The Uses of Infinity—Emergence and Reduction Reconciled? J. Butterfield

1. Introduction

I take emergence as novel and robust behaviour. But with two claims, I try to pour oil on troubled waters.

(1) Emergence is compatible with logical reduction. I will give four examples, each with a parameter $N = \infty$; (N is the number of degrees of freedom). And for each: choosing a salient weaker theory using finite N blocks the reduction.

(2) Supervenience is a red herring.

The examples are similar to—but more striking than!—continuous models of fluids (e.g. sound, flow), in that they model a finite system, i.e a system with finitely many degrees of freedom (atomism!), as infinite.

Three obvious justifications of $N = \infty$, which are shared with such models of fluids.

1: Mathematical convenience: impossible to overstate!

2: Elimination of finitary effects; (cf. how transient effects die out as time tends to infinity)

3: Empirical success: the proof of the pudding is in the eating.

So, agreed: $\infty - 10^{23} = \infty$! But if $f(10^{23}) \approx f(10^{46})$ etc, then it is good, and sometimes indispensible, to model with $f(\infty)$.

But in our examples, the limit yields something qualitatively new—often, new mathematical structure.

1.) Agreed: reduction can fail ... and in various ways. Two interesting ways:

A: a derivation can be correct but only by appealing to concepts/properties, and assumptions/facts about them, that really/originally belong to the reduced theory/phenomenon; (philosophical issues about explanation and the identity of properties);

B: a derivation can be irreparably wrong, because in fact the underlying theory is wrong; (two conjectural examples: Prigogine on phase transitions; Leggett on superfluid ^{3}He.)

2.) "But N is actually finite!"

I propose an analogy with David Lewis' view: that a counterfactual is made true by some intrinsic aspects of the actual world—but in order to state those aspects, it is indispensible to mention other worlds, similar in certain respects to the actual one.

2. Emergence as novelty and robustness

Emergence as properties/behaviour that are both novel and robust relative to some comparison class: especially one given by a theory of the micro-details.

Novelty and robustness are liable to be ambiguous, even controversial or subjective: even for a fixed comparison class. But nevermind!

So emergence is not just 'good' variables and-or approximation schemes. Here "good" is ambiguous between:

small in number and autonomous (uncoupled equations);

easily calculated with;

suited to *given* problem;

insightful, eg by suggestiveness for other theory.

These ideas, about both good variables/schemes and emergence, seem independent of philosophers' proposals that emergence is either:
(i) failure of reduction (in logicians' sense of *definitional extension*); or
(ii) failure of supervenience (also called: determination).

3. Emergence vs. reduction? No!
Beware the contrasting jargons! A philosopher says

> Theory $T_b \equiv T_{\text{bottom/basic/best}}$ reduces $T_t \equiv T_{\text{top/tangible/tainted}}$; or T_t reduces to, is reducible to, T_b

whereas a physicist says

> T_b reduces to T_t (typically in some limit of some parameter of T_b).

Using philosophers' jargon: T_t is logically reduced to T_b by being shown to be logically derivable from T_2—almost always together with some (judiciously chosen!) definitions. (*Definitional extension.*)

A theory could describe novel and robust properties/behaviour, while being a definitional extension of another. So the power of reduction is stronger than commonly admitted.

Example: equilibrium classical statistical mechanics (CSM). For the measures used (micro-canonical etc.) are definable from micro-mechanics (including Lebesgue measure on the phase space).

But examples are often sensitive to your exact choice of theories. Here: Does CSM include the ergodic hypothesis? That is not in general provable from micro-mechanics!

Our examples will involve reduction (using a strong enough mathematical language) with a parameter $N = \infty$; which is lost if you consider (naturally enough!) finite N.

Agreed: philosophers traditionally require more of reduction than just definitional extension. Recall the first way that a putative derivation may fail: it has to appeal to concepts/properties, and assumptions/facts about them, that really/originally belong to the reduced theory.

But in our examples, the theory T_b with $N = \infty$ *is* rich enough to contain the concepts/properties that the derivation needs. So we set aside these philosophical issues about e.g. explanation and the identity of properties.

