Duality and Emergence

Sebastian De Haro

University of Cambridge and University of Amsterdam

Based on joint work with: J.N. Butterfield, D. Dieks, J. van Dongen

Quantum Gravity: Physics and Philosophy

IHES, Paris, 25 October 2017

Introduction: Duality and Emergence

Dualities have become standard tools for theory construction in theoretical physics (see the talk by C. Bachas).

Dualities also connect well with two philosophical topics: **theoretical equivalence** and **physical equivalence**.

In the physics literature, duality also often appears related to **emergence**. When a weakly-coupled theory (GR in AdS, say) is dual to a strongly coupled theory (a CFT), it is often concluded that the weakly-coupled theory 'emerges' from the CFT.

In particular, it is often claimed that spacetime (and-or gravity) emerge.

This is a very interesting claim which deserves scrutiny.

Introduction: Duality and Emergence

Duality and emergence are two notions which seem to be closely connected, but are also in tension.

Duality: theoretical/formal equivalence between two theories.

Emergence: focus on novelty, hence on the *lack of equivalence* between two theories or phenomena.

Conceptual Questions (a philosopher's wish-list)

(1) **Construal:** how to best construe dualities?
How does duality relate to *theoretical equivalence*?
Can string-theoretic dualities be treated along with other cases of duality, such as *position-momentum* duality or *electric-magnetic* duality?

(2) **Illustration:** are string theory dualities exact (i.e. valid non-perturbatively)? Are there examples of exact dualities which illustrate the construal (1)?

(3) **Physicality:** under what conditions do dualities amount to cases of *physical equivalence*?

(4) **Emergence:** what is emergence? What is the relation between duality and emergence? Does spacetime emerge?

Aim of the Talk

Aim of the talk: to give a conceptual account of dualities and to relate it to the discussion of emergence.

(De Haro (2016, 2017), De Haro and Butterfield (2017), Dieks et al. (2015).)

Some (interesting!) things I will **not** do:

- Discuss the questions of physical (in)coherence, experiential spacetime, etc. (see the talks by D. Dieks and Y. Dolev).

- Discuss *emergence* in gauge-gravity dualities (see talk by G. Horowitz).

Outline

Part I. Duality

Duality and its related notions

We need working conceptions of: theory and interpretation.

Bare theory: a triple $T := \langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$, comprising a structured **state space**, $\mathcal{S}$, a structured set of **quantities**, $\mathcal{Q}$, and a **dynamics**, $\mathcal{D}$: together with a set of **rules for evaluating** quantities on the states. There are also **symmetries**, which are automorphisms $a : \mathcal{S} \to \mathcal{S}$ of the set of states (or as the dual maps on the set $\mathcal{Q}$ of quantities).

Interpreted theory: it adds, to a bare theory, an *interpretation*: i.e. a set of *structure-preserving partial maps* from the theory to the world. The interpretation fixes the reference of the terms in the theory. More precisely, an intepretation maps the theory, T, to a domain of application, D_W, within a possible world, W, i.e. it maps $I : T \to D_W$, preserving structure. Interpretations can be further restricted by imposing suitable conditions on the kinds of maps admissible (as I will do for emergence).

We of course normally work with interpreted theories. So, by a 'theory', I will mean an *interpreted theory*.

Our usage

Model: a homomorphic copy of a theory (i.e. representation, in the sense of representation theory).

A bare theory can be realized (I will say: modelled) in various ways: like the different representations of a group or algebra.

These models are in general *not* isomorphic, and 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*.

Our usage

Beware: the word 'model', as contrasted with 'theory', often connotes:

(i): a specific solution for the physical system concerned, whereas the 'theory' encompasses all solutions—and in many cases, for a whole class of systems;

(ii): an approximation, whereas the 'theory' deals with exact solutions;

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

Our use of 'model' rejects all three connotations.

Duality as surprising

We usually discover a duality in the context of studying, not a bare theory, but rather: two interpreted models of a bare theory.

For example, type IIB supergravity on $\text{AdS}_5 \times S^5$
and $\mathcal{N} = 4$, U(N) super-Yang-Mills.

Usually, we do not initially believe them to be isomorphic in any relevant sense. Or even: to be models of any single relevant theory (even of a bare one).

The **surprise** is to discover that they are such models—indeed are isomorphic ones. And the surprise is greater, the more detailed is the common structure.

Notation for theories and models

A notation for a model M that exhibits how M augments the structure of the theory T with specific structure, $\bar{M}$ say, of its own:

$$M = \langle T_M, \bar{M} \rangle \ . \tag{1}$$

The subscript on T_M reflects that the specific structure $\bar{M}$ is used to build the representation of T.

We call T_M, the 'part' of M that represents T, the **model root**.

Thus for a theory as a triple, $T = \langle \mathcal{S}, \mathcal{Q}, \mathcal{D} \rangle$: we write a model as a quadruple:

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

where $m := T_M := \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M \rangle$ is called the **model triple** (root).

The conception of duality

We propose that *a duality is an isomorphism between two model roots* (model triples). Recall that $M = \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M, \bar{M}\rangle =: \langle m, \bar{M}\rangle$, where $m := \langle \mathcal{S}_M, \mathcal{Q}_M, \mathcal{D}_M\rangle$ is the model root (model triple).

A **duality** between $M_1 = \langle \mathcal{S}_{M_1}, \mathcal{Q}_{M_1}, \mathcal{D}_{M_1}; \bar{M}_1\rangle$ and $M_2 = \langle \mathcal{S}_{M_2}, \mathcal{Q}_{M_2}, \mathcal{D}_{M_2}; \bar{M}_2\rangle$ requires isomorphism of the model roots: an isomorphism between (usually) Hilbert spaces:

$$d_{\mathcal{S}} : \mathcal{S}_{M_1} \to \mathcal{S}_{M_2} \;\text{ using } d \text{ for 'duality' ;} \tag{3}$$

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

$$d_{\mathcal{Q}} : \mathcal{Q}_{M_1} \to \mathcal{Q}_{M_2} \;\text{ using } d \text{ for 'duality' ;} \tag{4}$$

I will call the common model root of the models, i.e. the models defined up to isomorphism, the **common core** (if there are no non-isomorphic models to be considered, the common core will also be the *theory*).

