Lecture Notes on Quantum Phase Transitions

Draft: June 26, 2012

N.P. Landsman

Institute for Mathematics, Astrophysics, and Particle Physics
Radboud University Nijmegen
Heyendaalseweg 135
6525 AJ NIJMEGEN
THE NETHERLANDS

email: landsman@math.ru.nl
website: http://www.math.ru.nl/∼landsman/
tel.: 024-3652874
office: HG03.740

1 Basic definitions

In this course, both cassical and quantum lattice systems are defined on the standard lattice $\mathbb{Z}^d \subset \mathbb{R}^d$ in spatial dimension d. This infinite lattice contains many finite sublattices $\Lambda \subset \mathbb{Z}^d$, $|\Lambda| < \infty$, like $\Lambda = \Lambda_l = \{x = (x_1, \ldots, x_d) \in \mathbb{Z}^d \mid |x_i| \leq l \, \forall i\}$, where $l \in \mathbb{N}$. An expression like $\lim_{\Lambda \uparrow \mathbb{Z}^d} F(\Lambda)$ then simply means $\lim_{l \to \infty} F(\Lambda_l)$.

The following material may be found in far greater detail and mathematical rigour in [6, 7, 13] (classical theory) and [4, 11, 12] (quantum theory).

1.1 Classical spin-like systems on a lattice

1.1.1 Degrees of freedom and local observables

We write E for the "phase space" *per site* $x \in \mathbb{Z}^d$. This is typically a finite set, like $E = \{-1, 1\}$ for the Ising model (in arbitrary dimension d). Each possible configuration of the "spins" on Λ is given by a function $s : \Lambda \to E$. We write E^Λ for the set of all such functions, or "spin configurations"; this is the "phase space" of the system defined on Λ. It makes perfect sense to have $E^{\mathbb{Z}^d} = \{s : \mathbb{Z}^d \to E\}$ as well. We often write $s_x \equiv s(x)$, or $s_i \equiv s(i)$.

A classical observable *localized in* Λ is a function $f : E^\Lambda \to \mathbb{R}$. We write $A(\Lambda)$ for the set of all such functions. This is a vector space, since we can define $f + g$ by $(f + g)(s) = f(s) + g(s)$ and tf, $t \in \mathbb{R}$, by $(tf)(s) = tf(s)$, and even a commutative algebra, since we can define fg by $(fg)(s) = f(s)g(s)$. This makes sense for $\Lambda = \mathbb{Z}^d$, so that $A(\mathbb{Z}^d)$ is the set of all functions $f : E^{\mathbb{Z}^d} \to \mathbb{R}$ (N.B. $A(\mathbb{Z}^d) \neq A$ below!)

The following construction is extremely important:[1]

$$\text{if } \Lambda \subset \Lambda', \text{ there is a natural embedding } A(\Lambda) \hookrightarrow A(\Lambda').$$

Indeed, if we write $f' \in A(\Lambda')$ for the image of $f \in A(\Lambda)$, we put $f'(s') = f(s)$, where $s = s'_{|\Lambda}$ is the restriction of $s' : \Lambda' \to E$ to Λ (so that $s : \Lambda \to E$ as required). Seen as an element f' of the larger $A(\Lambda')$, elements f of $A(\Lambda)$ are characterized by the property that $f'(s') = f'(s'')$ for all $s', s'' \in E^{\Lambda'}$ that coincide on Λ. Thus observables in $A(\Lambda)$ are only sensitive to spin configurations inside Λ.

We now define the so-called *algebra of local observables*

$$A = \cup_{\Lambda \subset \mathbb{Z}^d, |\Lambda| < \infty} A(\Lambda) \tag{1.1}$$

as the set of all functions $f : E^{\mathbb{Z}^d} \to \mathbb{R}$ that lie in some $A(\Lambda)$, $|\Lambda| < \infty$, where we regard $A(\Lambda)$ as a subset of $A(\mathbb{Z}^d)$ under the above embedding $A(\Lambda) \hookrightarrow A(\mathbb{Z}^d)$.

[1] Mathematically, what happens here is this: a (continuous) map $\varphi : X \to Y$ induces a pullback $\varphi^* : C(Y, Z) \to C(X, Z)$, where $C(X, Z)$ denotes the set of (continuous) functions from X to Z (whatever it is), defined as $\varphi^* f = f \circ \varphi$, or $(\varphi^* f)(x) = f(\varphi(x))$. We apply this idea twice:

1. $X = \Lambda$, $Y = \Lambda'$ where $\Lambda \subset \Lambda'$, $\varphi : \Lambda \to \Lambda'$ is inclusion, and $Z = E$. In that case, $\varphi^* : E^{\Lambda'} \to E^\Lambda$ is just restriction to Λ, that is, $\varphi^* s = s_{|\Lambda}$. We write $\varphi^* = r$.

2. $X = E^{\Lambda'}$, $Y = E^\Lambda$, and $\varphi : E^{\Lambda'} \to E^\Lambda$ is the restriction map r of the previous point. Taking $Z = \mathbb{R}$, the pullback $r^* : A(\Lambda) \to A(\Lambda')$ is just the map $A(\Lambda) \hookrightarrow A(\Lambda')$ in the main text!

In other words, we have $A \subset A(\mathbb{Z}^d)$ as the subset of local observables, and $f \in A(\mathbb{Z}^d)$ lies in A iff there exists a *finite* sublattice $\Lambda \subset \mathbb{Z}^d$ such that $f(s) = f(s')$ for all spin configurations $s, s' : \mathbb{Z}^d \to E$ that coincide on Λ, i.e., for which $s_{|\Lambda} = s'_{|\Lambda}$. Of course, in that case f also lies in all $A(\Lambda')$, for $\Lambda \subset \Lambda' \subset \mathbb{Z}^d$ and $|\Lambda'| < \infty$.[2]

As an interesting example, take $d = 1$, replace $\mathbb{Z}$ by $\mathbb{N}$ for convenience, and take $E = \underline{2} = \{0, 1\}$. Then $\underline{2}^{\mathbb{N}}$ is the set of all binary sequences, and the "local observables" among the functions $f : \underline{2}^{\mathbb{N}} \to \mathbb{R}$ are those functions that depend on finitely many bits only (that is, $f \in A$ iff there exists a finite subset $S \subset \mathbb{N}$ such that $f(s) = f(s')$ whenever $s_i = s'_i$ for all $i \in S$).

1.1.2 Hamiltonian

Hamiltonians are typically well defined only for finite sublattice $\Lambda \subset \mathbb{Z}^d$; for example, for the Ising model we have

$$h_\Lambda(s) = -J \sum_{\langle ij \rangle_\Lambda} s_i s_j - B \sum_{i \in \Lambda} s_i, \tag{1.2}$$

where $J > 0$, $B \geq 0$, and the sum over $\langle ij \rangle_\Lambda$ denotes summing over nearest neighbours in Λ. Clearly, replacing Λ by $\mathbb{Z}^d$ would make $h_{\mathbb{Z}^d}(s)$ infinite for most s. Hence we would like to define some local Hamiltonian $h_\Lambda \in A(\Lambda)$ for each finite Λ. To do so uniformly in Λ, we first define an *interaction* Φ as an assignment $X \mapsto \Phi(X)$, where $X \subset \mathbb{Z}^d$ is finite and $\Phi(X) \in A(X)$. If $X \subset Y$ and we wish to regard $\Phi(X)$ an an element in $A(Y)$ through the inclusion $A(X) \subset A(Y)$, we sometimes indicate this explicitly by writing $\Phi(X)_Y \in A(Y)$. We then define $h_\Lambda \in A(\Lambda)$ by

$$h_\Lambda = \sum_{X \subset \Lambda} \Phi(X)_\Lambda, \tag{1.3}$$

where the the sum is over all subsets X of Λ. This looks like a large sum, but in practice only a few subsets X contribute. For example, to reproduce the Ising Hamiltonian (1.2), we put $\Phi(X) = 0$ whenever X has more than two elements or when X is a pair of "not nearest neighbours"; the only nonvanishing terms are $\Phi(\{i\}) : s \mapsto -Bs_i$, and $\Phi(\{i, j\}) : s \mapsto -Js_i s_j$ if i and j are nearest neighbours.

The prescription (1.3) only involves spins inside Λ; in the literature, this is called a Hamiltonian with *free boundary conditions*. Another, perhaps physically more

[2] If we identify $E^{\mathbb{Z}^d}$ with the infinite Cartesian product $\prod_{x \in \mathbb{Z}^d} E$ and equip the latter with the product topology (in which it is compact for finite E), then $A \subset C(E^{\mathbb{Z}^d}, \mathbb{R})$, and A is even dense w.r.t. the supremum-norm $\|f\|_\infty = \sup_{s \in E^{\mathbb{Z}^d}} |f(s)|$ on $C(E^{\mathbb{Z}^d}, \mathbb{R})$. The fact that this norm is finite for each $f \in A$ follows because f is continuous and $E^{\mathbb{Z}^d}$ is compact, but it may also be seen directly: provided $f \in A(\Lambda)$, the number of spin configurations s for which $f(s)$ can vary is finite (viz. $|E|^{|\Lambda|}$), since f only depends on the spins inside Λ. So the supremum is not really taken over an infinite numer of $s \in E^{\mathbb{Z}^d}$, but only over a finite number $s \in E^\Lambda$. This suggests taking the closure of A in the sup-norm, an operation which slightly enlarges A and yields the *quasilocal observables* A_{ql}. Mathematics freaks will be interested to know that the complexification of the latter, namely $C(E^{\mathbb{Z}^d}, \mathbb{C})$, is a commutative C*-algebra.

realistic possibility is to fix a spin configuration $b \in E^{\mathbb{Z}^d}$, and define $h^b_\Lambda \in A(\Lambda)$ by

$$h^b_\Lambda = \sum_{X \subset \mathbb{Z}^d, |X|<\infty, X\cap\Lambda\neq\emptyset} \Phi(X)^b_\Lambda. \tag{1.4}$$

This involves some new notation $\Phi(X)^b_\Lambda$, which means the following. In principle, $\Phi(X) \in A(X)$ is a function on E^X. We now turn $\Phi(X)$ into a function $\Phi(X)^b_\Lambda$ on E^Λ (so that h^b_Λ is a function on E^Λ as required): for given $s : \Lambda \to E$ and given $b : \mathbb{Z}^d \to E$ we define $s' : X \to E$ by putting $s' = s$ on $X \cap \Lambda$ and $s' = b$ on the remainder of X (which is $X \cap \Lambda^c$, with $\Lambda^c = \mathbb{Z}^d \backslash \Lambda$). Then

$$\Phi(X)^b_\Lambda(s) = \Phi(X)(s'). \tag{1.5}$$

Physically, this simply means that those spins outside Λ that interact with spins inside Λ are set at a fixed value determined by the boundary condition b. For example, consider the Ising model in $d = 1$. If we take $\Lambda = \{2, 3\}$, then from (1.3) we obtain $h_\Lambda = -Js_2s_3 - B(s_2 + s_3)$; spins outside Λ do not contribute. From (1.4), on the other hand, we obtain $h^b_\Lambda = h_\Lambda - J(b_1s_2 + s_3b_4)$. Although the boundary condition b is arbitrary, one may think of simple choices like $b_i = 1$ or -1 for each i.

For later use (and greater insight), we rewrite (1.4) as a difference between Hamiltonians with free boundary conditions. To do so, for given finite Λ we pick some finite $\Lambda' \supset \Lambda$ large enough that it contains all spins outside Λ that interact with spins inside Λ (provided this is possible). Writing $h_\Lambda(s|b) \equiv h^b_\Lambda(s)$, this yields

$$\begin{aligned} h_\Lambda(s|b) &= h_{\Lambda'}(s, b) - h_{\Lambda'\backslash\Lambda}(b) && (1.6)\\ &= \sum_{X'\subset\Lambda'} \Phi(X')_{\Lambda'}(s, b) - \sum_{Y\subset\Lambda'\backslash\Lambda} \Phi(Y)_{\Lambda'\backslash\Lambda}(b). && (1.7)\end{aligned}$$

Analogous to (1.5), the notation $\Phi(X')_{\Lambda'}(s, b)$ here means $\Phi(X')_{\Lambda'}(s')$, for the function $s' : \Lambda' \to E$ that on $\Lambda \subset \Lambda'$ coincides with $s : \Lambda \to E$, whilst on $(\Lambda'\backslash\Lambda) \subset \Lambda'$ it coincides with the restriction of b to $\Lambda'\backslash\Lambda$. Thus we may also write

$$h_\Lambda(s|b) = \lim_{\Lambda'\uparrow\mathbb{Z}^d} (h_{\Lambda'}(s, b) - h_{\Lambda'\backslash\Lambda}(b)), \tag{1.8}$$

realizing that neither $h_{\mathbb{Z}^d}(s, b)$ nor $h_{\mathbb{Z}^d\backslash\Lambda}(b)$ makes sense.

Exercise 1.1 *Write down $h_\Lambda(s|b)$ for the Ising model in arbitrary dimension.*

Exercise 1.2 *Define* periodic *boundary conditions for local Hamiltonians defined by arbitrary interactions Φ and special sublattices of the form $\Lambda = \Lambda_l$.*

For example, the Ising model in $d = 1$ would have local Hamiltonians with periodic boundary conditions of the type

$$h^{pbc}_{\{1,2,\ldots,n\}}(s) = J(s_1s_n + \sum_{i=1}^{n-1} s_is_{i+1}) - B\sum_{i=1}^{n} s_i. \tag{1.9}$$

1.2 Quantum spin-like systems on a lattice

1.2.1 Hilbert spaces defined by lattices

The quantum analogue of the (finite) classical phase space E per site is a (finite-dimensional) Hilbert space H; e.g., for the Ising model, we simply have $H = \mathbb{C}^2$.

For finite Λ, the quantum analogue of the space E^Λ is the *tensor product*

$$H(\Lambda) = \otimes_{x\in\Lambda} H_x, \tag{1.10}$$

where $H_x = H$ for each x. We will define the tensor product assuming that $\dim(H) < \infty$, starting with the "classical" function space $H^\Lambda = \{\psi : \Lambda \to H\}$. Each $\psi \in H^\Lambda$ defines a map $\hat{\psi} : H^\Lambda \to \mathbb{C}$ by[3]

$$\hat{\psi}(\varphi) = \prod_{x\in\Lambda} \langle \varphi(x), \psi(x)\rangle_H. \tag{1.11}$$

Such maps form a complex vector space, since we may add maps $\hat{\psi}_1$ and $\hat{\psi}_2$ by putting $(\hat{\psi}_1 + \hat{\psi}_2)(\varphi) = \hat{\psi}_1(\varphi) + \hat{\psi}_2(\varphi)$, and for $z \in \mathbb{C}$ we define $z\hat{\psi}$ by $(z\hat{\psi})(\varphi) = z\hat{\psi}(\varphi)$. This vector space is called $H(\Lambda)$. To turn it into a Hilbert space, we first define an inner product on the 'basic' maps by

$$\langle \hat{\psi}_1, \hat{\psi}_2\rangle_{H(\Lambda)} = \prod_{x\in\Lambda} \langle \psi_1(x), \psi_2(x)\rangle_H, \tag{1.12}$$

and subsequently extend this to all elements of $H(\Lambda)$ by (sesqui)linearity.