4. Emergence as supervenience? No!
Supervenience is, in the first instance, a relation between families of properties: viz. that total matching of any two entities as regards one family of properties (called the *subvening* family, say $\mathcal{B}$) implies their total matching as regards the other family (the *supervening* family, say $\mathcal{T}$).

Making this idea precise shows that it is a weakening of definitional extension (and so connect to the idea of supervenience as a relation between theories). Namely: a weakening that allows one or more of the definitions (of a property $T \in \mathcal{T}$ in terms of the various

$B \in \mathcal{B}$) to be infinitely long. That is: for each $T \in \mathcal{T}$, supervenience allows an unstructured, open-ended, infinity of "ways to be T". There is no "control" on the infinite disjunction; in particular, no kind of limit is taken.

(1): I think this open-endedness makes supervenience too weak a concept to be scientifically useful. Anyway, in our examples:
Though we will see some supervenience relations hold good, they will be unenlightening— since there is no connection between their infinite disjunction and the limits, especially $N \to \infty$, crucial to the example.

(2): Our examples of emergence with reduction (especially: logical reduction, i.e. definitional extension) imply that:
(i): Emergence is not "mere" supervenience; i.e. emergence is not definitional extension, with the requirement that for at least one $T \in \mathcal{T}$, the definition is infinitely long; and that
(ii): Emergence is not some sort of failure of supervenience.

These points, (1) and (2), mean that supervenience is a red herring.

5. States, quantities and limits
The best approach to discussing emergence is to consider limiting relations between theories (in general: for some states, some quantities, some parameter-values)—*regimes*). So emergence will often concern regimes for "large" systems.

Theories T_0 and T_κ postulate state-spaces Γ_0 and Γ_κ, and sets (algebras) of quantities $\mathcal{A}_0$ and $\mathcal{A}_\kappa$. Think of κ as a real parameter labelling a "version" of a generic theory: e.g. $\kappa \equiv \hbar$ if the generic theory is quantum mechanics, and T_0 is classical mechanics.
So there are two main kinds of limiting relation: about states and quantities.

For all, or maybe just some, states $s_0 \in \Gamma_0$, there is a sequence of states $s_\kappa \in \Gamma_\kappa$ such that $\mathrm{s}_\kappa \longrightarrow s_0$.

For all, maybe some, quantities $A_0 \in \mathcal{A}_0$, there is a sequence of quantities $A_\kappa \in \mathcal{A}_\kappa$ such that $\mathrm{A}_\kappa \longrightarrow A_0$.

Since the state-spaces/algebras can have different mathematical structures, both arrows $\rightarrow$ need to be clarified. (Especially so, for the quantum-classical case! Cf. quantization theory.)

In general: we expect the $\rightarrow$ s to mesh in that:
at (appropriate) s_0, A_0, the values obey the corresponding relation:
$$\mathrm{A}_\kappa(s_\kappa) \longrightarrow A_0(s_0);$$
and maybe also this relation is preserved under time-evolution.

Surely we should not mind too much which regimes—combinations of states, quantities, and parameter-values—to call 'emergent'. It depends on which ideas, of 'novelty' etc., are emphasised, and so is in part ambiguous/subjective.

6. The method of arbitrary functions

The idea: A roulette wheel, with an unknown regime of biassing in the spin and friction. Suppose we assume:

(i): there are *very many alternating* arcs of red and black;

(ii): whatever the unknown details of the biassing might be, the biassing favoured and disfavoured *large* segments, i.e. segments each of which contains many red and black segments.

(iii): within one of these large segments, the bias is "smooth": adjacent arcs get a similar bias.

Then we can be confident that each long-run frequency is about 50%.

So in general, we expect: Let a sample space (X, μ) be partitioned into two subsets, say R and B, in a very intricate/filamentous way. Then for any probability density function f that is not too "wiggly" (say: whose derivative is bounded: $\mid f' \mid < M$) the probabilities of R and B are about equal: that is

$$\int_R f d\mu \approx \int_B f d\mu \approx \frac{1}{2}. \tag{0.1}$$

And we expect: that, for any bound M on the derivative of the density f, as the partition becomes "more filamentous", the difference from exact equiprobability (and so to both probabilities equalling $\frac{1}{2}$) also goes to 0.