(i) the values of all quantities match:

$$\langle Q_1, s_1\rangle_1 = \langle d_{\mathcal{Q}}(Q_1), d_{\mathcal{S}}(s_1)\rangle_2 \ , \ \ \forall \ Q_1 \in \mathcal{Q}_{M_1}, s_1 \in \mathcal{S}_{M_1} . \qquad (5)$$

(and the same for maps involving more arguments, e.g. $\mathcal{S} \times \mathcal{S} \to \mathbb{C}$)

(ii) $d_{\mathcal{S}}$ is equivariant for the two triples' dynamics, $D_{S:1}, D_{S:2}$, in the Schrödinger picture; and $d_{\mathcal{Q}}$ is equivariant for the two triples' dynamics, $D_{H:1}, D_{H:2}$, in the Heisenberg picture:

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

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

A logically weak, but physically strong notion

Our notion of duality is logically weak:
there is duality whenever two model roots 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.

But this worry is blocked by the requirement that duality must map *model roots* (triples).
So, an isomorphism relative to a structure smaller than the model root **does not count as a duality**. Not any old isomorphism counts as duality!
(The choice of a model root is a physical choice: roughly, as in 'the uninterpreted degrees of freedom the model is able to describe'.)

Thus duality is an *isomorphism between uninterpreted, but physical, models.*

—A logically weak, but physically strong notion

Example: consider two models with state spaces:

$$\begin{aligned} M_0 &:= \langle \mathbf{J}, \cdot\,, \cdot\,; \cdot\,\rangle \\ M &:= \langle \mathbf{J}\otimes\mathbf{K}, \cdot\,, \cdot\,; \cdot\,\rangle\,, \end{aligned} \tag{6}$$

(where $\mathbf{J}$ = vector space acted upon by the irrep of su(2) with spin j.)

M_0 and M are both representations of su(2) and they share $\mathbf{J}$ as a common structure. Yet they are *not dual* to each other, because *their state spaces are not isomorphic*.

In order for M_0 and M to be dual, $\mathbf{K}$ would have to be part of M's specific structure. So, define:

$$M' := \langle \mathbf{J}, \cdot\,, \cdot\,; \mathbf{K}, \cdot\,\rangle \tag{7}$$

Now M_0 and M' *are* dual, because their state spaces are the same, regardless their different specific structures. But still $M \neq M'$, despite their sharing the same 'data': because their state spaces are different. They are *theoretically inequivalent.*

A logically weak, but physically strong notion

Two ways to construct a theory out of its models:

(1) **Reconstruction theorems** (cf. Majid (1991)): e.g. an Abelian group can be reconstructed from its set of representations (again an Abelian group), using Pontryagin's duality theorem.

Incidentally, Pontryagin duality, taken on face value, is *not* a duality, in our sense (for it is not an isomorphism of the groups).
But one can construct a suitable model root, of which it *is* a case of self-duality (with a *single* isomorphism, rather than a triple of them).

(2) **'Taking the union and modding out':** reconstruct a theory from a set of dual models (even if a only subset of all models) by 'abstracting from the specific structure of the models', i.e. by identifying elements of the models which are related by isomorphism.

I will illustrate roughly this procedure, of "**finding a common core**", in gauge-gravity duality.

The analogy with symmetry

People say: *a duality is like a symmetry, but at the level of a theory*. That 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 endorse this analogy. The interesting questions, for both sides, concern the different ways to be 'the same' or 'matching'.

A symmetry a (I write a for 'automorphism') carries a state s in a state space $\mathcal{S}$ to another state $a(s)$, where s and $a(s)$ assign the same values to all the quantities in some salient, usually large, set of quantities.

Hence the question: under what conditions, do s and $a(s)$ represent the **very same physical state of affairs**?

Similar issues arise for dualities:

Duality and Interpretation

I have suggested that duality is a special case of **theoretical equivalence**, i.e. for the most part a formal equivalence—an isomorphism between model roots, in our case: model triples.

The question of **physical equivalence** arises when we consider *interpreted theories or models*.

External and Internal Interpretations

To answer questions about physical equivalence, we will need to further distinguish two kinds of interpretations:

(1) **External interpretation:** maps as above, from the *model*, $M = (m_T, \bar{M})$, to the world. There are two basic options:

(1a) **The specific structure gets interpreted:** i.e. the maps map *both* the model root *and* the specific structure,i.e. $I_{\text{ext}} : M \to D_W$.

(1b) The specific structure does *not* get interpreted, but the interpretation **couples the model to another, already interpreted** theory, e.g. a theory of measurement, i.e. $I_{\text{ext}} : m_T \times T_{\text{meas}} \to D_W$.

(2) **Internal interpretation:** the maps map from the theory, and nothing else (equivalently, they map from the model root, neglecting the specific structure), i.e. $I_{\text{int}} : T \to D_W$.

Physical (In-)Equivalence: Two Interpretative Cases

Since the conception of duality is formal, it certainly allows the idea of '**distinct but isomorphic sectors of reality**'—namely as the codomains of the interpretation maps on the two sides of the duality. On such an (external) interpretation, the models are then **physically inequivalent**.

But this of course does not forbid the other sort of case: where the two models are **physically equivalent**, i.e. do describe '**the same sector of reality**'. This is modelled by the (internal) interpretation maps having the same images in their codomain:

$$\begin{array}{ccc} \mathcal{Q}_{M_1} & \xrightarrow{d_q} & \mathcal{Q}_{M_2} \\ I^1_{\text{Int}} \searrow & & \swarrow I^2_{\text{Int}} \\ & \mathcal{Q}_{\text{Int}} & \end{array}$$

Figure: The two sides of the duality describe 'the same sector of reality'.

