

# A Schema for Duality, Illustrated by Bosonization

Sebastian De Haro and Jeremy Butterfield

*Trinity College, Cambridge, CB2 1TQ, United Kingdom  
Department of History and Philosophy of Science, University of Cambridge  
Free School Lane, Cambridge CB2 3RH, United Kingdom*

sd696@cam.ac.uk, jb56@cam.ac.uk

Friday 22 Dec 2017

## Abstract

In this paper we present a schema for describing dualities between physical theories (Sections 2 and 3), and illustrate it in detail with the example of bosonization: a boson-fermion duality in two-dimensional quantum field theory (Sections 4 and 5).

The schema develops proposals in De Haro (2016, 2016a): these proposals include construals of notions related to duality, like representation, model, symmetry and interpretation. The aim of the schema is to give a more precise criterion for duality than has so far been considered.

The bosonization example, or boson-fermion duality, has the feature of being *simple* yet *rich enough* to illustrate the most relevant aspects of our schema, which also apply to more sophisticated dualities. The richness of the example consists, mainly, in its concern with two non-trivial *quantum field theories*: including massive Thirring-sine-Gordon duality, and non-abelian bosonization.

This prompts two comparisons with the recent philosophical literature on dualities:—

(a) Unlike the standard cases of duality in quantum field theory and string theory, where only specific simplifying limits of the theories are explicitly known, the boson-fermion duality is known to hold *exactly*. This exactness can be exhibited explicitly.

(b) The bosonization example illustrates both the cases of isomorphic and *non-isomorphic* models: which we believe the literature on dualities has not so far discussed.

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
1.1	The schema . . . . .	4
1.2	Bosonization and other dualities . . . . .	7
1.2.1	Examples of dualities . . . . .	8
1.2.2	Bosonization introduced . . . . .	9
<b>2</b>	<b>The Schema Introduced: Theories and Models</b>	<b>12</b>
2.1	Duality's scientific importance . . . . .	12
2.2	Theories and Models . . . . .	15
2.2.1	Bare theories . . . . .	15
2.2.2	Models . . . . .	16
2.2.3	Notations for models; model roots and model triples . . . . .	16
2.3	Interpretations of theories and models . . . . .	18
2.4	Isomorphisms: defining theories by abstraction from models . . . . .	21
<b>3</b>	<b>Duality and Symmetry</b>	<b>23</b>
3.1	Symmetries of theories and their models . . . . .	24
3.1.1	Symmetries of uninterpreted theories and models . . . . .	24
3.1.2	Interpretations respect symmetries . . . . .	28
3.2	Duality as isomorphism of models . . . . .	30
3.2.1	Duality as isomorphism . . . . .	31
3.2.2	A logically weak but physically strong notion of duality . . . . .	34
3.2.3	Duality and interpretation . . . . .	36
3.2.4	Combining duality and symmetries . . . . .	39
<b>4</b>	<b>The Basic Boson-Fermion Duality</b>	<b>43</b>
4.1	The free, massless boson . . . . .	44
4.2	The free, massless Dirac fermion . . . . .	49
<b>5</b>	<b>Boson-Fermion Duality Illustrates the Schema</b>	<b>51</b>
5.1	The duality dictionary . . . . .	51
5.2	Two isomorphic model triples . . . . .	53
5.2.1	States . . . . .	53
5.2.2	Quantities . . . . .	54
5.2.3	Dynamics . . . . .	56
5.2.4	Interpretation . . . . .	57
5.3	Defining a theory from the two model triples . . . . .	58
5.3.1	Two vulnerabilities . . . . .	58
5.3.2	Defining the theory . . . . .	60
5.4	Further abstraction . . . . .	61
5.5	General boson-fermion dualities . . . . .	63
5.5.1	Duality between the Thirring model and the sine-Gordon model . . . . .	63
5.5.2	Non-abelian boson-fermion dualities . . . . .	65

Envoi	66
Appendix. Some Elements of Conformal Field Theory	67
References	68

## 1 Introduction

In this paper we present a schema for describing dualities between physical theories (Sections 2 and 3). Then we illustrate it in detail with the example of bosonization: a boson-fermion duality in two-dimensional conformal field theories (Sections 4 and 5).

Before we introduce these two parts in turn (Sections 1.1, 1.2), we briefly set our project in the context of the legacy of Hilbert’s work, a hundred and more years ago, on the foundations of physics and its axiomatisation—work which it is an honour to commemorate. This legacy is of course so broad and deep that we can only touch on it. We will confine ourselves to recalling two Hilbertian ideas about the role of axiomatizing a theory (either mathematical or physical): ideas which obviously relate to our project, and which we will return to in Sections 2.4 and 5.4.

The background for both ideas, indeed for all Hilbert’s work in axiomatisation (such as his axiomatisation in 1899 of Euclidean geometry, and his choosing as the sixth Problem in his famous 1900 ‘To Do’ list, the axiomatisation of mechanics and geometry) was, of course, the development of formal methods, in particular axiomatic studies, in all of mathematics from about 1850.<sup>1</sup>

First, there is the idea that an axiom system can be realized, i.e. made true, by very different models. (Recall Hilbert’s famous remark that ‘one must be able to say at all times—instead of points, straight lines and planes—tables, chairs and beer-mugs’.) We shall see that duality, in the sense nowadays used in physics, gives illustrations of this idea. Indeed, very *vivid* illustrations. For duality, in physicists’ current jargon, involves there being two theories that look very different (not just in their formulation and concepts, but also apparently in the topics they are about) that are in some sense equivalent. In particular, there is a ‘dictionary’ that pairs off the concepts in one theory with those in the other. Thus in our example of bosonization, one theory will describe fermions, while the other describes bosons: very different field-contents. So duality illustrates this Hilbertian idea: but on a grand scale! For now, it is entire physical theories that are the very different realizations of some *common core* axiom system. Indeed, as explained in Section 1.1: we shall call the two sides, i.e. the items that are dual to each other, *models* (viz. models of the single common core), rather than *theories*. So this usage echoes the Hilbertian idea.<sup>2</sup>

---

<sup>1</sup>Again, we can only touch on the vast literature. For Hilbert’s Problems of 1900, cf. e.g. Gray (2000), Grattan-Guinness (2000). For the sixth Problem, cf. Corry (1999, 2004, 2006, 2018), and Stachel (2018). For some context for Hilbert’s famous ‘beer-mug’ remark, cf. Kennedy (1972). Finally, we note that Gray (2008) makes an interesting case that this broad development represented a rise of ‘modernism’, in a sense analogous to that in art and literature.

<sup>2</sup>The Hilbertian idea has, of course, other important facets: for example, in fostering the idea that an axiom system—or more generally, a doctrine expressed in language—‘implicitly defines’ its terms. This

Second, Hilbert sees the activity of axiomatisation, *not* as giving a theory its final form and so best undertaken when (one hopes!) it is fully understood, but as worthwhile even when we recognize that the theory is far from its final form. For it is worthwhile precisely in order to deepen our understanding of the theory. Again this idea has been very influential. In physics, the best known example of its influence is no doubt von Neumann’s monumental treatise on quantum mechanics (1932), which over the decades has spawned so many axiomatic studies of quantum mechanics, most directly the quantum logic approach. But also in philosophy, the idea was very influential. Reichenbach and other logical empiricists saw axiomatisation as the way by which philosophers could clarify scientific theories (and in particular, distinguish their factual and conventional contents—a project that, for the logical empiricists, was the distinctive task of philosophy). Thus we think of our own project—to formulate in general, almost formal, terms, the notion of duality (Sections 2 and 3), and to illustrate this in bosonization (Sections 4 and 5)—as an exercise in the tradition of this idea.<sup>3</sup>

In Section 1.1, we briefly introduce our notions of theory and model, and of duality as an isomorphism between models. We motivate our usage and compare the notion of duality to the analogous notion of symmetry. In Section 1.2, we introduce our main example, of bosonization, and compare this example to other examples used in the literature on dualities.

## 1.1 The schema

The schema develops proposals in De Haro (2016, 2016a). Like other authors, we take duality to be a suitable relation of equivalence between physical theories. The main features of our schema are that:

- 1): we distinguish uninterpreted theories, which we call *bare theories*, from interpreted theories;
- 2): we emphasize that, wholly independently of issues of interpretation, a bare theory can have many realizations, which we call *models*;
- 3): we take duality to be an isomorphism between two models of a single bare theory.

Of these three features, it is 2) and 3) that are the distinctive ones. For several authors also define duality in terms of uninterpreted theories. This has the advantage of making verdicts of duality not beholden to semantic issues, and so less vague or even controversial. And it allows cases of duality without any sort of physical or semantic equivalence—which certainly occur, e.g. Kramers-Wannier duality between the high and low temperature regimes of the statistical mechanics of a lattice. But features 2) and 3) make duality an equivalence (formally: an isomorphism) between items that are not only

---

has been very influential in the foundations of logic and mathematics, beginning with Hilbert’s debate with Frege. It has also of course been contested: in the face of non-categoricity, the claim to ‘define’ terms by a body of doctrine containing them is questionable.

<sup>3</sup>For some ‘post-Hilbert’ history of axiomatisation as ‘deepening the foundations’, cf. Stöltzner (2000, 2001). But we should add that we do not endorse the logical empiricist project of distinguishing, once and for all, the factual and conventional parts of a theory: our misgivings are essentially those of Putnam (1962).

uninterpreted, but also *more specific* than an uninterpreted theory: viz. realizations—which we will call ‘models’—of a (single) bare theory. A prototypical example is: taking a bare theory to be an abstract algebra of quantities (maybe also equipped with a dynamics, viz. as a 1-parameter group of automorphisms, and a set of abstract states, i.e. rules for evaluating (i.e. assigning values to) the quantities): a model or realization is given by a representation (in the mathematical sense) of the algebra, together with a realization of the rules for evaluating the quantities, for the representation in question, i.e. a set of maps to the relevant field, of complex or real numbers.

We shall say ‘model’ rather than ‘realization’, not least for brevity. But we should disavow, here at the outset, some misleading connotations of the word ‘model’. Indeed, there are three misleading connotations. The word ‘model’, as contrasted with ‘theory’, often connotes:

(i): a specific solution (at a single time: or for all times, i.e. a possible history) for the physical system concerned, whereas the ‘theory’ encompasses all solutions—and in many cases, for a whole class of systems;

(ii): an approximation, in particular an approximate solution, whereas the ‘theory’ deals with exact solutions;

(iii): being part of the physical world (in particular, being empirical, and-or observable) that gives the interpretation, whereas the ‘theory’ is of course not part of the physical world, and so stands in need of interpretation.

So we stress that our use of ‘model’ rejects all three connotations. As we just said: for us, a model is a specific realization—one might say ‘version’—of a theory. But it is a version of a bare, i.e. uninterpreted, theory, and the version is itself bare, i.e. uninterpreted. So a model adds details—we shall say: ‘specific structure’—to its bare theory. But these details are *not* a matter of specifying: (i) a solution or history of the system; or (ii) approximation(s); or (iii) interpretation(s). Rather, the extra details are extra mathematical structure: just like a representation of a group or an algebra has extra details or structure, beyond that of the group or algebra of which it is a representation.<sup>4</sup>

Indeed, for clarity later on, we should distinguish two broad kinds of extra detail or structure that a model adds. Again, group representations provide obvious—and countless—examples.

(A): The ‘concreteness’ of a specific mathematical object: such as  $GL(n, \mathbb{C})$ , the general linear group over  $\mathbb{C}^n$ , or any subgroup of it—any of which is a ‘concrete’, not abstract, group. (Agreed, the concrete vs. abstract contrast is flexible; but this will not matter for anything that follows.)

---

<sup>4</sup>We agreed that for our notion, the word ‘model’ has disadvantages. But note that other words also have disadvantages. For example: ‘formulation’ connotes that any two formulations of a theory are ‘notational variants’, i.e. fully equivalent: they say exactly the same thing about the world. But that is far from true for our notion (and this matches the connotations of ‘model’): for us, two models of a bare theory are in general not isomorphic, and not in any sense equivalent; and so typically, it is surprising to find two isomorphic models, i.e. to find a duality. Other examples: ‘realization’, ‘instance’ and ‘instantiation’ connote being part of the physical world, as in ‘the mechanism/hardware which realizes some specific function/software’, or ‘the object is an instance/instantiation of the predicate’—which is the misleading connotation (iii) above.

Notice that in theoretical physics, the use of model is, roughly, between: (a) our use, and (b) (ii) and (iii) above: e.g. the ‘massive Thirring model’ or the ‘sine-Gordon model’.

(B): The fact that the mathematical notion of representation requires homomorphism, not isomorphism: i.e. it allows non-injectivity and non-surjectivity. Thus two representations of an abstract group  $G$  can be non-isomorphic as groups—i.e. different, even as described in only abstract group-theoretic terms—to  $G$ ; and of course also, non-isomorphic to each other.

Of course, these kinds (A) and (B) of ‘extra detail’ usually occur together: just think of how every abstract group can be represented by the trivial one element subgroup of  $GL(n, \mathbb{C})$ , the  $n \times n$  identity matrix.

Our picture is therefore of a bare theory, that can be realized (we will say: modelled) in various ways: like the different representations of an abstract group or algebra. And these models are in general *not* isomorphic, since they differ from one another in their specific structure: like inequivalent representations of a group. But we say: *when the models are isomorphic, we have a duality.*

In Sections 2 and 3, we will develop this view of duality (with Section 2 dealing with theories, models and interpretations, and Section 3 with symmetries). We end this Subsection with two further remarks about our schema. The first motivates our usage of ‘theory’ and ‘model’; the second compares duality with that more familiar topic, symmetry.

(1): *Motivating our usage:*—Dualities in physics give a rationale for our usage of ‘theory’ and ‘model’, as introduced above. (This rationale does not depend on the contrast between interpreted and uninterpreted (bare) theories; and so we temporarily set that aside.) Recall that in both physics and philosophy of physics, ‘theory’ is usually taken as something like a set of differential equations, and ‘model’ is usually taken as something like a solution to such a set. But a duality often shows us that what we first considered as distinct theories can, or should, be seen as the same theory, in two guises. Agreed, that is very rough speaking: which will of course be clarified in what follows. But for now, we only need the point that this kind of surprising discovery prompts us to move our usage of ‘theory’ “one level up”. After all: if two sets of differential equations somehow express the same theory, then a theory cannot be identified with such a set. Besides: if we thus move our usage of ‘theory’ one level up, we can still keep the usual intuitive idea of how ‘theory’ and ‘model’ are related—viz. that a model is a realization, or instance, of a theory—by correspondingly moving our usage of ‘model’ one level up. And this is what we have proposed.

To sum up: the broad and widely-agreed idea, that in physics a duality often suggests that the two theories concerned, though they look different, are in fact ‘the same’, motivates our proposed usage of ‘theory’ and ‘model’.

(2): *Analogy with symmetry:*—The analogy is (as is often remarked) that ‘a duality is like a symmetry, but at the level of a theory’. Here, and for the rest of this Subsection, we will *temporarily* set aside our jargon just announced, of ‘theory’ vs. ‘model’. We will temporarily join the literature’s usual jargon of taking a theory to be interpreted, and a model to be—not a ‘version’ of the theory with some specific structure of its own—but a solution (or representative of a solution) of the theory.

That is, the analogy is: while a symmetry carries a state to another state that is ‘the

same’ or ‘matches it’, a duality carries a theory to another theory that is ‘the same’ or ‘matches it’. We will endorse this analogy. So the interesting questions, for both sides of the analogy, will concern the different ways to make precise ‘the same’ or ‘matches’. We give details (respecting our proposed ‘theory’ vs. ‘model’ usage!) in Sections 3.1 and 3.2. But the questions about making precise ‘the same’/‘matches’ can be introduced as follows.

A symmetry  $a$  (we write  $a$  for ‘automorphism’) carries a state  $s$  in a state space  $\mathcal{S}$  to another state  $a(s)$ : thanks to  $a$  being a symmetry, the two states  $s$  and  $a(s)$  assign the same values to all the quantities (i.e. magnitudes) in some salient, usually large, set of quantities. The question then arises: do  $s$  and  $a(s)$  represent the very same physical state of affairs, or scenario—or in philosophers’ jargon: the same possible world?

The answer, in full generality, is of course: ‘No’. That is: not always. But for a large enough set of quantities being preserved; and in particular for a theory that is a ‘toy cosmology’ (i.e. a theory whose system of interest is a cosmos, with no external environment, so that there are no relational quantities whose values are *not* preserved by  $a$ ): there is a tradition of answering ‘Yes’.

Debate then ensues about:

- (i) what are the general conditions for the ‘Yes’ answer being correct? and
- (ii) what does the ‘Yes’ answer imply about the propriety of—perhaps even the requirement of—moving to a reduced formalism, i.e. one in which states are taken as the orbits, in the given formalism, of the action of the symmetry  $a$ ?<sup>5</sup>

So, turning to our topic of dualities: we endorse the analogy. We will say, roughly speaking, that: a theory  $T$  is mapped by duality  $d$  to a theory  $d(T)$  which is ‘the same’ as  $T$ . This will be made precise in various ways. But it is worth stressing now, in line with the three features 1), 2) and 3) we listed at the start of this Subsection, that:

(a): We take theories to be initially uninterpreted: so it will not follow from the existence of a duality map  $d$  that  $T$  and  $d(T)$  are wholly equivalent (‘state the very same propositions’).

(b): We make explicit the interpretation of a theory’s formalism: so there will be interpretation maps  $I$  acting on both the theory  $T$  and its dual  $d(T)$ .

(c): For a given theory, we distinguish different realizations of it, which we call ‘models’. Duality is an isomorphism between such models: an isomorphism that is often surprising since the models, despite their common core, “look different”.

## 1.2 Bosonization and other dualities

In this Subsection, we motivate our choice of bosonization as the illustration of our schema. We first sketch a spectrum of examples of dualities (Section 1.2.1). Then we describe how bosonization strikes a balance between mathematical rigour and physical interest, and introduce its main features (Section 1.2.2).

---

<sup>5</sup>A bit more precisely: states would be taken as the union of the orbits for all the symmetries for which the ‘Yes’ answer is true. For recent work on the debate about (i) and (ii), cf. Caulton (2015), Dewar (2016) and Weatherall (2015).

### 1.2.1 Examples of dualities

Recent philosophical literature on duality and theoretical equivalence has dealt with three main kinds of examples:—

(a): equivalence between models (in our sense!) formulated in *first-order* (maybe *many-sorted*) *logic*: e.g. definitional equivalence, Morita equivalence, and-or categorical equivalence (e.g. Barrett and Halvorson (2016, 2016a));

(b): categorical equivalence of models (in our sense) of *classical theories* (e.g. Weatherall (2015), Teh (2016));

(c): dualities between models (in our sense) of *quantum theories* whose classical descriptions are very different (e.g. Dieks et al. (2015), De Haro (2017), De Haro et al. (2017), Fraser (2017), Huggett (2017), Rickles (2017), Matsubara (2013)).

The classification (a)-(c) is arranged in increasing order of physical (not mathematical!) sophistication. Consequently, there is also decreasing mathematical rigour, as one moves from kind (a) to kind (c):

Kind (a): These examples have the advantage of being very simple, in their reliance on first-order logic only: and so, the notions of equivalence in question can be defined rigorously. But in their simplicity, the notions developed, and the examples given, generally do not seem to have sufficient structure that they could describe in detail the sorts of examples that physicists would be interested in. (At any rate, the authors cited do not describe how such logical models can illustrate even the simplest physical models of, say, classical, source-free Maxwell theory: which is, of course, not to claim that this is impossible!).

Kind (b): These examples include some important models of *classical theories*, such as Newtonian gravitation, general relativity, and Yang-Mills theory models. But these examples also have some limitations. (1): To physicists, the example is, typically, not surprising (e.g. Newtonian gravitation being equivalent to *geometrized* Newtonian gravitation). (2): When it *is* surprising (e.g. Teh (2016)), it is not a case of equivalence, but rather of analogy. Furthermore, (3): categorical equivalence has been criticised by Barrett and Halvorson (2016a) for being ‘too liberal’. In our view, the element of ‘surprise’ (see Section 2.1) seems to come with models of *quantum theories*, i.e. examples of (c):

Kind (c): Typical examples of this kind are dualities between very different-looking models of *quantum field theories*, or of *string theories* (cf. Witten (2018)).

This explains the recent interest, shown by both physicists and philosophers of physics, in such dualities. Physicists tend to view dualities as powerful epistemic statements: the epistemic gain being both mathematical and physical. As the mathematical aspect: mirror symmetry is the prime example.<sup>6</sup> Michael Atiyah has characterised the discovery of mirror symmetry as ‘spectacular’: since it established a new link between complex geometry and symplectic geometry, later proven (in one of its simplified versions) by

---

<sup>6</sup>Despite the name of ‘symmetry’, mirror symmetry falls under what we here call a duality. That mirror symmetry is a case of duality—of two different models, rather than a single model, being related—is uncontroversial, and reflected in the literature.



mathematicians (2009, p. 83). As to the physical aspect: gauge-gravity duality is an important example, which has led to both new theoretical developments in quantum gravity, and to new experimental results and ideas (like the explanation of the shear viscosity-to-entropy ratio in a quark-gluon plasma, and recent applications to cosmology: cf. e.g. Ammon and Erdmenger (2015), De Haro et al. (2016)). Other examples are T-duality (related to mirror symmetry) and S-duality: which falls under the same class of dualities as our bosonization example, viz. exchanging Noether charges and topological charges (Castellani (2017)). An important idea of these dualities is that it is the models of the *quantum theories* which are equivalent, while their classical limits are very disparate (differing in the number or the size of the dimensions, the matter content, etc.). These two aspects—physical and mathematical—will be developed in Section 2.1’s discussion of the *scientific importance* of dualities.

But there is a second reason these dualities interest philosophers of physics: which the recent literature has emphasised. Namely, these dualities obviously bear on philosophical questions such as the distinction between theoretical and physical equivalence, emergence (of spacetime, and-or other entities), and realism vs. structuralism. We will return to these questions in another paper.

Agreed: examples of kind (c) also have limitations, as follows. (1): The models (in our sense) are mathematically very difficult; and typically, no exact formulation of the models that are dual is yet known. So the duality, e.g. in the case of gauge-gravity duality, is still—despite all the favourable evidence, in various limits etc.—a conjecture. (2): The physics involved is not yet established, since the models involved either deal with quantum gravity situations (a regime of energies about which experiment has so far given no direct clues: cf. Smolin (2018)), or involve simplifying assumptions about the world, typically a high degree of symmetry (e.g. supersymmetric quantum field theory models).

### 1.2.2 Bosonization introduced

It is clear, in the light of Section 1.2.1, that to illustrate our schema, we should choose an example that judiciously balances the desiderata: on the one hand, (i): mathematical precision and established physics, as in kinds (a) and (b); on the other, (ii): scientific importance, as in kind (c). As we will see in detail in Sections 4 and 5: bosonization or, more precisely, boson-fermion duality in two dimensions, is just such an example.

As to (i): Boson-fermion duality allows a treatment of the quantum theory model that does not need to rely on techniques of approximation such as perturbation theory. For this reason, boson-fermion duality explicitly illustrates our schema: the common core *theory* can be formulated according to our construal in Section 2.2.1, and the two sides of the duality are *models* in the sense of Section 2.2.2. The physics of these models is not speculative, and these 1+1-dimensional models describe systems that can be realised in the lab, e.g. as one-dimensional spin chains (Giamarchi (2003: Chapter 2), Altland and Simons (2010: Sections 4.3 and 9.4.4)).

As to (ii): The scientific importance of the duality is witnessed by three facts. (1): Bosonization involves rich models of *quantum field theories*, and not just classical theories; (2) it is an active area of physics (see e.g. Gogolin et al. (2004), Kopietz (2008)); and (3) it illustrates the surprise that we discuss in Section 2.1, viz. by relating a model of bosons

and a model of fermions.

Bosonization was discovered in two papers by Coleman (1975) and Mandelstam (1975), which built on previous work on the sine-Gordon and Thirring models (Dell’Antonio et al. (1972)). Coleman discovered that the sine-Gordon model (a scalar field whose interaction potential is the cosine of the field) in 1+1 dimensions was equivalent to the *charge-zero sector* of the massive Thirring model (a massive Dirac fermion field with quartic interaction) in 1+1 dimensions.<sup>7</sup> ‘Charge-zero sector’ here refers to the restriction of the physical quantities of the model to pairs of fermionic fields. Thus Coleman wrote:

‘... under the assumption that one can only make particle-antiparticle pairs out of the vacuum, not single particles ... For massless particles in two dimensions, it is quite possible to make a pair that never separates. Such a pair consists of two particles moving in the same direction. The wave functions do not spread; they just move on steadily at the speed of light, and the particles never get away from each other [since there is no other direction in which they could turn]. If the particles had a mass, or if the world were of greater than two dimensions, this would not be possible.’ (p. 2094).

While Coleman’s analysis was perturbative, Mandelstam constructed a map which was exact, and went both ways. Not only could a boson be mapped to a pair of fermions; but also the map could be inverted, so as to map a single fermion to a coherent state of bosons. The construction was non-perturbative, i.e. it did not use perturbation theory. This is related to the fact, already recognised by Coleman (1975: p. 2088), that all divergences that occur in perturbation theory can be removed by normal-ordering the Hamiltonian. Mandelstam also did a canonical treatment of the model, working out canonical commutation relations between the fields and the currents constructed from them, using regularisation and renormalisation. Therefore, boson-fermion duality was proven to be *exact*.

There are three significant features of this duality: features that both (a) justify our claim, two paragraphs above, that boson-fermion duality balances the desiderata (i) and (ii), and (b) bear out the conceptual relevance of the example.

(A) *The duality is exact.* That is: it is valid for all physically interesting values of the parameters, and it does not require the use of perturbation theory. In this respect, boson-fermion duality is closer to kinds (a) and (b) than kind (c). Yet the models related as duals are non-trivial (because massive, or massless, and interacting!: see Section 5.5) quantum field theory models—and in *that* sense they are in kind (c)!

(B) *The duality goes both ways.* It relates boson operators to fermion operators, and vice versa. This is, of course, surprising, since these two kinds of operators have very different properties: both mathematically (e.g. different statistics) and physically (they describe particles with distinct properties). We will explain, in Section 4, how two models with such disparate formulations can nevertheless be isomorphic to each other.

(C) *The duality maps the weak-coupling regime of one model to the strong-coupling*

---

<sup>7</sup>See Section 5.5.1; for the simple, free case, see Section 4. Here of course we adopt the usual theoretical physics usage of ‘model’: cf. the end of footnote 4.

regime of the other, and vice versa (Coleman (1975: p. 3027)):

$$\frac{g}{\pi} = \frac{4\pi}{\beta^2} - 1 . \quad (1)$$

Here,  $\beta$  is the coupling constant of the bosonic model (the sine-Gordon model) and  $g$  is the coupling constant of the fermionic model (the Thirring model: for more details, see Section 5.5.1). Clearly, when  $\beta \rightarrow 0$ ,  $g \rightarrow \infty$ . This attests to the physical richness and, indeed, the non-trivial character of the duality. This weak coupling/strong coupling correspondence has later been found to be a feature of most dualities of kind (c), i.e. dualities in models of quantum field theory and string theory: especially S-duality and gauge-gravity duality.<sup>8</sup> This richness is the main reason why physicists are interested in dualities: since they can learn about the strong-coupling regime of one model (where perturbation theory cannot be used effectively nor reliably, so that the model is in general much harder to deal with) from the weak-coupling regime of the other model (where perturbation theory is usually a good guide). For more discussion of how Eq. (1) contributes to scientific importance, see Section 2.1-(2).