It is convenient to write $\otimes_{x\in\Lambda}\psi(x)$ for $\hat{\psi}$, so that the elements of $H(\Lambda)$ are linear combinations of the former expressions. Indeed, we obtain an orthonormal basis of $H(\Lambda)$ by letting $\psi(x)$ vary over an arbitrary orthonormal basis of H, for each $x \in \Lambda$. If $H = \mathbb{C}^n$, this yields $n^{|\Lambda|}$ basis vectors, so that, recalling the fact that the dimension of a Hilbert space equals the cardinality of some orthonormal basis,

$$\dim(H(\Lambda)) = \dim(H)^{|\Lambda|}. \tag{1.13}$$

The following exercise is very important for the physical interpretation of $H(\Lambda)$. In preparation: for any countable set S, define a Hilbert space $\ell^2(S)$ as the space of functions $f : S \to \mathbb{C}$ that satisfy $\sum_{s\in S} |f(s)|^2 < \infty$,[4] with inner product

$$\langle f, g\rangle = \sum_{s\in S} \overline{f(s)} g(s). \tag{1.14}$$

Exercise 1.3 *Suppose $|E| = n$, so that we may assume $E = \{1, 2, \ldots, n\} \equiv \underline{n}$, and suppose $H = \mathbb{C}^n$. Show that $H(\Lambda)$ as in (1.10) is unitarily equivalent to $\ell^2(E^\Lambda)$.*

Under this equivalence, elements of $H(\Lambda)$ may be interpreted as "wavefunctions" whose argument is a classical spin configuration $s \in E^\Lambda$ (that is, $s : \Lambda \to E$).

[3] Our inner product is linear in the *second* variable, and $\langle \varphi, \psi\rangle \equiv \langle \varphi|\psi\rangle$. Also, $\langle \varphi, a\psi\rangle \equiv \langle \varphi|a|\psi\rangle$.

[4] This convergence condition is irrelevant if $S = E^\Lambda$ with $|\Lambda| < \infty$, in which case S is finite.

1.2.2 Local quantum observables

We define the algebra of quantum observable localized in Λ as

$$A(\Lambda) = B(H(\Lambda)), \tag{1.15}$$

where $B(K)$ stands for the algebra of all bounded operators on some Hilbert space K; in the case at hand, $H(\Lambda)$ is finite-dimensional, so that any operator (= linear map) is bounded. As in the classical case, whenever $\Lambda \subset \Lambda'$, there is a natural embedding $A(\Lambda) \hookrightarrow A(\Lambda')$, given by "adding unit matrices". More precisely, $B(H(\Lambda))$ may be constructed just like $H(\Lambda)$ itself, i.e., by starting with $B(H)^\Lambda$. Any $a \in B(H)^\Lambda$, that is, any map $a : \Lambda \to B(H)$, defines an operator $\hat{a}$ on $H(\Lambda)$ by first defining its action on elementary tensors by

$$\hat{a}\hat{\psi} = \otimes_{x\in\Lambda} a(x)\psi(x), \tag{1.16}$$

and extending linearly to arbitrary vectors in $H(\Lambda)$. We may write $\hat{a} = \otimes_{x\in\Lambda} a(x)$ and reconstruct $B(H(\Lambda))$ as the complex vector space spanned by all such elementary operators. Our injection $B(H(\Lambda)) \hookrightarrow B(H(\Lambda'))$, then, is given by linear extension of $\hat{a} \mapsto \hat{a}'$, where $\hat{a}'(x') = a(x)$ whenever $x' = x \in \Lambda \subset \Lambda'$, and $\hat{a}'(x') = 1$ otherwise. In other words, we expand $\otimes_{x\in\Lambda} a(x)$ in $A(\Lambda)$ to $\otimes_{x'\in\Lambda'} a'(x')$ in $A(\Lambda')$ by adding unit matrices at all $x' \in \Lambda'\backslash\Lambda$. The classical definition (1.1) of the algebra A of local observables may then be repeated literally, *mutatis mutandis*.[5]

In the classical case, we say that an observable $f \in A(\Lambda)$ is positive if $f(s) \geq 0$ for each $s \in E^\Lambda$. Since A is the union of all the $A(\Lambda)$, this also defines positivity of classical observables in A. Similarly, we have a notion of positivity in the quantum algebra of observables $A(\Lambda)$, saying that $a \geq 0$ iff $\langle\psi, a\psi\rangle \geq 0$ for all $\psi \in H(\Lambda)$. This notion propagates into A, too. Also, in both the classical and the quantum cases each $A(\Lambda)$ has a unit: in the classical case this is the function $1 : s \mapsto 1$ for all s, which survives all inclusions so as to become the unit of A. In the quantum case, the operator $\otimes_{x\in\Lambda} a(x)$ with $a(x) = 1$ for each x is the unit of each $A(\Lambda)$, persisting to A. The key *difference* between classical and quantum theory, of course, is that in the latter case the algebras $A(\Lambda)$, and hence also A, are *noncommutative*.

The definition of interactions and local quantum Hamiltonians is exactly the same as in the classical case, now using the quantum meaning (1.15) of $A(\Lambda)$. For example, the Hamiltonian of the quantum Ising model (in transverse magnetic field)

$$\hat{h}_\Lambda = -J \sum_{\langle ij\rangle_\Lambda} \sigma_i^z \sigma_j^z - B \sum_{i\in\Lambda} \sigma_i^x, \tag{1.17}$$

really means the following: a single term like σ_i^x stands for the operator $\otimes_{k\in\Lambda} a(k)$ in $H(\Lambda)$ which has $a(i) = \sigma^x$ (i.e., the first Pauli matrix) and $a(k) = 1$ for all $k \neq i$. Similarly, $\sigma_i^z \sigma_j^z$ denotes the operator $\otimes_{k\in\Lambda} a(k)$ in $H(\Lambda)$ which has $a(i) = \sigma^z$, $a(j) = \sigma^z$, and $a(k) = 1$ for all $j \neq k \neq i$. As we have seen, such elementary operators may be freely added to obtain further operators in $B(H(\Lambda))$, and the local Hamiltonian $\hat{h}_\Lambda$ is a shining example of this.

[5] In the quantum case, one may also define a norm on A by using the operator norm on each $A(\Lambda) \subset A$. Unlike each $A(\Lambda)$, the ensuing A is not complete in this norm, and, as in the classical case, it may be completed into the C*-algebra of quasi-local observables.

1.3 States

We start with the 'usual' definitions for finite systems, and later generalize these to infinite systems, using the above formalism. This generalization is strictly necessary for the study of phase transitions, since these cannot even occur in finite systems.

1.3.1 Ground states of finite systems

A ground state of a classical system of the type studied above, restricted to a finite lattice $\Lambda \subset \mathbb{Z}^d$, is simply a spin configuration $s_0 \in E^\Lambda$ that minimizes the local Hamiltonian h_Λ, cf. (1.3), or its counterpart (1.4). That is, we must have

$$h_\Lambda(s_0) \leq h_\Lambda(s) \tag{1.18}$$

for all $s \in E^\Lambda$. For example, the Ising model (1.2) has a unique ground state for $B > 0$, namely $s_0(x) = 1$ for all $x \in \Lambda$, whereas it has two ground states $s_0^\pm$ for $B = 0$, given by $s_0^\pm(x) = \pm 1$ for all x.

Similarly, a ground state of a quantum spin-like system on a finite lattice Λ is given by a unit vector $\psi_0 \in H(\Lambda)$ that minimizes the quantum Hamiltonian $\hat{h}_\Lambda$, i.e.,

$$\langle \psi_0, \hat{h}_\Lambda \psi_0 \rangle \leq \langle \psi, \hat{h}_\Lambda \psi \rangle \tag{1.19}$$

for all unit vectors $\psi \in H(\Lambda)$. Equivalently (at least for finite-dimensional $H(\Lambda)$), ψ_0 is an eigenstate of $\hat{h}_\Lambda$ with the lowest eigenvalue.

We will see later on that the quantum Ising model has a unique ground state for $0 < B < B_c$, but for $B = 0$ the model is essentially classical (since all operators in the Hamiltonian commute) and hence it has two degenerate ground states.

Exercise 1.4 *Write down the ground states of the quantum Ising model for $B = 0$: both as vectors in $H(\Lambda)$ and as vectors in $\ell^2(E^\Lambda)$; cf. Exercise 1.3.*

1.3.2 Mixed states

For the purposes of statistical physics the notion of a state has to be revised. According to Ludwig Boltzmann (or, mathematically speaking, David Ruelle), a state of a classical system (in the above sense) localized on Λ is a *probability distribution* on E^Λ, i.e., a function $p : E^\Lambda \to [0, 1]$ (or, given (1.20), $p \geq 0$ pointwise) such that

$$\sum_{s \in E^\Lambda} p(s) = 1. \tag{1.20}$$

Let us note that a point s_0 of E^Λ yields a probability distribution $p_{s_0} = \delta_{s_0}$ on E^Λ, defined by $\delta_{s_0}(s) = 0$ if $s \neq s_0$, and $\delta_{s_0}(s_0) = 1$. Writing $\mathcal{P}(E^\Lambda)$ for the set of all probability distributions on E^Λ, we therefore have an embedding

$$E^\Lambda \hookrightarrow \mathcal{P}(E^\Lambda), \;\; s_0 \mapsto \delta_{s_0}. \tag{1.21}$$

States of the form δ_{s_0} are called *pure*, all other states being *mixed*.

Similarly, according to Lev Landau (or, mathematically speaking, John von Neumann), a state of a quantum system (in the above sense) localized on Λ is a *density matrix* on $H(\Lambda)$, that is, an operator $\hat{\rho} \in B(H(\Lambda))$ satisfying $\hat{\rho} \geq 0$ and

$$\operatorname{Tr} \hat{\rho} = 1. \tag{1.22}$$

Since $\hat{\rho} \geq 0$ implies $\hat{\rho}^* = \hat{\rho}$, we may equivalently define a density matrix as a hermitian matrix with non-negative eigenvalues summing up to 1.

Once again, the original notion of a state as a unit vector in $H(\Lambda)$ is actually a special case of the above notion, at least, if we realize that ψ and $z\psi$ define the same state for any $z \in \mathbb{C}$ with $|z| = 1$ (that is, states are defined only "up to a phase"). Namely, we may pass from a unit vector ψ to a density matrix

$$\hat{\rho} = |\psi\rangle\langle\psi|, \tag{1.23}$$

where the general expression of the form $|\psi\rangle\langle\varphi|$, for vectors ψ and φ in some Hilbert space K, denotes the operator on K that maps a vector χ to $\langle\psi, \chi\rangle\psi$ (here physicists would probably want to write $\psi\rangle$ for ψ, etc., so that, quite neatly if not tautologically, $|\psi\rangle\langle\varphi|$ maps $|\chi\rangle$ to $|\psi\rangle\langle\varphi|\chi\rangle$). The expression (1.23) is just the orthogonal projection onto the (one-dimensional) linear span of ψ, and hence density operators $\hat{\rho}$ of this type are characterized by the equation $\hat{\rho}^2 = \hat{\rho}$ (abstractly, a projection on a Hilbert space K is an operator p satisfying $p^2 = p^* = p$, and the dimension of its image is $\dim(pK) = \operatorname{Tr} p$, so that a density matrix that is simultaneously a projection must have one-dimensional range).

For reasons to become clear later, we denote the set of all density operators on $H(\Lambda)$ by $\mathcal{S}(H(\Lambda))$. We also write $\mathbb{P}H(\Lambda)$ for the set of rays in $H(\Lambda)$, i.e., the set of unit vectors up to a phase. The construction (1.23) then yields an injection

$$\mathbb{P}H(\Lambda) \hookrightarrow \mathcal{S}(H(\Lambda)), \tag{1.24}$$

which is the quantum counterpart of (1.21). Once again, states of the form (1.23) are called *pure*, all other states being *mixed*. Here is a nice illustration.

Exercise 1.5 *1. Show that any density matrix on $\mathbb{C}^2$ can be parametrized as*

$$\hat{\rho} = \hat{\rho}(x, y, z) = \tfrac{1}{2} \begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix}, \tag{1.25}$$

where $(x, y, z) \in \mathbb{R}^3$ satisfy $x^2 + y^2 + z^2 \leq 1$ (these form the three-ball B^3).

2. Show that the pure states, i.e., the density matrices of the form (1.23)*, exactly correspond to the case $x^2 + y^2 + z^2 = 1$.*

This example falls into place if we use some convexity theory (due to Hermann Minkowski). A *convex* set C in a (real or complex) vector space V is a set for which the line segment between any two points in C entirely lies in C. In other words, C is convex if the following condition holds: if $\rho \in C$ and $\sigma \in C$, then $t\rho + (1-t)\sigma \in C$ for all $t \in [0, 1]$. Examples: $C = [0, 1]$ in $V = \mathbb{R}$, $C = B^3$ in $V = \mathbb{R}^3$.

Exercise 1.6 *1. Show that the classical state space $\mathcal{P}(E^\Lambda)$ is convex.*

2. Show that the quantum state space $\mathcal{S}(H(\Lambda))$ is convex.

The *extreme boundary* $\partial_e C$ of a convex set C is the set of all *extreme points* $\omega \in C$, defined by the following condition: if $\omega = t\rho + (1-t)\sigma$ for some $0 < t < 1$ and $\rho, \sigma \in C$, then $\rho = \sigma = \omega$. In other words, an extreme point cannot lie on a line segment inside C (except as an endpoint). As first proposed by von Neumann, in physics the pure states are precisely the extreme points of the state space. This idea is justified by seeing pure states as states about which we have maximal information; mixed states, on the other hand, are obtained by combining pure states with weights corresponding to (subjective) probabilities. Indeed, iterating the definition of a convex set it follows that if points ω_i lie in C $(i = 1, \ldots, n)$ and probabilities $p_i \in [0, 1]$ sum to 1 (as they should), then $\sum_i p_i \omega_i$ lies in C. Conversely, one may ask if all points in some convex set may be written as weighted sums of pure states. This turns out to be the case if C is compact, and if we allow suitably convergent infinite sums in the mixing operation. Clearly, under these conditions the pure state space cannot be empty.[6]

Exercise 1.7 *1. Let $M_2(\mathbb{C}) \to \mathbb{C}$ be the set of 2×2 complex matrices. Each density matrix $\hat{\rho}$ on $\mathbb{C}^2$ defines a map $\omega : M_2(\mathbb{C}) \to \mathbb{C}$ by*

$$\omega(a) = \mathrm{Tr}\,(\hat{\rho} a). \tag{1.26}$$

(a) Show that ω is linear (trivial), that $\omega(1) = 1$ (almost trivial), and that $\omega(a) \geq 0$ for all $a \geq 0$ (easy).

