoff independent quantities that are independent of the details of the regularization scheme used during calculations. In short, according to the mask-of-ignorance approach, the use of renormalization is justified by its instrumental utility in allowing us to get finite reliable quantities in QFT calculations despite the approximate nature of our current theory.

But then Teller picks up another interesting question: can we justify *renormalizability* as a constraint on theory construction? He offers one way of answering the question. It starts with regarding the current approximate theory as ‘incomplete’ in a sense that it should be supplemented by ‘measured’ parameters just as Hooke’s law is supplemented by spring constants. In theory construction, we also assign ‘measured’ values to the parameters of a theory, hoping that someday we might be able to calculate them from the ‘correct theory’. But then if a theory requires infinitely many parameters to be assigned in this way, the theory is no use. That’s why we require our approximate theory to be renormalizable.

Teller then quickly connects the issue of renormalizability with that of whether a given QFT model is insensitive to the details of what happens at very high energies. His explanation is rather intuitive.²⁷ Conceptually it is one thing to insist that a model have finite experimentally determined parameters (which seems reasonable if the model is to be ‘workable’), and it is another thing to insist that a model should be insensitive to the details of very high-energy physics. If the ‘correct theory’ turns out to be completely describable by a finite number of

²⁷) Teller 1994, p. 168

parameters, our approximate model may have to be sensitive to the details of the ‘correct theory’, but still have finite experimentally determined parameters.

A more satisfactory explanation of this connection could be sought from noticing a *fact of the matter* that some of the successful (though approximate) models in QFT are insensitive to the details of the regularization scheme used for their renormalization. Then we realize that we can characterise this fact of the matter by a theoretical condition, renormalizability. In fact this is the route condensed matter physicists follow in their use of the RG method. They start from an empirical fact of universality in critical phenomena: various different physical systems (or their models) share most of their critical behaviours. Then they notice that these models have a finite number of relevant ‘knobs’ under a well-chosen RG transformation \mathcal{R} . Here the connection is physically much deeper: the finiteness of relevant ‘knobs’ *explains* why models in a universality class described by the same fixed point for \mathcal{R} could be so different in their details, but still share many critical behaviours.

Huggett and Weingard criticise Teller for focusing too much on perturbative renormalization.²⁸ They highlight that the renormalization *group* is more appropriate than perturbative expansions for the discussion of renormalization in QFT. As the RG method is exact and non-perturbative, one can study the validity of perturbative results by checking whether a given theory is exactly renormalizable or not. For instance the exact RG analysis can distinguish physical QFTs in the sense that they can be regarded as just approximations to true physics in small lattice from unphysical ones such as quantum electrodynamics (QED).²⁹ Basically they want to separate the problem of ‘infinities’ in QFT, to which Teller pays much attention, from the issue of whether a given theory is exactly renormalizable by the RG method. Still they seem to endorse a view similar to Teller’s mask-of-ignorance approach. Approving Georgi’s ‘mild’ version of the effective field approach, they say that very successful theories such as quantum electrodynamics (QED) are irreducibly phenomenological. They are not correct at very high energies, and should be regarded as ‘effective’ field theories in a sense that they are valid only at low enough energies.

²⁸) Huggett and Weingard 1996, p. 312, my italics

²⁹) Huggett and Weingard 1995, pp. 184-5

The problematic part of Huggett and Weingard's view on the RG method appears when they say, "two theories with different cutoffs can describe *the same* physics."³⁰ It is not at all clear what they mean by 'the same physics'. But clearly they mean by it a strong enough thesis to imply "since the RG is defined to relate physically equivalent theories, every theory, with varying cut-off, on our unstable manifold is equivalent to the continuum physics of the limit."³¹ Interestingly they seem to think that this analysis is true of all the RG applications both in QFT and in CMP; they make their points using examples from QFT as well as those from CMP such as Kadanoff's blockspin transformation.

It is easy to see why this cannot be generally correct. Let me get their point straight first. The RG procedure in QFT starts with a phenomenological theory that is valid only at sufficiently low energies, and takes it closer and closer to the unknown 'correct theory' that is valid at any arbitrarily high energies. Therefore the relation between two could be thought of as that of approximation: the 'correct theory' is more fundamental and the phenomenological theory is only an effective theory. On the other hand, two different phenomenological theories can be taken to the same fixed point by the RG transformation. In this case, providing the RG transformation is mathematically *exact*, we can say that two phenomenological theories are physically equivalent with each other. When the fixed point describes a continuum physics, we could indeed say that they are both equivalent to a continuum physics.

In CMP, the RG procedure starts with a model and transforms it into another model. Here there is little sense to saying that the latter is more fundamental than the former. It is just a different model with a different Hamiltonian. Even the model described by a fixed point is not specially privileged: it is just a hypothetical model for which a one more time RG transformation changes nothing. As we have seen before, the really important information about the critical phenomena comes from the *near* fixed-point behaviour of the RG transformation, not from this hypothetical model.

In fact this model does not have much physical content in itself. Strictly speaking we need to apply a suitably defined RG transformation *infinitely* many times in order to arrive at the model, averaging out *all* scales of interactions so to speak. But clearly that doesn't mean that all the models described by fixed points are models with no interactions, and therefore

³⁰) Huggett and Weingard 1995, p. 179

physically indistinguishable from each other. (That would mean there is only one universality class!) Remember that by taking the blockspin transformation in Figure 5.2, we reduce the linear dimension of the system to one third of its original size. Repeating the blockspin transformation many times, we would sooner or later reduce the system to nothing, or wouldn't we? The trick is that we implicitly take the thermodynamics limit of each blockspin system, and thereby effectively deal with infinite systems. In short the fixed point doesn't have any *representational* value: we do not know how to write down the Hamiltonian for it, and we surely do not think that the hypothetical model describes the *essence* of critical phenomena. What is important is the scaling relation of free energy, and this can be obtained from the near fixed-point behaviour of the RG transformation.

