Chapter 10
Spontaneous Symmetry Breaking

As we shall see, the undeniable natural phenomenon of ***spontaneous symmetry breaking*** (SSB) seems to indicate a serious mismatch between theory and reality. This mismatch is well expressed by what is sometimes called ***Earman's Principle***:

> 'While idealizations are useful and, perhaps, even essential to progress in physics, a sound principle of interpretation would seem to be that no effect can be counted as a genuine physical effect if it disappears when the idealizations are removed.' (Earman, 2004, p. 191)

To describe the various examples apparently violating Earman's Principle (and hence the link between theory and reality) in a general way (so general even that it will encapsulate the measurement problem), it is convenient to install a definition:

Definition 10.1. Asymptotic emergence *is the conjunction of three conditions:*

1. *A **higher-level theory** H (which is often called a **phenomenological theory** or a **reduced theory**) is a limiting case of some **theory!lower-level** L (often called **fundamental theory** or a **reducing theory**).*
2. *Theory H is well defined and understood by itself (typically predating L).*
3. *Theory H has features that cannot be explained by L, e.g. because L does not have any property inducing those feature(s) in the pertinent limit to H.*

In connection with SSB (as item 3.) we will look at the following pairs (H, L):

- – H is classical mechanics (notably of a particle on the real line $\mathbb{R}$);
  – L is quantum mechanics (on the pertinent Hilbert space $L^2(\mathbb{R})$);
  – The limiting relationship between the two theories is as described in §7.1 (notably by the continuous bundle of C*-algebras (7.17) - (7.19) for $n = 1$).
- – H is classical thermodynamics of a spin system;
  – L is statistical mechanics of a quantum spin system on a *finite* lattice;
  – Their limiting relationship is as described in §8.6 (cf. Theorem 8.4).
- – H is statistical mechanics of an *infinite* quantum spin system;
  – L is statistical mechanics of a quantum spin system on a *finite* lattice;
  – The limiting relationship between H and L is given in §8.6 (cf. Theorem 8.8).

Of course, there are many other interesting example of (apparent) asymptotic emergence not treated in this book, such as geometric optics (as H) versus wave optics (as L), where the new feature of H would be the *absence* of interference of light rays—foreshadowing the measurement problem of quantum mechanics!— or hydrodynamics (as H) versus molecular dynamics (as L), where the new feature is irreversibility. Perhaps space-time asymptotically emerges from quantum gravity.

The "unexplained" features of H mentioned in the third part of Definition 10.1 are often called **emergent**, although this term has to be used with great care. Its meaning here reflects the original use of the term by the so-called "British Emergentists" (whose pioneer was J.S. Mill), as expressed in 1925 by C.D. Broad:

> 'The characteristic behaviour of the whole *could* not, even in theory, be deduced from the most complete knowledge of the behaviour of its components, taken separately or in other combinations, and of their proportions and arrangements in this whole. This is what I understand by the 'Theory of Emergence'. I cannot give a conclusive example of it, since it is a matter of controversy whether it actually applies to anything.' (Broad, 1925, p. 59)

In quotations like these, the notion "emergence" is meant to be the very opposite of the idea of "reduction" (or "mechanicism", as Broad called it); in fact, for many authors this opposition seems to be the principal attraction of emergence. In principle, two rather different notions of reduction then lead (contrapositively) to two different kinds of emergence, which are sometimes mixed up but should be distinguished:

1. The reduction of a whole (i.e., a composite system) to its parts;
2. The reduction of a theory H to a theory L.

In older literature concerned with the reduction of biology to chemistry (challenged by Mill) and of chemistry to physics (still contested by Broad), the first notion also referred to wholes consisting of a *small* number of particles. That notion of emergence seems a lost cause, since, as noted by Hempel,

> 'the properties of hydrogen include that of forming, if suitably combined with oxygen, a compound which is liquid, transparent, etc.' (Hempel, 1965, p. 260)

A similar comment applies to e.g. the tertiary structure of proteins, but also to cases of emergence such as ant hills, slime mold, and even large cities (Johnson, 2001), all of which are actually fascinating success stories for *reductionism*.

More recently, the apparent possibility that *very large* assemblies of parts might give rise to emergent properties of the corresponding wholes has become increasingly popular, both in physics and in the philosophy of mind (where consciousness has been proposed as an emergent property of the brain). In physics, the modern discussion on emergence in physics was initiated by P.W. Anderson, who in a famous essay from 1972 called 'More is different' emphasized the possibility of emergence in very large systems (surprisingly, Anderson actually avoids the term 'emergence', instead speaking of 'new laws' and 'a whole new conceptual structure'). In particular, Anderson claimed SSB to be an example (if not *the* example) of emergence, duly adding that one really had to take the $N \to \infty$ limit. Thus at least in physics, the interesting case for emergence in the first (i.e. whole-part) sense arises if the 'whole' is strictly infinite, as in the thermodynamic limit of quantum statistical mechanics.

This example confirms that 1. and 2. often go together, but they do not always do: the classical limit of quantum mechanics is a case of pure theory reduction.

A clear description of emergence has also been given by Jaegwon Kim:

1. *Emergence of higher-level properties:* All properties of higher-level entities arise out of the properties and relations that characterize their constituent parts. Some properties of these higher, complex systems are "emergent", and the rest merely "resultant". Instead of the expression "arise out of", such expressions as "supervene on" and "are consequential upon" could have been used. In any case, the idea is that when appropriate lower-level conditions are realized in a higher-level system (that is, the parts that constitute the system come to be configured in a certain relational structure), the system will necessarily exhibit certain higher-level properties, and, moreover, that no higher-level property will appear unless an appropriate set of lower-level conditions is realized. Thus, "arise" and "supervene" are neutral with respect to the emergent/resultant distinction: both emergent and resultant properties of a whole supervene on, or arise out of, its microstructural, or micro-based, properties. The distinction between properties that are emergent and those that are merely resultant is a central component of emergentism. As we have already seen, it is standard to characterize this distinction in terms of predictability and explainability.

2. *The unpredictability of emergent properties:* Emergent properties are not predictable from exhaustive information concerning their "basal conditions". In contrast, resultant properties are predictable from lower-level information.

3. *The unexplainability/irreducibility of emergent properties:* Emergent properties, unlike those that are merely resultant, are neither explainable nor reducible in terms of their basal conditions.' (Kim, 1999, p. 21, italics added)

Similarly, Silberstein (2002) states (paraphrased) that a higher-level theory H:

'bears predictive/explanatory emergence with respect to some lower-level theory L if L cannot *replace* H, if H cannot be *derived* from L [i.e., L cannot reductively explain H], or if L cannot be shown to be *isomorphic* to H.'

A key point here is Kim's no. 1: not even "emergentists" deny that the whole consists of its parts, or, in asymptotic emergence, that the higher-level theory H in fact originates from the lower-level theory L. The essence of emergence, then, would be that H nonetheless has "acquired" properties *not* reducible to L. One possibility for this to happen could be that the (allegedly) emergent property of H refers to some concept that does not even make sense in L, such as the experience of pain, which is hard to make sense of at a neural level, but another possibility, which is indeed the one relevant to physics and especially to SSB, is that some particular concept possessed by H (such as SSB) is admittedly *defined* within L, but *banned*.

In describing the relationship between H and L we have to be clear about the difference between *approximations* and *idealizations*. Following Norton (2012):

- An ***approximation*** is an inexact description of a target system.
- An ***idealization*** is a fictitious system, distinct from the target system, some of whose properties provide an inexact description of aspects of the target system.

Thus idealizations also provide approximations, but as systems they stand on their own and are defined independently of the target system. In our cases, the target system is a real physical system such as a ferromagnet or a quantum particle, which

is supposed to be described exactly by theory L, i.e., the lower-level theory. In fact, L is a family of theories parametrized by $1/N$ ($N \in \mathbb{N}$) or $\hbar \in (0, 1]$, and our real material relates to some very small value of this parameter (which may also be seen as a certain regime of L, seen as a single, unparametrized theory).

The pertinent theory H is an idealization in the above sense, through which one approximates very large systems by infinite ones and highly semi-classical ones (where $\hbar$ is very small) by classical ones (where $\hbar = 0$). It is in this setting that asymptotic emergence would violate Earman's Principle and hence would blast the relationship between theory and reality: the abstract point (made concrete for SSB earlier on) is that if some real property of a real system is described by H but is not approximated in any sense by L in any regime (as is the threat with SSB), although H is supposed to be a limit of L, then the latter theory L fails to describe the real system it is supposed to describe, whereas this systems *is* described by the theory H, which portrays fictitious systems. This marks a difference with other cases of emergence, where H (including some "whole") is not an idealization but a real system itself (as might be the case with consciousness and other examples from neuroscience and the philosophy of mind). Thus our discussion does not apply to such cases.

The tension between SSB and Earman's Principle has not quite gone unnoticed in the philosophy of physics literature. For example, Liu and Emch (2005) first write that it is a mistake to regard idealizations as acts of '*neglecting the negligible*' (p. 155, which already appears to deny Earman's Principle), and continue by:

> "The broken symmetry in question is *not reducible* to the configurations of the microscopic parts of any *finite* systems; but it should *supervene* on them in the sense that for any two systems that have the exactly (sic) duplicates of parts and configurations, both will have the same spontaneous symmetry breaking in them because both will behave identically in the limit. In other words, the result of the macroscopic limit is determined by the non-relational properties of parts of the finite system in question." (Liu & Emch, 2005, p. 156)

It is not easy to make sense of this, but the authors genuinely seem to believe in asymptotic emergence and hence they (again) appear to deny Earman's Principle. Another suggestion, made by Ruetsche, is to modify Earman's Principle to:

> 'No effect predicted by a non-final theory can be counted as a genuine physical effect if it disappears from that theory's successors.' (Ruetsche, 2011, p. 336)

For example, the theory L explaining SSB should not be quantum statistical mechanics but quantum field theory (which has an infinite number of ultraviolet degrees of freedom even in finite volume, and hence in principle allows SSB). This does make sense within physics, but, as Ruetsche herself notices, her principle 'has the pragmatic shortcoming that we can't apply it until we know what (all) successors to our present theories are.' With due respect, we will describe a rather different way out, based on unexpectedly implementing ***Butterfield's Principle***, which is a corollary to Earman's Principle that removes the reduction-emergence opposition:

> 'there is a weaker, yet still vivid, novel and robust behaviour that occurs before we get to the limit, i.e. for finite N. And it is this weaker behaviour which is physically real.' (Butterfield, 2011, p. 1065)

To do so, we now turn our attention to specific (classes of) models of SSB.

10.1 Spontaneous symmetry breaking: The double well

The simplest example of SSB is undoubtedly the equation $x^2 = 1$ (where $x \in \mathbb{C}$), which is invariant under a $\mathbb{Z}_2$ symmetry given by $x \mapsto -x$. Its solutions $x = \pm 1$, then, do not share this symmetry; instead $\mathbb{Z}_2$ acts nontrivially on the solution space.

Another example that is simple at least compared to quantum spin systems is provided by elementary quantum mechanics. Thus we are now in the context of the first of the three pairs (H, L) listed in the preamble to this chapter, where, in detail:

-   H is classical mechanics of a particle moving on the real line, with associated phase space $\mathbb{R}^2 = \{(p, q)\}$ and ensuing C*-algebra of observables $A_0 = C_0(\mathbb{R}^2)$;
-   L is the corresponding quantum theory, with a C*-algebra of observables $A_\hbar$ ($\hbar > 0$) taken to be the compact operators $B_0(L^2(\mathbb{R}))$ on the Hilbert space $L^2(\mathbb{R})$;
-   The relationship between H and L is given by the continuous bundle of C*-algebras (7.17) - (7.19), for $n = 1$, notably in the classical limit $\hbar \to 0$.

At the level of states, the passage to the classical limit $\hbar \to 0$ of any $\hbar$-dependent wave-function $\psi_\hbar \in L^2(\mathbb{R})$, if it exists, is described via the associated probability measure $\mu_{\psi_\hbar}$ on $\mathbb{R}^2$, which is defined by (7.31); in other words,

$$\mu_{\psi_\hbar}(\Delta) = \int_\Delta \frac{d^n p \, d^n q}{2\pi\hbar} |\langle \phi_\hbar^{(p,q)}, \psi_\hbar \rangle|^2 \ (\Delta \subset \mathbb{R}^{2n}), \tag{10.1}$$

where the (Schrödinger) ***coherent states*** $\phi_\hbar^{(p,q)} \in L^2(\mathbb{R})$ are given by (7.27), i.e.,

$$\phi_\hbar^{(p,q)}(x) = (\pi\hbar)^{-n/4} e^{-ipq/2\hbar} e^{ipx/\hbar} e^{-(x-q)^2/2\hbar}. \tag{10.2}$$

In terms of the associated vector states $\omega_{\psi_\hbar}$ on the C*-algebra $B_0(L^2(\mathbb{R}))$, one has

$$\omega_{\psi_\hbar}(Q_\hbar^B(f)) = \langle \psi_\hbar, Q_\hbar^B(f)\psi_\hbar \rangle = \int_{\mathbb{R}^{2n}} d\mu_\psi(p, q) \, f(p, q), \tag{10.3}$$

where $f \in C_0(\mathbb{R}^2)$. We then say that the wave-functions $\psi_\hbar$ have a classical limit if

$$\lim_{\hbar \to 0} \int_{\mathbb{R}^{2n}} d\mu_\psi \, f = \int_{\mathbb{R}^{2n}} d\mu_0 \, f, \tag{10.4}$$

for any $f \in C_0(\mathbb{R}^2)$, where μ_0 is some probability measure on $\mathbb{R}^2$. Seen as a state ω_0 on the classical C*-algebra of observables $C_0(\mathbb{R}^2)$, the probability measure μ_0 is regarded as the classical limit of the family $\omega_{\psi_\hbar}$ of states on the C*-algebra $B_0(L^2(\mathbb{R}))$ of quantum-mechanical observables. This family is continuous in the sense that the function $\hbar \mapsto \omega_{\psi_\hbar}(\sigma(\hbar))$ from $[0,1]$ to $\mathbb{C}$ is continuous for every continuous cross-section σ of the given bundle of C*-algebras. An example of such a continuous cross-section is $\sigma(0) = f$ and $\sigma(\hbar) = Q_\hbar^B(f)$, for any $f \in C_0(\mathbb{R}^2)$), cf. (C.550) - (C.551), and indeed this example reproduces (10.4), which after all is just

$$\lim_{\hbar \to 0} \omega_{\psi_\hbar}(Q_\hbar^B(f)) = \omega_0(f) \ (f \in C_0(\mathbb{R}^2)). \tag{10.5}$$

First, let us illustrate this formalism for the ground state of the one-dimensional harmonic oscillator. Taking $m = 1/2$ and $V(x) = \frac{1}{2}\omega^2 x^2$ in the usual Hamiltonian

$$h_\hbar = -\hbar^2 \frac{d^2}{dx^2} + V(x), \tag{10.6}$$

it is well known that the ground state is unique and that its wave-function, i.e.,

$$\psi_\hbar(x) = \left(\frac{\omega}{2\pi\hbar}\right)^{1/4} e^{-\omega x^2/4\hbar}, \tag{10.7}$$

is a Gaussian, peaked above $x = 0$. As $\hbar \to 0$, this ground state has a classical limit, namely the Dirac measure μ_0 concentrated at the origin $(p = 0, q = 0)$, i.e.,

$$\lim_{\hbar \to 0} \int_{\mathbb{R}^{2n}} d\mu_{\psi_\hbar} f = f(0,0) \quad (f \in C_0(\mathbb{R}^2)). \tag{10.8}$$

This is just the unique ground state of the corresponding classical Hamiltonian

$$h_0(p,q) = p^2 + V(q), \tag{10.9}$$

seen as a point in the phase space $\mathbb{R}^2$ minimizing h_0, reinterpreted as a probability measure on phase space as explained in the context of Theorem 3.3. Note that we kept the mass fixed at $m = 1/2$, but instead we could have kept $\hbar$ fixed and take the limit $m \to \infty$ instead of $\hbar \to 0$; cf. the preamble to Chapter 7.

The same features hold for the *an*harmonic oscillator (with small $\lambda > 0$), i.e.,

$$V(x) = \frac{1}{2}\omega^2 x^2 + \frac{1}{4}\lambda x^4. \tag{10.10}$$

However, a new situation arises for the symmetric double-well potential

$$V(x) = -\frac{1}{2}\omega^2 x^2 + \frac{1}{4}\lambda x^4 + \frac{1}{4}\omega^4/\lambda = \frac{1}{4}\lambda(x^2 - a^2)^2, \tag{10.11}$$

where $a = \omega/\sqrt{\lambda} > 0$ (assuming $\omega > 0$ as well as $\lambda > 0$). This time, the ground state of the classical Hamiltonian is doubly degenerate, being given by the points $(p = 0, q = \pm a) \in \mathbb{R}^2$, with ensuing Dirac measures $\mu_0^\pm$ given by

$$\int_{\mathbb{R}^{2n}} d\mu_0^\pm f = f(0, \pm a). \tag{10.12}$$

But it is a deep and counterintuitive fact of quantum theory that the corresponding quantum Hamiltonian (10.6) with (10.11) has a unique ground state. Indeed:

Theorem 10.2. *Let $V \in L^2_{\mathrm{loc}}(\mathbb{R}^m)$ be positive and suppose that $\lim_{|x| \to \infty} V(x) = \infty$. Then $-\Delta + V$ has a nondegenerate (and strictly positive) ground state.*

Roughly speaking, the proof is based on an infinite-dimensional version of the Perron–Frobenius Theorem in linear algebra (applied to $\exp(-th_\hbar)$ rather than to the Hamiltonian $h_\hbar$ itself, so that the largest eigenvalue of the former corresponds to the smallest eigenvalue of the latter, i.e., the energy of the ground state).

And yet there are two quantum-mechanical shadows of the classical degeneracy:

- The wave-function $\psi_\hbar^{(0)}$ of the ground state (which by a suitable choice of phase may be taken to be real) is positive definite and has two peaks, above $x = \pm a$, with exponential decay $|\psi_\hbar^{(0)}(x)| \sim \exp(-1/\hbar)$ in the classically forbidden region.
- Energy eigenfunctions (and the associated eigenvalues) come in pairs.

In what follows, we will be especially interested in the first excited state $\psi_\hbar^{(1)}$, which like $\psi_\hbar^{(0)}$ is real, but has one peak *above* $x = a$ and another peak *below* $x = -a$. See Figure 10.1. The eigenvalue splitting (or "gap") vanishes exponentially in $-1/\hbar$ like

$$\Delta_\hbar \equiv E_\hbar^{(1)} - E_\hbar^{(0)} \sim (\hbar\omega/\sqrt{\tfrac{1}{2}e\pi}) \cdot e^{-d_V/\hbar} \ (\hbar \to 0), \tag{10.13}$$

where the typical WKB-factor is given by

$$d_V = \int_{-a}^{a} dx \sqrt{V(x)}. \tag{10.14}$$

Also, the probability density of each of the wave-functions $\psi_\hbar^{(0)}$ or $\psi_\hbar^{(1)}$ contains approximate δ-function peaks above *both* classical minima $\pm a$. See Figure 10.2, displayed just for $\psi_\hbar^{(0)}$, the other being analogous. We can make the correspondence between the *nondegenerate* pair $(\psi_\hbar^{(0)}, \psi_\hbar^{(1)})$ of low-lying quantum-mechanical wave-functions and the pair (μ_0^+, μ_0^-) of *degenerate* classical ground states more transparent by invoking the above notion of a classical limit of states. Indeed, in terms of the corresponding algebraic states $\omega_{\psi_\hbar^{(0)}}$ and $\omega_{\psi_\hbar^{(1)}}$, one has

$$\lim_{\hbar \to 0} \psi_\hbar^{(0)} = \lim_{\hbar \to 0} \psi_\hbar^{(1)} = \mu_0^{(0)}, \tag{10.15}$$

$$\mu_0^{(0)} \equiv \tfrac{1}{2}(\mu_0^+ + \mu_0^-), \tag{10.16}$$

where $\mu_0^\pm$ are the pure classical ground states (10.12) of the double-well Hamiltonian. To see this, one may consider numerically computed Husimi functions, as shown in Figure 10.3 (just for $\psi_\hbar^{(0)}$, as before). From this, it is clear that the *pure* (algebraic) quantum ground state $\psi_\hbar^{(0)}$ converges to the *mixed* classical state (10.16).

In contrast, the localized (but now time-dependent) wave-functions

$$\psi_\hbar^\pm = \frac{\psi_\hbar^{(0)} \pm \psi_\hbar^{(1)}}{\sqrt{2}}, \tag{10.17}$$

which of course define pure states as well, converge to *pure* classical states, i.e.,

$$\lim_{\hbar \to 0} \psi_\hbar^\pm = \mu_0^\pm. \tag{10.18}$$

In conclusion, one has SSB in H, but at first sight the underlying theory L seems to forbid it. Yet we will now show that (10.17) - (10.18), will save Earman's Principle.
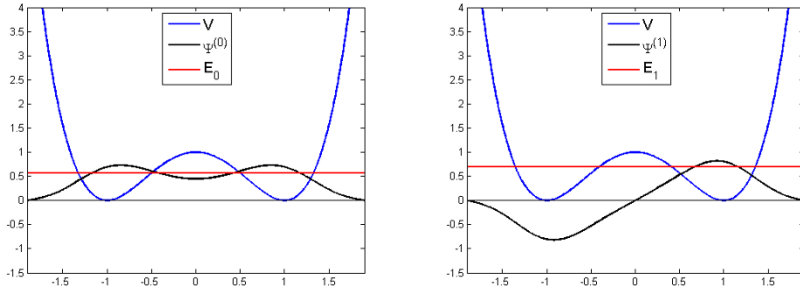
Fig. 10.1 Double-well potential with ground state $\psi_{\hbar=0.5}^{(0)}$ and first excited state $\psi_{\hbar=0.5}^{(1)}$.
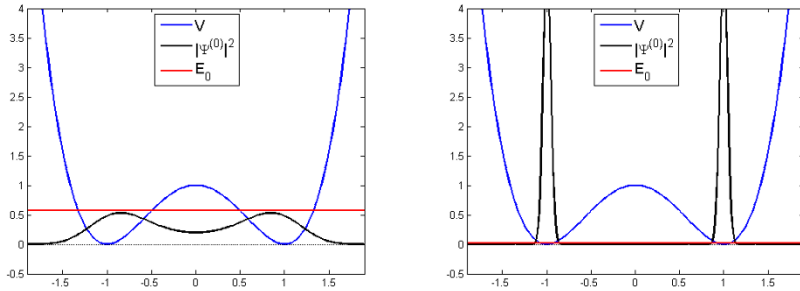

Fig. 10.2 Probability densities for $\psi_{\hbar=0.5}^{(0)}$ (left) and $\psi_{\hbar=0.01}^{(0)}$ (right).
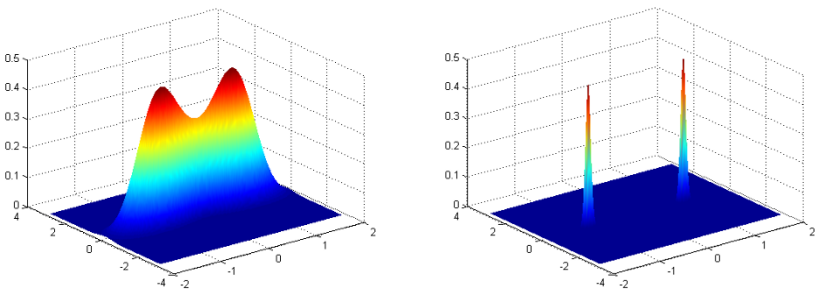

Fig. 10.3 Husimi functions for $\psi_{\hbar=0.5}^{(0)}$ (left) and $\psi_{\hbar=0.01}^{(0)}$ (right).

10.2 Spontaneous symmetry breaking: The flea

Regarding the doubly-peaked ground state $\psi_\hbar^{(0)}$ of the symmetric double well as the quantum-mechanical counterpart of a hung parliament, the analogue of a small party that decides which coalition is formed is a tiny *asymmetric* perturbation δV of the potential. Indeed, the following spectacular phenomenon in the theory of Schrödinger operators was discovered in 1981 by Jona-Lasinio, Martinelli and Scoppola. In view of the extensive (and very complicated) ensuing mathematical literature, we just take it as our goal to explain the main idea in a heuristic way.

Replace V in (10.6) by $V + \delta V$, where δV (i.e., the "flea") is assumed to:

1. Be real-valued with fixed sign, and C_c^∞ (hence bounded) with connected support not including the minima $x = a$ or $x = -a$;
2. Satisfy $|\delta V| >> e^{-d_V/\hbar}$ for sufficiently small $\hbar$ (e.g., by being independent of $\hbar$);
3. Be localized not too far from at least one the minima, in the following sense. First, for $y, z \in \mathbb{R}$ and $A \subset \mathbb{R}$, we extend the notation (10.14) to

$$d_V(y,z) = \left| \int_y^z dx \sqrt{V(x)} \right|; \tag{10.19}$$

$$d_V(y,A) = \inf\{d_V(y,z), z \in A\}. \tag{10.20}$$

Second, we introduce the symbols

$$d_V' = 2 \cdot \min\{d_V(-a, \operatorname{supp} \delta V), d_V(a, \operatorname{supp} \delta V)\}; \tag{10.21}$$

$$d_V'' = 2 \cdot \max\{d_V(-a, \operatorname{supp} \delta V), d_V(a, \operatorname{supp} \delta V)\}. \tag{10.22}$$

The localization assumption on δV is that one of the following conditions holds:

$$d_V' < d_V < d_V''; \tag{10.23}$$

$$d_V' < d_V'' < d_V. \tag{10.24}$$

In the first case, the perturbation is typically localized either on the left or on the right edge of the double well, whereas in the second it resides on the middle bump (symmetric perturbations are excluded by 3, as these would satisfy $d_V' = d_V''$).

Under these assumptions, the ground state wave-function $\psi_\hbar^{(\delta)}$ of the perturbed Hamiltonian (which had two peaks for $\delta V = 0$!) localizes as $\hbar \to 0$, in a direction which *given that localization happens* may be understood from energetic considerations. For example, if δV is positive and is localized to the right, then the relative energy in the left-hand part of the double well is lowered, so that localization will be to the left. See Figures 10.4 - 10.6. Eqs. (10.17) - (10.18) then yield Butterfield's Principle (with $N \rightsquigarrow 1/\hbar$), so that also Earman's Principle is saved: the essence of the argument is that (at least in the presence of a flea-perturbation) SSB is already foreshadowed in quantum mechanics *for small yet positive $\hbar$*, if only approximately.

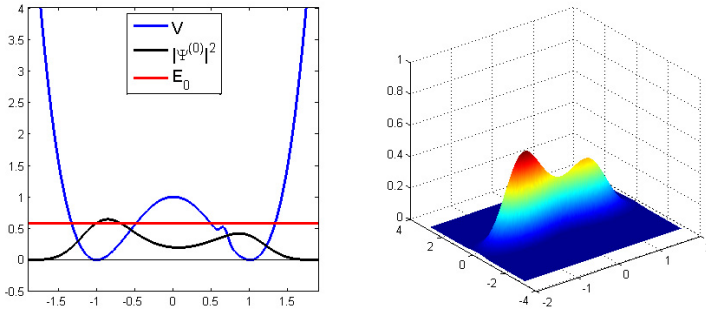**Fig. 10.4** Flea perturbation of ground state $\psi_{\hbar=0.5}^{(\delta)}$ with corresponding Husimi function. For such relative large values of $\hbar$, little (but some) localization takes place.
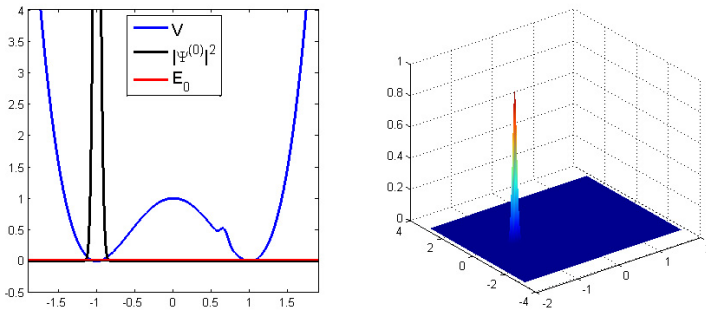

Fig. 10.5 Same at $\hbar = 0.01$. For such small values of $\hbar$, localization is almost total.
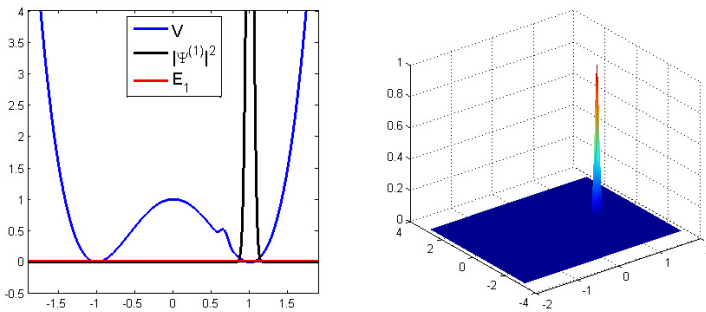

Fig. 10.6 First excited state for $\hbar = 0.01$. Note the opposite localization area.