Indeed, Poincare proved for the roulette wheel, where X is the circle, and the intricate partitioning of X is just dividing it into N equal intervals, labelled alternatingly 'red' and 'black':

> For any $M \in \mathbb{R}$, for all density functions f with derivative bounded by M, $\mid f' \mid < M$: as $N =$ the number of arcs goes to infinity:
> $\int_R f\, d\mu \equiv$ prob(Red) $\rightarrow \frac{1}{2}$; and $\int_B f\, d\mu \equiv$ prob(Black) $\rightarrow \frac{1}{2}$.

The idea: any biassing regime, no matter how "wiggly" (i.e. sensitive to angular position), can be "washed out" so as to give equiprobability, by a sufficiently intricate partition (a sufficiently large N).

7. Emergent probabilities: with reduction—and without

Equiprobability in the limit of infinite N is the novel and robust behaviour. Certainly, it is 'robust' in the sense of invariant under the choice of a density function from a wide class.

We have 'logical reduction', in as strong a sense as you could demand. For we here take T_t to be just that statement of equiprobability in the limit of infinite N, and T_b to be model of the wheel, including both

(i) the postulation of various possible density functions f and

(ii) consideration of the infinite limit $N = \infty$.

We can also easily see "the other side of the coin": how the emergent behaviour, i.e. equiprobability, is frustrated if we confine ourselves to finitary T_b. For:

> For any finite N, no matter how large, equiprobability will fail, as badly as you may care to require, for a sufficiently "wiggly" density function, i.e. a sufficiently position-sensitive biassing regime.

That is, we have:

> For all positive integers N, for all $\varepsilon > 0$, there is $M \in \mathbb{R}$ and a density function f with $| f' | < M$ such that: $\int_R f \, d\mu \equiv$ prob(Red) $> 1 - \varepsilon$.

So here is emergence *without* reduction. Since this weaker finitary T_b theory is salient, one is tempted to speak of irreducibility.

Supervenience? Indeed: for any finite N, the long-run frequency in any sequence of spins (including an infinite sequence) supervenes on all the details of the wheel and its many spinnings. (And similarly if one defined a infinitary $N = \infty$ model of a wheel.)

But such supervenience claims seem trivial and useless, presumably because:
(a): there is no control on the infinity (infinite disjunction) they are concerned with, because no kind of limit is taken;
(b): their infinity makes no connection with the limits, infinitely many spinnings and-or $N \to \infty$, that the example is concerned with.

8. Wider still, and wider: the idea of a mathematical function from 1600-2007
Fractals form a recent episode in a grand narrative stretching across 400 years of the history of mathematics: viz. successive generalizations of the notion of function, so as to better address problems in physics. This narrative throws up philosophical issues about the applicability of mathematics to the physical world; prompting me to praise the attitude of Euler—which is happily endorsed by Thomasina Coverly, of Tom Stoppard's play *Arcadia*.

A spectrum of meanings of 'function':—
From strong meanings: e.g. the function must be calculable, or expressible in a formula built from an elite minority of functions, or analytic, or ...
to weak meanings: e.g. any assignment to each of set of arguments (inputs), of a unique value (output).

From 1600, mathematics has traversed this spectrum: often prompted by physics.

Philosophy is written in this immense book that stands ever open before our eyes (I speak of the Universe), but it cannot be read if one does not first learn the language and recognize the characters in which it is written. It is written in mathematical language, and the characters are triangles, circles and other geometrical figures, without the means of which it is humanly impossible to understand a word; without these philosophy is a confused wandering in a dark labyrinth. (Galileo, *Il Saggiatore* 1623, *Opere VI*, 197.)