(b) Conversely, show that any linear map $\omega : M_2(\mathbb{C}) \to \mathbb{C}$ that satisfies the latter two conditions is necessarily of the form (1.26), where $\hat{\rho}$ is some density matrix on $\mathbb{C}^2$.

This exercise leads to a unified picture of states of classical and quantum systems. We start with the algebra of observables A, or any local subalgebra $A(\Lambda)$ thereof, and define a state as a linear map $\omega : A \to \mathbb{R}$ (classically) or $\omega : A \to \mathbb{C}$ (quantummy) that satisfies the two conditions in the previous exercise, i.e. $\omega(1) = 1$ and $\omega(a) \geq 0$ for all $a \geq 0$. It immediately follows that the set of all states thus defined is convex.[7]

The physical interpretation of a state is simply that $\omega(a)$ is the expectation value of an observable $a \in A$ in the state ω; in other words, a state is now regarded as a rule that tells any observable what its expectation value is. In classical physics,[8] the Riesz–Markov Theorem of measure theory shows that this notion of a state is

[6] This is the Krein–Milman Theorem from functional analysis. A simple example of a convex set with empty extreme boundary is the three-ball without its boundary, which indeed is non-compact!

[7] It is also compact as a subset of the dual space A^* of continuous linear functionals on A, but only if A^* is equipped with the so-called weak-star topology.

[8] Here it is crucial that classical observables in A are *localized*, which implies that they are *continuous* functions on $E^{\mathbb{Z}^d}$. Otherwise, the Riesz–Markov Theorem does not apply.

equivalent to the one we had, i.e., any state ω is given by a probability distribution p, according to

$$\omega(f) = \sum_{s \in E^{\mathbb{Z}^d}} p(s) f(s) \equiv \langle f \rangle_p. \tag{1.27}$$

In quantum physics, a generalization of Exercise (b) above shows that for finite-dimensional Hilbert spaces the new notion of a state just recovers the old notion of a density matrix, in that any state is given by (1.26). The true value of our new definition of a state emerges for infinite quantum systems: since (unlike its subalgebras $A(\Lambda)$) the algebra A of local observables no longer acts on any Hilbert space, the whole notion of a density matrix becomes obscure.

1.3.3 Equilibrium states of finite systems

Arguably the most interesting states in physics are *equilibrium states*, defined with respect to some temperature T (whose deeper mening shall remain obscure in these notes). We first defines such states locally, i.e., in a finite lattice $\Lambda \subset \mathbb{Z}^d$.

Classically, given an interaction Φ and the ensuing family of local Hamiltonians h_Λ, we define the local *energy* for each Λ as a function $\mathcal{E}_\Lambda : \mathcal{P}(E^\Lambda) \to \mathbb{R}$ of the classical states on E^Λ, i.e., of the probability distributions on E^Λ, by

$$\mathcal{E}_\Lambda(p) = \sum_{s \in E^\Lambda} p(s) h_\Lambda(s). \tag{1.28}$$

Of course, this is just the expectation value of the Hamiltonian in the state p, cf. (1.27). The local *entropy* $S_\Lambda : \mathcal{P}(E^\Lambda) \to \mathbb{R}$ is a more subtle concept; rather than the expectation value of some (local) observable, it yields a property of the probability distribution itself. With Boltzmann's constant k_B, we have

$$S_\Lambda(p) = -k_B \sum_{s \in E^\Lambda} p(s) \ln(p(s)). \tag{1.29}$$

Note that $S_\Lambda(p) \geq 0$, with equality iff p is a pure state.

Finally, the local *free energy* $\mathcal{F}^\beta_\Lambda : \mathcal{P}(E^\Lambda) \to \mathbb{R}$ is then defined as

$$\mathcal{F}^\beta_\Lambda = \mathcal{E}_\Lambda - T S_\Lambda, \tag{1.30}$$

where $\beta = 1/k_B T$. A *local equilibrium state*, then, is a probability distribution p^β_Λ that minimizes the free energy (for fixed T). Boltzmann's solution is given by

$$p^\beta_\Lambda(s) = (Z^\beta_\Lambda)^{-1} e^{-\beta h_\Lambda(s)}; \tag{1.31}$$

$$Z^\beta_\Lambda = \sum_{s' \in E^\Lambda} e^{-\beta h_\Lambda(s')}. \tag{1.32}$$

Exercise 1.8 *Show that $\mathcal{F}^\beta_\Lambda(p) \geq -\beta^{-1} \ln Z^\beta_\Lambda$ for all p, with equality iff $p = p^\beta_\Lambda$.*

Here "for all p" of course means for all $p \in \mathcal{P}(E^\Lambda)$. It follows that there is a unique local equilibrium state for each T, with ensuing *free energy in equilibrium*

$$F^\beta_\Lambda = \mathcal{F}^\beta_\Lambda(p^\beta_\Lambda) = -\beta^{-1} \ln Z^\beta_\Lambda. \tag{1.33}$$

Note that none of the above expressions makes sense for $\Lambda = \mathbb{Z}^d$, but one might hope that the corresponding intensive quantities (like $f^\beta_\Lambda = F^\beta_\Lambda / |\Lambda|$) have a limit.

The corresponding quantum-mechanical expressions are the same, *mutatis mutandis*. In particular, the energy $\hat{\mathcal{E}}_\Lambda$, entropy $\hat{S}_\Lambda$, and free energy $\hat{\mathcal{F}}^\beta_\Lambda$ are now functions on the space $\mathcal{S}(H(\Lambda))$ of the density matrices on $H(\Lambda)$. We have

$$\hat{\mathcal{E}}_\Lambda(\hat{\rho}) = \mathrm{Tr}\,(\hat{\rho}\hat{h}_\Lambda); \tag{1.34}$$

$$\hat{S}_\Lambda(\hat{\rho}) = -k_B\mathrm{Tr}\,(\hat{\rho}\ln\hat{\rho}); \tag{1.35}$$

$$\hat{\mathcal{F}}^\beta_\Lambda = \hat{\mathcal{E}}_\Lambda - T\hat{S}_\Lambda, \tag{1.36}$$

$$\hat{Z}^\beta_\Lambda = \mathrm{Tr}\,e^{-\beta\hat{h}_\Lambda}. \tag{1.37}$$

If we define a local equilibrium state as a density matrix $\hat{\rho}^\beta_\Lambda$ that minimizes the free energy (for fixed T), the unique solution is given by the density matrix

$$\hat{\rho}^\beta_\Lambda = (\hat{Z}^\beta_\Lambda)^{-1}e^{-\beta\hat{h}_\Lambda}. \tag{1.38}$$

Exercise 1.9 *Show that* $\hat{\mathcal{F}}^\beta_\Lambda(\hat{\rho}) \geq -\beta^{-1}\ln\hat{Z}^\beta_\Lambda$ *for all* $\hat{\rho}$*, with equality iff* $\hat{\rho} = \hat{\rho}^\beta_\Lambda$.

What remains to be done, however, is to define ground states and equilibrium states for infinite systems.

1.3.4 Ground states of infinite classical systems

The classical case is easy: with local Hamiltonians h_Λ (or h^b_Λ, in case of a fixed boundary condition b) defined by a single interaction Φ according to (1.3) (or (1.4)), a *ground state* for Φ is simply a point $s_0 \in E^{\mathbb{Z}^d}$, i.e., a function $s_0 : \mathbb{Z}^d \to E$, whose restriction $(s_0)_{|\Lambda}$ to Λ minimizes h_Λ (or h^b_Λ), for each finite $\Lambda \subset \mathbb{Z}^d$. In the Ising model in any d with $B = 0$ and free boundary condition, this gives the usual two ground states (in which all spins are either "up" or "down").

Some authors (e.g., [6], however, use a slightly different notion for Hamiltonians (1.3) determined by free boundary conditions: they say that $s_0 \in E^{\mathbb{Z}^d}$ is a ground state for a given interaction Φ if, writing $h^{s_0}_\Lambda$ for (1.4) with $b = s_0$, the condition

$$h^{s_0}_\Lambda(s_0) \leq h^{s_0}_\Lambda(s) \tag{1.39}$$

holds for all finite $\Lambda \subset \mathbb{Z}^d$ and all $s \in E^{\mathbb{Z}^d}$ *that coincide with* s_0 *outside* Λ. In other words, s_0 itself acts as its own boundary condition b and this boundary condition is fixed for all s that compete with s_0 in minimizing the local Hamiltonian $h^{b=s_0}_\Lambda$. This definition opens the possibility of *domain walls*. For example, in the Ising model in $d = 1$ with $B = 0$, this definition admits ground states in which infinite chains of "spin up" alternate with infinite chains of "spin down", and similarly in higher d.

Ground states may not exist and if they do, they may not be unique. Let us, therefore, briefly look at the *set of ground states* (for some fixed interaction). If this set has at least two elements, say $s^{(1)}_0$ and $s^{(2)}_0$, then for any $t \in (0,1)$ we may form the mixed state $p_0 = ts^{(1)}_0 + (1-t)s^{(2)}_0$, reinterpreted as a probability distribution on $E^{\mathbb{Z}^d}$ assigning probability t to $s_0 = s^{(1)}_0$, probability $(1-t)$ to $s = s^{(2)}_0$, and probability zero to all other points of $E^{\mathbb{Z}^d}$. Restricting ourselves to free boundary conditions for simplicity, this state satisfies

$$\langle h_\Lambda\rangle_{p_0} \leq \langle h_\Lambda\rangle_p \tag{1.40}$$

for all probability distributions p on $E^{\mathbb{Z}^d}$. Hence we may relax the definition of a ground state so as to admit mixed states, i.e., probability distributions p on $E^{\mathbb{Z}^d}$, and say that $p_0 \in \mathcal{P}(E^{\mathbb{Z}^d})$ is a ground state if (1.40) holds for any $p \in \mathcal{P}(E^{\mathbb{Z}^d})$. It follows that the set $\mathcal{G}(\Phi)$ of ground states of a given interaction Φ is a convex set, whose extreme points are the "pure ground states". The above discussion suggests that these are pure states according to our original definition, that is, we would like to identify $\partial_e \mathcal{G}(\Phi)$ with $\mathcal{G}(\Phi) \cap \partial_e \mathcal{P}(E^{\mathbb{Z}^d}) = \mathcal{G}(\Phi) \cap E^{\mathbb{Z}^d}$. Under suitable hypotheses on Φ this is correct, and we may unambiguously talk about "pure ground states".

1.3.5 Ground states of infinite quantum systems

The above definition of a classical ground state suggests that also in the quantum case we may define a ground state of an infinite system as a state ω whose localization ω_Λ to any finite volume $|\Lambda| < \infty$ (i.e., ω_Λ is the restriction of $\omega : A \to \mathbb{C}$ to $A(\Lambda) \subset A$) is a ground state for $\hat{h}_\Lambda$. Surprisingly, such a naive definition would be inappropriate because of the superposition principle.

For example, we will see later on that in any finite volume Λ and $0 < B < B_c$, the quantum Ising model (1.17) has a unique ground state Ψ_0^B, as opposed to the case $B = 0$, where it has two degenerate ground states $\Psi_\pm^{B=0}$, namely the obvious ones with either all spins up or all spins down. Seen as states in the Hilbert space $\ell^2(E^\Lambda)$, the functions $\Psi_\pm^{B=0}$ are given by $\Psi_\pm^{B=0} = \delta_{s_\pm}$, i.e., $\Psi_\pm^{B=0}(s) = 0$ for all $s \neq s_\pm$ and $\Psi_\pm^{B=0}(s_\pm) = 1$, where $s_\pm(x) = \pm 1$ for all $x \in \Lambda$. Roughly speaking, Ψ_0^B peaks above both s_+ and s_-, like the wave function of the ground state of a symmetric double well potential, or, indeed, like the state of Schrödinger's Cat.

However, in infinite volume the symmetry between s_+ and s_- (or the one $\sigma_i^z \mapsto -\sigma_i^z$ in the Hamiltonian (1.17)) will be broken, so that, as in the finite-volume model with $B = 0$, there are two different ground states, one with all spins up and the other with all spins down.[9] The point, then, is that the restriction of either of those states to finite Λ obviously fails to be of the above "Schrödinger Cat" form, so that it cannot be a ground state of $\hat{h}_\Lambda$.

The correct definition of a ground state relies on the existence of the Heisenberg equation in infinite volume. Recall that in finite volume, this equation reads

$$\frac{da(t)}{dt} = i[\hat{h}_\Lambda, a(t)]. \tag{1.41}$$

Setting $t = 0$, this defines a map $\delta_\Lambda : A(\Lambda) \to A(\Lambda)$ by $\delta_\Lambda(a) = i[\hat{h}_\Lambda, a]$, which is a so-called *derivation*.[10] We now assume that for each $a \in A$ (i.e., $a \in A(\Lambda)$ for some

[9] One way of seeing this is that tunneling between the two classical ground states in finite Λ is suppressed by $\sim \exp(-|\Lambda|)$.

[10] For any algebra A, a derivation is a linear map $\delta : A \to A$ such that $\delta(ab) = \delta(a)b + a\delta(b)$. In classical physics A is a commutative algebra of functions on phase space, and the derivative (w.r.t. either time or some spatial variable) provides an example of a derivation. In quantum physics, as first recognized by Dirac, taking the commutator defines a derivation, as in $\delta(a) = i[h, a]$. The factor i is useful in case that $h^* = h$, because in hat case we have $\delta(a^*) = \delta(a)^*$. Such a derivation is called *symmetric*, *hermitian*, or *self-adjoint*.

finite Λ) the following limit exists:

$$\delta(a) = i \lim_{\Lambda \uparrow \mathbb{Z}^d} [\hat{h}_\Lambda, a]. \tag{1.42}$$

If the interaction Φ has short range, in that spins within Λ only interact with a finite number of spins (within Λ or elsewhere), this will be certainly by the case, because

$$[A(\Lambda_1), A(\Lambda_2)] = 0 \text{ if } \Lambda_1 \cap \Lambda_2 = \emptyset, \ A(\Lambda_i) \subset A. \tag{1.43}$$

More precisely, this *locality* property somewhat symbolically states that if $a \in A(\Lambda_1)$ and $b \in A(\Lambda_2)$, then $[a, b] = 0$. Indeed, although the sum in (the quantum analogue of) (1.3) has increasingly many terms as $\Lambda \uparrow \mathbb{Z}^d$, for fixed $a \in A(\Lambda)$ in (1.42) only finitely many terms will contribute to the commutator.

Exercise 1.10 *Prove*(1.43) *from the definition of* $A(\Lambda) = B(H(\Lambda))$.