In more detail, for the perturbed ground state we have (subject to assumptions 1–3):

$$\frac{\psi_\hbar^{(\delta)}(a)}{\psi_\hbar^{(\delta)}(-a)} \sim e^{\mp d_V/\hbar} \quad (\pm \delta V > 0, \, \mathrm{supp}(V) \subset \mathbb{R}^+); \tag{10.25}$$

$$\frac{\psi_\hbar^{(\delta)}(a)}{\psi_\hbar^{(\delta)}(-a)} \sim e^{\pm d_V/\hbar} \quad (\pm \delta V > 0, \, \mathrm{supp}(V) \subset \mathbb{R}^-), \tag{10.26}$$

with the opposite localization for the perturbed first excited state (so as to remain orthogonal to the ground state). A more precise version of the energetics used above is as follows. The ground state tries to minimize its energy according to the rules:

- The cost of localization (if $\delta V = 0$) is $\mathcal{O}(e^{-d_V/\hbar})$.
- The cost of turning on δV is $\mathcal{O}(e^{-d_V'/\hbar})$ when the wave-function is delocalized.
- The cost of turning on δV is $\mathcal{O}(e^{-d_V''/\hbar})$ when the wave-function is localized in the well around $x_0 = \pm a$ for which $d_V(x_0, \mathrm{supp}\, \delta V) = d_V''$.

In any case, these results only depend on the support of δV, but not on its size: this means that the tiniest of perturbations may cause collapse in the classical limit.

Although the collapse of the perturbed ground state for small $\hbar$ is a mathematical theorem, it remains enigmatic. Indeed, despite the fact that in quantum theory the localizing effect of the flea is enhanced for small $\hbar$, the corresponding classical system has no analogue of it. Trivially, a classical particle residing at one of the two minima of the double well at zero (or small) velocity, i.e., in one of its degenerate ground states, will not even notice the flea; the ground states are unchanged. But even under a stochastic perturbation, which leads to a nonzero probability for the particle to be driven from one ground state to the other in finite time (as some form of classical "tunneling", where in this case the necessary fluctuations come from Brownian motion), the flea plays a negligible role. For example, in the case at hand the standard ***Eyring–Kramers formula*** for the mean transition time reads

$$\langle \tau \rangle \cong \frac{2\pi}{\sqrt{V''(a)V''(0)}} e^{V(0)/\varepsilon}, \tag{10.27}$$

where ε is the parameter in the Langevin equation $dx_t = -\nabla V(x_t)dt + \sqrt{2\varepsilon}dW_t$, in which W_t is standard Brownian motion. Clearly, this expression only contains the height of the potential at its maximum and its curvature at its critical points; most perturbations satisfying assumptions 1–3 above do not affect these quantities.

The instability of the ground state of the double-well potential under "flea" perturbations as $\hbar \to 0$ is easy to understand (at least heuristically) if one truncates the infinite-dimensional Hilbert space $L^2(\mathbb{R})$ to a two-level system. This simplification is accomplished by keeping only the lowest energy states $\psi_\hbar^{(0)}$ and $\psi_\hbar^{(1)}$, in which case the full Hamiltonian (10.6) with (10.11) is reduced to the 2×2 matrix

$$H_0 = \tfrac{1}{2} \begin{pmatrix} 0 & -\Delta \\ -\Delta & 0 \end{pmatrix}, \tag{10.28}$$

with $\Delta > 0$ given by (10.13). Dropping $\hbar$, the eigenstates of H_0 are given by

$$\varphi_0^{(0)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \varphi_0^{(1)} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \tag{10.29}$$

with energies $E_0 = -\frac{1}{2}\Delta$ and $E_1 = \frac{1}{2}\Delta$, respectively; in particular, $E_1 - E_0 = \Delta$. If

$$\varphi_0^{\pm} = \frac{\varphi_0^{(0)} \pm \varphi_0^{(1)}}{\sqrt{2}}, \tag{10.30}$$

as in (10.17), then

$$\varphi_0^{+} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \varphi_0^{-} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{10.31}$$

Hence in this approximation φ_0^{+} and φ_0^{-} play the role of wave-functions (10.17) localized above the classical minima $x = +a$ and $x = -a$, respectively, with classical limits $\mu_0^{\pm}$. The "flea" is introduced as follows. If its support is in $\mathbb{R}^{+}$, we put

$$\delta_{+}V = \begin{pmatrix} 0 & 0 \\ 0 & \delta \end{pmatrix}, \tag{10.32}$$

where $\delta \in \mathbb{R}$ is a constant. A perturbation with support in $\mathbb{R}^{-}$ is approximated by

$$\delta_{-}V = \begin{pmatrix} \delta & 0 \\ 0 & 0 \end{pmatrix}. \tag{10.33}$$

Without loss of generality, take the latter (a change of sign of δ leads to the former). The eigenvalues of $H^{(\delta)} = H_0 + \delta_{-}V$ are $E_0 = E_{-}$ and $E_1 = E_{+}$, with energies

$$E_{\pm} = \frac{1}{2}(\delta \pm \sqrt{\delta^2 + \Delta^2}), \tag{10.34}$$

and normalized eigenvectors

$$\varphi_\delta^{(0)} = \frac{1}{\sqrt{2}} \left(\delta^2 + \Delta^2 + \delta\sqrt{\delta^2 + \Delta^2} \right)^{-1/2} \begin{pmatrix} \Delta \\ \delta + \sqrt{\delta^2 + \Delta^2} \end{pmatrix}; \tag{10.35}$$

$$\varphi_\delta^{(1)} = \frac{1}{\sqrt{2}} \left(\delta^2 + \Delta^2 - \delta\sqrt{\delta^2 + \Delta^2} \right)^{-1/2} \begin{pmatrix} \Delta \\ \delta - \sqrt{\delta^2 + \Delta^2} \end{pmatrix}. \tag{10.36}$$

Note that $\lim_{\delta \to 0} \varphi_\delta^{(i)} = \varphi_0^{(i)}$ for $i = 0, 1$. Now, if $\hbar \to 0$, then $|\delta| >> \Delta$, in which case $\varphi_\delta^{(0)} \to \varphi_0^{\pm}$ for $\pm\delta > 0$ (and starting from (10.32) instead of (10.33) would have given the opposite case, i.e., $\varphi_\delta^{(0)} \to \varphi_0^{\mp}$ for $\pm\delta > 0$). Thus the ground state localizes as $\hbar \to 0$, which resembles the situation (10.25) - (10.26) for the full double-well.

   In conclusion, in the (practically unavoidable) presence of asymmetric "flea" perturbations, *explicit* (rather than spontaneous) symmetry breaking already takes place for positive $\hbar$, so that Butterfield's Principle holds, and hence also Earman's.

10.3 Spontaneous symmetry breaking in quantum spin systems

Before discussing SSB in quantum spin systems, we return to ground states and KMS states as discussed in the generality of §§9.4–9.6. Starting with the former, it is natural to ask whether ground states are pure, as would be expected on physical grounds; indeed, this question goes to the heart of SSB. Proposition 9.20 implies that ground states (for given dynamics) form a compact convex subset $S(A)_\infty$ of the total state space $S(A)$; the notation $S_\infty(A)$ (rather than e.g. $S_0(A)$) will be motivated shortly by the analogy with equilibrium states. It would be desirable that

$$\partial_e S_\infty(A) = S_\infty(A) \cap \partial_e S(A), \tag{10.37}$$

in which case extreme ground states are necessarily pure. This will indeed be the case in the simple models we study in this book, but it is provably the case in general only under additional assumptions, such as **weak asymptotic abeliannnes** of the dynamics, i.e., $\lim_{t\to\infty} \omega([\alpha_t(a),b]) = 0$ for all $a,b \in A$. A weaker sufficient condition for (10.37) is that $\pi_\omega(A)'$ be commutative (which is the case if ω is pure).

We are now in a position to define SSB, at least in the context of ground states.

Definition 10.3. *Suppose we have a (topological) group G and a (continuous) homomorphism $\gamma: G \to \mathrm{Aut}(A)$, which is a symmetry of the dynamics in that*

$$\alpha_t \circ \gamma_g = \gamma_g \circ \alpha_t \quad (g \in G, t \in \mathbb{R}). \tag{10.38}$$

The G-symmetry is said to be **spontaneously broken** *(at temperature $T = 0$) if*

$$(\partial_e S_\infty(A))^G = \emptyset, \tag{10.39}$$

and **weakly broken** *if $(\partial_e S_\infty(A))^G \neq \partial_e S_\infty(A)$, i.e., there is at least one $\omega \in \partial_e S_\infty(A)$ that fails to be G-invariant (although invariant extreme ground states may exist).*

Here $\mathscr{S}^G = \{\omega \in \mathscr{S} \mid \omega \circ \gamma_g = \omega \,\forall g \in G\}$, defined for any subset $\mathscr{S} \subset S(A)$, is the set of G-invariant states in $\mathscr{S}$. Assuming (10.37), eq. (10.39) means that there are no *pure* G-invariant ground states. This by no means implies that there are no G-invariant ground states at all, quite to the contrary: for compact, or, more generally, amenable groups G, one can always construct G-invariant ground states by averaging over G, exploiting the fact that if G is a symmetry of the dynamics, then each affine homeomorphism γ_g^* of $S(A)$ (defined by $\gamma_g^*(\omega) = \omega \circ \gamma_g$) maps $S_\infty(A)$ to itself. Definition 10.3 therefore implies that if SSB occurs, then one has a dichotomy:

- *Pure ground states are not invariant, whilst invariant ground states are not pure.*

Definition 10.4. *We call a G-symmetry* **spontaneously broken** *at inverse temperature $\beta \in (0,\infty)$ if there are no G-invariant extreme β-KMS states, i.e.,*

$$(\partial_e S_\beta(A))^G = \emptyset, \tag{10.40}$$

and **weakly broken** *if there is at least one non-G-invariant extreme KMS state.*

By Theorem 9.31 we may replace *extreme β-*KMS* states by *primary β-*KMS* states,
so that, similarly to ground states, SSB at nonzero temperature means that:

- *Primary* KMS *states are invariant, whilst invariant* KMS *states are not primary.*

For the next result, please recall Definition 9.10 and Theorem 9.11.

Proposition 10.5. *Let A be a quasi-local C*-algebra of the kind* (8.130) *and sup-
pose the given G-action γ commutes not only with time translations α_t but also
with space translation τ_x. If $\gamma_g^* \omega \neq \omega$ for some $\omega \in \partial_e S_\beta(A)$ and $g \in G$, then the
automorphism γ_g cannot be unitarily implemented in the* GNS*-representation π_ω.*

This is true also at $\beta = \infty$, i.e., for ground states.

Proof. This is an obvious corollary of Proposition 9.13 and Theorems 9.14 and
9.31: if γ_g were implementable by a unitary u_g, then $u_g \Omega_\omega \neq \Omega_\omega$ (not even up to
a phase), since $\gamma_g^* \omega \neq \omega$. But in that case, since $\tau_x \circ \gamma_g = \gamma_g \circ \tau_x$ for each $x \in \mathbb{Z}^d$,
we would have $u_x u_g = u_g u_x$ and hence $u_x(u_g \Omega_\omega) = u_g \Omega_\omega$. Thus $u_g \Omega_\omega$ would be
another translation-invariant ground state, contradicting Theorem 9.14.   □

This result is worth mentioning, since some authors *define* SSB through the con-
clusion of this proposition, that is, they call a symmetry γ_g (spontaneously) broken
by some state ω iff γ_g cannot be unitarily implemented in π_ω. This definition seems
physically dubious, however, because quantum spin systems may have ground states
ω that are not *G*-invariant but in which nonetheless all of *G* is unitarily imple-
mentable (in such states translation invariance has to be broken, of course). For
example, the Ising model in $d = 1$ with ferromagnetic nearest-neighbour interaction
and vanishing external magnetic field (where $G = \mathbb{Z}_2$) has an infinite number of
such ground states, in which a "domain wall" separates infinitely many "spins up"
to the left from infinitely many "spins down" to the right. Although this model has a
unique KMS state at any nonzero temperature, such ground states (and perhaps anal-
ogous states at $\beta \neq \infty$ in different models, so far understood only heuristically) seem
far from pathological and play a major role in modern condensed matter physics.
Hence we trust this alternative definition only if the states it singles out also satisfy
Definition 10.3 or 10.4, for which Proposition 10.5 gives a sufficient condition: for
translation-invariant states and symmetries on quasi-local algebras, our definition of
SSB through (10.40) is compatible with the one based on unitary implementability.

This is fortunate, since the physicist's notion of an *order parameter*, through
which at least *weak* SSB may be detected, is tailored to translation-invariant states:

Definition 10.6. *Let A be a quasi-local C*-algebra A as in* (8.130)*, with symmetry
group G. A **(strong) order parameter** in A is an n-tuple $\phi = (\phi_1, \ldots, \phi_n) \in A^n$ for
which $\omega(\phi) = 0$ if (and only if) ω is G-invariant, for any $\mathbb{Z}^d$-invariant state ω on A.*

An order parameter defines an accompanying vector field $x \mapsto \phi(x)$ by $\phi_i(x) =
\tau_x(\phi)$. Since ω is translation-invariant, $\omega(\phi) = 0$ is equivalent to $\omega(\phi(x)) = 0$ for
all x. In the Ising model, with $G = \mathbb{Z}_2$, $\sigma_3(0)$ is an order parameter, which can
be extended to a strong one $\phi = (\sigma_2(0), \sigma_3(0))$. In the Heisenberg model, where
$G = SO(3)$, the triple $(\sigma_1(0), \sigma_2(0), \sigma_3(0))$ provides a strong order parameter.

Theorem 10.7. *Suppose that ϕ is a (strong) order parameter, as in Definition 10.6. Then a G-invariant and translation-invariant* KMS *state $\omega \in S_\beta(A)^G$ (including $\beta = \infty$, i.e., a ground state) displays weak* SSB—*in the sense that at least one of the components in its extremal decomposition fails to be G-invariant—if (and only if) the associated two-point function exhibits* **long-range order**, *in that*

$$\lim_{x \to \infty} \omega \left(\sum_{i=1}^{n} \phi_i(0)^* \phi_i(x) \right) > 0. \tag{10.41}$$

Proof. The "if" part of the theorem is equivalent to the vanishing of the limit in question in the *absence* of SSB. Let (9.132) be the extremal decomposition of ω. If (almost) each extreme state φ is invariant, then $\omega'(\phi_i(x)) = 0$ for all i by definition of an order parameter, and similarly $\omega'(\phi_i(x)^*) = \overline{\omega'(\phi_i(x))} = 0$. Interchanging $\lim_{x \to \infty}$ with the integral over $\partial S_\beta(A)$ (which is allowed because μ is a probability measure), and using (9.30) then shows that the left-hand side of (10.41) vanishes.

To avoid difficult measure-theoretic aspects of the extremal decomposition theory, and also for pedagogical purposes, we prove the "only if" part only in the case

$$\omega = \int_G dg \, \omega_g', \tag{10.42}$$

weakly, where $\omega' \in \partial S_\beta(A)$ and $\omega_g' = \gamma_g^* \omega'$. Since the expression

$$\omega_g'(\sum_{i=1}^{n} \phi_i(0)^* \phi_i(x))$$

is independent of $g \in G$ (by definition of an order parameter), we may replace ω_g' by ω' in the expression for ω; the term $\int_G dg$ then factors out and is equal to unity. Thus we may replace ω in (10.41) by ω'. Since ω' is a primary state, we may now use (9.30) once again, so that the left-hand side of (10.41) becomes $\sum_{i=1}^{n} |\omega'(\phi_i)|^2$. By assumption, ω' is not G-invariant, so that (by definition of a strong order parameter) at least one of the terms $|\omega'(\phi_i)|$ is nonzero. $\qquad\square$

If G is compact, for any C*-algebra A, invariant KMS states (including ground states) can always be constructed via (9.133), provided, of course, KMS states (or ground states) exist in the first place. Fortunately, existence can be shown in the following way. Let A be a quasi-local C*-algebra à la (8.130), in which:

1. $\dim(H) < \infty$ (and hence also $\dim(H_\Lambda) < \infty$ for any finite $\Lambda \subset \mathbb{Z}^d$);
2. Dynamics is defined locally on each algebra $A_\Lambda = B(H_\Lambda)$ via (9.40) and (9.41), i.e., with free boundary conditions, having a global limit α as in Theorem 9.15.

In that case, by Corollary 9.27 each C*-algebra A_Λ has a unique β-KMS state ω_Λ^β, given by the local Gibbs state (9.96). However, if $\Lambda^{(1)} \subset \Lambda^{(2)}$, then the restriction of the β-KMS state $\omega_{\Lambda^{(2)}}^\beta$ to $A_{\Lambda^{(1)}} \subset A_{\Lambda^{(2)}}$ is not given as naively expected, namely by the β-KMS state $\omega_{\Lambda^{(1)}}^\beta$, because the former involves boundary terms.

Fortunately, this complication may be overcome, since at least *for models with short-range forces* (cf. Theorem 9.15) one may put

$$\omega_G^\beta(a) = \lim_{N \to \infty} \omega_{\Lambda_N}^\beta(a), \tag{10.43}$$

where Λ_N is defined in (8.153). This limit exists for $a \in \cup_\Lambda A_\Lambda$, from which ω^β extends by continuity to all of A, on which it is a β-KMS state (cf. Theorem 10.10).

Alternatively, by the Hahn–Banach Theorem (in the form of Corollary B.41) combined with Lemma C.4 (which guarantees that any Hahn–Banach extension of a state remains a state), each local Gibbs state ω_Λ^β on $A_\Lambda \subset A$ extends, in a non-unique way, to a state $\hat{\omega}_\Lambda^\beta$ on A. This gives a net of states $(\hat{\omega}_\Lambda^\beta)$ on A indexed by the finite subsets Λ of $\mathbb{Z}^d$; one may also work with sequences $(\hat{\omega}_{\Lambda_N}^\beta)$. Since A has a unit, its state space $S(A)$ is a *compact* convex set, so the above net (or sequence) has at least one limit point, or, equivalently, has at least one convergent subnet (or subsequence), which—despite its potential lack of uniqueness in two respects, i.e. the choice of the extensions $\hat{\omega}_\Lambda^\beta$ and the choice of a limit point—one might write as

$$\hat{\omega}^\beta = \lim_{\Lambda \nearrow \mathbb{Z}^d} \hat{\omega}_\Lambda^\beta. \tag{10.44}$$

Without proof, we quote the relevant technical result (assuming 1–2 above):

Proposition 10.8. *Each limit state $\hat{\omega}^\beta$ is a β-KMS state (i.e. for the dynamics α).*

Anticipating the existence of SSB in models, one should now feel a little uneasy:

- It follows from Corollary 9.27 that (at fixed β) there is a unique KMS state on each local algebra A_Λ for the given local dynamics $\alpha_t^{(\Lambda)}$, namely the local Gibbs state ω_Λ^β on A_Λ. If—as is the case in all our examples—the globally broken G-symmetry is induced by local automorphisms $\gamma_g^{(\Lambda)} : A_\Lambda \to A_\Lambda$ that commute with the local dynamics $\alpha_t^{(\Lambda)}$, then each local Gibbs state is G-invariant: this follows explicitly from G-invariance of the local Hamiltonian h_Λ and the formulae (9.96) - (9.98), or, more abstractly, from the fact if ω_Λ^β were not invariant under all $\gamma_g^{(\Lambda)}$, it would not be unique (as its translate $\omega_\Lambda^\beta \circ \gamma_g^{(\Lambda)}$ would be another KMS state).
- And yet (in case of SSB) there exist non-invariant (and hence non-unique) KMS states on A, which are even limits in the sense of (10.44) of the above invariant (and hence unique) local KMS states on A_Λ!
- Real samples are finite and hence are described by the local algebras A_Λ, with their *unique* invariant equilibrium states ω_Λ^β. Yet finite samples do display SSB, e.g., ferromagnetism (broken $\mathbb{Z}_2$-symmetry), superconductivity (broken $U(1)$).
- Therefore, the theory that *should* describe SSB in real materials, namely the finite theory A_Λ, apparently fails to do so (as it seems to forbid SSB), whereas the idealized theory A, which describes strictly *infinite* systems and in those systems allows SSB, in fact turns out to describe key properties of *finite* samples.

10.4 Spontaneous symmetry breaking for short-range forces

We continue our discussion of SSB in quantum spin systems, especially of the construction of global KMS states in the previous section, see (10.44) and preceding text. Recall that each finite system A_Λ has a unique β-KMS state ω^β_Λ, namely the local Gibbs state (9.96), but that these states are incompatible for different Λ's, in that, if $\Lambda^{(1)} \subset \Lambda^{(2)}$, then the restriction of $\omega^\beta_{\Lambda^{(2)}}$ to $A_{\Lambda^{(1)}} \subset A_{\Lambda^{(2)}}$ is not given by $\omega^\beta_{\Lambda^{(1)}}$ because of boundary terms. To correct for this, one introduces the **surface energy**

$$b_{\Lambda^{(1)},\Lambda^{(2)}} = \sum_{X \subseteq \Lambda^{(2)}: X \cap \Lambda^{(1)} \neq \emptyset, X \cap \Lambda^c_1 \neq \emptyset} \Phi(X), \tag{10.45}$$

with ensuing **interaction energy**

$$b_\Lambda = \lim_{\Lambda^{(2)} \nearrow \mathbb{Z}^d} b_{\Lambda,\Lambda^{(2)}} = \sum_{X \cap \Lambda \neq \emptyset, X \cap \Lambda^c \neq \emptyset} \Phi(X), \tag{10.46}$$

provided this limit exists (which it does for short-range forces). Now perturb $\omega^\beta_{\Lambda^{(2)}}$ by replacing $h_{\Lambda^{(2)}}$ in (9.96) - (9.98) (with $\Lambda \rightsquigarrow \Lambda^{(2)}$) by $h_{\Lambda^{(2)}} - b_{\Lambda^{(1)},\Lambda^{(2)}}$. Denoting this modification of $\omega^\beta_{\Lambda^{(2)}}$ by $\omega^\beta_{\Lambda^{(1)},\Lambda^{(2)}}$, we obtain (10.47), which implies (10.48):

$$\omega^\beta_{\Lambda^{(1)},\Lambda^{(2)}} = \omega^\beta_{\Lambda^{(1)}} \otimes \omega^\beta_{\Lambda \setminus \Lambda^{(1)}}; \tag{10.47}$$

$$(\omega^\beta_{\Lambda^{(1)},\Lambda^{(2)}})|_{A_{\Lambda^{(1)}}} = \omega^\beta_{\Lambda^{(1)}}. \tag{10.48}$$

If (10.46) exists, we may likewise perturb any t-invariant state ω on A to $\tilde{\omega}_\Lambda$, i.e.,

$$\tilde{\omega}_\Lambda(a) = \frac{\langle e^{-\beta(h_\omega - \pi_\omega(b_\Lambda))/2}\Omega_\omega, \pi_\omega(a)e^{-\beta(h_\omega - \pi_\omega(b_\Lambda))/2}\Omega_\omega\rangle}{\|e^{-\beta(h_\omega - \pi_\omega(b_\Lambda))/2}\Omega_\omega\|^2}, \tag{10.49}$$

where $\Lambda \subset \mathbb{Z}^d$ is finite, h_ω is defined as in (9.51) - (9.52), and Ω_ω is in the domain of the unbounded operator $\exp(-\beta(h_\omega - \pi_\omega(b_\Lambda))/2)$; the reason is that $\pi_\omega(b_\Lambda)$ is bounded, whereas $\exp(-\beta h_\omega/2)\Omega_\omega = \Omega_\omega$ (since $h_\omega\Omega_\omega = 0$). For example,

$$(\tilde{\omega}^\beta_{\Lambda^{(2)}})_{\Lambda^{(1)}} = \omega^\beta_{\Lambda^{(1)},\Lambda^{(2)}}, \tag{10.50}$$

where $\omega = \omega^\beta_{\Lambda^{(2)}}$ is a Gibbs state on $A = A_{\Lambda^{(2)}}$, as in Theorem 9.24 (with $\Lambda \rightsquigarrow \Lambda^{(2)}$). Indeed, using (9.114) - (9.117) and the relation $h_\omega = h_{\Lambda^{(2)}} - Jh_{\Lambda^{(2)}}J$, where the operator J is defined in (9.124), we compute the numerator in (10.49) as

$$\text{Tr}\left(\left(\left(e^{-\beta(h_{\Lambda^{(2)}} - Jh_{\Lambda^{(2)}}J - b_\Lambda)/2}e^{-\beta h_{\Lambda^{(2)}}/2}\right)^* ae^{-\beta(h_{\Lambda^{(2)}} - Jh_{\Lambda^{(2)}}J - b_\Lambda)/2}e^{-\beta h_{\Lambda^{(2)}}/2}\right)\right)$$

$$= \text{Tr}\left(e^{-\beta(h_{\Lambda^{(2)}} - b_\Lambda)}a\right), \tag{10.51}$$

since $Jh_{\Lambda^{(2)}}J$ commutes with $h_{\Lambda^{(2)}} - b_{\Lambda}$. This subsequently gives

$$e^{-\beta(h_{\Lambda^{(2)}} - Jh_{\Lambda^{(2)}}J - b_{\Lambda})/2} = e^{-\beta(h_{\Lambda^{(2)}} - b_{\Lambda})/2} e^{\beta Jh_{\Lambda^{(2)}}J};$$
$$e^{\beta Jh_{\Lambda^{(2)}}J} e^{-\beta h_{\Lambda^{(2)}}/2} = e^{-\beta h_{\Lambda^{(2)}}/2} e^{\beta h_{\Lambda^{(2)}}/2} = 1_H. \qquad (10.52)$$

Likewise, the denominator in (10.49) equals $\mathrm{Tr}\,(\exp(-\beta(h_{\Lambda^{(2)}} - b_{\Lambda})))$.

Eqs. (10.50) and (10.48) suggest that if $\omega = \omega^{\beta}$ is a β-KMS state, then although ω^{β} itself does not localizes to a Gibbs state ω_{Λ}^{β} on A_{Λ}, its perturbed version $\tilde{\omega}_{\Lambda}^{\beta}$ does. Under assumptions 1–2 stated in §10.3, i.e., in the situation of Theorem 9.15 with $\dim(H) < \infty$, this motivates the following quantum analogue of the DLR approach to classical equilibrium states, i.e., of Definition 9.23:

Definition 10.9. *For fixed inverse temperature $\beta \in \mathbb{R}\backslash\{0\}$ and fixed interaction Φ, a **Gibbs state** ω^{β} on a quasi-local algebra A with dynamics given by some potential Φ is an α_t-independent state such that for each finite region $\Lambda \subset \mathbb{Z}^d$ one has*

$$\tilde{\omega}_{\Lambda}^{\beta} = \omega_{\Lambda}^{\beta} \otimes \omega'_{\Lambda^c}, \qquad (10.53)$$

where ω_{Λ}^{β} is the local Gibbs state (9.96) on A_{Λ} and ω'_{Λ^c} is some state on A_{Λ^c}.

Theorem 10.10. *Under assumptions 1–2 in §10.3, and if in addition the subspace $D = \cup_{\Lambda}A_{\Lambda} \subset A$ is a core for the derivation (9.54) (i.e., the closure of δ defined on D is δ as defined in Proposition 9.19), then Gibbs states coincide with KMS states.*

The proof is rather technical and so we omit it. It follows that if $\omega^{\beta} \in S_{\beta}(A)$, then

$$(\tilde{\omega}_{\Lambda}^{\beta})_{|A_{\Lambda}} = \omega_{\Lambda}^{\beta}. \qquad (10.54)$$

Even so, we still need to define in precisely which sense the net $((\tilde{\omega}_{\Lambda}^{\beta})_{|A_{\Lambda}})_{\Lambda}$ converges to ω_{Λ} (or when perhaps even the net $(\omega_{\Lambda}^{\beta})$ converges to ω_{Λ}); for simplicity we take $\Lambda = \Lambda_N$ as in (8.153), and just consider sequences indexed by N (rather than nets). To this end, let $(\omega_{1/N})_N$ be a sequence of states with $\omega_{1/N} \in S(A_{\Lambda_N})$. As in Definition 8.24, given some $\omega_0 \in S(A)$ (if it exists), we say that

$$\lim_{N \to \infty} \omega_{1/N} = \omega_0 \qquad (10.55)$$

iff for any sequence $(a_{1/N})_N$ in A with $a_{1/N} \in A_{\Lambda_N} \subset A$ that converges to $a \in A$ one has

$$\lim_{N \to \infty} \omega_{1/N}(a_{1/N}) = \omega_0(a). \qquad (10.56)$$

For example, if we take $\omega_0 \in S(A)$ and define $\omega_{1/N} = \omega_{0|A_{\Lambda_N}}$, then (10.55) holds by continuity of ω_0 (as $\|\omega_0\| = 1$), which implies that $\lim_{N \to \infty} \omega_0(a_{1/N}) = \omega_0(a)$.

It follows from the comments preceding Definition 8.24 that the above notion (10.55) - (10.56) of convergence is the same as the one given by (8.164), so that it is similar to the convergence of states we defined for the other two classes of examples of listed earlier, viz. classical mechanics (cf. §10.1) and thermodynamics.