When does physical equivalence obtain?

Thus we get the following **criterion for physical equivalence**:

$$I^1_{\mathrm{Int}} = I^2_{\mathrm{Int}} \circ d \ . \tag{8}$$

I will say that '**interpretation commutes with duality**'
(in the sense that the corresponding diagram commutes).

This analysis simply formalises what we mean by the phrase 'physical equivalence', as sameness of reference of two models.

But are there significant cases, in physics, in which such physical equivalence in fact obtains?

Interpretations like the ones above ('mapping the same sector of reality') can be constructed, given two conditions:

Two conditions for physical equivalence

(1) **Internal interpretation:** the interpretation maps the model root (the theory) but not the specific structure of the model. The idea here is that 'no other facts, external to the model root, should determine the codomain of the interpretation', i.e. 'the interpretation should start from the model roots, and nothing else'.

(2) **Unextendability:** we should be considering 'models of the whole world'. To claim the sameness of the two sectors of reality described by the theory, we need to ensure that 'there is no more for the two models to describe', so that 'extending the theory beyond its given domain of application' will not be able to distinguish between the two models (cf. Leibniz's principle of the identity of indiscernibles).

Interpretation: a quick example

AdS/CFT, if exact, contains sufficiently rich models, whose common core supports an internal interpretation.

And these are 'models of the whole world', not extendable beyond their domain of application: and so, on this internal interpretation, the two models are physically equivalent.

On the other hand, the application of gauge-gravity duality to a **quark-gluon plasma** (described as a five-dimensional black hole) is only approximate, and does not map the whole world: and so, there is no physical equivalence in that case.

I now discuss the **duality** in more detail
(and independently of matters of interpretation):

Gauge-gravity duality

The schema can be illustrated in **gauge-gravity dualities**: they relate $(d+1)$-dimensional string theories (models) to d-dimensional quantum field theories (QFT models).

We **do not have an exact (or non-perturbative) definition** of the models, or of the duality. Having a rigorous definition of it would almost be like 'proving' the duality.

But we can get important insights about the theory, and thus illustrate the schema, by considering the **semi-classical limit**
(i.e. strong 't Hooft coupling, cf. Bachas' talk).

Suppose two such models are theoretically equivalent. Are they also dual, in the sense just discussed? What is their **common core**?

Gauge-gravity duality (pure gravity: no matter)

The gravity theory is defined under two **boundary conditions** for the metric and the stress-energy tensor, defined at spacelike infinity:

(i) A boundary condition for the **metric** at infinity, which is defined up to conformal transformations.

We have a d-dimensional conformal manifold, $\mathcal{M}$, at infinity, with a conformal class of metrics, $[g]$. This is identified with the manifold on which the CFT, with its conformal class of metrics, is defined.

Thus the pair $(\mathcal{M}, [g])$ is part of the **common core.**

Gauge-gravity duality

The asymptotic symmetry algebra associated with the gravity model is the d-dimensional **conformal algebra**, and the representations of this algebra form the set of admissible states belonging to the Hilbert space, $|s\rangle_{\mathcal{M},[g]} \in \mathcal{H}$.

(ii) A boundary condition is also required for the asymptotic value of the **canonical momentum**, Π_g, conjugate to the metric on the boundary, evaluated on all the states, $\langle s|\Pi_g|s\rangle$. This choice further constrains the states in $\mathcal{H}$: it determines a subset of states of the conformal algebra.

The simplest (and usual) choice, $\langle s|\Pi_g|s\rangle = 0$, preserves the full conformal symmetry.

Gauge-gravity duality

In the case of interest (pure matter on the gravity side), the only operator turned on in the CFT is the **stress-energy tensor**, T_{ij}.

The duality dictionary tells us that: $\Pi_g \equiv T_{ij}$.

Thus, the two models share the d-dimensional conformal manifold $\mathcal{M}$ with its conformal class of metrics $[g]$, the conformal algebra, and the structure of operators.

These determine the values of the infinite set of **correlation functions**:

$$ {}_{\mathcal{M},[g]}\langle s|\ T_{i_1 j_1}(x_1)\cdots T_{i_n j_n}(x_n)\ |s\rangle_{\mathcal{M},[g]}\ , \tag{9}$$

which agree between the two models (cf. C. Bachas' talk), with the approximations made (and for a subset of states).

This discussion can be generalised to include other states and operators.

An Exact Example: Bosonization

The Bosonic Model

The basic case: The duality, in two (Euclidean) dimensions, between:
(1): the free, massless bosonic scalar field; and
(2): the free, massless Dirac fermion.

Use complex coordinates, $z = x^0 + ix^1$, $\bar{z} = x^0 - ix^1$, parametrising $\mathbb{C} \cong \mathbb{R}^2$. Then the massless Klein-Gordon equation takes the form:

$$\partial\bar{\partial}\Phi = 0 \tag{10}$$

(with $\partial := \partial/\partial z$, $\bar{\partial} := \partial/\partial\bar{z}$). The general classical solution is the sum of a holomorphic and an anti-holomorphic function:

$$\Phi(z,\bar{z}) = \phi(z) + \bar{\phi}(\bar{z}) \ . \tag{11}$$

Symmetries of the equations of motion (the action): (i) **conformal** transformations; (ii) **affine current algebra** transformations.

(i): In two dimensions, the conformal group is *infinite*-dimensional. It is parametrised by arbitrary holomorphic and anti-holomorphic functions:

$$z \quad \rightarrow \quad z' = f(z) \ , \qquad \bar{z} \rightarrow \bar{z}' = \bar{f}(\bar{z}) \ . \tag{12}$$

The corresponding algebra will yield, under quantization, the Virasoro algebra.