But since 1600, the mathematical description of nature has had to go far beyond triangles and circles. I give one example: d'Alembert vs. Euler, on vibrating strings; (with a rich legacy, e.g. weak solutions).

d'Alembert (1747) describes the displacement $f(x,t)$ of a vibrating string by:

$$\frac{\partial^2 f}{\partial t^2} = a^2 \, \frac{\partial^2 f}{\partial x^2} \quad . \tag{0.2}$$

Can this describe a plucked string, with a "corner"?

d'Alembert says: No. Euler replies (1748): Yes. The analysis should be generalised 'so that the initial shape of the string can be set arbitrarily ... either regular and contained in a certain equation, or irregular and mechanical'. (Euler proposes to mean by 'function' what we now call a continuous function with piecewise continuous first and second derivative.)

Truesdell: 'Euler's refutation of Leibniz [i.e.: Leibniz's claim that natural phenomena can all be described by what we now call analytic functions] was the greatest advance in scientific methodology in the entire century.'

Mark Wilson suggests the following taxonomy (and rejection of what he calls 'Lazy Optimism').

> Optimism: *every real-life physical structure can be expected to have a direct representative in mathematics.*
> Lazy optimism: *optimism is guaranteed.*
> Honest optimism: *optimism can fail, but has a good track-record.*
> Mathematical opportunism: *few physical structures can be expected to be described adequately by mathematics.*

So Euler is an honest optimist. So is Thomasina Coverly. She is prompted by natural forms, such as bluebells, to generalize the idea of function so as to describe them—and, unlike her tutor Septimus Hodge, she is optimistic about succeeding.

> Thomasina: ... Each week I plot your equations dot for dot, x's against y's in all manner of algebraical relation, and every week they draw themselves as commonplace geometry, as if the world of forms were nothing but arcs and angles. God's truth, Septimus, if there is an equation for a bell, then there must be an equation for a bluebell, and if a bluebell, why not a rose? Do we believe nature is written in numbers?
>
> Septimus: We do.
>
> Thomasina: Then why do your equations only describe the shapes of manufacture?
>
> Septimus: I do not know.
>
> Thomasina: Armed thus, God could only make a cabinet.
>
> Septimus: He has mastery of equations which lead into infinities where we cannot follow.
>
> Thomasina: What a faint heart! We must work outward from the middle of the maze. We will start with something simple. I will plot this leaf and deduce its equation. You will be famous for being my tutor when Lord Byron is dead and forgotten.

9. Fractals in terms of self-similarity and scaling dimension

Fractals make sense of the idea that a set of spatial points, i.e a subset of $\mathbb{R}$ or $\mathbb{R}^2$ or $\mathbb{R}^3$, can have a dimension that is *not* an integer. I approach this through the 'scaling dimension'.

In the late nineteenth century, mathematicians realized that one can define objects, whose features *cannot* be arbitrarily well approximated by using sufficiently small straight line-segments (the methods of the calculus). Examples include the Cantor set (1872) and the Koch snowflake (1906). Because the calculus failed, such objects were regarded as pathological.

But such objects *are* self-similar. Though they are not describable as unions of short line-segments, and their features cannot be treated in terms of limits of such unions, they *are* unions of shrunken copies of *themselves*—just as much as a piecewise-straight line is.

So we use this self-similarity to study these objects. When we apply this idea to the topic of dimension, we find that such objects have a dimension that is not an integer.

Note first that squares and cubes are self-similar: they can be described as unions of shrunken copies of themselves. In particular, the area of a square (volume of a cube) can be defined as the limit of the sum of the areas (volumes) of arbitrarily small squares (cubes) decomposing the given square (cube).

Besides, the description of a square or cube as a union of smaller squares or cubes brings out the idea of dimension as an exponent in a formula for the measure of the original: the area of the square, or volume of the cube. Thus:–

A square with edge l is the union of l^2 unit squares. For example, consider a square whose edge is $l = 2$ units long. It is the union of $2^2 = 4$ unit squares. A square whose edge is $l = 3$ units long is the union of $3^2 = 9$ unit squares.

A cube with edge l is the union of l^3 unit squares. For example, consider a cube whose edge is $l = 2$ units long. It is the union of $2^3 = 8$ unit squares. A cube whose edge is $l = 3$ units long (think of the Rubik cube!) is the union of $3^3 = 27$ unit squares.