We denote the restriction of some global KMS state ω^β (defined on A) to $A_{\Lambda_N} \subset A$ by $\omega^\beta_{1/N}$, whereas as usual we write $\omega^\beta_{\Lambda_N}$ for the unique local Gibbs state on A_{Λ_N}. Keeping Definition 8.24 and Proposition 8.25 in mind, the situation is as follows:

1. Any KMS state ω^β equals the limit ω^β_0 of its restrictions $\omega^\beta_{1/N}$ (i.e. to A_{Λ_N}).
2. Each state $\omega^\beta_{1/N}$ differs from the local Gibbs state $\omega^\beta_{\Lambda_N}$ (even if ω^β is unique).
3. The local Gibbs states $\omega^\beta_{\Lambda_N}$ typically converge to a KMS state ω^β_G, cf. (10.43).
4. In models with symmetry, this global Gibbs state ω^β_G is invariant (like the $\omega^\beta_{\Lambda_N}$).

The first claim follows from the argument given after (10.55). The second is the contrapositive to (10.54) and has been explained in §10.3: although the states $\omega^\beta_{1/N}$ and $\omega^\beta_{\Lambda_N}$ are both of local Gibbs type, their Hamiltonians differ from h_{Λ_N} by the boundary term b_Λ. The third claim cannot be proved in general, but in models with short-range forces it holds in both forms (10.43) and (10.55) - (10.56). In such models the G-symmetry is local, i.e., G acts on each A_Λ through unitaries

$$u_g^{(\Lambda)} = \otimes_{x \in \Lambda} u_g(x); \tag{10.57}$$

$$\gamma_g^{(\Lambda)}(a_\Lambda) = u_g^{(\Lambda)} a (u_g^{(\Lambda)})^* \quad (a_\Lambda \in A_\Lambda, g \in G), \tag{10.58}$$

where $u_g(x) \in B(H_x)$, leaving each local Hamiltonian h_Λ and hence each local Gibbs state $\omega^\beta_{\Lambda_N}$ invariant. If $a \in A$ is local, i.e., $a \in \cup_\Lambda A_\Lambda$, then

$$\gamma_g(a) = \lim_{N \to \infty} \gamma_g^{(\Lambda_N)}(a_N), \tag{10.59}$$

followed by continuous extension to $a \in A$, so that, assuming (10.55),

$$\omega_0(\gamma_g(a)) = \lim_{N \to \infty} \omega_{1/N}(\gamma_g(a_N)) = \lim_{N \to \infty} \omega_{1/N}(\gamma_g^{(\Lambda_N)}(a_N)) = \lim_{N \to \infty} \omega_{1/N}(a_N) = \omega_0(a),$$

since $\omega_{1/N} \circ \gamma_g^{(\Lambda_N)} = \omega_{1/N}$ by assumption. Thus the global Gibbs state ω^β_G inherits the G-invariance of its local approximants $\omega^\beta_{\Lambda_N}$. In case of SSB, the restrictions $\omega^\beta_{1/N}$ of some non-invariant extreme KMS state ω^β determine ω^β, so that in principle SSB is detectable through the local states $\omega^\beta_{1/N}$. It would be question-begging to construct the latter from the global states ω^β, though, so Butterfield's Principle (and hence in its wake Earman's Principle) holds only if we can show how and why the states of sufficiently large yet *finite* systems A_{Λ_N} tend to $\omega^\beta_{1/N}$ rather than to $\omega^\beta_{\Lambda_N}$.

Unfortunately, showing any of this in specific models at finite (inverse) temperature $0 < \beta < \infty$ is pretty complicated. For example, in the quantum Ising model (9.42) in $d = 1$, KMS states are unique for any B, so that for SSB one must go to $d \geq 2$. In that case, it can be shown from Theorem 10.7 that for $B = 0$, below some critical temperature (i.e. for $\beta > \beta_c$) the $\mathbb{Z}_2$ symmetry defined in (10.68) below is broken, but this takes considerable effort and is beyond the scope of this book.

10.5 Ground state(s) of the quantum Ising chain

It is much simpler to put $\beta = \infty$ and hence turn to the *ground state(s)* of the quantum Ising model (9.42) in $d = 1$, which is manageable. The interesting case is $B > 0$, with $J = 1$ and free boundary conditions, so that for $\Lambda = \Lambda_N$ (with N even), we have

$$h_N = -\sum_{x \in \Lambda_N} (\sigma_3(x)\sigma_3(x+1) + B\sigma_1(x)); \tag{10.60}$$

$$\Lambda_N = \{-\tfrac{1}{2}N, \ldots, \tfrac{1}{2}N - 1\}; \tag{10.61}$$

$$H_{\Lambda_N} = H_N = \otimes_{x \in \Lambda_N} H_x; \tag{10.62}$$

$$H_x = \mathbb{C}^2 \ (x \in \Lambda_N), \tag{10.63}$$

where the operator $\sigma_i(x)$ acts as the Pauli matrix σ_i on H_x and as the unit matrix 1_2 elsewhere. This model describes a chain of N immobile spin-$\frac{1}{2}$ particles with ferromagnetic coupling in a transverse magnetic field (it is a special case of the so-called XY-model, to which similar conclusions apply). The local Hamiltonians h_N define time evolution on the local algebras

$$A_{\Lambda_N} \equiv A_N = B(H_N) \tag{10.64}$$

by (9.40), i.e.,

$$\alpha_t^{(N)}(a_N) = e^{ith_N} a_N e^{-ith_N} \ (a \in A_N), \tag{10.65}$$

which by Theorem 9.15 defines a time evolution on the quasi-local C*-algebra

$$A = \overline{\bigcup_{N \in \mathbb{N}} A_N}^{\|\cdot\|} = \bigotimes_{x \in \mathbb{Z}} B(H_x), \tag{10.66}$$

namely by regarding the unitaries $\exp(ith_N) \in A_N \subset A$ as elements of A and putting

$$\alpha_t(a) = \lim_{N \to \infty} e^{ith_N} a e^{-ith_N} \ (a \in A), \tag{10.67}$$

which exists (although the sequence $(\exp(ith_N))_N$ in A does not converge in A).

For any $B \in \mathbb{R}$, the quantum Ising chain has a $\mathbb{Z}_2$-symmetry given by a 180-degree rotation around the x-axis, locally implemented by the unitary operator $u(x) = \sigma_1(x)$, which at each $x \in \Lambda_N$ yields $(\sigma_1, \sigma_2, \sigma_3) \mapsto (\sigma_1, -\sigma_2, -\sigma_3)$, since $\sigma_i \sigma_j \sigma_i^* = -\sigma_j$ for $i \neq j$. Thus $u(x)$ sends each $\sigma_3(x)$ to $-\sigma_3(x)$ but maps each $\sigma_1(x)$ to itself. As in (10.57), this symmetry is implemented by the unitary operator

$$u^{(N)} = \otimes_{x \in \Lambda_N} \sigma_1(x) \tag{10.68}$$

on H_N, which satisfies $[h_N, u^{(N)}] = 0$, or, equivalently,

$$u^{(N)} h_N (u^{(N)})^* = h_N. \tag{10.69}$$

The ensuing $\mathbb{Z}_2$-symmetry is given by the automorphism $\gamma^{(N)}$ of A_N defined by

$$\gamma^{(N)}(a) = u^{(N)}a(u^{(N)})^* \quad (a \in A_N), \tag{10.70}$$

which induces a global automorphism $\gamma \in \mathrm{Aut}(A)$ as in (10.59), i.e.,

$$\gamma(a) = \lim_{N \to \infty} u^{(N)}a(u^{(N)})^* \quad (a \in A), \tag{10.71}$$

which limit once again exists despite the fact that the sequence $u^{(N)}$ has no limit in A. Thus $\mathbb{Z}_2$-invariance of the model follows from the local property

$$\alpha_t^{(N)} \circ \gamma^{(N)} = \gamma^{(N)} \circ \alpha_t^{(N)}, \tag{10.72}$$

which in the limit $N \to \infty$ gives

$$\alpha_t \circ \gamma = \gamma \circ \alpha_t \quad (t \in \mathbb{R}). \tag{10.73}$$

Since $\gamma^2 = \mathrm{id}_A$, we have an action of the group $\mathbb{Z}_2 = \{-1, 1\}$ on A, where the nontrivial element (i.e., $g = -1$) is sent to γ. By (10.72) this group acts on the set $S_\infty(A_N)$ of ground states of A_N relative to the dynamics $\alpha^{(N)}$, and by (10.73) the same is true for the set $S_\infty(A)$ of ground states of the corresponding infinite system for α (and analogously for β-KMS states). These sets may be described as follows.

Theorem 10.11. *1. For any $N < \infty$ and $B = 0$ the ground state of the quantum Ising model* (10.60) *is doubly degenerate and breaks the $\mathbb{Z}_2$ symmetry of the model.*

2. For $N < \infty$ and any $B > 0$ the ground state $\omega_{1/N}^{(0)}$ is unique and hence $\mathbb{Z}_2$-invariant.

3. At $N = \infty$ with magnetic field $0 \leq B < 1$, the model has a doubly degenerate translation-invariant ground state $\omega_0^\pm$, which again breaks the $\mathbb{Z}_2$ symmetry.

4. At $N = \infty$ and $B \geq 1$ the ground state is unique (and hence $\mathbb{Z}_2$-invariant).

5. Recall Definition 8.24. For $0 \leq B < 1$ the states $(\omega_{1/N}^{(0)})_{N \in \mathbb{N}}$ (as in no. 2) with

$$\omega_0^{(0)} = \tfrac{1}{2}(\omega_0^+ + \omega_0^-) \tag{10.74}$$

form a continuous field of states on the continuous bundle $A^{(q)}$; in particular,

$$\lim_{N \to \infty} \omega_{1/N}^{(0)} = \omega_0^{(0)}. \tag{10.75}$$

The two ground states in no. 1 and no. 3 are tensor products of $|\uparrow\rangle$ and $|\downarrow\rangle$, respectively (where $\sigma_3|\uparrow\rangle = |\uparrow\rangle$ and $\sigma_3|\downarrow\rangle = -|\downarrow\rangle$), so that $\sigma_3(0)$ is an order parameter in the sense of Definition 10.6. In no. 4, on the other hand, each spin aligns with the magnetic field in the x-direction, so that the ground state is an infinite tensor product of states $|\rightarrow\rangle$, where $\sigma_1|\rightarrow\rangle = |\rightarrow\rangle$, and this time $\sigma_1(0)$ is an order parameter.

Case no. 2 becomes more transparent if we realize the Hilbert space H_N as $\ell^2(S_N)$, where S_N is the set of all spin configurations s on N sites, that is,

$$s : \{-\tfrac{1}{2}N, -\tfrac{1}{2}N+1, \ldots, \tfrac{1}{2}N-1\} \to \{-1, 1\}.$$

In terms of the eigenvectors $|1\rangle \equiv |\uparrow\rangle$ and $|-1\rangle \equiv |\downarrow\rangle$ of σ_3, and the orthonormal basis $(\delta_s)_{s \in S_N}$ of $\ell^2(S_N)$ (where $\delta_s(t) = \delta_{st}$), a suitable unitary equivalence

$$v_N : \ell^2(S_N) \to H_N \tag{10.76}$$

is given by linear extension of

$$v_N \delta_s = |s(-\tfrac{1}{2}N) \cdots s(\tfrac{1}{2}N-1)\rangle, \ s,t \in S_N. \tag{10.77}$$

For example, the state $|1\cdots 1\rangle$ corresponds to $\delta_{s_\uparrow}$, where $s_\uparrow(x) = 1$ for all x, and analogously $s_\downarrow(x) = -1$ for the state $|-1\cdots -1\rangle$. Using $\ell^2(S_N)$, we may talk of localization of states in spin configuration space (similar to localization of wave-functions in $L^2(\mathbb{R}^n)$), in the sense that some $\psi \in \ell^2(S_N)$ may be peaked on just a few spins configurations. Provided $0 < B < 1$ this is indeed the case for the unique ground state in case no. 2, which is similar to the ground state of the double-well potential discussed in §§10.1–10.2, replacing $\mathbb{R}$ by S_N (and $\hbar > 0$ by $1/N$).

Theorem 10.11 and related results used below, such as eq. (10.82), follow from the exact solution of the model for both $N < \infty$ and $N = \infty$, to be discussed in §§10.6–10.7. This solution is rather involved, but a rough picture of the various ground states may already be obtained from a classical approximation in the spirit of §8.1. This approximation assumes that the spin-1/2 operators $\tfrac{1}{2}\sigma_i$ are replaced by their counterparts for spin $n \cdot \tfrac{1}{2}$, upon which one takes the limit $n \to \infty$. In this limit, the spin operators are turned into the corresponding coordinate functions on the coadjoint orbit $\mathcal{O}_{1/2} \subset \mathbb{R}^3$ for $SU(2)$, which is the two-sphere $S^2_{1/2}$ with radius $r = 1/2$. In principle, this should be done for each of the N spins separately, yielding a classical Hamiltonian h_c that is a function on the N-fold cartesian product of $S^2_{1/2}$ with itself. However, if we *a priori* assume translation invariance of the classical ground state, only one such copy remains. Using spherical coordinates

$$(x = \tfrac{1}{2}\sin\theta\cos\phi, y = \tfrac{1}{2}\sin\theta\sin\phi, z = \tfrac{1}{2}\cos\theta), \tag{10.78}$$

the ensuing trial Hamiltonian becomes just a function on $\mathcal{O}_{1/2}$, given by

$$h(\theta,\phi) \approx -(\tfrac{1}{2}\cos^2\theta + B\sin\theta\cos\phi). \tag{10.79}$$

Minimizing gives $\cos\phi = 1$ and hence $y = 0$ for any B, upon which

$$h(\theta) \approx -(\tfrac{1}{2}\cos^2\theta + B\sin\theta) \tag{10.80}$$

yields the phase portrait of Theorem 10.11 for $N = \infty$, as follows. For $0 \le B < 1$, the global minimum is reached at the two different solutions $\theta_\pm$ of $\cos\theta_\pm = B$, with ensuing spin vectors

$$\mathbf{x}_\pm(B) = (\tfrac{1}{2}B, 0, \pm\tfrac{1}{2}\sqrt{1-B^2}), \tag{10.81}$$

starting at $\mathbf{x}_\pm(0) = (0,0,\pm\tfrac{1}{2})$ and merging at $B = 1$ to $\mathbf{x}_+(1) = \mathbf{x}_-(1) = (\tfrac{1}{2},0,0)$. This remains the unique ground state for $B \ge 1$, where all spins align with the field.

In the regime $0 < B < 1$ with large but finite N, one finds a far-reaching analogy between the double-well potential and and the quantum Ising chain, namely:

- The ground state of (10.60) is doubly peaked in spin configuration space, similar to its counterpart for the double-well potential in real configuration space.
- One has convergence to localized ground states (10.15) - (10.16) for the quantum Ising chain and (10.74) - (10.75) for the double well.
- For the energy difference $\Delta_N = E_N^{(1)} - E_N^{(0)}$ between the first excited state and the ground state one has (10.17) - (10.18) for the double well, and

$$\Delta_N \approx (1 - B^2)B^N \ (N \to \infty),\tag{10.82}$$

for the quantum Ising chain. Thus both models show exponential decay, i.e. of (10.82) in N as $N \to \infty$, and of (10.13) in $1/\hbar$ as $\hbar \to 0$.

It should be mentioned that *exponential* decay of the energy gap seems a low-dimensional luxury, which is not really needed for SSB. All that counts is that $\lim_{N \to \infty} \Delta_N = 0$, which guarantees that the first excited state is asymptotically degenerate with the ground state, so that appropriate linear combinations like $\omega_0^\pm$ can be formed that converge to the degenerate symmetry-breaking *pure* (and hence physical) ground states (or *extreme* and hence physical KMS states) of the limit system, which are localized and stable (as is clear from the double well). The fact that in the two models at hand only *one* excited state participates in this mechanism is due to the simple $\mathbb{Z}_2$ symmetry that is being broken; SSB of continuous symmetries requires a large number of low-lying states that are asymptotically degenerate with the ground state and hence also with each other—one speaks of a *thin* energy spectrum).

The existence of low-lying excited states may be proved abstractly (i.e., in a model-independent way), as follows. For $N < \infty$, let $\psi_N^{(0)}$ be the ground state (assumed unique) of some model defined on $\Lambda_N \subset \mathbb{Z}^d$, and let ϕ be an order parameter (cf. Theorem 10.7) with accompanying vector field $\Phi_N = \sum_{x \in \Lambda_N} \phi(x)$; in the quantum Ising chain, we take $\phi = \sigma_1$. Then the key assumptions are expressed by

$$\langle \psi_N^{(0)}, \Phi_N \psi_N^{(0)} \rangle = 0;\tag{10.83}$$

$$\langle \Phi_N \psi_N^{(0)}, \Phi_N \psi_N^{(0)} \rangle \geq C_1 \cdot N^2 \ (N \to \infty, C_1 > 0);\tag{10.84}$$

$$\|[[\Phi_N, h_N], \Phi_N]\| \leq C_2 \cdot N \ (N \to \infty, C_2 > 0).\tag{10.85}$$

The first states that the ground state is symmetric, the second enforces long-range order, as in (10.41), and the third follows from having short-range forces. A simple computation then shows that the unit vector $\tilde{\psi}_N^{(1)} = \Phi_N \psi_N^{(0)} / \|\Phi_N \psi_N^{(0)}\|$ satisfies

$$\langle \tilde{\psi}_N^{(1)}, h_N \tilde{\psi}_N^{(1)} \rangle - \langle \psi_N^{(0)}, h_N \psi_N^{(0)} \rangle \leq C_2/(C_1 N) \ (N \to \infty).\tag{10.86}$$

Since $\tilde{\psi}_N^{(1)}$ is orthogonal to $\psi_N^{(0)}$ by (10.83), the variational principle for eigenvalues (note that h_N has discrete spectrum, as $\dim(H_{\Lambda_N}) < \infty$) then gives $\Delta_N \leq C_2/(C_1 N)$, so that Δ_N vanishes as $N \to \infty$, though perhaps not as quickly as (10.82) indicates.

10.6 Exact solution of the quantum Ising chain: $N < \infty$

The solution of the quantum Ising chain is based on a transformation to fermionic variables. Let H be a Hilbert space and let $F_-(H)$ be its **fermionic Fock space**, i.e.,

$$F_-(H) = \oplus_{k=0}^\infty H_-^k, \tag{10.87}$$

where $H^0 = \mathbb{C}$, and for $k > 0$ the Hilbert space $H_-^k = e_-^{(k)} H^k$ is the totally antisymmetrized k-fold tensor product of H with itself, see also §7.7. Here the projection $e_-^{(k)} : H^k \to H^k$ is defined by linear extension of

$$e_-^{(k)} f_1 \otimes \cdots \otimes f_k = \frac{1}{k!} \sum_{p \in \mathfrak{S}_k} \mathrm{sgn}(p) f_{p(1)} \otimes \cdots \otimes f_{p(k)}, \tag{10.88}$$

where $\mathfrak{S}_k$ is the permutation group on k objects, and $\mathrm{sgn}(p)$ is $+1/-1$ if p is an even/odd permutation. With the (total) Fock space $F(H) = \oplus_{k=0}^\infty H^k$ we have $F_-(H) = e_- F(H)$, where $e = \sum_k e_-^{(k)}$ (strongly) is a projection. For $f \in H$ we define the (unbounded) **annihilation operator** $a(f)$ on $F(H)$ by (finite) linear extension of

$$a(f) f_1 \otimes \cdots \otimes f_k = \sqrt{k} \langle f, f_1 \rangle_H \otimes \cdots \otimes f_k, \tag{10.89}$$

for $k > 0$, with $a(f)z = 0$ on $H^0 = \mathbb{C}$. This gives the adjoint $a(f)^* \equiv a^*(f)$ as

$$a^*(f) f_1 \otimes \cdots \otimes f_k = \sqrt{k+1} f \otimes f_1 \otimes \cdots \otimes f_k. \tag{10.90}$$

For each $f \in H$, we then define the following operators on $F_-(H)$:

$$c(f) = e_- a(f) e_-; \tag{10.91}$$
$$c^*(f) = e_- a^*(f) e_-. \tag{10.92}$$

Note that the map $f \mapsto c(f)$ is *antilinear* in f, whereas $f \mapsto a^*(f)$ is *linear* in f. It follows that $c^*(f) = c(f)^*$, that each operator $c(f)$ and $c(f)$ on $F_-(H)$ is bounded with $\|c(f)\| = \|c^*(f)\| = \|f\|$, and the **canonical anticommutation relations** hold:

$$[c(f), c^*(g)]_+ = \langle f, g \rangle_H \cdot 1_{F_-(H)}; \tag{10.93}$$
$$[c(f), c(g)]_+ = [c^*(f), c^*(g)]_+ = 0. \tag{10.94}$$

Thus we may define $\mathrm{CAR}(H)$ as the C*-algebra within $B(F_-(H))$ generated by all $c(f)$, where $f \in H$. This is called the C*-algebra of **canonical anticommutation relations** over H, which have constructed in its defining representation on $F_-(H)$. Choosing an orthonormal basis (e_i) of H and writing $c(e_i) = c_i$ etc. clearly yields

$$[c_i, c_j^*]_+ = \delta_{ij} \cdot 1_{F_-(H)}; \tag{10.95}$$
$$[c_i, c_j]_+ = [c_i^*, c_j^*]_+ = 0. \tag{10.96}$$

If $\dim(H) = N < \infty$, then $\mathrm{CAR}(H) = B(F_-(H))$. First, a dimension count yields

$$F_-(\mathbb{C}^N) = \oplus_{k=0}^N H_-^k \cong \mathbb{C}^{2^N} \cong \otimes^N \mathbb{C}^2. \tag{10.97}$$

By Theorem C.90, the C*-algebra $\mathrm{CAR}(H)$ acts irreducibly on $F_-(H)$, so that

$$\mathrm{CAR}(\mathbb{C}^N) \cong M_{2^N}(\mathbb{C}). \tag{10.98}$$

This is already nontrivial for $N = 1$. In that case, $F_-(\mathbb{C}) = \mathbb{C} \oplus \mathbb{C} = \mathbb{C}^2$, and

$$c = \sigma_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}; \tag{10.99}$$

$$c^* = \sigma_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \tag{10.100}$$

where $\sigma_\pm = \frac{1}{2}(\sigma_1 \pm i\sigma_2)$. This realization explicitly shows that

$$\mathrm{CAR}(\mathbb{C}) = M_2(\mathbb{C}). \tag{10.101}$$

To generalize this to $N > 1$, we introduce a lattice (or chain) $\underline{N} = \{1, \ldots, N\}$, and for each $x \in \underline{N}$ we define operators c_x, c_x^* by the **Jordan–Wigner transformation**

$$c_x = e^{\pi i \sum_{y=1}^{x-1} \sigma_+(y)\sigma_-(y)} \sigma_-(x) = \left(\prod_{y=1}^{x-1} (-\sigma_3)(y) \right) \cdot \sigma_-(x); \tag{10.102}$$

$$c_x^* = e^{-\pi i \sum_{y=1}^{x-1} \sigma_+(y)\sigma_-(y)} \sigma_+(x) = \left(\prod_{y=1}^{x-1} (-\sigma_3)(y) \right) \cdot \sigma_+(x), \tag{10.103}$$

where $x > 1$, and $c_1 = \sigma_1^-$ and $c_1^* = \sigma_1^+$ (here $\sigma_\pm(x) = \frac{1}{2}(\sigma_1(x) \pm i\sigma_2(x))$ etc.). These operators satisfy (10.95) - (10.96); the second expression on each line follows because the operators $\sigma_+(y)\sigma_-(y)$ commute for different sites y, and

$$e^{\pi i \sigma^+ \sigma^-} = -\sigma_3. \tag{10.104}$$

Furthermore, since

$$c_x^* c_x = \sigma_+(x)\sigma_-(x) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}(x); \tag{10.105}$$

$$c_x c_x^* = \sigma_-(x)\sigma_+(x)) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}(x), \tag{10.106}$$

the inverse of the Jordan–Wigner transformation is given by

$$\sigma_-(x) = e^{-\pi i \sum_{y=1}^{x-1} c_y^* c_y} c_x; \tag{10.107}$$

$$\sigma_+(x) = c_x^* e^{\pi i \sum_{y=1}^{x-1} c_y^* c_y}. \tag{10.108}$$

We return to the quantum Ising model (10.60) with free boundary conditions, where we relabel the sites as $\{1,\dots,N\}$, as above, and change to the Hamiltonian

$$h_N^{\mathrm{QI}} = -\tfrac{1}{2}\left(\sum_{x=1}^{N-1}\sigma_1(x)\sigma_1(x+1) + \lambda\sum_{x=1}^{N}\sigma_3(x)\right), \qquad (10.109)$$

where, in order to avoid notational confusion with the operator B in (10.111) below, we henceforth replace $B \rightsquigarrow \lambda$. In terms of the unitary operator $u = \sqrt{1/2}(1_2 + i\sigma_2)$ on $\mathbb{C}^2$ and hence $u^{(N)} = \otimes_{x=1}^{N}u(x)$ on $\otimes^N\mathbb{C}^2$, we have $u^{(N)}h_N(u^{(N)})^* = h_N'$.

Using (10.102) - (10.103), up to an additive constant $\lambda N \cdot 1_N$ we omit, we find

$$h_N^{\mathrm{QI}} = -\sum_{x=1}^{N}(\lambda c_x^* c_x + \tfrac{1}{2}(c_x^* - c_x)(c_{x+1}^* + c_{x+1})), \qquad (10.110)$$

so we now show how to diagonalize quadratic fermionic Hamiltonians of the type

$$h_N = -\sum_{x,y=1}^{N}\left(A_{xy}c_x^* c_y + \tfrac{1}{2}B_{xy}(c_x^* c_y^* - c_x c_y)\right), \qquad (10.111)$$

where A and B are real $N \times N$ matrices, with $A^* = A$ and $B^* = -B$. Indeed, taking

$$A = \tfrac{1}{2}(S + S^*) + \lambda \cdot 1_N; \qquad (10.112)$$
$$B = \tfrac{1}{2}(S - S^*), \qquad (10.113)$$

recovers (10.110), where $S : \mathbb{C}^N \to \mathbb{C}^N$ is the **shift operator**, defined by

$$Sf(x) = f(x+1); \qquad (10.114)$$
$$S^*f(x) = f(x-1). \qquad (10.115)$$

By convention, $f(N+1) = f(0) = 0$ (i.e., $Sf(N) = S^*f(0) = 0$ for any $f \in \mathbb{C}^N$); in terms of the standard basis (v_x) of $\mathbb{C}^N$ we have $Sv_1 = 0$ and $Sv_x = v_{x-1}$ for $x = \{2,\dots,N\}$, and likewise $S^*v_N = 0$ and $Sv_x = v_{x+1}$ for $x = \{1,\dots,N-1\}$.

The smart thing to do now turns out to be diagonalizing the $2N \times 2N$-matrix

$$M = \begin{pmatrix} A & B \\ -B & -A \end{pmatrix}, \qquad (10.116)$$

which by a unitary transformation may be brought into the simpler form

$$M' = \begin{pmatrix} \sqrt{1/2} & -\sqrt{1/2} \\ \sqrt{1/2} & \sqrt{1/2} \end{pmatrix}\begin{pmatrix} A & B \\ -B & -A \end{pmatrix}\begin{pmatrix} \sqrt{1/2} & \sqrt{1/2} \\ -\sqrt{1/2} & \sqrt{1/2} \end{pmatrix} = \begin{pmatrix} 0 & C \\ C^* & 0 \end{pmatrix}, \qquad (10.117)$$

where $C = A + B$. For example, for the model (10.111) we simply have

$$C = S + \lambda \cdot 1_N. \qquad (10.118)$$

The equations for the eigenvalues ε_k and eigenvectors of M', i.e.,

$$M' \begin{pmatrix} \varphi_k \\ \psi_k \end{pmatrix} = \varepsilon_k \begin{pmatrix} \varphi_k \\ \psi_k \end{pmatrix} \tag{10.119}$$

where $\varphi_k, \psi_k \in \mathbb{C}^N$, are equivalent to both the coupled system of equations

$$C \psi_k = \varepsilon_k \varphi_k; \tag{10.120}$$
$$C^* \varphi_k = \varepsilon_k \psi_k; \tag{10.121}$$
$$C = A + B, \tag{10.122}$$

where the eigenvalues ε_k are real (since $M^* = M$), and to the uncoupled version

$$CC^* \varphi_k = \varepsilon_k^2 \varphi_k; \tag{10.123}$$
$$C^* C \psi_k = \varepsilon_k^2 \psi_k; \tag{10.124}$$
$$CC^* = A^2 - B^2 - [A, B]; \tag{10.125}$$
$$C^* C = A^2 - B^2 + [A, B]. \tag{10.126}$$

Without loss of generality we may (and will) assume that the φ_k, ψ_k are unit vectors in $\mathbb{C}^N$, so that the corresponding unit vector in $\mathbb{C}^{2N}$ is $(\varphi_k, \psi_k)/\sqrt{2}$. Furthermore, since C (or M) is a matrix with real entries and the ε_k are real, by a suitable choice of phase we may (and will) also arrange that φ_k, ψ_k have real components. Finally, it follows from (10.120) - (10.120) that $(-\varphi_k, \psi_k)$ is an eigenvector of C with eigenvalue $-\varepsilon_k$, so that the unitary transformation U' that diagonalizes M', i.e.,

$$(U')^{-1} M' U' = \begin{pmatrix} -E & 0 \\ 0 & E \end{pmatrix}, \tag{10.127}$$

where $E = \mathrm{diag}(\varepsilon_1, \ldots, \varepsilon_N)$, takes the form

$$U' = \frac{1}{\sqrt{2}} \begin{pmatrix} \varphi & -\varphi \\ \psi & \psi \end{pmatrix}, \tag{10.128}$$

where φ is the $N \times N$ matrix $(\varphi_1, \ldots, \varphi_N)$, seeing each vector φ_i as a column, etc. Combined with (10.117), we obtain

$$U^{-1} M U = \begin{pmatrix} -E & 0 \\ 0 & E \end{pmatrix}; \tag{10.129}$$

$$U = \tfrac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} \varphi & -\varphi \\ \psi & \psi \end{pmatrix} = \tfrac{1}{2} \begin{pmatrix} \psi + \varphi & \psi - \varphi \\ \psi - \varphi & \psi + \varphi \end{pmatrix} \equiv \begin{pmatrix} u & v \\ v & u \end{pmatrix}, \tag{10.130}$$

where we introduced $N \times N$ matrices

$$u = \tfrac{1}{2}(\psi + \varphi); \tag{10.131}$$
$$v = \tfrac{1}{2}(\psi - \varphi). \tag{10.132}$$

Using orthonormality and completeness of both the (φ_k) and the (ψ_k), one obtains

$$u^*u + v^*v = 1_H; \tag{10.133}$$
$$u^*v + v^*u = 0; \tag{10.134}$$
$$uu^* + vv^* = 1_H; \tag{10.135}$$
$$uv^* + vu^* = 0. \tag{10.136}$$

Of course, u and v are far from unique, as they depend on both the ordering and the phases of the vectors φ_k and ψ_k. In partial remedy of the former ambiguity we assume that $0 \leq \varepsilon_0 \leq \varepsilon_1 \leq \cdots \leq \varepsilon_N$ (which can be arranged by a suitable ordering as well as choice of sign of the eigenvectors φ_k). Towards the latter, we already agreed that both the φ_k and ψ_k are real, so that also our matrices u and v have real entries.