(ii): These are translations of the field by holomorphic or anti-holomorphic functions,

$$\Phi(z,\bar{z}) \quad \rightarrow \quad \Phi(z,\bar{z}) + \varphi(z) \ , \qquad \Phi(z,\bar{z}) \rightarrow \Phi(z,\bar{z}) + \bar{\varphi}(\bar{z}) \ . \tag{13}$$

These transformations generalise the invariance of the action under constant shifts $\Phi \rightarrow \Phi + \varphi_0$, and are specific to two dimensions. The corresponding algebra will yield, under quantization, an affine Lie (aka: Kac-Moody) algebra.

The conserved currents for the affine **current algebra transformations**:

$$J(z) := \partial\phi(z) \ , \quad \bar{J}(\bar{z}) := \bar{\partial}\bar{\phi}(\bar{z}) \ . \tag{14}$$

The conserved currents for the **conformal transformations** are the (holomorphic/anti-holomorphic) components of the stress-energy tensor:

$$T(z) = -\frac{1}{2}\,\partial\phi\,\partial\phi = -\frac{1}{2}\,J^2(z) \ , \quad \bar{T}(\bar{z}) = -\frac{1}{2}\,\bar{\partial}\phi\,\bar{\partial}\bar{\phi} = -\frac{1}{2}\,\bar{J}^2(\bar{z}) \ . \tag{15}$$

Under quantization: (i) One defines (15) by *normal ordering*:

$$T(z) = -\frac{1}{2}\ :J(z)\,J(z):\ , \quad \bar{T}(\bar{z}) = -\frac{1}{2}\ :\bar{J}(\bar{z})\,\bar{J}(\bar{z}):\ , \tag{16}$$

with the affine currents still given by (14), but now as operator equations.

(ii) Write the holomorphic and anti-holomorphic parts of the field as Laurent series. The affine currents and the stress-energy tensor:

$$\begin{aligned} J(z) &= \sum_{n\in\mathbb{Z}} \frac{J_n}{z^{n+1}} \,, \qquad \bar{J}(\bar{z}) = \sum_{n\in\mathbb{Z}} \frac{\bar{J}_n}{z^{n+1}} \\ T(z) &= \sum_{n\in\mathbb{Z}} \frac{L_n}{z^{n+2}} \,, \qquad \bar{T}(\bar{z}) = \sum_{n\in\mathbb{Z}} \frac{\bar{L}_n}{\bar{z}^{n+2}} \,. \end{aligned} \tag{17}$$

The J_n are linear in the creation and annihilation operators of the holomorphic field ϕ, and the L_n are quadratic in the J_n's.

Finally, we get the **algebras** satisfied by J_n and L_n. The result is:

$$\begin{aligned}
[L_m, L_n] &= (m-n)\, L_{m+n} + \frac{c}{12}\, n(n^2-1)\, \delta_{(m+n)0} \\
[J_m, J_n] &= -m\, \delta_{(m+n)0} \\
[L_m, J_n] &= -n\, J_{(m+n)0}\ .
\end{aligned} \tag{18}$$

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, the central charge $c = 1$.

The first line is the **Virasoro algebra**.

The second line is the (level $k = 1$, abelian) **Kac-Moody algebra**.

The third line comes from the fact that the L_m's are quadratic in the J_n's.

The algebra (18) is called the **enveloping Virasoro algebra** (with $c = 1$ and $k = 1$).

This is the central result we need. For, together with (16) and the mode expansions (17), the tensor product of the holomorphic enveloping Virasoro algebra of the affine Lie algebra (18) and its anti-holomorphic copy contains all of the information about the quantum model:

For the **state-spaces** are now the vector spaces on which this algebra acts, endowed with an appropriate semi-positive norm, and **all the quantities**, and the **dynamics**, are constructed from the operators satisfying the algebra.

The fermionic model

The massless Dirac fermion can be decomposed as $\Psi =: (\psi, \tilde{\psi})$, where ψ and $\tilde{\psi}$ are chiral (Weyl) fermions, called left- and right-chiral, respectively.

We 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.

The action takes the form of the sum of two copies of a single Weyl-Majorana fermion, conventionally called χ (so $\chi = \psi_{1,2}$). For a single **Weyl-Majorana fermion**:

$$S_{\rm WM} = \frac{1}{8\pi}\int \mathrm{d}^2 z\,\left(\chi\,\bar{\partial}\chi + \tilde{\chi}\,\partial\tilde{\chi}\right). \tag{19}$$

The action is invariant under two sets of **symmetries**:
(i) **Conformal transformations:** $z \to f(z)$, $\bar{z} \to \bar{f}(\bar{z})$, the same transformations on the complex plane as in the bosonic model;
(ii) **Left-holomorphic-chiral and right-anti-holomorphic-chiral transformations:**

$$\begin{aligned} \psi \to \psi' &= e^{i\,\alpha(z)}\,\psi \\ \tilde{\psi} \to \tilde{\psi}' &= e^{i\,\tilde{\alpha}(\bar{z})}\,\tilde{\psi}\ . \end{aligned} \tag{20}$$

The definition of conserved currents, and quantization, proceeds similarly to the bosonic model. We have holomorphically and anti-holomorphically conserved currents associated with the symmetries (i) and (ii):

$$\begin{aligned} J(z) &= \ :\psi^\dagger\psi:\ , \qquad \bar{J} =: \tilde{\psi}^\dagger\,\tilde{\psi}: \\ T(z) &= -\frac{1}{2} : J(z)\,J(z) := -\frac{1}{2} : \left(\psi^\dagger\,\partial\psi - \partial\psi^\dagger\,\psi\right) : \ , \end{aligned} \tag{21}$$

with a similar expression for $\bar{T}$ in terms of $\tilde{\psi}$.

The resulting algebra is the very same as in the bosonic case, i.e. the **enveloping Virasoro algebra**—Virasoro algebra with $c = 1$ with abelian affine algebra at level $k = 1$ (and its anti-holomorphic copy)!