Features (A) and (B) are needed in order that the example illustrate our schema with mathematical precision. We will spell this out in Section 5. Indeed, we believe this is the first conceptual and technical exposition in the philosophical literature of a duality combining the physical interest of kind (c), with features (A)-(B).

Feature (C) relates to another important topic relating to dualities, viz. that of *emergence*. Indeed, a recent theme in the philosophy of physics literature has been the close connection between duality and emergence.<sup>9</sup> A framework for understanding the connection between dualities and emergence was developed in De Haro (2016): it was argued that the two notions (duality as isomorphism, and emergence as novel and robust behaviour relative to a comparison class), while close to each other, also exclude one another. But we must leave the topic of emergence for another occasion.

Feature (B) also prompts the question of *fundamentality*. Coleman himself wrote:

‘I am led to conjecture a form of duality, or nuclear democracy in the sense of Chew, for this two-dimensional theory. A single theory has two equally valid descriptions in terms of Lagrangian field theory: the massive Thirring model and the quantum sine-Gordon equation. The particles which are fundamental in one description are composite in the other: In the Thirring model, the fermion is fundamental and the boson a fermion-antifermion bound state; in the sine-Gordon equation, the boson is fundamental and the fermion a coherent bound state’ (1975: p. 2096).

The issue of fundamentality in boson-fermion duality, and in electric-magnetic duality, has been addressed by Castellani (2017). Her account is, in philosophers’ jargon, deflationary. That is: she argues that all the manifestations of the fields (as bosonic or as fermionic; as electric or magnetic) are ontologically equally fundamental: ‘What the duality specifically implies here, concerns, not mutual composition of the particles, but

---

<sup>8</sup>In T-duality and mirror symmetry, it is not the size of the *couplings* that is inverted by the duality map but, roughly speaking, the sizes of the *spaces*.

<sup>9</sup>On the connection between duality and emergence, see: Dieks et al. (2015), Rickles (2013), Teh (2013).

rather their different modes of appearance when considering the different classical limits of the quantum theory, i.e. the dual perspectives’ (2017: Section 3.3).

Our construal of duality as an isomorphism, in Sections 2 and 3, is in agreement with such a deflationary account. For the content of the theory will be taken to be based on the *common core* of the models: and this common core includes *both* bosons and fermions, on an equal footing. We will discuss some of these issues in Section 3.2.3, but we will not emphasise this point: (for it was worked out in detail for a general duality, and illustrated for gauge-gravity duality, in De Haro (2016a: Section 1), under the heading of ‘physical equivalence’).

## 2 The Schema Introduced: Theories and Models

In this Section and the next, we develop the treatments of theory, model, interpretation, symmetry and duality, given in De Haro (2016: Section 1, 2016a: Section 1) (and foreshadowed in De Haro, Teh, and Butterfield (2016)). This Section deals with theory, model and interpretation; Section 3 will deal with symmetry and duality itself.

We begin with the scientific importance of dualities, and the comparison of duality with gauge (Section 2.1). Then we further specify our notions of theory and model (Section 2.2). Then we discuss: interpretations (Section 2.3), representations and isomorphisms (Section 2.4).

### 2.1 Duality’s scientific importance

Recall from Section 1, our overall proposal. A bare theory can be realized (we will say: modelled) in various ways, like the different representations of an abstract algebra. These models are in general *not* isomorphic, since they differ from one another in their specific structure. But when they are isomorphic, we have a duality.

To develop this proposal, we begin with four clarifying remarks. Each remark leads in to the next. The first three defend our taking duality as a notion that is both logically weak and independent of a theory’s interpretation. The first is, in effect, just the point that ‘duality’ is a term of art; so one can choose how to use it: and our choice of a logically weak definition makes for a *strong* physical notion! But the second and third are substantive—about the scientific importance of dualities. The fourth remark is a contrast with the notion of *gauge*.

(1): *A logically weak but physically strong definition:*— We agree that at first sight, it looks profligate to say that there is duality whenever two models are isomorphic. For it means there are countless dualities. For example: if a group or an algebra, endowed with a set of rules for evaluating quantities, can be a bare theory, any two isomorphic representations will yield a duality, as long as the isomorphism preserves the values of the quantities. Accordingly, the notion of duality is sometimes narrowed by adding physical conditions, not just on ‘bare theory’, but also on the isomorphism between models, e.g. by requiring the isomorphism to link the weak and strong coupling regimes of the two models (see Eq. (1)). But we will maintain in (2) and (3) below that it is best to leave

‘duality’ broadly defined, as we have done: with such extra conditions being articulated in individual cases as the need arises. As we will see in Section 3.2.2, the strengthening will be given by the kind of physical degrees of freedom that one wishes to describe. And so, our notion of duality will be physically strong. In particular, it cannot be argued that two given models which share some structure are dual, unless the common structure is exactly equal to what the models regard as physical. In short: this apparently profligate verdict can be accepted.

(2): *Duality as surprising*:— So far we have spoken mainly of logico-semantic issues, and ignored epistemological ones: we have said what a duality is, but not how surprising and fruitful it can be. Our case-study, in Sections 4 et seq., will of course bring out these issues. It is surprising indeed to learn that a theory we thought of as having as its quantum particles fermions also contains bosons—and even more surprising to learn that conversely the theory can be presented in the first place as having bosons, and then shown to contain the fermions with which we first began. For the moment, we note three clarifying comments—which are suggested by phrases like ‘a theory we thought of’, and ‘the theory can be presented’. Each comment leads in to the next.

(i): We usually discover a duality in the context, not of a bare theory, but of an interpreted theory; for of course we work with interpreted theories.<sup>10</sup>

(ii): Indeed, we usually work with what we have called ‘a model of the theory’, indeed an interpreted model. That is: usually, before the duality is discovered, we have two interpreted models (usually called ‘physical theories’!) which we do not believe to be isomorphic in any relevant sense.

(iii): Usually, we do not initially believe the two models are models of any single relevant theory (even of a bare one: i.e. even if we let ourselves completely suspend our antecedent interpretation of the models). The surprise is to discover that they are such models—indeed are isomorphic ones.

The word ‘relevant’ in (ii) and (iii) signals the fact that of course ‘isomorphism’, ‘model’ and ‘theory’ are very flexible words. For example: almost any two items can be considered isomorphic, i.e. as having a common structure, under a weak enough construal of ‘structure’. Thus physicists might well in some specific context notice that the two models in question are both groups, or both algebras. But they rightly do not announce this as discovering a duality: not even if they also notice that the two groups (or algebras) are isomorphic. They set it aside as irrelevant, since the abstract notion of group or of algebra is so general that having it identified as a bare theory in common between the models is scientifically useless.

On the contrary, what is surprising, and scientifically valuable, is to find very specific, *not* general, structures in common between different models: especially when

(a) the models as presented (so: as interpreted) are very disparate, and-or

---

<sup>10</sup>Agreed, pure mathematicians sometimes work with uninterpreted theories; and duality is a grand theme in mathematics, just as it is in physics. But although comparing duality in mathematics and in physics would be a very worthwhile project, we set it aside. Cf. Corfield (2017).

(b) the common structure is not only detailed (like ‘10-dimensional semisimple Lie group’, as against ‘group’) but amounts to an isomorphism of that detailed structure (like ‘isomorphic as 10-dimensional semisimple Lie groups’).

As noted above, what will give physical theories their specificity, thus making duality a more powerful tool than its logically weak definition might make it seem, is the fact that physical theories, even bare ones, come with sets of maps from groups and algebras to appropriate fields (in the mathematical, not physical, sense!), i.e. maps that assign values to the physical quantities. These maps are defined at the level of the abstract structure, but must also be instantiated in each of the models (according to the relevant sense of instantiation, as either ‘representation’ or ‘realization’: cf. Section 2.2.2). And this set of maps is usually so rich, that it often suffices to reconstruct a model. And so, the fact that duality preserves these maps can be very non-trivial, and surprising, especially when combined with (a)-(b) above.

This discussion of (a)-(b) returns us to (1) above. We doubt that there can be a general characterization of when the models as presented are disparate enough, and-or the discovered isomorphism is detailed enough, for scientific importance. Instead, one can only articulate in any specific case how the disparity and-or the details are enough: e.g. because the isomorphism links the weak and strong coupling regimes of the two models. So it is not worth trying to tighten the *definition* of ‘duality’ with conditions beyond the logically weak ones we advocate. One just needs to use one’s judgment about which cases count as scientifically important enough to analyse.

(3): *Examples*:— The conclusion of (2) is supported by some famous examples of duality in physics. Apart from boson-fermion duality, which we already introduced in Section 1.2, it is worth illustrating this with two other examples.

(A): Gauge-gravity duality. In this case, the models differ in the dimensions they assign to spacetime ( $d$  in the gravity model,  $d-1$  in the gauge model), in their field content and classical equations of motion (Einstein’s equations coupled to matter in the gravity model, the Yang-Mills equations in the gauge model), and in much more. In this case, the common core consists only in a class of asymptotic operators and a conformal class of  $(d-1)$ -dimensional metrics. Of course, it is very surprising to learn that a gauge theory model in  $d-1$  dimensions, and a model of quantum gravity in  $d$  dimensions, despite their very disparate guises, nevertheless have the same common core, and represent the same theory. See De Haro (2016, 2016a) for a discussion in the context of our schema.

(B): Electric-magnetic, or S-duality. This relates two models by mapping the electric charges of one model to the magnetic charges of the other. Furthermore, it does so by mapping a small electric charge to a large magnetic charge, analogously to (1) (since the charges play the role of couplings, in gauge models). Nevertheless, the common structure is the same in the two models, i.e. the quantum theory is invariant under the replacement of one gauge group by its dual.

(4): *A contrast with ‘gauge’*:— This discussion of dualities’ scientific importance brings out a contrast between our treatment of duality, and the notion of gauge. Physicists sometimes make remarks like: ‘two dual theories are like different gauge formulations of a single theory’. We agree that this remark is *analogous* to our view: indeed, in two ways.

(i): A gauge formulation of a theory has specific structure (viz. the gauge variables) going beyond that mandated by the ideas (gauge-invariant ideas!) of the theory; just like for us, a model has specific structure going beyond that mandated by the bare theory.

(ii): The idea of gauge as ‘descriptive redundancy’ means that two gauge formulations of a single theory must ‘say the same thing’; just like we say that in a duality, two models are isomorphic, and so (if interpreted: could) ‘say the same thing’.

But we submit that this is *only* an analogy. There are two differences. First, we want to allow for cases where the two duals are not physically equivalent (as in Kramers-Wannier duality, mentioned above): *pace* the suggestion in (ii). Second (and more importantly), the extra structure in a model is usually *not* gauge, i.e. descriptively redundant: think of how the extra structure in a representation of a group usually carries physical information (e.g. a representation of the Poincare group carrying mass and spin information). Again, as stressed in (2) above: the surprising and scientifically important discovery is that in two models, with apparently very disparate structures, there is in fact an exact correspondence of structures. We shall return to these two differences in Section 3.2.4, comment (3).

## 2.2 Theories and Models

In this subsection, we add details about the ideas of a *bare theory* and its *models*, already introduced in Section 1.

### 2.2.1 Bare theories

Following De Haro (2016, 2016a), we take a *bare theory* to be a triple  $T := \langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$  comprising a structured set of states, a structured set of quantities, and a dynamics: together of course with the rules for evaluating the physical quantities on the states. (We will later discuss symmetries, which we will take as automorphisms  $a : \mathcal{S} \rightarrow \mathcal{S}$  of the set of states; or as the dual maps on the set  $\mathcal{Q}$  of quantities.)

Two immediate points of clarification:—

(1): We stress that, despite the physical connotations of the words ‘states’, ‘quantities’, and ‘dynamics’, a bare theory is *not* interpreted physically. Interpretation comes later (cf. Section 2.3). Thus it will help to think of a bare theory as given by an entirely abstract algebra of quantities, together with a similarly abstract state-space and dynamics. For example, the quantities might be (the self-adjoint elements of) an abstract C\*-algebra  $A$ , the state-space might be determined by  $A$ , viz. as the positive linear functionals on  $A$ , and the dynamics might be an arbitrary automorphism of  $A$ .

(2): Indeed, it will help to think of a bare theory yet more generally. The reason is that most of what we need to say throughout Section 2, about theories and accordingly about their models, is independent of taking a bare theory as a triple—even an uninterpreted one *a la* (1). It depends only on a bare theory having two features:

(a) being uninterpreted, yet ready to be interpreted as a physical theory (hence the idea of the abstract set of states, quantities etc. ‘standing ready’ for interpretation);

(b) being augmentable, i.e. able to be supplemented with extra (again: uninterpreted) structure, in various ways, yielding different realizations, which we will call ‘models’: (to which we will turn directly, in Section 2.2.2).

Clearly, a theory does not need to be a triple  $\langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$  in order to have features (a) and (b). It could, for example, be a theory in logicians' traditional sense of a deductively closed set of formulas in a formal language: such a theory is uninterpreted, and can be augmented in many ways, for example just by adding an extra set of formulas and then closing under deduction. (Here we again connect with the Hilbertian and logical empiricist tradition of formulating physical theories axiomatically, mentioned at the start of Section 1. But we must postpone discussion till Section 2.4.)

### 2.2.2 Models

So we define a *model*  $M$  of a bare theory  $T$  to be a realization of  $T$ . The word 'realization' can be taken in two senses:—

(i): In the broad sense of a mathematical instantiation: i.e. a mathematical entity having the structure of the theory, and usually of course some specific structure of its own (cf. (2)(b) in Section 2.2.1). Thus if  $T$  is a theory in the logicians' sense of a deductively closed set of formulas, a realization is an entity that in some sense 'satisfies' all the formulas of  $T$ , and usually of course some formulas of its own. So any deductively closed superset of  $T$  would count as a realization of  $T$ ; but so of course would any model of  $T$  in logicians' usual sense of 'model'.

(ii): In the mathematical sense of 'representation', as in representation theory. This requires  $T$  to have some structure, so that a representation is a homomorphism (of that structure) from  $T$  to some given, structured object. Since homomorphism need not be isomorphism, the homomorphism's range—the structured object that represents—may have only a 'coarse-grained' version of  $T$ 's structure. But like in (i): since the representing object is 'given', it will also have specific structure of its own. (Recall the two kinds, (A) and (B), of 'extra detail' in Section 1.1.) Again, the obvious examples are when  $T$  is an abstract group or algebra, endowed with a set of rules for evaluating quantities, and a model is a group/algebra representation: for example a subgroup of the general linear group on a complex vector space,  $GL(n, \mathbb{C})$ , endowed with a set of maps to the complex numbers, invariant under similarity transformations, e.g. the group characters.

In our specific example of duality, bosonization (Section 4 et seq.), we will use the second more specific notion, i.e. representation, (ii). But again: much of what we say in this Section needs only (i): the first, more general, sense of realization. And we believe that this notion applies more generally, to the dualities in quantum field theory and string theory: gauge-gravity duality (De Haro (2016a)), mirror symmetry, T-duality, and S-duality; cf. Sections 2.4 and 5.4.

### 2.2.3 Notations for models; model roots and model triples

It is helpful to have a schematic notation for models that exhibits how they augment the structure of a theory with specific structure of their own. This will also introduce some jargon which will be important for us.

One's first thought is to write the model as the ordered pair of the theory and its specific structure,  $\bar{M}$  say:  $M = \langle T, \bar{M} \rangle$ . But we need to reflect the fact that (in almost all cases) the way that a model incorporates the theory's structure is not by simply containing



a ‘copy’ of the theory ‘beside’ its specific structure  $\bar{M}$ : but instead, by using  $\bar{M}$  to build a realization—in one or other of Section 2.2.2’s two senses—of the theory’s structure. Again, the obvious example of group representations illustrates. We should not think of a matrix representation of, say, the symmetric group  $S_N$  as containing a copy of  $S_N$  ‘beside’ its specific structure of a vector space  $V$  and  $N!$  linear maps on  $V$ ; (or maybe less than  $N!$  maps—recall that a representation need only be a homomorphism). Rather,  $V$  and the (upto!)  $N!$  linear maps realize, give a ‘concrete copy’ of,  $S_N$ .

Similarly for examples of dualities in physics, including our example of bosonization. In a typical physics example, the specific structure  $\bar{M}$  consists of a set of fields, endowed with a set of symmetries, a dynamics for the fields, and a set of states of the fields. (So here, fields play the role of quantities in our conception of theories as triples, though of course not all fields are observable: in quantum theories, self-adjoint.) These fields etc. are used to build a ‘concrete copy’ of the bare theory’s structure (or maybe a ‘coarse-grained’ copy). In our own example: a concrete copy of the enveloping affine Lie algebra or Kac-Moody algebra (see Eq. (41) in the Appendix), which is the algebra of which both the bosons and the fermions are representations.

So a better notation that reflects how  $\bar{M}$  is used to build a realization of  $T$  is to write:  $M = \langle T_M, \bar{M} \rangle$ . The occurrence of  $T$  in the notation encodes that the model  $M$  is indeed a model of  $T$ . But the subscript  $M$  on  $T$  reflects that the specific structure  $\bar{M}$  is used to realize  $T$ . In short,  $T_M$  is not ‘given before’  $M$  itself: rather,  $T_M$  realizes  $T$ —in one or other of Section 2.2.2’s two senses—by making use of  $\bar{M}$ .

So one should not think of the model as an ordered pair made from two already-given items  $T_M$  and  $\bar{M}$ . Rather, the decomposition  $M = \langle T_M, \bar{M} \rangle$  is conceptual.<sup>11</sup>

It will be convenient to have a word for  $T_M$ , the ‘part’ of  $M$  that realizes  $T$ . We call it the **model root**. It will also be convenient to have a notation for the model root that does not mention  $T$ , just for simplicity in discussions where it is obvious that one theory  $T$  is in question. We use  $m$ . Thus a theory  $T$  can have various models and model roots,  $M_i$  and  $m_i$ , where  $i$  is in some index set  $I$ . This notation  $m$ , i.e. without mention of  $T$ , will be useful in Section 2.4.

This discussion carries over intact to the more detailed conception of a theory as a triple,  $T = \langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$ . We write a model, not as a pair, but as a quadruple:

$$M = \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M, \bar{M} \rangle =: \langle m, \bar{M} \rangle, \quad (2)$$

where  $m := T_M := \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M \rangle$  will be dubbed the **model triple**, as well as *model root*. As before,  $\bar{M}$  is the specific structure that distinguishes one model, and so model root/triple, from another; and it is  $\bar{M}$  that is used to build the model triple.

This jargon of model root, and model triple, will be important in relation to our proposed schema for duality, that a duality is an isomorphism of models. For this isomorphism is of course isomorphism as regards the bare theory’s structure, rather than any other structure: in particular, this isomorphism sets aside the specific structure (even

---

<sup>11</sup>Besides, in physics examples the actual mathematical structure of a model is often very rich, e.g. in gauge-gravity dualities the metric of a  $(d+1)$ -dimensional spacetime will belong to the specific structure  $\bar{M}$ , but this metric will be a *fibration* over the metric of a  $d$ -dimensional spacetime, which belongs to  $T_M$ : see De Haro (2016: Section 2.1, 2016a: Section 2.2).

though the model’s realizing the bare theory is built from its specific structure). Accordingly, we will often talk of duality as an isomorphism of model roots; and also, when a theory is conceived as a triple, as an isomorphism of model triples.

Finally, it will also be convenient (especially in Section 3.1) to have notation for a model considered in itself, not by comparison with the bare theory of which it is a model. A model is of course *itself* also a triple of a set of states, quantities and a dynamics: i.e. its own states etc., not that of the bare theory. And we will again use the overbar to indicate what is specific to the model. So we write:  $M = \langle \bar{\mathcal{S}}, \bar{\mathcal{Q}}, \bar{\mathcal{D}} \rangle$ .

This prompts the question: what is the relation between the unbarred items  $\mathcal{S}, \mathcal{Q}, \mathcal{D}$  that make up a bare theory  $T$ , and the barred items  $\bar{\mathcal{S}}, \bar{\mathcal{Q}}, \bar{\mathcal{D}}$ ? We will not attempt a general answer to this: we will not need one, and indeed we doubt that there is one. But bearing in mind the prototypical cases of representations of groups or algebras (e.g. representations of the Poincare or conformal algebras which add a specific field content), the tempting, broad answer is that the barred items “are bigger”/“have more structure”. (This is another aspect of the analogy between duality and gauge, despite the contrasts we stressed in remark (4) at the end of Section 2.1: both for two dual models and for a gauge formulation of a theory, there is the intuition of “being bigger”/“having more structure”.)

In the example of boson-fermion duality, as we will see in Sections 4.1 and 4.2, the theory consists of a specific enveloping algebra of the affine algebra (see Eq. (41) in the Appendix), together with a representation space for this algebra and appropriate functionals to the real numbers, which represent the quantities that (in the interpreted theory) become the physical quantities. But the models contain a complex variable  $z$ , which represents two-dimensional spacetime. Using this variable, one can construct, from the operators of the algebra, the bosonic (Klein-Gordon) or fermionic (Dirac fermion) field of the bosonic and the fermionic model, respectively. We shall argue in Section 5.4:(c) that, at the level of the bare theory, the specific dependence of the fields on the spacetime variable  $z$  has no particular significance, and is not needed: so that this suggests the formulation of a more general theory, with less structure.

## 2.3 Interpretations of theories and models

So far, so abstract! Both theories and models have so far been bare, i.e. uninterpreted. We now sketch how we envisage their physical interpretations. We will adapt the elementary ideas of the Frege-Carnap-Lewis framework for semantics (Frege 1892, Carnap 1947, Lewis 1970). This will be easy work: two interpretation maps,  $I_{\text{Int}}$  and  $I_{\text{Ext}}$ , will map from our theories and models, to ‘meanings’ and to ‘the world’ respectively. These maps will later be useful for discussing symmetries (in Section 3.1.2).

We recall that according to the Frege-Carnap-Lewis framework:—

(i) A word gets assigned: first, an intension (Carnap’s word: Frege’s is ‘Sinn’, or in English, ‘sense’: roughly, ‘linguistic meaning’); and second, an extension (Carnap’s word: Frege’s is ‘Bedeutung’, or in English, ‘reference/referent’: roughly, ‘the object or worldly item mentioned’). Roughly speaking: our interpretation maps,  $I_{\text{Int}}$  and  $I_{\text{Ext}}$ , will assign intensions and extensions, respectively.

(ii) A word's intension is assigned to it, once and for all (making the simplifying idealization that all words are univocal, and their linguistic meanings do not change). But a word's extension is assigned to it, relative to a possible world and to other features of the context of use that together determine the reference. For example, the reference of 'the tallest Swede alive today' depends not just on the possible world, but also on the day of use. And in general, the set of features that together determine the reference is large and open-ended (cf. Lewis 1980).

(iii) A singular term (such as a proper name 'Aristotle' or a definite description 'the capital of Denmark') has as its extension, its bearer (in these examples: the man, the city Copenhagen); while a one-place predicate (such as '...walks') has as its extension its set of instances (the set of walkers at the possible world and time in question); while similarly, a two-place predicate (such as '... loves ...') has as its extension *its* set of instances, i.e. the set of ordered pairs where the first loves the second (at the possible world and time in question), e.g. ⟨Romeo, Juliet⟩; and so on for predicates with three or more places.

(iv) Compositional rules describe how to assign intensions and extensions to grammatical phrases and thus to complete sentences, in terms of the intensions and extensions of the component words. For example, 'John walks' is assigned, at a world  $W$  and time  $t$ , the extension True (as against False) iff the reference of 'John' for  $(W, t)$  is in the set of walkers for  $(W, t)$ .

To adapt these ideas, (i) to (iv), to our theories and models, there are two points to bear in mind. They seem to be stumbling blocks, or sources of confusion. But they are easily surmounted and dispelled.

(a): Notice how the ideas above express *contingency* and *transience*: by postulating a background set of possible worlds and times, they secure that a sentence's truth-value can be contingent (vary across the worlds) and transient (vary across the times). In philosophical discussion of physics, contingency and transience are often expressed in a corresponding way: the theory has many solutions, and typically a solution changes with time. This is often expressed using the word *model*: a state at a time, or a temporal sequence of states (a trajectory through the state-space) is called 'a model of the theory'. Thus recall that this was connotation (i) at the start of Section 1.1. The stumbling block is of course that since Section 1, we have reserved 'model' for a very different use: for what many would call 'specific theory', i.e. for a notion that encompasses many solutions throughout time. But we take it that one can surmount this stumbling block, and avoid confusion, just by recognising our stipulated usage.

(b): The ideas, (i) to (iv), were of course developed to give semantics for language about ordinary objects, such as people and towns, like Aristotle, Romeo and Copenhagen. But when one considers one of Section 2.2's theories or models, one is hard pressed to find mention of objects: at least, of objects in the plural. For undoubtedly, 'most of the talk' in the theory or model is about the various states and quantities, about which so many details are given. But these are surely not objects, but rather properties. Namely, properties of the one object—the physical system itself—being theorized about. Agreed: if the system is composite, one naturally regards its component systems as objects in their

own right.<sup>12</sup> But the main point remains: most of the talk in a theory or model is about (numerically quantifiable) properties, and their intricate (quantitative) relationships, not about objects. But this second stumbling block is, like the first, minor. For nothing prohibits interpretations using just one object, viz. the system; or just the system and its component subsystems. (Besides, though we will not go into details: nothing prohibits interpretations treating as objects what are in fact properties; cf. e.g. Lewis (1970a: 429).)

Bearing in mind (a) and (b), we can now spell out how interpretation maps,  $I_{\text{Int}}$  and  $I_{\text{Ext}}$ , assign intensions and extensions, respectively—once we are given a bare theory  $T$ , or a bare model  $M$ . We will discuss the assignment to an element of the set of quantities: i.e. for  $T$ , an element of  $\mathcal{Q}$  in the triple  $T = \langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$ ; and for  $M$ , an element of  $\bar{\mathcal{Q}}$  in the triple  $M = \langle \bar{\mathcal{S}}, \bar{\mathcal{Q}}, \bar{\mathcal{D}} \rangle$ . (Recall from the end of Section 2.2.3 that  $\bar{\mathcal{Q}}$  is all the quantities in the model, and is intuitively ‘larger’ than  $\mathcal{Q}_M$ , which is the realization of  $T$ ’s  $\mathcal{Q}$  in the quadruple  $M = \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M, \bar{M} \rangle$ .) But to stress that this element is uninterpreted/abstract, we will call it  $a$ , not  $Q$ . It will be obvious how the corresponding assignments get made for an element interpreted as a state, or as a dynamics.

For  $I_{\text{Int}}$ , the idea is:  $I_{\text{Int}}$  assigns to an element  $a$  of  $T$  or  $M$ , a physical quantity understood in general terms. For example,  $a$  could be an element of an abstract  $C^*$ -algebra, abelian for classical mechanics and non-abelian for quantum mechanics. And  $I_{\text{Int}}(a)$  could be the quantity, position: which in philosophical terms, is a property with numerically measurable degrees. Or perhaps  $I_{\text{Int}}(a)$  is, more specifically, position in the  $x$  direction, using such-and-such point as spatial origin, and using axes and length-unit thus and so.