We now explain the purpose of diagonalizing M in (10.116) using u and v.

Proposition 10.12. *Let* u *and* v *be operators on a Hilbert space H, where* u *is* linear *and* v *is* anti-linear. *Let $c(f)$ and $c^*(f)$ be the operators (10.91) - (10.92), satisfying the CAR (10.93) - (10.94). Define the* **Bogoliubov transformation**

$$\eta(f) = c(uf) + c^*(vf); \tag{10.137}$$
$$\eta^*(f) = c^*(uf) + c(vf), \tag{10.138}$$

which extends to a linear map $\alpha : \mathrm{CAR}(H) \to \mathrm{CAR}(H)$, where $\eta(f) = \alpha(c(f))$ etc. Then α is a homomorphism of C*-algebras, or, equivalently, one has the CAR

$$[\eta(f), \eta^*(g)]_+ = \langle f, g \rangle_H \cdot 1_H; \tag{10.139}$$
$$[\eta(f), \eta(g)]_+ = [\eta^*(f), \eta^*(g)]_+ = 0, \tag{10.140}$$

iff u *and* v *satisfy (10.133) - (10.134), with $u \rightsquigarrow$ u, $v \rightsquigarrow$ v. Moreover, α is invertible (and hence defines an automorphism of $\mathrm{CAR}(H)$) iff in addition (10.135) - (10.136) are valid (again with with $u \rightsquigarrow$ u, $v \rightsquigarrow$ v), in which case the inverse is*

$$c(f) = \eta(u^*f) + \eta^*(v^*f); \tag{10.141}$$
$$c^*(f) = \eta^*(u^*f) + \eta(v^*f). \tag{10.142}$$

Note that anti-linearity of v is needed to make $f \mapsto \eta(f)$ anti-linear, like $f \mapsto c(f)$. With respect to a base (e_i) of H, the transformations (10.137) - (10.142) reads

$$\eta_i = \sum_j (\bar{u}_{ji} c_j + v_{ji} c_j^*); \tag{10.143}$$

$$\eta_j^* = \sum_j (u_{ji} c_j^* + \bar{v}_{ji} c_j); \tag{10.144}$$

$$c_i = \sum_j (u_{ij} \eta_j + \bar{v}_{ij} \eta_j^*); \tag{10.145}$$

$$c_i^* = \sum_j (\bar{u}_{ij} \eta_j^* + v_{ij} \eta_j). \tag{10.146}$$

Proof. The proof is a straightforward computation. □

In comparison with the preceding diagonalization process, where $H = \mathbb{C}^N$, we notice that in this process u and v were both linear, whereas in Proposition 10.12 u is linear whereas v is antilinear. This difference is easily overcome by taking $\mathsf{u} = u$ and $\mathsf{v} = Jv$, where $J : \mathbb{C}^N \to \mathbb{C}^N$ is the anti-linear map $Jf(x) = \overline{f(x)}$, so that J is a ***conjugation*** in being an anti-linear map that satisfies $J^* = J^{-1} = J$.

Returning to our generic Hamiltonian (10.111), a straightforward computation using (10.145) - (10.146), (10.116), (10.129), and (10.133) - (10.136) yields

$$h_N = \sum_{k=1}^N \varepsilon_k \eta_k^* \eta_k, \tag{10.147}$$

up to a (computable) constant, where we recall that $\varepsilon_k \geq 0$ $(k = 1, \ldots, N)$. Note that h_N is still defined on the fermionic Fock space $F_-(\mathbb{C}^N)$, as h_N is a (complicated) *quadratic* expression in the operators c_i and c_i^* on $F_-(\mathbb{C}^N)$. The point is that (as a consequence of Proposition 10.12) the η_k and η_k^* also satisfy the CAR, i.e.,

$$[\eta_i, \eta_j^*]_+ = \delta_{ij} \cdot 1_{F_-(H)}; \tag{10.148}$$
$$[\eta_i, \eta_j]_+ = [\eta_i^*, \eta_j^*]_+ = 0. \tag{10.149}$$

Theorem 10.13. *Let $A = \mathrm{CAR}(\mathbb{C}^N)$ be the CAR-algebra over $H = \mathbb{C}^N$ with dynamics $\alpha_t(a) = e^{ith_N} a e^{-ith_N}$ given by (10.111) and hence by(10.147). Then α has a unique (and hence pure and symmetric) ground state ω_0, specified by the property*

$$\pi_{\omega_0}(\eta(f))\Omega_{\omega_0} = 0 \ (f \in \mathbb{C}^N). \tag{10.150}$$

Proof. Recall that α defines a derivation $\delta : \mathrm{CAR}(\mathbb{C}^N) \to \mathrm{CAR}(\mathbb{C}^N)$ defined by (9.54), which in the case at hand is simply by $\delta(a) = i[h_N, a]$ (since A is finite-dimensional, δ is bounded and hence defined everywhere). Using the identity

$$[ab, c] = a[b, c]_+ - [c, a]_+ b, \tag{10.151}$$

as well as the relations (10.148) - (10.149), we obtain $\delta(\eta_k) = -i\varepsilon_k \eta_k$, and hence

$$-i\omega_0(\eta_k^* \delta(\eta_k)) = -\omega_0(\eta_k^* \eta_k). \tag{10.152}$$

The condition $-i\omega_0(a^* \delta(a)) \geq 0$, i.e., eq. (9.56) from Proposition (9.20), therefore implies that $\omega_0(\eta_k^* \eta_k) \leq 0$, and hence $\omega_0(\eta_k^* \eta_k) = 0$ by positivity of ω_0. Since $F_0(H)$ is finite-dimensional and $A \cong B(F_0(H))$, cf. (10.98), we may assume ground state(s) to be pure and normal, i.e., there is some unit vector $\psi_0 \in F_-(H)$ with $\omega(a) = \langle \psi_0, a\psi_0 \rangle$ for each $a \in A$. Hence $\langle \psi_0, \eta_k^* \eta_k \psi_0 \rangle = 0$, which enforces

$$\eta_k \psi_0 = 0 \ (k = 1, \ldots, N). \tag{10.153}$$

This property makes ψ_0 unique up to a phase. Indeed, together with (10.148) - (10.149), eq. (10.153) implies the values of all one- and two-point functions, i.e.,

$$\omega_0(\eta(f)) = \omega_0(\eta^*(f)) = 0; \tag{10.154}$$

$$\omega_0(\eta^*(f)\eta(g)) = \omega_0(\eta^*(f)\eta^*(g)) = \omega_0(\eta(f)\eta(g)) = 0; \tag{10.155}$$

$$\omega_0(\eta(f)\eta^*(g)) = \langle f,g\rangle_H. \tag{10.156}$$

Furthermore, the value of ω_0 on any product of an odd number of $\eta(f)$ and $\eta^*(g)$ vanishes; for an even number the value $\omega_0(\prod_{i=1}^n \eta(f_i)\prod_{j=1}^n \eta^*(g_j))$ it is given by

$$\sum_{p=1}^n (-1)^{n-p}\omega_0(\eta(f_1)\eta^*(g_p)\omega_0\left(\prod_{i=2}^n \eta(f_i)\prod_{j=1,j\neq p}^n \eta^*(g_j)\right).$$

Hence (10.153) gives ω_0 on all of $\mathrm{CAR}(\mathbb{C}^N)$. Since $\mathrm{CAR}(H) = B(F_-(H))$, this fixes ψ_0 up to a phase. Eq. (10.150) is just a fancy way of rewriting (10.153).   □

By construction, the ground state energy of (10.147) is zero. In connection with our approach to SSB via Butterfield's Principle it is of interest to compute the energy ε_1 of the first excited state. This may be done from (10.120) - (10.121) with (10.122) and the specific expression (10.118) for the quantum Ising chain. Thus we solve

$$\lambda\psi_k(x) + \psi_k(x+1) = \varepsilon_k\varphi_k(x) \quad (x=1,\dots,N, \psi_k(N+1)=0); \tag{10.157}$$

$$\lambda\varphi_k(x) + \varphi_k(x-1) = \varepsilon_k\psi_k(x) \quad (x=1,\dots,N, \varphi_k(0)=0). \tag{10.158}$$

A solution of this system (with real wave-functions and positive energy) is given by

$$\varphi_k(x) = C(-1)^k \sin(q_k(x-N-1)); \tag{10.159}$$

$$\psi_k(x) = -C\sin(q_k x); \tag{10.160}$$

$$\varepsilon_k = \sqrt{1+\lambda^2+2\lambda\cos(q_k)}, \tag{10.161}$$

where $C > 0$ is a normalization constant, and q_k should be solved from

$$(N+1)q_k = (k-1)\pi + \arctan\left(\frac{\sin q_k}{\lambda + \cos q_k}\right). \tag{10.162}$$

For example, for $\lambda = 0$ (i.e. no transverse magnetic field) we obtain $q_k = k\pi/N$, where $k = 1,\dots,N$. For $\lambda > 1$ there is a unique real solution q_k for each k, too, and even as $N \to \infty$ there is an energy gap $\varepsilon_k > 0$ for each k. For $0 < \lambda < 1$, however, there is a *complex* solution $q_1 = \pi + i\rho$, where $\rho \in \mathbb{R}$ is a solution to

$$\tanh((N+1)\rho) = \frac{\sinh\rho}{\cosh\rho - \lambda}. \tag{10.163}$$

As $N \to \infty$, we find $\rho = -\ln(\lambda) - (1-\lambda^2)\lambda^{2(N-1)}$. Eq. (10.161) then gives

$$\varepsilon(q_1) \approx (1-\lambda^2)\lambda^N \quad (N \to \infty), \tag{10.164}$$

which, recalling that $E_N^{(1)} = \varepsilon_1$ and $E_N^{(0)} = 0$ and hence $\Delta_N = \varepsilon_1$, confirms (10.82).

10.7 Exact solution of the quantum Ising chain: $N = \infty$

The (two-sided) infinite quantum Ising chain is described by the C*-algebra

$$F = \mathrm{CAR}(\ell^2(\mathbb{Z})); \tag{10.165}$$

one may also consider a one-sided chain, but it lacks translation symmetry. By the construction at the beginning of the previous section, F is isomorphic to the infinite tensor product $A = M_2(\mathbb{C})^\infty$. We consider F to be generated by the operators $c_x^\pm$ ($x \in \mathbb{Z}$), where $c_x^- \equiv c_x$ and $c_x^+ \equiv c_x^*$. In this notation, the CAR (10.95) - (10.96) read

$$[c_x^\pm, c_y^\mp]_+ = \delta_{xy}; \tag{10.166}$$

$$[c_x^\pm, c_y^\pm]_+ = 0. \tag{10.167}$$

Although the local Hamiltonians (10.111) do not have a limit as $N \to \infty$, as explained in §10.5 they do generate a time-evolution on F in the sense of a continuous homomorphism $\alpha : \mathbb{R} \to \mathrm{Aut}(F)$ via (10.65) and (10.67); see also Theorem 9.15.

Let us first extend the approach in the previous section to $N = \infty$, in which case $\mathbb{C}^N$ is replaced by $H = \ell^2(\mathbb{Z})$, assuming the theory has already been brought into fermionic form with local Hamiltonians (10.111) (as we will see, it is this step, i.e., the Jordan–Wigner transformation, that marks the difference between $N < \infty$ and $N = \infty$). Thus we define operators $A : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ and $B : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ as the obvious extensions of the $N \times N$ matrices A and B to operators on $\ell^2(\mathbb{Z})$, and similarly $S : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ is the "full" shift operator, defined by $(Sf)(x) = f(x+1)$. Instead of the somewhat clumsy explicit solution procedure sketched in the previous section for $N < \infty$, we may now simply rely on the Fourier transformation

$$\mathscr{F} : \ell^2(\mathbb{Z}) \to L^2([-\pi, \pi]); \tag{10.168}$$

$$(\mathscr{F}f)(k) \equiv \hat{f}(k) = \sum_{x \in \mathbb{Z}} e^{-ikx} f_j; \tag{10.169}$$

$$(\mathscr{F}^{-1}\hat{f})(x) \equiv f(x) = \int_{-\pi}^{\pi} \frac{dk}{2\pi} e^{ikx} \hat{f}(k), \tag{10.170}$$

which diagonalizes A and B to operators $\hat{A}, \hat{B} : L^2([-\pi, \pi]) \to L^2([-\pi, \pi])$. For the quantum Ising Hamiltonian (10.110) these are given by the multiplication operators

$$\hat{A}\hat{\psi}(k) = -(\cos k + \lambda)\hat{\psi}(k); \tag{10.171}$$

$$\hat{B}\hat{\psi}(k) = -i\sin k\,\hat{\psi}(k). \tag{10.172}$$

For fixed k, the eigenvalues and eigenvectors of the 2×2 matrix

$$M_k = \begin{pmatrix} -(\cos k + \lambda) & -i\sin k \\ i\sin k & \cos k + \lambda \end{pmatrix}, \tag{10.173}$$

are $\pm\varepsilon_k$, given by (10.161) with $q_k \rightsquigarrow k$. It is then routine to find a unitary 2×2 matrix $U_k = \begin{pmatrix} u_k & v_k \\ v_k & u_k \end{pmatrix}$ that diagonalizes M_k in the sense that $U_k^{-1} M_k U_k = \begin{pmatrix} -\varepsilon_k & 0 \\ 0 & \varepsilon_k \end{pmatrix}$.
Fourier transforming these multiplication operators back to $\ell^2(\mathbb{Z})$ then yields an operator U on $\ell^2(\mathbb{Z}) \oplus \ell^2(\mathbb{Z})$ that satisfies (10.129). This yields a unique ground state ω_0 characterized by a property like (10.150) or (10.153), where

$$\eta(\hat{f}) = \int_{-\pi}^{\pi} \frac{dk}{2\pi} \hat{f}(k)(u_k \hat{c}_k + v_k \hat{c}_{-k}^*); \tag{10.174}$$

$$\hat{c}_k = \sum_{j \in \mathbb{Z}} e^{-ijk} c_j; \tag{10.175}$$

$$\hat{c}_k^* = \sum_{j \in \mathbb{Z}} e^{ijk} c_j^*. \tag{10.176}$$

In summary, one-dimensional fermionic models with quadratic Hamiltonians like (10.111) have a unique ground state even at $N = \infty$. Thus one wonders where SSB in the quantum Ising chain could possibly come from. We will answer this question.

Almost every argument to follow relies on $\mathbb{Z}_2$-symmetry. In general, a $\mathbb{Z}_2$-action on a C*-algebra A corresponds to an automorphism $\theta : A \to A$ such that $\theta^2 = \mathrm{id}_A$, i.e. θ represents the nontrivial element of $\mathbb{Z}_2$. For example, define $\theta : F \to F$ by

$$\theta(c_x^{\pm}) = -c_x^{\pm} \quad (j \in \mathbb{Z}), \tag{10.177}$$

which is an example of a Bogoliubov transformation (cf. Proposition 10.12) and hence extends to an automorphism of F (which implies that $\theta(1_F) = 1_F$). Clearly, $\theta^2 = \mathrm{id}_F$, and in addition each local Hamiltonian (10.111) is invariant under θ; by implication, so is the dynamics α, i.e., $\alpha_t \circ \theta = \theta \circ \alpha_t$ for all $t \in \mathbb{R}$.

A C*-algebra A carrying a $\mathbb{Z}_2$-action decomposes as

$$A = A_+ \oplus A_-; \tag{10.178}$$
$$A_{\pm} = \{a \in A \mid \theta(a) = \pm a\}, \tag{10.179}$$

where the ***even*** part A_+ is a subalgebra of A, whereas the ***odd*** part A_- is not: one has $ab \in A_+$ for a, b both in either A_+ or A_-, and $ab \in A_-$ if one is in A_+ and the other in A_-. For example, if $A = B(H)$ for some Hilbert space H and $w : H \to H$ is a unitary operator satisfying $w^2 = 1$ (and hence $w^* = w$), then

$$\theta(a) = waw^* (= waw) \tag{10.180}$$

defines a $\mathbb{Z}_2$-action on A. In that case, A_+ and A_- consist of all $a \in A$ that commute and anticommute with w, respectively, that is,

$$A_{\pm} = \{a \in A \mid aw \mp wa = 0\}. \tag{10.181}$$

In case of (10.165) with (10.177), the subspace F_+ (F_-) is just the linear span of all products of an even (odd) number of $c_j^{\pm}$'s.

Let us move to Theorem C.90 and reconsider the proof of the claim that if $\pi_\omega(A)' \neq \mathbb{C} \cdot 1$, then ω is mixed. If the commutant $\pi_\omega(A)'$ is nontrivial, then it contains a nontrivial projection $e_+ \in \pi_\omega(A)'$. It then follows that $e_+ \Omega_\omega \neq 0$: for if $e_+ \Omega_\omega = 0$, then $ae_+ \Omega_\omega = e_+ a\Omega_\omega = 0$ for all $a \in A$, so that $e_+ = 0$, since π_ω is cyclic. Similarly, $e_- \Omega_\omega \neq 0$ with $e_- = 1_H - e_+$, so we may define the unit vectors

$$\Omega_\pm = e_\pm \Omega_\omega / \|e_\pm \Omega_\omega\|, \tag{10.182}$$

and the associated states $\omega_\pm(a) = \langle \Omega_\pm, \pi_\omega(a)\Omega_\pm \rangle$ on A. This yields a convex decomposition $\omega = \lambda \omega_+ + (1-\lambda)\omega_-$, with $\lambda = \|\Omega_-\|^2$. Since $\lambda \neq 0, 1$ and $\omega_+ \neq \omega_-$, it follows that ω is mixed. The associated reduction is effected by writing

$$H = H_+ \oplus H_-; \tag{10.183}$$
$$H_\pm = e_\pm H, \tag{10.184}$$

in that A (more precisely, $\pi_\omega(A)$) maps each subspace $H_\pm$ into itself. Now pass from the projections $e_\pm$ to the operator $w = e_+ - e_-$, which by construction satisfies

$$w^* = w^{-1} = w. \tag{10.185}$$

In particular, w is unitary. Conversely, if some unitary w satisfies $w^2 = 1_H$, then

$$e_\pm = \tfrac{1}{2}(1_H \pm w) \tag{10.186}$$

are projections satisfying $e_+ + e_- = 1_H$, giving rise to the decomposition (10.184). Group-theoretically, this means that one has a unitary $\mathbb{Z}_2$-action on $H \equiv H_\omega$, in which the nontrivial element of $\mathbb{Z}_2 = \{-1, 1\}$ is represented by w. The decomposition (10.184) then simply means that $\mathbb{Z}_2$ acts trivially on H_+ (in that both group elements are represented by the unit operator) and acts nontrivially on H_- (in that the nontrivial element is represented by *minus* the unit operator). In conclusion, one has a $\mathbb{Z}_2$ perspective on the reduction of H_ω, and instead of a projection $e \in \pi_\omega(A)'$ one may equivalently look for an operator $w \in \pi_\omega(A)'$ that satisfies (10.185).

Proposition 10.14. *Suppose A carries a $\mathbb{Z}_2$-action θ and consider a state $\omega : A \to \mathbb{C}$ that is $\mathbb{Z}_2$-invariant in the sense that $\omega(\theta(a)) = \omega(a)$ for all $a \in A$. We write this as $\theta^* \omega = \omega$, with $\theta^* \omega = \omega \circ \theta$. Then there is a unitary operator $w : H_\omega \to H_\omega$ satisfying $w^2 = 1_H$, $w\Omega = \Omega$, and and $w\pi_\omega(a)w^* = \pi_\omega(\theta(a))$ for each $a \in A$.*

Cf. Corollary 9.12. In this situation, we obtain a decomposition of $H \equiv H_\omega$ according to (10.183), where the projections $e_\pm$ are given by (10.186), so that, equivalently,

$$H_\pm = \{\psi \in H \mid w\psi = \pm\psi\} = A_\pm \Omega_-. \tag{10.187}$$

In terms of the decomposition (10.178), it is easily seen that each subspace $H_\pm$ is stable under A_+, whereas A_- maps $H_\pm$ into $H_\mp$. We denote the restriction of $\pi_\omega(A_+)$ to $H_\pm$ by $\pi_\pm$, so that a $\mathbb{Z}_2$-invariant state θ on A not just gives rise to the GNS-representation π_ω of A on H_ω, but also induces two representations $\pi_\pm$ of the even part A_+ on $H_\pm$. This leads to a refinement of Theorem C.90:

Theorem 10.15. *Suppose A carries a $\mathbb{Z}_2$-action θ, and let $\omega : A \to \mathbb{C}$ be a $\mathbb{Z}_2$-invariant state. With the above notation, suppose the representation $\pi_+(A_+)$ on H_+ is irreducible. Then also the representation $\pi_-(A_+)$ on H_- is irreducible, and there are the following two possibilities for the representation $\pi_\omega(A)$ on $H = H_+ \oplus H_-$:*

1. *$\pi_\omega(A)$ is irreducible (and ω is pure) iff $\pi_+(A_+)$ and $\pi_-(A_+)$ are inequivalent.*
2. *$\pi_\omega(A)$ is reducible (and ω is mixed) iff $\pi_+(A_+)$ and $\pi_-(A_+)$ are equivalent.*

Proof. The proof of this theorem is much more difficult than one would expect (given its simple statement), so we restrict ourselves to the easy steps, as well as to two examples illustrating each of the two possibilities. To start with the latter:

1. $A = M_2(\mathbb{C})$, with $\theta(a) = \sigma_3 a \sigma_3$; note that $\sigma_3^2 = 1$ and $\sigma_3^* = \sigma_3$. Then

$$A_+ = \left\{ \begin{pmatrix} z_+ & 0 \\ 0 & z_- \end{pmatrix}, z_\pm \in \mathbb{C} \right\} \equiv D_2(\mathbb{C}); \qquad (10.188)$$

$$A_- = \left\{ \begin{pmatrix} 0 & z_1 \\ z_2 & 0 \end{pmatrix}, z_1, z_2 \in \mathbb{C} \right\}, \qquad (10.189)$$

where $D_n(\mathbb{C})$ denotes the C*-algebra of diagonal $n \times n$ matrices. Take $\Omega = (1,0)$, with associated state

$$\omega(a) = \langle \Omega, a\Omega \rangle, \qquad (10.190)$$

where $a \in M_2(\mathbb{C})$. It follows from §2.4 that the associated GNS-representation $\pi_\omega(A)$ is just (equivalent to) the defining representation of $M_2(\mathbb{C})$ on $H_\omega = \mathbb{C}^2$, in which the cyclic vector Ω_ω of the GNS-construction is Ω itself. Since $\sigma_3 \Omega = \Omega$, the state defined by (10.190) is $\mathbb{Z}_2$-invariant, and the unitary operator w in Proposition 10.14 is simply $w = \sigma_3$. Hence the decomposition (10.183) of $H = \mathbb{C}^2$ is simply $\mathbb{C}^2 = \mathbb{C} \oplus \mathbb{C}$, i.e.,

$$H_+ = \{(z,0), z \in \mathbb{C}\}; \qquad (10.191)$$
$$H_- = \{(0,z), z \in \mathbb{C}\}. \qquad (10.192)$$

Of course, we then have $H_\pm = A_\pm \Omega$. Identifying $H_\pm \cong \mathbb{C}$, this gives the one-dimensional representations $\pi_\pm(D_2(\mathbb{C}))$ as

$$\pi_\pm \begin{pmatrix} z_+ & 0 \\ 0 & z_- \end{pmatrix} = z_\pm, \qquad (10.193)$$

which are trivially inequivalent. Hence by Theorem 10.15 the defining representation of $M_2(\mathbb{C})$ on $\mathbb{C}^2$ is irreducible, as it should be.

2. $A = D_2(\mathbb{C})$, with

$$\theta(\mathrm{diag}(z_+, z_-)) = \mathrm{diag}(z_-, z_+), \qquad (10.194)$$

where we have denoted the matrix in (10.188) by $\mathrm{diag}(z_+, z_-)$. This time,

$$A_\pm = \{\mathrm{diag}(z, \pm z), z \in \mathbb{C}\}. \qquad (10.195)$$

We once again define a $\mathbb{Z}_2$-invariant state ω by (10.190), but this time we take

$$\Omega = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{10.196}$$

Hence

$$H_\pm = \{(z, \pm z), z \in \mathbb{C}\}.. \tag{10.197}$$

We may now identify each $A_\pm$ with $\mathbb{C}$ under the map $\mathrm{diag}(z, \pm z) \mapsto z$ from $A_\pm$ to $\mathbb{C}$. Similarly, we identify each each subspace $H_\pm$ with $\mathbb{C}$ under the map $H_\pm \to \mathbb{C}$ defined by $(z, \pm z) \mapsto z$. Under these identifications, we have two one-dimensional representations $\pi_\pm$ of the C*-algebra $\mathbb{C}$ on the Hilbert space $\mathbb{C}$, given by $\pi_\pm(z) = z$. Clearly, these are equivalent: they are even identical. Hence by Theorem 10.15 the defining representation of $D_2(\mathbb{C})$ on $\mathbb{C}^2$ is reducible, as it should be: the explicit decomposition of $\mathbb{C}^2$ in $D_2(\mathbb{C})$-invariant subspaces is just the one (10.191) - (10.192) of the previous example.

The first-numbered claim of Theorem 10.15 is relatively easy to prove from Theorem C.90. Suppose $\pi_\pm(A_+)$ are inequivalent and take $b \in \pi_\omega(A)'$: we want to show that $b = \lambda \cdot 1$ for some $\lambda \in \mathbb{C}$. Relative to $H = H_+ \oplus H_-$, we write

$$b = \begin{pmatrix} b_{++} & b_{+-} \\ b_{-+} & b_{--} \end{pmatrix}, \tag{10.198}$$

where the four operators in this matrix act as follows:

$$b_{++} : H_+ \to H_+, b_{+-} : H_- \to H_+, \; b_{-+} : H_+ \to H_-, \; b_{--} : H_- \to H_-. \tag{10.199}$$

Since $A_+ \subset A$, we also have $b \in \pi_\omega(A_+)'$. The condition $[b, a] = 0$ for each $a \in A_+$ is equivalent to the four conditions

$$[b_{++}, \pi_+(a)] = 0; \tag{10.200}$$
$$[b_{--}, \pi_-(a)] = 0; \tag{10.201}$$
$$\pi_+(a)b_{+-} = b_{+-}\pi_-(a); \tag{10.202}$$
$$\pi_-(a)b_{-+} = b_{-+}\pi_+(a). \tag{10.203}$$

We now use the fact (which we state without proof) that, as in group theory, the irreducibility and inequivalence of $\pi_\pm(A_+)$ implies that there can be no nonzero operator $c : H_+ \to H_-$ such that $c\pi_+(a) = \pi_-(a)c$ for all $a \in A_+$, and vice versa. Hence $b_{+-} = 0$ as well as $b_{-+} = 0$. In addition, the irreducibility of $\pi_\pm(A_+)$ implies that $b_{++} = \lambda_+ \cdot 1_{H_+}$ and $b_{--} = \lambda_- \cdot 1_{H_1}$. Finally, the property $[b, a] = 0$ for each $a \in A_-$ implies $\lambda_+ = \lambda_-$. Hence $b = \lambda \cdot 1$, and $\pi_\omega(A)$ is irreducible.