Therefore, also the **state-space** is the same, since it is constructed as the vector space of semi-positive norm, on which the algebra acts.

Notice that: (A) the irreducible unitary (highest-weight) $k = 1$ representations of the affine Lie algebra are unique up to unitary equivalence—and
(B) there is a corresponding unitary equivalence for the general enveloping Virasoro algebra, i.e. the affine Lie algebra with the Virasoro algebra (Eq. (18)), given the same central charge, level, and underlying Lie algebra.

This **correspondence between the algebras** is the basis for the duality map, d.

The dictionary

The **duality map** d pairs: (a) the affine current algebra currents, and (b) the stress-energy tensors. So it relates the currents of the two models. We add subscripts 'B' and 'F' for 'bosonic' and 'fermionic':

$$J_{\mathrm{B}}(z) = \partial\phi(z) \quad \leftrightarrow \quad J_{\mathrm{F}}(z) = :\psi^\dagger(z)\,\psi(z):$$
$$T_{\mathrm{B}}(z) = -\frac{1}{2}\,:\partial\phi(z)\,\partial\phi(z): \quad \leftrightarrow \quad T_{\mathrm{F}}(z) = -\frac{1}{2}\,:\left(\psi^\dagger(z)\,\partial\psi(z) - \partial\psi^\dagger(z)\,\psi(z)\right):$$

and similarly for the anti-holomorphic currents. The two sides satisfy the **same algebra**.

That only fermion bilinears appear is expected. But surprisingly, the bosonic field to which a single fermion corresponds is—not some kind of 'square root' of the boson—but an exponential:

$$\begin{aligned} :e^{i\,\phi(z)}: \quad &\leftrightarrow \quad \psi(z)\,, \qquad :e^{-i\phi(z)}: \;\leftrightarrow\; \psi^\dagger(z) \\ :e^{-i\bar\phi(\bar z)}: \quad &\leftrightarrow \quad \tilde\psi(\bar z)\,, \qquad :e^{i\,\bar\phi(\bar z)}: \;\leftrightarrow\; \tilde\psi^\dagger(\bar z)\,. \end{aligned} \tag{22}$$

Two isomorphic model triples

Not any old isomorphism will do. We need **isomorphism of triples**!

But we almost already have that. We complete the picture as follows:

1. **Quantities:** There is an isomorphism betwen all the quantities, because these are constructed from the currents, which satisfy the same algebra.
2. **States:** This isomorphism also induces an isomorphism between the states, because these are representations of the algebra (vector spaces on which the algebra elements act).
3. **Dynamics:** The Hamiltonians are the same as well (same choice, viz. the 00-component of the stress-energy tensor: but of course the whole stress-energy tensor matches).

Therefore, also all the numeric values of all the correlation functions, $\langle s|Q|s'\rangle$, agree between the bosons and the fermions.

—Bosonization respects symmetries

The theory has two built-in **symmetries**: (i) the conformal group, whose generator on the fields is the stress-energy tensor, (ii) an affine current symmetry algebra, generated by the J-currents.

(i) The **conformal group** is the symmetry group of the background spacetime, and it is represented in the same way in the two models, viz. by the stress-energy tensor, for which there is an isomorphism.

(ii) The **affine current symmetry algebra** is represented very differently in the two models! As translations of the field in the bosonic model $(\Phi(z,\bar{z}) \to \Phi(z,\bar{z}) + \varphi(z))$, and as left- and right-chiral symmetry transformations $(\psi \to e^{i\alpha(z)}\,\psi)$ in the fermionic model. Yet their corresponding currents (the J-currents) satisfy the same algebra, and so these symmetries act in the same way on the states and the quantities.

The symmetries were used in restricting the class of admissible quantities, and they restrict this class in the same way in the two models.

Generalisations of bosonization

- Thirring model-sine Gordon duality (massive fermion-boson field with potential cos($\beta\phi$)). In this case, the exponential map mapping the boson to the fermion is appropriately generalised to include the parameter β and the mass.
- Bosonization with non-abelian symmetries: between fermions with non-abelian gauge groups and WZW models.

Part II. Emergence

Emergence: Motivation

There is a large recent philosophical and physical literature on emergence. But the notion of emergence remains vague, which limits its applicability.

Butterfield (2011) has clarified aspects of emergence: its independence from related notions like reduction and supervenience. Yet the notions involved (**novelty** and **robustness**) are informal, even vague.

Another source of confusion is the lack of a good account of the distinction between **ontological** and **epistemic** emergence. Though philosophers already have an understanding of this distinction, we are rarely told what the distinction consists in, and how it plays out in specific examples.

The consensus in the philosophy of physics literature seems to be that '**emergence and reduction are independent**'.
But just how this can be has not been made sufficiently clear.
The tension is similar to that existing between emergence and duality.

Emergence: Motivation

The examples of **emergence of space** in the literature are usually not very explicit.

A good **conceptual framework for emergence** could be of importance for the analysis of concrete examples of emergence in physics, and in particular for the development of theories that do not have an obvious spacetime interpretation at the fundamental level.

One would wish to analyse emergence as a property of an **interpreted theory**, rather than as a technical property of the bare theory.

Such an account should help clarify emergence of spacetime in the construction of new theories (and new observations!).

Emergence

My starting point is as follows: the working conceptions of **theory and of interpretation**, introduced for duality, can also be used to describe emergence.

The analysis of emergence will be analogous to that of duality, where we replace 'duality' by a different map.

This map should capture the idea of 'coarse-graining a theory to describe a novel and-or specific physical situation or system'.

This will give us a "mechanism" for emergence, of a specific (and pervasive) kind .

(Cf. also: Butterfield (2011), Fletcher (2016), Landsman (2013).)

Approximative Emergence

I will advocate a view in which emergence is not in the bare, but in the *interpreted theory*.