Contrary to Huggett and Weingard, the RG in CMP does not relate physically equivalent theories. The models belonging to the same universality class do not share all the physically important features. They share physical features *relevant* to their critical phenomena. In particular they usually do not share *bulk* properties of a stable phase such as density of liquid phase or elasticity of solid phase. Even about critical phenomena, they do not share *all* relevant physical features. There are non-universal critical phenomena as well: most notably the critical temperature is different from model to model within the same universality class. The RG in CMP picks up the universality class of critical phenomena among models, not an equivalent class of all the physical features among models.

Why is there a difference? The basic reason is that the RG in QFT is supposed to study how the same system can be theorised at different energies. Then you might be able to prove that under a certain RG transformation, two different theories are physically equivalent in the sense that they can be understood as two different representations of the same system probed at different energies. In CMP, the RG is supposed to study why many different systems share some remarkable characteristics in critical phenomena, that is to provide an explanation of universality and scaling relations. Then you might be able to explain why among potentially infinitely many physical features of a given model, only a few of them are relevant to the major characterisation of the critical phenomena and therefore define the universality class. The RG methods are employed in different ways in the two fields in order to achieve

³¹) Huggett and Weingard 1995, p. 182

different goals. Despite its similarity in mathematical formalism, the RG of CMP has substantial theoretical differences from the RG of QFT.

In fact the RG of CMP is motivated by different considerations from that of QFT. First of all, the problem of dealing with undesirable infinities in calculations does not even arise in CMP. One generic way of getting infinities in QFT is through some integration up to an infinite energy/momentum scale (*ultra-violet divergence*). This corresponds to, in real space, the integration up to an arbitrary small length scale. In generic critical phenomena models (or many other applications of the RG method in CMP), we start from a certain lattice structure with a finite lattice spacing like the one in Figure 5.1. Here when we do the RG transformation, we take a chunk of spins and transform it into a single blockspin. As I pointed out before, the system becomes smaller before taking the thermodynamic limit. Figuratively speaking, we partially integrate out towards bigger and bigger linear dimension, which is the opposite direction from that of the RG in QFT. Due to this feature of the RG in CMP, the infinities arising in CMP are not directly related with the RG, but with the existence of phase transitions and ultimately with taking the thermodynamic limit. That is why the infinities and divergences in CMP, such as infinite correlation lengths or the divergence of specific heat, are called *infrared divergence*.

In QFT, it is hard to justify a particular choice of cutoff over another, so we want our theory to be cutoff independent in the end of calculations. On the other hand, there is a *natural* cutoff in the RG of CMP. By construction, there is nothing between lattice points; therefore it is meaningless to talk about interactions shorter than the lattice spacing. Consequently, the lattice spacing of a given model gives a natural choice of cutoff when we apply the RG method to the model.

Without undesirable infinities to eliminate, what kind of considerations could motivate the development of the RG method in CMP? The answer is easy to guess from our discussion in section 2. There are many empirical facts about the properties of many-body physical systems in the world: the stability of bulk properties, the universal critical behaviour many radically different systems share when they change their bulk phases, etc. Condensed matter physicists have developed many useful concepts in order to understand these empirical facts: correlation length, long range order, scaling relations, etc. The RG method in CMP is another development in this physicists' effort to understand many-body phenomena.

More specifically the RG method allows us to see why only a few features in a given model matter as regards its critical behaviour, while many other features not. The RG method also allows us to calculate in principle many quantities such as critical exponents, which are otherwise only empirically obtainable. These are great achievements of the RG method in CMP. But still the method is motivated by a longstanding tradition of many-body modelling, not by the problem of infinities. This explains why the real space RG is, despite its disadvantage in calculation, more prominent in CMP, rather than the momentum space RG that is almost universally adopted in QFT. What I have said here is not only historically correct about the introduction of the RG method into CMP, but also true to condensed matter physicists' practice these days.³²

Finally, let us consider how appropriate the mask-of-ignorance view is for interpreting the RG method in CMP. There is certainly an element of ignorance in the RG method in CMP. For instance, except for a few simple cases, we do not know how to calculate renormalized Hamiltonians on a given RG flow from (5.11). But this doesn't mean that we do not have definite models on the RG flow of CMP in the same sense that we do not know the 'correct theory' at high enough energies in QFT. We do have a definite analytical expression given by (5.11) for each renormalized Hamiltonian. Every model on the RG flow is fully specified, even though we cannot write it down explicitly. Here the ignorance involved is more or less that of classical statistical mechanics. But there is one aspect that the RG method in CMP shares with the mask-of-ignorance view of the RG method in QFT: it has methodological nature. Adopting the RG method in QFT is a methodological decision to get reliable, effective theories about mid-range energy phenomena in order to evade our ignorance of the 'correct theory'. Employing the RG method in CMP is a methodological decision to extract the relevant features of critical phenomena taking a 'rough' look at many-body systems. That is to say, adopting the RG method in investigating critical phenomena means pursuing a research strategy of understanding a large part of critical phenomena without even trying to calculate the practically uncalculable partition function of the model.

5.3.3 The RG in CMP as a *method* rather than as a theory

Now briefly I want to argue that the RG in CMP should be understood as method rather than theory. In fact these two ways of referring to the RG can be found even in a same text,³³ so it might be just a matter of terminology. But when I emphasise the RG as method, what I want to say is the following: (1) The RG is not an algorithm. (2) There is no common set of features shared by all RG applications in CMP.

First, about the non-algorithmic character of the RG. As we have seen in section 5.3, there is no prescription for building an appropriate RG transformation given a model. A consequence of this is that when we do not find any fixed point from a RG transformation, we cannot conclude that the model has no fixed point. Perhaps the RG transformation is badly designed, and a better designed RG transformation might give us a fixed point. But recall that the existence of a fixed point is directly related to the existence of a phase transition in the model. Presumably, whether a given model has a phase transition or not has a definite answer. Still, a mere application of RG cannot decide this matter.