Two comments are in order, here. First: We thus envisage a ‘Platonic realm’ of numerically measurable properties as the codomain of the function  $I_{\text{Int}}$ . But our ‘realism’ about quantities is milder than it might appear; and anyway, nothing in what follows will depend on it. In particular, (i): nothing will depend on our taking intensions to encode conventional choices such as spatial origin, axis-direction and length-unit. Besides, (ii): we do not need ‘trans-theoretic identity’ for quantities. That is, a quantity like position need not be ‘the very same quantity’ in different theories: especially if they are radically different, e.g. classical mechanics and quantum mechanics. Despite the single word ‘position’,  $I_{\text{Int}}(a)$  can be different quantities, according as  $a$  is in an abelian, or non-abelian,  $C^*$ -algebra.

Second: Note that we do not need to be precise about the exact domain of definition of  $I_{\text{Int}}$ . For we are only sketching how we envisage interpretation proceeding; and there is of course a great deal of convention about how to formalize both  $T$  and  $M$ , and thus

---

<sup>12</sup>Also agreed: it is common, and mathematically natural, to consider the set  $\mathcal{S}$  of all states, with quantities as extra structure on  $\mathcal{S}$ : e.g. in classical mechanics, as real-valued functions on the phase space  $\mathcal{S}$ , and in quantum mechanics, as linear operators on the Hilbert space  $\mathcal{S}$ . But this does *not* make it compulsory to treat states as the basic objects in a semantics of a physical theory. For it is equally legitimate, though less common in textbooks, to start with the set  $\mathcal{Q}$  of quantities, and take states as extra structure on  $\mathcal{Q}$ . And the legitimacy of both these approaches shows that *au fond*, a state is an assignment of numerical values to all quantities; and *mutatis mutandis* a quantity is an assignment of numerical values to all states. For now, the point is just that if one is asked to classify states and quantities in either of the philosophical categories of ‘object’ and ‘property’, undoubtedly one should classify both states and quantities as properties.

about what is the exact domain of definition of  $I_{\text{Int}}$ .

For  $I_{\text{Ext}}$ , the idea is similar, except that we need to allow for the fact that extensions are assigned relative to a possible world and to all the other features of the context of use that together determine reference. Interpreting any theory or model  $T$  or  $M$  means taking it to be used to describe some empirical phenomena: i.e. taking there to be, in an appropriate possible world, a context of use with a rich enough set of features to determine reference for the elements of  $T$  or  $M$ .<sup>13</sup>

For example, let  $a$  be again an element of an abstract C\*-algebra: say, an abelian one, because  $T$  is a bare theory that is to be interpreted as classical mechanics. Suppose that the possible world  $W$  contains two classical point particles; and we take  $T$  to be used in a context sufficiently rich that  $a$  successfully refers to the position of the more massive particle; or perhaps more specifically, its position in the  $x$  direction, using such-and-such place in  $W$  as spatial origin, and using axes and length-unit thus and so. Then relative to  $W$  and this assumed context of use,  $I_{\text{Ext}}(a)$  is defined to be: the heavier particle's position.

## 2.4 Isomorphisms: defining theories by abstraction from models

So far, we have taken a theory  $T$  as given, and then considered its models  $M$ . In this Subsection, we note that one can argue in the opposite direction: i.e. one can approach defining a theory starting from a class of models. The idea is a widespread one: to define a notion (here, a theory) as those features in common between a suitably varied class of examples (here, models).<sup>14</sup> This is what earlier we called the ‘common core’ of the models (in Sections 1.1-(c), and 1.2.2, 2.1-(3)). We will begin within our previous perspective, i.e. with a theory as given, to introduce notation; then we will sketch how to define a theory, ‘arguing in the opposite direction’.

So recall from Sections 2.2.2 and 2.2.3 that a bare theory  $T$  can have various models and model roots,  $M_i$  and  $m_i$ , where  $i$  is in some index set  $I$ . These ‘realize’ the theory, in the senses (‘instantiate’ or ‘represent’) discussed in Section 2.2.2. But they are in general not isomorphic to each other, nor to the theory: we in general do not have  $m_i \cong m_j \cong T$ . And even if there *is* an isomorphism: it is not an *identity*, because each  $m_i$  is a realization of  $T$  built using the model's specific structure  $\bar{M}_i$ : it is not a ‘pure copy’ of  $T$ .

Supposing there is an isomorphism between a model root  $m_i$  and the theory  $T$ , we denote it by  $f_i : m_i \rightarrow T$ ,  $i \in I$ . Of course,  $m_i$  is in general not a single set, but an  $n$ -tuple or family of sets: often endowing some base set in the family with structure, e.g. the

---

<sup>13</sup>We say ‘appropriate’ so as to signal that of course, for any  $T$  or  $M$ , not every possible world has a context rich enough to determine reference for all  $T$ 's or  $M$ 's elements. Indeed: for many worlds, all their contexts will determine reference for none of  $T$ 's or  $M$ 's elements. For example, take  $T$  or  $M$  to be supersymmetric theories, and a world with no supersymmetric physics. So whatever our precise definition of the domain (i.e. set of arguments) of the map  $I_{\text{Ext}}$ , the map will surely be partial, i.e. undefined on some, maybe the majority, of its arguments. But that is no problem. Formal semantics and philosophy of language in the Frege-Carnap-Lewis framework have long had various proposals for how to treat words and phrases that lack extensions (called ‘bearerless terms’); and these proposals can be adapted to our  $T$  or  $M$ .

<sup>14</sup>This line of thought is not only widespread, but has a long tradition: for thousands of years in philosophical accounts of abstraction; and for a hundred and fifty years in mathematics, with e.g. Frege's proposal to define notions as equivalence classes (e.g. a direction as an equivalence class of parallel lines).

structure of an algebra. In particular, with a theory taken as a triple,  $T = \langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$ , each  $m_i$  is a triple  $m_i = \langle \mathcal{S}_{M_i}, \mathcal{Q}_{M_i}, \mathcal{D}_{M_i} \rangle$ . So the domain and codomain of  $f_i$  are triples, so that  $f_i$  is actually a triple of maps that map respectively: states to states, quantities to quantities, and dynamics to dynamics.<sup>15</sup> We will, however, not need to indicate this in our notation.

Likewise, let the maps  $f_{ij} : m_i \rightarrow m_j$  ( $i, j \in I$ ) denote isomorphisms—when they exist—between two model roots. But again, our notation will not need to indicate the precise domains and codomains of these maps.

Now in the opposite direction. Suppose we are given, not a theory  $T$ , but a set of models indexed by  $I$ : we write this set as  $\{M_i : i \in I\}$ . Similarly, we write the set of model roots indexed by  $I$  as  $\{m_i : i \in I\}$ . Thus we are not assuming that these models, or these model roots, are given as realizing a theory. But we do assume that the model roots (and the models) have some kind of structure, so that it makes sense to say that some pair of model roots is isomorphic with respect to that kind of structure. The idea now is to define a theory as ‘what is in common’ among a suitable set of model roots; where ‘what is in common’ will be expressed as an equivalence class of an appropriate equivalence relation.

The most obvious implementation of this idea is to define equivalence as just the given notion of isomorphism that we assumed applies to the model roots. Thus we might define two model roots  $m_i, m_j$  ( $i, j \in I$ ) to be equivalent,  $m_i \sim m_j$ , just in case there is an isomorphism  $f_{ij} : m_i \rightarrow m_j$  between them. Then the proposal would be: a theory is an equivalence class under this equivalence relation. Recall for example the way in which Frege (1884: Sections 64-67), defined a *direction* as an equivalence class of straight lines under the relation of being parallel.

But if we apply this proposal to model roots, it has the trivial consequence that all model roots of a theory thus defined *are* isomorphic. And as we have discussed above, we must allow a theory to have non-isomorphic model roots.<sup>16</sup>

So the obvious implementation of the initial idea stumbles. If we want to define a theory as ‘what is in common’ among model roots, and express ‘what is in common’ as an equivalence class, then we need a more judicious—no doubt, a weaker—choice of equivalence relation than isomorphism. We will not go into details about how to make this choice. We doubt that there are general rules. But we note that it will be guided by two desiderata:

---

<sup>15</sup>The first two maps will be isomorphisms, the last an equivariance condition: we will say more about this in Section 3.2.1.

<sup>16</sup>Similarly, if we apply this proposal to models: all models of a theory thus defined *are* isomorphic. Of course: we expect that since a model  $M$  has specific structure  $\bar{M}$  going beyond its model root  $m$ , isomorphism for models will in general be stronger—i.e. lead to smaller equivalence classes—than does isomorphism of model roots. But in this Section we will not need to linger on this model vs. model root contrast. For our main concern is defining a theory using isomorphism of model roots. As we will argue below (at the end of this Subsection), there is a natural constraint that model roots must be *sufficiently varied*. For mistaking the presence of accidental similarities between the model roots one happens to have at hand for unnecessary similarities between *all* the model roots of the theory one is trying to define, leads to unnecessarily, or perhaps undesirably, restrictive theories. This is *a fortiori* true of the models  $M_i$  of the theory: since the specific structure  $\bar{M}_i$  is *specific* to  $M_i$ , and so in general not shared with the another model.

(i): our previous understanding (maybe partial understanding) of the theory we are trying to define: the choice is meant to pick out just the structure of the theory we intuitively intend;

(ii): our expectation that the model roots will end up being representations, in the mathematical sense, of the theory thus defined (and so, in general, non-isomorphic to each other).

We will close by mentioning another constraint. It is partly independent of the question of choosing an equivalence relation. For it is about the membership of the class of models or model roots one begins with, rather than the judicious choice one must make of an equivalence relation over it.

Namely: if this whole approach to defining theories is to work—i.e. is to define theories of the kind we intended in our original conception—then the model roots must be sufficiently varied that there are no ‘accidental commonalities’ between them, which would then be inadvertently encoded in the theory defined as an equivalence class—thereby limiting the theory’s possibilities of representation. The point is familiar, e.g. from Frege’s example. If one imagines the lines can be coloured, then Frege’s approach to defining direction needs the lines to vary in colour sufficiently. For if all the lines in a given parallelism equivalence class were the same colour, Frege’s definition of the corresponding direction, viz. as that class, would inadvertently be ambiguous between (i) the direction, as we originally intended it, and (ii) the unintended common colour. Hence our mentioning, in the opening paragraph of this Subsection, ‘a suitably varied class of examples’.

We will not go into details about this constraint. But we stress that obviously, it is substantive. For recall how on the original ‘theory-first’ approach in Section 2.2, we stressed that a model root  $m = T_M$  is not a ‘pure copy’ of the theory  $T$ , but is built from the model  $M$ ’s specific structure  $\bar{M}$ . So on the present ‘reverse’, or ‘models-first’, approach: the danger is that if a class of model roots is not sufficiently varied, they may have considerable specific structure in common (like colour for Frege’s lines)—which will therefore be inadvertently encoded in theories defined as equivalence classes of model roots. We will come back to this point in Section 5.4.<sup>17</sup>

### 3 Duality and Symmetry

In this Section, we first develop our treatment of symmetry (Section 3.1). This, together with our discussion in Section 2, sets us up to present (at last!) our schema for duality as isomorphism of models (Section 3.2).

---

<sup>17</sup>Agreed: one does not always get to ‘choose’ one’s model roots (or models), and so this constraint cannot always be implemented. Thus there is judgment involved in this process of abstraction, viz. of (i) how many, and how varied, the model roots should be, to provide representations of one’s theory, and (ii) how to make the distinction, for a given model, between model root and specific structure (since part of the specific structure of a model could be mistaken for e.g. additional information about the theory). We therefore maintain that this reverse approach, from model roots to theory, is not deductive but inductive—which brings us back to our Hilbertian theme from Section 1. It only stops when one is happy with the theory—based on whatever independent criteria one uses to judge one’s theory and models.

### 3.1 Symmetries of theories and their models

We mentioned symmetry in Section 1.1’s closing analogy between a symmetry mapping a state to a state that is ‘the same’, and a duality mapping a theory to a theory that is ‘the same’. But we now can say more about symmetries, using: (A) our distinction between theories and models (from Section 2.2); and (B) our interpretation maps (from Section 2.3). We take up these topics in Sections 3.1.1 and 3.1.2, respectively.

About (A), our main point will be that the symmetries of a bare (or indeed, an interpreted) theory, and the symmetries of a model of it, are in general overlapping, but distinct, sets. In particular, the symmetries of a theory can be a proper subset of the symmetries of its model: and beware—this inclusion is in the opposite direction from that for the other, more common, use of ‘model’, viz. as a solution, or representative of a solution, of a theory.

About (B), our main point will be that a symmetry must ‘commute’ in an appropriate sense with interpretation. Both these points, and our other ones, will be uncontroversial.

#### 3.1.1 Symmetries of uninterpreted theories and models

Recall the usual conception of symmetry as a map  $a$  on states that preserves the values of a salient, usually large, set of quantities: the state  $s$  and the image-state  $a(s)$  have the same values for quantities. This prompts three immediate comments.

(i): Agreed, it is also usual to think of a symmetry as a map on quantities that preserves values on a salient, usually large, set of states: i.e. for a given state, the value of the argument-quantity equals the value of the image-quantity. But there is no conflict here: the two conceptions are related by duality—in the mathematical, not physical, sense! That is: one map is the (mathematical) dual of the other.

In more detail: given any map  $a : \mathcal{S} \rightarrow \mathcal{S}$ , we can define its *dual map* (not to be confused with a ‘duality map!’) on quantities,  $a^* : \mathcal{Q} \rightarrow \mathcal{Q}$ , by requiring that for any  $s \in \mathcal{S}$  and  $Q \in \mathcal{Q}$ :  $\langle a^*(Q), s \rangle := \langle Q, a(s) \rangle$ . And similarly, starting with quantities: given any map  $a : \mathcal{Q} \rightarrow \mathcal{Q}$ , we say that its dual map on states,  $a^* : \mathcal{S} \rightarrow \mathcal{S}$ , is defined by requiring for all arguments:  $\langle Q, a^*(s) \rangle := \langle a(Q), s \rangle$ .

(ii): Recall the question at the end of Section 1.1: do a state  $s$ , and its image  $a(s)$  under a symmetry, represent the same physical state of affairs (‘possible world’)? Our answer there was, roughly: ‘In general, No: but the ‘Yes’ cases are a natural focus of interest’. We will return to this when discussing interpretation, in Section 3.1.2.<sup>18</sup>

---

<sup>18</sup>For the moment, we just note that it is also common to think that a symmetry as a map on states is ‘active’, i.e. the image-state must be a different physical state of affairs (so the question’s answer is ‘No’), while a symmetry as a map on quantities is ‘passive’, i.e. the image-quantity and the argument-quantity (each with their common value) describe the single given physical state (so that now the question’s answer is ‘Yes’).

We will *deny* this. There is no universal association of symmetry as a map on states as ‘active’, and symmetry as a map on quantities as ‘passive’. The reason lies, essentially, in the distinction between a mathematical state and a physical state: (in the jargon of ‘gauge’, the latter is a gauge-equivalence class of the former). That is: we of course concede that a symmetry as a map on states is ‘active’, in the sense that it changes the states. That is a tautology: (except for the degenerate case where the symmetry is given as being the identity map!). But this concession does not imply that a symmetry as a map on



(iii): We have discussed symmetries as preserving values. But it is common to also require that a symmetry ‘preserves the dynamics’. Taking a symmetry as a map on states, this means, roughly: if a sequence of states is possible under the dynamics, so is the sequence of image-states. That is: if a possible time-evolution is represented by a temporal sequence of states, with the state at time  $t$  being  $s(t)$  (a ‘Schrödinger picture’ of time-evolution), then the sequence  $a(s(t))$  of states is also possible.<sup>19</sup> A corresponding definition can be given for when we take a symmetry as a map on quantities, and use a ‘Heisenberg picture’ of time-evolution as given by a sequence of quantities, i.e. by the sequence of their values on a single ‘fixed’ state.

These comments (i)-(iii) bring out two points. The second is longer: it addresses symmetries of *models*, pointing out that these are in general overlapping but distinct (as abstract groups) from symmetries of theories; and that in the special case where a model triple is isomorphic to the bare theory, the symmetries of the theory can be a proper subset of the symmetries of its model.

First: it is clear that discussing symmetries returns us to Section 2.2’s more detailed conception of a theory, even a bare one, as a triple comprising a set of states, a set of quantities and a dynamics:  $T := \langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$ , together with a set of rules for evaluating quantities.

Second: it is clear that comments (i)-(iii) carry over exactly to *models* in our sense, viz. a realization (‘a more detailed version’) of a theory  $T$ , with specific structure of its own. Recall that Section 2.2.3 introduced two notations for models in this sense. Both notations will be useful in what follows: the first notation immediately, the second notation in the next Subsection. (The second notation was simpler: since a model is itself a triple of its own sets of states, quantities and a dynamics, we wrote:  $M = \langle \bar{\mathcal{S}}, \bar{\mathcal{Q}}, \bar{\mathcal{D}} \rangle$ . It means we can define and discuss ‘symmetry of a model’ just as we did symmetries of theories, e.g. as an automorphism of the state-space  $\bar{\mathcal{S}}$  preserving values of a large salient subset of  $\bar{\mathcal{Q}}$ .)

Section 2.2.3’s first notation distinguishes the realization of the theory’s triple from the specific structure  $\bar{M}$ , and gives a subscript  $M$  to the former to signal that it is built out of the latter:  $M = \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M, \bar{M} \rangle$  (Eq. (2)). We also wrote this as  $M =: \langle m, \bar{M} \rangle$ , where  $m := T_M := \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M \rangle$  is the *model triple*. This notation brings out that for any theory  $T$  and any of its models  $M$ , there is a natural condition for a symmetry  $a$  of  $T$  to be itself realized in  $M$ : for it to have, so to speak, a ‘shadow’ in the model  $M$ , i.e. in

---

states must change the physical state of affairs represented: for the states in question could yet be ‘merely’ mathematical. That is: one still needs a further argument why a difference of these states must imply a difference of physical state (and thus why the question’s answer is ‘No’). This distinction, between a mathematical and a physical symmetry, was labelled, in De Haro et al. (2017: Section 2), with the label (Redundant); and in De Haro (2016a: Section 1.1.2.b), as (Physical)-(Redundant). It also roughly corresponds to the distinction, in Caulton (2015), between an ‘analytic’ and a ‘synthetic’ symmetry.

<sup>19</sup>If the dynamics is deterministic, we can write  $s(t) = D_{t,t_0}(s(t_0))$  where  $D_{t,t_0}$  represents the deterministic dynamics; and then ‘preserving the dynamics’ is equivalent to the commutation i.e. equivariance condition,  $a(s(t)) \equiv a(D_{t,t_0}(s(t_0))) = D_{t,t_0}(a(s(t_0)))$ .

$$\begin{array}{ccc}
\mathcal{S} & \xrightarrow{a} & \mathcal{S} \\
\downarrow \theta & & \downarrow \theta \\
\mathcal{S}_M & \xrightarrow{a_M} & \mathcal{S}_M
\end{array}$$

Figure 1: Commutativity diagram of the symmetry  $a$  with the representation map  $\theta$ .

$$\begin{array}{ccc}
G_1 & \xrightarrow{a_1} & G_1 \\
\downarrow \theta & & \downarrow \theta \\
G_2 & \xrightarrow{h} & G_2
\end{array}$$

Figure 2: Commutativity of group automorphism  $a_1$  with group homomorphism  $\theta$ .

the model triple. This condition is that a diagram should commute, and is not automatic.

To state this condition, however, we need a bit more notation about realization. We will write it as a map  $\theta$ . And we will suppose that in  $T$  we treat symmetries as maps on states, so that  $a : \mathcal{S} \rightarrow \mathcal{S}$  preserves the value of all quantities in a salient subset, say  $\mathcal{Q}^0$ , of the set of all quantities  $\mathcal{Q}$ . Then in the usual case where ‘realization’ means ‘representation’, we can take  $\theta$  as an appropriate structure-preserving map: from  $\mathcal{S}$  in the theory  $T$  itself, to  $\mathcal{S}_M$  in the representing model triple  $m = T_M = \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M \rangle$ . Then the condition in question—that the symmetry  $a$  is itself realized in  $M$ —is that there should be a map  $a_M : \mathcal{S}_M \rightarrow \mathcal{S}_M$ , such that the diagram in Figure 1 commutes.

To convey this idea less abstractly, think of the simplest case. Let the bare theory be just a group  $G_1$ ,<sup>20</sup> with an automorphism  $a_1 : G_1 \rightarrow G_1$ ; and suppose a group  $G_2$  represents  $G_1$  thanks to the existence of a homomorphism  $\theta : G_1 \rightarrow G_2$ . So  $G_2 \cong G_1/\ker \theta$ . For there to be a homomorphism of  $G_2$ ,  $h : G_2 \rightarrow G_2$  (even homomorphism: let alone automorphism), that realizes  $a_1$  (counts as  $a_1$ ’s ‘shadow’ in  $G_2$ ) requires commutation: i.e. for all  $g_1 \in G_1$ ,  $\theta(a_1(g_1)) = h(\theta(g_1))$ . Or as a diagram, see: Figure 2.

In the special case where the model triple is isomorphic to the bare theory, this discussion of course simplifies. Then the map  $\theta$  will be an isomorphism, and the map  $a_M$  (or  $h$  in the toy example of groups) will trivially exist, and both the above diagrams will trivially commute. In this case, a symmetry of the theory has a ‘duplicate’, or ‘replica’, in the symmetries of the model—so that in effect, the symmetries of the theory are a *subset* of the symmetries of the model. We say ‘in effect’ just because of the different domains of definition:  $\mathcal{S}$  vs.  $\mathcal{S}_M$ . Apart from this ‘in effect’, there are two comments to make about this special, simplified, case.

(1): First, we note that on the other use of ‘model’ as an individual solution of

---

<sup>20</sup>Together with a set of maps to the real numbers, to express evaluation of the quantities. But for simplicity we ignore these maps for the moment.

a theory, a model is in general *less* symmetric than its theory—as is often remarked, with the buzz-word ‘symmetry-breaking’. A solution of a dynamics with a spherically symmetric Hamiltonian need not be spherically symmetric; a cubical crystal lattice with one particular placing of its lattice points, and one particular orientation of its edges, can be a solution of a dynamics that is translation-invariant and isotropic; and so on. So: the subset-inclusion in our case above is in the opposite direction from that holding for the other use of ‘model’.

(2): Besides, ‘subset’ here will usually mean *proper subset*. That is: a model’s specific structure  $\bar{M}$ —its ‘content’ that goes beyond its being a model/realization of  $T$ —will mean the model has symmetries additional to those that are ‘duplicates’ of the symmetries of  $T$ . And we expect that if these additional symmetries are well-defined on the model triple, or if they naturally induce a symmetry there, that symmetry is trivial, i.e. just the identity map on the model triple. Our prototypical cases of representations of a group or algebra give examples. Perhaps the simplest is as follows. Let  $T$  be the real numbers  $\mathbb{R}$ ; and let  $M$  be the complex numbers  $\mathbb{C}$  which of course represents  $\mathbb{R}$  as the real axis, i.e. the complex numbers with zero imaginary part,  $\{z \in \mathbb{C} \mid z = x + i0, x \in \mathbb{R}\}$ . So this latter set, the real axis, is like the model triple. Then  $M$  has the symmetry of complex conjugation  $z \mapsto \bar{z}$  which is indeed well-defined on the real axis: but there, it is just the identity map.

And there are examples in interesting cases of dualities. In gauge-gravity dualities, De Haro (2017a) showed that a certain subgroup of the diffeomorphism group of the gravity model of the theory (roughly, the diffeomorphisms which preserve the asymptotic boundary conditions) was ‘invisible’ to the gauge model of the theory, in the sense of not representing any difference on that model: and so these diffeomorphisms are not in the common core between the two models, and they are trivially represented on the theory. The same verdict was made in De Haro (2016a: §2.2.3) for the ‘gauge symmetries’ of the gauge side of the duality. These are not visible on the gravity side: they are symmetries of the formulation of the gauge model of the theory, and are trivially represented on the theory.

To sum up this discussion of the symmetries of a bare theory, and those of its models, and of its model-triples: there are really three points here:

(i): A bare theory  $T$  is realized—typically: represented in the mathematical sense—by one of its model triples,  $m$ . The model  $M$  then consists of  $m$  and some specific structure  $\bar{M}$ ; (cf. Section 2.2.3). And representation requires only a homomorphism, not an isomorphism. Hence our articulating in this Section the condition—in terms of a commuting diagram—for a symmetry of  $T$  to be itself realized in  $m$ .

(ii): And even if in some given case, the representation is an isomorphism, i.e. the representing model triple is isomorphic to the theory  $T$ , so that any symmetry  $a$  of  $T$  will indeed have a ‘duplicate’ or ‘replica’ symmetry in the model triple: still, we must expect that the *model* (as against the model triple) has its own specific structure  $\bar{M}$ . And this specific structure may have symmetries that  $m$ , and the theory  $T$ , ‘knows nothing of’: (cf. comment (2) just above).

(iii): Furthermore, different models, and therefore model triples, of a bare theory are in general *not isomorphic* (as we also discussed in Section 2.4). However, our example of boson-fermion duality, in Section 4, will not illustrate this in detail: i.e. all the model

triples *will* be isomorphic. (But see the comments in Section 5.4.)

### 3.1.2 Interpretations respect symmetries

Recall from Section 2.3 that once we are given a bare theory  $T$ , or a bare model  $M$ , the interpretation maps  $I_{\text{Int}}$  and  $I_{\text{Ext}}$  assign intensions and extensions (respectively) to, for example, an element  $a$  of the set of quantities: i.e. for  $T$ , an element of  $\mathcal{Q}$  in the triple  $T = \langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$ ; and for  $M$ , an element of  $\bar{\mathcal{Q}}$  in the triple  $M = \langle \bar{\mathcal{S}}, \bar{\mathcal{Q}}, \bar{\mathcal{D}} \rangle$ . (And similarly for assigning intensions and extensions to states and dynamics.) Thus  $I_{\text{Int}}$  assigns to an element  $a$  of  $T$  or  $M$ , a physical quantity, e.g. position (or more specifically, position in the  $x$  direction, using such-and-such spatial origin), understood in general terms. Similarly,  $I_{\text{Ext}}$  assigns  $a$  an extension relative to a possible world  $W$  and the other features of the context of use that together determine reference. For example,  $W$  and the context of use may determine that  $a$  is assigned the position of the more massive of two classical point particles that are in  $W$ .

We now propose that these interpretation maps should satisfy appropriate meshing conditions whereby they form commuting diagrams with symmetry maps. The reason is simply that this reflects what one usually means by ‘interpretation’ of the formalism of a physical theory. And this is so whether ‘formalism of a physical theory’ is (in our senses) a bare theory, or a bare model; and whether ‘interpretation’ refers to intension or extension.