If the limit in (1.42) exists for some interaction Φ and ensuing local Hamiltonians $\hat{h}_\Lambda$, we define a *ground state* as a state $\omega_0 : A \to \Lambda$ that for all $a \in A$ satisfies

$$-i\omega_0(a^* \delta(a)) \geq 0. \tag{1.44}$$

To justify this definition, let us assume we have a Hamiltonian $\hat{h}$ on some finite-dimensional Hilbert space H. By adding a constant if necessary, we may assume that its lowest eigenvalue of $\hat{h}$ is $E_0 = 0$, so that $\hat{h} \geq 0$ in the usual sense that

$$\langle \psi, \hat{h}\psi \rangle \geq 0 \tag{1.45}$$

for all $\psi \in H$. If some unit vector ψ_0 satisfies

$$\hat{h}\psi_0 = 0, \tag{1.46}$$

so that it is a ground state in the usual sense, then by (1.45) with $\psi = a\psi_0$ and (1.46) the associated state (in the algebraic sense)

$$\omega_0(a) = \langle \psi_0, a\psi_0 \rangle \tag{1.47}$$

has the property

$$-i\omega_0(a^* \delta(a)) = \langle \psi_0, a^*(ha - ah)\psi_0 \rangle = \langle a\psi_0, ha\psi_0 \rangle \geq 0.$$

Exercise 1.11 *Show, conversely, that for any unit vector* $\psi \in H$ *that does* not *satisfy* $\hat{h}\psi = 0$, *the associated state* $\omega(a) = \langle \psi, a\psi \rangle$ *fails the condition* (1.44).

Finally, we note that the discussion on the *set* of classical ground states in §1.3.4 may be repeated almost *verbatim*: the set of ground states of a quantum system is a compact convex set, whose extreme points are pure states under reasonable conditions on the interaction. As we shall see, this is no longer the case for equilibrium states, where the extreme points correspond to "pure thermodynamic phases".

1.3.6 Equilibrium states of infinite classical systems

Neither the local Hamiltonians (1.3) nor the local partition functions (1.32) have a limit as $\Lambda \uparrow \mathbb{Z}^d$. The correct way to define equilibrium states of infinite classical systems was given in 1968 by Dobrushin and independently by Lanford and Ruelle.

To explain their solution, we need to recall *conditional probabilities*. So far, we have used *probability distributions* p on $S = E^\Lambda$ or $S = E^{\mathbb{Z}^d}$, which associate a number $p(s) \in [0,1]$ to each $s \in S$, subject to the condition $\sum_s p(s) = 1$. More generally, a *probability measure* on a discrete set S is a function $P : \mathsf{P}(S) \to [0,1]$ (where $\mathsf{P}(S)$ is the *power set* of S, i.e., the set of all subsets of S), satisfying

$$P(S) = 1; \tag{1.48}$$

$$P(A \cup B) = P(A) + P(B) \text{ if } A \cap B = \emptyset \ (S \text{ finite}); \tag{1.49}$$

$$P(\cup_i A_i) = \sum_i P(A_i) \text{ if } A_i \cap A_j = \emptyset \, (i \neq j) \ (S \text{ infinite}), \tag{1.50}$$

where $(A_i)_i$ is any countable family of subsets of S. Here $s \in S$ is often called an *outcome* (of some stochastic process), whereas $A \subset S$ is called an *event*. Clearly, a probability distribution p on S gives rise to a probability measure P on S by

$$P(A) = \sum_{s \in A} p(s), \tag{1.51}$$

whilst a probability measure P on S induces a probability distribution p on S by

$$p(s) = P(\{s\}). \tag{1.52}$$

If $P(B) > 0$, the *conditional probability* of A given B is defined by

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{1.53}$$

Now take some finite $\Lambda \subset \mathbb{Z}^d$, and pick a spin configuration $s : \Lambda \to E$ as well as a boundary condition $b : \Lambda^c \to E$. These defines events $\mathsf{s} \subset E^{\mathbb{Z}^d}$ and $\mathsf{b} \subset E^{\mathbb{Z}^d}$ by

$$\mathsf{s} = \{s'' \in E^{\mathbb{Z}^d} \mid s''_{|\Lambda} = s\}; \tag{1.54}$$

$$\mathsf{b} = \{s''' \in E^{\mathbb{Z}^d} \mid s'''_{|\Lambda^c} = b\}, \tag{1.55}$$

whose intersection $\mathsf{s} \cap \mathsf{b} = \{s'\}$ consists of the single spin configuration $s' : \mathbb{Z}^d \to E$ that coincides with s on Λ and coincides with b on Λ^c, or $s'_{|\Lambda} = s$ and $s'_{|\Lambda^c} = b$.

Dobrushin, Lanford, and Ruelle, then, proposed that an equilibrium state of an infinite (spin-like) system is given by a probability distribution p^β on $E^{\mathbb{Z}^d}$ whose associated conditional probabilities for any finite Λ, s, and b as above, are given by

$$P^\beta(\mathsf{s}|\mathsf{b}) = (Z^\beta(b))^{-1} e^{-\beta h_\Lambda(s|b)}, \tag{1.56}$$

where P^β is defined in terms of p^β by (1.51), $h_\Lambda(s|b)$ is given by (1.8), and

$$Z^\beta(b) = \sum_{s \in E^\Lambda} e^{-\beta h_\Lambda(s|b)}. \tag{1.57}$$

Exercise 1.12 *Let $\Lambda' \supset \Lambda$ be finite, but large enough that spins in Λ do not interact with spins outside Λ'. Show that the probability distribution $p^\beta_{\Lambda'}$, defined as in* (1.31)*, satisfies* (1.56) *if $\mathbb{Z}^d$ is replaced by Λ' in the explanation after* (1.53)*.*

1.3.7 Equilibrium states of infinite quantum systems

Attempting to define an equilibrium state of an infinite quantum system as a state whose restriction to finite volume is an equilibrium state of the ensuing finite system is unsatisfactory for the same reason as for ground states. The correct definition once again relies on the possibility of defining dynamics in infinite volume, but now we assume that for each $a \in A$ the limit

$$a(t) = \lim_{\Lambda \uparrow \mathbb{Z}^d} e^{it\hat{h}_\Lambda} a e^{-it\hat{h}_\Lambda} \tag{1.58}$$

exists. Although mathematically speaking this condition is slightly different from the existence of the limit in (1.42), as before (and for the same reason) it is satisfied by short-range interactions. We assume this is the case, so that $a(t)$ exists.

Roughly speaking, a *KMS-state* (named after Kubo, Martin, and Schwinger) on A at fixed inverse temperature $\beta \in \mathbb{R}$ is a state $\omega : A \to \mathbb{C}$ that for all $a, b \in A$ and all $t \in \mathbb{R}$ satisfies

$$\omega(a(t)b) = \omega(ba(t + i\beta)). \tag{1.59}$$

This definition is correct for finite systems i.e., for $A(\Lambda)$ instead of A, but for infinite systems the following more precise formulation is needed: A KMS-state at inverse temperature $\beta \in \mathbb{R}$ is a state ω on A with the following property:

1. For any $a, b \in A$, the function $F_{a,b} : t \mapsto \omega(ba(t))$ from $\mathbb{R}$ to $\mathbb{C}$ has an analytic continuation to the strip $S_\beta = \{z \in \mathbb{C} \mid 0 \leq \mathrm{Im}\,(z) \leq \beta\}$, where it is holomorphic in the interior and continuous on the boundary $\partial S_\beta = \mathbb{R} \cup (\mathbb{R} + i\beta)$;

2. The boundary values of $F_{a,b}$ are related, for all $t \in \mathbb{R}$, by

$$F_{a,b}(t) = \omega(ba(t)); \tag{1.60}$$
$$F_{a,b}(t + i\beta) = \omega(a(t)b). \tag{1.61}$$

This precise definition shows a typical phenomenon for quantum statistical mechanics: time is no longer real, but takes values in the strip S_β. For $\beta \to \infty$, i.e., $T \to 0$, this strip becomes the entire upper half plane in $\mathbb{C}$. In the opposite limit $\beta \to 0$, or $T \to \infty$, a KMS-state obviously becomes a trace, in that $\omega(ab) = \omega(ba)$.

Exercise 1.13 *1. Define a state ω_Λ^β on $A(\Lambda) = B(H(\Lambda))$ by*

$$\omega_\Lambda^\beta(a) = \mathrm{Tr}\,(\hat{\rho}_\Lambda^\beta a), \tag{1.62}$$

where the density matrix $\hat{\rho}_\Lambda^\beta$ is given by (1.38)*. Show that ω_Λ^β satisfies* (1.59)*.*

2. Conversely, show that $\hat{\rho}_\Lambda^\beta$ is the only density matrix whose associated state (1.26) *satisfies* (1.59)*.*

3. For arbitrary operators a and b and state ω, define

$$L_{ab}(t) = i\omega([a(t), b]); \tag{1.63}$$
$$C_{ab}(t) = \omega([a(t), b]_+), \tag{1.64}$$

where $[a,b]_+ = ab + ba$. *Prove the* fluctuation-dissipation theorem:[11]

If ω *is an equilibrium state (i.e., a KMS-state) at inverse temperature* β*, then*

$$\hat{L}_{ab}(\omega) = i \tanh(\tfrac{1}{2}\beta\omega)\hat{C}_{ab}(\omega), \tag{1.65}$$

where the Fourier transform $\hat{f}$ *of an arbitrary function* f *of* t *is defined as*

$$\hat{f}(t) = \int_{-\infty}^{\infty} dt\, e^{-i\omega t} f(t). \tag{1.66}$$

It follows from this exercise that for finite systems the KMS-condition is equivalent to the condition that a state minimizes the free energy, so that the KMS-condition characterizes thermal equilibrium states. As first proposed by Haag, Hugenholtz, and Winnink in 1967, the KMS-condition *defines* thermal equilibrium states also in infinite systems, where the free energy is infinite. This definition has proven its values in all subsequent studies of physical models. It has also led to the correct definition of *pure thermodynamic phases*, namely as the extreme points of the compact convex *set of KMS-states* K_β (at fixed temperature).[12]

The original relationship between equilibrium states and the free energy remains valid for infinite systems, in the sense that KMS-states minimize the *intensive free energy* $\mathsf{f}^\beta : \mathcal{S}(A) \to \mathbb{R}$ (where $\mathcal{S}(A)$ is the (compact convex) set of all states on A), given by

$$\mathsf{f}^\beta(\omega) = \lim_{\Lambda \uparrow \mathbb{Z}^d} \frac{1}{|\Lambda|} \mathcal{F}_\Lambda^\beta(\omega_{|\Lambda}), \tag{1.67}$$

where $\mathcal{F}_\Lambda^\beta$ is given by (1.30), and we assume that the limit exists. Consequently, if ω^β is any KMS-state, and

$$F_\Lambda^\beta = -\beta^{-1} \ln \hat{Z}_\Lambda^\beta, \tag{1.68}$$

we have

$$f^\beta := \lim_{\Lambda \uparrow \mathbb{Z}^d} \frac{1}{|\Lambda|} F_\Lambda^\beta = \mathsf{f}^\beta(\omega^\beta). \tag{1.69}$$

There are many other characterizations and good properties of KMS-states, for which we refer to the literature [4, 11, 12].

[11] Through the linear response theory of Kubo, the function L_{ab} is related to the influence of "dissipative" external influences on the system, whereas C_{ab} is the two-point function for equilibrium fluctuations. The fluctuation-dissipation theorem is even equivalent to the KMS-condition.

[12] In contrast to ground states, extreme points $\omega \in \partial_e K_\beta$ are never pure states.

2 Answers to selected exercises

Exercise 1.3. We first recall that if some Hilbert space K has an orthonormal basis $(e_s)_{s\in S}$, (assumed to be finite or at most countable), then there is a unitary operator $U : K \to \ell^2(S)$. Indeed, we simply define U by linear extension of $Ue_s = \delta_s$, where $\delta_s(t) = \delta_{st}$. In other words, $U(\sum_{s\in S} c_s e_s) = c$ with $c(s) = c_s$, where $\sum_{s\in S} |c_s|^2 < \infty$.

Let $K = H(\Lambda)$ and $S = E^\Lambda$, with $E = \{1, 2, \dots, n\}$. In terms of the standard basis $(e_1, \dots e_n)$ of $H = \mathbb{C}^n$ (any other basis might be used here, too), S now labels the specific orthonormal basis $(e_s)_{s\in S}$ of $H(\Lambda)$ defined by $e_s = \otimes_{x\in\Lambda} e_s(x)$, where $e_s(x) = e_{s(x)}$; recall that $s \in S$ is a function $s : \Lambda \to E$.

Combining everything, we see that $U : H(\Lambda) \to \ell^2(S)$ defined by linear extension of $e_s \mapsto \delta_s$, or, explicitly, $U(\sum_{s\in E^\Lambda} c_s \otimes_{x\in\Lambda} e_s(x)) = c$, with $c(s) = c_s$, is unitary.

Exercise 1.8. We need to show that $\mathcal{F}^\beta_\Lambda(p) \geq -\beta^{-1} \ln Z^\beta_\Lambda$ with equality iff $p = p^\beta_\Lambda$, or, using (1.30), (1.28), and (1.29), that

$$\sum_{s\in E^\Lambda} p(s)(h_\Lambda(s) + \beta^{-1} \ln p(s)) + \beta^{-1} \ln Z^\beta_\Lambda \geq 0. \tag{2.70}$$

Using (1.31), for each $s \in E^\Lambda$ we obtain

$$h_\Lambda(s) = -\beta^{-1} \ln Z^\beta_\Lambda - \beta^{-1} \ln p^\beta_\Lambda(s). \tag{2.71}$$

Substituting this in (2.70), using $\sum_s p(s) = 1$, omitting the ensuing prefactor β^{-1}, and noting that $p^\beta_\Lambda(s) > 0$ for all s, the inequality (2.70) to be proved becomes

$$\sum_{s\in E^\Lambda} p(s) \ln \left(\frac{p(s)}{p^\beta_\Lambda(s)} \right) \geq 0. \tag{2.72}$$

Hence we need to prove the inequality

$$\sum_{s\in E^\Lambda} p^\beta_\Lambda(s) \cdot \left(\frac{p(s)}{p^\beta_\Lambda(s)} \right) \ln \left(\frac{p(s)}{p^\beta_\Lambda(s)} \right) \geq 0, \tag{2.73}$$

with equality iff $p(s) = p^\beta_\Lambda(s)$ for all s.

Let us now note that the function $f(x) = x \ln x$ is strictly convex for all $x \geq 0$, that is, for any finite set of numbers $p'(s) \in (0, 1)$ with $\sum_s p'(s) = 1$ and any set of positive real numbers $(x_s)_s \geq 0$, we have

$$\sum_s p'(s) f(x_s) \geq f(\sum_s p'(s) x_s), \tag{2.74}$$

with equality iff all numbers x_s are the same. Applying this with $p'(s) = p^\beta_\Lambda(s)$ and $x_s = p(s)/p^\beta_\Lambda(s)$, so that $p'(s)x_s = p(s)$ and hence $\sum_s p'(s)x_s = \sum_s p(s) = 1$, which makes the right-hand side of (2.74) vanish since $\ln(1) = 0$, finally leads to (2.73). Equality arises iff $p(s)/p^\beta_\Lambda(s)$ equals the same numer c for all s; summing over all s forces $c = 1$, so that one has equality iff $p(s) = p^\beta_\Lambda(s)$ for all s, as desired.