This indeterminacy makes it hard to call the RG in CMP a theory. A theory should have unambiguous implications about the targeted phenomena no matter whether we can actually write down those implications or not. The RG method in CMP has no such implications. It cannot tell whether there is a phase transition in a given model. What it can do is to give us a framework with lots of heuristics by which we can investigate the question. The RG method itself does not say much about the critical phenomena. It talks about how to establish a set of effectively equivalent (with respect to relevant features) models given a starting model, and how to get valuable information out of the set. This is what I see as one of the major characteristics of the methods in physics.

A typical scientific theory describes or claims something about the world. In order to do that, it should have a well-defined set of conditions shared by all of its applications. You can apply the classical kinetic theory of gases to nitrogen, oxygen and even to water, so long as each application satisfies the conditions given by the theory, such as very weak interaction among constituents, low density, etc. The RG method in CMP does not have a well-defined set of conditions. In earlier sections I mainly discuss an intuitively clear application of the RG

³²) See Domb 1996 for the canonical evaluation of the significance of the RG method in CMP. Domb's view shared by many other authors, including Cardy 1996, Fisher 1999, Goldenfeld 1992 and Uzmov 1993.

³³) For instance, Goldenfeld 1992.

method in CMP: a blockspin transformation in real space. Here the RG deals with the renormalization in physical space. It connects a starting model with its blockspin models generated by physical space rescaling. But this feature is not necessary, although it is very widely used in CMP. When we study percolation phenomena, such as draining of water into soil, using RG, what we renormalize is not along the physical space, but along time. Given a model, we generate renormalized models that correspond to the later stages of the model. In this case, a fixed point describes a kind of stabilised percolation state. Here only certain conditions of the RG method are kept, while others are either loosened or not met. This kind of RG, so-called *dynamic* RG, is as legitimate an application of the RG ideas as the real-space RG and the momentum-space RG. The success of the dynamic RG in CMP shows that the RG in CMP better be understood as method rather than theory.

5.3.4 The RG method and the MF method in CMP: *Co-operation* rather than *Replacement*

Now I will argue very briefly that the MF method is co-operated with, rather than replaced by, the RG method in CMP. First of all, let me make the following facts clear. It is simply *wrong* to say that the MF method is actually not used anymore in CMP, at least in an advanced research level, and all the theoretical investigation is done by the RG method. The MF method is quite widely used for research purpose. Many important calculations have been done by the MF method. Moreover, new techniques have been developed for the last 20 years in order to meet various difficulties arising from the applications of the method.³⁴ That is to say, the MF method is very much an active research method, widely used and valued. So in this sense, the RG method does not replace the MF method.

But we need to be careful at this point because there is another sense of replacement. We often say that Newtonian mechanics is replaced by the special theory of relativity. But that does not bar us from using Newtonian mechanics in various areas including mechanics itself, nor from developing new techniques in Newtonina mechanics to tackle new problems arising in its theoretical and engineering applications. What we mean by 'replacement' in this context is that we could do all the calculations and explanations Newtonian mechanics do using special theory of relativity. It might be much more cumbersome to do the relevant calculations, and sometimes we would have less intuitive explanations from special theory of

³⁴) Cf. Uzunov 1993

relativity. (After all we do not 'feel' like living in the four-dimensional Minkowski spacetime.) But special theory of relativity makes Newtonian mechanics *redundant* in principle.

Now this is certainly not true of the RG method and the MF method. The RG method does not make the MF method redundant. On the contrary, in order to give us a full explanation of critical phenomena, the RG method should be supplemented by the MF method. If we recall how the RG method works, the reason is clear. The RG method is specially aimed to find out the near fixed-point behaviour of the RG transformation \mathcal{R} . From this, we can obtain the scaling relations of the free energy of a given model and get the explanation of the *phase transition*. But in any of these results, we have no clue about the nature of phases *away* from phase transition. This information should be available if we want to understand the physical implications of the change from one phase to another phase. The MF method can give this information by investigating the model away from the phase transition point. Then we can combine this result with the RG result and get the whole picture of phase transitions. In this way the RG method and the MF method complement each other in our quest for the explanation of critical phenomena.