But I will restrict to bare theories with the following property:

Approximation: an asymmetric relation, Approx, among the theories in a given family.

More precisely, there is a surjective and non-injective **approximative map**, denoted by:

$$\text{Approx} : T_{\text{b}} \to T_{\text{t}} \quad \text{(basic vs. top theory).} \tag{23}$$

The map approximates T_{b} by the approximating theory, T_{t}.

Approximative Emergence

The approximative map Approx can be (a combination of) one of the following:

(i) **A mathematical limit:** some variable of the theory is taken to a special value, e.g. $\hbar \to 0$ or $N \to \infty$. In that case, $T_{\text{t}} = \lim_{\hbar, 1/N \to 0} T_{\text{b}}$.

(ii) **Comparing different physical situations or systems:** one may approximate one physical situation, or system, by another that resembles it. For instance, one compares situations with different values of the action (in addition to sending $\hbar \to 0$).

(iii) **Mathematical approximations:** one compares theories mathematically: perhaps numerically, or in terms of some parameters of approximation. For example, as when in a Taylor expansion one drops the terms after a certain order of interest.

Approximative Emergence

Approximative emergence: if two theories, T_b and T_t, are related by an approximative map and if, *in addition*, T_t emerges from T_b, then we have approximative emergence. In other words, *the map itself does not give us emergence* (recall: we are not seeking to describe emergence as a formal relation between two bare theories). The consensus view:

Emergence, informally (Butterfield, 2011): 'I shall take emergence to mean: properties or behaviour of a system which are **novel** and **robust** relative to some appropriate comparison class. Here 'novel' means something like: 'not definable from the comparison class', and maybe showing features (maybe striking ones) absent from the comparison class'. And 'robust' means something like: 'the same for various choices of, or assumptions about, the comparison class'... I shall also put the idea in terms of theories, rather than systems: a theory describes properties or behaviour which are novel and robust relative to what is described by some other theory with which it is appropriate to compare.'

Emergence as novel reference

Norton (2012) is a perspicacious discussion of the difference between idealisation and approximation. He summarises the main difference between the two as an answer to the question:

'Do the terms involve novel reference?'

According to Norton, only idealisations introduce reference to a **novel system**—where 'system' here may refer to a real or a fictitious system, a mathematical object, etc. In the case of *idealization*, there is a 'limit system' that realizes the 'limit properties'. In the case of *approximation*, the limit system either does not exist, or is not accurate enough to describe the system under study.

(There is an important distinction in the way one can use 'reference' here for the interpretative map: as an **intension** or as an **extension**. But I will not dwell on these details.)

Emergence, of what kind?

(i) **Epistemic emergence:** if the two theories, $T_{\rm b}$ and $T_{\rm t}$, describe the same 'sector of reality', i.e. the same world and domain of application, then the codomains of their interpretations must coincide:

$$\begin{array}{ccc} T_{\rm b} & \xrightarrow{\text{Approx}} & T_{\rm t} \\ I_{\rm b} \searrow & & \swarrow I_{\rm t} \\ & D & \end{array}$$

Figure: The two interpretations describe 'the same sector of reality', so that $I_{\rm b} = I_{\rm t} \circ \text{Approx}$.

If, in addition to $I_{\rm b} = I_{\rm t} \circ \text{Approx}$, there is emergence, we say this is **epistemic emergence**. The emergence is at the level of the *bare theory*, not in the interpretation, and so it is only epistemic: it reflects *novelty of description*.
(I will not analyse further what the conditions are for novelty and robustness in epistemic emergence.)

Emergence as novel reference: second interpretative case

(ii) **Ontological emergence:** there is novel reference (cf. Norton (2012)), i.e. the two interpretations refer to different domains of application (even different worlds). In this case, $I_{\mathrm{b}} \neq I_{\mathrm{t}} \circ \text{Approx}$, and:

$$\begin{array}{ccc} T_{\mathrm{b}} & \xrightarrow{I_{\mathrm{b}}} & D \\ \text{Approx} \downarrow & & \not\updownarrow \\ T_{\mathrm{t}} & \xrightarrow{I_{\mathrm{t}}} & D' \end{array}$$

Figure: The failure of interpretation and approximation to commute ($I_{\mathrm{b}} \neq I_{\mathrm{t}} \circ \text{Approx}$) gives rise to different interpretations, possibly with different domains of application D, D'.

In other words, **interpretation and approximation fail to commute.**

The Case Study: Random Matrix Models

Motivation from quantum gravity

Random matrix models are examples of purely algebraic structures (with very minimal geometrical structure) within which a **two-dimensional Riemann surface** emerges, in the 't Hooft limit (N large but $g^2 N$ fixed).

RMM illustrate the framework for emergence presented here, they are mathematically precise, and so they can shed light on what we mean by **emergence of space out of non-spatial structures**.

(The example only addresses emergence of space not emergence of time or gravity.)

RMM are well-known in **QFT**: motivated by 't Hooft's large-N of QCD, Brézin, Itzykson, Parisi and Zuber (1978) used them to model QFT.

They were also applied in **two-dimensional quantum gravity**.
(Cf. Di Francesco (1995).)

Random Matrix Models: quantum gravity motivation

- Dijkgraaf and Vafa (2002): RMM calculate the effective superpotential of a large class of **four-dimensional gauge theories** with minimal supersymmetry. Non-perturbative results derived from this perspective: e.g. the Seiberg-Witten solution of $\mathcal{N} = 2$ theories.

- RMM can be embedded in string theory: the **open topological string theory**, on a class of six-dimensional Calabi-Yau manifolds, reduces to a RMM. In the large-N limit, the RMM describes the **closed topological string theory** on a related, but different, class of Calabi-Yau manifolds (a theory of topological gravity!). So, RMM realise simple versions of **open-closed string duality**.

- They form some of the original motivation for group field theory.