3 Mean Field Theory

Exercise 3.1 *1. Check (3.8) on p. 25 of Parisi.*

2. *Verify Parisi's assertion that the free energy $\Phi = \Phi(m_i)$ in (3.8) is minimized iff $m_i = m$ for all $i \in \Lambda$, for a suitable $m \in [-1, 1]$. Here Parisi's $\Phi[P]$ is our $\mathcal{F}^\beta_\Lambda(p)$.*

3. *From this result, assuming that $m_i = m$ for all $i \in \Lambda$, verify that for small m:*

$$f(\beta, h) = -hm + (\tfrac{1}{2}T - DJ)m^2 + \frac{T}{12}m^4 + \mathcal{O}(m^6). \tag{3.75}$$

Parisi's (3.16) is the *Bogoliubov inequality* for the exact free energy:

$$F \leq F_0 + \langle h - h_0 \rangle_0, \tag{3.76}$$

where (for fixed Λ and β) for the sake of readability we have omitted all suffixes Λ and β, so that F equals F^β_Λ as defined in (1.33), (1.31), and (1.32), F_0 is defined by (1.33) with h in (1.31) and (1.32) replaced by any "trial Hamiltonian" h_0, and

$$\langle a \rangle_0 = \sum_{s \in E^\Lambda} p_0(s) a(s), \tag{3.77}$$

with p_0 as defined in (1.31) and (1.32) with, once again, h replaced by h_0.

Exercise 3.2 *Show that $F_0 + \langle h - h_0 \rangle_0 = \mathcal{F}^\beta_\Lambda(p_0)$, cf. (1.30), and argue that (given theses lecture notes) Parisi's convexity proof of (3.76) is unnecessary.*

4 Low- and High-Temperature expansions

4.1 Low T

To begin with, we take the Hamiltonian for the Ising model in zero external field:

$$h_\Lambda(s) = -J \sum_{\langle ij \rangle_\Lambda} s_i s_j. \tag{4.78}$$

Introducing the set $\mathcal{B}(\Lambda)$ of all nearest-neighbour pairs within Λ, we typically write b for some specific pair $\{i, j\} \in \mathcal{B}(\Lambda)$, and accordingly, $s(b) = s_i s_j$. Also, define a map $\gamma : E^\Lambda \to \mathbb{P}(\mathcal{B}(\Lambda))$, where $\mathbb{P}(X)$ is the power set of some set X (that is, the set of all subsets of X; N.B. we write $|X|$ for the number of elements of a set X), by

$$\gamma(s) = \{b \in \mathcal{B}(\Lambda) \mid s(b) = -1\}. \tag{4.79}$$

We may then rewrite the partition function (1.32) in finite $\Lambda \subset \mathbb{Z}^D$ as

$$\begin{aligned} Z^\beta_\Lambda &= \sum_{s \in E^\Lambda} e^{-\beta h_\Lambda(s)} = \sum_{s \in E^\Lambda} e^{\beta J \sum_{\langle ij \rangle_\Lambda} s_i s_j} = \sum_{s \in E^\Lambda} e^{\beta J \sum_{b \in \mathcal{B}(\Lambda)} s(b)} \\ &= \sum_{s \in E^\Lambda} e^{\beta J \sum_{b \in \mathcal{B}(\Lambda)} (s(b) - 1 + 1)} = e^{\beta J |\mathcal{B}(\Lambda)|} \sum_{s \in E^\Lambda} e^{\beta J \sum_{b \in \mathcal{B}(\Lambda)} (s(b) - 1)} \\ &= e^{\beta J D |\Lambda|} \sum_{s \in E^\Lambda} e^{-2\beta J |\gamma(s)|} = 2 e^{\beta J D |\Lambda|} \sum_{B \in \gamma(E^\Lambda)} e^{-2\beta J |B|}. \end{aligned} \tag{4.80}$$

With $t = \exp(-4\beta J$, it follows from (4.80) and (1.69) that, taking into account the empty set $B = \emptyset$, the intensive free energy is given by

$$f^\beta = -JD - \beta^{-1} \lim_{\Lambda \uparrow \mathbb{Z}^d} \frac{1}{|\Lambda|} \ln \left(1 + \sum_{B \in \gamma(E^\Lambda)} t^{|B|/2} \right). \tag{4.81}$$

Expanding the logarithm as $\ln(1+x) = x - x^2/2 + x^3/3 - \cdots$ yields the *low-temperature expansion* of the free energy. This expansion is difficult, because the sum over B interferes with the power series expansion of ln.

Exercise 4.1 *1. Reproduce Parisi's (4.6) on p. 48 from* (4.81).

2. Generalize the derivation of (4.81) *for Hamiltonians of the form*

$$h_\Lambda(s) = -\sum_{b \in \mathcal{B}(\Lambda)} J(b) s(b), \tag{4.82}$$

where $J : \mathcal{B}(\Lambda) \to \mathbb{R}$ is some function.

4.2 High T

It is instructive to start with a naive high-T expansion for the Ising model, that is,

$$\begin{aligned} Z_\Lambda^\beta &= \sum_{s \in E^\Lambda} e^{\beta J \sum_{\langle ij \rangle_\Lambda} s_i s_j} = \sum_s \prod_{\langle ij \rangle_\Lambda} e^{\beta J s_i s_j} \qquad (4.83) \\ &= \sum_s \prod_{\langle ij \rangle_\Lambda} \sum_{n=0}^\infty \frac{(\beta J)^n}{n!} s_x^n s_y^n = \sum_s \prod_{b \in \mathcal{B}(\Lambda)} \sum_{n=0}^\infty \frac{(\beta J)^n}{n!} s(b)^n \\ &= \sum_s \sum_{\nu \in \mathbb{N}^{\mathcal{B}(\Lambda)}} \prod_b \frac{(\beta J)^{\nu_b}}{\nu_b!} s(b)^{\nu_b} = \sum_\nu \frac{(\beta J)^{\sum_b \nu_b}}{\prod_b \nu_b!} \sum_s \prod_{x \in \Lambda} s_x^{\nu(x)} \\ &= 2^{|\Lambda|} \sum_\nu{}' \frac{(\beta J)^{\sum_b \nu_b}}{\prod_b \nu_b!} \equiv 2^{|\Lambda|} \sum_{\mathcal{G}} w(\mathcal{G}) (\beta J)^{|\mathcal{G}|}. \qquad (4.84) \end{aligned}$$

Here $\nu(x) = \sum_{b \ni x} \nu_b$, the restricted sum $\sum_\nu'$ is over all configurations $\nu : \mathcal{B}(\Lambda) \to \mathbb{N}$ for which $\nu(x)$ is even for each $x \in \Lambda$, and the final sum is over all topologically different graphs $\mathcal{G}$ in Λ with associated weights $w(\mathcal{G})$ and length $|\mathcal{G}|$. In this case, a graph is just a collection of lines drawn between nearest-neigbour vertices $\langle ij \rangle$ of Λ, subject to the rules that any number $\nu_{\langle ij \rangle} \in \mathbb{N}$ of lines may be drawn (including zero), and that the number of lines terminating at each vertex i must be even (including zero). The weight $w(\mathcal{G})$ is the product of $\prod_{b \in \mathcal{G}} \nu_b!$ and the number of distinct ways a graph of the given topological type may be drawn inside Λ.

For example, the empty graph has weight $w(\emptyset) = 1$. The graph $\circ = \circ$ consisting of two lines between some nearest-neighbour pair has weight $w(\circ = \circ) = \frac{1}{2} D|\Lambda|$. The idea, then, is that for high T (= low β) only small graphs (i.e., graphs for which $|\mathcal{G}|$ is small) contribute significantly to the expansion (4.84).

Exercise 4.2 *1. Compute the coefficient of β^4 in Z_Λ^β in any D.*

2. Compute the free energy to order β^4 and verify that $\ln Z_\Lambda^\beta \sim |\Lambda|$ to this order.

A simpler high-T expansion for the Ising model arises if we introduce the variable $t = \tanh(\beta J)$, cf. [**?**, **?**]. The result is

$$Z_\Lambda^\beta \cdot 2^{-|\Lambda|}(\cosh(\beta J))^{-D|\Lambda|} = \sum_{\nu \in \underline{2}^{\mathcal{B}(\Lambda)}}{}' t^{\sum_b \nu_b} \equiv \sum_{\mathcal{G}} w(\mathcal{G}) t^{|\mathcal{G}|}. \tag{4.85}$$

where $\underline{2} = \{0, 1\}$, so that this time only graphs consisting of single lines between nearest-neigbour vertices contribute, still subject to the rule that $\nu(x)$ (i.e., the number of lines terminating at vertex x–now either zero or one) be even for each $x \in \Lambda$. This implies that only closed paths contribute, and $w(\mathcal{G})$ just counts the number of ways $\mathcal{G}$ may be drawn. Apart from the empty graph, this process starts at $|\mathcal{G}| = 4$, for which only one type of graph exists, namely the elementary square or 'plaquette'. For order $|\mathcal{G}| = 8$ and higher, both connected and disconnected graphs contribute (even to the free energy).

Exercise 4.3 *Compute* (4.85) *to order t^4, compute the corresponding free energy to order β^4, and verify consistency with the previous exercise.*

4.3 Gaussian model

See Parisi [**?**], §4.2.

Exercise 4.4 *1. Find the 'critical' value of β for which the $N \times N$ ($N = |\Lambda|$) matrix $A = \mathbb{I} - \beta J$ defining the Gaussian model loses positive definiteness.*

2. Prove the last equation in (4.14), p. 50, of Parisi.

5 Renormalization Group

Theory: Mussardo Ch. 8.

Exercise 5.1 *This exercise is about the classical Ising model in $D = 1$, with the usual variables $s_i = \pm 1$, $i = 1, \ldots, N$ (assumed even), $K = \beta J$ and $h = 0$, so that*

$$-\beta h_N(s) = K \sum_{i=1}^{N} s_i s_{i+1}. \tag{5.86}$$

1. *Show that*

$$\sum_{s_i = \pm 1, i\, even} e^{-\beta h_N(s)} = e^{-\beta h'_{N/2}(s')}, \tag{5.87}$$

where

$$-\beta h'_{N/2}(s') = \sum_{i=1}^{N/2} (K' s'_i s'_{i+1} + K' + \ln(2)), \tag{5.88}$$

and, in terms of $L = \exp(2K)$, so that $L \in [1, \infty]$ as $K \in [0, \infty]$,

$$L' = \tfrac{1}{2}(L + L^{-1}). \tag{5.89}$$

2. *Show that the fixed point $L_* = 1$ (i.e., $L'_* = L_*$) is attractive, whereas the fixed point $L_* = \infty$ is repelling (so that the renormalization flow is from $L = \infty$ towards $L = 0$, or from $T = 0$ to $T = \infty$).*

3. *Assuming $N/3 \in \mathbb{N}$, repeat this analysis summing over all spins except nos. 2, 5, 8, 11, ..., using periodic boundary conditions (in other words, of each block spin of three, keep the middle one and sum over the rest). Show, e.g., that*

$$-\beta h'_{N/3}(s') = \sum_{i=1}^{N/3} (K' s'_i s'_{i+1} + g(K)), \tag{5.90}$$

for suitable $g(K)$, where this time, writing $M = \tanh(K)$, we have

$$M' = M^3. \tag{5.91}$$

4. *Assuming $N/(2l+1) \in \mathbb{N}$, generalize the previous approach from block spins sized 3 to block spin sized $2l + 1$, $l \in \mathbb{N}$. The result looks like (5.90), but with*

$$M' = M^{2l+1}. \tag{5.92}$$

Exercise 5.2 *For general coupling constants $K = (K_1, \ldots, K_n)$, write the RG transformation induced by a block spin consisting of b original spins by $K' = R_b(K)$. For some fixed point K_*, that is, $R_b(K_*) = K_*$, denote the eigenvalues of the $n \times n$ matrix $R'_b(K_*)$ by $\lambda_i(b)$. Show that $\lambda_i(b_1)\lambda_i(b_2) = \lambda_i(b_1 b_2)$, so that $\lambda_i(b) = b^{y_i}$ for some 'critical index' y_i (all critical exponents are expressible in terms of these y_i).*

6 Exact solution of the quantum Ising chain

6.1 Fermionic Fock space and CAR

Let H be a Hilbert space and let $F_-(H)$ be the corresponding fermionic Fock space, i.e.,

$$F_-(H) = \oplus_{n=0}^{\infty} H_-^{\otimes n}, \tag{6.93}$$

where $H^{\otimes 0} = \mathbb{C}$, and for $n > 0$ we have

$$H_-^{\otimes n} = P_-^{(n)} H^{\otimes n} \tag{6.94}$$

is the totally antisymmetrized n-fold tensor product of H with itself. Here the projection $P_-^{(n)} : H^{\otimes n} \to H^{\otimes n}$ is defined by linear extension of

$$P_-^{(n)} f_1 \otimes \cdots \otimes f_n = \frac{1}{n!} \sum_{\pi \in S_n} \epsilon(\pi) f_{\pi(1)} \otimes \cdots \otimes f_{\pi(n)}, \tag{6.95}$$

where S_N is the permutation group on n objects and $\epsilon(\pi)$ is $+1/-1$ if π is an even/odd permutation. On the usual Fock space

$$F(H) = \oplus_{n=0}^{\infty} H^{\otimes n}, \tag{6.96}$$

for each $f \in H$ we define the usual *annihilation operator* $a(f)$ by linear extension of

$$a(f) f_1 \otimes \cdots \otimes f_n = \sqrt{n} (f, f_1)_H \otimes \cdots \otimes f_n, \tag{6.97}$$

for $n > 0$, with $a(f)z = 0$ on $H^{\otimes 0} = \mathbb{C}$, with adjoint $a(f)^* \equiv a^*(f)$ given by

$$a^*(f) f_1 \otimes \cdots \otimes f_n = \sqrt{n+1} f \otimes f_1 \otimes \cdots \otimes f_n. \tag{6.98}$$

Exercise 6.1 *Show that* $(\Psi, a^*(f)\Phi)_{F(H)} = \overline{(\Phi, a(f)\Psi)}_{F(H)}$ *for all* $\Phi, \Psi \in F(H)$.

For each $f \in H$, we then define the following operators on the fermionic Fock space:

$$c(f) = P_- a(f) P_-; \tag{6.99}$$

$$c^*(f) = P_- a^*(f) P_-. \tag{6.100}$$

Exercise 6.2 *Show that* $c^*(f) = c(f)^*$ *and prove the anticommutation relations*

$$[c(f), c^*(g)]_+ = (f, g)_H \cdot 1; \tag{6.101}$$

$$[c(f), c(g)]_+ = 0; \tag{6.102}$$

$$[c^*(f), c^*(g)]_+ = 0. \tag{6.103}$$

Of course, choosing an orthonormal basis (e_i) of H and writing $c(e_i) = c_i$ etc. yields

$$[c_i, c_j^*]_+ = \delta_{ij} \cdot 1; \tag{6.104}$$

$$[c_i, c_j]_+ = 0; \tag{6.105}$$

$$[c_i^*, c_j^*]_+ = 0. \tag{6.106}$$

The algebra $CAR(H)$ of *Canonical Anticommutation Relations* over H is the operator algebra (technically, the C*-algebra) generated by the operators $c(f)$ and $c^*(f)$, $f \in H$, on $F_-(H)$.