Figure 5.1: The 2D-Ising model

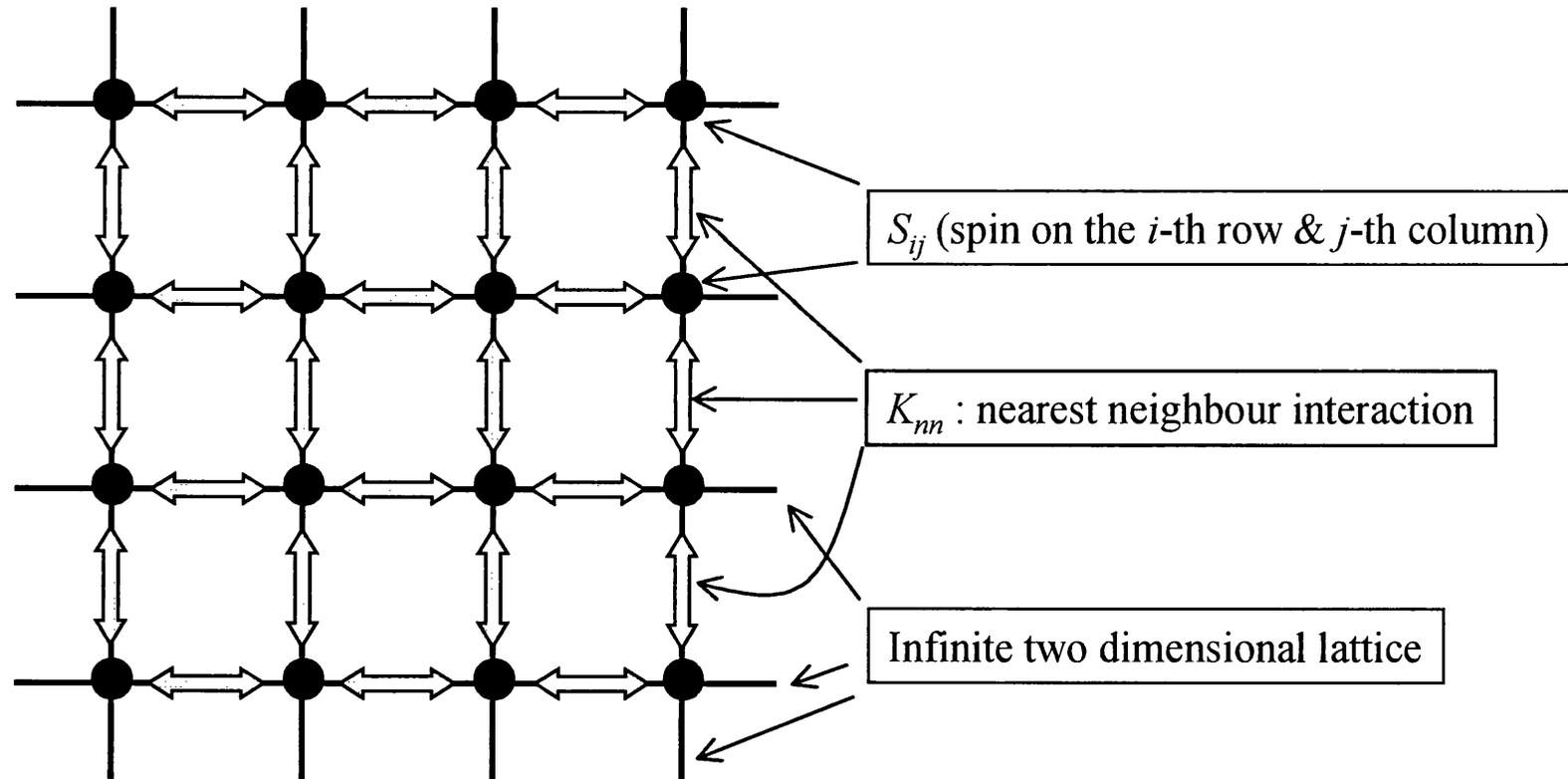


Figure 5.2: Block-spin transformation (2D-Ising model)