To prove the second-numbered claim of Theorem 10.15, let $\pi_+(A_+) \cong \pi_-(A_+)$, so by definition (of equivalence) there is a unitary operator $v : H_- \to H_+$ such that

$$v\pi_-(a) = \pi_+(a)v, \forall a \in A_+. \tag{10.204}$$

Extend v to an operator $w : H \to H$ by

$$w = \begin{pmatrix} 0 & v \\ v^* & 0 \end{pmatrix}. \qquad (10.205)$$

It is easy to verify from (10.204) that $[w, \pi(a)] = 0$ for each $a \in A_+$. To check that the same is true for each $a \in A_-$, one needs the difficult analytical fact that w is a (weak) limit of operators of the kind $\pi(a_n)$, where $a_n \in A_-$, which also implies that $w^* \pi(a) \in \pi(A_+)''$. Since $\pi(A_+)''' = \pi(A_+)'$ and $w \in \pi(A_+)'$, we obtain $[w^* \pi(a), w] = 0$ for each $a \in A_-$. But for unitary operators w this is the same as $[w, \pi(a)] = 0$. So $w \in \pi(A)'$, and hence $\pi(A)$ is reducible by Theorem C.90. $\square$

In determining the ground state(s) of the quantum Ising chain, we will apply Theorem 10.15 to the C*-algebra (10.87). This application relies on the representation theory of F. For the moment we leave the Hilbert space H general, equipped though with a conjugation $J : H \to H$. It turns out to be convenient to use the **self-dual formulation of the CAR**, which treats c and c^* on an equal footing. Define

$$K = H \oplus H, \qquad (10.206)$$

whose elements are written as $h = (f, g)$ or $h = f \dotplus g$, with inner product

$$\langle h_1, h_2 \rangle_K = \langle f_1, f_2 \rangle_H + \langle g_1, g_2 \rangle_H. \qquad (10.207)$$

We then introduce a new operator in CAR(H), namely the **field**

$$\Phi(h) = c^*(f) + c(Jg), \qquad (10.208)$$

which is *linear* in $h = f \dotplus g$, because the antilinearity of $c(f)$ in f is canceled by the antilinearity of J. This yields the anti-commutation relations

$$[\Phi^*(h_1), \Phi(h_2)]_+ = \langle h_1, h_2 \rangle_K, \qquad (10.209)$$

but be aware that generally $[\Phi^*(h_1), \Phi^*(h_2)]_+$ and $[\Phi(h_1), \Phi(h_2)]_+$ do not vanish. Indeed, in terms of the antilinear operator $\Gamma : K \to K$, defined by

$$\Gamma = \begin{pmatrix} 0 & J \\ J & 0 \end{pmatrix} \qquad (10.210)$$

we have the following expression for the adjoint $\Phi(h)^* \equiv \Phi^*(h)$:

$$\Phi^*(h) = \Phi(\Gamma h). \qquad (10.211)$$

If we identify $f \in H$ with $f \dotplus 0 \in K$, we may reconstruct c and c^* from Φ through

$$c^*(f) = \Phi(f); \qquad (10.212)$$
$$c(f) = \Phi(\Gamma f). \qquad (10.213)$$

Bogoliubov transformations now take an extremely elegant form. For any unitary operator S on K that satisfies $[S, \Gamma] = 0$, we define the transform Φ_S of Φ by

$$\Phi_S(h) = \Phi(Sh), \tag{10.214}$$

with associated creation- and annihilation operators (where $H \ni f \equiv f + 0$, as above)

$$c_S^*(f) = \Phi_S(f); \tag{10.215}$$
$$c_S(f) = \Phi_S^*(f). \tag{10.216}$$

To see the equivalence with the original formulation of the Bogoliubov transformation, note that for unitary S, the condition $[S, \Gamma] = 0$ is equivalent to the structure

$$S = \begin{pmatrix} u & vJ \\ Jv & JuJ \end{pmatrix}, \tag{10.217}$$

where $u : H \to H$ is linear, $v : H \to H$ is antilinear, and u and v satisfy (10.133) - (10.134). Moreover, from (10.137) - (10.138) we obtain

$$c_S(f) = \eta(f); \tag{10.218}$$
$$c_S^*(f) = \eta^*(f). \tag{10.219}$$

An interesting class of pure states on $\mathrm{CAR}(H)$ arises as follows.

Theorem 10.16. *There is a bijective correspondence between:*

- *Projections $e : K \to K$ that (apart form the properties $e^2 = e^* = e$) satisfy*

$$\Gamma e \Gamma = 1_K - e; \tag{10.220}$$

- *States ω_e on F that satisfy*

$$\omega_e(\Phi(h)^* \Phi(h)) = \langle h, eh \rangle \ \forall h \in K. \tag{10.221}$$

Such a state ω_e is automatically pure (so that the corresponding GNS*-representation π_e is irreducible), and is explicitly given by*

$$\omega_e(\Phi(h_1) \cdots \Phi(h_{2n+1})) = 0; \tag{10.222}$$
$$\omega_e(\Phi(h_1) \cdots \Phi(h_{2n})) = \sum_{p \in \mathfrak{S}_{2n}}' \mathrm{sgn}(p) \prod_{j=1}^{n} \langle eh_{\mathrm{sgn}(2j)}, \Gamma h_{\mathrm{sgn}(2j-1)} \rangle, \tag{10.223}$$

the sum Σ' is over all permutations p of $1, \ldots, 2n$ such that

$$p(2j-1) < p(2j); \tag{10.224}$$
$$p(1) < p(3) < \cdots < p(2n-1). \tag{10.225}$$

We omit the proof. Note that (10.221) is a special case of (10.223), because of (10.211). States like ω_e, which are determined by their two-point functions, are called **quasi-free**; the ground state ω_0 on $\mathrm{CAR}(\mathbb{C}^N)$ constructed in the previous section is an example (one also has mixed quasi-free states, e.g. certain KMS states).

As a warm-up, we reconstruct the ground state of the free fermionic Hamiltonian on F using the above formalism. That is, we assume that h_N in (10.111) reads

$$h_N = \sum_{x=-N/2}^{N/2-1} \varepsilon_x c_x^* c_x, \qquad (10.226)$$

initially defining dynamics on $F_N = \mathrm{CAR}(\mathbb{C}^N)$. In that case, the projection e_0 onto the second copy of $H = \mathbb{C}^N$ in K, i.e.

$$e_0 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \qquad (10.227)$$

reproduces the ground state $\omega_0(a) = \langle 0|a|0\rangle$, where $|0\rangle$ is the vector $1 \in \mathbb{C}$ in $F_-(H)$, such that $c(f)|0\rangle = 0$ for all $f \in H$. This also works for $N = \infty$, i.e., we construct dynamics on $\mathrm{CAR}(\ell^2(\mathbb{Z}))$ from the local Hamiltonians (10.226) as indicated at the beginning of this section, and use the same formula for e_0, this time with $H = \ell^2(\mathbb{Z})$.

In the more general case (10.111), we replace e_0 in (10.227) by

$$e_0^{(S)} = S e_0 S^{-1}, \qquad (10.228)$$

where S is given by (10.217), in which for $N < \infty$ the operators u and v were constructed in (10.131) - (10.132). This time, the associated state $\omega_{e_0^{(S)}} \equiv \omega_S$ is the state called ω_0 in Theorem 10.13. As explained at the beginning of this section, this procedure even works for $N = \infty$ and hence $H = \ell^2(\mathbb{Z})$.

Having understood fermionic models with quadratic Hamiltonians, what remains to be done now is to reformulate the original quantum Ising chain, defined in terms of the local spin matrices $\sigma_i(x)$, in terms of the fermionic variables c_x and c_x^*. For finite N this was done through the Jordan–Wigner transformation (10.102) - (10.103). This time we need a similar isomorphism between A and F, where

$$A = \otimes_{j \in \mathbb{Z}} M_2(\mathbb{C}); \qquad (10.229)$$
$$F = \mathrm{CAR}(\ell^2(\mathbb{Z})), \qquad (10.230)$$

and hence we would need to start the sums in the right-hand side of (10.102) - (10.103) at $j = -\infty$. At first sight this appears to be impossible, though, because operators like $\exp(\pi i \sum_{y=-\infty}^{x-1} \sigma_+(y)\sigma_-(y))$ do not lie in A (whose elements have infinite tails of 2×2 unit matrices). Fortunately, this problem can be solved by adding a formal operator T to A, which plays the role of the "tail"

$$\text{``}T = e^{\pi i \sum_{y=-\infty}^{0} \sigma_+(y)\sigma_-(y))\text{''}}. \qquad (10.231)$$

This formal expression (to be used only heuristically) suggests the relations:

$$T^2 = 1; \tag{10.232}$$

$$T^* = T; \tag{10.233}$$

$$TaT = \theta_-(a), \tag{10.234}$$

where $\theta_- : A \to A$ is a $\mathbb{Z}_2$-action defined by (algebraic) extension of

$$\theta_-(\sigma_\pm(y)) = -\sigma_\pm(y) \ (y \leq 0); \tag{10.235}$$

$$\theta_-(\sigma_\pm(y)) = \sigma_\pm(y) \ (y > 0); \tag{10.236}$$

$$\theta_-(\sigma_3(y)) = \sigma_3(y) \ (y \in \mathbb{Z}); \tag{10.237}$$

$$\theta_-(\sigma_0(y)) = \sigma_0(y) \ (y \in \mathbb{Z}), \tag{10.238}$$

where $\sigma_0 = 1_2$. Formally, define an algebra extension

$$\hat{A} = A \oplus A \cdot T, \tag{10.239}$$

with elements of the type $a + bT$, $a, b \in A$, and algebraic relations given by (10.232) - (10.233). That is, we have

$$(a + bT)^* = a^* + \theta_-(b^*)T; \tag{10.240}$$

$$(a + bT) \cdot (a' + b'T) = aa' + b\theta_-(b') + (ab' + b\theta_-(a'))T. \tag{10.241}$$

Within $\hat{A}$, the correct version of (10.102) - (10.103) may now be written down as

$$c_x^\pm = Te^{\mp \pi i \sum_{y=x}^0 \sigma_+(y)\sigma_-(y)} \sigma_x^\pm \ (x < 1); \tag{10.242}$$

$$c_x^\pm = T\sigma_1^\pm; \tag{10.243}$$

$$c_x^\pm = Te^{\mp \pi i \sum_{y=1}^{x-1} \sigma_+(y)\sigma_-(y)} \sigma_x^\pm \ (x > 1), \tag{10.244}$$

with formal inverse transformation given by

$$\sigma_\pm(x) = Te^{\pm \pi i \sum_{y=x}^0 c_y^+ c_y^-} c_x^\pm \ (x < 1); \tag{10.245}$$

$$\sigma_\pm(x) = Tc_1^\pm; \tag{10.246}$$

$$\sigma_\pm(x) = Te^{\pm \pi i \sum_{y=1}^{x-1} \sigma_+(y)\sigma_-(y)} \sigma_\pm(x) \ (x > 1), \tag{10.247}$$

where this time we regard T as an element of the extended fermionic algebra

$$\hat{F} = F \oplus F \cdot T, \tag{10.248}$$

satisfying the same rules (10.232) - (10.234), but now in terms of a "fermionic" $\mathbb{Z}_2$-action $\theta_y : F \to F$ given by extending the following action on elementary operators:

$$\theta_-(c_y^\pm) = -c_y^\pm \ (y \leq 0); \tag{10.249}$$

$$\theta_-(c_y^\pm) = c_y^\pm \ (y > 0). \tag{10.250}$$

$$\tag{10.251}$$

Because of T, the Jordan–Wigner transformation does not give an isomorphism $A \cong F$, but it does give an isomorphism $\hat{A} \cong \hat{F}$. More importantly, if, having already defined the $\mathbb{Z}_2$-action θ on F by (10.177), we define a similar $\mathbb{Z}_2$-action on A by

$$\theta(\sigma_{\pm}(y)) = -\sigma_{\pm}(y) \ (y \in \mathbb{Z}); \tag{10.252}$$

$$\theta(\sigma_3(y)) = \sigma_3(y) \ (y \in \mathbb{Z}); \tag{10.253}$$

$$\theta(\sigma_0(y)) = \sigma_0(y) \ (y \in \mathbb{Z}), \tag{10.254}$$

and decompose $A = A_+ \oplus A_-$ and $F = F_+ \oplus F_-$, according to this action, cf. (10.178), we have isomorphisms

$$A_+ \cong F_+; \tag{10.255}$$

$$A_- \cong F_- T; \tag{10.256}$$

$$A \cong F_+ \oplus F_- T. \tag{10.257}$$

For given dynamics (10.111), suppose ω_0^A is a $\mathbb{Z}_2$-*invariant* ground state on A. Then ω_0^A also defines a $\mathbb{Z}_2$-invariant ground state ω_0^F on F by (10.255) and $\omega_0^F(f) = 0$ for all $f \in F_-$. Conversely, a $\mathbb{Z}_2$-invariant ground state ω_0^F on F defines a state ω_0^A on A by (10.255) and $\omega_0^A(a) = 0$ for all $a \in A_-$. But F has a unique ground state, so:

- Either ω_0 is pure on A, in which case it is the unique ground state on A;
- Or ω_0 is mixed on A, in which case $\omega_0 = \frac{1}{2}(\omega_0^+ + \omega_0^-)$, where $\omega_0^{\pm}$ are pure but transform under the above $\mathbb{Z}_2$-action θ as $\omega_0^{\pm} \circ \theta = \omega_0^{\mp}$.

Theorem 10.15 gives a representation-theoretical criterion deciding between these possibilities, but to apply it we need some information on the restriction of $\mathbb{Z}_2$-invariant quasi-free pure states on F to its even part F_+. The abstract setting involves a $\mathbb{Z}_2$-action W on K that commutes with Γ (so that W is unitary, $W^2 = 1$, and $[\Gamma, W] = 0$), which induces a $\mathbb{Z}_2$-action θ on F by linear and algebraic extension of $\theta(\Phi(h)) = \Phi(Wh)$. A quasi-free state ω_e, defined according to Theorem 10.16 by a projection $e : K \to K$ that satisfies (10.220), is then $\mathbb{Z}_2$-invariant iff $[W, e] = 0$.

In our case, this simplifies to $\theta(\Phi(h)) = -\Phi(h)$, so that $W = -1$, and every projection commutes with W. In any case, with considerable effort one can prove:

Lemma 10.17. *Given some $\mathbb{Z}_2$-action W on K, as well as a projection $e : K \to K$ satisfying (10.220), such that $[W, \Gamma] = [W, e] = 0$:*

1. *The quasi-free state ω_e of Theorem 10.16 is $\mathbb{Z}_2$-invariant (i.e., $\omega_e \circ \theta = \omega_e$);*
2. *The corresponding GNS-representation space $H_e \equiv H_{\omega_e}$ for $F = F_+ \oplus F_-$ decomposes as $H_e = H_e^+ \oplus H_e^-$, with $H_e^{\pm} = \overline{F_{\pm}\Omega_e}$. Each subspace $H_e^{\pm}$ is stable under $\pi_e(F_+)$, and the restriction $\pi_e^{\pm}$ of $\pi(F_+)$ to $H_e^{\pm}$ is irreducible.*

Theorem 10.15 then leads to a lemma, which also summarizes the discussion so far.

Lemma 10.18. *1. For given $\mathbb{Z}_2$-invariant dynamics, let ω_0^F be the (unique, $\mathbb{Z}_2$-invariant) ground state on $F = F_+ \oplus F_-$. Under $F_+ \subset F$ the associated GNS-representation space H_0 decomposes as $H_0 = H_0^+ \oplus H_0^-$, with $H_0^{\pm} = \overline{F_{\pm}\Omega_0}$, and we denote the restriction of $\pi_0(F_+)$ to $H_0^{\pm}$ by $\pi_0^{\pm}$. Then $\pi_0^{\pm}(F_+)$ are irreducible.*

2. *Regard ω_0^F also as a state ω_0^T on $F_+ \oplus F_- T$ by putting $\omega_0^T(a) = 0$ for all $a \in F_- T$, and similarly as a state ω_0^A on A by invoking (10.255) and putting $\omega_0^A(a) = 0$ for all $a \in A_-$. Let $H_0^T = H_+^T \oplus H_-^T$ be the GNS-representation space of $F_+ \oplus F_- T$ defined by ω_0^T, where $H_+^T = \overline{F_+\Omega}$ and $H_-^T = \overline{F_- T\Omega}$. Here H_+^T and H_-^T are stable under F_+; we denote the restriction of F_+ to $H_\pm^T$ by $\pi_\pm^T$, so that $\pi_+^T \cong \pi_0^+$.*

    a. *Then ω_0^A is a ground state on A. Any $\mathbb{Z}_2$-invariant ground state on A arises in this way (via F), so that there is a unique $\mathbb{Z}_2$-invariant ground state on A.*

    b. *The state ω_0^A is pure on A iff the irreducible representations $\pi_+^T(F_+)$ (or $\pi_0^+(F_+)$) and $\pi_-^T(F_+)$ are inequivalent.*

It turns out to be difficult to directly check the (in)equivalence of $\pi_\pm^T(F_+)$. Fortunately, we can circumvent this problem by passing to yet another (irreducible) representation of F_+. We first enlarge F to a new algebra

$$\hat{F} = F \oplus FT = F_+ \oplus F_- \oplus F_+ T \oplus F_- T, \tag{10.258}$$

and extend the state ω_0^F on F to a state $\hat{\omega}_0$ on $\hat{F}$ by putting $\hat{\omega}_0(FT) = 0$, so that $\hat{\omega}_0$ is nonzero only on $F_+ \subset \hat{F}$. Let $\hat{\pi}_0$ be the associated GNS-representation of $\hat{F}$ on the Hilbert space $\hat{H}_0 = \overline{\hat{F}\hat{\Omega}}$. Under $\hat{\pi}(F_+)$ this space decomposes as

$$\hat{H}_0 = \overline{F_+\hat{\Omega}_0} \oplus \overline{F_-\hat{\Omega}_0} \oplus \overline{F_+ T\hat{\Omega}_0} \oplus \overline{F_- T\hat{\Omega}_0}, \tag{10.259}$$

with corresponding restrictions $\hat{\pi}_\pm(F_+)$ and $\hat{\pi}_\pm^T(F_+)$; more precisely, $\hat{\pi}_\pm$ is the restriction of $\hat{\pi}(F_+)$ to $\overline{F_\pm\hat{\Omega}_0}$, whilst $\hat{\pi}_\pm^T$ is is the restriction of $\hat{\pi}(F_+)$ to $\overline{F_\pm T\hat{\Omega}_0}$. Clearly, $\hat{\pi}_\pm(F_+)$ is the same as $\pi_0^\pm(F_+)$, and $\hat{\pi}_-^T(F_+)$ is just our earlier $\pi_-^T(F_+)$, but $\hat{\pi}_+^T(F_+)$ is new. To understand the latter, we rewrite (10.259) as

$$\hat{H}_0 = H_0 \oplus \hat{H}_0^T; \tag{10.260}$$

$$H_0 = \overline{F_+\hat{\Omega}_0} \oplus \overline{F_-\hat{\Omega}_0} \cong \overline{F_+\Omega_0} \oplus \overline{F_-\Omega_0}; \tag{10.261}$$

$$\hat{H}_0^T = \overline{F_+ T\hat{\Omega}_0} \oplus \overline{F_- T\hat{\Omega}_0}, \tag{10.262}$$

the point being that $\hat{\pi}(F)$ evidently restricts to both H_0 and $\hat{H}_0^T$. We know the action of $\hat{\pi}(F)$ on H_0 quite well: it is the representation induced by the ground state ω_0. As to $\hat{H}_0^T$, we define a state $\hat{\omega}_0^T$ on F by

$$\hat{\omega}_0^T(a) = \langle \hat{\pi}(T)\hat{\Omega}_0, \hat{\pi}(a)\hat{\pi}(T)\hat{\Omega}_0 \rangle_{\hat{H}_0} = \langle \hat{\Omega}_0, \hat{\pi}(\theta_-(a))\hat{\Omega}_0 \rangle_{\hat{H}_0}, \tag{10.263}$$

where the second equality follows from (10.234). Comparing H_0 and $\hat{H}_0$, for all $b \in F$ (and hence especially for $b = \theta_-(a)$) we simply have

$$\langle \hat{\Omega}_0, \hat{\pi}(b)\hat{\Omega}_0 \rangle_{\hat{H}_0} = \hat{\omega}_0(b) = \omega_0^F(b), \tag{10.264}$$

so that $\hat{\omega}_0^T = \omega_0^F \circ \theta_- \equiv \theta_-^* \omega_0^F$. Decomposing the GNS-representation space $H_{\theta_-^* \omega_0^F}$ of $\pi_{\theta_-^* \omega_0^F}(F)$ as $H_{\theta_-^* \omega_0^F} = H_{\theta_-^* \omega_0^F}^+ \oplus H_{\theta_-^* \omega_0^F}^-$, it follows that $\hat{\pi}_+^T(F_+)$ is the restriction

of $\pi_{\theta^* \omega_0^F}(F_+)$ to $H^+_{\theta^* \omega_0^F}$. Therefore, the representation $\hat{\pi}(F)$ restricted to $\hat{H}^T_0$ is the GNS-representation $\pi_{\theta^* \omega_0^F}(F)$, so that in turn $\hat{\pi}^T_+(F_+)$ is $\pi_{\theta^* \omega_0^F}(F_+)$, restricted to $H^+_{\theta^* \omega_0}$. Hence, further to (10.260) - (10.262), we obtain the decomposition

$$\hat{\pi}(F) \cong \pi_{\omega_0^F}(F) \oplus \pi_{\theta^* \omega_0^F}(F). \tag{10.265}$$

The point is that for the quantum Ising chain Hamiltonian (10.110), we have:

Lemma 10.19. *1. For each $\lambda \neq \pm 1$, we have $\pi_{\omega_0^F}(F) \cong \pi_{\theta^* \omega_0^F}(F)$.*

2. If this holds, then the representations $\pi_0^+(F_+) \equiv \pi^+_{\omega_0^F}(F_+)$ and $\pi^T_-(F_+)$ are inequivalent iff the representations $\pi^+_{\omega_0^F}(F_+)$ and $\pi^+_{\theta^* \omega_0^F}(F_+)$ are equivalent.

3. For each $\lambda \neq \pm 1$, the ground state ω_0^A is pure on A iff the representations $\pi_{\omega_0^F}(F_+)$ and $\pi_{\theta^* \omega_0^F}(F_+)$ are equivalent.

The first claim follows from Theorem 10.20 below. The third follows from Lemma 10.18 and the previous claims. The second claim is proved by repeatedly applying Theorem 10.15 to $\hat{\pi}(\hat{F})$. Given this lemma, the real issue now lies in comparing $\pi_{\omega_0^F}$ and $\pi_{\theta^* \omega_0^F}$, both as representations of F (as they are defined) and as representations of $F_+ \subset F$. This can be settled in great generality by first looking at Theorem 10.16, and thence, recalling the positive-energy projection (10.228), realizing that

$$\pi_{\omega_0^F} = \pi_{e_0^{(S)}}; \tag{10.266}$$

$$\pi_{\theta^* \omega_0^F} = \pi_{W_- e_0^{(S)} W_-}. \tag{10.267}$$

Here $W_- : K \to K$ is the $\mathbb{Z}_2$-action on K defining the $\mathbb{Z}_2$-action θ_- on F as explained above Lemma 10.17; specifically, W_- is the direct sum of two copies of $w_- : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$, defined by $w_-(f_j) = f_j$ $(j > 0)$ and $w_-(f_j) = -f_j$ $(j \leq 0)$.

Subsequently, without proof we invoke a basic result on the CAR-algebra:

Theorem 10.20. *Let e and e' be projections on K that satisfy (10.220). Then:*

1. $\pi_e(F) \cong \pi_{e'}(F)$ iff $e - e' \in B_2(K)$;
2. $\pi_e^+(F_+) \cong \pi_{e'}^+(F_+)$ iff $e - e' \in B_2(K)$ and $\dim(eK \cap (1 - e')K)$ is even.

If the first condition is satisfied, the dimension in the second part is finite, so that one may indeed say it is even or odd. From Lemmas 10.18 and 10.19 and Theorem 10.20, we finally obtain the phase structure of the infinite quantum Ising chain:

Theorem 10.21. *The unique $\mathbb{Z}_2$-invariant ground state ω_0 of the Hamiltonian (10.110) is pure (and hence forms the unique ground state) iff both of the following hold:*

$$e_0^{(S)} - W_- e_0^{(S)} W_- \in B_2(K); \tag{10.268}$$

$$\dim(e_0^{(S)} K \cap (1 - W_- e_0^{(S)} W_-)K) \text{ is even.} \tag{10.269}$$

This is true for all λ with $|\lambda| \geq 1$. If $|\lambda| < 1$, then $\omega_0 = \frac{1}{2}(\omega_0^+ + \omega_0^-)$, where $\omega_0^\pm$ are pure and transform under the $\mathbb{Z}_2$-action θ as $\omega_0^\pm \circ \theta = \omega_0^\mp$.

10.8 Spontaneous symmetry breaking in mean-field theories

We are now going to study SSB in so-called **mean-field theories**: these are quantum spin systems with Hamiltonians like the **Curie–Weiss-model** for ferromagnetism:

$$h_\Lambda^{CW} = -\frac{J}{2|\Lambda|} \sum_{x,y\in\Lambda} \sigma_3(x)\sigma_3(y) - B\sum_{x\in\Lambda}\sigma_1(x), \qquad (10.270)$$

where $J > 0$ scales the spin-spin coupling, and B is an external magnetic field. Similar to the quantum Ising model, (10.270) has a $\mathbb{Z}_2$-symmetry $(\sigma_1,\sigma_2,\sigma_3) \mapsto (\sigma_1,-\sigma_2,-\sigma_3)$, which at each site x is implemented by $u(x) = \sigma_1(x)$. This model differs from its short-range counterpart (9.42), i.e, the quantum Ising model, or the Heisenberg model (9.44), in that every spin now interacts with every other spin. It falls into the class of **homogeneous mean-field theories**, which are defined by a single-site Hilbert space $H_x = H = \mathbb{C}^n$ and local Hamiltonians of the type

$$h_\Lambda = |\Lambda|\tilde{h}(T_0^{(\Lambda)}, T_1^{(\Lambda)}, \ldots, T_{n^2-1}^{(\Lambda)}). \qquad (10.271)$$

Here $T_0 = 1_n$, and the matrices $(T_i)_{i=1}^{n^2-1}$ in $M_n(\mathbb{C})$ form a basis of the real vector space of traceless self-adjoint $n\times n$ matrices; the latter may be identified with i times the Lie algebra $\mathfrak{su}(\mathfrak{n})$ of $SU(n)$, so that $(T_0, T_1, \ldots, T_{n^2-1})$ is a basis of i times the Lie algebra $\mathfrak{u}(\mathfrak{n})$ of the unitary group $U(n)$ on $\mathbb{C}^n$. In those terms, we define

$$T_i^{(\Lambda)} = \frac{1}{|\Lambda|}\sum_{x\in\Lambda} T_i(x), \qquad (10.272)$$

Finally, $\tilde{h}$ is a polynomial (which is sensitive to operator ordering). For example, to cast (10.270) (with $J = 1$) in the form (10.271), take $n = 2$, $T_i = \frac{1}{2}\sigma_i (= 1, 2, 3)$, and

$$\tilde{h}^{CW}(T_1, T_2, T_3) = -2(T_3^2 + BT_1). \qquad (10.273)$$

The assumptions of Theorem 9.15 do not hold now, and indeed the local dynamics (9.40) fails to converge to global dynamics on the quasi-local C*-algebra A defined by (8.130). Fortunately, it does converge to a global dynamics on the C*-algebra $C(S(B))$, where $B = M_n(\mathbb{C})$ is the single-site algebra. In order to describe the limiting dynamics of (homogeneous) mean-field models as $\Lambda \nearrow \mathbb{Z}^d$, we equip the state space $S(B)$ with the Poisson structure (8.52), which we now elucidate.

For unital C*-algebras B, we may regard $S(B)$ as a w^*-compact subspace of either the complex vector space B^* or the real vector space B_{sa}^*; in the latter case we regard states as linear maps $\omega : B_{sa}^* \to \mathbb{R}$ that satisfy $\omega(1_B) = 1$ and $\omega(a^2) \geq 0$ for each $a \in B_{sa}$. If $B = M_n(\mathbb{C})$, which is all we need, we may furthermore identify B_{sa}^* with $i\mathfrak{u}(\mathfrak{n})^*$, and since the value of each state $\omega \in S(M_n(\mathbb{C}))$ is fixed on $T_0 = 1_B \in i\mathfrak{u}(\mathfrak{n})$, it follows that $S(M_n(\mathbb{C}))$ is a compact convex subset of $i\mathfrak{su}(\mathfrak{n})^*$. In that case, the Poisson bracket (8.52) on $S(M_n(\mathbb{C}))$ is none other than the restriction of (minus) the canonical Lie-Poisson bracket on $\mathfrak{su}(\mathfrak{n})^* \cong i\mathfrak{su}(\mathfrak{n})^*$ to $S(M_n(\mathbb{C}))$, cf. (3.98) - (3.99).