6.2 Jordan–Wigner transformation

The following fact is of great importance.

Proposition 6.3 *If* $\dim(H) = N < \infty$*, then*

$$F_-(\mathbb{C}^N) = \oplus_{n=0}^N H_-^{\otimes n} \cong \mathbb{C}^{2^N}, \tag{6.107}$$

and

$$CAR(\mathbb{C}^N) \cong M_{2^N}(\mathbb{C}). \tag{6.108}$$

Exercise 6.4 *Prove* (6.107) *by a dimension count and prove* (6.108) *via* Schur's Lemma*: if* $A \subset M_k(\mathbb{C})$ *is an algebra of matrices (i.e.,* A *is closed under linear operations, multiplication, and taking adjoints) containing the unit, then* $A = M_k(\mathbb{C})$ *iff the only matrices that commute with all* $a \in A$ *are multiples of the identity.*

This result is important to us, because also

$$\otimes^N \mathbb{C}^2 \cong \mathbb{C}^{2^N}, \tag{6.109}$$

so that by Proposition 6.3,

$$F_-(\mathbb{C}^N) = \otimes^N \mathbb{C}^2; \tag{6.110}$$

$$CAR(\mathbb{C}^N) \cong M_{2^N}(\mathbb{C}). \tag{6.111}$$

This is already nontrivial for $N = 1$, in which case $F_-(\mathbb{C}) \cong \mathbb{C}^2$, and, with $a \equiv a(1)$,

$$a = \sigma^- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}; \tag{6.112}$$

$$a^* = \sigma^+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \tag{6.113}$$

gives a realization of $CAR(\mathbb{C})$, with the usual notation $\sigma^\pm = \frac{1}{2}(\sigma^x \pm i\sigma^y)$.

To generalize this to arbitrary $N > 1$, we put a copy of the Pauli matrices on each site i of a chain $\underline{N} = \{1, 2, \ldots, N\}$, denoted by σ_i^μ, $\mu = 0, 1, 2, 3$, with $\sigma^0 = 1$. Following [8], for each $i = 1, 2, \ldots, N-1$ we then define operators

$$c_i = e^{\pi i \sum_{j=1}^{i-1} \sigma_j^+ \sigma_j^-} \sigma_i^-; \tag{6.114}$$

$$c_i^* = e^{-\pi i \sum_{j=1}^{i-1} \sigma_j^+ \sigma_j^-} \sigma_i^+, \tag{6.115}$$

along with $c_1 = \sigma_1^-$ and $c_1^* = \sigma_1^+$. Since

$$c_i^* c_i = a_i^* a_i = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{6.116}$$

the inverse transformation is given by

$$a_i = e^{-\pi i \sum_{j=1}^{i-1} c_j^* c_j} c_i; \tag{6.117}$$

$$a_i^* = c_i^* e^{\pi i \sum_{j=1}^{i-1} c_j^* c_j}. \tag{6.118}$$

More concretely, since the operators $\sigma_j^+\sigma_j^-$ commute for different sites j, and

$$e^{\pi i\sigma^+\sigma^-} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{6.119}$$

the so-called *Jordan–Wigner transformation* (6.114) - (6.115) is simply given by

$$c_i = \prod_{j=1}^{i-1} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}_j \cdot \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}_i; \tag{6.120}$$

$$c_i^* = \prod_{j=1}^{i-1} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}_j \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}_i. \tag{6.121}$$

Exercise 6.5 *1. Defining*

$$u = \begin{pmatrix} \sqrt{1/2} & \sqrt{1/2} \\ -\sqrt{1/2} & \sqrt{1/2} \end{pmatrix}, \tag{6.122}$$

and subsequently $u_{(N)} = \otimes_{i=1}^N u_i$*, show that*

$$u_{(N)} h_N u_{(N)}^* = h'_N, \tag{6.123}$$

where

$$h_N = -J\sum_{i=1}^N \sigma_i^z\sigma_{i+1}^z - \Gamma\sum_{i=1}^N \sigma_i^x; \tag{6.124}$$

$$h'_N = -J\sum_{i=1}^N \sigma_i^x\sigma_{i+1}^x - \Gamma\sum_{i=1}^N \sigma_i^z. \tag{6.125}$$

2. Show that

$$h'_N = \tfrac{1}{2}\Gamma N - \Gamma\sum_{i=1}^N c_i^*c_i - \tfrac{1}{4}J\sum_{i=1}^N (c_i^* - c_i)(c_{i+1}^* + c_{i+1}) \tag{6.126}$$

$$+\tfrac{1}{4}J(c_N^* - c_N)(c_1^* + c_1)(e^{\pi i\sum_{j=1}^N c_j^*c_j} + 1). \tag{6.127}$$

For large N, the term on the second line can be neglected (as it is bounded, unlike the first) [8]. So we now show how to diagonalize quadratic fermionic Hamiltonians of the type

$$h_N = \sum_{i,j=1}^N \left(A_{ij}c_i^*c_j + \tfrac{1}{2}B_{ij}(c_i^*c_j^* - c_ic_j)\right), \tag{6.128}$$

where A and B are real $N \times N$ matrices, with $A^T = A$ and $B^T = -B$. This is done by a so-called Bogoliubov transformation, whose abstract theory is as follows.

6.3 Bogoliubov transformation

The passage (6.129) - (6.130) from operators c, c^* satisfying the CAR, to new operators η, η^* that also satisfy the CAR, as expressed in the following theorem, is called a *Bogoliubov transformation*.

Theorem 6.6 *Let u and v be operators on a Hilbert space H, where u is linear and v is anti-linear (i.e., $v(\lambda\Psi) = \overline{\lambda}v(\Psi)$ for $\lambda \in \mathbb{C}$ and $\Psi \in H$). Let $c(f)$ and $c^*(f)$ be the operators* (6.99) - (6.100)*, satisfying the CAR* (6.101) - (6.103)*. Define*

$$\eta(f) = c(uf) + c^*(vf); \tag{6.129}$$
$$\eta^*(f) = c^*(uf) + c(vf). \tag{6.130}$$

Then the canonical anticommutation relations

$$[\eta(f), \eta^*(g)]_+ = (f,g)_H \cdot 1; \tag{6.131}$$
$$[\eta(f), \eta(g)]_+ = 0; \tag{6.132}$$
$$[\eta^*(f), \eta^*(g)]_+ = 0. \tag{6.133}$$

hold if and only if u and v satisfy

$$uv^* + vu^* = v^*u + u^*v = 0; \tag{6.134}$$
$$u^*u + v^*v = uu^* + vv^* = 1. \tag{6.135}$$

Exercise 6.7 *Prove this theorem.*

Taking into account that $c(f)$ is antilinear in f, whereas $c^*(f)$ is linear, with respect to a base (e_i) of H the Bogoliubov transformation (6.129) - (6.130) looks like

$$\eta_i = \sum_j (\overline{u}_{ji} c_j + v_{ji} c_j^*); \tag{6.136}$$
$$\eta_j^* = \sum_j (u_{ji} c_j^* + \overline{v}_{ji} c_j). \tag{6.137}$$

6.4 Dynamics

With $H = \mathbb{C}^N$ and $c_k := c(e_k)$ and $c_k^* := c^*(e_k)$, defined in terms of the usual orthonormal base $(e_k)_{k=1}^N$ of $\mathbb{C}^N$, take the free Hamiltonian h on $F_-(\mathbb{C}^N)$ defined by

$$h = \sum_{k=1}^N \varepsilon_k c_k^* c_k, \tag{6.138}$$

where $\varepsilon_k \geq 0$ for each k. Note that this is the second quantization of the single-particle Hamiltonian $h^{(1)} : \mathbb{C}^N \to \mathbb{C}^N$ given by (linear extension of) $h^{(1)} e_k = \varepsilon_k e_k$. In this simple situation, the Heisenberg equations of motion for c and c^* can be easily solved, with the result that for arbitrary $f \in C^N$ we have

$$c(f)(t) \equiv e^{ith} c(f) e^{-ith} = c(e^{ith^{(1)}} f); \tag{6.139}$$
$$c^*(f)(t) \equiv e^{ith} c^*(f) e^{-ith} = c^*(e^{ith^{(1)}} f). \tag{6.140}$$

In other words, we have

$$c(f)(t) = c(f(t)), \tag{6.141}$$

where $f(t)$ solves the (anti) Schrödinger equation ($-i$ instead of i)

$$-i\frac{df(t)}{dt} = h^{(1)}f, \tag{6.142}$$

with initial condition $f(0) = f$. As a special case, we obviously have

$$c_k(t) = e^{-it\varepsilon_k}c_k; \tag{6.143}$$

$$c_k^*(t) = e^{it\varepsilon_k}c_k^*. \tag{6.144}$$

This remains true if H is infinite-dimensional, as long as the Hamiltonian h on $F_-(H)$ is the second quantization of some self-adjoint single-particle Hamiltonian $h^{(1)} : H \to H$ (possibly even unbounded). One way to define what this means is:

Exercise 6.8 *Given some self-adjoint single-particle Hamiltonian $h^{(1)} : H \to H$ (assumed bounded for convenience), show that there is a self-adjoint operator $h : F_-(H) \to F_-(H)$, unique up to an additive constant, such that* (6.139) - (6.140) *hold. Also show that* (6.138)*,* (6.143)*,* (6.144) *is a special case of this construction.*

However, most fermionic Hamiltonians are not in this form, even if they are quadratic in the c and c^*. In that case, h is diagonalizable by a Bogoliubov transformation, which means that there is a different fermionic Fock space (i.e., the one for which $|0\rangle$, the unit 1 of $H^{\otimes 0} = \mathbb{C}$) satisfies $\eta(f)|0\rangle$ for all $f \in H$) on which h is the second quantization of some $h^{(1)}$. The simplest formalism to accomplish Bogoliubov transformations for infinite-dimensional H is as follows.

6.5 Self-dual formalism

The self-dual formulation of the CAR, due to Araki, treats c and c^* on equal footing. The advantage of this formalism is that formulae like (6.139) - (6.140) hold even if h is not (already) of the form (6.138). The basic trick is to double H into $K = H \oplus H$, with elements written $h = (f, g)$ or $h = f\dot{+}g$, indicating that f lies in the first copy of H, whereas g lies in the second. The inner product in K is given by

$$\langle h_1, h_2\rangle_K = \langle f_1, f_2\rangle_H + \langle g_1, g_2\rangle_H. \tag{6.145}$$

We also need a conjugation on H, that is, an antilinear map $J : H \to H$ satisfying $J^* = J$ and $J^2 = 1$. For $H = \ell^2(\mathbb{Z})$, the simplest example is

$$Jf_i = \overline{f_i}. \tag{6.146}$$

However, on

$$\hat{H} = L^2([-\pi, \pi]) \tag{6.147}$$

it turns out to be appropriate to use the Fourier transform of (6.146), which is

$$\hat{J}\hat{\psi}(k) = \overline{\hat{\psi}(-k)}. \tag{6.148}$$

We then introduce the "field" (often called B instead of Φ)

$$\Phi(h) = c^*(f) + c(Jg), \tag{6.149}$$

which is *linear* in $h = f\dot{+}g$, because the antilinearity of $c(f)$ in f is canceled by the antilinearity of J. This yields

$$[\Phi^*(h_1), \Phi(h_2)]_+ = \langle h_1, h_2 \rangle_K, \tag{6.150}$$

but 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

$$\Gamma = \begin{pmatrix} 0 & J \\ J & 0 \end{pmatrix} \tag{6.151}$$

we have

$$\Phi^*(h) \equiv \Phi(h)^* = \Phi(\Gamma h). \tag{6.152}$$

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

$$c^*(f) = \Phi(f); \tag{6.153}$$
$$c(f) = \Phi(\Gamma f). \tag{6.154}$$

The self-dual formulation of Bogoliubov-transformations is now extremely elegant, as follows: for any unitary operator S on K that satisfies $[S, \Gamma] = 0$, we define a Bogoliubov-transformed field

$$\Phi_S(h) = \Phi(Sh), \tag{6.155}$$

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

$$c_S^*(f) = \Phi_S(f); \tag{6.156}$$
$$c_S(f) = \Phi_S^*(f). \tag{6.157}$$

Exercise 6.9 *Show that S necessarily has the form*

$$S = \begin{pmatrix} u & vJ \\ Jv & JuJ \end{pmatrix}, \tag{6.158}$$

where $u : H \to H$ is linear and $v : H \to H$ is antilinear, and u and v satisfy (6.134) *and* (6.135)*. Subsequently, prove that $c_S(f)$ and $c_S^*(f)$ coincide with $\eta(f)$ and $\eta^*(f)$ in* (6.129) *and* (6.130)*, respectively.*

6.6 Diagonalizing the Hamiltonian

The point of all this is that we can now diagonalize the $N \to \infty$ limit of the Hamiltonian (6.128), or rather, for a two-sided chain (assuming N even)

$$h_N = \sum_{i,j=-N/2}^{N/2-1} \left(A_{ij} c_i^* c_j + \tfrac{1}{2} B_{ij} (c_i^* c_j^* - c_i c_j) \right), \tag{6.159}$$

This limit exist in the sense that, for $H = \ell^2$ and any $a \in CAR(H)$, the derivation

$$\delta(a) = i \lim_{N\to\infty} [h_N, a] \tag{6.160}$$

is well defined, whilst also the Heisenberg-picture time evolution

$$a(t) = \lim_{N\to\infty} e^{ith_N} a e^{-ith_N} \tag{6.161}$$

exists. Combining this, the operator Heisenberg equation

$$\frac{da(t)}{dt} = i \lim_{N\to\infty} [h_N, a(t)] \tag{6.162}$$

is well defined and has a unique solution (6.161) subject to $a(0) = a$. The point is:

Exercise 6.10 *Show that for any $h \in K$ one has (cf.* (6.141) - (6.142)*)*

$$\Phi(h)(t) = \Phi(h(t)), \tag{6.163}$$

where $h(t)$ solves the 'doubled' (anti) Schrödinger equation

$$-i\frac{dh(t)}{dt} = h_D^{(1)} h, \tag{6.164}$$

with initial condition $h(0) = h$. Here the 'doubled' one-particle Hamiltonian operator $h_D^{(1)} : K \to K$ is given by

$$h_D^{(1)} = \begin{pmatrix} A & B \\ -B & -A \end{pmatrix}, \tag{6.165}$$

where $A : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ and $B\ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ are the obvious extensions of the $N \times N$ matrices A and B to operators on $\ell^2(\mathbb{Z})$ (assuming the entries A_{ij} and B_{ij} are defined whenever $(i, j) \in \mathbb{Z} \times \mathbb{Z}$).

Exercise 6.11 *Recall the finite-N Hamiltonian for the quantum Ising chain:*

$$\hat{h}_N^I = -\sum_{j=-N/2}^{N/2-1} \left(\sigma_j^x \sigma_{j+1}^x + \lambda \sigma_j^z\right), \tag{6.166}$$

where we (conventionally) have put $J = 1$ and interchanged the x- and z-axes.