- But here we will study the RMM on its own, stripped from its field-theoretic or quantum gravity interpretations.

Random Matrix Models

Defined by three objects: an $N \times N$ matrix Φ (usually: self-adjoint), an integration measure w on the set of matrices, and a polynomial of degree $n+1$ in Φ (the potential). The basic quantity is the **partition function:**

$$\mathcal{Z} = \int_{\Gamma_N} w \ \exp\left(-\frac{1}{g^2}\,\mathrm{Tr}\,W(\Phi)\right) \ , \qquad W(x) = \sum_{\alpha=0}^{n+1} g_\alpha\, x^\alpha \ . \qquad (24)$$

The partition function is invariant under a symmetry group G_N (usually, $\mathrm{U}(N)$), which acts as: $\Phi \mapsto S\,\Phi\,S^{-1}$. So, the integral reduces to an integral over eigenvalues (Brézin, Itzykson, Parisi, Zuber (1978)):

$$\begin{aligned} Z_N &= \int_{\gamma\times\cdots\times\gamma} \left(\prod_{i=1}^{N} \mathrm{d}\lambda_i\right) e^{-N^2 S_N} \\ S_N &= \frac{1}{g^2 N^2}\sum_{j=1}^{N} W(\lambda_j) - \frac{2}{N^2}\sum_{i<j}\log(\lambda_i-\lambda_j) \end{aligned} \qquad (25)$$

The Saddle Point Approximation

The integral can be evaluated in the **saddle point approximation:**

$$W'(\lambda_i) - 2g^2 \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j} = 0 \; . \tag{26}$$

The second term is induced by the integration measure. It gives a repulsive 'Coulomb force'. Introduce the **resolvent**:

$$\omega_N(\lambda) = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\lambda_i - x} \; . \tag{27}$$

From this quantity one can obtain the **eigenvalue density**, and it will allow us to solve the random matrix model.

Classically $g = 0$, each eigenvalue lies at an **extremum** of its potential, $W'(\lambda_i) = 0$. Thus we have a collection of points on the complex plane. But as the 't Hooft coupling $g^2 N$ grows, the Coulomb repulsion becomes stronger, and **the eigenvalues spread out** over a curve segment.

Solving the Random Matrix Model

The resolvent solves the **loop equation**, which in the large-N limit has the solution $(\omega(x) := \lim_{N\to\infty} \omega_N(x))$:

$$\omega(x) = -\frac{1}{2\mu}\, W'(x) + \frac{1}{2\mu}\sqrt{W'(x)^2 - f(x)}\ . \tag{28}$$

W' is a polynomial of degree n, and f is of degree $n-1$. We can rewrite this in terms of the singular part of ω:

$$y^2 = W'(x)^2 - f(x)\ \Rightarrow\ y = \sqrt{\prod_{\beta=1}^{2n}(x - a_\beta)}\ . \tag{29}$$

$\{a_\beta\}$ are **branch points**, with branch cuts on $[a_{2\alpha-1}, a_{2\alpha}]$ $(\alpha = 1, \ldots, n)$. The above is called the **spectral curve** of the random matrix model. It is a compact, hyperelliptic **Riemann surface** of genus $n-1$.

Solving the Random Matrix Model

The **density of eigenvalues** can be solved from the above:

$$\rho(x) = \begin{cases} \frac{1}{4\pi g^2 N}\sqrt{f(x) - W'(x)^2} & \text{if } \ x \in \text{supp}\,(\rho) \\ 0 & \text{if } \ x \notin \text{supp}\,(\rho) \end{cases} . \tag{30}$$

The support is the union of cuts, $\mathcal{C} = \cup_{\alpha=1}^{n}[a_{2\alpha-1}, a_{2\alpha}]$.

Wigner's semicircle law. For a quadratic potential, $W(x) = \frac{1}{2}\,x^2$, there is a single cut, and $\rho(x) = \frac{1}{4\pi g^2 N}\sqrt{8g^2 N - x^2}$, Wigner's semicircle law. The length of the cut is $2\sqrt{8g^2 N}$.

The Matrix Model Solution in the 't Hooft approximation

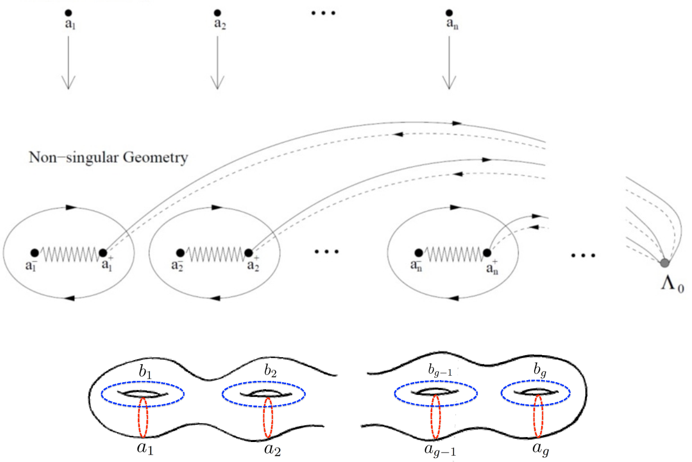

Figure: From a set of singular points to a smooth Riemann surface as $g^2 N \gg 1$.

Emergence in Random Matrix Models

Recall the approximative map, Approx. In this case, this map is the 't Hooft approximation ($g^2N \gg 1$): $\text{Approx}\,(T_N) =: T_\infty$, where T_∞ is the 't Hooft approximation to the random matrix model. I will argue that:

$$T_\infty := \text{Approx}\,(T_N) \neq T_N \quad \text{(for any } N) \,. \tag{31}$$

Consequently, also the **domains of the interpretated theories** differ:

$$D_\infty \neq D_N \quad \text{(for any } N) \,. \tag{32}$$

Eqs. (31)-(32) are not hard to show. The (interpreted) **state space** at finite N is the set of N **eigenvalues**. But this sequence of state spaces differs, for any N, from T_∞'s state space, i.e. the **complex plane** (which the dynamics constrains to be a Riemann surface with a prescribed number of handles).