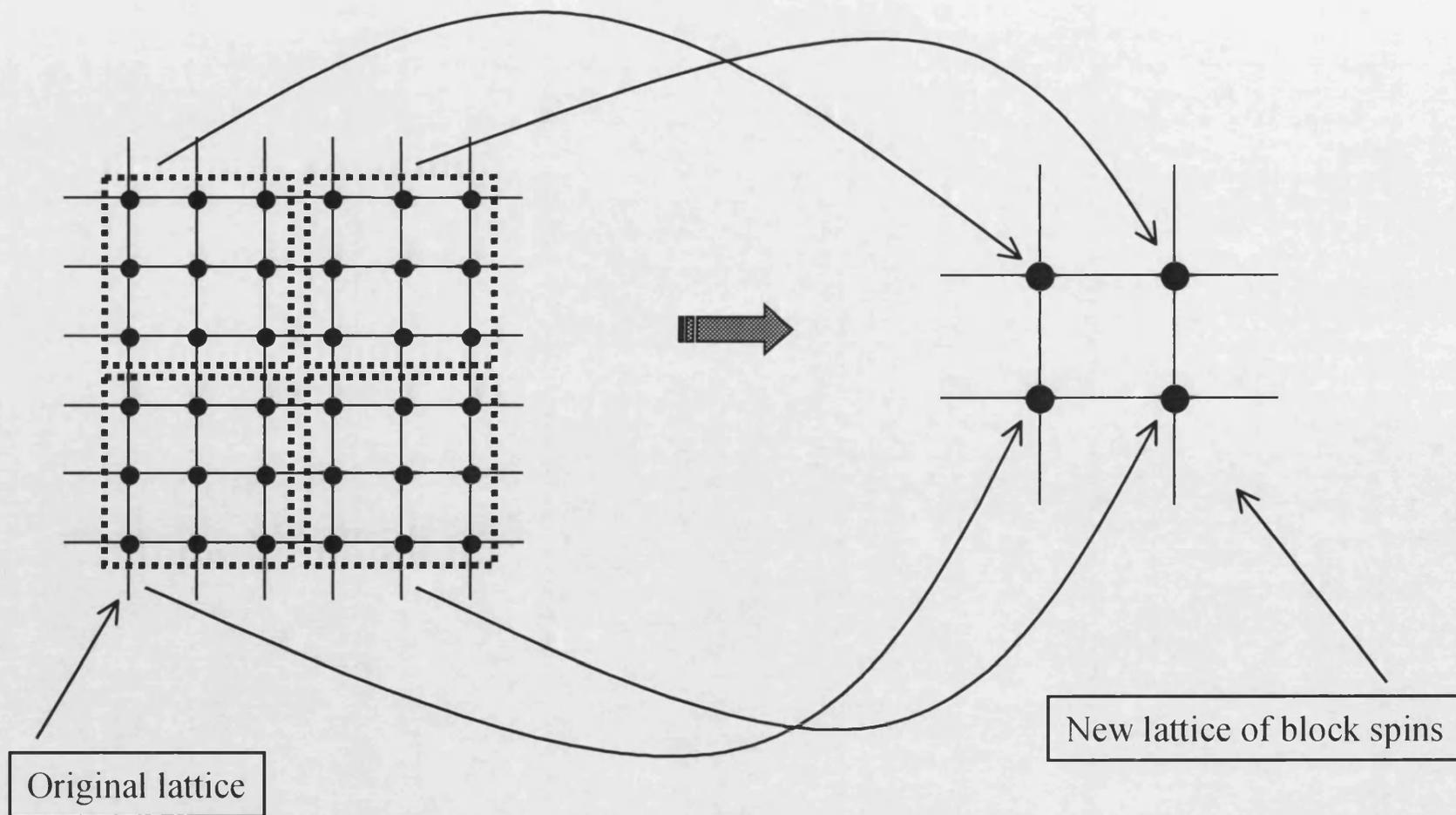


Figure 5.3: Generation of next-nearest neighbour couplings

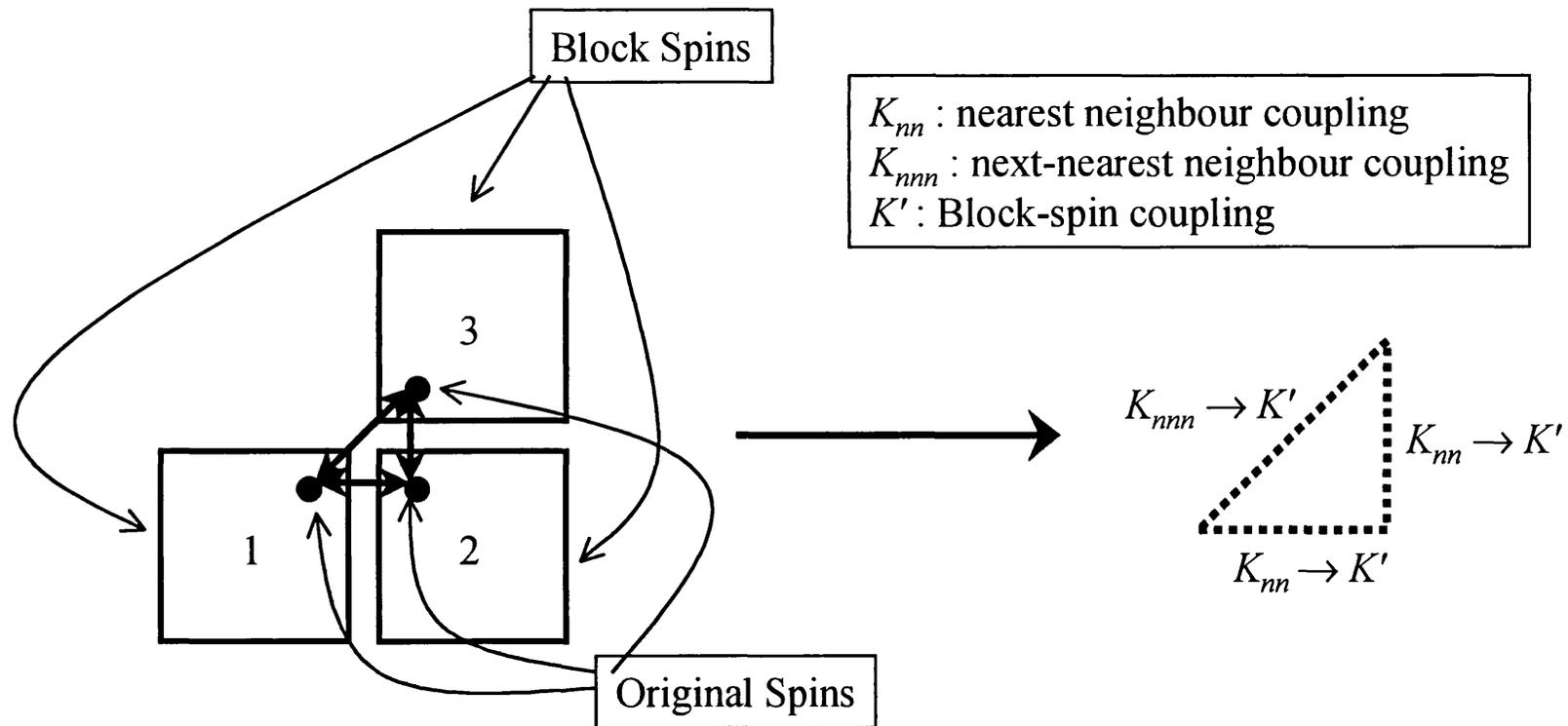
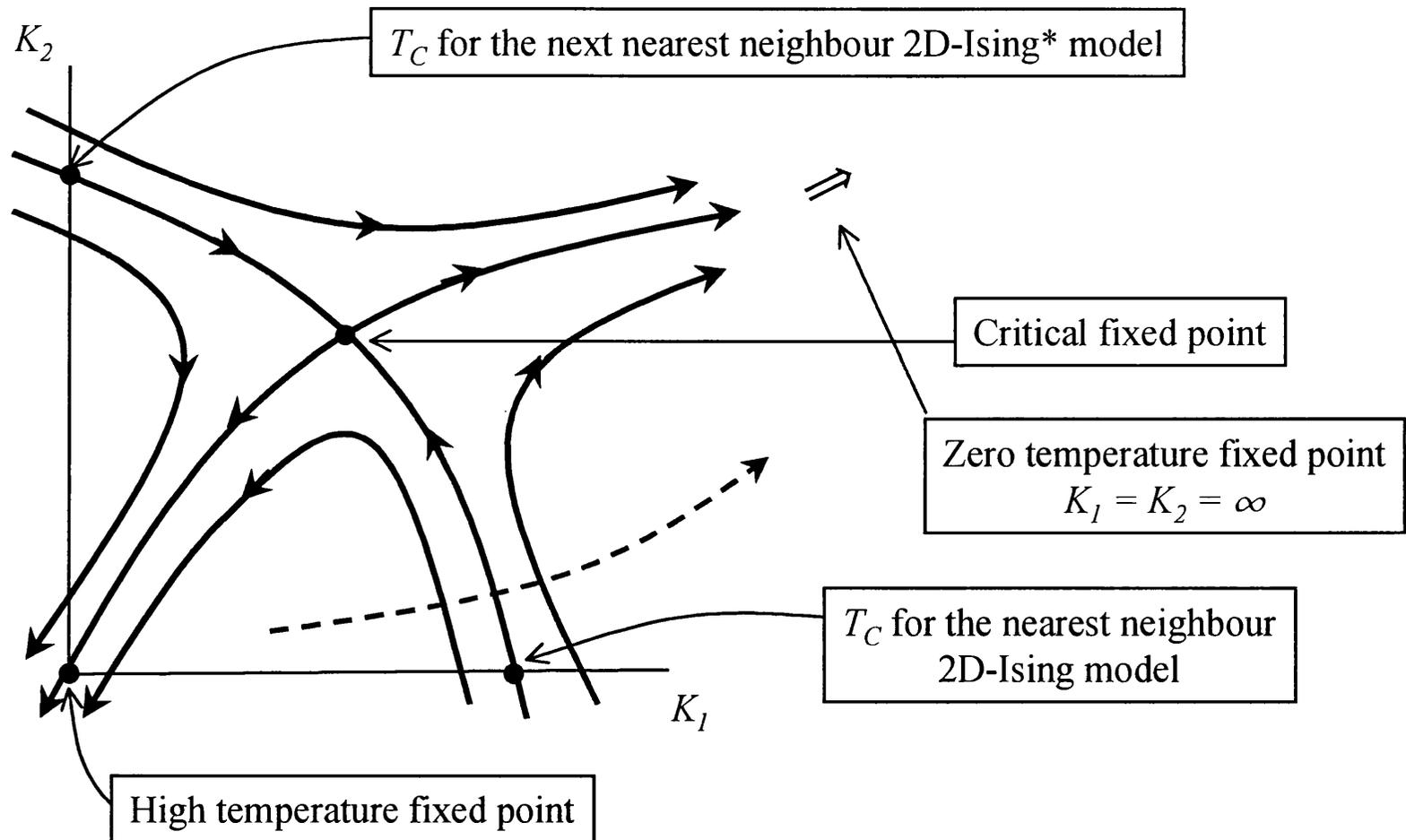