Thus we see that in the formula that gives the volume of an object (i.e. the number of unit blocks in the object), the dimension occurs in the exponent. More precisely:

$$\text{number of unit blocks in object with edge } l = l^{\text{dimension of object}}. \tag{0.3}$$

So the main idea will now be that, applying the idea of eq. 0.3 to more "pathological" objects, we find that such objects have non-integral dimensions. I will consider just: the Cantor set and the Koch snowflake.

10. The Cantor set

The Cantor set C (1872) is a subset of the unit interval $[0, 1] \subset \mathbb{R}$. It is defined as the intersection of infinitely many other subsets, which we will call 'stages', labelled 0, 1, 2,...

So we think of the unit interval $[0, 1]$ as stage 0. After stage 0, each later stage is obtained by deleting the open middle third of each closed interval of its predecessor stage. More precisely, ...

Stage 0 is $[0, 1]$. Stage 1 is $[0, 1]$, minus its open middle third. That is: stage 1 is

$[0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. Then stage 2 is defined by deleting the open middle third of each of $[0, \frac{1}{3}]$ and $[\frac{2}{3}, 1]$. So stage 2 consists of four disjoint closed intervals: it is the set $[0, \frac{1}{9}] \cup [\frac{2}{9}, \frac{1}{3}] \cup [\frac{2}{3}, \frac{7}{9}] \cup [\frac{8}{9}, 1]$. And so on. See figure!

Thus stage n is the union of 2^n intervals, each interval being of length $(\frac{1}{3})^n$. So the total length of stage n is $2^n \times (\frac{1}{3})^n \equiv (\frac{2}{3})^n$. So as n goes to infinity, the length of stage n goes to 0.

C is defined to be the intersection of all the stages.

Now we apply to C the idea of scaling dimension. So think of C as a scale 1 object: it is the unit block of "Cantor type". Now observe that C is the union of two shrunken copies of itself, each smaller by a factor of 3. This observation can be "reproduced" at the next scale up. That is: we can define the "Cantor type" object of scale 3, call it C', as the set that results from applying the infinite 'delete and take intersection' process to $[0, 3]$, rather than (as above) to $[0, 1]$.

Just as our original C is the union of two shrunken copies of itself, each smaller by a factor of 3, so also is C'. That is: C' is the union of two shrunken copies of itself, i.e. of two copies of C: of two unit-size Cantor sets. Now we apply the idea of eq. 0.3, getting

$$\text{number of unit Cantor sets in Cantor object of scale } 3 \equiv 2 = 3^{\text{dimension of } C}. \tag{0.4}$$

So what power of 3 is equal to 2? At this point, we recall some elementary logarithms! The fundamental idea of logarithms is that for any real numbers $a, b, c \in \mathbb{R}$, $a^b = c$ is *rewritten* as $\log_a c = b$. Thus $10^2 = 100$ is equivalent to $\log_{10} 100 = 2$. It follows that for any base a

$$b = c^{(\log_a b / \log_a c)} \tag{0.5}$$

so that: $2 = 3^{(\log 2/ \log 3)}$.

That is: the dimension of C is $\log 2/ \log 3$: which is about 0.63.

Another example: the Koch snowflake. This is different in that the Koch snowflake K is not itself the union of similar smaller snowflakes. But each "side" of K is the union of four smaller similar curves, each smaller by a factor 3. So applying again the idea of eq. 0.3, we get:

$$4 = 3^{\text{dimension of Koch}}. \tag{0.6}$$

Since $4 = 3^{(\log 4/ \log 3)}$, we have:

$$\text{dimension of Koch} = \frac{\log 4}{\log 3} \approx 1.26. \tag{0.7}$$

10. Emergent dimensions: with reduction, and without

Non-integer dimensions are novel. And they are 'robust' in various senses:
(i): the scaling dimension is the same for congruent objects, and for enlarged and reduced versions of them;
(ii): there are various cousin novel notions of dimension which can take non-integer values. Broadly speaking: when both the scaling dimension and the cousin are defined, they agree.

So indeed: we have 'emergent dimensions'.