Thus to take a very simple example: suppose that a state describes three classical point particles forming a scalene triangle, stationary in an absolute Newtonian space, and (for more simplicity) that the particles do not interact; and suppose the bare theory, or model, at issue has spatial translation and rotation as symmetries. Taking symmetries as maps on states (as usual): these suppositions mean that a spatially translated and-or rotated state has the same values as the given state, for a large and salient set of quantities. So far, these suppositions are, officially, at the level of a bare theory or model—though of course words like ‘point particles’, ‘scalene triangle’ and ‘Newtonian space’ suggest interpretation. And indeed: reading these suppositions, one tends to read them as interpreted, i.e. to unconsciously apply the interpretation maps  $I$ —where  $I$  is short for both  $I_{\text{Int}}$  and  $I_{\text{Ext}}$ . In any case: applying these maps to the two bare states, say  $s$  and  $T(s)$  (‘T’ for ‘transform’ or ‘translate and-or rotate’), one concludes that for the act of interpretation to respect the given symmetries, the interpretations  $I(s)$  and  $I(T(s))$  (where again:  $I$  is short for both  $I_{\text{Int}}$  and  $I_{\text{Ext}}$ ) must have the same values for a large and salient set of quantities—namely, of course, for the interpretations of the bare quantities. This is exactly the condition that symmetry and interpretation form a commuting diagram. But to write such diagrams down, we will need a bit more notation.<sup>21</sup>

The reasons we need more notation are as follows. So far:

- (1): We did not spell out that the domain of each of  $I_{\text{Int}}$  and  $I_{\text{Ext}}$  would include

---

<sup>21</sup>Agreed, to impose this commutation condition for every symmetry and every interpretation is contentious. It seems best justified when we envisage that the bare theory or model describes the whole universe; so that for the example of three point particles, there are no other material bodies in the universe. But in this paper, we do not need to assess exactly when the commutation condition is justified. See De Haro (2016a: §1.3-§1.4), especially the condition called ‘unextendability’.

states and quantities and dynamics; so that each of  $I_{\text{Int}}$  and  $I_{\text{Ext}}$  is really a triple of three maps from states to states, from quantities to quantities, and from dynamics to dynamics. (This is just like an isomorphism  $f_i$  in Section 2.4 being really a triple of such maps; cf. footnote 15.)

(2): We have not introduced notation for the codomains of the interpretation maps: for what one might call the ‘realm of intension’, or ‘meanings’, for  $I_{\text{Int}}$ , and for what one might call the ‘realm of extension’, or the ‘world’, for  $I_{\text{Ext}}$ .

(3): Nor did we introduce a notation for symmetry maps defined on the ‘realm of intension’ or on the ‘realm of extension’. Indeed, we did not yet do this: neither (a) on the states therein (generally understood, in the realm of intension, and specific to a particular physical system, in the realm of extension), analogous to the symmetry maps  $a : \mathcal{S} \rightarrow \mathcal{S}$  on the states of a bare theory; nor (b) on the quantities therein (generally understood, in the realm of intension, and specific to a particular physical system, in the realm of extension), analogous to the symmetry maps  $a^* : \mathcal{Q} \rightarrow \mathcal{Q}$  on the quantities of a bare theory (which are mathematical duals of the maps  $a : \mathcal{S} \rightarrow \mathcal{S}$ ).

To avoid a lot of extra notation, we shall only act on (2) and (3) above. That is:

(2’): We now denote the codomains of the interpretation maps  $I_{\text{Int}}$  and  $I_{\text{Ext}}$  by, respectively: ‘Sinn’ (in honour of Frege’s German word for the realm of intension) and ‘Bed’ (short for ‘Bedeutung’, which was Frege’s word for referent, such as the bearer of a name).

(3’): We confine ourselves to treating symmetries as maps on states (treating them as maps on quantities would be parallel). So we now denote a symmetry on the states in ‘Sinn’ as  $a^{\text{Int}}$ , where the superscript  $\text{Int}$  corresponds to the subscript in  $I_{\text{Int}}$ . And we now denote a symmetry on the states in ‘Bed’ as  $a^{\text{Ext}}$ , where the superscript  $\text{Ext}$  corresponds to the subscript in  $I_{\text{Ext}}$ .

Putting (2’) and (3’) together, we write a state-space in the realm of intension as  $\mathcal{S}_{\text{Sinn}}$ , and a state-space in the realm of extension as  $\mathcal{S}_{\text{Bed}}$ . So we write:  $a^{\text{Int}} : \mathcal{S}_{\text{Sinn}} \rightarrow \mathcal{S}_{\text{Sinn}}$ ; and we write  $a^{\text{Ext}} : \mathcal{S}_{\text{Bed}} \rightarrow \mathcal{S}_{\text{Bed}}$ .

But we shall not act on (1) above: the notation would be cumbersome, and without compensating advantages. In short: acting only on (2) and (3) above—i.e. introducing ‘Sinn’ and ‘Bed’, with symmetry maps  $a^{\text{Int}}$  and  $a^{\text{Ext}}$  respectively—is enough to enable us to draw the required commuting diagrams.

We spell these out: first for the realm of intension, then for the realm of extension. In each case, we first draw the diagram for a bare theory  $T$  taken as a triple,  $T = \langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$ ; and then draw the diagram for a bare model  $M$  taken as a triple  $M = \langle \bar{\mathcal{S}}, \bar{\mathcal{Q}}, \bar{\mathcal{D}} \rangle$ . (Note that these diagrams do not reflect Section 3.1.1’s discussion about the overlap but distinctness (in general) of symmetries of a theory  $T$ , and symmetries of its model  $M$ .)

Thus for the realm of intension, we have: for a bare theory  $T$  with state-space  $\mathcal{S}$ , the diagram in Figure 3. For a bare model  $M$  with state-space  $\bar{\mathcal{S}}$ , we have the diagram in Figure 4.

Strictly speaking, we should in diagrams in Figure 4 and Figure 3 distinguish two interpretation maps—both of which we have in fact written as  $I_{\text{Int}}$ —according as the domain of definition is the state-space of a bare theory or of a bare model. But like in

$$\begin{array}{ccc}
\mathcal{S} & \xrightarrow{a} & \mathcal{S} \\
\downarrow I_{\text{Int}} & & \downarrow I_{\text{Int}} \\
\mathcal{S}_{\text{Sinn}} & \xrightarrow{a^{\text{Int}}} & \mathcal{S}_{\text{Sinn}}
\end{array}$$

Figure 3: Commutativity of the symmetry  $a$  with the interpretation map  $I_{\text{Int}}$  for  $T$ .

$$\begin{array}{ccc}
\bar{\mathcal{S}} & \xrightarrow{a} & \bar{\mathcal{S}} \\
\downarrow I_{\text{Int}} & & \downarrow I_{\text{Int}} \\
\bar{\mathcal{S}}_{\text{Sinn}} & \xrightarrow{a^{\text{Int}}} & \bar{\mathcal{S}}_{\text{Sinn}}
\end{array}$$

Figure 4: Commutativity of the symmetry  $a$  with the interpretation map  $I_{\text{Int}}$  for  $M$ .

comment (1) above: the precision is not worth the burden of extra notation. And similarly of course, for the next two diagrams about the realm of extension.

For the realm of extension, we have, at some given possible world  $W$  and context of use sufficiently rich to determine references (i.e. to avoid the interpretation maps being undefined on the given arguments): for a bare theory  $T$  with state-space  $\mathcal{S}$ , the diagram in Figure 5.

For a bare model  $M$  with state-space  $\bar{\mathcal{S}}$ , we have the diagram in Figure 6.

### 3.2 Duality as isomorphism of models

We turn in this last Subsection to our proposal that a duality is an isomorphism of models of a bare theory. To be precise: it is *an isomorphism of model triples* of a bare theory. Indeed, after all the stage-setting of the previous Subsections (!), the proposal is straightforward. We first give its details, using the notations we have established (Section 3.2.1). In Section 3.2.2, we argue that our notion of duality is logically weak but physically strong. Then in Sections 3.2.3 and 3.2.4, we turn to how duality relates to the topics of Sections 2.3 and 3.1: interpretations and symmetries.

$$\begin{array}{ccc}
\mathcal{S} & \xrightarrow{a} & \mathcal{S} \\
\downarrow I_{\text{Ext}} & & \downarrow I_{\text{Ext}} \\
\mathcal{S}_{\text{Bed}} & \xrightarrow{a^{\text{Ext}}} & \mathcal{S}_{\text{Bed}}
\end{array}$$

Figure 5: Commutativity of the symmetry  $a$  with the interpretation map  $I_{\text{Ext}}$  for  $T$ .

$$\begin{array}{ccc}
\bar{\mathcal{S}} & \xrightarrow{a} & \bar{\mathcal{S}} \\
\downarrow I_{\text{Ext}} & & \downarrow I_{\text{Ext}} \\
\bar{\mathcal{S}}_{\text{Bed}} & \xrightarrow{a^{\text{Ext}}} & \bar{\mathcal{S}}_{\text{Bed}}
\end{array}$$

Figure 6: Commutativity of the symmetry  $a$  with the interpretation map  $I_{\text{Ext}}$  for  $M$ .

### 3.2.1 Duality as isomorphism

Our basic idea is that a duality is an isomorphism between two triples, each comprising a state-space endowed with appropriate structure, a set of quantities endowed with appropriate structure, and a dynamics, consistent with that structure.<sup>22</sup> ‘Isomorphism between triples’ is of course short for a triple of maps: an isomorphism between the two state-spaces, and isomorphism between the sets (almost always: algebras, cf. footnote 22) of quantities, and an equivariance condition on the dynamics.<sup>23</sup> In addition, the isomorphism must commute with the symmetries of the theory, as sketched in Section 3.1.

More important is the question of *which* kinds of triples are related by duality. Recalling our distinction between bare theories and their more specific models, the answer is clear: *a duality relates two model triples of a single bare theory.*

The crucial point here is that the model triple is separated from the model’s own specific structure, and expresses only the model’s realizing (typically: representing in the mathematical sense) the bare theory. Recall the notation from Eq. (2) in Section 2.2.3:  $M = \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M, \bar{M} \rangle =: \langle m, \bar{M} \rangle$ , where  $m := T_M := \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M \rangle$  is the model triple. We emphasised already in Sections 2.2.2 and 2.2.3 that two model triples are in general *not* isomorphic to each other, nor to the bare theory. So the assertion of duality is substantive: it asserts that two model triples are in fact isomorphic.

But this is *not* to say that the two *models*, each ‘considered in their entirety’, are

---

<sup>22</sup>As we mentioned in Section 2.2.1, ‘appropriate structure’ here refers to: (i) the structure of the sets of spaces and quantities, (ii) the rules for evaluating quantities, (iii) the structure which the dynamics satisfies, (iv) the set of symmetries of the theory. We can now be more specific about these, for the examples of quantum theories, which will illustrate our schema: (ia) the set of states will be a separable Hilbert space; (ib) the quantities will be elements (normally the self-adjoint, renormalisable elements) of an algebra; (ii) the rules for evaluating quantities are maps to the appropriate field: for most quantum theories, the inner product on the Hilbert space, and the usual rules for evaluating matrix elements; (iii) dynamical evolution will usually be a (unitary) map, satisfying appropriate commuting diagrams with the other maps in the theory; (iv) the group of symmetries will comprise the automorphisms of the algebra: and possibly additional symmetries, on the states and on the quantities. For classical theories, these comments get modified in familiar ways: e.g. (ia) would say that the set of states is a manifold, with structure appropriate to e.g. Lagrangian or Hamiltonian mechanics.

<sup>23</sup>Our proposal does not depend on the formulation of models as triples. A model root can be presented in many different forms, and the isomorphism should then preserve the corresponding structure. Even for triples, one can envisage isomorphisms which do not respect the triple structure, though they map the model roots isomorphically. Compare Section 3.2.2. But it will suffice for our purposes to restrict to model roots defined as triples, whose structure is preserved by the duality.

isomorphic. They each have their own specific structure, and are (in almost all cases) *not* isomorphic. Recall our other notation from Section 2.2.3 for models ‘considered in their entirety’:  $M = \langle \bar{\mathcal{S}}, \bar{\mathcal{Q}}, \bar{\mathcal{D}} \rangle$ . Indeed, their being non-isomorphic is usually part of what makes the duality surprising and (if Nature is kind to us!) empirically fruitful, i.e. of scientific importance. And ‘the more non-isomorphic’—i.e. the more disparate the two models, considered in their entirety, are—the more surprising, and (one hopes) empirically fruitful, is the duality (cf. (2) and (3) in Section 2.1).<sup>24</sup>

We now introduce some notation for dualities as isomorphisms between model triples. This will require first giving:

- (1) some new notation for the value of a quantity on a state, and
- (2) a more detailed discussion of dynamics (in both the ‘Schrödinger’ and ‘Heisenberg’ pictures).

Both (1) and (2) can be given wholly independently of our distinctions (i) between theories and their models, and (ii) between interpreted and uninterpreted theories. So for the moment, please consider a generic triple of a state-space, a set of quantities, and a dynamics:  $\langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$ .<sup>25</sup>

(1): Suppose we are given a set of states  $\mathcal{S}$ , a set of quantities  $\mathcal{Q}$  and a dynamics  $\mathcal{D}$ :  $\langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$ . We will write  $\langle Q, s \rangle$  for the value of quantity  $Q$  in state  $s$ . This prompts two further general points.

First: it is common to think of a state  $s \in \mathcal{S}$  as a maximal specification of the instantaneous properties of the system in question; and a quantity  $Q \in \mathcal{Q}$  as a numerically measurable property of it. In effect, this makes states and quantities nothing but assignments of values to each other. Second: for classical physics, one naturally takes quantities as real-valued functions on states, so that  $\langle Q, s \rangle := Q(s) \in \mathbb{R}$  is the system’s possessed or intrinsic value of the quantity; and for quantum physics, one naturally takes quantities as linear operators on a Hilbert space of states, so that  $\langle Q, s \rangle := \langle s | \hat{Q} | s \rangle \in \mathbb{R}$  is the system’s Born-rule expectation value of the quantity. But for quantum physics it is often important to consider the non-diagonal matrix elements of a given quantity/operator  $\hat{Q}$ , without requiring this to be adequately encoded in the Born-rule expectation values of various other quantities. So for a quantum theory—as in the bosonization example of Section 4 et seq.!—we should understand a value written schematically as  $\langle Q, s \rangle$  to also represent all the matrix elements  $\langle s_1 | \hat{Q} | s_2 \rangle$ . Thus  $\langle Q, s \rangle$  is a short-hand for an expression

---

<sup>24</sup>We should put this last point more precisely, since our notion of bare theory is logically weak, with even a group or an algebra, together with a set of maps to the real numbers, counting as a legitimate bare theory: (cf. (1) in Section 2.1). And it is in general not surprising, nor likely to be empirically fruitful, to learn that two very disparate models are both groups, or both algebras: (unless the maps to the real numbers are so disparate that the existence of an isomorphism is not easy to guess). Thus the point here, more precisely, is that, for a given degree of detail or logical strength in the bare theory (and the more, the better!): the more disparate its models (considered in their entirety), the more surprising, and one hopes fruitful, is their both realizing the bare theory. That is: the more surprising is the duality.

<sup>25</sup>This simpler idea of a triple was used in our earlier—cruder!—discussion of duality: cf. De Haro et al. (2017: Section 3.2). There, the simplicity engendered no errors, since our general description of duality was but a preamble to a specialist topic: an assessment of gauge symmetries in gauge-gravity duality.



like  $\langle Q; s_1, s_2 \rangle := \langle s_1 | \hat{Q} | s_2 \rangle$ ,<sup>26</sup> i.e.  $Q$  is regarded as a map:  $\mathcal{S} \times \mathcal{S} \rightarrow \mathbb{C}$ .

(2) We turn to the dynamics  $\mathcal{D}$ , i.e. a specification of how the values of quantities change over time. We will keep the discussion very simple. First, we assume the dynamics is deterministic: also in quantum theories, despite the threat of Schrödinger’s cat. Then it can be presented in two ways, for which we adopt the quantum terminology, viz. the ‘Schrödinger’ and ‘Heisenberg’ pictures; (though the ideas occur equally in classical mechanics: for example the remark, frequent in the textbooks, that in Hamiltonian mechanics time-evolution can be regarded as a sequence of canonical transformations, is in effect a statement of the Heisenberg picture). But we shall not need to distinguish otherwise between the different detailed formalisms for dynamics, such as Hamiltonian vs. Lagrangian, and the path-integral. Besides, we will adopt for simplicity the Schrödinger picture.

So we say:  $D_S$  is an action of the real line  $\mathbb{R}$  representing time on  $\mathcal{S}$ . There is an equivalent Heisenberg picture of dynamics with  $D_H$ , an action of  $\mathbb{R}$  representing time on  $\mathcal{Q}$ . The pictures are related by, in an obvious notation:

$$D_S : \mathbb{R} \times \mathcal{S} \ni (t, s) \mapsto D_S(t, s) =: s(t) \in \mathcal{S} \text{ iff } D_H : \mathbb{R} \times \mathcal{Q} \ni (t, Q) \mapsto D_H(t, Q) =: Q(t) \in \mathcal{S} \quad (3)$$

where for all  $s \in \mathcal{S}$  considered as the initial state, and all quantities  $Q \in \mathcal{Q}$ , the values of physical quantities at the later time  $t$  agree in the two pictures:

$$\langle Q, s(t) \rangle = \langle Q(t), s \rangle. \quad (4)$$

With the notations and notions of remarks (1) and (2) in hand, we can now present the notation for dualities as isomorphisms between model triples. Let  $M_1, M_2$  be two models, with model triples  $m_1 = \langle \mathcal{S}_{M_1}, \mathcal{Q}_{M_1}, \mathcal{D}_{M_1} \rangle$  and  $m_2 = \langle \mathcal{S}_{M_2}, \mathcal{Q}_{M_2}, \mathcal{D}_{M_2} \rangle$ . We can suppose that  $M_1, M_2$  are both models of a bare theory  $T$ . Or we can proceed in the ‘opposite direction’ discussed in Section 2.4: that is, we can suppose that  $M_1, M_2$  are given independently of a bare theory  $T$ , but their model triples (model roots in the more general language of Section 2.4) are isomorphic. Either way, the notation for dualities is as follows.

To say that the model triples  $m_1, m_2$  are isomorphic is to say, in short, that: there are isomorphisms between their respective state-spaces and sets of quantities, that (i) make values match, and (ii) are equivariant for the two triples’ dynamics (in the Schrödinger and Heisenberg pictures, respectively). We now spell this out. Though retaining the  $M$ s in the subscripts is cumbersome, we will do so, in order to emphasise our main conceptual point: that duality is a relation between model triples in our sense—it is *not* between theories, or between generic triples  $\langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$  as in remarks (1) and (2).

Thus we say:— A duality between  $m_1 = \langle \mathcal{S}_{M_1}, \mathcal{Q}_{M_1}, \mathcal{D}_{M_1} \rangle$  and  $m_2 = \langle \mathcal{S}_{M_2}, \mathcal{Q}_{M_2}, \mathcal{D}_{M_2} \rangle$  requires:<sup>27</sup>

an isomorphism between the state-spaces (almost always: Hilbert spaces, or for classical theories, manifolds):

$$d_s : \mathcal{S}_{M_1} \rightarrow \mathcal{S}_{M_2} \text{ using } d \text{ for ‘duality’ ;} \quad (5)$$

<sup>26</sup>Therefore duality will imply unitary equivalence.

<sup>27</sup>See footnote 23 and Section 3.2.2 for a brief discussion of more general cases.

$$\begin{array}{ccc}
\mathcal{S}_{M_1} & \xrightarrow{d_s} & \mathcal{S}_{M_2} \\
\downarrow D_{S:1} & & \downarrow D_{S:2} \\
\mathcal{S}_{M_1} & \xrightarrow{d_s} & \mathcal{S}_{M_2}
\end{array}
\qquad
\begin{array}{ccc}
\mathcal{Q}_{M_1} & \xrightarrow{d_q} & \mathcal{Q}_{M_2} \\
\downarrow D_{H:1} & & \downarrow D_{H:2} \\
\mathcal{Q}_{M_1} & \xrightarrow{d_q} & \mathcal{Q}_{M_2}
\end{array}$$

Figure 7: Equivariance of duality and dynamics, for states and quantities.

and an isomorphism between the sets (almost always: algebras) of quantities

$$d_q : \mathcal{Q}_{M_1} \rightarrow \mathcal{Q}_{M_2} \text{ using } d \text{ for 'duality' ;} \quad (6)$$

such that: (i) the values of quantities match:

$$\langle Q_1, s_1 \rangle_1 = \langle d_q(Q_1), d_s(s_1) \rangle_2, \quad \forall Q_1 \in \mathcal{Q}_{M_1}, s_1 \in \mathcal{S}_{M_1}. \quad (7)$$

and: (ii)  $d_s$  is equivariant for the two triples' dynamics,  $D_{S:1}, D_{S:2}$ , in the Schrödinger picture; and  $d_q$  is equivariant for the two triples' dynamics,  $D_{H:1}, D_{H:2}$ , in the Heisenberg picture: see Figure 7.

Eq. (7) appears to favour  $m_1$  over  $m_2$ ; but in fact does not, thanks to the maps  $d$  being bijections.

It is already clear that a duality reduces to a symmetry, in the case where there is just one model, and one model triple, at issue, i.e.  $M_1 = M_2$  and  $m_1 = m_2$ . We shall return to this topic in Section 3.2.4. First, we turn to the questions (i) whether our notion of duality is too weak (Section 3.2.2) and (ii) how it relates to Section 2.3 topic of interpretation (Section 3.2.3).

### 3.2.2 A logically weak but physically strong notion of duality

In Section 2.1, we admitted that our definition of duality is logically weak because there is a duality whenever two models are isomorphic. Thus one might worry that, whenever two given models share some common structure smaller than the model triples, they are dual with respect to the substructures they share. In this Section, we argue that this worry is unfounded, for models that purport to describe physical systems—which is our concern in this paper. Thus the notion of duality is physically stronger than it would at first seem. The point will be to distinguish between a purely formal model vs. a physical (although uninterpreted) model. Our schema is intended for the latter: and it is only the model triples, not their specific structure, that is physically significant.

We will illustrate this in an example. Consider, for simplicity, the following model, based on the  $\mathfrak{su}(2)$  algebra:

$$M_0 = \langle \mathcal{S}_0, \mathcal{Q}_0, \mathcal{D}_0 \rangle := \langle \mathcal{S}_J, U(\mathfrak{su}(2)), C_2(J) \rangle. \quad (8)$$

$\mathcal{S}_J$  is here the Hilbert space of irreducible representations of  $\mathfrak{su}(2)$  with total quantum number  $J$ .  $U(\mathfrak{su}(2))$  is the universal enveloping algebra of  $\mathfrak{su}(2)$ , i.e. roughly, all powers of the algebra elements, quotiented by the algebra relations. The dynamics is the Hamiltonian of the model, which we take to be  $C_2(J)$ , the second Casimir.

Let us compare  $M_0$  with a model  $M$  with which it shares a common structure, and which is based on the  $\mathfrak{su}(2) \otimes \mathfrak{su}(2)$  algebra. The state space is:  $\mathcal{S} = \mathcal{S}_J \otimes \mathcal{S}_K$ , where  $\mathcal{S}_J$  and  $\mathcal{S}_K$  are the state spaces of the first and the second  $\mathfrak{su}(2)$ , respectively.  $J$  is the total quantum number of the first  $\mathfrak{su}(2)$ , and  $K$  is the total quantum number of the second  $\mathfrak{su}(2)$ . We take the dynamics to be given by the sum of the Casimirs of the two  $\mathfrak{su}(2)$ 's,  $\mathcal{D} = C_2(J) + C_2(K)$ . This model is written as:

$$M = \langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle = \left\langle \mathcal{S}_J \otimes \mathcal{S}_K, U(\mathfrak{su}(2) \otimes \mathfrak{su}(2)), C_2(J) + C_2(K) \right\rangle, \quad (9)$$

where  $U$  again indicates the universal enveloping algebra.

$M_0$  and  $M$  are of course both representations of  $M_0$ , i.e.  $M$  is a representation of  $\mathfrak{su}(2)$  with as ‘extra structure’ the second  $\mathfrak{su}(2)$ . In fact,  $M_0$  is isomorphic to  $M$  for the trivial representation with  $K = 0$ , i.e.  $M_0 \cong M|_{K=0}$ .

But  $M$  and  $M_0$  are *not* dual for arbitrary values of  $K$  (as in Eq. (9)). To see this, we rewrite  $M$  in a way which makes explicit the common structure they share, i.e. the first  $\mathfrak{su}(2)$ . So, define:

$$M' := \langle m, \bar{M} \rangle, \quad (10)$$

where  $m$  contains the first  $\mathfrak{su}(2)$ , and the specific structure  $\bar{M}$  contains the second  $\mathfrak{su}(2)$ . Explicitly,  $m = \langle \mathcal{S}_J, U(\mathfrak{su}(2)), C_2(J) \rangle$  and  $\bar{M} = \langle \mathcal{S}_K, U(\mathfrak{su}(2)), C_2(K) \rangle$ . Thus,  $m$  and  $\bar{M}$  are both triples, and they are both isomorphic to  $M_0$ , in particular  $m \cong M_0$ .

We can summarise the above definitions introducing the following short notation:

$$\begin{aligned} M_0 &\cong m = \mathbf{J} \\ M &= \mathbf{J} \otimes \mathbf{K} \\ M' &= \langle \mathbf{J}, \mathbf{K} \rangle, \end{aligned} \quad (11)$$

where  $\mathbf{J}$  is short for the first factor of the tensor product and  $\mathbf{K}$  for the second.

To reconstruct  $M = \mathbf{J} \otimes \mathbf{K}$  from  $M' = \langle \mathbf{J}, \mathbf{K} \rangle$ , one takes the tensor products of the states in  $\mathbf{J}$  and  $\mathbf{K}$ , takes all the products of the quantities, and adds up the dynamics, to reproduce Eq. (9).

If we were to say that  $M = \mathbf{J} \otimes \mathbf{K}$  and its model quadruple counterpart,  $M' = \langle \mathbf{J}, \mathbf{K} \rangle$ , are ‘the same’, then since it is true that  $M' = \langle \mathbf{J}, \mathbf{K} \rangle$ 's model triple and  $M_0 \cong \mathbf{J}$  are isomorphic: it would follow that  $M = \mathbf{J} \otimes \mathbf{K}$  and  $M_0$  would be dual in the relevant sense.

But notice that  $M'$  is *not* the same as  $M$ , nor is it isomorphic to it, because they differ in what they regard as physically significant (cf. De Haro (2017: p.5)). Only the first model triple,  $\mathbf{J}$ , is physically significant in  $M'$ , whereas both model triples are physically significant in  $M$ . Thus there can be no isomorphism between  $M$  and  $M'$  as candidate models of physics, for they differ in their physical content.

In other words, the difference is in a tensor product model presented as such (i.e.  $M = \mathbf{J} \otimes \mathbf{K}$ ) vs. a model quadruple that has as model triple the first factor  $\mathbf{J}$ , and as its specific structure the second factor  $\mathbf{K}$ .

This argument reinforces the point that it is not necessary, nor desirable, to define *bare theories* as equivalence classes of models. This means that the condition to have

a *duality* is only that two model triples be isomorphic: but the model triples need not be isomorphic to the bare theory, only homomorphic to it. And since, as we have just seen, the isomorphism between dual models is essentially unique (i.e. it is not possible to weaken the isomorphism to get dual structures, without changing their physical content), there is no gain in requiring that duality must also involve the theory. If one starts with a bare theory which is weaker than two isomorphic model triples that represent it, it may be possible to strengthen it so as to match the two model triples: but there is no gain in this. So, it is best to keep the notion of *bare theory* physically weak, and the notion of *duality* physically strong.