For example, for $n = 2$ we have $S(M_2(\mathbb{C})) \cong B^3 \subset \mathbb{R}^3$ by Proposition 2.9, i.e.,

$$\omega_{(x,y,z)}(a) = \mathrm{Tr}\left(\rho(x,y,x)a\right) \ ((x,y,z) \in B^3, a \in M_2(\mathbb{C})); \qquad (10.274)$$

$$\rho(x,y,z) = \tfrac{1}{2} \begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix}. \qquad (10.275)$$

We also have $\mathfrak{su}(2)^* \cong \mathbb{R}^3$ upon the choice of the basis $(T_i = \tfrac{1}{2}\sigma_i)$, $i = 1,2,3$, of $i\mathfrak{su}(2)$, which means that $\theta_{(x,y,z)} \in i\mathfrak{su}(2)^*$ maps (T_1, T_2, T_3) to (x,y,z) (where this time $(x,y,z) \in \mathbb{R}^3$), cf. §5.8. If we now regard the matrices T_i as functions $\hat{T}_i$ on B^3 by $\hat{T}_i(\omega) = \omega(T_i)$, we find that the corresponding functions on B^3 are given by

$$\hat{T}_1(x,y,z) = \tfrac{1}{2}x, \ \ \hat{T}_2(x,y,z) = \tfrac{1}{2}y, \ \ \hat{T}_3(x,y,z) = \tfrac{1}{2}z. \qquad (10.276)$$

The corresponding Poisson brackets (8.52) are $\{T_1, T_2\} = -2T_3$ etc., i.e., $\{x,y\} = -2z$ etc.; this is -2 times the bracket defined in (3.43) or (3.97) - (3.98). This factor 2 could have been avoided by moving to the three-ball with radius $r = 1/2$ instead of $r = 1$, whose boundary is the coadjoint orbit $\mathcal{O}_{1/2}$ naturally associated to spin-$\tfrac{1}{2}$.

We now return to our continuous bundle of C*-algebras $A^{(c)}$ of Theorem 8.4, of course in the slightly adapted form appropriate to quantum spin systems, see §8.6. In particular, we recall that $A_0^{(c)} = C(S(B))$ and $A_{1/N}^{(c)} = B(H_{\Lambda_N})$, cf. (8.157) - (8.158), and hence we see the limit $N \to \infty$ as a specific way of taking the limit $\Lambda \nearrow \mathbb{Z}^d$ along the hypercubes Λ_N. Symmetric and quasi-symmetric sequences $(a_{1/N})_{N \in \mathbb{N}}$ are defined as explained after (8.161). The following observation is fundamental.

Theorem 10.22. *Let $B = M_n(\mathbb{C})$. If $(a_{1/N})_{N \in \mathbb{N}}$ and $(b_{1/N})_{N \in \mathbb{N}}$ are symmetric sequences with limits a_0 and b_0 as defined by (8.46), respectively (so that $(a_{1/N})_{N \in \mathbb{N}}$ and $(b_{1/N})_{N \in \mathbb{N}}$ are continuous sections of the continuous bundle $A^{(c)}$), then the sequence*

$$\left(\{a_0,b_0\}, i[a_1,b_1], \ldots, i|\Lambda_N|[a_{1/N},b_{1/N}], \cdots\right) \qquad (10.277)$$

defines a continuous section of $A^{(c)}$. In particular, for each $\omega \in S(B)$ we have

$$i \lim_{N \to \infty} \omega^{|\Lambda_N|}(|\Lambda_N|[a_{1/N},b_{1/N}]) = \{a_0,b_0\}(\omega). \qquad (10.278)$$

Proof. The proof is a straightforward combinatorial exercise, and we just mention the simplest case where $d = 1$ and $a_{1/N} = S_{1,N}(a_1)$ and $b_{1/N} = S_{1,N}(b_1)$, where $a_1 \in B$ and $b_1 \in B$, cf. (8.39). Then $a_0 = \hat{a}_1$, $b_0 = \hat{b}_1$, and similarly to (8.45) we find

$$[S_{1,N}(a_1), S_{1,N}(b_1)] = \frac{1}{N}S_{1,N}([a_1,b_1]), \qquad (10.279)$$

Using (8.52), we find that (10.277) is equal to $(i\widehat{[a_1,b_1]}, \ldots, S_{1,N}([a_1,b_1]), \ldots)$. Since $\omega^N(S_{1,N}([a_1,b_1])) = \omega([a_1,b_1])$, the left-hand side of (10.278) is therefore equal to $i\omega([a_1,b_1])$, which by (8.52) equals the right-hand side. $\qquad\square$

In other words, although the sequence of commutators $[a_{1/N}, b_{1/N}]$ converges to zero (which is why $A_0^{(c)}$ has to be commutative!), the rescaled commutators $iN[a_{1/N}, b_{1/N}]$ converge to the macroscopic observable $\{a_0, b_0\} \in C(S(B))$. This reconfirms the analogy between the limit $N \to \infty$ and the limit $\hbar \to 0$ of Chapter 7, see especially Definitions 7.1 and 8.2. With $B = M_n(\mathbb{C})$, Theorem 10.22 implies the central result about the macroscopic (and hence classical!) dynamics of mean-field theories:

Corollary 10.23. *Let $(h_{1/N})_{N \in \dot{\mathbb{N}}}$ be a continuous section of $A^{(c)}$ defined by a symmetric sequence, and let $(a_{1/N})_{N \in \dot{\mathbb{N}}}$ be an arbitrary continuous section of $A^{(c)}$ (i.e. a quasi-symmetric sequence). Then, writing $h_{1/N} = h_{\Lambda_N}$ for clarity, the sequence*

$$\left(a_0(t), e^{ih_{\Lambda_1} t} a_1 e^{-ih_{\Lambda_1} t}, \cdots e^{ih_{\Lambda_N} t} a_{1/N} e^{-ih_{\Lambda_N} t}, \cdots \right), \tag{10.280}$$

where $a_0(t)$ is the solution of the equations of motion on $S(M_n(\mathbb{C}))$ with classical Hamiltonian h_0 and Poisson bracket (8.52), defines a continuous section of $A^{(c)}$.

In other words, the Heisenberg dynamics on $A_{\Lambda_N} = B(H_{\Lambda_N})$ defined by the quantum Hamiltonians h_{Λ_N} converges to the classical dynamics on the Poisson manifold $S(M_n(\mathbb{C}))$ that is generated by their classical limit, viz. the Hamiltonian h_0.

For example, since the operators $T_i^{(\Lambda)}$ form symmetric sequences, so do Hamiltonians of the type (10.271). The limit $h_0 \in C(S(M_n(\mathbb{C})))$ of the family (h_Λ) in (10.271) is simply obtained by replacing the operators $T_i^{(\Lambda)}$ in the function $\tilde{h}$ by the functions $\hat{T}_i$ on $S(M_n(\mathbb{C}))$. Equivalently, one may replace the $T_i^{(\Lambda)}$ by the canonical coordinates (θ_i) of $i\mathfrak{su}(n)^*$ dual to the basis $(T_1, \ldots, T_{n^2-1})$ of $i\mathfrak{su}(n)^*$, i.e., $\theta_i(T_j) = \delta_{ij}$, and restricting the ensuing function on $i\mathfrak{su}(n)^*$ to $S(M_n(\mathbb{C})) \subset i\mathfrak{su}(n)^*$.

Using (10.276), for the Curie–Weiss model (10.270) with $J = 1$ this gives

$$h_0^{CW}(x, y, z) = -\tfrac{1}{2} z^2 - Bx. \tag{10.281}$$

The ground states of this Hamiltonian are simply its minima, viz.

$$\mathbf{x}_\pm = (B, 0, \pm\sqrt{1 - B^2}) \ (0 \le B < 1); \tag{10.282}$$

$$\mathbf{x} = (1, 0, 0)) \ (B \ge 1), \tag{10.283}$$

all of which lie on the boundary S^2 of B^3. Note that the points $\mathbf{x}_\pm$ coalesce as $B \to 1$, where they form a saddle point. Modulo our use of radius $r = 1$ instead of $r = 1/2$, this result coincides with (10.81) for classical limit of the quantum Ising model.

We now turn to symmetry and its possible breakdown. Suppose there is some subgroup of $U(n)$, typically the image of a unitary representation $g \mapsto u_g$ of a compact group G on $\mathbb{C}^n$, under which $\tilde{h}(T_0, T_1, \ldots, T_{n^2-1})$ in (10.271) satisfies

$$\tilde{h}(T_0, u_g T_1 u_g^*, \ldots, u_g T_{n^2-1} u_g^*) = \tilde{h}(T_0, T_1, \ldots, T_{n^2-1}) \ (g \in G). \tag{10.284}$$

For example, in the Curie–Weiss model one has $G = \mathbb{Z}_2$, whose nontrivial element is represented by σ_1. For (10.271) itself this implies $u^{(N)} h_N (u^{(N)})^* = h_N$, cf. (10.69).

Hence also in homogeneous mean-field models we obtain the structure (10.57), (10.58), and (10.59) familiar from the case of short-range forces. For the limit theory this implies that the classical Hamiltonian h_0 on $S(M_n(\mathbb{C}))$ is invariant under the coadjoint action of $G \subset U(n)$ on $i\mathfrak{su}(\mathfrak{n})^*$, restricted to $S(M_n(\mathbb{C})) \subset i\mathfrak{su}(\mathfrak{n})^*$: in the Curie–Weiss model this "classical shadow" of the $\mathbb{Z}_2$ symmetry of the quantum theory is simply the map $(x,y,z) \mapsto (x,-y,-z)$ on B^3.

In the regime $0 < B < 1$, the degenerate ground states of this model break this symmetry. In contrast, it can be shown from the Perron–Frobenius Theorem (which applies since both σ_3 and σ_1 are real matrices) that for $B > 0$ each quantum-mechanical Hamiltonian (10.270) has a unique ground state $\psi_N^{(0)}$. Being unique, this vector must share the invariance of h_N under the permutation group $\mathfrak{S}_N$, so that

$$\psi_N^{(0)} = \sum_{n_+=0}^{N} c(n_+/N)|n_+,n_-\rangle, \qquad (10.285)$$

where $|n_+,n_-\rangle$ is the totally symmetrized unit vector in $\otimes^N \mathbb{C}^2$ with n_+ spins up and $n_- = N - n_+$ spins down, and $c : \{0, 1/N, 2/N, \ldots, (N-1)/N, 1\} \to [0,1]$ is some function such that $\sum_{n_+} c(n_+/N)^2 = 1$ (we may assume $c \geq 0$ by the Perron–Frobenius Theorem). The asymptotic behaviour of c as $N \to \infty$ has been studied, and as expected, c to converges pointwise to $c(0) = c(1) = \sqrt{1/2}$ and $c(x) = 0$, and zero elsewhere (at $B = 0$ one of course has either $c(0) = 1$ or $c(1) = 1$ for all N).

Thus we encounter a familiar headache: the "higher-level" theory $C(S(M_n(\mathbb{C})))$ at $N = \infty$ breaks the $\mathbb{Z}_2$ symmetry, whereas the "lower-level" quantum theories $B(H_{\Lambda_N})$ ($N < \infty$) do not, although the former should be a limiting case of the latter. Indeed, the situation for the Curie–Weiss model in the regime $0 < B < 1$ is exactly analogous to the double-well potential as well as to the quantum Ising model in the same regime: if the two degenerate ground states $\mathbf{x}_\pm \in B^3$ of h_0^{CW} are reinterpreted as Dirac measures $\delta_\pm$ on B^3, which in turn are seen as (pure) states $\omega_\pm$ on the classical algebra of observables $C(S(M_2(\mathbb{C})))$, then (10.74) holds, *mutatis mutandis*.

The resolution of this problem through the restoration of Butterfield's Principle should also be the same as for the previous two cases: there is a first excited state $\psi_N^{(1)}$ such that as $N \to \infty$, the energy difference with the ground state approaches zero and one has approximate symmetry breaking as in (10.75)). Alas, for the Curie–Weiss model so far only numerical evidence is available supporting this scenario.

Equilibrium states of homogeneous mean-field models at any inverse temperature $0 < \beta < \infty$ exist, despite the fact that in such models time-evolution α_t on the infinite system A (and hence the KMS condition characterizing equilibrium states) is ill-defined (unless one passes to certain representations of A, which would be question-begging). Instead, one invokes the quasi-local C*-algebra A, cf. (8.130), and *in lieu* of KMS states looks for limit points $\hat{\omega}^\beta \in S(A)$ of the local Gibbs states $\omega_{\Lambda_N}^\beta$ defined by (9.96) as $N \to \infty$; see (10.44) and surrounding discussion. Proposition 10.8 does not apply now, but Theorem 8.9 does: since each local Hamiltonian h_{Λ_N} is permutation-invariant (because each $T_i^{(\Lambda_N)}$ is), so is each local Gibbs state $\omega_{\Lambda_N}^\beta$, and accordingly, each w^*-limit point of this sequence must share this property.

As in (8.174), from the quantum De Finetti Theorem 8.9 we therefore have:

$$\hat{\omega}^{\beta} = \int_{S(M_n(\mathbb{C}))} d\mu_{\beta}(\theta) \left(\omega_{\theta}^{\beta}\right)^{\infty},$$
(10.286)

for some probability measure μ_{β} on the single-spin state space $S(M_n(\mathbb{C}))$. By Proposition 8.28, this measure may also be regarded as a limit of the local Gibbs states, but now regarded as a state on the limit algebra $A_0^{(c)} = C(S(M_n(\mathbb{C})))$ rather than as a state on $A_0^{(q)} = A$. By the same token, each state ω_{θ}^{β} in the decomposition (10.286) is a pure state on $A_0^{(c)}$ (though seen as a state on $M_n(\mathbb{C})$ it will be mixed!). The states ω_{θ}^{β} are computed as follows. Given a classical Hamiltonian h_0 computed from (10.271) as explained after Corollary 10.23, for each point $\theta = (\theta_0, \ldots, \theta_{n^2-1}) \in i u(n)^*$ we define a new self-adjoint operator $\hat{h}_{\theta} \in M_n(\mathbb{C})$ by

$$\hat{h}_{\theta} = h_0(\theta) \cdot 1_n + \sum_{i=0}^{n^2-1} \frac{\partial h_0}{\partial \theta_i}(\theta) \cdot T_i.$$
(10.287)

For example, in the Curie–Weiss model, from (10.273) we have

$$h_0^{CW}(\theta) = -2(\theta_3^2 + B\theta_1);$$
(10.288)

$$\hat{h}_{\theta}^{CW} = h_0^{CW}(\theta) - 2\theta_3\sigma_3 - B\sigma_1.$$
(10.289)

Eq. (10.287) has the following origin. Let ω be any state on A for which the strong limit $T_i^{(\omega)}$ of each operator $\pi_{\omega}(T_i^{(\Lambda_N)})$ on H_{ω} exists as $N \to \infty$ (for example, as in the proof of Theorem 8.16 one may show that this is the case when ω is a permutation-invariant state of A). It easily follows that $T_i^{(\omega)}$ lies in the algebra at infinity for π_{ω}, and hence in the center of $\pi_{\omega}(A)''$, cf. §8.5. If, in addition, ω is primary, then

$$T_i^{(\omega)} = \theta_i \cdot 1_{H_{\omega}};$$
(10.290)

$$\theta_i = \lim_{N\to\infty} \omega(T_i^{(\Lambda_N)}).$$
(10.291)

Under these assumptions, we compute the commutator

$$[\pi_{\omega}(h_{\Lambda_N}), \pi_{\omega}(a)] = \sum_i \frac{\partial h_0}{\partial \theta_i} \left(T_0^{(\Lambda)}, \ldots, T_{n^2-1}^{(\Lambda)}\right) \cdot \sum_{x \in \Lambda_N} [\pi_{\omega}(T_i(x)), \pi_{\omega}(a)] + O\left(\frac{1}{|\Lambda_N|}\right),$$

where $a \in \cup_{\Lambda} A_{\Lambda}$, and $O(1/|\Lambda_N|)$ denotes a finite sum of (multiple) commutators between some power of $T_i^{(\Lambda)}$ and operators that are (norm-) bounded in N. For example, for the Curie–Weiss model the $O(1/|\Lambda_N|)$ term is a multiple of

$$\sum_{x \in \Lambda_N} [[\pi_{\omega}(\sigma_3(x)), \pi_{\omega}(a)], \sigma_3^{(\Lambda_N)}].$$
(10.292)

Since a is local, all commutators $\sum_{x \in \Lambda_N} [\pi_\omega(T_i(x)), \pi_\omega(a)]$ are in $\pi_\omega(A)$, so that further commutators à la (10.292) vanish as $N \to \infty$. Also, in this limit the terms $T_i^{(\Lambda)}$ in the argument of $\sum_i \partial h_0 / \partial \theta_i$ assume their c-number values θ_i, so that

$$\lim_{N \to \infty} [\pi_\omega(h_{\Lambda_N}), \pi_\omega(a)] = [h_\omega, \pi_\omega(a)], \tag{10.293}$$

where formally (i.e. on a suitable domain) we have an ω-dependent Hamiltonian

$$h_\omega = \sum_{x \in \mathbb{Z}^d} \pi(\hat{h}_\theta(x)), \tag{10.294}$$

where the θ_i depend on ω via (10.291). Also, for each $a \in A$ one has strong limits

$$\lim_{N \to \infty} \pi_\omega \left(e^{ih_{\Lambda_N} t} a e^{-ih_{\Lambda_N} t} \right) = e^{ih_\omega t} \pi(a) e^{-ih_\omega t}. \tag{10.295}$$

Hence in the limit $N = \infty$ (provided it makes sense, which it does under the stated assumptions), the original mean-field Hamiltonian (10.271) with its homogeneous long-range forces converges to a sum of single-body Hamiltonians, in which the original forces between the spins have been incorporated into the parameters θ_i.

Returning to (10.286), for any $\beta = T^{-1}$, we now determine ω_θ^β from the *Ansatz*

$$\omega_\theta^\beta(a) = \frac{\mathrm{Tr}\left(e^{-\beta \hat{h}_\theta} a \right)}{\mathrm{Tr}\left(e^{-\beta \hat{h}_\theta} \right)}, \tag{10.296}$$

where θ is found by by solving the ***self-consistency equation***

$$\omega_\theta^\beta = \theta. \tag{10.297}$$

As explained after Corollary 10.23, here $\omega_\theta^\beta : M_n(\mathbb{C})_{\mathrm{sa}} \to \mathbb{R}$ is defined by its values on $i\mathfrak{su}(\mathfrak{n})$ and hence should be seen as a map $i\mathfrak{su}(\mathfrak{n}) \to \mathbb{R}$, like $\theta \in \mathfrak{su}(\mathfrak{n})^*$, so that (10.297) consists of $n^2 - 1$ equations $\omega_\theta^\beta(T_i) = \theta_i$ ($i = 1, \ldots, n^2 - 1$). Alternatively, one may extend θ from $i\mathfrak{su}(\mathfrak{n})$ to $i\mathfrak{u}(\mathfrak{n})$ by prescribing $\theta(1_n) = 1$, and subsequently extend it further to $M_n(\mathbb{C})$ by complex linearity. Clearly, the constant $h_0(\theta)$ in (10.287) drops out of (5.152) and may be ignored in solving (10.297).

For example, if we take (10.289) with $B = 0$, then (10.297) forces $\theta_1 = \theta_2 = 0$, whereas the magnetization $2\theta_3 \equiv m = \omega_\theta^\beta(\sigma_3)$ satisfies the famous ***gap equation***

$$\tanh(\beta m) = m. \tag{10.298}$$

For any β this has a solution $m = 0$, i.e., $\theta = 0$ in B^3, which corresponds to the tracial state $\omega(a) = \frac{1}{2}\mathrm{Tr}(a)$ normally associated with infinite temperature (i.e., $\beta = 0$). This state is evidently $\mathbb{Z}_2$-invariant. For $T \geq T_c = 1/4$ (i.e. $\beta \leq 4$) this is the only solution. For $T < T_c$ (or $\beta > 4$), two additional solutions $\pm m_\beta$ (with $m_\beta > 0$) appear, which break the $\mathbb{Z}_2$ symmetry. For $B > 0$ computations become tedious, but for $\beta \to \infty$, where ω_θ^β converges to the ground state of $\hat{h}_\theta$, one recovers our earlier conclusions.

Proposition 10.24. *The self-consistency equation* (10.297) *has at least one solution.*

Proof. This follows from Brouwer's Fixed Point Theorem (stating that any continuous map f from a compact compact set $K \subset \mathbb{R}^k$ to itself has a fixed point), applied to $K = S(M_n(\mathbb{C}))$ and $f(\theta) = \omega_\theta^\beta$, where $\theta \in S(M_n(\mathbb{C}))$, as just explained. $\square$

The key result on equilibrium states of homogeneous mean-field theories, then, is:

Theorem 10.25. *Let h_Λ in* (10.271) *define a homogeneous mean-field theory with compact symmetry group G. The sequence $(\omega_{\Lambda_N}^\beta)$ of local Gibbs states defined by* (9.96) *and* (10.271) *has a unique G-invariant limit point $\hat{\omega}^\beta$, whose decomposition into primary states is given by* (10.286). *The G-invariant probability measure μ_β is concentrated on some G-orbit in $S(M_n(\mathbb{C}))$, and the states ω_θ^β on $M_n(\mathbb{C})$ are given by* (10.296), *with Hamiltonians $\hat{h}_\theta$ defined by* (10.287), *where θ satisfies* (10.297).

Proof. We just sketch the proof, which is based on the Quantum De Finetti Theorem 8.9. Each operator $T_i^{(\Lambda_N)}$ is permutation-invariant, which property is transferred first to each local Hamiltonian h_{Λ_N}, thence to each local Gibbs state $\omega_{\Lambda_N}^\beta$ defined by h_{Λ_N}, and finally to each limit point of this sequence. As already noted, Theorem 8.9 then gives the decomposition (10.286), which by Theorem 8.29 (whose assumption holds in mean-field models) also gives the primary decomposition of $\hat{\omega}^\beta$ (i.e., each state $(\omega_\theta^\beta)^\infty$ is primary on the quasi-local algebra A). By our earlier argument centered on (10.294) - (10.295), time-evolution is implemented in the GNS-representation induced by such a state. An important step in the proof—which we omit because it requires various reformulations of the KMS condition we have not discussed—is that $(\omega_\theta^\beta)^\infty$ satisfies the KMS condition with respect to the dynamics (10.295). This, in turn, implies (10.296), which, by definition of θ through (10.290) - (10.291), gives the self-consistency condition (10.297). The proof is completed by a tricky argument (which again uses alternatives to the KMS condition) to the effect that if some ω_θ^β breaks the G-symmetry, the probability measure μ_β on the G-orbit in $S(M_n(\mathbb{C}))$ through ω_θ^β induced by the normalized Haar measure on G, defines the only possible limit point of the local Gibbs states, and hence must be unique. $\square$

Thus SSB can be detected by solving (10.297) and checking if the ensuing state(s) ω_θ^β on $M_n(\mathbb{C})$ is (are) G-invariant. As we have seen, in the Curie–Weiss model this is the case for $\beta \le 4$, whereas for $\beta > 4$ the measure μ_β in (10.286) is given by

$$\mu_\beta = \tfrac{1}{2}(\delta_{(0,0,m_\beta/2)} + \delta_{(0,0,-m_\beta/2)}), \tag{10.299}$$

where $\delta_\theta(f) = f(\theta)$. In such cases, since each local Gibbs state is invariant, one faces the (by now) familiar threat to Earman's Principle. In response, we expect Butterfield's Principle to be restored through the introduction of asymmetric flea-type perturbations to h_Λ that are localized in spin configuration space, although at nonzero temperature all excited states (rather than just the first) will start to play a role, and the precise details of the "flea" scenario remain to be settled.

10.9 The Goldstone Theorem

So far, we have only discussed the simplest of all symmetry groups, namely $G = \mathbb{Z}_2$, which is both finite and abelian. Although it will not change our picture of SSB, for the sake of completeness (and interest to foundations) we also present a brief introduction to continuous symmetries, culminating in the Goldstone Theorem and the Higgs mechanism (which at first sight contradict each other and hence require a very careful treatment). The former results when the broken symmetry group G is a Lie group, whereas the latter arises when it is an infinite-dimensional gauge group.

Let us start with the simple case $G = SO(2)$, acting on $\mathbb{R}^2$ by rotation. This induces the obvious action on the classical phase space $T^*\mathbb{R}^2$, i.e.,

$$R(p,q) = (Rp, Rq), \tag{10.300}$$

cf. (3.94), as well as on the quantum Hilbert space $H = L^2(\mathbb{R}^2)$, that is,

$$u_R \psi(x) = \psi(R^{-1}x). \tag{10.301}$$

Let us see what changes with respect to the action of $\mathbb{Z}_2$ on $\mathbb{R}$ considered in §10.1. We now regard the double-well potential V in (10.11) as an $SO(2)$-invariant function on $\mathbb{R}^2$ through the reinterpretation of x^2 as $x_1^2 + x_2^2$. This is the **Mexican hat potential**. Thus the classical Hamiltonian $h(p,q) = p^2/2m + V(q)$, similarly with $p^2 = p_1^2 + p_2^2$, is $SO(2)$-invariant, and the set of classical ground states

$$\mathscr{E}_0 = \{(p,q) \in T^*\mathbb{R}^2 \mid p = 0, q^2 = a^2\} \tag{10.302}$$

is the SO(2)-orbit through e.g. the point $(p_1 = p_2 = 0, q_1 = a, q_2 = 0)$. Unlike the one-dimensional case, the set of ground states is now connected and forms a circle in phase space, on which the symmetry group $SO(2)$ acts. The intuition behind the Goldstone Theorem is that a particle can freely move in this circle at no cost of energy. If we look at mass as inertia, such motion is "massless", as there is no obstruction. However, this intuition is only realized in quantum field theory. In quantum mechanics, the ground state of the Hamiltonian (10.6) (now acting on $L^2(\mathbb{R}^2)$) remains unique, as in the one-dimensional case. In polar coordinates (r, ϕ) we have

$$h_\hbar = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \phi^2} \right) + V(r), \tag{10.303}$$

with $V(r) = \frac{1}{4}\lambda (r^2 - a^2)^2$. With

$$L^2(\mathbb{R}^2) \cong L^2(\mathbb{R}^+) \otimes \ell^2(\mathbb{Z}) \tag{10.304}$$

under Fourier transformation in the angle variable, this becomes

$$h_\hbar \psi(r,n) = \left(-\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} - \frac{n^2}{r^2} \right) + V(r) \right) \psi(r,n). \tag{10.305}$$

Since $\hbar^2 n^2/2mr^2$ is positive, the ground state $\psi_\hbar^{(0)}$ has $\psi_\hbar^{(0)}(r,n) = 0$ for all $n \neq 0$, and hence it is $SO(2)$-invariant, since the $SO(2)$-action on $L^2(\mathbb{R}^2)$ becomes

$$u_\theta \psi(r,n) = \exp(in\theta)\psi(r,n), \tag{10.306}$$

after a Fourier-transform. Indeed, from a group-theoretical point of view, the unitary isomorphism (10.304) is nothing but the decomposition

$$L^2(\mathbb{R}^2) \cong \bigoplus_{n \in \mathbb{Z}} H_n, \tag{10.307}$$

where $H_n = L^2(\mathbb{R}^+)$ for all n, but with $\phi_n \in H_n$ transforming under $SO(2)$ as

$$u_\theta \phi_n(r) = \exp(in\theta)\phi_n(r) \; (\theta \in [0,2\pi]). \tag{10.308}$$

The $SO(2)$-invariant subspace of $L^2(\mathbb{R}^2)$, then, is precisely the space H_0 in which $\psi_\hbar^{(0)}$ lies. This is analogous to the situation occurring in one dimension higher (i.e. $\mathbb{R}^3$) with e.g. the hydrogen atom: in that case, the symmetry group is $SO(3)$, and $L^2(\mathbb{R}^3)$ decomposes accordingly as

$$L^2(\mathbb{R}^3) \cong \bigoplus_{j \in \mathbb{N}} H_j; \tag{10.309}$$

$$H_j = L^2(\mathbb{R}^+) \otimes \mathbb{C}^{2j+1}. \tag{10.310}$$

The ground state for a spherically symmetric potential, then, lies in H_0 and is $SO(3)$-invariant. For our purposes the relevant comparison is with the one-dimensional case: the decomposition of $L^2(\mathbb{R})$ under the natural $\mathbb{Z}_2$-action $u_{-1}\psi(x) = \psi(-x)$ is

$$L^2(\mathbb{R}) = H_0 \oplus H_1 \tag{10.311}$$

$$H_i = \{\psi \in L^2(\mathbb{R}) \mid \psi(x) = (-1)^i \psi(-x)\}, \; i = 0,1. \tag{10.312}$$

This time, H_+ is the $\mathbb{Z}_2$-invariant subspace containing the ground state $\psi_\hbar^{(0)}$. Being $\mathbb{Z}_2$-invariant, $\psi_\hbar^{(0)}$ is has peaks above both classical minima $\pm a$; in fact, $\psi_\hbar^{(0)}$ is real-valued and strictly positive. The ground state of the corresponding two-dimensional system, seen as an element of $L^2(\mathbb{R}^2)$, is just this wave-function $\psi_\hbar^{(0)}$ extended from $\mathbb{R}$ to $\mathbb{R}^2$ by rotational invariance. Hence the ground state remains real-valued and strictly positive, with peaks about the circle of classical minima in $\mathbb{R}^2$.