If we take:
as T_b: the rich modern theory of scaling dimension (and its cousins), including as a sub-theory the traditional theory of dimension (integer-valued: mostly, a topological theory), and
as T_t: the assignment of non-integral dimensions to objects like C, K, then:
clearly, we have reduction. T_b contains T_t!

But if T_b is just the traditional theory of dimension, there is no reduction. And because this weaker theory is salient, one is tempted to speak of irreducibility.

Supervenience? As for the method of arbitrary functions:
Indeed: for any finite N, the dimension supervenes on the point-set constitution of the object (subset of $\mathbb{R}^n$) concerned. And similarly for $N = \infty$: the scaling dimension thus supervenes.

But such supervenience claims seem trivial and useless, presumably because:
(a): there is no control on the infinity (infinite disjunction) they are concerned with, because no kind of limit is taken;
(b): their infinity makes no connection with the limit, $N \to \infty$, that the example is concerned with.

So much for pure mathematics. But fractals have many empirical applications, especially in physics—and not just computer graphics! So we have to ask...

11. Is Fractal Geometry the Geometry of Nature?

We should distinguish two questions. First: *do fractals describe the geometry, in the everyday sense, of naturally occurring ("natural history") objects?*

It sure looks like it! Thus Valentine in *Arcadia* says: 'If you knew the algorithm and fed it back say ten thousand times, each time there'd be a dot somewhere on the screen. You'd never know where to expect the next dot. But gradually you'd start to see this shape, because every dot would be inside the shape of this leaf. It wouldn't be a leaf, it would be a mathematical object. But yes ...'

Yes? But a fractal has an infinite sequence of intricate but similar structure on ever smaller length-scales. A leaf and its ilk surely does not! So if 'the geometry of nature' means the complete and accurate description of the shapes and sizes of material objects (using the usual euclidean geometry of physical space), the answer to our question is definitely: 'No'.

Indeed, this negative verdict can be generalized. Usually, a fractal dimension D is calculated from a relation, between a property P and the resolution r, of the general form

$$P = kr^{f(D)} \tag{0.8}$$

where k is just a prefactor, and f is a simple function of D.

But a 1998 survey of a hundred *Physical Review* articles showed that at most three (and on average 1.3) orders of magnitude were probed.

Example: you measure the length of a coastline with a metre-stick, a decimetre-stick, a centimetre-stick, and a millimetre-stick. The average spread of length-scales that were probed was 1.3 orders of magnitude; i.e. from a minimum probed length L to about $13L$.

Agreed: It remains striking that:

(i) the result of an experimental resolution analysis often yields a power law, with D non-integral—as occurs in fractals.

(ii): It is heuristically better to have the suggestive language and results of fractal geometry, than just the bare power law.

(ii) can be strengthened. Recall that ever since Lagrange introduced configuration space, physical theories have made indispensable use of various spaces, especially state-spaces, often equipped with a rich structure that surely deserves the name 'geometry'. In our most successful such theories, such a space is surely 'physically real', and so its geometry deserves to be called (a) 'geometry of nature'. So if we ask instead a second question,

> *Do some of our best physical theories use fractals to describe certain subsets of their abstract spaces, in particular attributing a non-integer dimension to such objects?*

the answer is: *Yes*. Two examples:

(1): Statistical mechanics describes aspects of some processes (phase transitions, like boiling and freezing) with scale-free (regimes of) theories, involving power-law behaviour on all scales, and fractals.

(2): In classical mechanics, there are physically important fractal sets. The famous Lorenz attractor was recently proven to have fractal dimension.

To sum up this example: Again, we see the morals:

(i): the large finite is well-modelled by the infinite;

(ii): the infinite brings new mathematical structure: here non-integer dimension;

(iii): the three obvious justifications from Section **1** apply: mathematical convenience, elimination of finitary effects, and empirical success;

(iv): Section **1**'s proposed analogy with David Lewis' view about counterfactuals also applies: to state the truth about the finite it is indispensible to mention the infinite;

(v): various supervenience claims (which for Section **10**, I have not spelt out) hold—but they are trivial, or at least scientifically useless; and finally, our *main moral*:

(vi) there is a reduction: the emergent non-integer dimensions are reducible to a sufficiently rich theory that takes the infinite limit.