We end with a contrast of the notions of duality in physics and mathematics. In mathematics, just as in physics, ‘duality’ does not have a fixed meaning; however, all the examples of duality involve just *two* theories. More precisely, the duality operation generates the two-element group  $\mathbb{Z}_2$ . This is not so in the physics literature, where duality can involve more than two models, and the duality group can be rich. For example, the S-duality group of electric-magnetic duality is  $\text{SL}(2, \mathbb{Z})$ , and string theories realize so-called U-duality groups, which involve orthogonal and exceptional groups. Our schema allows for dualities among many models, and so it is closer to the notion in physics. This also strengthens the analogy between duality and symmetry, mentioned in Section 1.1-(2).

### 3.2.3 Duality and interpretation

So far, our discussion of interpretation has concerned a *single* theory or model. Thus recall that Section 2.3 introduced interpretation maps  $I_{\text{Int}}$  and  $I_{\text{Ext}}$  in a rather informal way, as mapping from a bare i.e. uninterpreted theory or a bare model, to the realm of intension (‘Sinn’), or to the realm of extension (‘Bed’), respectively. Then Section 3.1.2 laid out how  $I_{\text{Int}}$  and  $I_{\text{Ext}}$  are to mesh with symmetry maps. This amounted to a commutation condition, i.e.  $I_{\text{Int}}$  and  $I_{\text{Ext}}$  forming a commuting diagram with symmetry maps, which for simplicity we only considered as defined on state-spaces: either on a bare state-space, or on a state-space in the realm of intension (‘Sinn’), or on a state-space in the realm of extension (‘Bed’). (Cf. the diagrams in Figures 3 to 6.) But again, everything in Section 3.1.2 concerned a *single* theory or model.

Since duality is about relations between theories/models, there is, at first sight, little to say about duality and interpretation. That is: interpretation should simply proceed independently on the two sides of the duality—for example, we just require the interpretation-symmetry commuting diagram on both sides of the duality. Indeed: we said already at the start of Section 1.1 that in some cases of duality, the two sides were clearly not—nor intended to be—physically or semantically equivalent: e.g. the high and low temperature regimes in Kramers-Wannier duality. And our definition of duality as formal (viz. an isomorphism of model triples) certainly allows this idea of ‘distinct but isomorphic sectors of reality’—namely as the codomains of the interpretation maps on the two sides of the duality.

This verdict—‘there is little to say’—is true, so far as it goes. And of course, it does not forbid the other sort of case: where the two sides of the duality *are* physically/semantically equivalent, i.e. do describe ‘the same sector of reality’. In our schema, this would be modelled by the interpretation maps on the two sides having the same images/values in

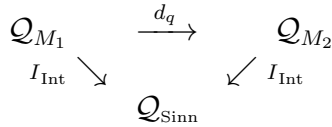


Figure 8: The two sides of the duality describe ‘the same sector of reality’, in the realm of intension.

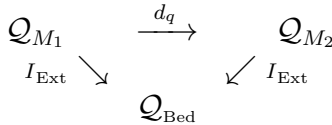


Figure 9: The two sides of the duality describe ‘the same sector of reality’, in the realm of extension.

their codomain—so as to give a *triangular*, rather than *square*, commuting diagram. We shall spell this out as regards the interpretation of (bare) quantities: similar diagrams could of course be drawn for states.

For (bare) quantities being mapped by  $I_{\text{Int}}$  into the realm of intension ‘Sinn’, the two sides of a duality describing ‘the same sector of reality’ amounts to the diagram in Figure 8.

Similarly: for (bare) quantities being mapped by  $I_{\text{Ext}}$  into the realm of extension ‘Bed’—relative to some given possible world  $W$  with a context rich enough to determine references, of course—the two sides of a duality describing ‘the same sector of reality’ amounts to Figure 9.

So far, so straightforward. But the above verdict is a bit quick: there are two further points to make.

(1): *What determines equivalence?*— First, there is the question what determines whether the two sides of a duality are physically/semantically equivalent, i.e. describe the same ‘sector of reality’. De Haro (2016) and Dieks et al. (2015)) have argued that the choice between these two options should depend on whether the models in question are given what was called an ‘internal’ or an ‘external’ interpretation. The idea is that for the ranges of the interpretation maps to be distinct, there must be other facts, external to the triples themselves and our use of them, that determine the distinct ranges. Typically, these facts will be other pieces of physics to which the system described by each model triple is coupled—with different pieces of physics on the two sides of the duality. This coupling ‘breaks the symmetry’ between the two sides, and secures that the two model triples are about distinct, albeit isomorphic, subject matters (‘sectors of reality’). In the proposed jargon: the coupling provides an ‘external interpretation’ of the model triple. On the other hand: sometimes we propose a physical theory as a putative theory of the

whole universe, i.e. as a putative cosmology, so that according to the theory there are no physical facts beyond those about the system (viz. universe) it describes. If in such a case, there is a duality—which in our framework, means there are two isomorphic model triples, each putatively describing the whole universe—then there can be no such coupling to other pieces of physics. (Gauge-gravity duality provides, of course, a putative example of such a duality between theories of the universe.) An interpretation of each triple must therefore be what was labelled an ‘internal interpretation’; and this prompts the conclusion that the two triples describe the very same ‘sector of reality’. That is: the interpretation maps have the same range; and there is a triangular diagram, as in Figures 8 and 9.

(2): *Interpreting the specific structure*:— Second, there is more to say about the interpretation of a model’s specific structure, especially in the latter sort of case, i.e. two sides of a duality describing the same ‘sector of reality’.

Recall that a model  $M$  is more than the model triple  $m$ , by which it realizes a bare theory, and which relates to another model triple in a duality.  $M$  also has a specific structure  $\bar{M}$ : as we stressed at the start of Section 3.2.1, this structure is *not* related by the duality to ‘the other side’. But the specific structure  $\bar{M}$  *does* get interpreted—it supplies arguments for the interpretation maps  $I_{\text{Int}}$  and  $I_{\text{Ext}}$ —just as much as the model triple  $m$  gets interpreted.<sup>28</sup> This was emphasised by the other notation for models introduced at the end of Section 2.2.3: viz. a model is itself, like a bare theory, a triple.  $M$ ’s states and quantities, being specific to  $M$ , are in general ‘bigger’/‘more structured’ than the states and quantities of the bare theory that  $M$  models/realizes. We wrote them with a ‘bar’: thus  $M = \langle \bar{\mathcal{S}}, \bar{\mathcal{Q}}, \bar{\mathcal{D}} \rangle$ . And recall that then Section 2.3 took elements of these ‘bigger’ sets  $\bar{\mathcal{S}}, \bar{\mathcal{Q}}$  as arguments for the interpretation maps  $I_{\text{Int}}$  and  $I_{\text{Ext}}$ .

So the first point to make is: our discussion of duality has so far ignored the specific structures  $\bar{M}_1$  and  $\bar{M}_2$  on the two sides, even though they do get interpreted. This silence is presumably no problem in a case where we agree that the two sides are not physically or semantically equivalent. In such a case, the interpretation of  $\bar{M}_1$  and  $\bar{M}_2$  just means there are physical facts on each of the two sides, additional to the facts that are isomorphic with (a subset of) facts on the other side.<sup>29</sup> Thus for the case of Kramers-Wannier duality, the obvious examples of such non-matching physical facts would be facts about the value of the temperature: high on one side, and low on the other. In short: these additional facts (in the realms of intension and extension, respectively) are: on one side, in the ranges  $I_{\text{Int}}(\bar{M}_1)$  and  $I_{\text{Ext}}(\bar{M}_1)$ ; and on the other side, in the ranges  $I_{\text{Int}}(\bar{M}_2)$  and  $I_{\text{Ext}}(\bar{M}_2)$ .

But what about the other sort of case: where the two sides *do* describe ‘the same

---

<sup>28</sup>At least, this is what we would in general expect. Agreed, one might interpret a model without interpreting *all* of the specific structure  $\bar{M}$ : recall footnote 13 on the need to allow the interpretation maps to be partial, i.e. to deliver no value for certain arguments. For more details, see Section 1.1.2.a of De Haro (2016a).

<sup>29</sup>In Section 3.2.2, we emphasised the fact that only the model triples, and not the specific structure, are physically significant. When we now consider external interpretations that do give a physical meaning to the specific structure, we have to say that these interpretations change the physical content of the model (its physical degrees of freedom). This is correct, because external interpretations do not need to preserve the structure of the model as a quadruple.

sector of reality’? Is it really satisfactory to say that there are physical facts that:

(a) are additional to those facts described by the isomorphic model triples, i.e. those ‘caught’ by the duality/the common bare theory; yet also

(b) fall into two such disparate subsets: one subset expressed by  $\bar{M}_1$  and the other subset expressed by  $\bar{M}_2$ ?

In short: this world-picture, combining (i) a set of facts expressed by the two sides in the same way (though this sameness may be not obvious—the duality can be surprising), and (ii) two other sets of facts expressed in very different ways by the two sides, is surprising: and maybe it is odd, or unsatisfactory ...

We saw examples of this in comment (3) of Section 2.1, for example in gauge-gravity duality. Here, the set of facts that are the common core, *a la* (i), consists only in a class of asymptotic operators and a conformal class of  $(d - 1)$ -dimensional metrics. And the sets of facts *a la* (ii) include, on the bulk side, gravity (such as: Einstein’s equations coupled to matter) in  $d$  dimensions expressed by  $\bar{M}_1$ , and on the boundary side, a conformal field theory (such as: the Yang-Mills equations) in  $d - 1$  dimensions expressed by  $\bar{M}_2$ . See DeHaro (2016: Section 2.1, 2016a: Section 2.2) for a discussion in the context of our schema.

For our example of bosonization, we will see in Section 5.2.4 how external interpretations map to different (sets of) worlds, while the internal interpretation maps to the same (set of) world(s). For the latter interpretation to be possible, we will see that the worlds must contain both bosonic and fermionic facts. So, the internal interpretation does not efface the distinction between bosons and fermions, but distinguishes them and identifies them in the world.

### 3.2.4 Combining duality and symmetries

In this Subsection, we turn to the relations between dualities and symmetries. There are three comments to make. They are not controversial. Indeed, they simply gather some threads from discussions in previous Sections. The first makes the obvious comparison between dualities and symmetries, and notes the conditions for a duality to reduce to being a symmetry. The second is about a duality preserving a symmetry of its model-triples; and so returns us to the contrast between the symmetries of a bare theory, and those of its model-triples. The third returns us to the contrast between duality and gauge, discussed in comment (4) at end of Section 2.1.

(1): *Making the comparison precise:—*

Earlier (at the end of Section 1.1) we announced that we would endorse a basic analogy between duality and symmetry: ‘a duality is like a symmetry, but at the level of theory’, so that while a symmetry carries e.g. a state into a ‘matching’ state, a duality carries a theory into a ‘matching’ theory.

Indeed, we endorse this analogy—allowing of course for the shift of words from ‘theory’ to ‘model-triple’. This endorsement is clear from:

(a): our discussion of symmetries of theories, taken as triples, and symmetries of their models, and their model-triples (Section 3.1.1); and

(b): our definition of duality as an isomorphism of model-triples that (i) makes the

values of quantities match, and (ii) is equivariant for the two triples' dynamics (cf. Eq. 7 and Figure 7 at the end of Section 3.2.1).

In particular (as mentioned at the end of Section 3.2.1): a duality reduces to a symmetry, in the case where there is just one model, and one model triple, at issue, i.e.  $M_1 = M_2$  and  $m_1 = m_2$ . Spelling this out will use the notion of a dual map (in the pure mathematical sense!), introduced in (i) at the start of Section 3.1.1. Recall that this notion is defined by the pairing whereby states  $s \in \mathcal{S}$  and quantities  $Q \in \mathcal{Q}$  assign each other a value  $\langle Q, s \rangle$ . Namely: given any map  $a : \mathcal{S} \rightarrow \mathcal{S}$ , we said that its dual map on quantities,  $a^* : \mathcal{Q} \rightarrow \mathcal{Q}$  is defined by requiring that for any  $s \in \mathcal{S}$  and  $Q \in \mathcal{Q}$ :  $\langle a^*(Q), s \rangle := \langle Q, a(s) \rangle$ . And similarly, starting with quantities: given any map  $a : \mathcal{Q} \rightarrow \mathcal{Q}$ , we said that its dual map on states,  $a^* : \mathcal{S} \rightarrow \mathcal{S}$  is defined by requiring for all arguments:  $\langle Q, a^*(s) \rangle := \langle a(Q), s \rangle$ .

Thus suppose there is just one model triple at issue. Then  $d_s$  is an automorphism of  $\mathcal{S}_{M_1} \equiv \mathcal{S}_{M_2}$ , i.e. of the state-space in the one model triple; and similarly, for  $d_q$  on  $\mathcal{Q}_{M_1} \equiv \mathcal{Q}_{M_2}$ . So duality's condition (i), that the values of quantities match (Eq. (7)), becomes the condition

$$\langle Q_1, s_1 \rangle_1 = \langle d_q(Q_1), d_s(s_1) \rangle_1, \quad \forall Q_1 \in \mathcal{Q}_{M_1}, s_1 \in \mathcal{S}_{M_1}. \quad (12)$$

But  $d_q$  induces a dual map  $d_q^*$  on states, such that:  $\langle d_q(Q_1), d_s(s_1) \rangle_1 = \langle Q_1, d_q^*(d_s(s_1)) \rangle_1$ . So we conclude that  $(d_q^* \circ d_s) : \mathcal{S}_{M_1} \rightarrow \mathcal{S}_{M_1}$  is a symmetry (written, as usual for us, as a map on states rather than quantities). For we have:

$$\langle Q_1, s_1 \rangle_1 = \langle d_q(Q_1), d_s(s_1) \rangle_1 = \langle Q_1, d_q^*(d_s(s_1)) \rangle_1. \quad (13)$$

Finally, the same verdict—that for a single theory, duality reduces to symmetry—applies to dynamics, i.e. to dynamical symmetries. That is: if a duality concerns just one model triple, then Section 3.2.1's condition (ii) for duality—that the duality map is equivariant for the two triples' dynamics (i.e.  $d_s$  is equivariant for Schrödinger dynamics, and  $d_q$  is equivariant for Heisenberg dynamics)—reduces to the condition that the duality is also a dynamical symmetry: for example, that  $d_s$  is a dynamical symmetry represented as a map on states.

(2): *On duality preserving a symmetry*:—

It is straightforward to confirm that on Section 3.2.1's definition of duality, a duality preserves any symmetry of its model triples. There are two points here. First: there is a commuting square diagram of isomorphisms. Second: there is the issue of the values of a quantity being equal on a given state, and on its transform under a symmetry. The first point will lead in to the second.

First: The duality maps  $d_s, d_q$  are not only bijections, but isomorphisms:  $d_s : \mathcal{S}_{M_1} \rightarrow \mathcal{S}_{M_2}$ , and  $d_q : \mathcal{Q}_{M_1} \rightarrow \mathcal{Q}_{M_2}$ . And although we did not have to spell out the exact structures of  $\mathcal{S}_{M_i}, \mathcal{Q}_{M_i}$  that these isomorphisms are to preserve (but cf. footnote 22), it is obvious from the fact that 'is isomorphic to' is both a symmetric and a transitive relation, that the following diagram, with  $a$  understood to be any automorphism of  $\mathcal{S}_{M_1}$ , commutes (cf. Figure 10).



$$\begin{array}{ccc}
\mathcal{S}_{M_1} & \xrightarrow{a} & \mathcal{S}_{M_1} \\
\downarrow d_s & & \downarrow d_s \\
\mathcal{S}_{M_2} & \longrightarrow & \mathcal{S}_{M_2}
\end{array}$$

Figure 10: Commutativity of duality and symmetry for states.

$$\begin{array}{ccc}
\mathcal{Q}_{M_1} & \xrightarrow{a} & \mathcal{Q}_{M_1} \\
\downarrow d_q & & \downarrow d_q \\
\mathcal{Q}_{M_2} & \longrightarrow & \mathcal{Q}_{M_2}
\end{array}$$

Figure 11: Commutativity of duality and symmetry for quantities.

And of course, this diagram of isomorphisms is just what we mean by saying a duality  $d$  preserves an automorphism of the state-space  $\mathcal{S}_{M_1}$  in its domain model triple, and preserves  $\mathcal{S}_{M_1}$ 's structure. Namely,  $d$  carries the automorphism—a map  $a$  on  $\mathcal{S}_{M_1}$ —to a corresponding automorphism of states in the codomain (indeed: range) model triple. The diagram defines this corresponding automorphism, i.e. the map forming the fourth side of the square:  $d_s \circ a \circ (d_s)^{-1} : \mathcal{S}_{M_2} \rightarrow \mathcal{S}_{M_2}$ .

There is obviously a corresponding point about quantities, as against states. Since  $d_q$  is required to be an isomorphism of quantities, the following diagram, with  $a$  now understood to be any automorphism of  $\mathcal{Q}_{M_1}$ , must commute, cf. Figure 11.

And again, this diagram is just what we mean by saying a duality  $d$  preserves an automorphism of the quantities in its domain model triple, and preserves  $\mathcal{Q}_{M_1}$ 's structure. Namely,  $d$  carries the automorphism—a map  $a$  on  $\mathcal{Q}_{M_1}$ —to a corresponding automorphism of quantities in the codomain (indeed: range) model triple. The diagram defines this corresponding automorphism:  $d_q \circ a \circ (d_q)^{-1} : \mathcal{Q}_{M_2} \rightarrow \mathcal{Q}_{M_2}$ .

Second: But in physics, the notion of symmetry involves more than the notions of automorphism of the state-space, and of the set (usually algebra) of quantities. It involves the pairing whereby states  $s$  and quantities  $Q$  assign each other a value:  $\langle Q, s \rangle$ . For these values (for a large and salient set of quantities, though usually not *all* quantities) must be preserved under the symmetry.

But satisfying this is automatic, for a duality as defined at the end of Section 3.2.1. That is: For a duality to respect this aspect of symmetry was already built in to our definition of duality: namely in condition (i), that the values are equal between states and quantities that correspond by the duality. Recall Eq. (7)), which we here repeat:

$$\langle Q_1, s_1 \rangle_1 = \langle d_q(Q_1), d_s(s_1) \rangle_2, \quad \forall Q_1 \in \mathcal{Q}_{M_1}, s_1 \in \mathcal{S}_{M_1}. \quad (14)$$

Finally (and just like at the end of (1) above): the same verdict—that a duality

$$\begin{array}{ccc}
\mathcal{S}_{M_1} & \xrightarrow{a} & \mathcal{S}_{M_1} \\
\downarrow D_{t,t_0} & & \downarrow D_{t,t_0} \\
\mathcal{S}_{M_1} & \xrightarrow{a} & \mathcal{S}_{M_1}
\end{array}$$

Figure 12: Commutativity of symmetry and dynamics.

$$\begin{array}{ccc}
\mathcal{S}_{M_2} & \longrightarrow & \mathcal{S}_{M_2} \\
\downarrow d_s^{-1} & & \downarrow d_s^{-1} \\
\mathcal{S}_{M_1} & \xrightarrow{a} & \mathcal{S}_{M_1} \\
\downarrow D_{t,t_0} & & \downarrow D_{t,t_0} \\
\mathcal{S}_{M_1} & \xrightarrow{a} & \mathcal{S}_{M_1} \\
\downarrow d_s & & \downarrow d_s \\
\mathcal{S}_{M_2} & \longrightarrow & \mathcal{S}_{M_2}
\end{array}$$

Figure 13: Commutativity of duality, symmetry, and dynamics.

preserves any symmetry of its model triples—applies to dynamics, i.e. to dynamical symmetries. Recall from footnote 19 (in Section 3.1.1) that a dynamical symmetry is a commutation i.e. equivariance condition. So for the Schrödinger picture of dynamics, the diagram for the ‘first’ side of a duality, i.e.  $m_1 = \langle \mathcal{S}_{M_1}, \mathcal{Q}_{M_1}, \mathcal{D}_{M_1} \rangle$ , is, with  $a$  the dynamical symmetry, as in Figure 12.

So we now compose this diagram with Figure 10, which represents that a duality preserves a symmetry. But since in Figure 12, the ‘first’ side, ‘1’, of the duality occurs twice, on both top and bottom rows, we now need to compose Figure 12 with Figure 10 twice: both on its bottom row; and also on its top row (with the duality arrow in Figure 10 reversed). The resulting diagram (Figure 13) shows that the duality isomorphism on state-spaces  $d_s$  carries the dynamical symmetry  $a$  on the ‘1’ side of the duality, to a dynamical symmetry on the ‘2’ side: namely, the symmetry  $d_s \circ a \circ d_s^{-1}$  (cf. either the top or bottom square). The Schrödinger picture dynamics on  $\mathcal{S}_{M_2}$  is (reading down the columns in the Figure):  $d_s \circ D_{t,t_0} \circ d_s^{-1}$ .

So much by way of showing that a duality always preserves a symmetry of its model triples. In conclusion, we should emphasise again the three summarising comments, (i) to (iii), at the end of Section 3.1.1 about the contrasts between the symmetries of a bare theory, and those of its models, and of its model-triples.

(3): *The contrast between duality and gauge:—*

Finally, we should briefly return to our comment (4) at the end of Section 2.1. We said there that, although there is some truth in the common remark that two dual theories

are like gauge formulations of a single theory, there are two important differences. Sometimes the two duals are agreed to *not* be physically equivalent (as in Kramers-Wannier duality). And anyway, the specific structure in a model is usually not gauge, in the sense of descriptively redundant.

Our discussion since Section 2.1 reinforces this comment. For we have seen in more detail the idea of specific structure in a model—starting with our notation,  $\bar{M}$ , from Section 2.2.3. And by relating duality as isomorphism of model triples to our interpretation maps, we saw that duality allows, but does not entail, physical equivalence (cf. Section 3.2.3).

Besides, we have also seen a more specific contrast between duality and gauge, that was not foreshadowed in comment (4) at the end of Section 2.1. Namely, we noted in (2) at the end of Section 3.1.1 that if a symmetry of a model’s specific structure—a symmetry of  $\bar{M}$ —is well-defined on the model triple, we expect it to be trivial, i.e. the identity map, there. This point implies that we would in general expect gauge (i.e. descriptively redundant) structure to *not* be carried across intact by a duality. And indeed: this has been illustrated in detail in gauge-gravity dualities. De Haro (2017a) has shown that a certain subgroup of the diffeomorphism group of the gravity model of the theory (roughly, the diffeomorphisms which preserve the asymptotic boundary conditions) is ‘invisible’ to the gauge model of the theory, in the sense of not representing any difference on that model, and so being trivially represented on the theory (the common core). Similarly for the ‘gauge symmetries’ of the gauge side of the duality. These are not visible on the gravity side: they are symmetries of the specific structure of the gauge model, and are trivially represented on the theory (the common core); De Haro (2016a: Section 2.2.3), De Haro et al. (2017: Section 5.2).

## 4 The Basic Boson-Fermion Duality

Boson-fermion duality will be our main example of the schema developed in Sections 2 and 3: for, as we discussed in Section 1.2, it lies in the middle of the spectrum between (i) mathematical precision and established physics, (ii) scientific importance.

Boson-fermion duality is a vast field, still an active area of research today, especially in its three- and four-dimensional versions.<sup>30</sup> But even just in two dimensions, there is a large number of examples, which we will discuss in the next Section. In this Section, we will start with the basic case: that is, the equivalence between the free, massless scalar field, and the free, massless Dirac fermion, in two Euclidean dimensions.<sup>31</sup> This case already exhibits all of the interesting and non-trivial features of the more involved dualities, and so we will analyse it in some detail.

Our exposition will be necessarily brief, and will focus on those aspects that best illustrate the schema. Thus we will downplay many other important physical and mathematical aspects of this duality, such as: the existence of classical and quantum soliton solutions in these theories, the integrability of the equations, the notions of Noether and

---

<sup>30</sup>Cf. e.g. Gogolin (2004), Kopietz (2008), Kouneiher (2018).

<sup>31</sup>The results in Minkowski signature are readily obtained by a Wick rotation, as we discuss in Section 4.1.

topological charges, the connection with QCD and monopoles. Neglecting these important topics is the price we pay for focussing on illustrating a conceptual schema.

This Section is introductory. We here collect the technical results (especially, about the symmetries, their associated Noether currents, and the algebras the currents generate) that will allow us to illustrate our schema. Section 4.1 introduces the free, massless boson. Section 4.2 introduces the free, massless Dirac fermion. (In Section 5, we will present the basic boson-fermion duality, and show how it exemplifies our duality schema from Sections 2 and 3.)

Our exposition will mainly follow Ginsparg (1990) and Frishman and Sonnenschein (2010).

## 4.1 The free, massless boson

In this subsection, we study the free, massless boson. We analyse its symmetries, write down the associated Noether currents, and give relevant details about its quantization, in particular we give the algebra that the Noether currents satisfy. This algebra will be the starting point of the comparison, in Section 5, with the fermionic model (where we will also justify how the bosonic and fermionic models are ‘models’, in our sense of Sections 2.2.2 and 2.2.3).

In two-dimensional quantum field theory, it is very useful to work in complex coordinates. We will use the coordinates  $z = x^0 + ix^1$ ,  $\bar{z} = x^0 - ix^1$  parametrising  $\mathbb{C} \cong \mathbb{R}^2$ , the complex and Euclidean planes, respectively.

We first discuss the classical field and its symmetries, in two points, (i)-(ii), below. A free, massless scalar field  $\Phi$  satisfies the massless Klein-Gordon equation, which in the complex coordinates chosen in the previous paragraph takes the form:  $\partial\bar{\partial}\Phi = 0$  (here, and elsewhere, we use the short-hand  $\partial = \partial/\partial z$ ,  $\bar{\partial} = \partial/\partial\bar{z}$ ). The general solution to this equation allows for much more general functions than it does in higher dimensions, and this is the root of the richness of two-dimensional quantum field theory. The general solution is the sum of a holomorphic and an anti-holomorphic function:<sup>32</sup>

$$\Phi(z, \bar{z}) = \phi(z) + \bar{\phi}(\bar{z}) . \tag{15}$$

The holomorphic and anti-holomorphic functions  $\phi(z)$  and  $\bar{\phi}(\bar{z})$  are often called the left-, respectively right-moving parts of  $\Phi$ . This is because a holomorphic function depends only on  $z = x^0 + ix^1$ : and after Wick rotation  $x^1 \rightarrow -ix$  (with  $x^0 = t$ ), the holomorphic part of  $\Phi$  induces a function of  $t+x$ . For any fixed value of  $t+x$ , this indeed gives motion to the left (the speed is always negative at fixed  $t+x$ ); whereas  $\bar{z} \rightarrow t-x$ , and thus  $\bar{\phi}(\bar{z})$  induces a function of  $t-x$ , which is right-moving (the speed is always positive at any fixed  $t+x$ ).