Let us recall the situation for $d = 1$ (cf. §10.1). The first excited state $\psi_\hbar^{(1)}$ lies in H_1; it is real-valued, like $\psi_\hbar^{(0)}$, but since it has to satisfy $\psi_\hbar^{(1)}(-x) = -\psi_\hbar(x)$, it cannot be positive. Indeed, with a suitable choice of phase, $\psi_\hbar^{(1)}$ has one positive peak above a and the same peak but now negative below $-a$. Then the wave-function

$$\psi_\hbar^\pm = (\psi_\hbar^{(0)} \pm \psi_\hbar^{(1)})\sqrt{2}, \tag{10.313}$$

is peaked above $\pm a$ alone (i.e., the negative peak of $\pm \psi_\hbar^{(1)}$ below $\mp a$ exactly cancels the corresponding peak of $\psi_\hbar^{(0)}$). The classical limit of $\psi_\hbar^{(0)}$ comes out as the mixed state $\frac{1}{2}(\omega_0^+ + \omega_0^-)$, where $\omega_0^\pm = (p = 0, \pm a)$, but each state $\psi_\hbar^\pm$ has the pure state $\omega_0^\pm$ as its classical limit. The latter are ground states, and hence in particular they are time-independent, because the energy difference $E^{(1)} - E^{(0)}$ between $\psi_\hbar^{(1)}$ and $\psi_\hbar^{(0)}$ vanishes (even exponentially fast) as $\hbar \to 0$.

A similar but more complicated situation arises in $d = 2$. The role of the pair

$$\left(\psi_\hbar^{(0)} \in H_0, \psi_\hbar^{(1)} \in H_1 \right)$$

is now played by an infinite tower of unit vectors

$$\left(\psi_\hbar^{(n)} \in H_n, n \in \mathbb{Z} \right),$$

where $\psi_\hbar^{(n)}$ is the lowest energy eigenstate (for $h_\hbar$ in (10.305)) in $H_n \subset L^2(\mathbb{R}^2)$. The analogue of the states $\psi_\hbar^\pm$ for $d = 1$ involves a limit which heuristically is like

$$\lim_{N \to \infty} \psi_\hbar^{(N,\theta)} = \frac{1}{\sqrt{2N+1}} \sum_{n=-N}^{N} u_\theta \psi_\hbar^{(n)}, \tag{10.314}$$

but this limit does not exist in $L^2(\mathbb{R}^2)$. As in §10.1, we instead rely on the technique explained around (10.4), which makes the unit vectors $\psi_\hbar^{(N,\theta)}$ converge to some probability measure $\mu_\hbar^\theta$ on $\mathbb{R}^2$ as $N \to \infty$. In the subsequent limit $\hbar \to 0$, one obtains a probability measure μ_0^θ concentrated on a suitable point in the orbit of classical ground states (10.302). Similarly, in the same sense the ground state $\psi_\hbar^{(0)}$ converges to a probability measure supported by all of $\mathcal{E}_0$.

To the extent that there is a Goldstone Theorem in classical mechanics, it would state that motion in the orbit $\mathcal{E}_0$ is free. That is, at fixed $(r = a, p_r = 0)$, where p_r is the radial component of momentum, one has an effective Hamiltonian

$$h_a(p_\phi, \phi) = \frac{p_\phi^2}{2ma^2}, \tag{10.315}$$

whose time-independent states $(p_\phi = 0, \phi_0)$ for arbitrary $\phi_0 \in [0, 2\pi)$ yield the ground states of the system, and whose "excited states"

$$(p_\phi(t), \phi(t)) = \left(p_\phi(0), \phi(0) + \frac{p_\phi(0)t}{ma^2} \right) \tag{10.316}$$

give motion along the orbit $\mathcal{E}_0$ with effective mass ma^2, whose energy converges to zero as $p_\phi \to 0$. However, since massless particles (whose existence is the main conclusion of the usual Goldstone Theorem) are not defined in classical mechanics, we now turn to relativistic field theory (with which we assume some familiarity).

We now illustrate SSB in classical field theory through a simple example, where the symmetry group is $G = SO(N)$, but whenever write things down in such a way that the generalization to arbitrary scalar field theories is obvious. Suppose we have N real scalar fields $\varphi \equiv (\varphi_1, \ldots, \varphi_N)$, on which $SO(N)$ acts in the defining representation on $\mathbb{R}^N$. Following the physics literature, from now on we sum over repeated indices like i and μ (**Einstein summation convention**). Let the Lagrangian

$$\mathscr{L} = \tfrac{1}{2}\partial_\mu \varphi_i \partial^\mu \varphi_i - V(\varphi), \tag{10.317}$$

contain an $SO(N)$-invariant potential V, typically of the form (with $\varphi^2 \equiv \sum_{i=1}^N \varphi_i^2$)

$$V(\varphi) = -\frac{m^2}{2}\varphi^2 + \frac{\lambda}{4}\varphi^4, \tag{10.318}$$

where $\lambda > 0$, but m^2 may have either sign. If $m^2 < 0$, the minimum of V lies at $\varphi = 0$, but if $m^2 > 0$ the minima form the $SO(N)$-orbit through

$$\varphi^c = (v, 0, \cdots, 0); \tag{10.319}$$
$$v \equiv m/\sqrt{\lambda} = \|\varphi^c\|. \tag{10.320}$$

The idea is that the physical fields are excitations of the "vacuum state" φ^c, so that, instead of φ, as the appropriate "small oscillation" field one should use

$$\chi(x) = \varphi(x) - \varphi^c. \tag{10.321}$$

Consequently, the potential is expanded in a Taylor series for small χ as

$$V(\varphi) = V(\varphi^c) + \tfrac{1}{2}V_{ij}''\chi_i\chi_j + O(\chi^3); \tag{10.322}$$
$$V_{ij}'' \equiv \frac{\partial^2 V}{\partial \varphi_i \partial \varphi_j}(\varphi^c). \tag{10.323}$$

Note that the linear term vanishes because $V'(\varphi^c) = 0$. We now use the $SO(N)$-invariance of V, i.e., $V(g\varphi) = V(\varphi)$ for all $g \in SO(N)$. For $T_a \in \mathfrak{g}$ (i.e. the Lie algebra of G, realized by anti-symmetric traceless $N \times N$ matrices) this yields

$$\frac{d}{dt}V(e^{tT_a}\varphi)_{t=0} = 0 \Leftrightarrow \frac{\partial V(\varphi)}{\partial \varphi_i}(T_a)_{ij}\varphi_j = 0. \tag{10.324}$$

Differentiation with respect to φ_k and putting $\varphi = \varphi^c$ then gives

$$V_{ik}''(T_a)_{ij}\varphi_j^c = 0. \tag{10.325}$$

In general, let $H \subset G$ be the stabilizer of φ^c, i.e., $g \in H$ iff $g\varphi^c = \varphi^c$. In our example (10.318) - (10.319), we evidently have $H = SO(N-1)$. Then $T_a\varphi^c = 0$ for all generators T_a of the Lie algebra $\mathfrak{h}$ of H, so that there are

$$M \equiv \dim(G) - \dim(H) = \dim(G/H) = \dim(G \cdot \varphi^c) \tag{10.326}$$

linearly independent null eigenvectors of V'' (seen as an $N \times N$ matrix). This number equals the dimension of the submanifold of $\mathbb{R}^N$ where V assumes its minimum. In our example we have $M = N - 1$, since $\dim(SO(N)) = \frac{1}{2}N(N-1)$. We now perform an affine field redefinition, based on an affine coordinate transformation in $\mathbb{R}^N$ that diagonalizes the matrix V''. The original (real) fields were $\varphi = (\varphi_1, \ldots, \varphi_N)$, and the new (real) fields are $(\chi_1, \theta_2, \cdots, \theta_N)$, with

$$\chi_1 = \varphi_1 - v, \tag{10.327}$$

as in (10.321), and the **Goldstone fields** are defined, also in general, by

$$\theta_a = \frac{1}{v}\langle T_a \varphi^c, \varphi \rangle = \frac{1}{v}(T_a)_{ij}\varphi_j^c \varphi_i. \tag{10.328}$$

Here $\langle \cdot, \cdot \rangle$ denotes the inner product in $\mathbb{R}^N$, and we have chosen a basis of $\mathfrak{g}$ in which the elements $(T_1, \ldots, T_{\dim(H)})$ form a basis of $\mathfrak{h}$, completed by M further elements $(T_{\dim(H)+1}, \ldots T_{\dim(G)+1})$, so as to have basis of $\mathfrak{g}$. The index a in (10.328), then, runs from $\dim(H) + 1$ to $\dim(G)$, so that there are M Goldstone fields, cf. (10.326). In our running example, this number was shown to be $M = N - 1$, and in view of (10.319), the field $\theta_a = (T_a)_{i1}\varphi_i$ is a linear combination of the φ_2 till φ_N.

The simplest example is $N = 2$, with potential (10.318) and $m^2 > 0$. With the single generator $T = -i\sigma_2$, we obtain $\theta = \varphi_2$. Since $V'' = \mathrm{diag}(2m^2, 0)$, we see that the mass term $-\frac{1}{2}m^2\varphi_1^2$ in (10.318) (with $\varphi^2 = \varphi_1^2 + \varphi_2^2$) changes from the "wrong" sign $-m^2$ to the 'right' sign $+2m^2$ in (10.322), whilst $-\frac{1}{2}m^2\varphi_2^2$ in (10.318) disappears, so that the field θ comes out to be massless. Indeed, this is the point of the introduction of the Goldstone fields: in view of (10.325) and (10.328), the Goldstone fields do not occur in the quadratic term in (10.322) and hence they are massless, in satisfying a field equation of the form $\partial_\mu \partial^\mu \theta_a = \cdots$, where $\cdots$ does not contain any term linear in any field. This proves the **classical Goldstone Theorem**:

Theorem 10.26. *Suppose that a compact Lie group $G \subset SO(N)$ acts on N real scalar fields $\varphi = (\varphi_1, \ldots, \varphi_N)$, leaving the potential V in the Lagrangian* (10.317) *invariant. If G is spontaneously broken to an unbroken subgroup $H \subset G$ (in the sense that the stability group of some point φ^c in the G-orbit minimizing V is H), then there are at least* $\dim(G/H)$ *massless fields, i.e., there is a field transformation*

$$(\varphi_1, \ldots, \varphi_N) \mapsto (\chi_1, \ldots, \chi_{N-M}, \theta_1, \ldots, \theta_M) \ (M = \dim(G) - \dim(H)), \tag{10.329}$$

that is invertible in a neighborhood of $\varphi = \varphi^c$, such that the potential $V(\varphi)$, re-expressed in the fields χ and θ, has no quadratic terms in θ.

The local invertibility of the field redefinition around $\varphi^c \neq 0$ is crucial; in our example, where $\chi \equiv \chi_1 = \varphi_1 - v$ and $\theta_a = T_{i1}^a \varphi_i$, this may be checked explicitly.

An alternative proof of Theorem 10.26 uses nonlinear Goldstone fields, viz.

$$\varphi(x) = e^{\frac{1}{v}\theta_a(x)T_a}(\varphi^c + \chi(x)), \tag{10.330}$$

where the sum over a (implicit in the Einstein summation convention) ranges from 1 to M, $v = \|\varphi^c\|$, and the fields $\chi = (\chi_1, \ldots, \chi_{N-M})$ are chosen orthogonal (in $\mathbb{R}^N$) to each $T_a \varphi_c$, $a = 1, \ldots, M$, and hence to the θ_a. Provided that the generators of $SO(N)$ (and hence of $G \subset SO(N)$) have been chosen such that

$$\langle T_a \varphi^c, T_b \varphi^c \rangle = v^2 \delta^{ab}, \tag{10.331}$$

the fields θ^a defined by (10.330) coincide with the fields in (10.328) up to quadratic terms in χ and θ; to see this, expand the exponential and also use the fact that both $\langle T_a \varphi^c, \varphi^c \rangle$ and $\langle T_a \varphi^c, \chi \rangle$ vanish. This transformation is only well defined if $v \neq 0$, i..e., if SSB from G to H occurs, and its existence implies the Goldstone Theorem 10.26, for by (10.330) and G-invariance, $V(\varphi)$ is independent of θ.

The Goldstone Theorem can be derived in quantum field theory, but in the spirit of this chapter we will discuss it rigorously for quantum spin systems. Far from considering the most general case, we merely treat the simplest setting. We assume that A is a quasi-local C*-algebra given by (8.130), with $H = \mathbb{C}^n$. Furthermore:

1. The group of space translations $\mathbb{Z}^d$ acts on A by automorphisms τ_x, and so does the group $\mathbb{R}$ of time translations by automorphisms α_t commuting with the τ_x (cf. §9.3); we often write $\alpha_{(x,t)}$ for $\alpha_t \circ \tau_x$ as well as $a(x,t)$ for $\alpha_t \circ \tau_x(a)$.
2. A compact Lie group G acts on $H = \mathbb{C}^n$ through a unitary representation u and hence acts on on A by automorphisms γ_g as in (10.58) - (10.59), such that

$$\gamma_g \circ \alpha_{(x,t)} = \alpha_{(x,t)} \circ \gamma_g \quad ((x,t) \in \mathbb{Z}^d \times \mathbb{R}, g \in G). \tag{10.332}$$

3. There exists a pure translation-invariant ground state ω.
4. One has SSB in that $\omega \circ \gamma_g \neq \omega$ for all $g \in G_a \subset G$, where

$$G_a = \{\exp(sT_a), s \in \mathbb{R}, T_a \in \mathfrak{g}\}. \tag{10.333}$$

5. There is an n-tuple $\varphi = (\varphi_1, \ldots, \varphi_n)$ of local operators $\varphi_\alpha \in M_n(\mathbb{C})$ that transforms under G by $\varphi \mapsto u_g \varphi u_g^* = \gamma_g(\varphi)$, and defines an order parameter ϕ_a by

$$\phi_a = \delta_a \varphi \equiv \frac{d}{ds} \left(\gamma_{\exp(sT_a)}(\varphi) \right)_{|s=0}, \tag{10.334}$$

at least for SSB of G_a (as above) in that, cf. Definition 10.6,

$$\omega(\delta_a \varphi) \neq 0. \tag{10.335}$$

6. Writing $j_a^0 = iu'(T_a) \in M_n(\mathbb{C})$, it follows that $\delta_a \varphi = -i[j_a^0, \varphi]$, and hence that

$$\delta_a \varphi(x) = -i \lim_{\Lambda \nearrow \mathbb{Z}^d} \sum_{y \in \Lambda} [j_a^0(y), \varphi(x)] \quad (x \in \mathbb{Z}^d), \tag{10.336}$$

since by (8.132) (i.e., Einstein locality) only the term $y = x$ will contribute. Physicists then wish to define a charge by $Q_a = \sum_{y \in \mathbb{Z}^d} j_a^0(y)$ and write (10.336) as $\delta_a \varphi(x) = -i[Q_a, \varphi(x)]$, but Q_a does not exist precisely in the case of SSB!

Eq. (10.336) motivates the crucial assumption for the Goldstone Theorem, viz.

$$\omega(\delta_a\varphi(x,t)) = -i \lim_{\Lambda \nearrow \mathbb{Z}^d} \sum_{y\in\Lambda} \omega([j_a^0(y),\varphi(x,t)]) \ (x\in\mathbb{Z}^d, t\in\mathbb{R}), \qquad (10.337)$$

which incorporates the condition that the sum over y converge absolutely.
Although (10.337) at first sight *softens* (10.336) in turning an operator equation
into a numerical one, in fact (10.337) decisively *sharpens* (10.336) by involving
the time-dependence of φ, whose propagation speed should be sufficiently small
for enabling the limit in (10.336) to catch up with the limit in (10.337). As such,
eq. (10.337) is satisfied with short-range forces, but the Meissner effect in su-
perconductivity and the closely related Higgs mechanism in gauge theories (both
of which circumvents the Goldstone Theorem) are possible precisely because in
those cases (10.337) fails (at least in physical gauges, see also §10.10).
7. Finally, we make two assumptions just for convenience, namely

$$\varphi_\alpha(x)^* = \varphi_\alpha(x); \qquad (10.338)$$
$$\omega(\varphi_\alpha(x)) = 0. \qquad (10.339)$$

If these are not the case, one could simply take real and imaginary components
of φ_α and/or redefine φ_α as $\tilde\varphi_\alpha = \varphi_\alpha - \omega(\varphi_\alpha)\cdot 1_A$, so that $\omega(\tilde\varphi_\alpha(x)) = 0$.

The Goldstone Theorem provides information about the joint-energy momentum
spectrum of the theory at hand. To define this notion, we exploit the fact that from
assumption no. (3) and Corollary 9.12 we obtain a unitary representation u_ω of
the (locally compact) abelian space-time translation group $A = \mathbb{Z}^d \times \mathbb{R}$ on the GNS-
representation space H_ω induced by ω. The SNAG-Theorem C.114 applied to A, with
dual $\hat{A} = \mathbb{T}^d \times \mathbb{R}$ (cf. Proposition C.108), then yields a projection-valued measure

$$e_\omega : \mathscr{B}(\mathbb{R}\times\mathbb{T}^d) \to \mathscr{P}(H_\omega), \qquad (10.340)$$

as a map from the Borel sets in $\mathbb{R}\times\mathbb{T}^d$ to the projection lattice in $B(H_\omega)$, such that

$$1_{H_\omega} = \int_{\mathbb{T}^d}\int_0^\infty de(E,k); \qquad (10.341)$$
$$u_\omega(y,t) = \int_{\mathbb{T}^d}\int_0^\infty de(E,k)\, e^{i(Et-y\cdot k)} \ (y\in\mathbb{Z}^d, t\in\mathbb{R}). \qquad (10.342)$$

Here $k = (k_1,\ldots,l_d)$, $y\cdot k = \sum_{i=1}^d y_i k_i$, and we have reduced the integration range over
E (which *a priori* would be $\mathbb{R}$) to $\mathbb{R}^+$. Indeed, by Stone's Theorem we have $u_\omega(t) = \exp(ith_\omega)$, where $\sigma(h_\omega) \subset [0,\infty)$ because ω is a ground state by assumption, and
the support of e is evidently contained in $\mathbb{Z}^d \times \sigma(h_\omega)$ (cf. Definition A.16).

Definition 10.27. *The **joint energy-momentum spectrum** $\sigma(h_\omega, p_\omega)$ of a space-
time invariant state ω (i.e., $\omega \circ \alpha_{(x,t)} = \omega$, $(x,t)\in\mathbb{Z}^d \times \mathbb{R}$) is the support of the
projection-valued measure e_ω associated to the GNS-representation π_ω, i.e., the
smallest closed set $\sigma(h_\omega, p_\omega) \subset \mathbb{T}^d \times \mathbb{R}$ such that $e((\mathbb{T}^d \times \mathbb{R})\backslash\sigma(h_\omega, p_\omega)) = 0$.*

The notation $\sigma(h_\omega, p_\omega)$ is purely symbolic here, since (as opposed to the continuum case) the group $\mathbb{Z}^d$ of spatial translations is discrete and hence has no generators p_ω.

Since $u_\omega(x,t)\Omega_\omega = \Omega_\omega$, the origin $(0,0)$ certainly lies in $\sigma(h_\omega, p_\omega)$, with

$$e_\omega(0,0) = |\Omega_\omega\rangle\langle\Omega_\omega|, \tag{10.343}$$

which by Theorem 9.14 is the unique $\mathbb{T}^d \times \mathbb{R}$-invariant state in H_ω. Denoting this contribution to e_ω by $e_\omega^{(0)}$, in many physical theories one has $e_\omega = e_\omega^{(0)} + e_\omega^{(1)} + \cdots$, where $e_\omega^{(1)}$ is supported on the graph of some continuous function $k \mapsto \varepsilon_k \geq 0$, i.e.,

$$\{(k, \varepsilon_k), k \in \mathbb{T}^d\} \subset \sigma(h_\omega, p_\omega) \subset \mathbb{T}^d \times \mathbb{R}. \tag{10.344}$$

The joint energy-momentum spectrum may be studied in part by considering

$$
\begin{aligned}
f(\varepsilon, p) &= \sum_{y \in \mathbb{Z}^d} \int_{-\infty}^{\infty} dt\, e^{-i\varepsilon t + ip\cdot(x-y)} \omega([j_a^0(y), \varphi(x,t)]) \\
&= 2i \sum_{y \in \mathbb{Z}^d} \int_{-\infty}^{\infty} dt\, e^{-i\varepsilon t + ip\cdot(x-y)} \mathrm{Im}\langle\Omega_\omega, \pi_\omega(j_a^0(0))e^{ith_\omega}u_\omega(y)\pi_\omega(\varphi_\alpha(0))\Omega_\omega\rangle \\
&= \int_{\mathbb{T}^d} \int_0^{\infty} \left(\langle\Omega_\omega, \pi_\omega(j_a^0(0))de_\omega(E,k)\,\pi_\omega(\varphi_\alpha(0))\Omega_\omega\rangle\delta(\varepsilon - E)\delta(p - k) \right. \\
&\quad \left. - \langle\Omega_\omega\pi_\omega(\varphi_\alpha(0))de_\omega(E,k)\,\pi_\omega(j_a^0(0))\Omega_\omega\rangle\delta(\varepsilon + E)\delta(p + k) \right), \tag{10.345}
\end{aligned}
$$

i.e., the Fourier transform of the two-point function defined by j_a^0 and φ, which is a distribution on the dual group $\mathbb{T}^d \times \mathbb{R}$; for the third equality we used a distributional version of the Fourier inversion formula (C.382). For example, if we replace $e_\omega(E,k)$ by $e_\omega^{(1)}(E,k)$, then, since $e_\omega^{(1)}$ is absolutely continuous with respect to Haar measure $d^d k$ on $\mathbb{T}^d$, we see that $f(\varepsilon, p)$ is proportional to $\delta(\varepsilon - \varepsilon_p)$.

Theorem 10.28. *Under assumptions 1–7 (notably (10.337) and* SSB *of some continuous symmetry), the Hamiltonian h_ω has continuous spectrum starting at zero and hence has no gap. If there is an excitation spectrum $e_\omega^{(1)}$ as explained above, with*

$$\int \langle\Omega_\omega, \pi_\omega(j_a^0(0))de_\omega^{(1)}(E,k)\,\pi_\omega(\varphi_\alpha(0))\Omega_\omega\rangle \neq 0, \tag{10.346}$$

then the continuous function $k \mapsto \varepsilon_k$ defining the spectrum satisfies $\varepsilon_0 = 0$.

Proof. Since the sum in (10.337) converges absolutely, the Fourier transform $\check{f}(t, p)$ of $y \mapsto \omega([j_a^0(y), \varphi(x,t)])$ in y alone is continuous in p, and by (10.337) we have

$$i\omega(\delta_a\varphi(x,t)) = \check{f}(t,0). \tag{10.347}$$

By (10.332), the left-hand side is independent of x and t, hence the Fourier transform $f(\varepsilon, 0)$ of the right-hand side in t is proportional to $\delta(\varepsilon)$. Since (10.343) does not contribute to f by (10.339), the calculation (10.345) shows that $f(\varepsilon, 0) = 0$ if $\sigma(h_\omega)$ has a gap. But $f(\varepsilon, 0) \neq 0$ by (10.335), and so $\sigma(h_\omega)$ has no gap. Similarly, for the final claim note that $f(\varepsilon, 0) \sim \delta(\varepsilon - \varepsilon_0)$ as well as $f(\varepsilon, 0) \sim \delta(\varepsilon)$. $\qquad\square$

10.10 The Higgs mechanism

We proceed to a discussion of SSB in gauge theories, especially with an eye on the Higgs Mechanism, which plays a central role in the Standard Model of high-energy physics (whose empirical confirmation was more or less finished with the discovery of the Higgs boson at CERN, announced on July 4, 2012).

We look at the **Abelian Higgs Model**, given by the Lagrangian

$$\mathscr{L} = -\tfrac{1}{4}F_A^2 + \tfrac{1}{2}\langle D_\mu^A \varphi, D_\mu^A \varphi \rangle - V(\varphi), \tag{10.348}$$

where $\varphi = (\varphi_1, \varphi_2)$ is a scalar doublet, the usual electromagnetic field strength is

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{10.349}$$

in terms of which $F_A^2 = F_{\mu\nu}F^{\mu\nu}$, and the covariant derivative is

$$D_\mu^A \equiv \partial_\mu - eA_\mu \cdot T = \partial_\mu \cdot 1_2 + ieA_\mu \cdot \sigma_2. \tag{10.350}$$

Here e is some coupling constant, identified with the unit of electrical charge. We still assume that V only depends on $\|\varphi\|^2 = \langle \varphi, \varphi \rangle$ and hence is $SO(2)$-invariant.

The novel situation compared to (10.317) and the like is that, whereas (10.317) is invariant under *global* $SO(2)$ transformations, the Lagrangian (10.348) is invariant under *local* $SO(2)$ **gauge transformations** that depend on x, namely

$$\varphi(x) \mapsto e^{\alpha(x)\cdot T}\varphi(x) = \begin{pmatrix} \cos\alpha(x) & -\sin\alpha(x) \\ \sin\alpha(x) & \cos\alpha(x) \end{pmatrix} \cdot \begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix}; \tag{10.351}$$

$$A_\mu(x) \mapsto A_\mu(x) + \frac{1}{e}\partial_\mu\alpha(x). \tag{10.352}$$

We say that the **local gauge group** $\mathscr{G} = C^\infty(\mathbb{R}^d, U(1))$ acts on the space of fields (A, φ) by (10.351) - (10.352). Now suppose V has a minimum at some constant value $\varphi^c \neq 0$. In that case, any field configuration

$$\varphi(x) = \exp(\alpha(x) \cdot T)\varphi^c; \tag{10.353}$$
$$A_\mu(x) = (1/e)\partial_\mu\alpha(x)) \ (\alpha \in \mathscr{G}), \tag{10.354}$$

minimizes the action. Hence the possible "vacua" of the model comprise the (infinite-dimensional) orbit $\mathscr{V}$ of the gauge group through $(A = 0, \varphi = \varphi^c)$. Note that $D_\mu^A \varphi = 0$ for $(A, \varphi) \in \mathscr{V}$, i.e., φ is *covariantly* constant along the vacuum orbit (whereas for global symmetries it is constant full stop). Relative to the (arbitrary) choice $(0, \varphi^c) \in \mathscr{V}$, we then introduce real fields χ and θ, called the **Higgs field** and the **would-be Goldstone boson**, respectively, by (10.330), which now simply reads

$$\begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix} = e^{\frac{1}{v}\theta(x)\cdot T} \cdot \begin{pmatrix} v + \chi(x) \\ 0 \end{pmatrix}. \tag{10.355}$$

After this redefinition of the scalar fields, the Lagrangian (10.348) becomes

$$\mathscr{L} = -\tfrac{1}{4}F_B^2 + \tfrac{1}{2}\partial_\mu \chi \partial^\mu \chi + \tfrac{1}{2}e^2(v+\chi)^2 B_\mu B^\mu - V(v+\chi,0), \qquad (10.356)$$

where $B_\mu = A_\mu - (1/ev)\partial_\mu \theta$, and $F_B^2 = F_{\mu\nu}F^{\mu\nu}$ for $F_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu$. This describes a vector boson B with mass term $\tfrac{1}{2}m_B^2 B_\mu B^\mu$, with $m_B^2 = \tfrac{1}{2}e^2 v^2 > 0$ (as opposed to the massless vector field A), and a scalar field χ with mass term $\tfrac{1}{2}m_\chi^2 \chi^2$, with $m_\chi^2 = (\partial^2 V/\partial \phi_1^2)_{|(v,0)} > 0$ (since V supposedly has a minimum at $\varphi^c = (v,0)$).

This is the **_Higgs mechanism_**: the gauge field becomes massive, whilst the massless ("would-be") Goldstone boson disappears from the theory: it is (allegedly) "eaten" by the gauge field. Thus the scalar degree of freedom θ that seems lost is recovered as the longitudinal component of the massive vector field (which for a gauge field would have been an unphysical gauge degree of freedom, see below).

In the description just given, the Higgs mechanism in classical field theory is seen as a consequence of SSB. Remarkably, there is an alternative account of the Higgs mechanism, according to which it has nothing to do with SSB! Namely, we now perform a field redefinition analogous to (10.355) etc. straight away, viz.

$$\begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix} = e^{\theta(x)\cdot T} \cdot \begin{pmatrix} \rho(x) \\ 0 \end{pmatrix}; \qquad (10.357)$$

$$A_\mu = B_\mu + (1/e)\partial_\mu \theta. \qquad (10.358)$$

This transformation is defined and invertible in a neighbourhood of any point $(\rho_0, \theta_0, B_0$, where $\rho_0 > 0$, $\theta_0 \in (-\pi, \pi)$, and B_0 is arbitrary. Each of these new fields is gauge-invariant: for the gauge transformation (10.351) becomes

$$\theta(x) \mapsto \theta(x) + \alpha(x); \qquad (10.359)$$

$$\rho(x) \mapsto \rho(x), \qquad (10.360)$$

and in view of (10.352), B does not transform at all. The Lagrangian becomes

$$\mathscr{L} = -\tfrac{1}{4}F_B^2 + \tfrac{1}{2}\partial_\mu \rho \partial^\mu \rho + \tfrac{1}{2}e^2 \rho^2 B_\mu B^\mu - V(\rho), \qquad (10.361)$$

with $V(\rho) \equiv V(\rho,0)$. This is a Lagrangian without any internal symmetries at all (not even $\mathbb{Z}_2$, since $\rho > 0$), but of course one can still look for classical vacua that minimize the energy and hence the potential $V(\rho)$. If $\rho = 0$ is the absolue minimum, then the above field redefinition is _a fortiori_ invalidated, but if $V'(v) = 0$ for some $v > 0$, we proceed as before, introducing a Higgs field $\chi(x) = \rho(x) - v$, and recovering the Lagrangian (10.356). This once again leads to the Higgs mechanism.