1. *Show that for this Hamiltonian*

$$A = \tfrac{1}{2}(\mathsf{S} + \mathsf{S}^*) - \lambda; \tag{6.167}$$
$$B = \tfrac{1}{2}(\mathsf{S} - \mathsf{S}^*), \tag{6.168}$$

where $\mathsf{S} : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ is the shift operator*, defined on a sequence (f_j) by*

$$\mathsf{S} f_j = f_{j+1}; \tag{6.169}$$
$$\mathsf{S}^* f_j = f_{j-1}. \tag{6.170}$$

Note that in terms of the standard basis (e_j) of $\ell^2(\mathbb{Z})$ we have $\mathsf{S}e_j = e_{j-1}$ and $\mathsf{S}^* e_j = e_{j+1}$.

2. Show that a Fourier transformation $U : \ell^2(\mathbb{Z}) \to L^2([-\pi,\pi])$, given by

$$(Uf)(k) \equiv \hat{f}(k) = \sum_{j\in\mathbb{Z}} e^{-ijk} f_j; \tag{6.171}$$

$$(U^{-1}\hat{f})_j \equiv f_j = \int_{-\pi}^{\pi} \frac{dk}{2\pi}\, e^{ijk}\hat{f}(k), \tag{6.172}$$

diagonalizes A and B to $\hat{A}, \hat{B} : L^2([-\pi,\pi]) \to L^2([-\pi,\pi])$, given by

$$\hat{A}\hat{\psi}(k) = ((\cos k) - \lambda)\hat{\psi}(k); \tag{6.173}$$

$$\hat{B}\hat{\psi}(k) = -i \sin k\, \hat{\psi}(k). \tag{6.174}$$

3. For fixed k, show that the eigenvalues and eigenvectors of the 2×2 matrix

$$M_k = \begin{pmatrix} (\cos k) - \lambda & -i \sin k \\ i \sin k & -(\cos k - \lambda), \end{pmatrix} \tag{6.175}$$

are $\pm\varepsilon_k$, with

$$\varepsilon_k = \sqrt{1 + \lambda^2 - 2\lambda \cos k}. \tag{6.176}$$

4. Show that the unitary 2×2 matrix

$$U_k = \begin{pmatrix} u_k & v_k \\ v_k & u_k \end{pmatrix}; \tag{6.177}$$

$$u_k = N_k \sin k; \tag{6.178}$$

$$v_k = iN_k(\varepsilon_k + \lambda - \cos k), \tag{6.179}$$

$$N_k = (1 + (\varepsilon_k + \lambda)^2 - 2(\varepsilon_k + \lambda)\cos k)^{-1/2} \tag{6.180}$$

$$= (\sin^2 k + (\varepsilon_k + \lambda - \cos k)^2)^{-1/2}. \tag{6.181}$$

diagonalizes M_k in the sense that

$$U_k^{-1} M_k U_k = \begin{pmatrix} \varepsilon_k & 0 \\ 0 & -\varepsilon_k \end{pmatrix}. \tag{6.182}$$

5. With $\hat{K} = \hat{H} \oplus \hat{H}$, where $\hat{H} = L^2([-\pi,\pi])$, we now turn M_k and U_k into multiplication operators M and U on $\hat{K}$ in the obvious way, so that

$$M = \hat{h}_D^{(1)} = \begin{pmatrix} \hat{A} & \hat{B} \\ -\hat{B} & -\hat{A} \end{pmatrix}, \; U = \begin{pmatrix} u & v \\ v & u \end{pmatrix}, \tag{6.183}$$

with $\hat{A}$ and $\hat{B}$ defined by (6.173) - (6.174)*, and $u\hat{\psi}(k) = u_k\psi(k)$, etc. Prove*

$$U^{-1} M U = \begin{pmatrix} \varepsilon & 0 \\ 0 & -\varepsilon \end{pmatrix}, \tag{6.184}$$

where $\varepsilon : \hat{H} \to \hat{H}$ is the multiplication operator defined by $\varepsilon\hat{\psi}(k) = \varepsilon_k\hat{\psi}(k)$.

6.7 Ground states for the CAR-algebra

The fermionic Fock space $F_-(H)$ has a special state $|0\rangle$ for which $\langle 0|0\rangle = 1$ and $c(f)|0\rangle = 0$, and all states in $F_-(H)$ are obtained by repeatedly acting on $|0\rangle$ with the $c^*(f)$ and taking linear combinations (and limits). More abstractly, the unit vector $|0\rangle$ defines a state $\omega_0 : CAR(H) \to \mathbb{C}$ on $CAR(H)$ in the abstract sense by

$$\omega_0(a) := \langle 0|a|0\rangle, \;\; a \in CAR(H). \tag{6.185}$$

This state turns out to be a ground state for the simple time-evolution given by (6.143) - (6.144), or, equivalently, for the limiting dynamics defined by the finite-volume Hamiltonians (6.138) (though the sum should run as indicated below).

Exercise 6.12 *1. With $h_N = \sum_{k=-N/2}^{N/2-1} \varepsilon_k c_k^* c_k$ and $\delta(a) := i[h_N, a]$ as a derivation $\delta : CAR(\mathbb{C}^N) \to CAR(\mathbb{C}^N)$, show that ω_0 is a ground state à la (1.44).*

2. Show that there exists no other pure ground state on $CAR(\mathbb{C}^N)$, and hence no other ground state whatsoever. You may use the fact that the construction

$$\psi(a) = \langle \Psi a, \Psi\rangle, \tag{6.186}$$

for some unit vector $\Psi \in F_-(H)$, yields all pure states on $CAR(\mathbb{C}^N)$.

Now let $N \to \infty$, that is, replace $\mathbb{C}^N$ by the Hilbert space $\ell^2(\mathbb{Z})$ of square-summable sequences of complex numbers. In that case, the algebra $CAR(\ell^2)$ and the fermionic Fock space $F_-(\ell^2)$ are still perfectly well defined, but a Hamiltonian like (6.138) will generally be unbounded. Nonetheless, the derivation

$$\delta(a) := i \lim_{N\to\infty} [h_N, a], \tag{6.187}$$

is well defined, but only on operators a that are finite sums of products of finitely many c's and $c's$. Of course, in this case (1.44) should only hold on such operators.

Exercise 6.13 *Repeat the previous exercise for this situation.*

From the point of view of the self-dual formalism, the above discussion merely covers the case where $h_D^{(1)}$ is already diagonal on $K = H \oplus H$. In that case, the state $|0\rangle$ is characterized by the property $c(f)|0\rangle = 0$ for all $f \in H$. In general, however, we need to deal with the case (6.165). According to the previous two sections (especially (6.184), (6.158), and (6.146), and Exercise 6.9), the ground state is now given by (6.185), where this time the Fock space vacuum state is characterized by $\eta(\hat{f})|0\rangle = 0$ for all $f \in \hat{H}$, with $\eta(\hat{f}) = \int_{-\pi}^{\pi} \frac{dk}{2\pi} \eta_k \hat{f}(k)$ (and hence $\eta^*(\hat{f}) = \int_{-\pi}^{\pi} \frac{dk}{2\pi} \eta_k^* \overline{\hat{f}(k)}$), where

$$\eta_k = u_k \hat{c}_k + v_k \hat{c}_{-k}^*; \qquad \eta_k^* \overline{u}_k \hat{c}_k^* + \overline{v}_k \hat{c}_{-k}; \tag{6.188}$$

$$\hat{c}_k = \sum_{j\in\mathbb{Z}} e^{-ijk} c_j; \qquad \hat{c}_k^* = \sum_{j\in\mathbb{Z}} e^{ijk} c_j^*, \tag{6.189}$$

with u_k, v_k given by (6.178) - (6.179). It follows that $[\eta(\hat{f}), \eta^*(\hat{g})]_+ = \langle \hat{f}, \hat{g}\rangle$ (inner product in $\hat{H}$) and $[\eta^{\pm}(\hat{f}), \eta^{\pm}(\hat{g})]_+ = 0$, so that also

$$\omega_0(\eta(\hat{f})\eta^*(\hat{g})) = \langle 0|\eta(\hat{f})\eta^*(\hat{g})|0\rangle = \langle \hat{f}, \hat{g}\rangle; \tag{6.190}$$

$$\omega_0(\eta^*(\hat{f})\eta(\hat{g})) = \langle 0|\eta^*(\hat{f})\eta(\hat{g})|0\rangle = 0. \tag{6.191}$$

6.8 The GNS-construction

Algebras like $CAR(H)$ and $\otimes^N M_2(\mathbb{C})$, also for $\dim(H) = \infty$ and $N = \infty$, are examples of so-called C^*-algebras. In what follows, we will not give a formal definition of such algebras but just use the facts that one can add and multiply the elements of such algebras (which are to be thought of as operators), and that they have an involution $a \mapsto a^*$, which is an abstract version of the hermitian conjugate for matrices or, more generally, for operators on Hilbert space. In a more mathematical account, the fact that a C^*-algebra forms a Banach space plays a role, which is responsible for the fact that C^*-algebras can always be realized as algebras as *bounded* operators on some Hilbert space. In any case, C^*-algebras make sense without reference to a Hilbert space, although their construction or definition typically starts from some Hilbert space, like the fermionic Fock space $F_-(H)$ for $CAR(H)$, or $\otimes^N \mathbb{C}^2$ for $\otimes^N M_2(\mathbb{C})$ (where $N < \infty$). As will be clear from what follows, in many respects C^*-algebras behave like groups, in that they are defined abstractly, upon which one may look for concrete representations on vectors spaces, notably on Hilbert spaces.

Definition 6.14 *A **representation** of a C^*-algebra A on a Hilbert space H is a linear map $\pi : A \to B(H)$ such that $\pi(ab) = \pi(a)\pi(b)$, and $\pi(a^*) = \pi(a)^*$, $\forall a, b \in A$.*

- *A representation $\pi : A \to B(H)$ is called* **cyclic** *if there is a vector $\Omega \in H$ for which $\overline{\pi(A)\Omega} = H$;*[13] *in other words, each $\Psi \in H$ is the limit of a sequence $\pi(a_n)\Omega$ in H, where $a_n \in A$. In that case, Ω is called a* **cyclic vector** *for π.*

Physically, the idea behind cyclicity would be that each state arises by 'filling up' the ground state Ω with 'excitations' $\pi(a)\Omega$. There is a beautiful connection between cyclic *representations* of A and *states* on A, given by the **GNS-construction**.[14] In quantum physics, this construction provides the bridge between the usual Hilbert space formalism and the abstract C^*-algebraic approach, so it is quite important.

Theorem 6.15 *Let $\omega : A \to \mathbb{C}$ be a state on a C^*-algebra A. There exists a cyclic representation π_ω of A on a Hilbert space H_ω with cyclic unit vector Ω_ω such that*

$$\omega(a) = \langle \Omega_\omega, \pi_\omega(a)\Omega_\omega \rangle \quad \forall a \in A. \tag{6.192}$$

The idea is to construct H_ω from A and subsequently define π_ω by left-multiplication:

1. Define a sesquilinear form $\langle -, - \rangle_0$ on A by

$$\langle a, b \rangle_0 := \omega(a^* b). \tag{6.193}$$

This form almost defines an inner product on A, except that it may not be positive definite (i.e., it might be that $\omega(a^*a) = 0$ for some $a \neq 0$, so that $\langle a, a \rangle_0 = 0$). Hence we remove the null space

$$N_\omega = \{a \in A \,|\, \omega(a^* a) = 0\} \tag{6.194}$$

[13] Here $\overline{\pi(A)\Omega}$ is the closure of the linear span of the subset $\{\pi(a)\Psi, a \in A, \Psi \in H\}$ of H.

[14] Named after three founding fathers of the field: I.M. Gelfand, M. Naimark, and I.E. Segal.

by forming the quotient A/N_ω. The form

$$\langle [a], [b] \rangle := \omega(a^* b) \tag{6.195}$$

on A/N_ω (where $[a]$ is a modulo elements of the null space N_ω) is positive definite by construction and hence defines an inner product $\langle -, - \rangle$ on the complex vector space A/N_ω. The Hilbert space H_ω, then, is the completion of A/N_ω in the corresponding norm.

2. The representation $\pi_\omega(A)$ is initially defined on $A/N_\omega \subset H_\omega$ by

$$\pi_\omega(a)[b] := [ab]. \tag{6.196}$$

It is trivial that π_ω is linear and satisfies $\pi_\omega(ab) = \pi_\omega(a)\pi_\omega(b)$; to prove that $\pi_\omega(a)^* = \pi_\omega(a^*)$, take inner products with vectors $[b]$ and $[c]$ in A/N_ω. The technical point of the proof, which we omit, is that each $\pi_\omega(a)$ is well defined and *bounded* on A/N_ω, so that it may be extended to all of H_ω by continuity.

3. If A has a unit, which is the case in all our examples (but which is not part of the official definition of a C^*-algebra), define $\Omega_\omega = [1]$; then (6.192) follows by a simple computation:

$$\langle \Omega_\omega, \pi_\omega(a)\Omega_\omega \rangle = \langle [1], [a1] \rangle = \omega(1^* a 1) = \omega(a).$$

We will usually drop the index ω and also omit the symbol $\pi \equiv \pi_\omega$, so the the ω-dependence of the entire construction is hidden in the definition of the inner product on H. If we don't bother about the fact that for infinite-dimensionla H, the space $A\Omega$ is just dense in H (and so strictly speaking does not coincide with H), we can write vectors in H as $b\Omega$ for some $b \in A$, so that the representation of A on H is just given by $a(b\Omega) = ab\Omega$, and the inner product on H is simply

$$\langle a\Omega, b\Omega \rangle = \omega(a^* b). \tag{6.197}$$

For example, take $A = M_n(\mathbb{C})$, with a state necessarily of the form

$$\omega(a) = \mathrm{Tr}\,(\rho a), \tag{6.198}$$

for some density matrix ρ (as we have seen). Writing N for N_ω, etc., it follows that

$$N = \{a \in A \mid \mathrm{Tr}\,(\rho a^* a) = 0\}. \tag{6.199}$$

Exercise 6.16 *In this example, assume for simplicity that $\rho = \sum_{i=1}^n \lambda_i |e_i\rangle\langle e_i|$ with respect to the standard basis (e_i) of $\mathbb{C}^n$ (otherwise, change basis), Compute N, A/N, and the representation $\pi(A)$ on A/N in the following special cases:*

1. *$\lambda_j = 1$ for some j, so that $\rho = |e_j\rangle\langle e_j|$ is pure;*

2. *$\lambda_i > 0$ for all i (as in an equilibrium state $\rho = Z^{-1}\exp(-\beta h)$, where $Z = \mathrm{Tr}\,\exp(-\beta h)$ for some Hamiltonian $h : \mathbb{C}^n \to \mathbb{C}^n$).*

6.9 Irreducible representations and pure states

As for group representation, there are natural notions of equivalence and irreducibility for representations of C^*-algebras.