The equations of motion (equivalently, the classical action) have two sets of symmetries which will be the starting point of our set-up (once they are generalised to symmetries of

---

<sup>32</sup>Classically, we may indeed require the solutions to be holomorphic and anti-holomorphic functions. Quantum mechanically, there are singularities which are both inevitable and the source of interesting physics, as we will see. Thus we will allow  $\phi(z)$  and  $\bar{\phi}(\bar{z})$  to have isolated singularities, hence we will allow them to be meromorphic and anti-meromorphic functions (operators, in the quantum version of the model), respectively.

the quantum version of the model):

(i) **Conformal transformations.** In two dimensions, the action of a massless scalar field is invariant under a large group of coordinate transformations, namely conformal transformations: these are scale transformations with a variable scale factor, such that angles are preserved. In complex variables  $z, \bar{z}$ , the conformal transformations are parametrised by arbitrary holomorphic and anti-holomorphic functions:

$$z \rightarrow z' = f(z), \quad \bar{z} \rightarrow \bar{z}' = \bar{f}(\bar{z}). \quad (16)$$

This is the two-dimensional version of the conformal transformations. Unlike the conformal group in higher dimensions, which has a finite number of generators (the generators of the Poincaré group plus additional generators of conformal transformations), the above transformations form a group whose corresponding algebra has an infinite number of generators. After taking into account quantum effects, this will be the celebrated Virasoro algebra.

It is easy to see that the above transformations contain, in terms of the Euclidean coordinates  $x^\mu = (x^0, x^1)$ , in particular: (a) constant translations,  $x^\mu \rightarrow x^\mu + a^\mu$ , (b) SO(2) rotations (in Minkowski signature, these are SO(1, 1) Lorentz transformations), which in complex coordinates induce a U(1) action, (c) dilations (scale transformations)  $z' = \lambda z$ ,  $\bar{z}' = \bar{\lambda} \bar{z}$  ( $\lambda \in \mathbb{R}$ ).

(ii) **Affine current algebra transformations.** These are translations of the field by holomorphic or anti-holomorphic functions,

$$\Phi(z, \bar{z}) \rightarrow \Phi(z, \bar{z}) + \varphi(z), \quad \bar{\Phi}(z, \bar{z}) \rightarrow \bar{\Phi}(z, \bar{z}) + \bar{\varphi}(\bar{z}). \quad (17)$$

Again, these transformations generalise the invariance of the action under constant shifts  $\Phi \rightarrow \Phi + \varphi_0$ , and are specific to two dimensions.

The conserved currents associated with these two sets of symmetries are obtained through the Noether procedure. The currents for the affine current algebra transformations (17) are, up to an overall constant:

$$J(z) := \partial\phi(z), \quad \bar{J}(\bar{z}) := \bar{\partial}\bar{\phi}(\bar{z}), \quad (18)$$

and they are anti-holomorphically, respectively holomorphically conserved in virtue of their (anti-) holomorphicity. These currents are called ‘affine currents’ because, in the quantum version of the model, they generate an affine Lie algebra or Kac-Moody algebra.

The conserved currents associated with the conformal transformations (16) are the (holomorphic and anti-holomorphic) components of the stress-energy tensor:

$$T(z) = -\frac{1}{2} \partial\phi \partial\phi = -\frac{1}{2} J^2(z), \quad \bar{T}(\bar{z}) = -\frac{1}{2} \bar{\partial}\phi \bar{\partial}\phi = -\frac{1}{2} \bar{J}^2(\bar{z}). \quad (19)$$

The fact that the components of the stress-energy tensor can be written as squares of the affine algebra currents will be important upon quantisation, since it will link together the

Virasoro and the Kac-Moody algebras. In the quantum case, the right-hand side of (19) will contain the normally ordered product, and the relation is then called the *Sugawara construction*.

Our next task is to quantise the model. There are two well-known ways to quantise this model (which we briefly discuss in what follows), but we will settle for a third one, which is the more ‘modern approach’, called ‘radial quantisation’. It is well suited to our perspective because it exploits the conformal symmetry group, and delivers the conformal algebra in the way we want it for Section 5.

One can adopt conventional canonical quantization, where one writes down canonical commutation relations for the fields and their canonically conjugate momenta. One can realize these commutation relations by choosing a Fock-space representation of the fields, and writing down the algebra of the creation and annihilation operators, which (in the present case of two dimensions) depend on one-dimensional momenta. One can then write down a Hamiltonian in terms of the creation and annihilation operators, and set up the physical states in the Fock space. See e.g. Frishman and Sonnenschein (2010: Section 1.6).

The model can also be readily quantised using path integral quantisation. See e.g. Frishman and Sonnenschein (2010: Section 1.9). Here, we will adopt the third method, *radial quantisation*, as follows (ibid. Section 1.7).

This approach is based on the complex coordinates  $z, \bar{z}$ . First, we will use the conformal symmetry group to choose more convenient coordinates: mapping  $\xi := x^0 + ix^1 \mapsto z := e^\xi = e^{x^0 + ix^1}$  (for more details, see the Appendix). ‘Radial’ then refers to the fact that, after this conformal transformation, the equal-time slices  $x^0 = \text{constant}$ , used for canonical quantisation, become circles of constant radius. The line integrals over space which appear in physical quantities, such as charges, then turn into contour integrals on the complex plane. We will not spell out the details of this procedure, of which there are many good reviews (see e.g. Section 2.2 of Ginsparg (1990) or Lüst et al. (1989: Section 4.1)), but the key technique we will mention is the use of radial ordering to define the order of operators which are integrated over a contour in the complex plane.

Remember that our aim for the quantum model of the boson is to obtain its algebra of operators. To this end, we will use the short-distance behaviour of a distinguished set of fields (so-called ‘primary fields’: cf. next paragraph). The algebra is indeed encoded in the short-distance behaviour of the products of the primary fields among each other and with the stress-energy tensor. This short-distance behaviour of products is called the ‘operator product expansion’. To this we now turn.

So let us consider a *primary field*  $\Psi(z, \bar{z})$ . Primary fields are defined by their transformation properties under conformal transformations (see Eq. (40) in the Appendix).<sup>33</sup> It can be readily shown that the short-distance behaviour of the product of the stress-energy

---

<sup>33</sup>They are called ‘primary’ because all other fields, which are called ‘descendants’, can be obtained from them, through successive application of derivatives. See De Haro et al. (2016: Section 3).

tensor with any primary field is:

$$\begin{aligned}
T(z) \Psi(w, \bar{w}) &= \frac{h}{(z-w)^2} \Psi(w, \bar{w}) + \frac{1}{z-w} \partial_w \Psi(w, \bar{w}) + \text{finite terms}, \quad (z \rightarrow w) \\
\bar{T}(\bar{z}) \Psi(w, \bar{w}) &= \frac{\bar{h}}{(\bar{z}-\bar{w})^2} \Psi(w, \bar{w}) + \frac{1}{\bar{z}-\bar{w}} \partial_{\bar{w}} \Psi(w, \bar{w}) + \text{finite terms}, \quad (\bar{z} \rightarrow \bar{w})
\end{aligned}
\tag{20}$$

where  $h, \bar{h}$  are the conformal weights of the primary field  $\Psi$ , i.e. the powers with which a field scales under meromorphic, respectively anti-meromorphic transformations (16), as in Eq. (40) of the Appendix.<sup>34</sup> The conformal weights are usually denoted as  $(h, \bar{h})$ , and they encode how a field transforms under dilatations.<sup>35</sup> For example, taking  $\Psi = \partial\phi(w)$ , i.e. the (derivative of the) massless left-moving part of  $\phi$ , we have  $h = 1$ , since this field is a vector, hence the conformal weight is  $(1, 0)$ . Obviously, this is our primary field of interest in this model. As we will see below, after Eq. (22), the stress-energy tensor is *not* a primary field. On the other hand (cf. Section 4.2), a Weyl-Majorana fermion  $\Psi = \chi(w)$  has conformal weight  $(\frac{1}{2}, 0)$ . The anti-holomorphic bosonic field  $\bar{\phi}$  has conformal weight  $(0, 1)$ ; and the anti-holomorphic Weyl-Majorana fermion has  $(0, \frac{1}{2})$ .

The expansion (20) is called an *operator product expansion*. Operator product expansions are important because they tell us almost everything we need to know about a field: in particular, we will derive from them *the algebra of operators* of the model. The form of (20) is a direct consequence of the quantum nature of the operators, together with the assumption of  $\Psi$ 's being a primary field, i.e. that it satisfies (40). It can be shown that the operator product expansion is equivalent to the canonical commutation relations of the modes of the fields. Eq. (20) thus encodes the short-distance behaviour of  $\Psi$ , and is often taken to be an alternative *definition* of a primary field  $\Psi$ .

In the classical model, the stress-energy tensor was given, in Eq. (19), by the squares of operators evaluated at the same point. In the quantum version of the model, this gives rise to divergences which need to be (and can be) renormalised. For the two-dimensional quantum field theories which we consider in this paper, the divergences are renormalised, to all orders, by the addition of a *single* counterterm to the action; alternatively, it suffices to define expressions such as (19) by *normal ordering*:

$$T(z) = -\frac{1}{2} : J(z) J(z) : , \quad \bar{T}(\bar{z}) = -\frac{1}{2} : \bar{J}(\bar{z}) \bar{J}(\bar{z}) : , \tag{21}$$

where the affine currents are still given, in the free bosonic scalar case, by (18), now as operator equations. The normal ordering is denoted by the colons. For the details of the normal ordering procedure, see e.g. Ginsparg (1990: Section 2.3).

Similarly to (20), the operator product expansion of the stress-energy tensor with itself can be worked out:

$$T(z) T(w) = \frac{c/2}{(z-w)^4} + \frac{2}{(z-w)^2} T(w) + \frac{1}{z-w} \partial T(w) . \tag{22}$$

---

<sup>34</sup>As remarked in footnote 32, in the quantum case we allow for (anti-) *meromorphic*, rather than (anti-) holomorphic, operators.

<sup>35</sup>In more detail:  $h + \bar{h}$  is the eigenvalue of the dilatation operator, and  $h - \bar{h}$  is the eigenvalue of the (Euclidean) rotation operator. Hence, the conformal weights contain information about the mass and the Euclidean spin of a field.

Here,  $c = 1$  for the free scalar (19), and it is called the *central charge*. Comparing the second term with (20), we see that  $h = 2$ , i.e. the stress-energy tensor  $T$  is a conformal field of weight  $(2, 0)$ , as expected from dimensional analysis. However, compared to (20), this expansion has, in addition, the first term. So  $T(z)$  is not a primary field in the sense of Eq. (20), nor does it transform as in Eq. (40). The term proportional to the central charge  $c = 1$  is obtained making use of the normalized expression (21). For another way to calculate the central charge, see Frishman and Sonnenschein (2010: Section 1.10).

In the same way, the operator product expansion of two affine currents (18), as well as that between  $T$  and  $J$ , can be calculated, with the following result:

$$\begin{aligned} J(z) J(w) &= \frac{1}{(z-w)^2} + \text{finite terms}, \quad (z \rightarrow w) \\ T(z) J(w) &= \frac{1}{(z-w)^2} J(w) + \frac{1}{z-w} \partial J(w). \end{aligned} \quad (23)$$

The second equation between  $T$  and  $J$  is simply (20) applied to the special case  $h = 1$ , owing to the fact that  $J = \partial\phi$  is indeed a primary field of weight  $(1, 0)$ .

Like in the canonical formalism, where a classical expansion of the field translates, quantum mechanically, into Fock modes: also in radial quantization it is convenient to introduce modes, which will give rise to creation and annihilation operators of the fields upon quantisation. That this is possible is ensured by (15), which says that the field can be decomposed into meromorphic and anti-meromorphic parts. In particular, we can do a Laurent expansion of the meromorphic and anti-meromorphic parts of the field. Thus, in virtue of (18) and (21), the affine currents and the stress-energy tensor will have their own Laurent expansions:

$$\begin{aligned} J(z) &= \sum_{n \in \mathbb{Z}} \frac{J_n}{z^{n+1}}, & \bar{J}(\bar{z}) &= \sum_{n \in \mathbb{Z}} \frac{\bar{J}_n}{\bar{z}^{n+1}} \\ T(z) &= \sum_{n \in \mathbb{Z}} \frac{L_n}{z^{n+2}}, & \bar{T}(\bar{z}) &= \sum_{n \in \mathbb{Z}} \frac{\bar{L}_n}{\bar{z}^{n+2}}. \end{aligned} \quad (24)$$

Overall factors of  $1/z$  and  $1/z^2$  can be extracted from  $J$ , respectively  $T$ : a fact which will be useful because these currents have  $h = 1$  and  $h = 2$ , respectively, and so  $J_0$  and  $L_0$ , as defined by (24), will have special physical significance. The summation range is infinite and therefore this normalization can always be reached, by a simple translation of  $n$ .

Because the currents satisfy (18) and (21), it is clear that  $J_n$  are linear in the creation and annihilation operators of the meromorphic field  $\phi$ , and the  $L_n$  are quadratic in (an infinite sum of) the  $J_n$ 's. This fact will be built into our considerations in what follows, though we do not work it out explicitly.

Finally, we find the algebras satisfied by  $J_n$  and  $L_n$ . Again these can be found either by canonical quantization or, in line with the methods we have used so far, they can be obtained directly from the operator product expansions (23) and (22), combined with the radial ordering prescription for the operators which contain circle integrals, mentioned above. Writing  $\delta_{m+n}$  as short for  $\delta_{(m+n)0}$  (defined as usual to be 1 or 0 according as



$m + n = 0$  or  $m + n \neq 0$ ): the result is:

$$\begin{aligned} [L_m, L_n] &= (m - n) L_{m+n} + \frac{c}{12} n(n^2 - 1) \delta_{m+n} \\ [J_m, J_n] &= -m \delta_{m+n} \\ [L_m, J_n] &= -n J_{m+n} . \end{aligned} \tag{25}$$

and the same algebra is satisfied by the  $\bar{J}_n$  and  $\bar{L}_n$ . Barred and unbarred quantities commute with each other. In the case at hand,  $c = 1$ .

We recognise the first line as the celebrated Virasoro algebra, as expected from the fact that the classical theory has a conformal symmetry group. In the second line, one may recognise the level  $k = 1$ , abelian Kac-Moody algebra. The third line is obtained from the operator product expansion between  $T$  and  $J$  in (23), and it makes the total algebra into the semi-direct product of the abelian Kac-Moody algebra and the Virasoro algebra. The algebra (25) is called the *enveloping Virasoro algebra* (with  $c = 1$  and  $k = 1$ ). The general enveloping algebra of the affine Lie algebra is given in (41) in the Appendix.<sup>36</sup>

This, i.e. Eq. (25), is the central result from the physics literature which we have been seeking, and will use in Section 5. For, together with (21) and the mode expansions (24), the tensor product of the holomorphic enveloping algebra of the affine Lie algebra (25) and its anti-holomorphic copy contains all of the information about the quantum version of the model. This is because the states now live in the vector space on which this algebra acts (a Hilbert space), and the quantities and the dynamics are constructed from the operators satisfying the algebra. We will show this in Section 5.

At this point, we notice that the generators  $L_{\pm}$  and  $L_0$  span an  $\text{SL}(2, \mathbb{R})$  subalgebra. Together with the generators  $\bar{L}_{\pm}$  and  $\bar{L}_0$ , which satisfy the same algebra, they form the conformal algebra of  $\text{SL}(2, \mathbb{C})$ , which is the symmetry group of the vacuum of this model.

## 4.2 The free, massless Dirac fermion

In this subsection, we consider our second model: of a free, massless Dirac fermion in two dimensions. Our goal is to rederive the infinite-dimensional algebra (25) in this model. We will follow similar steps as in Section 4.1: we will derive the classical symmetries, quantise the model, and obtain operator product expansions. (The discussion of how this accommodates to the sense of ‘model’, in Sections 2.2.2 and 2.2.3, is postponed to Section 5.2.)

The massless Dirac fermion is a two-component, complex spinor which can be decomposed as  $\Psi =: (\psi, \tilde{\psi})$ , where  $\psi$  and  $\tilde{\psi}$  are chiral (Weyl) fermions, called left- and right-chiral, respectively. The action is:

$$S_{\text{Dirac}} = \frac{1}{4\pi} \int d^2x \bar{\Psi} \not{\partial} \Psi = \frac{1}{4\pi} \int d^2z \left( \psi^\dagger \bar{\partial} \psi + \tilde{\psi}^\dagger \partial \tilde{\psi} \right), \tag{26}$$

---

<sup>36</sup>Notice that (25) satisfies the property that the level  $k$  can be changed by a rescaling of  $J$ . Thus, in the simple case we are dealing with here, in which the affine Lie algebra is based on the commutative Lie group  $\text{U}(1)$ , the level has no real meaning. This is not important for us, since we will not use it: rather, our analysis in Section 5 will be based on the fact that we are here dealing with a special case of the general enveloping algebra of the affine Lie algebra.

where in the middle expression we used the real form of the action, and in the last expression we used complex coordinates. The equations of motion imply that  $\psi$  and  $\tilde{\psi}$  are, respectively, holomorphic and anti-holomorphic (respectively meromorphic, in the quantum version of the model).

It is customary to further decompose the Dirac fermions into real, i.e. Majorana fermions, as follows:  $\Psi = \frac{1}{\sqrt{2}}(\Psi_1 + i\Psi_2)$ . In terms of the chiral (Weyl) fermions, we get  $\psi = \frac{1}{\sqrt{2}}(\psi_1 + i\psi_2)$ , where  $\psi_{1,2}$  are Weyl-Majorana fermions, and so the action takes the form of the sum of two copies of a single Weyl-Majorana fermion, which is conventionally called  $\chi$  (where  $\chi = \psi_{1,2}$ ). The action for a single Weyl-Majorana fermion is:

$$S_{\text{WM}} = \frac{1}{8\pi} \int d^2z (\chi \bar{\partial}\chi + \tilde{\chi} \partial\tilde{\chi}), \quad (27)$$

and again  $\chi, \tilde{\chi}$  are (if extended from the real line to the entire complex plane) meromorphic and anti-meromorphic, respectively.

Like in the case of the free, massless boson studied in Section 4.1, this action is invariant under two sets of symmetries:

(i) **Conformal transformations:**  $z \rightarrow f(z)$ ,  $\bar{z} \rightarrow \bar{f}(\bar{z})$ , the same transformations on the complex plane that we found in the bosonic model, Section 4.1.

(ii) **Left-holomorphic-chiral and right-anti-holomorphic-chiral transformations:** which act on the Weyl-Majorana fermions as follows:

$$\begin{aligned} \psi &\rightarrow \psi' = e^{i\alpha(z)} \psi \\ \tilde{\psi} &\rightarrow \tilde{\psi}' = e^{i\tilde{\alpha}(\bar{z})} \tilde{\psi}. \end{aligned} \quad (28)$$

Quantisation now proceeds similarly to what we did for the free, massless boson. One can do canonical quantisation (see Frishman and Sonnenschein (2010: Section 2.12)) or again use circle integration and radial ordering to calculate operator product expansions, and get from them the commutation relations for the generators.

First, let us notice that the definition of primary fields, Eq. (40) in the Appendix, applies equally well to fermions as it does to bosons. Also, the stress-energy tensor for fermions can be introduced, analogously to Eq. (19) (as we do below).

The Weyl-Majorana fermion is a primary field, as one finds from its operator product expansion with the stress-energy tensor (20). Like before, we have conserved currents (i.e. annihilated by  $\bar{\partial}$  for the meromorphic, and by  $\partial$  for the anti-meromorphic current) associated with the sets of symmetries (i) and (ii) above:

$$\begin{aligned} J(z) &= : \psi^\dagger \psi : , & \bar{J} &= : \tilde{\psi}^\dagger \tilde{\psi} : \\ T(z) &= -\frac{1}{2} : J(z) J(z) := -\frac{1}{2} : (\psi^\dagger \partial\psi - \partial\psi^\dagger \psi) : , \end{aligned} \quad (29)$$

with a similar expression for  $\bar{T}$  in terms of  $\tilde{\psi}$ .  $\chi$  has conformal dimension  $(\frac{1}{2}, 0)$ , and  $\tilde{\chi}$  has conformal dimension  $(0, \frac{1}{2})$ . The central charge (22) is  $c = \frac{1}{2}$  (and so, a Dirac fermion, which is the sum of two Majorana-Weyl fermions, has twice this amount, i.e.  $c = 1$ ).

Notice that despite the half-integral values of the conformal dimensions, the currents  $J(z), \bar{J}(\bar{z})$  in (29) actually have the same conformal dimensions as the currents (21) in the bosonic model. This is because they are now *quadratic* in the fermions.

As it turns out, the operator product expansion of  $J$  with itself is identical to that in (23). Consequently, because  $T$  in (29) satisfies the Sugawara construction (cf. after Eq. (19)), the product expansions of  $T$  with itself and of  $T$  and  $J$  are also identical to those in (22) and in (23). Now since also here,  $J$  and  $T$  are meromorphic operators, we can expand them in coefficients  $J_n$  and  $L_n$ , as in (24). The resulting algebra is thus the very same as in the bosonic case, (25), i.e. the semi-direct product of the Virasoro algebra with  $c = 1$  with the abelian affine algebra at level  $k = 1$  (and its anti-meromorphic copy)! Therefore, also the state-space will be the same, since it is constructed as the Hilbert space on which the algebra acts.

The implication of this basic fact—the agreement of the two models on their algebra of currents—will be explored in the next Section. We will see that it naturally leads to the existence of a duality, and thus to the formulation of a theory comprising the two models.

## 5 Boson-Fermion Duality Illustrates the Schema

In this Section, we show how Section 4’s bosonic and fermionic models (models in our sense!) illustrate the schema set out in Sections 2 and 3, leading up to the definition of duality in Section 3.2. In Section 5.1, we state the basic ‘dictionary’ of the duality, mapping fields, currents and stress-energy tensors. Then Section 5.2 builds on this, to show that the two models are isomorphic, in exactly the sense of our schema: i.e. as regards the whole trio of states, quantities and dynamics. Then in Section 5.3, we return to Section 2.4’s theme: of defining a theory by abstraction from its models. We first note some special features of our case-study, in particular that it has just two isomorphic models; and then we define a theory, a common core, from the two models. In Section 5.4, we discuss other ways one might define a theory from these models, i.e. so as to have them be representations of it. Finally, we briefly discuss generalizations of our case-study: to include massive particles, and to include non-abelian degrees of freedom (Section 5.5).

### 5.1 The duality dictionary

In Section 4, we derived the algebraic structures of the bosonic and the fermionic models. In particular, we saw that both models satisfy the enveloping algebra of the affine algebra (Eq. (25) and Eq. (41) in the Appendix) with  $c = 1$  and  $k = 1$ , for the quantum currents corresponding to the Noether symmetries.

At this point, Frishman and Sonnenschein (2010: Section 6.1) conclude that the two theories are equivalent. They write:

‘Due to the uniqueness of the irreducible unitary  $k = 1$  representation of the affine Lie algebras, and the fact that the infinite-dimensional algebraic structure fully determines the theories, we conclude that *in two space-time dimensions the theories of massless free scalar field and Dirac field are equivalent*. The equivalence implies that every operator

of one model should have a partner in the other model, in such a way that the operator product expansions of these operators should be identical.’ (p. 133: *Italics in the original*).

While we basically agree with this statement, one of its clauses—that the infinite-dimensional algebraic structure fully determines the theories (in our jargon form Section 2.2: the models)—has not been proven, and requires some explanation and justification. Doing that is our plan for this subsection and the next. In this subsection, we give the basic dictionary that Frishman and Sonnenschein refer to in the quoted passage. (Then in Section 5.2.1-5.2.3, we will argue that this infinite-dimensional algebraic structure indeed is sufficient to fully specify what we have called a model.)

The duality dictionary is given by the correspondence of the bosonic affine current algebra currents with the corresponding fermionic currents, and between the stress-energy tensors, as follows (cf. Frishman and Sonnenschein (2010: Section 6.1))<sup>37</sup>:

$$\begin{aligned} J_{\text{B}}(z) = \partial\phi(z) &\leftrightarrow J_{\text{F}}(z) = : \psi^\dagger(z) \psi(z) : \\ T_{\text{B}}(z) = -\frac{1}{2} : \partial\phi(z) \partial\phi(z) : &\leftrightarrow T_{\text{F}}(z) = -\frac{1}{2} : (\psi^\dagger(z) \partial\psi(z) - \partial\psi^\dagger(z) \psi(z)) : , \end{aligned} \quad (30)$$

and similarly for the anti-meromorphic currents. We have already seen that both sides satisfy the same operator product expansion, and therefore they satisfy the same algebra.

Notice that only fermion bilinears appear in (30). This is what we expect, since a single boson ( $h = 1$ ) should correspond to a pair of fermions ( $h = \frac{1}{2}$  and  $\bar{h} = \frac{1}{2}$ ). But it is not a priori clear which bosonic field a single fermion should correspond to. One would here expect to take some kind of ‘square root’ of the boson. But, surprisingly, the dictionary turns out to extend to a single fermion field as follows:

$$\begin{aligned} : e^{i\phi(z)} : &\leftrightarrow \psi(z) , & : e^{-i\phi(z)} : &\leftrightarrow \psi^\dagger(z) \\ : e^{-i\bar{\phi}(\bar{z})} : &\leftrightarrow \tilde{\psi}(\bar{z}) , & : e^{i\bar{\phi}(\bar{z})} : &\leftrightarrow \tilde{\psi}^\dagger(\bar{z}) , \end{aligned} \quad (31)$$

and again the operator product expansions agree. At first, this is a very surprising result, reminiscent of the construction of a coherent state. However, the operator product expansion shows that the conformal dimension of  $: e^{i\alpha\phi(z)} :$  is  $h = \alpha^2/2$ , so the above dictionary indeed reproduces  $h = \frac{1}{2}$  and  $\bar{h} = \frac{1}{2}$  for the left- and right-chiral Dirac fermions, respectively. This is indeed a purely quantum result, with no straightforward classical analogue.

One of the most surprising features of (31) is that  $\phi$  satisfies canonical commutation relations, while  $\psi$  satisfies canonical *anti*-commutation relations. How can this be? To calculate the anticommutator of two  $\psi$ ’s, one uses the formula  $e^A e^B = e^{[A,B]} e^B e^A$ , which holds when  $[A, B]$  is a c-number. Using this formula to evaluate the commutator of two exponentials, and using the canonical commutation relations for the boson and its conjugate canonical momentum, one finds that, indeed, the left- and right-chiral fermions anti-commute!

---

<sup>37</sup>The dictionary thus relates (18) and (21) to (29). We here add the subscripts ‘B’ and ‘F’ for ‘bosonic’ and ‘fermionic’, respectively.

## 5.2 Two isomorphic model triples

In Section 5.1, we gave the ‘dictionary’ between the bosonic and fermionic models. This is a bijection between the basic operators of the theory (the fields). But there is, of course, more to duality than this. We need to show that the two models are *isomorphic*, as triples. Recall our second sense of ‘model’, in Section 2.2.2: as a representation of a theory. So a model is a triple (what we called the ‘model triple’), together with some specific structure. And the model triple was not a ‘pure copy’ of the theory, but a representation of it using the specific structure. The next three subsections deal with the triples of states, quantities, and dynamics. In each of the subsections we show the existence of an isomorphism between the states, quantities, and dynamics of the two models (in the third case, an equivariance relation). This will justify that the boson-fermion equivalence is indeed a duality, in the sense of our schema. The final subsection considers the interpretation of this duality.