Figure 5.4: RG flows and Fixed Points



BIBLIOGRAPHY

- Achinstein, P. and Hannaway, O. (eds.) 1985, *Experiments and Observation in Modern Science*, Boston, MA: MIT Press.
- Anderson, P.W. 1963, *Concepts in Solids*, Reading, MA: W.A. Benjamin.
- Anderson, P. W. 1984, *Basic Notions of Condensed Matter Physics*, Reading, MA: Addison-Wesley.
- Anderson, P. W. 1994, *A Career in Theoretical Physics*, Singapore: World Scientific.
- Ashcroft, N. W. and Mermin, N. D. 1976, *Solid State Physics*, New York: Holt, Rinehart & Winston.
- Auyang, S. Y. 1998, *Foundations of Complex-system Theories in Economics, Evolutionary Biology and Statistical Physics*, Cambridge: Cambridge Univ. Press.
- Bailer-Jones, Daniela 1999, 'Tracing the Development of Models in the Philosophy of Science', in Magnani, Nersessian and Thagard 1999.
- Bailer-Jones, Daniela 2000, 'Modelling Extended Extragalactic Radio Sources', *Studies in History and Philosophy of Modern Physics* **31**: 49-74.
- Balzer, W., Moulines, C.U. and Sneed, J.D. 1987, *An Architectonic for Science: The Structuralist Program*, Dordrecht: Reidel.
- Batterman, R. W. 1997, "'Into the Mist': Asymptotic Theories on a Caustic', *Stud. Hist. Phil. Mod. Phys.* **28**: 395-413.
- Batterman, R. W. 1998, 'Why Equilibrium Statistical Mechanics Works: Universality and the Renormalization Group', *Philosophy of Science* **65**: 183-208.
- Batterman, R. W. 2000, "Multiple Realizability and Universality", *British Journal for Philosophy of Science* **51**: 115-145.
- Brown, L. M. (ed.) 1993, *Renormalization: From Lorentz to Landau (and Beyond)*, New York: Springer-Verlag.
- Brush, S.G. 1967, 'History of the Lenz-Ising model', *Reviews of Modern Physics* **30**: 883-93.
- Brush, S.G. 1983, *Statistical Physics and the Atomic Theory of Matter: from Boyle and Newton to Landau and Onsager*, Princeton, NJ: Princeton Univ. Press.

- Cao, T. Y. 1998, *Conceptual Developments of 20th Century Field Theories*, Cambridge: Cambridge Univ. Press.
- Cao, T. Y. (ed.) 1999, *Conceptual Foundations of Quantum Field Theory*, Cambridge: Cambridge Univ. Press.
- Cao, T. Y. and Schweber, S. 1993, 'The Conceptual Foundations and Philosophical Aspects of Renormalization Theory', *Synthese* 97: 32-108.
- Cardy, J. 1996, *Scaling and Renormalization in Statistical Physics*, Cambridge: Cambridge Univ. Press.
- Cartwright, N. 1983, *How the Laws of Physics Lie*, Oxford: Clarendon Press.
- Cartwright, N. 1989, *Nature's Capacities and their Measurement*, Oxford: Clarendon Press.
- Cartwright, N. 1999a, 'Models and the Limits of Theory: Quantum Hamiltonians and the BCS Models of Superconductivity', in Morgan and Morrison 1999, pp. 241-281.
- Cartwright, N. 1999b, *The Dappled World: A Study of the Boundaries of Science*, Cambridge: Cambridge Univ. Press.
- Cartwright, N., Shomar, T. and Suarez, M. 1995, 'The Tool-Box of Science', *Poznan Studies in the Philosophy of Science* 44: 137-149.
- Chaikin, P.M. and Lubensky, T.C. 1995, *Principles of Condensed Matter Physics*, Cambridge: Cambridge Univ. Press.
- Chang, H. 1999, 'History and Philosophy of Science as a Continuation of Science by Other Means', *Science and Education* 8: 413-425.
- Chang, H. 2000, 'Spirit, Air and Quicksilver: the Search for the 'Real' Scale of Temperature', *Historical Studies in the Physical and Biological Sciences*.
- Chang, H. forthcoming, 'How to Take Realism Beyond Foot-stamping', *Philosophy*.
- Craig, Edward (ed.) 1998, *Routledge Encyclopedia of Philosophy*, version 1.0, London and New York: Routledge.
- Darden, Lindley (ed.) 1997, *PSA 1996*, Vol. 2, Philosophy of Science Association.
- Davies, Paul (ed.) 1989, *The New Physics*, Cambridge: Cambridge Univ. Press.
- Domb, C. (ed.) 1972, *Phase Transitions and Critical Phenomena*, vol. 1, London: Academic Press.