This can be generalized to the nonabelian case; since it suffices to explain the idea, we just discuss the $SU(2)$ case. In (10.348), the scalar field $\varphi = (\varphi_1, \varphi_2)$ is now complex, forming an $SU(2)$ doublet, the brackets $\langle \cdot, \cdot \rangle$ now denote the inner product in $\mathbb{C}^2$, the nonabelian gauge field is $A = A^a \sigma_a$ (where the Pauli matrices σ_a, $a = 1, 2, 3$, form a self-adjoint basis of the Lie algebra of $SU(2)$), with associated field strength $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + g[A_\mu, A_\nu]$ and covariant derivative $D_\mu^A = \partial_\mu + igA_\mu$.

With $F_A^2 = F_{\mu\nu}^a F_a^{\mu\nu}$, the Lagrangian (10.348) is invariant under the transformations

$$\varphi(x) \mapsto e^{i\alpha_a(x)\sigma_a(x)}\varphi(x); \tag{10.362}$$

$$A_\mu(x) \mapsto e^{i\alpha_a(x)\sigma_a(x)}(A_\mu(x) - (i/g)\partial_\mu)e^{-i\alpha_a(x)\sigma_a(x)}. \tag{10.363}$$

The definition of the gauge-invariant fields B and ρ à la (10.357) - (10.358) is now

$$\begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix} = e^{i\theta_a(x)\cdot\sigma_a} \cdot \begin{pmatrix} \rho(x) \\ 0 \end{pmatrix}; \tag{10.364}$$

$$A_\mu(x) = e^{i\theta_a(x)\sigma_a(x)}(B_\mu(x) - (i/g)\partial_\mu)e^{-i\theta_a(x)\sigma_a(x)}, \tag{10.365}$$

which leads, *mutatis mutandis*, to the very same Lagrangian (10.361).

As a compromise between these two derivations of the Higgs mechanism, it is also possible to fix the gauge by picking the representative (φ, A) in each $\mathscr{G}$-orbit for which $\varphi_2(x) = 0$ and $\varphi_1(x) > 0$; note that this so-called **unitary gauge** is ill-defined if $\varphi_1(x) = 0$. Calling this unique representative (ρ, B), we are again led to (10.361).

Gauge field theories are **constrained systems**, in which the *apparent* degrees of freedom in the Lagrangian are not the *physical* ones. For free electromagnetism, the Lagrangian is $\mathscr{L}(A) = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, with $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. In terms of the gauge-invariant fields $E_i = F_{i0} = \partial_i A_0 - \partial_0 A_i$ and $\mathbf{B} = \nabla \times \mathbf{A}$, Maxwell's equations

$$\nabla \cdot \mathbf{E} = 0; \tag{10.366}$$

$$\partial \mathbf{E}/\partial t = \nabla \times \mathbf{B}; \tag{10.367}$$

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}; \tag{10.368}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{10.369}$$

then arise as follows: eqs. (10.366) and (10.367) correspond to the Euler–Lagrange equation for A_0 and A_i, respectively, whereas (10.368) and (10.369) immediately follow from the definitions of $\mathbf{B}$ and $\mathbf{E}$ in terms of A. The Maxwell equations are in Hamiltonian form, with canonical momenta $\Pi_\mu = \partial\mathscr{L}/\partial\dot{A}_\mu$; this yields $\Pi_i = -E_i$, as well as the **primary constraint** $\Pi_0 = 0$. Nonetheless, the canonical Hamiltonian

$$h = \int d^3x \left(\Pi_\mu(x)\dot{A}_\mu(x) - \mathscr{L}(x) \right) = \int d^3x \left(\tfrac{1}{2}\mathbf{E}^2(x) + \tfrac{1}{2}\mathbf{B}^2(x) - A_0(x)\nabla \cdot \mathbf{E}(x) \right)$$

is well defined. In the Hamiltonian formalism, Gauss' Law resurfaces as the **secondary constraint** stating that the primary constraint be preserved in time, viz.

$$\dot{\Pi}_0(x) = -\frac{\delta h}{\delta A_0(x)} = \nabla \cdot \mathbf{E}(x) \equiv 0. \tag{10.370}$$

Since

$$\frac{d}{dt}\nabla \cdot \mathbf{E}(x) = -\partial_i(\delta h/\delta A_i(x)) = -\partial_i(\Delta A_i - \partial_i \nabla \cdot \mathbf{A}) = 0, \tag{10.371}$$

there are no "tertiary" constraints. Thus we have canonical phase space variables $(\mathbf{E}, \mathbf{A})$ and (Π_0, A_0), subject to (10.366) and to $\Pi_0(x) = 0$ for each $x \in \mathbb{R}^3$, i.e.,

$$\Pi_0(\lambda_0) \equiv \int d^3x \, \Pi_0(x) \lambda_0(x) = 0; \tag{10.372}$$

$$\Pi(\lambda) \equiv \int d^3x \, \nabla \cdot \mathbf{E}(x) \lambda(x) = 0, \tag{10.373}$$

for all (reasonable) functions λ_0 and λ on $\mathbb{R}^3$. The constraints (10.372) - (10.373) are **first class** in the sense of Dirac, which means that their Poisson brackets are equal to existing constraints (or zero). In the Hamiltonian formalism, the role of the space-time dependent gauge transformations of the Lagrangian theory is played by the canonical transformations generated by the first class constraints, i.e.,

$$\delta_{\lambda_0} A_0(x) = \{\Pi_0(\lambda_0), A_0(x)\} = \lambda_0(x); \tag{10.374}$$

$$\delta_{\lambda_0} A_i(x) = \delta_{\lambda_0} E_i(x) = 0; \tag{10.375}$$

$$\delta_\lambda \mathbf{A}(x) = \nabla \lambda(x); \tag{10.376}$$

$$\delta_\lambda \mathbf{E}(x) = 0; \tag{10.377}$$

$$\delta_\lambda A_0(x) = 0. \tag{10.378}$$

The holy grail of the Hamiltonian formalism is to find variables that are both *gauge invariant* and *unconstrained*. In our case, $A_\mu = (A_0, \mathbf{A})$ are unconstrained but gauge variant, whilst $\Pi_\mu = (\Pi_0, -\mathbf{E})$ are gauge invariant but constrained! Now write some vector field $\mathbf{V}$ as $\mathbf{V} = \mathbf{V}^L + \mathbf{V}^T$, where $\mathbf{V}^L = \Delta^{-1}\nabla(\nabla \cdot \mathbf{V})$ is the longitudinal component, so that $V_i^T = (\delta_{ij} - \Delta^{-1}\partial_i\partial_j)V_j$ is the transverse part. Then the physical variables of free electromagnetism are $\mathbf{A}^T$ and $\mathbf{E}^T$. The physical Hamiltonian

$$h = \tfrac{1}{2} \int d^3x \, (\mathbf{E}^T \cdot \mathbf{E}^T - \mathbf{A}^T \cdot \Delta \mathbf{A}^T), \tag{10.379}$$

then, is well defined on the physical (or reduced) phase space, which is the subset of all (A_μ, Π_μ) where the constraints (10.373) hold, modulo gauge equivalence.

After this preparation, we now revisit the abelian Higgs model as a constrained Hamiltonian system. It is convenient to combine the two real scalar fields φ_1 and φ_2 into a single complex scalar field $\varphi = (\varphi_1 + i\varphi_2)/\sqrt{2}$, and treat φ and its complex conjugate $\overline{\varphi}$ as independent variables. The Lagrangian (10.348) then becomes

$$\mathscr{L} = -\tfrac{1}{4}F_A^2 + \overline{D_\mu^A \varphi} \cdot D_\mu^A \varphi - V(\varphi, \overline{\varphi}), \tag{10.380}$$

with $D_\mu^A \varphi = (\partial_\mu - ieA_\mu)\varphi$, etc. The conjugate momenta Π_μ to A_μ are the same as for free electromagnetism, i.e., $\Pi_0 = 0$ and $\Pi_i = -E_i$, and for φ we obtain

$$\pi = \partial \mathscr{L}/\partial \dot{\varphi} = \overline{D_0^A \varphi}; \tag{10.381}$$

$$\overline{\pi} = \partial \mathscr{L}/\partial \dot{\overline{\varphi}} = D_0^A \varphi. \tag{10.382}$$

The associated Hamiltonian h is equal to

$$\int d^3x \left(\tfrac{1}{2}\mathbf{E}^2 + \tfrac{1}{2}\mathbf{B}^2 - A_0(\nabla \cdot \mathbf{E} - j_0) + \overline{\pi}\pi + \overline{D_i^A \varphi} \cdot D_i^A \varphi + V(\varphi, \overline{\varphi}) \right), \quad (10.383)$$

where $j_0 = ie(\pi\varphi - \overline{\pi}\overline{\varphi})$ is the zero'th component of the Noether current. Hence the primary constraint remains $\Pi_0 = 0$, but the secondary constraint picks up an additional term and becomes $\nabla \cdot \mathbf{E} = j_0$ (which remains Gauss' law!). The physical (i.e., gauge invariant and unconstrained) variables can be computed as

$$\varphi_A = e^{ie\Delta^{-1}\nabla\cdot\mathbf{A}}\varphi, \quad \overline{\varphi}_A = e^{-ie\Delta^{-1}\nabla\cdot\mathbf{A}}\overline{\varphi}; \quad\quad (10.384)$$

$$\pi_A = e^{-ie\Delta^{-1}\nabla\cdot\mathbf{A}}\varphi, \quad \overline{\pi}_A = e^{ie\Delta^{-1}\nabla\cdot\mathbf{A}}\overline{\pi}, \quad\quad (10.385)$$

plus the same transverse fields $\mathbf{A}^T$ and $\mathbf{E}^T$, as in free electromagnetism. In terms of the transverse covariant derivative $D_i^T = \partial_i - ieA_i^T$, the physical Hamiltonian h is

$$\int d^3x \left(\tfrac{1}{2}(\mathbf{E}^T \cdot \mathbf{E}^T - \mathbf{A}^T \cdot \Delta\mathbf{A}^T - j_0^A \Delta^{-1} j_0^A) + \overline{\pi}_A \pi_A + \overline{D_i^T \varphi_A} \cdot D_i^T \varphi_A + V(\varphi_A, \overline{\varphi}_A) \right).$$
$$(10.386)$$

The third term in (10.386) is the Coulomb energy, in which the charge density

$$j_0^A = ie(\pi_A \varphi_A - \overline{\pi}_A \overline{\varphi}_A) \quad\quad (10.387)$$

is the same as j_0 (since the latter is gauge invariant). Remarkably, the physical field variables carry a residual *global U(1)*-symmetry, viz.

$$\varphi_A \mapsto \exp(i\alpha)\varphi_A; \quad\quad (10.388)$$

$$\pi_A \mapsto \exp(-i\alpha)\pi_A; \quad\quad (10.389)$$

$$\overline{\varphi}_A \mapsto \exp(-i\alpha)\overline{\varphi}_A; \quad\quad (10.390)$$

$$\overline{\pi}_A \mapsto \exp(i\alpha)\overline{\pi}_A, \quad\quad (10.391)$$

and no change for $\mathbf{A}^T$ and $\mathbf{E}^T$, under which the Hamiltonian (10.386) is invariant.

If V has a minimum at $\varphi = \overline{\varphi} = v$, we recover the Higgs mechanism: redefining

$$\varphi_A = \exp(i\theta/v)(v + \chi), \quad\quad (10.392)$$

and complex conjugate, and the reintroduction of the longitudinal components

$$A_i^L = -(1/ev)\partial_i\theta; \quad E_i^L = -ev\Delta^{-1}\partial_i\pi_\theta, \quad\quad (10.393)$$

of the gauge field and its conjugate momentum, the Hamiltonian (10.386) becomes

$$\tfrac{1}{2}\int d^3x \left(\mathbf{E}^2 + \mathbf{B}^2 + \pi_\chi^2 + \partial_i\chi\partial_i\chi + \frac{(\nabla \cdot \mathbf{E})^2}{e^2 v^2} + e^2 v^2 \mathbf{A}^2 + V(v + \chi) \right), \quad (10.394)$$

where $\mathbf{A} = \mathbf{A}^T + \mathbf{A}^L$ and $\mathbf{E} = \mathbf{E}^T + \mathbf{E}^L$. This describes a massive vector field, and the would-be Goldstone boson θ has disappeared, as befits the Higgs mechanism!

It is fair to say that the Higgs mechanism in quantum field theory—and more generally, the notion of SSB in gauge theories—is poorly understood. Indeed, the entire quantization of gauge theories is not well understood, except at the perturbative level or on a lattice. The problems already come out in the abelian case with $d = 3$. The main culprit is Gauss' Law $\nabla \cdot \mathbf{E} = j_0$. One would naively expect this constraint to remain valid in quantum field theory as an operator equation, and this is indeed the case in so-called physical gauges like the Coulomb gauge (i.e. $\partial_i A_i = 0$). If we now look at condition (10.337) in §10.9, which for $G = U(1)$ and for example $\delta \varphi_1 = \varphi_2$ and $\delta \varphi_2 = -\varphi_1$ for a charged field $\varphi = (\varphi_1 + i\varphi_2)/\sqrt{2}$, or $\delta \varphi = i\varphi$, reads

$$\lim_{\Lambda \nearrow \mathbb{R}^3} \int_\Lambda d^3 y \, \omega([j_0(y,0), \varphi_\alpha(x,t)]) = -i\omega(\delta \varphi_\alpha(x,t)), \qquad (10.395)$$

then it is clear that (10.395) can only hold if charged fields are nonlocal. For by Gauss' Law the commutator $[j_0(y,0), \varphi_\alpha(x,t)]$ equals $[\nabla \cdot \mathbf{E}(0,y), \varphi_\alpha(x,t)]$, and by Gauss'(!) Theorem in vector calculus, all contributions to the left-hand side of (10.395) come from terms $[E_i(0,y), \varphi_\alpha(x,t)]$, with $y \in \partial \Lambda$ (i.e., the boundary of Λ). These must remain nonzero if $\Lambda \nearrow \mathbb{R}^3$, at least if (10.395) holds. On the other hand, such nonlocality must be enforced by massless fields, which idea leads to one of the very few rigorous result about the Higgs mechanism (in the continuum):

Theorem 10.29. *In the Coulomb gauge the following conditions are equivalent:*

- *The electromagnetic field $\mathbf{A}$ is massless;*
- *Eq. (10.395) holds for any field φ_α;*
- *The charge operator $Q = \lim_{\Lambda \uparrow \mathbb{R}^3} \int_\Lambda d^3 y \, j_0(y,0)$ exists (on some suitable domain in H_ω containing Ω_ω) and satisfies $Q\Omega_\omega = 0$.*

Hence (contrapositively), SSB of $U(1)$ by the state ω is only possible if $\mathbf{A}$ is massive. In that case, the Fourier transform of the two-point function $\langle 0|\varphi_\alpha(x,x_0) j_0^a(y,y_0)|0\rangle$ (cf. the proof of the Goldstone Theorem 10.28 in §10.9) has a pole at the mass of $\mathbf{A}$.

This theorem indeed yields the Higgs mechanism for say the abelian Higgs model in a specific physical gauge: note that the idea that the would-be Goldstone boson is eaten by the gauge field is already suggested by Gauss' Law, through which (minus) the canonical momentum $\mathbf{E}$ to $\mathbf{A}$ acquires j_0 as its longitudinal component; that is, the very same field that creates the Goldstone boson from the ground state.

In covariant gauges, all fields remain local, but (10.395) is rescued by the gauge-fixing term added to the Lagrangian. For example, adding $\mathcal{L}_{gf} = -(1/2\xi)(\partial_\mu A^\mu)^2$ to (10.348) leads to an equation of motion $\partial_\mu F_v^\mu = j_v - \partial_v \partial_\mu A^\mu$, so that (discarding all surface terms by locality), one obtains

$$-i\omega(\delta \varphi_\alpha(x,t)) = \int_{\mathbb{R}^3} d^3 y \, \omega([\partial_0^2 A_0(y,0), \varphi_\alpha(x,t)]). \qquad (10.396)$$

In the proof of the Goldstone Theorem, the massless Goldstone bosons do emerge, but they turn out to lie in some "unphysical subspace" of H_ω (which, for local gauges, is not a Hilbert space but has zero- and negative norm states).

Notes

In a philosophical context, the notion of ***emergence*** is usually traced to J.S. Mill (1843), who drew attention to 'a distinction so radical, and of so much importance, as to require a chapter to itself', namely the one between what Mill calls the principle of the 'Composition of Causes', according to which the joint effect of several causes is identical with the sum of their separate effects, and the negation of this principle. For example, in the context of his overall materialism, Mill believed that although all 'organised bodies' are composed of material parts,

> 'the phenomena of life, which result from the juxtaposition of those parts in a certain manner, bear no analogy to any of the effects which would be produced by the action of the component substances considered as mere physical agents. To whatever degree we might imagine our knowledge of the properties of the several ingredients of a living body to be extended and perfected, it is certain that no mere summing up of the separate actions of those elements will ever amount to the action of the living body itself.'
> Mill (1952 [1843], p. 243)

Mill launched what is now called ***British Emergentism*** (Stephan, 1992; McLaughlin, 2008; O'Connor & Wong, 2012), a school of thought which seems to have ended with C.D. Broad, who has our sympathy over Mill because of the doubt he expresses in our quotation in the preamble. Among the British Emergentists, the most modern views seem to have been those of S. Alexander, who, as paraphrased in O'Connor & Wong (2012), was committed to a view of emergence as

> 'the appearance of novel qualities and associated, high-level causal patterns which cannot be directly expressed in terms of the more fundamental entities and principles. But these patterns do not supplement, much less supersede, the fundamental interactions. Rather, they are macroscopic patterns running through those very microscopic interactions. Emergent qualities are something truly new (...), but the world's fundamental dynamics remain unchanged.'

Alexander's idea that emergent qualities 'admit no explanation' and had 'to be accepted with the "natural piety" of the investigator foreshadowed the later notion of ***explanatory emergence***. Indeed, philosophers distinguish between *ontological* and *epistemological* reduction or emergence, but ontological emergence seems a relic from the days of vitalism and other immature understandings of physics and (bio)chemistry (including the formation of chemical compounds, which Broad and some of his contemporaries still saw as an example of emergence in the strongest possible sense, i.e., falling outside the scope of the laws of physics). Recent literature, including the present chapter, is concerned with epistemological emergence, of which explanatory emergence is a branch. For example, Hempel wrote:

> 'The concept of *emergence* has been used to characterize certain phenomena as 'novel', and this not merely in the psychological sense of being unexpected, but in the theoretical sense of being unexplainable, or unpredictable, on the basis of information concerning the spatial parts or other constituents of the systems in which the phenomena occur, and which in this context are often referred to as "wholes".'   (Hempel, 1965, p. 62)

See also Batterman (2002), Bedau & Humpreys (2008), Norton (2012), Silberstein (2002), Wayne & Arciszewski (2009), and many other surveys of emergence.

§10.1. **Spontaneous symmetry breaking: The double well**

The facts we use about the double-well Hamiltonian may be found in Garg (2000) or Landau & Lifshitz (1977) at a heuristic level (but with correct conclusions), or, rigorously, in Reed & Simon (1978), Simon (1985), Helffer (1988), and Hislop & Sigal (1996). Theorem 10.2 is Theorem XIII.47 in Reed & Simon (1978).

§10.2. **Spontaneous symmetry breaking: The flea**

The flea perturbation and its effect on the ground state were first described in Jona-Lasinio, Martinelli, & Scoppola (1981a,b), who used methods from stochastic mechanics. See also Claverie & Jona-Lasinio (1986). Using more conventional methods, their results were reconfirmed and analyzed further by e.g. Combes, Duclos, & Seiler (1983), Graffi, Grecchi, & Jona-Lasinio (1984), Helffer & Sjöstrand (1985), Simon (1985), Helffer (1988), and Cesi (1989). The "Flea on the Elephant" terminology used by Simon (1985) motivated the title of Landsman & Reuvers (2013), who, as will be explained in the next chapter, identified the proper host animal as a cat. All pictures in this section are taken from the latter paper (and were prepared by the second author). For the Eyring–Kramers formula see Berglund (2011) for mathematicians or Hänggi, Talkner, & Borkovec (1990) for physicists.

§10.3. **Spontaneous symmetry breaking in quantum spin systems**

The translation-non-invariant ground states mentioned after Proposition 10.5 are discussed e.g. in Example 6.2.56 in Bratteli & Robinson (1997). See also Liu & Emch (2005), which was in important source for this section, and Ruetsche (2011) for a discussion of the definition of SSB through non-implementability. For order parameters see e.g. Sewell (2002), §3.3. A proof of Proposition 10.8 may be found in Bratteli & Robinson (1997), Proposition 6.2.15.

§10.4. **Spontaneous symmetry breaking for short-range forces**

The idea of SSB goes back to Heisenberg(1928). The C*-algebraic approach in quantum spin systems with short-range forces is reviewed in Bratteli & Robinson (1997); see also Nachtergaele (2007). Theorem 10.10 is due to Araki (1974); see also Simon (1993), Theorem IV.5.6, and Bratteli & Robinson (1997), Theorem 6.2.18. In Definition 10.9, Araki required Ω_ω to be separating for $\pi_\omega(A)''$ instead of ω to be α_t-invariant, but in the presence of (10.53) and hence (10.53) these conditions are equivalent. The fact that (for short-range forces) global Gibbs states defined by (10.43) satisfy the KMS condition follows from Theorem 10.10, but this was the starting point of Haag, Hugenholtz, & Winnink (1967); see Winnink (1972).

Uniqueness of KMS states for one-dimensional quantum spin systems with short-range forces at any positive temperature (which also holds for the classical case, e.g. the one-dimensional Ising model) has been proved by Araki (1975). See also Mattis (1965) and Altland & Simons (2010) for some of the underlying physical intuition.

§10.5. **Ground state(s) of the quantum Ising chain**

Theorem 10.11.1 was first established in Pfeuty (1970) by explicit calculation, based on Lieb, Schultz, & Mattis (1961). For more information on the quantum Ising model (also in higher dimension) see e.g. Karevski (2006), Sachdev (2011), Suzuki et al (2013), and Dutta et al (2015). Uniqueness of the ground state of the quantum Ising model with $B \neq 0$ holds in any dimension d, as first shown by Campanino,

Klein & Perez (1991) on the basis of Perron–Frobenius type arguments similar to those for Schrödinger operators. The singular case $B = 0$ leads to a violation of the strict positivity conditions necessary to apply the Perron–Frobenius Theorem, and this case indeed features a degenerate ground state even when $N < \infty$.

The overall picture of SSB described in this section arose from the work of Horsch & von der Linden (1988), Kaplan, Horsch, & von der Linden (1989), Kaplan, von der Linden, & Horsch (1990), and especially Koma & Tasaki (1993, 1994). See also van Wezel (2007, 2008), van Wezel & van den Brink (2007), and Fraser (2016).

The analogy between the quantum Ising chain and the double-well potential may not be surprising physically, since the latter was originally derived from the former: in potassium dihydrogen phosphate, i.e. KH_2PO_4, each proton of the hydrogen bond would reside in one of the two minima of an effective double-well potential originating in the oxygen atoms, if it were not for tunneling, parametrized by the field B, which at small values yields a symmetric ground state (De Gennes, 1963).

§10.6. **Exact solution of the quantum Ising chain: $N < \infty$**

The general set-up to this solution is due to Lieb, Schultz, & Mattis (1961), and was adapted to the quantum Ising by Pfeuty (1970), with further details by Karevski (2006). The complex solution q_0 was already noted by Lieb et al. The energy splitting in higher dimensions does not seem to be known, but Koma & Tasaki (1994, eq. (1.5)) expect similar behaviour as in $d = 1$.

§10.7. **Exact solution of the quantum Ising chain: $N = \infty$**

The solution described in this section is due to Araki & Matsui (1985), where further details may be found; this is a highlight of modern mathematical physics! Theorem 10.20 is due to Araki (1987), although such results have a long history going back to Shale & Stinespring (1964, 1965). For a very clear exposition see Ruijsenaars (1987). See also Evans & Kawahigashi (1998), Chapter 6.

The reason the one-sided chain $\Lambda = \mathbb{N}$ is problematic is that although the bosonic algebra $\otimes_{j \in \mathbb{N}} M_2(\mathbb{C})$ and its fermionic counterpart $CAR(\ell^2(\mathbb{N}))$ are well defined, and are isomorphic through the Jordan–Wigner transformation (10.102) - (10.103), the limiting dynamics has no simple form on either A or F, because the Fourier transform of $\ell^2(\mathbb{N})$ is the Hardy space $H^2(-\pi, \pi)$ of L^2-functions with positive Fourier coefficents, instead of the usual $L^2(-\pi, \pi)$. Unlike on L^2, The energies sgn_k of the fermionic quasiparticles do not define a multiplication operator on H^2.

§10.8. **Spontaneous symmetry breaking in mean-field theories**

The Poisson structure on $S(B)$ was introduced by Bona (1988) and more generally by Duffield & Werner (1992a); see also Bona (2000). Theorem 10.22 and Corollary 10.23 are due to Duffield & Werner (1992a). The symplectic leaves of the given Poisson structure on $S(B)$ (for which notion see e.g. Marsden & Ratiu (1994) or Landsman (1998a)) were determined by Duffield & Werner (1992a): Two states ρ and σ lie in the same symplectic leaf of $\mathscr{S}(B)$ iff $\rho(a) = \sigma(uau^*)$ for some unitary $u \in B$. If ρ and σ are pure, this is the case iff the GNS-representations $\pi_\rho(B)$ and $\pi_\sigma(B)$ are unitarily equivalent, cf. Thm. 10.2.6 in Kadison & Ringrose (1986). In general the implication holds only in one direction: if ρ and σ lie in the same leaf, then they have unitarily equivalent GNS-representations.

Our survey of equilibrium states of homogeneous mean-field models is based on Fannes, Spohn, & Verbeure (1980) and Bona (1989). For rigorous results on the Curie–Weiss model see Chayes et al (2008) and Ioffe & Levit (2013). Numerical evidence for the restoration of Butterfield's Principle may be found in Botet, Julien & Pfeuty (1982) and Botet & Julien (1982), which are up to $N \sim 150$, and Vidal et al (2004), which reaches $N = 1000$. Note that experimental samples have $N < 10$.

In the context of the BCS model of superconductivity in the strong coupling limit), the Hamiltonian, $\hat{h}_\theta$ in (10.287) or h_ω in (10.294) is called the **Bogoliubov–Haag Hamiltonian**, after Bogoliubov (1958) and Haag (1962). Further contributions to mean-field theories include Thirring & Wehrl (1967), Thirring (1968), Hepp (1972), Hepp & Lieb (1973), van Hemmen (1978), Rieckers (1984), Morchio & Strocchi (1987), Duffner & Rieckers (1988), Bona (1988, 1989, 2000), Unnerstall (1990a, 1990b), and Sewell (2002). For a nice proof of Theorem 10.25, which originates in Fannes, Spohn, &Verbeure (1980) and Bona (1989), see Gerisch (1993).

Even in the absence of a global KMS condition for $\hat{\omega}^\beta$, one is justified in interpreting the primary states $(\omega_\theta^\beta)^\infty$ as pure thermodynamic phases of the given infinite quantum system, whose thermodynamics is described by the "phase space" $S(M_n(\mathbb{C}))$. Though somewhat against the spirit of Bohrification (according to which the commutative C*-algebra $C(M_n(\mathbb{C}))$ is the right one to look at), the argument can be strengthened by enlarging A to $A \otimes C(M_n(\mathbb{C}))$ (where the choice of the tensor product does not matter, since $C(M_n(\mathbb{C}))$ is commutative and hence nuclear, see §C.13). This larger C*-algebra was introduced by Bona (1990), who proved:

Theorem 10.30. *1. There is a unique time-evolution α on $A \otimes C(M_n(\mathbb{C}))$ such that for any primary permutation-invariant state ω on A and $a \in A$ one (strongly) has*

$$\lim_{N \to \infty} \pi_\omega \left(e^{ith_{\Lambda_N}} a e^{-ith_{\Lambda_N}} \right) = \pi_\omega(\alpha_t(a)). \tag{10.397}$$

2. The states $\hat{\omega}^\beta$ and ω_θ^β in (10.286), which are defined on A, extend to the tensor product $A \otimes C(M_n(\mathbb{C}))$ as $\hat{\omega}^\beta \otimes \mu_\beta$ and $\omega_\theta^\beta \otimes \delta_\theta$, respectively, and as such satisfy the KMS condition at inverse temperature β with respect to the dynamics α.

§10.9. The Goldstone Theorem

There is a large amount of literature on the Goldstone Theorem, both heuristic and rigorous. The former started with Goldstone, Salam, & Weinberg (1962), whereas the latter originates in Kastler, Robinson, & Swieca (1966); see also Buchholz et al (1992). For a survey, see Strocchi (2008, 2012), whose approach (based on Morchio & Strocchi, 1987) we follow. See also Berzi (1979, 1981), Landau, Perez, & Wreszinski (1981), Fannes, Pule, & Verbeure (1982), and Wreszinski (1987).

§10.10. The Higgs mechanism

The original reference is Higgs (1964ab). Our discussion is based on Lusanna & Valtancoli (1996ab) and Struyve (2011), both of whom derive the physical variables in the abelian Higgs model. See also Rubakov (2002), Strocchi (2008), where Theorem 10.29 may be found, and Stöltzner (2014) for some history and sociology.