Also the quantities and the dynamics disagree. (One can choose *subsets* of states and quantities, so that there is *partial* reduction.)

Novelty of Interpretation

The interpretations are **qualitatively different:**

(Finite N) At any finite N, the state space is as a **discrete set of points** on $\mathbb{C}$. The degree of the potential, $n \in \mathbb{Z}$, measures the growth of the **force** between the eigenvalues with increasing distance.

($N \to \infty$) The interpretation of the state space of T_∞ is as a **Riemann surface**. The integer n characterises the genus. There is **no force interpretation**. No matter how large N, the interpretations for finite and infinite N refer to **different kinds of objects**. Interpretation and approximation do not commute, viz. $I_\infty \circ \text{Approx} \neq I_N$.

This can be summarised in the following diagram:

$$\begin{array}{ccc} T_\infty & \xrightarrow{I_\infty} & D_\infty \\ \text{Approx} \uparrow & & \nparallel \\ T_N & \xrightarrow{I_N} & D_N \end{array}$$

Figure: Failure of interpretation and approximation to commute.

Novelty and Robustness

The novelty of the approximating theory is conveniently summarised in the emergence of a collective field, the **master field**, which describes the states, quantities, and dynamics of T_∞. This field can be described either by the density of eigenvalues, or by the Riemann surface. This semi-classical field describes an infinite number of degrees of freedom.

So this is a case of **ontological emergence**, because of the **difference in interpretation**. But to have emergence, we also require **robustness** of the novel behaviour. Robustness concerns the relative independence of T_∞ from the details of T_N at $N \in \mathbb{N}$.

The density of eigenvalues $\rho(\lambda)$ is indeed **independent of the location of the individual eigenvalues** in T_N. ρ_N's support, for finite N, is a collection of N points on the complex plane. In the limit, ρ's support consists of a finite number n of cuts. The 'memory' of the original location of the eigenvalues is lost!

Space from a Theory without Space?

All approaches to quantum gravity seem to point to spacetime's emergence from some **non-geometric or pre-geometric structure**. And part of my aim was to try and make this idea precise, for space.

Question: is the random matrix model, at finite N, a completely non-geometric, or pre-geometric, model?

Answer: Yes, because the defining equations contain only eigenvalues and integrals over them: with no geometric structures.

Obvious objection: there is geometry in defining the N^2-dimensional space of matrices, and in choosing a contour of integration $\gamma \subset \mathbb{C}$!

But it is surely unreasonable to demand that space should emerge out of mathematical structures that are not, in any way, *mathematical* spaces. Most structures in physics *are* mathematical, even if not *physical*, spaces. Write a quantity $Q \in \mathbb{R}$, and you have a space!

Space from a Theory without Space?

The demand for no space should be *balanced*: as the demand that the set of states of the underlying theory be **non-geometrically interpreted**.

Random matrices *do* satisfy this more reasonable demand. The matrix Φ takes values in a space of *matrices*, of dimension N^2: not a space that can be interpreted geometrically in any way that would be relevant to physics. And a state is specified by a finite set of N eigenvalues. This is an **algebraic**, rather than geometric, structure.

The **contour of integration** γ is again not the kind of geometric structure we have in mind when we say 'space', in the context of random matrices—a Riemann surface with meromorphic functions on it, etc. The contour is at best a *pre-geometric* structure. For the Hermitian matrix model, the contour is simply $\mathbb{R}$.

Conclusion: Duality

We advocate the notion of duality as **isomorphism of model roots** (triples).

On this account, duality is a special case of **theoretical equivalence**.

The schema allows conceptions of a number of other notions, like **physical equivalence**.

The schema is illustrated by **boson-fermion duality** (exactly) and **gauge-gravity duality** (at large N).
Other examples worth exploring: electric-magnetic, T duality, S duality.

What gives the examples their specificity/scientific importance is:
(i) the structures preserved by the isomorphism;
(ii) their physical interpretation.

The framework can also be used to give an account of **emergence:**

Conclusion: Emergence

The notion of reference distinguishes ontological vs. epistemic emergence. The mark of ontological, approximative emergence is in the **lack of commutativity between approximation and interpretation**.

Interpretations are maps with specific domains of application: they are constrained by their empirical adequacy. Emergence is to be guided by as rigorous as possible physical interpretation—which is assessed with independent criteria.

Random matrix models exhibit approximative emergence of space in a straightforward way. At finite N, the theory is interpreted in terms of points and algebraic structures. In the 't Hooft limit, the theory is a classical theory with a master field: a **Riemann surface with its geometric structures**. Beware: limits are not essential to approximative emergence! (emergence is already visible at finite N).

The emergence is ontological because there is novel reference: **approximation and interpretation fail to commute**.

References

- (2016) 'Spacetime and Physical Equivalence'. Forthcoming in *Space and Time after Quantum Gravity*, Huggett, N. and Wüthrich, C. (Eds.)
- (2017) with Butterfield. 'A Schema for Duality, Illustrated by Bosonization', *Foundations of Mathematics and Physics one century after Hilbert*, Kouneiher, J. (Ed.). In press.
- (2017) 'Towards a Theory of Emergence for the Physical Sciences'.
- (2017) 'The Emergence of Space, Illustrated by Random Matrix Models'.
- (2015) 'Dualities and emergent gravity: Gauge/gravity duality', *SHPMP* 59.
- (2015) with Dieks, Dongen. 'Emergence in Holographic Scenarios for Gravity', *SHPMP* 52.
- (2017) 'Duality and Theoretical Equivalence'.
- (2016) with Mayerson, Butterfield. 'Conceptual Aspects of Gauge/Gravity Duality', *Foundations of Physics* 46.

Thank you!