- Domb, C. 1985, 'Critical Phenomena: A Brief Historical Survey', *Contemp. Phys.*, **26**: 49-72.
- Domb, C. 1996, *The Critical Point: A historical introduction to the modern theory of critical phenomena*, London: Taylor and Francis.
- Domb, C. and Green, M.S. (eds.) 1974, *Phase Transitions and Critical Phenomena*, vol. 3, London: Academic Press.
- Ehrlich, P. 1982, 'Negative, Infinite and Hotter than Infinite Temperatures', *Synthese* **50**: 233-277.
- Feshbach, H., Matsui, T. and Oleson, A. (eds.) 1988, *Niels Bohr: Physics and the World (Proceedings of the Niels Bohr Centennial Symposium)*, London: Harwood Academic.
- Feyerabend, Paul K. 1981, *Problems of Empiricism*, Philosophical Papers Vol. 2, Cambridge: Cambridge University Press.
- Feynman, R.P. 1972, *Statistical Mechanics: A Set of Lectures*, Reading, MA: W.A. Benjamin.
- Fine, A, Forbes, M. and Wessels, L. (eds.) 1990, *PSA 1990*, vol. 2, East Lansing, MI: Philosophy of Science Association.
- Fisher, M. E. 1981, 'Simple Ising Models Still Thrive', *Physica*, **106A**: 28-47.
- Fisher, M. E. 1988, 'Condensed Matter Physics: Does Quantum Mechanics Matter?', in Feshbach, Matsui and Oleson 1988, pp. 65-116.
- Fisher, M. E. 1999, 'Renormalization Group Theory: its basis and formulation in statistical physics' in Cao 1999, pp. 89-135.
- French, Peter A. et al. (eds.) 1993, *Midwest Studies in Philosophy XVIII: Philosophy of Science*, Notre Dame, IN: University of Notre Dame Press.
- Friedman, Michael 1974, 'Explanation and Scientific Understanding', in Pitt 1988.
- Gähde, U. 1995, 'Holism and the Empirical Claim of Theory-Nets', in Moulines 1995.
- Giere, R. 1988, *Explaining Science: A Cognitive Approach*, Chicago, IL: University of Chicago Press.
- Goldenfeld, N. 1992, *Lectures on Phase Transitions and the Renormalization Group*, Reading, MA: Addison-Wesley.
- Griffiths, R. 1972, "Rigorous Results and Theorems", p. 11, in Domb 1972.

- Hacking, Ian 1983, *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science*, Cambridge: Cambridge Univ. Press.
- Hartmann, S. 1999, 'Models and Stories in Hadron Physics', in Morgan and Morrison 1999, pp. 326-346.
- Hasok, Chang 1999, 'How to Take Realism Beyond Foot-Stamping', manuscript.
- Hempel, C. G. 1966, *Philosophy of Natural Science*, Englewood Cliffs, NJ: Prentice-Hall.
- Hoddeson, L. et al. 1992, 'Collective Phenomena', in Hoddeson, Braun, Teichmann and Weart 1992.
- Hoddeson, L, Braun, E., Teichmann, J. and Weart, S. (eds.) 1992, *Out of the Crystal Maze*, Oxford: Oxford Univ. Press.
- Huggett, N. and Weingard, R. 1995, 'The Renormalisation Group and Effective Field Theories', *Synthese* **102**: 171-194.
- Huggett, N. and Weingard, R. 1996, 'Critical Review: Paul Teller's *Interpretive Introduction to Quantum Field Theory*', *Philosophy of Science* **63**: 302-314.
- Hughes, R. I. G. 1993, 'Theoretical Explanation', in French et al. 1993, pp. 132-153.
- Hughes, R. I. G. 1996, 'The Semantic View of Theories', manuscript.
- Hughes, R. I. G. 1997, 'Models and Representation', in Darden 1997, pp. S325-S336.
- Hughes, R. I. G. 1999, 'The Ising Model, Computer Simulation, and Universal Physics', in Morgan and Morrison 1999, pp. 97-145.
- Kadanoff, L. P. 1966, 'Scaling Laws for Ising Models near T_c ', *Physica* **2**: 263-272.
- Kemeny, J. and Oppenheim, P. 1956, "On Reduction", *Philosophical Studies* **7**: 10-19.
- Kitcher, Philip 1981, 'Explanatory Unification', in Pitt 1988.
- Kitcher, Philip 1989, 'Explanatory Unification and the Causal Structure of the World', in Kitcher and Salmon 1989.
- Kitcher, Philip and Salmon, Wesley C. (eds.) 1989, *Minnesota Studies in the Philosophy of Science, Volume XIII: Scientific Explanation*, Minneapolis, MN: Univ. of Minnesota Press.
- Kondepudi, Dilip and Prigogine, Ilya 1998, *Modern Thermodynamics: From Heat Engines to Dissipative Structures*, New York: John Wiley & Sons.