### 5.2.1 States

We begin by showing that the state spaces of the two models are isomorphic. This will form the first item in our model triple. The state spaces of the bosonic and the fermionic models were introduced at the ends of Sections 4.1 and 4.2, as the Hilbert spaces obtained from the representations of the enveloping algebra of the affine Lie algebra (25) (which, for short, we shall also call the ‘enveloping algebra’). We give some more details here.

The enveloping algebra is realized in the bosonic and the fermionic cases in terms of different fields: the bosonic  $J$ -currents (18) and the fermionic  $J$ -currents (29) are defined in terms of different fields. Consequently, also the stress-energy tensors, given in the two cases by the Sugawara construction, differ, and so do the mode operators  $J_n$  and  $L_n$  which enter the enveloping algebra of the affine Lie algebra. In other words, we have two representations of the same algebra. Let us call these  $\mathcal{A}_B$  for the bosonic representation and  $\mathcal{A}_F$  for the fermionic representation.

The state spaces are, as mentioned in Sections 4.1 and 4.2, the Hilbert spaces obtained from the representations of the enveloping algebra, Eq. (25): or, better, its generalisation, Eq. (41) in the Appendix, to general underlying Lie group and general value of  $k$ . The vacuum state is obtained by requiring  $J_n^a |0\rangle = 0$ , for  $n \geq 0$ , where  $a$  labels the generators of the Lie algebra. Here, the generators  $J_n^a$  are the ladder operators of the algebra. This condition can also be understood physically as the regularity condition  $J(z)|0\rangle = 0$  at  $z = 0$  for the affine currents.

The non-trivial representations are constructed from the irreducible representations of the algebra, which are uniquely characterised by the highest-weight states (analogous to states of maximal  $J$  for  $SU(2)$ ), obtained by application of a primary field (cf. Eqs. (20) and (40)) to the vacuum. If we denote such a state by  $|l, \bar{l}\rangle$ , where  $l, \bar{l}$  are the representations, the remaining states in the representation (so-called ‘descendants’) are obtained by appropriately applying generators, and take the generic form:

$L_{-m_1} \cdots L_{-m_M} \bar{L}_{-\bar{m}_1} \cdots \bar{L}_{-\bar{m}_{\bar{M}}} J_{-n_1}^{a_1} \cdots J_{-n_N}^{a_N} \bar{J}_{-n_1}^{a_1} \cdots \bar{J}_{-n_{\bar{N}}}^{a_{\bar{N}}} |l, \bar{l}\rangle$ , for some integers  $M, \bar{M}, N, \bar{N}$ . For instance, on the fermionic model, expanding the fermion into modes:<sup>38</sup>  $\psi(z) = \sum_{r \in \mathbb{Z} + \nu} \frac{\psi_r}{z^{r+\frac{1}{2}}}$ , the  $\psi_{-n}$  operators are creation operators, and the  $\psi_n$  are annihilation op-

<sup>38</sup>Here,  $\nu = \frac{1}{2}$  for periodic boundary conditions, and  $\nu = 0$  for anti-periodic.

erators. The fermionic vacuum is defined as  $\psi_n |0\rangle = 0$  ( $n > 0$ ), and the states in the Hilbert space are of the type  $\psi_{-n_1} \cdots \psi_{-n_k} |0\rangle$ .

We will use the fact that *the irreducible unitary representations of the enveloping algebra, thus constructed, are unique up to unitary equivalence.* (See e.g. Frishman and Sonnenschein (2010, p. 133)).<sup>39</sup>

So, for each of the models, we construct a representation of the algebra on a Hilbert space,  $\mathcal{H}_B$  for the bosons and  $\mathcal{H}_F$  for the fermions; and these representations are unitarily equivalent. Let us denote the unitary operator in question  $U$ .

### 5.2.2 Quantities

Next, we show that also the quantities of the two models are isomorphic, thus providing the second item of the model triple. What are the relevant physical quantities? We already mentioned in Section 3.2.1 that we take the physical quantities to be all the renormalizable, self-adjoint operators constructed from a more basic set of quantities (which in a moment we will identify with the currents) and respecting the appropriate symmetries.

Let us look again at the two representations,  $\mathcal{A}_B$  and  $\mathcal{A}_F$ , of the enveloping algebra of the affine Lie algebra, which we discussed in Section 5.2.1. We can say more about these representations: for we already have the ‘dictionary’, Eq. (30)-(31), between the fields. At this point we know that this dictionary is a bijection, but we wish to find a condition for it to be an isomorphism. First: notice that this bijection induces a similar bijection between the respective mode operators, i.e. it induces a bijective map between the algebras,  $d : \mathcal{A}_B \rightarrow \mathcal{A}_F$ : ( $d$  for ‘duality’)

$$\begin{aligned} d(J_{n,B}) &= J_{n,F} \\ d(L_{n,B}) &= L_{n,F} . \end{aligned} \tag{32}$$

This is of course just the statement that we have two equivalent representations of the enveloping algebra.

From the previous section, we know that the representation spaces  $\mathcal{H}_B$  and  $\mathcal{H}_F$  of these algebras are constructed from the irreducible representations of highest weight, which are unique up to unitary equivalence. Since the construction of the representations of the states is the same for the two models (i.e. in terms of the algebra generators mapped by Eq. (32)), the algebra generators themselves must be compatible with the bijection which maps the highest weight representations on the two sides. So, for consistency we must require:

$$d(\cdots) = U^\dagger (\cdots) U , \tag{33}$$

where  $U$  is the *same* map used to map the states of the highest-weight representations in the previous subsection. If the maps  $U$  and  $d$  were not related *a la* (33), the structure of the theory would not be preserved by the duality (since an operator acting on a state would not be mapped to the the corresponding operator acting on the corresponding

---

<sup>39</sup>For more details, see Di Francesco et al. (1997: Chapter 14) or Kac (1990).

state). In short: once we have fixed the states to be mapped by  $U$ , as in Section 5.2.1, then the *same* transformation must map the quantities, as in (33).

But Eq. (33) is precisely the condition for the map to be an *isomorphism*, in addition to a bijection. And it now follows from the above that the modes of the currents of the two models are mapped as:

$$\begin{aligned} U^\dagger J_{n,B} U &= J_{n,F} \\ U^\dagger L_{n,B} U &= L_{n,F} . \end{aligned} \tag{34}$$

It also follows that the same relations hold for the currents (30) themselves:

$$\begin{aligned} U^\dagger J_B(z) U &= J_F(z) \\ U^\dagger T_B(z) U &= T_F(z) . \end{aligned} \tag{35}$$

The same map also maps the fermionic operators:

$$\begin{aligned} U^\dagger : e^{i\phi(z)} : U &= \psi(z) , & U^\dagger : e^{-i\phi(z)} : U &= \psi^\dagger(z) \\ U^\dagger : e^{-i\bar{\phi}(\bar{z})} : U &= \tilde{\psi}(\bar{z}) , & U^\dagger : e^{i\bar{\phi}(\bar{z})} : U &= \tilde{\psi}^\dagger(\bar{z}) . \end{aligned} \tag{36}$$

This is the formalization of (31), which was justified by checking its properties (including the correct statistics, but also all the correct correlation functions). Given (31), the factors of  $U$  follow from the fact that  $\psi(z)$  creates or annihilates a *fermionic* state in  $\mathcal{H}_F$ , and this state is obtained from the corresponding bosonic state in  $\mathcal{H}_B$  using the same map  $U$ . The operator creating or annihilating this state must therefore transform bilinearly in  $U$ . One can also check that the map between the affine currents, the first equation in (34), follows from (36).

We have discussed the maps between the operators of the enveloping algebra of the affine Lie algebra, and the corresponding fields. We now discuss the *physical* quantities,  $\mathcal{Q}_B$  and  $\mathcal{Q}_F$ , of self-adjoint operators respecting the relevant symmetries. In the fermionic model, all the self-adjoint operators are quadratic in the fermions: and taking into account the chiral symmetry algebra, they must necessarily be powers of the fermionic currents (35) (and their anti-chiral counterparts), appropriately normally-ordered. In fact, arbitrary analytic functionals of the currents are allowed. There can be a mixing between the meromorphic and anti-meromorphic sectors in the analytic functionals, but only such that the chiral symmetry algebra is preserved. Thus, the enveloping algebra indeed contains all of the information about the physical quantities:  $\mathcal{Q}_F$  consists of arbitrary analytic functionals in the fermionic currents (35) (and their anti-meromorphic counterparts), appropriately normally-ordered. The normal ordering automatically ensures that these correlation functions are well-defined.

One can in principle enlarge the set  $\mathcal{Q}_F$  to also contain the correlation functions of operators which violate the chiral symmetry algebra, evaluated on the same states, i.e. without changing the Lagrangian of the model. However, one is then changing the symmetries of the model triple  $f = \langle \mathcal{H}_F, \mathcal{Q}_F, \mathcal{D}_F \rangle$ , and hence one is defining a *new model triple* (and, consequently, if there is a duality for this larger class of models, one is defining a new

theory). We will return to this possibility when we discuss sine-Gordon/massive Thirring duality in Section 5.5.1.

In the bosonic model, there is a similar structure. Now it is not the mixing of *chiralities* of the fields which the symmetries forbid, but the appearance of the *undervived* field  $\phi(z)$  in  $\mathcal{Q}_B$ . Namely, the translation symmetry  $\Phi \rightarrow \Phi + a$  (or, more generally, the affine current symmetry algebra (17)), forbids the appearance of operators which depend on  $\phi(z)$  or  $\bar{\phi}(\bar{z})$  *directly*, i.e. as opposed to depending on them through their derivatives. When the model is written in terms of the meromorphic and anti-meromorphic parts of  $\Phi$ , translation symmetry is preserved iff all operators depend on the *derivatives* of the scalar field. This is precisely how the scalar field appears in the currents (35). Thus, once again, the affine current symmetry algebra of the model tells us that all the operators which are physical quantities of  $\mathcal{Q}_B$ , must be analytic functions of the two currents (and their anti-meromorphic counterparts).  $\mathcal{Q}_B$  consists of all possible analytic functionals of those currents. As for the renormalizability constraint on the physical quantities: as we mentioned before, normal ordering automatically takes care of getting the correct renormalized quantities.

To end this subsection, let us get a possible worry out of the way. Namely, one might be concerned that (36) seems to be introducing complex operators, such as  $: e^{i\phi(z)} :$ , into the set of physical quantities of the bosonic model, which is supposed to be entirely real. But this is not quite right. For (36) are in fact *not* physical operators, on our conception of the term. It is indeed true that  $: e^{i\phi(z)} :$  is not self-adjoint in the bosonic model, but then neither is  $\psi(z)$  self-adjoint in the fermionic model. Neither of these two operators are thus to be taken as physical on either of the two sides, even if we can build physical operators, such as the currents (30), by taking powers of them. Thus, none of the operators in (36) belong to either  $\mathcal{Q}_B$  or  $\mathcal{Q}_F$ . Rather we should think of such operators as *states* in the Hilbert space, using the state-operator correspondence. Indeed, recall, from Section 5.2.1, that the states in the fermionic Hilbert space are of the type:  $\psi_{-n_1} \cdots \psi_{-n_k} |0\rangle$ . So this explains how  $\psi(z)$  can be a physical operator—namely, it creates and annihilates physical states in  $\mathcal{H}_F$ —while it is not self-adjoint and does not belong to the set  $\mathcal{Q}_F$  of physical quantities.

Perhaps the most surprising aspect of the above is the fact that, in two dimensions, *the Hilbert space of the bosonic model contains fermionic states*, e.g. states with conformal weight  $(\frac{1}{2}, 0)$ . In fact, because the conformal weight of the operator  $: e^{i\alpha\phi(z)} :$  is  $h = \alpha^2/2$ , the Hilbert space contains a 1-parameter family of states with a continuous range of Euclidean spin values.<sup>40</sup> This feature of the quantised models is indeed surprising—and illustrates our theme of surprise announced in (2) of Section 2.1.

### 5.2.3 Dynamics

Finally, we discuss the equivariance of the dynamics of the two model triples. ‘Dynamics’ can be understood in different ways in different theories, and even in one formulation of

---

<sup>40</sup>Our use of the word ‘Euclidean spin’ here follows the jargon in the physics literature, for the eigenvalue under Euclidean rotations, as we mentioned in Section 4.1. It is questionable whether such jargon is justified by the physical interpretation in 1+1 dimensions. But we will not need to dwell on this point, since our main aim in this Section is formal.



a single theory. Think, for example, of the difference in the dynamics if it is formulated in the Heisenberg or in the Schrödinger pictures of a theory.

First of all, we have formulated our model triples in the Heisenberg picture: operators are generally time-dependent and the states are time-independent. Let us consider the bosonic model first. We are working in Euclidean spacetime, but when analytically continued to Minkowski spacetime the operator  $H_B := L_{0,B} + \bar{L}_{0,B}$  (which generates dilatations on the plane) is the generator of time translations, and is to be identified with the Hamiltonian. It is indeed the 00-component of the stress-energy tensor. The same is true in the fermionic model: the 00-component of the stress-energy tensor, which is the generator of time translations, is the operator  $H_F := L_{0,F} + \bar{L}_{0,F}$ . These two Hamiltonians are mapped to each other by the map  $U$  in (34) (and the anti-meromorphic version of that equation). Thus the dynamics is correctly preserved by the duality map: more precisely, it is equivariant with the unitary transformation.

We see that the requirement that the dynamics is correctly mapped between the two theories does not give any additional piece of the duality map, but simply follows from the other two: since all the quantities were already dual. This seems to be a general fact in dualities between quantum field theories, though we do not necessarily expect it to be the case for any duality. Namely, in a quantum field theory model, once we require that all states and physical quantities between the two theories map correctly, i.e. according to the *isomorphism* (33), then all correlation functions of the model automatically map correctly as well. But the correlation functions of a quantum field theory model exhaust all the dynamical information about the model, so that knowing all the correlation functions we can, in principle, reconstruct the dynamics.

Thus, an alternative way to formulate duality is in terms of the correlation functions. This is, in fact, how duality is usually formulated in the string theory literature. However, we find our own conception of duality more illuminating, because it allows us to map the states and the operators individually and directly, without having to unscramble this information from the full set of correlation functions. But either way, the main point is that the dynamics is correctly equivariant with the duality map.

#### 5.2.4 Interpretation

With our boson-fermion duality now in hand, we return to Section 3.2.3's theme that a duality should respect the interpretation maps introduced in Section 2.3. Our initial point in Section 3.2.3 was that since duality relates two model triples, and interpretation maps apply to model triples (strictly speaking: their components, such as the set of quantities), interpretation simply proceeds independently on the two sides of the duality. The range of these interpretation maps could be: either two distinct but isomorphic 'sectors of reality', or the very same sector of reality—in which case there is a triangular rather than square diagram, as in Figures 8 and 9. We also said that the choice between these cases was a matter of an 'internal' vs. an 'external' interpretation.

External interpretations for the models are straightforward to read off from the model triples: they are the bosonic, respectively fermionic, interpretations which the two models come with in the first place (i.e. in this case, their original historical interpretations). There is a natural map  $I_{\text{int}}$  which assigns intensions: for example, it maps a bare or

abstract bosonic state to the meaning ‘boson on a line with such-and-such properties’. And the map  $I_{\text{Ext}}$  assigns extensions. For example, it maps the abstract expectation value of a bosonic field to a measurable property of a specific boson that is fixed as the reference by the context of use. For example, the property might be the boson’s amplitude or probability of being at a specific place at a specific time: we say ‘amplitude or probability’, since we understand a value or expectation value, written schematically as  $\langle Q, s \rangle$ , to also represent all the matrix elements  $\langle s_1 | \hat{Q} | s_2 \rangle$  (cf. Section 3.2.1-(1)). In the fermionic model, there are similar interpretation maps for the fermionic model triple: where the codomains of the maps are now not the properties of bosons, but the properties of (say) electrons! Thus clearly, on the external interpretation the codomains of the maps do not need to agree. In philosophical jargon: the set of possible worlds where the bosons and the fermions appear need not be the same sets of worlds.

The internal interpretation abstracts from the specific structure of the two models, so that the two model triples receive the same interpretation. This interpretation is therefore best worked out at the level of the theory. Since the duality is exact, each model contains both bosonic and fermionic states, as we discussed in Section 5.2.2. This means that the world (or set of worlds) which is the codomain of our interpretation, should contain both bosons and fermions. The interpretation maps, applied to the models, must commute with the duality map  $d$  (or, equivalently,  $U$ ) in Eqs. (32)-(33). The interpretations will map  $J(z)$  and  $T(z)$  to a current and a stress-energy tensor, respectively; (as intensions or as extensions, as appropriate: in the ensuing discussion, we shall not make this distinction, since everything we say applies equally well to both kinds of maps). The internal interpretation also maps the abstract expectation value of a bosonic field (whether written as  $\phi(z)$  or as a fermion bilinear) to the amplitude or probability of some bosonic event. Likewise, the abstract expectation value (matrix element) of a fermionic field between two suitable given states will be mapped to the transition amplitude between two fermionic states.

### 5.3 Defining a theory from the two model triples

We now return to Section 2.4’s theme: of defining a theory by abstraction from its model triples. We first discuss how our duality’s having just two isomorphic model triples makes this enterprise vulnerable, in two ways (Section 5.3.1). Then we undertake to define a theory, a common core, from the two models; and discuss their specific structures (Section 5.3.2). Thus the general issues of Section 5.3.1—which return us to the Hilbertian themes at the start of Section 1—will lead in to the specific details of Section 5.3.2.

#### 5.3.1 Two vulnerabilities

At first sight, there might seem to be no issues about the definition of a theory (in our sense, from Sections 1.1 and 2.2.1) from a pair of model triples that are isomorphic: or indeed, from any set of two or more isomorphic model triples. Can we not simply define the theory as the structure of which the isomorphic model triples are isomorphic ‘concrete’ copies? More precisely: here we should clarify the phrase ‘as the structure of which’, in order to respect Section 2.2.2’s point that a model (in our sense!) usually realizes a

theory by being a representation of it, and representation allows mere homomorphism, rather than isomorphism. So the thought is: can we not simply define the theory as the structure,  $S$  say, of which the (two or more) isomorphic model triples are representations—as it happens, isomorphic ones? Talk of  $S$  thus defined as a theory will then engender admitting all its representations as models in Section 2.2.2’s sense.

We agree that this relaxed attitude is tenable. After all, ‘theory’ is a term of art: so one is at liberty to define it as one sees fit. But in both pure mathematics and theoretical physics, the abstraction (‘extraction’) of a general pattern or structure from a given set of examples is a matter of judgement, to be made in the light of one’s aims and intuitions, including the aim of representing the world as accurately as possible. This means that there are two related worries one might have: that is, two limitations to which the above strategy for defining a theory is vulnerable.

(i): Suppose that—as in our bosonization case-study—the given examples are isomorphic: where ‘isomorphic’ is itself a term of art, made precise by some mathematical structure we see exactly copied in the examples. Then one worries that there might be non-isomorphic examples which one has not thought of (has not ‘been given’) that are equally good examples of the general pattern one is trying to write down—that equally fit one’s aims and intuitions.

(ii): Suppose the given examples are not isomorphic; (again, in the sense we have in mind, especially initially). Still one worries that they might not be sufficiently varied, so that the pattern one writes down after considering them is too restrictive. That is: the pattern encodes aspect(s) that, given one’s aims and intuitions, are really accidental commonalities of the examples. Recall (from the end of Section 2.4) the example of colour as an artefact that could beset Frege’s definition of direction as an equivalence class of mutually parallel lines.

The general answer to these worries lies in the point, argued in Section 3.2.2, that for models which purport to describe the physical world, the distinction between what is in the triple and what is specific structure cannot be blurred.<sup>41</sup> So these worries can be set aside: for a give target system, typically only a single triple will provide a complete description—up to isomorphism, that is.

So much by way of rehearsing the general issues about defining a theory from models: specifically, from model triples. For the purposes of this paper, what matters is how these issues play out in our case-study, bosonization. Section 5.3.2 will give details about this. But to summarise:— The theory we will construct in Section 5.3.2 is the simplest one that

---

<sup>41</sup>As an example, consider the bosonic and fermionic models, but now weakened by the stipulation that the Virasoro algebra belong to the model triple, while the affine Kac-Moody algebra (and the third line in Eq. (25)) belongs to the specific structure. As we argued in Section 3.2.2, this stipulation *changes the physical content* of the models, and so it is not innocuous. The models thus obtained contain different numbers of (uninterpreted) physical degrees of freedom, and so cannot describe the bosons or the fermions of Section 4. This is because the boson and the fermion CFTs (even before they are physically interpreted) treat the Kac-Moody degrees of freedom not as ‘accidental commonalities’, in the sense of Section 2.4: but as physical, and related to the Virasoro generators by the Sugawara construction. (For example, if we drop the chiral symmetry on the fermionic model, we lose the reason to restrict to chiral quantities only: cf. Section 5.2.2; and likewise for the boson’s affine current symmetry algebra.) Thus the boson and fermion models are *not* dual, if based on just the Virasoro algebra. We thank Josh Hunt for bringing up this example.

can be constructed from our bosonic and fermionic model triples, using the isomorphism at hand from Section 5.2. But in line with this Subsection's comments, we make no claim that is the only (or even best) way to 'extract a pattern' from the model triples. Indeed, our final two Subsections, Sections 5.4 and 5.5, *will* consider other such ways: in particular, by taking into account other models.

### 5.3.2 Defining the theory

In this subsection, we collect our results from Section 5.2 about the two model triples, in order to define our *theory*. We have so far defined two *models*:—

(a) A bosonic model triple,  $b := \langle \mathcal{S}_B, \mathcal{Q}_B, \mathcal{D}_B \rangle$ , with:  $\mathcal{S}_B = \mathcal{H}_B$ , the Hilbert space which is constructed from the irreducible highest-weight representations of the enveloping algebra of the affine Lie algebra, represented on the bosons. 'Represented on the bosons' here means that the generators of the affine Lie algebra (25), which proceed from the Laurent expansion (24), are constructed from the *bosonic* fields (18) and (19) (and their anti-holomorphic counterparts). Thus,  $\mathcal{Q}_B$  is the set of normally-ordered analytic functions in the bosonic currents  $(J(z)_B, \bar{J}(\bar{z})_B, T(z)_B, \bar{T}(\bar{z})_B)$ ; and  $\mathcal{D}_B = L_{0,B} + \bar{L}_{0,B}$  the dilatation operator of the bosonic model.

(b) A fermionic model triple,  $f := \langle \mathcal{S}_F, \mathcal{Q}_F, \mathcal{D}_F \rangle$ , with:  $\mathcal{S}_F$  the same Hilbert space, represented on the fermions,  $\mathcal{H}_F$ ;  $\mathcal{Q}_F$  the set of normally-ordered analytic functions of the currents  $(J(z)_F, \bar{J}(\bar{z})_F, T(z)_F, \bar{T}(\bar{z})_F)$ ; and  $\mathcal{D}_F = L_{0,F} + \bar{L}_{0,F}$  the dilatation operator of the fermionic model.

What structures make up the 'specific structure'  $\bar{M}$ , of each of the models  $B := \langle b, \bar{B} \rangle$  and  $F := \langle f, \bar{F} \rangle$ , in the sense of our notation in Section 2.2.3, and especially Eq. (2)? On the bosonic side, the specific structure  $\bar{B}$  clearly contains the field  $\phi(z)$  (and functionals of it), together with the symmetry algebra (17) acting on it. This symmetry algebra will, however, manifest itself in the model triple through the affine currents and their algebra. Also the defining relations of the field (equation of motion, etc.) are specific to  $\bar{B}$ .

On the fermionic side, it is the field  $\psi(z)$ , with its chirality symmetry (and a different set of defining relations, equation of motion, etc.), which are parts of the specific structure  $\bar{F}$ . Though  $\psi(z)$  defines a state in the Hilbert space, as discussed in Section 5.2.2, thinking of this state as created by  $\psi(z)$ , i.e. a fermion with certain meromorphic and chirality properties, it is part of the specific structure. All the Hilbert space knows about this fermion is that there is a state of conformal weight  $(\frac{1}{2}, 0)$ .

We are now ready to discuss the *theory* which we can construct from these two models: by discussing the common core of the two models, i.e. the model triples (or roots), to which the theory is isomorphic:  $b \cong f \cong T$ . It is a theory based on four currents  $(J(z), \bar{J}(\bar{z}), T(z), \bar{T}(\bar{z}))$  of conformal dimensions  $(1, 0)$ ,  $(0, 1)$ ,  $(2, 0)$ ,  $(0, 2)$ , satisfying the enveloping algebra with  $c = 1$  and  $k = 1$ . Its states are the unitary representations of this algebra. The dynamics is given by singling out the Hamiltonian  $H = L_0 + \bar{L}_0$ .

*Sets of symmetries of the theory:* The theory has two built-in sets of symmetries: (i) a conformal group and (what we shall call), generated by the stress-energy tensor, (ii) an affine current symmetry algebra, generated by the  $J$ -currents.

(i) The conformal symmetry group, Eq. (16), is represented in the same way in the two models—since this is a symmetry group of the background spacetime. This is also related to the fact that the Hamiltonian, and hence the dilations, are related by equivariance between the two theories. Again:  $T(z)$ , the stress-energy tensor of the theory, is represented in both the models.

(ii) The affine current symmetry algebra is represented very differently in the two models! Namely, as affine current algebra transformations (Eq. (17) in Section 4.1(ii)) on the bosonic model, but as left- and right-chiral symmetry algebra (Eq. (28) in Section 4.2(ii)) on the fermionic model. This symmetry algebra restricts the kinds of physical quantities in the theory, as we have explicitly discussed in Section 5.2.2, in the same way. Yet the theory does not “see” any of the features from which the affine current algebra arises: which are very different in the two models, viz. the transformations of the fields Eq. (17) vs. (28).

We have discussed this duality in some detail because it is a good model to the more general, and technically involved, boson-fermion dualities in two-dimensional conformal field theory: to which we turn in Section 5.5.

## 5.4 Further abstraction

The discussion in Sections 5.2.2 and 5.3 illustrates our theme of the limitations of abstraction. We constructed our *theory* from two isomorphic model triples: the model triple of the theory was built from the representations of the enveloping algebra of the affine Lie algebra (25), which give a unique set of *states* (discussed in Section 5.2.1), and a basic set of *quantities*: the generators of the algebra themselves. The *dynamics* was a choice of a Hamiltonian among the quantities (Section 5.2.3). As we saw in Section 5.2.2, the *full* set of quantities  $\mathcal{Q}$  of the theory contains more than just the basic set: so that arbitrary analytic functionals of the currents ( $J(z), \bar{J}(\bar{z}), T(z), \bar{T}(\bar{z})$ ), and their derivatives, are allowed. Compare the discussion of the symmetry algebra (ii) in Section 5.3.

But we also know, from Sections 2.4 and 5.3.1, that there are no general rules, fixed once-and-for-all, for defining theories. So one asks: to what extent is our procedure above unique? It is of course unique if what one wishes to describe is a boson or a fermion, as given systems with known degrees of freedom. But the procedure of abstraction suggests three natural ways in which our theory might be modified. The first way would make for a more *restrictive* theory; the other two entail further abstraction, thus allowing for a more *general* one.

(a) The *conformal symmetry group* (Section 5.3.2-(i)) was used to form the enveloping algebra. Hence it is realised by the states and the quantities, in the sense that the states and quantities form representations of this symmetry group. But one can construct a new theory in which the class of operators is *reduced*: namely, by placing restrictions on the conformal transformation properties of the quantities. This leads to a theory with a smaller set of quantities (and subsequently to bosonic and fermionic models with smaller sets of quantities). Alternatively, one can reduce the space of states by placing similar restrictions on them.