12. Phase transitions

Statistical mechanics follows thermodynamics in representing phase transitions by singularities, viz. by non-analyticities of the free energy F ($\sim$ log of the partition function $Z \sim \int \exp(-\beta H)$).

But a non-analytic point cannot occur for the free energy of a finite system. So the thermodynamic limit is taken: $N :=$ number of particles, $V :=$ volume $\to \infty$ with $\rho = N/V$ fixed.

This infinite limit brings new mathematical structure: e.g. in Yang-Lee theory, the density of zeroes in $\mathbb{C}$. And again, we have three obvious justification: mathematical convenience, elimination of finitary effects (here: especially edge effects), and empirical success.

But: surely a boiling kettle is not infinite! So what exactly should we say about the finite kettle?

I endorse Mainwood's proposal: for systems with a well-defined thermodynamic limit, $F_N \to F_\infty$: phase transitions occur in the finite system iff F_∞ has non-analyticities.
(And if we wish, we can add: *and* if N is large enough, or the gradient of F_N is steep enough. This vagueness is acceptable.)

This definition justifies the application of the thermodynamic limit. It is vivid in specific effects, e.g. cross-over; and in simulations, e.g. results for a 10×10 square lattice can be very close to those for the thermodynamic limit.

Emergent phase transitions:
Our main moral, (vi) at the end of **11**, applies again. That is: the emergent phase transitions are reducible to a sufficiently rich theory that takes the appropriate infinite limit, but not to a theory of finite N.

13. Superselection in the $N \to \infty$ limit of quantum mechanics

Imagine that we compare two vectors, each of unit-length, with angle θ between them, by giving the pair the "score" $\cos\theta$. So if they are parallel, the score is $\cos 0 = 1$; but if they are not parallel, it is less than 1.

Imagine that we compare two sequences of unit vectors, each with N members, by taking the cos of the angle between corresponding members of the sequences, and multiplying all these. So we write:

$$\text{score}(< v_1, v_2, ..., v_N > , < u_1, u_2, ..., u_N >) := \cos\theta_{v_1 u_1} \cos\theta_{v_2 u_2} ... \cos\theta_{v_N u_N} \tag{0.9}$$

We now let N tend to infinity, and consider the limiting values of the score we have defined. We note two sorts of case.

(1): A pair of infinite sequences $< v_i >, < u_i >$ in which the vectors at corresponding positions i are *not* parallel, only for finitely many i. Then only finitely many factors in the score will be *different* from 1; infinitely many factors will be 1. So the total infinite product of numbers is a product of finitely many cosines each less than 1. This is some number less than 1. It *might* be zero: namely if at least one corresponding pair of vector v_i, u_i are at right angles to each other.

(2): A pair of infinite sequences $< v_i >, < u_i >$ in which the vectors at corresponding i are *not* parallel, for infinitely many i. So the total infinite product of numbers is a product including infinitely many numbers that are each less than 1. In general, this infinite product is zero (and even if there are *also* infinitely many factors each equal to 1).

These are the (happily!) elementary considerations underlying the emergence of superselection in the $N \to \infty$ limit of quantum mechanics.

(1) corresponds to two quantum mechanical states (two infinite sequences of unit-vectors) being in the same superselection sector. And

(2) corresponds to two quantum mechanical states being in different superselection sectors. We think of these two sectors as *orthogonal* (i.e. at right angles). This is neatly encoded in the infinite product being zero. For recall that for elementary unit vectors v, u: if v and u are at right angles, then $\cos\theta_{uv} = 0$.

14. Other examples

There are many other examples of novel and robust behaviour in a limit.

Some are related to our examples, e.g. KMS states in the thermodynamic limit.

Some involve continuous models of fluids:

1: plumes in convecting fluids;

2: drops falling from a leaking tap.

Some involve short-wave asymptotics:

1: the geometric optics limit ($\lambda \to 0$) of wave optics; and similarly

2: the classical limit ($\hbar \to 0$) of quantum mechanics; (or rather: some aspects of this limit! Which involves so much else, such as: coherent states, decoherence, the measurement problem ...)