- Krylov, N. 1979, *Works on the Foundations of Statistical Physics*, Princeton: Princeton Univ. Press.
- Kubo, R., Toda, M. and Hashitsume 1985, *Statistical Physics II: Nonequilibrium Statistical Mechanics*, Berlin: Springer-Verlag.
- Kuhn, T. S. 1970, *The Structure of Scientific Revolution*, 2nd edition, Chicago, IL: Chicago University Press.
- Lakatos 1970, 'Falsification and the Methodology of Scientific Research Programmes', in Lakatos and Musgrave 1970, pp. 91-196.
- Lakatos, Imre 1978, *The Methodology of Scientific Research Programmes*, Philosophical Papers Vol. 1, (eds.) John Worrall and Gregory Currie, Cambridge: Cambridge University Press.
- Lakatos, Imre and Musgrave, Alan (eds.) 1970, *Criticism and the Growth of Knowledge*, Cambridge: Cambridge University Press.
- Laymon, R. 1985, 'Idealizations and the Testing of Theories by Experimentation', in Achinstein and Hananway 1985.
- Laymon, R. 1989, 'Cartwright and the Lying Laws of Physics', *Journal of Philosophy* **86**: 353-372.
- Laymon, R. 1990, 'Computer Simulations, Idealization and Approximation', in Fine, Forbes and Wessels 1990.
- Laymon, R. 1998, 'Idealizations', in Craig 1998.
- Leggett, A. J. 1985, *Problems of Physics*, Oxford: Oxford Univ. Press.
- Leibniz, G.W. 1985 [1692] 'Critical Remarks Concerning the General Part of Descartes', in *Methodology and Other Philosophical Essays* (trans. and ed. by Paul Schrecker and Anne Martin Schrecker), New York: Macmillan.
- Liboff, Richard L. 1990, *Kinetic Theory: Classical, Quantum and Relativistic Descriptions*, London: Prentice-Hall.
- Lovesey, Stephen W. 1980, *Condensed Matter Physics: Dynamic Correlations*, Reading MA: W.A. Benjamin.
- Magnani, L., Nersessian, N.J. and Thargard, P. (eds.) 1999, *Model-Based Reasoning in Scientific Discovery*, New York: Kluwer Academic/Plenum Publishers.

- Maudlin, Tim 1994, *Quantum Non-Locality and Relativity: Metaphysical Intimations of Modern Physics*, Oxford: Blackwell.
- McMullin, Ernan 1885, 'Galilean Idealization', *Studies in History and Philosophy of Science* **16**: 247-273.
- Morgan, Mary S. and Morrison, Margaret (eds.) 1999, *Models as Mediators: Perspectives on Natural and Social Science*, Cambridge: Cambridge University Press.
- Moulines, C.U. (ed.) 1995, *Structuralist Theory of Science: Focal Issues, New Results*, Berlin: de Gruyter.
- Negele, J.W. and Orland, H. 1987, *Quantum Many-Particle Systems*, Reading MA: Addison-Wesley.
- Pitt, Joseph C. (ed.) 1988, *Theories of Explanation*, Oxford: Oxford Univ. Press.
- Redhead, Michael 1980, 'Models in Physics', *British Journal for the Philosophy of Science* **31**: 145-163.
- Risken, H. 1989, *The Fokker-Planck Equation: Methods of Solution and Application*, 2nd Edition, Berlin: Springer-Verlag.
- Rushbrooke, G.S., Baker Jr., George A. and Wood, P.J. 1974, 'Heisenberg Model', in Domb and Green 1974.
- Samuelson, Paul 1958, 'An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money', *Journal of Political Economy* **66**: 467-82.
- Savage, C. Wade (ed.) 1990, *Minnesota Studies in the Philosophy of Science XIV: Scientific Theories*, Minneapolis MN: University of Minnesota Press.
- Schweber, S. 1994, *QED and the Men Who Made It*, Princeton, NJ: Princeton Univ. Press.
- Sklar, Lawrence 1974, *Space, Time, and Spacetime*, Berkeley, CA: Univ. of California Press.
- Sklar, Lawrence 1993, *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics*, Cambridge: Cambridge University Press.
- Ter Harr, D. and Wergeland, H. 1966, *Elements of Thermodynamics*, Reading, MA: Addison-Wesley.

- Teller, P. 1994, *An Interpretive Introduction to Quantum Field Theory*, Princeton, NJ: Princeton Univ. Press.
- Teller, P. 1998, 'On Huggett and Weingard's Review of *An Interpretive Introduction to Quantum Field Theory*: Continuing the Discussion', *Philosophy of Science* **65**: 151-161.
- Thouless, David 1989, 'Condensed Matter Physics in Less than Three Dimensions', in Davies 1989.
- Toda, M., Kubo, R. and Saitô, N. 1983, *Statistical Physics I: Equilibrium Statistical Mechanics*, Berlin: Springer-Verlag.
- Uzunov, D. I. 1993, *Introduction to Theory of Critical Phenomena: Mean Field, Fluctuations and Renormalization*, Singapore: World Scientific.
- Van Fraassen, Bas C. 1980, *The Scientific Image*, Oxford: Clarendon Press.
- Von Plato, Jan 1994, *Creating Modern Probability: Its Mathematics, Physics and Philosophy in Historical Perspective*, Cambridge: Cambridge Univ. Press.
- Wise, M. N. and Brock, D. C. 1998, 'The Culture of Quantum Chaos', *Studies in History and Philosophy of Modern Physics*, **29B**: 369-390.
- Wu, F. Y. 1982, 'The Potts Model', *Review of Modern Physics*, **54**: 235-287.
- Yang, C.N. and Lee, T.D. 1952, 'Theory of Phase Transitions', *Physical Review* **87**: 404-11.