(b) The *affine current symmetry algebra* (Section 5.3.2-(ii)) limited the set of quantities

to those that can be constructed from the currents  $(J(z), \bar{J}(\bar{z}), T(z), \bar{T}(\bar{z}))$  (as mentioned at the start of this subsection). But one might decide that this is not a symmetry algebra one wishes to keep, for the physical system of interest; (for example, in the presence of mass terms, this symmetry algebra will be dropped). One then allows a larger class of, or even all, self-adjoint, renormalisable operators constructed from (31), not just the ones that preserve this symmetry algebra. In this way, one clearly gets a richer theory (with a larger set of operators and states), of which Section 4's two models are still representations.<sup>42</sup> And as we stressed in Section 3.2.2, it is a matter of physical judgment, which kinds of operators one needs to admit in order to describe the physics at hand. In particular: if one wants to add a mass, one is forced to generalise the theory in this way. The boson-fermion duality continues to hold, thanks to the existence of the maps Eq. (31) and (34). But we get a *more general class of theories*, which will not necessarily be each other's duals: the generality of the class depends on which additional set of operators one takes on board with the quantities. This will be illustrated explicitly in Section 5.5.

(c) Though the two models share the *spacetime coordinates*  $z, \bar{z}$ , these coordinates do not enter the basic considerations that led to building the states, quantities, and dynamics of the theory, in Sections 5.2.1-5.2.3. Indeed, the basic object of interest, in constructing the triple, is the algebra of the mode operators  $L_n$  and  $J_n$  (and powers of them): and these are spacetime-independent. Furthermore, these modes contain essentially the same information as the spacetime-dependent currents  $(J(z), \bar{J}(\bar{z}), T(z), \bar{T}(\bar{z}))$ , i.e. the latter can be reconstructed from the former through the Laurent expansions (24), which take identical form in the fermionic case. So, one might decide that  $z, \bar{z}$  are just book-keeping devices with no essential information about the theory. On this view, one can construct a theory just based on the algebra (25) and its representations, without its spatio-temporal realization. The bosonic and fermionic model triples then still form (spatio-temporal) representations of this algebra: but one can envisage the existence of other representations, which are not spatio-temporally realised. This would presumably give rise to non-isomorphic models, in the sense of Sections 2.4 and 5.3.1. While doing away with spacetime may seem a radical suggestion, it is not so uncommon: think e.g. of spin chains as possible models.

Point (a) is a straightforward modification of our theory, but also of the models. So it should not be seen as illustrating the limitations of abstraction, in the sense of Sections 2.4 and 5.3.1. Rather, it is a *method to obtain more restrictive theories*, by consistently strengthening the symmetry requirements of the models.

But points (b) and (c) do illustrate our remarks, in Sections 2.4 and 5.3.1, about the need for models to be 'sufficiently varied'. By taking, in (b) and (c), some of the symmetry shared by the models to be accidental, one gets a larger class of models, which is likely to include non-isomorphic ones.

In the next subsection, we will give such examples of dualities between isomorphic models which are more general: either because they have less symmetry, or because they have more fields, with additional symmetries.

---

<sup>42</sup>Notice that the ambiguity here is in the best definition of the *theory*, not of the *duality*: cf. footnote 41.

## 5.5 General boson-fermion dualities

In this Subsection, we briefly discuss two generalizations of our basic duality. The two generalizations are important for our discussion since they fulfill, even better than our basic duality does, Section 1.2's two desiderata for the choice of examples of dualities. Recall that these desiderata were: on the one hand, (i) an example should be mathematically precise and represent established physics; and on the other hand, (ii) an example should be scientifically important (as described more fully in Section 2.1).

As we will see, these desiderata will be amply fulfilled by this Subsection's two generalizations. The first generalization especially illustrates scientific importance. The second illustrates, not so much mathematical precision *per se*: but rather, *mathematical richness and generality*, thus showing that the boson-fermion duality described in this Section is not an isolated 'coincidence' that holds for free, massless models, but part of a very rich class of mathematically interesting (as well as rigorous!) isomorphic models. Thus we will here define a rich class of theories, each of which is an equivalence class of exactly two isomorphic model triples.

We will discuss the two kinds of generalization in turn, in the next two subsections. Section 5.5.1 discusses the duality between the massive Thirring model and the sine-Gordon model. Section 5.5.2 considers non-abelian versions of boson-fermion duality. In both cases, we must be brief and must suppress technical details.

### 5.5.1 Duality between the Thirring model and the sine-Gordon model

The basic boson-fermion duality can be extended to include mass terms for the fermions, and interaction terms for both fermions and bosons. This generalisation is important, because it shows that the duality is not a special property that only occurs in the free, massless case, in which the action is conformally invariant. Massive, and interactive, theories are also subject to duality. So this strengthens the scientific importance of duality: it brings duality into the 'real world'. In fact, the Thirring model-sine-Gordon duality is quite important in condensed matter systems. See Giamarchi (2003), Altland and Simons (2010).

The massless Thirring model generalises the free, massless Dirac fermion by the addition of a quartic interaction term for the fermions, with coupling constant  $g$ . This quartic interaction is built from the  $J$ -currents, and so preserves the chiral symmetry algebra described in (ii), Section 4.2. The model can be solved exactly, and the quantum theory is well-defined only for  $g > -\pi$  (Coleman (1975: p. 2094)).<sup>43</sup>

The fermionic mass term in the massive Thirring model explicitly breaks the chiral symmetry algebra described in (ii), Section 4.2. This is because the mass term which is added to the action of the Dirac fermion, Eq. (26), mixes the left- and right-chiral (Weyl) fermions. It takes the following form:

$$\Delta S_{\text{mass}} = m \left( \tilde{\psi}^\dagger \psi + \psi^\dagger \tilde{\psi} \right) . \quad (37)$$

---

<sup>43</sup>Looking at the relation between the couplings (1) (Section 1.2), this will correspond to the value  $\beta^2 < \infty$  of the bosonic coupling.

Remember that the action (26) did not mix the left-chiral Weyl fermion  $\psi, \psi^\dagger$  with the right-chiral Weyl fermion  $\tilde{\psi}, \tilde{\psi}^\dagger$ . The above mass term explicitly mixes the two, and so breaks the chiral symmetry algebra.

The mass term (37) can be translated into a bosonic term using a straightforward generalisation of the dictionary (31). The generalisation is straightforward in that it takes the same form, but now depends on a bosonic coupling  $\beta$ , which is related to the fermionic coupling through (37). The equation of motion of the sine-Gordon model is:

$$\square\phi + \frac{\mu^2}{\beta} : \sin(\beta\phi) := 0 , \quad (38)$$

and so the fermionic coupling is related, through (1), to the bosonic coupling  $\beta$ . Notice that, to linear order in  $\beta$ , the above reduces to a standard bosonic mass term, with mass  $\mu$ . We are here, however, interested in the exact model.

There are divergences that need to be renormalised. The bosonic term corresponding to (37) is:

$$\Delta S_{\text{sine-G}} = \frac{\mu^2}{\beta^2} : \cos \beta\phi : , \quad (39)$$

where  $\mu$  is a scale which originates in the normal ordering procedure, and appears in the boson-fermion dictionary as an overall multiplicative constant.<sup>44</sup>

The algebra underlying the model triple of this model is still the enveloping algebra of the affine Lie algebra with  $c = 1$ , but the level now depends on the coupling:  $k = \frac{\beta^2}{4\pi} = \frac{1}{1+g/\pi}$ . We see that in the limit of zero fermionic coupling, we reproduce the algebra at level 1.

Our discussion, in Section 5.2, of the isomorphic model triples, thus generalises to the massive Thirring and sine-Gordon models, with appropriate modifications. In both cases, the algebra is the enveloping algebra, now with a coupling-dependent level. Therefore, the discussion of the *states* is analogous to the one in Section 5.2.1: the Hilbert space consists of the highest-weight representations of the enveloping algebra, and their descendants. The *quantities* are constructed from a wider class of operators, compared to our discussion in Section 5.2.2. Namely, the quantities now include non-chiral operators (respectively, operators which break affine current algebra transformations: see Section 4.1-(ii) for bosons, Section 4.2 for fermions) built from the fields: such as (38) for the bosons, and (37) for the fermions. Finally, the dynamics is still given by the Hamiltonian, which is the zero component of the Virasoro generator  $L_0$ , in the bosonic or fermionic representation. Thus, we get a theory by abstraction from these two isomorphic triples, as outlined in Section 5.3.

As we remarked before, this generalised theory explicitly illustrates our comment (b) in Section 5.4, about the contingent nature of the chiral symmetry algebra. By allowing the theory to break the chiral symmetry algebra, we get a wider class of theories (which depend on the coupling and the mass): a class of which the basic free, massless case is just a special case.

---

<sup>44</sup>The scale  $\mu$  is already present in the massless theory. But it does not play any important role, since it is just an overall renormalisation constant.



### 5.5.2 Non-abelian boson-fermion dualities

As we mentioned in this Subsection's preamble, the free, massless bosons and fermions can be generalised in another direction (Witten (1984)): to include non-abelian degrees of freedom. We will not here provide any technical details, but we will simply list some of the important examples of dualities studied in the literature; all of which are conformal field theories, except for (c) and (e):

(a)  $N$  free Majorana (real) fermions, in an  $N$ -dimensional vector representation of  $O(N)$ . They are dual to the bosonic, Wess-Zumino-Witten (WZW) model with an  $O(N)$  symmetry group. The Wess-Zumino-Witten model is a model whose action is built from a bosonic group element (rather than an algebra element), in this case an  $N \times N$  matrix of  $O(N)$ . The bosonic and fermionic models are both invariant under the affine Lie algebra transformations of  $O_L(N) \times O_R(N)$  (for left- and right-action, respectively) at level  $k = 1$ . In both cases, the central charge is  $c = N/2$ .

(b)  $N$  free, massless Dirac (complex) fermions are dual to a bosonic WZW model with group  $U(N)$ . The two models satisfy the affine Lie algebra of  $SU_L(N) \times SU_R(N) \times U(1)$ , with central charge  $c = N$  and  $k = 1$ .

(c) Mass terms can be added to the Dirac fermions in case (b), with modifications in the dictionary and in the bosonic theory which are similar to the ones discussed in the previous subsection.

(d) The Majorana and Dirac fermion models can be endowed with colour and flavour charges. Majorana fermions with  $N_F$  flavours and  $N_C$  colours, thus transforming under the group  $[O(N_F) \times O(N_C)]_L \times [O(N_F) \times O(N_C)]_R$ , are dual to the Wess-Zumino-Witten action with two bosonic fields, taking values in  $O(N_F)$  and  $O(N_C)$ . In the same way,  $N_F \times N_C$  Dirac fermions can be expressed as the sum of two Wess-Zumino-Witten actions, and a third term for an additional field. The three bosonic fields take values in the group manifolds  $SU(N_F)$ ,  $SU(N_C)$ , and  $U(1)$ . Again, one finds two copies of the affine Lie algebra, now with levels different from one, viz.  $k = N_F$  and  $k = N_C$ , respectively. This duality can also be generalised to other gauge groups.

(e) Mass terms can be added to the theories (d) with flavour and colour charges, with appropriate modifications in the dictionary and in the bosonic theory, as before.

All of the above models end up having model triple structures which are constructed as representations of the enveloping algebra of the affine Lie algebra (41) in the Appendix, for various values of the level,  $k$ , and the central charge  $c$ , and for different Lie groups. Thus their Hilbert spaces are constructed from the highest-weight representations, which as we already saw in Section 5.2.1 are unique up to unitary transformations. Clearly, all of these theories can be given a set of states, as discussed in Section 5.2.1, of which the bosonic and fermionic models form two representations. Also, all of these theories are based on the same class of algebras, with generators  $J_n$  and  $L_n$  as in (25) for the bosonic case, and likewise for the fermionic case in Section 4.2 and their anti-holomorphic counterparts (see Eq. (41) in the Appendix). Thus, each of these theories can also be given a set of quantities, in the way discussed in Section 5.2.2. For the non-chiral theories (c) and (e), this set of quantities is enlarged by the addition of non-chiral quantities, as

we already discussed in detail in the abelian case in Section 5.5.1. Finally, the dynamics of these models is as discussed in Section 5.2.3.

In conclusion, this large class of examples, based on a general enveloping algebra of an affine Lie algebra, exemplifies our schema for duality in Sections 2 and 3, just as the basic case did in Section 5.2. Namely: a duality is an isomorphism between models. More specifically: it is an isomorphism between model triples; since models also have their own specific structures. And the theory obtained for each of the dualities accords with what we said in Section 2.4: equivalence classes of isomorphic model triples give rise to a theory which is itself a triple, in which the models' specific structures have been abstracted away. And finally: our comments about non-isomorphic models (in Sections 2.4, 5.3.1, and 5.4) are illustrated by the examples (c)-(e). For these models have less symmetry: the theory which then results is more general.

## Envoi

In the physics and philosophy of physics literature, a duality in physics is agreed to be a matter of two theories being in some sense 'the same'. In this paper, we have answered the question how this can be made precise, and illustrated our answer with a case-study: bosonization.

We have proposed that a duality is best understood formally, i.e. in terms of uninterpreted theories: hence our term, 'schema'. Namely: there is a bare theory—the common core of the two given theories—which has various models, among which are the two given theories. The duality then consists in the fact that these two models are isomorphic as regards the structure and notions given in the bare theory. (Thus each of the two models also has specific structure of its own, which is unmatched by the other; and the bare theory also has, in general, other non-isomorphic models.) Often, this isomorphism is a surprising fact, since the two given theories are presented in very different terms.

We spelt out this schema in detail, in Sections 2 to 3. Among the themes we emphasized are: (i) the distinction between theories and models, (ii) the role of interpretation, (iii) the relations between a duality and the symmetries of the two given theories, and (iv) the presence of non-isomorphic models in physics.

Then in Sections 4 and 5, we illustrated the schema with bosonization. This is a matter of a quantum field theory of bosons being in some sense 'the same' as a quantum field theory of fermions. Nowadays, many such boson-fermion pairings are known. Our discussion emphasized the simplest, and earliest, case, which is known to hold exactly: a duality between a free, massless bosonic quantum field theory, and a free, massless fermionic theory, both in two spacetime dimensions. But we ended with a brief overview of other examples: involving, in particular, interacting and massive theories. (And there are extensions to higher dimensions: see e.g. Kopietz (2008); as well as an experimental interest in these systems as realising e.g. one-dimensional spin chains: Giamarchi (2003: Chapter 2), Altland and Simons (2010: Sections 4.3 and 9.4.4).)

Our schema, and this illustration of it, of course leaves plenty of work still to be done. As to physics, one should seek other illustrations of the schema: maybe some of these

will prompt revision, or at least augmentation, of the schema. As to philosophy, one should ask what light this schema casts on philosophical debates about the interpretation of physical theories, and about such theories' equivalence. But we postpone these topics to another occasion.

## Acknowledgements

We thank Joseph Kouneiher, not least for his patience! We also thank an anonymous referee for comments, the participants at the Symmetries and Asymmetries in Physics conference in Hannover, and especially Josh Hunt for comments. SDH thanks several audiences: the British Society for the Philosophy of Science 2016 annual conference, the Oxford philosophy of physics group, LSE's Sigma Club, the Munich Center for Mathematical Philosophy, and DICE2016. SDH's work was supported by the Turner scholarship in Philosophy of Science and History of Ideas, held at Trinity College, Cambridge.

## Appendix. Some Elements of Conformal Field Theory

In Section 4.1, we used the notion of a primary field. A primary field of conformal weight  $(h, \bar{h})$  is defined to transform, under a conformal transformation (16), as follows:

$$\Phi(z, \bar{z}) \rightarrow \left(\frac{\partial f}{\partial z}\right)^h \left(\frac{\partial \bar{f}}{\partial \bar{z}}\right)^{\bar{h}} \Phi(f(z), \bar{f}(\bar{z})) . \quad (40)$$

This is in analogy with the transformation law for covariant tensors in ordinary QFTs: it takes the transformation property of the field under conformal transformations as defining for the class of primary fields. The physical significance of primary fields is discussed around Eq. (20).

In our analysis in Sections 4 and 5, an essential role was played by the enveloping Virasoro algebra (25), with  $c = 1$  and  $k = 1$ . This algebra is a special case of the following general enveloping algebra of an affine Lie algebra:

$$\begin{aligned} [L_n, L_m] &= (n - m) L_{n+m} + \frac{c}{12} n (n^2 - 1) \delta_{n+m} \\ [L_n, J_m^a] &= -m J_{n+m}^a \\ [J_n^a, J_m^b] &= i f_c^{ab} J_{n+m}^c + k n \delta_{ab} \delta_{n+m} . \end{aligned} \quad (41)$$

Here,  $c$  is the central charge and  $k$  is the level, and  $f_c^{ab}$  are the structure constants of the underlying Lie algebra of the affine Lie algebra. Notice that the above algebra contains, in the first line, the ordinary Virasoro algebra. And the last line is the affine Lie algebra. The middle line gives the commutation relation between generators of the two algebras.<sup>45</sup>

<sup>45</sup>For more on affine Lie algebras, see Di Francesco et al. (1997: Chapter 14) or Kac (1990).

## References

- Ammon, M. and Erdmenger, J. (2015), *Gauge/Gravity Duality: Foundations and Applications*, Cambridge: Cambridge University Press.
- Atiyah, M. F. (2007) “Duality in Mathematics and Physics”, lecture delivered at the Institut de Matemàtica de la Universitat de Barcelona: <http://www.imub.ub.es>.
- Atland, A. and Simons, B. (2010). *Condensed Matter Field Theory*. Cambridge: Cambridge University Press.
- Barrett, T.W. and Halvorson, H. (2016). “Glymour and Quine on Theoretical Equivalence”. *Journal of Philosophical Logic*, 45(5), pp. 467-483.  
<http://philsci-archive.pitt.edu/id/eprint/11341>
- Barrett, T.W. and Halvorson, H. (2016a). “Morita Equivalence”. *The Review of Symbolic Logic*, 9(3), pp. 556-582  
<http://philsci-archive.pitt.edu/id/eprint/11511>
- Carnap, R. (1947) *Meaning and Necessity*, Chicago: University of Chicago Press
- Castellani, E. (2017). “Duality and ‘particle’ democracy”. *Studies in History and Philosophy of Modern Physics*, 59, pp. 100-108.
- Caulton, A. (2015). “ The role of symmetry in the interpretation of physical theories”. *Studies in History and Philosophy of Modern Physics*, 52, pp. 153-162.
- Coleman, S. (1975). “Quantum sine-Gordon equation as the massive Thirring model”. *Physical Review D* 11 (8), pp. 2088-2097.
- Corfield, D. (2017). “Duality as a Category-Theoretic Concept”. *Studies in History and Philosophy of Modern Physics*, 59, pp. 55-61.
- Corry, L. (1999), ‘Hilbert and Physics 1900-1915’, in Gray J., ed. (1999) *The Symbolic Universe: geometry and physics 1890-1930*, Oxford University Press: pp. 145-188.
- Corry, L. (2004), *David Hilbert and the Axiomatization of Physics*, Springer-Science.
- Corry, L. (2006), ‘On the origins of Hilbert’s sixth problem: physics and the empiricist approach to axiomatization’, *Proceedings of the ICM 2006: Madrid Spain*, pp. 1697-1718.
- Corry, L. (2018), ‘Mie’s Electromagnetic Theory of Matter and the Background to Hilbert’s Unified Foundations of Physics’, this volume.
- De Haro, S. (2016). “Spacetime and Physical Equivalence”. In *Space and Time after*

- Quantum Gravity*, Huggett, N. and Wüthrich, C. (Eds.), to appear.
- De Haro, S. (2016a). “Duality and Physical Equivalence”. <http://philsci-archive.pitt.edu/id/eprint/12279> (the title of this preprint has changed).
- De Haro, S. (2017). “Dualities and emergent gravity: Gauge/gravity duality”. *Studies in History and Philosophy of Modern Physics*, 59, 2017, pp. 109-125.
- De Haro, S. (2017a). “Invisibility of Diffeomorphisms”. *Foundations of Physics*, 47 (11), 2017, pp. 1464-1497.
- De Haro, S., Mayerson, D., Butterfield, J.N. (2016). “Conceptual Aspects of Gauge/Gravity Duality”, *Foundations of Physics*, 46 (11), pp.1381-1425. doi: 10.1007/s10701-016-0037-4.
- De Haro, S., Teh, N., Butterfield, J.N. (2017). “Comparing Dualities and Gauge Symmetries”. *Studies in History and Philosophy of Modern Physics*, 59, pp. 68-80.
- Dell’Antonio, G.F., Frishman, Y., Zwanzinger, D. (1972). “Thirring Model in Terms of Currents: Solution and Light-Cone Expansions”. *Physical Review D* 6(4), pp. 988-1007.
- Dewar, N. (2016). “Sophistication about symmetries”, forthcoming in *British Journal for the Philosophy of Science*; available at: <http://philsci-archive.pitt.edu/id/eprint/12469>
- Dieks, D., Dongen, J. van, Haro, S. de (2015), “Emergence in Holographic Scenarios for Gravity”. *Studies in History and Philosophy of Modern Physics* 52 (B), pp. 203-216. doi: 10.1016/j.shpsb.2015.07.007.
- Di Francesco, P., Mathieu, P., Sénéchal, D. (1997). *Conformal Field Theory*. New York: Springer-Verlag.
- Fraser, D. (2017). “Formal and physical equivalence in two cases in contemporary quantum physics”. *Studies in History and Philosophy of Modern Physics*, 59, pp. 30-43. doi: 10.1016/j.shpsb.2015.07.005.
- Frege, G. (1884/1950) *The Foundations of Arithmetic*, (translated by J. L. Austin), Oxford: Blackwell
- Frege, G. (1892), ‘Über Sinn und Bedeutung’, *Zeitschrift für Philosophie und philosophische Kritik*, pp. 25-50; translated as ‘On Sense and reference’, in P.T. Geach and M. Black eds. (1960), *Translations from the Philosophical Writings of Gottlob Frege*, Oxford: Blackwell.
- Frishman, Y. and Sonnenschein, J. (2010). *Non-Perturbative Field Theory*, Cambridge: Cambridge University Press.

- Giamarchi, T. (2003), *Quantum Physics in One Dimension*, Oxford: Oxford University Press.
- Ginsparg, P. (1990). “Applied Conformal Field Theory”, in *Fields, Strings, Critical Phenomena: Proceedings*. E. Brezin and J. Zinn-Justin (Eds.). Amsterdam: North-Holland. hep-th/9108028.
- Gogolin, A.O., Nersesyan, A.A. and Tsvelik, A.M. (2004). *Bosonization and Strongly Correlated Systems*, Cambridge: Cambridge University Press.
- Grattan-Guinness, I. (2000), ‘A sideways look at Hilbert’s twenty-three problems of 1900’, *Notices of the American Mathematical Society* **47**, pp/ 752-757.
- Gray J. (2000), *The Hilbert Challenge*, Oxford University Press.
- Gray, J. (2008), *Plato’s Ghost: the modernist transformation of mathematics*, Princeton University Press.
- Gray, J. (2012), *Henri Poincare, a scientific biography*, Oxford University Press.
- Huggett, N. (2017), “Target space  $\neq$  space”. *Studies in the History and Philosophy of Modern Physics*, 59, pp. 81-88. doi:10.1016/j.shpsb.2015.08.007.
- Kac, V. (1990), *Infinite-Dimensional Lie Algebras*, Cambridge: Cambridge University Press.
- Kennedy, H. (1972), ‘The origins of modern axiomatics: from Pasch to Peano’, *The American Mathematical Monthly*, **79** pp. 133-136.
- Kopietz, P. (2008), ‘Bosonization of Interacting Fermions in Arbitrary Dimensions’. Lecture Notes in Physics Monographs, Vol. 48. Springer.
- Kouneiher, J. (2018), ‘QFT, dualities, integrable systems, and supersymmetry: where we stand today’, this volume.
- Lewis, D. (1970), ‘General semantics’, *Synthese* **22**, pp. 18-67; reprinted in his *Philosophical Papers: volume 1* (1983), Oxford University Press.
- Lewis, D. (1970a), ‘How to define theoretical terms’, *Journal of Philosophy* **67**: 427-446; reprinted in his *Philosophical Papers: volume 1* (1983), Oxford University Press.
- Lewis, D. (1980), ‘Index, context and content’, in S. Kanger and S. Ohman (eds.), *Philosophy and Grammar*, D. Reidel, pp. 79-100.
- Lüst, D. and Theisen, S. (1989). *Lectures on String Theory*. Lecture Notes in Physics,

346. Berlin: Springer-Verlag.

Mandelstam, S. (1975). “Soliton Operators for the Quantized sine-Gordon Equation”. *Physical Review D* 11 (10), pp. 3026-3030.

Matsubara, K. (2013). “Realism, underdetermination and string theory dualities”. *Synthese*, 190: 471-489.

Putnam, H. (1962) ‘The Analytic and the Synthetic’, in H. Feigl and G. Maxwell (eds.) *Minnesota Studies in the Philosophy of Science* vol. III, Minneapolis: University of Minnesota Press, pp. 358-397.

Rickles, D. (2013). “Duality and the emergence of spacetime”. *Studies in History and Philosophy of Modern Physics*, 44, pp. 312-320.

Rickles, D. (2017). “Dual theories: ‘same but different’ or different but same’?” *Studies in the History and Philosophy of Modern Physics*, 59, pp. 62-67. doi: 10.1016/j.shpsb.2015.09.005.

Smolin, L. (2018), ‘What are we missing in our search for quantum gravity’, this volume.

Stachel, J. (2018), ‘Einstein and Hilbert’, this volume.

Stöltzner, M. (2000), ‘Opportunistic Axiomatics Von Neumann on the Methodology of Mathematical Physics’, in *John Von Neumann and the Foundations of Quantum Physics*, M. Redei and M. Stöltzner (eds.): Springer: Volume 8 of the series Vienna Circle Institute Yearbook (2001) pp. 35-62.

Stöltzner, M. (2001), ‘How Metaphysical is “Deepening the Foundation”? Hahn and Frank on Hilbert’s Axiomatic Method’, in *History of Philosophy of Science*, M. Heidelberger and F. Stadler (eds.): Springer: Volume 9 of the series Vienna Circle Institute Yearbook (2001) pp. 245-262.

Teh, N.J. (2013). “Holography and emergence.” *Studies in History and Philosophy of Modern Physics*, 44(3), pp. 300-311.

Teh, N. (2016). “Gravity and Gauge”. *British Journal for the Philosophy of Science*, 67(2), pp. 497-530.

von Neumann, J. (1932), *Mathematical Foundations of Quantum Mechanics*: English translation published in 1955 by Princeton University Press.

Weatherall, J.O. (2015). “Categories and the Foundations of Classical Field Theories”. Forthcoming in *Categories for the Working Philosopher*, E. Landry (Ed). Oxford: Oxford University Press.

Witten, E. (1984). “Nonabelian bosonization in two dimensions”. *Communications in Mathematical Physics*, 92, pp. 455.

Witten, E. (2018), “What Every Physicist Should Know About String Theory”, this volume.