Definition 6.17 *Two representations $\pi_1 : A \to B(H_1)$ and $\pi_2 : A \to B(H_2)$ are called* **equivalent** *($\pi_1 \cong \pi_2$) if there is a unitary operator $u : H_1 \to H_2$ intertwining π_1 and π_2, in the sense that $\pi_2(a) = u\pi_1(a)u^*$ for all $a \in A$.*

Exercise 6.18 *Show that (for fixed A) $\cong$ is an equivalence relation, in that:*

1. *$\pi \cong \pi$ for each representation π;*
2. *If $\pi_1 \cong \pi_2$, then $\pi_2 \cong \pi_1$;*
3. *If $\pi_1 \cong \pi_2$ and $\pi_2 \cong \pi_3$, then $\pi_1 \cong \pi_3$.*

Definition 6.19 *A representation $\pi : A \to B(H)$ is called* **irreducible** *when H has no nontrivial closed subspaces stable under $\pi(A)$. In other words, if $K \subset H$ is a closed subspace such that $\pi(a)\Psi \in K$ for all $a \in A$, $\Psi \in K$, then $K = 0$ or $K = H$.*

If π is *not* irreducible, then we do have such a subspace K, and since $\pi(a^*) = \pi(a)^*$ it follows that also $K^\perp$ is stable under π. Thus we may *reduce* $H = K \oplus K^\perp$, with $\pi(A)K^{(\perp)} \subset K^{(\perp)}$. For example, the defining representation of $A = M_n(\mathbb{C})$ is irreducible, but its restriction to the algebra D_n of diagonal matrices is not: each subspace $\mathbb{C} \cdot e_i$ is stable under D_n.

Exercise 6.20 *Show that if $\pi_1 \cong \pi_2$ and π_1 is irreducible, then so is π_2.*

The following result is analogous to Schur's Lemma in group theory.

Theorem 6.21 *The following conditions on $\pi : A \to B(H)$ are equivalent:*[15]

1. *π is irreducible;*
2. *Each nonzero vector $\Omega \in H$ is cyclic;*
3. *$\pi(A)' = \mathbb{C} \cdot 1$, or, equivalently, $\pi(A)'' = B(H)$.*

To get some idea of the proof, if $\pi(A)' \neq \mathbb{C}\cdot 1$, then by advanced (functional) analysis $\pi(A)'$ contains a nontrivial projection p, and hence the image $K = pH$ of H under p is stable under A. This proves $\neg 3 \Rightarrow \neg 1$ and hence $1 \Rightarrow 3$.

There is a beautiful characterization of irreducibility of GNS-representations, which in the theory of phase transitions will be the main technique (in combination with the previous theorem) for proving that some ground state is pure or mixed.

[15]Here the **commutant** M' of a collection M of bounded operators consists of all bounded operators that commute with all elements of M, and the **bicommutant** M'' is simply the iterated commutant $(M')'$.

Theorem 6.22 *The GNS-representation $\pi_\omega(A)$ is irreducible iff ω is pure.*

Equivalently, ω is mixed iff there exists an operator $b \neq \lambda \cdot 1$ on H_ω that commutes with all operators $\pi_\omega(a)$ on H_ω (informally: $bac\Omega = abc\Omega$ for all $a, c \in A$).

Corollary 6.23 *A state ω on A is pure iff $\pi_\omega(A)' = \mathbb{C} \cdot 1$.*

We just prove the easy direction of the corollary: if $\pi_\omega(A)' \neq \mathbb{C} \cdot 1$, then ω is mixed.

So suppose the commutant $\pi_\omega(A)'$ is nontrivial. Again by functional analysis, it then contains a nontrivial projection $p_+ \in \pi_\omega(A)'$ (so $p_+ \neq 1$ and $p_+ \neq 0$). It then follows that $p_+\Omega_\omega \neq 0$: for if $p_+\Omega_\omega = 0$, then $ap_+\Omega_\omega = p_+a\Omega_\omega = 0$ for all $a \in A$, so that $p_+ = 0$, since π_ω is cyclic. Similarly, $p_-\Omega_\omega \neq 0$ with $p_- = 1 - p_+$, so we may define the unit vectors

$$\Omega_\pm := p_\pm\Omega_\omega / \|p_\pm\Omega_\omega\|, \tag{6.200}$$

and the associated states

$$\omega_\pm(a) := \langle \Omega_\pm, \pi_\omega(a)\Omega_\pm \rangle \tag{6.201}$$

on A. This yields

$$\omega = \lambda\omega_+ + (1 - \lambda)\omega_-, \tag{6.202}$$

with

$$\lambda = \|\Omega_-\|^2. \tag{6.203}$$

Since $\lambda \neq 0, 1$ and $\omega_+ \neq \omega_-$, it follows that ω is mixed. The associated reduction is effected by writing

$$H = H_+ \oplus H_-; \tag{6.204}$$

$$H_\pm = p_\pm H, \tag{6.205}$$

in that A (more precisely, $\pi_\omega(A)$) maps each subspace $H_\pm$ into itself. Q.E.D.

Exercise 6.24 *Continue the previous Exercise 6.4:*

1. *Prove that the state $\rho = |e_j\rangle\langle e_j|$ on $A = M_n(\mathbb{C})$ is pure by computing the commutant of the corresponding (GNS) representation;*
2. *If $\lambda_i > 0$ for all i in $\rho = \sum_{i=1}^n \lambda_i |e_i\rangle\langle e_i|$, prove that ρ is mixed by computing the commutant of the corresponding (GNS) representation.*

Exercise 6.25 *Continue the previous Exercise 6.16: prove that the defining representation of $CAR(H)$ on the fermionic Fock space $F_-(H)$ is irreducible, for any Hilbert space H (and not just for $H = \mathbb{C}^N$).*

6.10 $\mathbb{Z}_2$-actions

The $\pm$ notation above is explained by passing from the projections $p_\pm$ to the operator

$$w = p_+ - p_-, \tag{6.206}$$

so that

$$w^* = w; \;\; w^2 = 1. \tag{6.207}$$

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

$$p_\pm = \tfrac{1}{2}(1 \pm w) \tag{6.208}$$

are projections satisfying $p_+ + p_- = 1$, giving rise to the decomposition (6.205). 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 (6.205) 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 H_- (in that the nontrivial element is represented by *minus* the unit operator).

Thus instead of a projection $p \in \pi_\omega(A)'$, one may equivalently look for an operator $w \in \pi_\omega(A)'$ that satisfies (6.207).

Where group actions (i.e., unitary representations) on Hilbert spaces should be familiar, it may be less familiar (but equally useful) to consider group actions on (C*) algebras. Such actions are defined in terms of so-called automorphisms.

Definition 6.26 *An* automorphism *of A is an invertible linear map $\theta : A \to A$ satisfying $\theta(ab) = \theta(a)\theta(b)$ and $\theta(a^*) = \theta(a)^*$.*

For example, if $u \in A$ is unitary, in that $uu^* = u^*u = 1$, then $\theta(a) = uau^*$ defines an automorphism (check).

Definition 6.27 *A $\mathbb{Z}_2$-action on A is an automorphism $\theta : A \to A$ with $\theta^2 = \mathrm{id}$.*

Similarly to the Hilbert space decomposition (6.204) under a $\mathbb{Z}_2$-action, an algebra A carrying a $\mathbb{Z}_2$-action decomposes as

$$A = A_+ \oplus A_-; \tag{6.209}$$

$$A_\pm = \{a \in A \mid \theta(a) = \pm a\}. \tag{6.210}$$

Here the so-called *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 consists of all bounded operators on some Hilbert space H and $w : H \to H$ is a untitary operator satisfying $w^2 = 1$, then

$$\theta(a) = waw^* \,(= waw) \tag{6.211}$$

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

$$A_\pm = \{a \in A \mid aw \mp wa = 0\}. \tag{6.212}$$

As another example, take the Fermion algebra

$$F = CAR(\ell^2(\mathbb{Z})), \tag{6.213}$$

generated by the operators $c_j^{\pm}$, $j \in \mathbb{Z}$, $c_j^- \equiv c_j$, $c_j^+ \equiv c_j^*$, subject to the usual CAR $[c_i^{\pm}, c_j^{\mp}]_+ = \delta_{ij}$, $[c_i^{\pm}, c_j^{\pm}]_+ = 0$. Here we may define $\theta : F \to F$ by

$$\theta(c_j^{\pm}) = -c_j^{\pm},\ j \in \mathbb{Z}, \tag{6.214}$$

extended to all $a \in F$ by the defining properties of an automorphism (which, for example, imply $\theta(1) = 1$). In this case, F_+ (F_-) is just the linear span of all products of an even (odd) number of $c_j^{\pm}$'s.

There is some sort of a converse to the construction (6.211) of a $\mathbb{Z}_2$-action.

Theorem 6.28 *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 $w^2 = 1$, $w\Omega = \Omega$, and and $w\pi_\omega(a)w^* = \pi_\omega(\theta(a))$ for each $a \in A$.*

Informally: $waw^* = \theta(a)$ on H. The idea of the proof is to define w on vectors of the type $b\Omega \equiv \pi_\omega(b)\Omega_\omega$ (and thence on all vectors in H by continuity) by

$$wb\Omega := \theta(b)\Omega. \tag{6.215}$$

Taking $b = 1$ already gives $w\Omega = \Omega$, and $w^2 = 1$ (and hence invertibility of w, in that $w^{-1} = w$) follows from $\theta^2 = \mathrm{id}$. Finally, unitarity follows from the computation

$$\begin{aligned} \langle wa\Omega, wb\Omega\rangle &= \langle \theta(a)\Omega, \theta(b)\Omega\rangle = \langle \Omega, \theta(a)^*\theta(b)\Omega\rangle = \langle \Omega, \theta(a^*)\theta(b)\Omega\rangle \\ &= \langle \Omega, \theta(a^*b)\Omega\rangle = \omega(\theta(a^*b)) = \omega(a^*b) = \langle a\Omega, b\Omega\rangle. \end{aligned}$$

In this situation, we obtain a decomposition of $H \equiv H_\omega$ according to (6.204), where the projections $p_\pm$ are given by (6.208), so that, equivalently,

$$H_\pm = \{\Psi \in H \mid w\Psi = \pm\Psi\} = \overline{A_\pm\Omega}, \tag{6.216}$$

the bar denoting closure. In terms of the decomposition (6.209), 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 6.22 [3]:

Theorem 6.29 *Suppose A carries a $\mathbb{Z}_2$-action θ as well as a $\mathbb{Z}_2$-invariant state $\omega : A \to \mathbb{C}$. 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 hence ω is pure) iff $\pi_\pm(A_+)$ are inequivalent;*
2. *$\pi_\omega(A)$ is reducible (and hence ω is mixed) iff $\pi_\pm(A_+)$ are equivalent.*

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
$$\theta(a) = \sigma_3 a \sigma_3; \tag{6.217}$$
note that $\sigma_3^2 = 1$ and $\sigma_3^* = \sigma_3$. Then
$$A_+ = \left\{ \begin{pmatrix} z_+ & 0 \\ 0 & z_- \end{pmatrix}, z_\pm \in \mathbb{C} \right\} \equiv D_2; \tag{6.218}$$
$$A_- = \left\{ \begin{pmatrix} 0 & z_1 \\ z_2 & 0 \end{pmatrix}, z_1, z_2 \in \mathbb{C} \right\}, \tag{6.219}$$
where we have introduced the notation D_n for the algebra of diagonal $n \times n$ matrices. Take $\Omega = (1, 0)$, with associated state
$$\omega(a) = \langle \Omega, a\Omega \rangle, \tag{6.220}$$
where $a \in M_2(\mathbb{C})$. It follows from Exercise 6.16 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 (6.220) is $\mathbb{Z}_2$-invariant, and the unitary operator w in Theorem 6.28 is simply $w = \sigma_3$. Hence the decomposition (6.204) of $H = \mathbb{C}^2$ is simply $\mathbb{C}^2 = \mathbb{C} \oplus \mathbb{C}$, i.e.,
$$H_+ = \{(z, 0), z \in \mathbb{C}\}; \tag{6.221}$$
$$H_- = \{(0, z), z \in \mathbb{C}\}. \tag{6.222}$$
Do check that $H_\pm = A_\pm\Omega$. Identifying $H_\pm \cong \mathbb{C}$, this gives the one-dimensional representations $\pi_\pm(D_2)$ as
$$\pi_\pm \begin{pmatrix} z_+ & 0 \\ 0 & z_- \end{pmatrix} = z_\pm, \tag{6.223}$$
which are trivially inequivalent. Hence by Theorem 6.29 the defining representation of $M_2(\mathbb{C})$ on $\mathbb{C}^2$ is irreducible, as it should be.

2. $A = D_2$, with
$$\theta(\mathrm{diag}(z_+, z_-)) = \mathrm{diag}(z_-, z_+), \tag{6.224}$$
where we have denoted the matrix in (6.218) by $\mathrm{diag}(z_+, z_-)$. This time,
$$A_\pm = \{\mathrm{diag}(z, \pm z), z \in \mathbb{C}\}. \tag{6.225}$$
We once again define a $\mathbb{Z}_2$-invariant state ω by (6.220), but this time we take
$$\Omega = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{6.226}$$

Hence

$$H_\pm = \{(z, \pm z), z \in \mathbb{C}\}.. \tag{6.227}$$

We may now identify each $A_\pm$ with $\mathbb{C}$ under the map $A_\pm \to \mathbb{C}$ given by $\mathrm{diag}(z, \pm z) \mapsto z$. 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 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 6.29 the defining representation of D_2 on $\mathbb{C}^2$ is reducible, as it should be: the explicit decomposition of $\mathbb{C}^2$ in D_2-invariant subspaces is just the one (6.221) - (6.222) of the previous example.

Exercise 6.30 *Check all claims in these examples and understand every step.*

The first numbered claim of Theorem 6.29 is relatively easy to prove from Theorem 6.22. 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

$$b = \begin{pmatrix} b_{++} & b_{+-} \\ b_{-+} & b_{--} \end{pmatrix}, \tag{6.228}$$

where the four operators in this matrix act as follows:

$$b_{++} : H_+ \to H_+, b_{+-} : H_- \to H_+,\ b_{-+} : H_+ \to H_-,\ b_{--} : H_- \to H_-. \tag{6.229}$$

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

$$[b_{++}, \pi_+(a)] = 0;\ [b_{--}, \pi_-(a)] = 0; \tag{6.230}$$

$$\pi_+(a)b_{+-} = b_{+-}\pi_-(a);\ \pi_-(a)b_{-+} = b_{-+}\pi_+(a). \tag{6.231}$$

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 6.29, let $\pi_+(A_+) \cong \pi_-(A_+)$, so by definition (of equivalence) there is a unitary operator $v : H_- \to H_+$ such that

$$v\pi_-(a) = \pi_+(a)v, \forall\, a \in A_+. \tag{6.232}$$

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

$$w = \begin{pmatrix} 0 & v \\ v^* & 0 \end{pmatrix}. \tag{6.233}$$

It is easy to verify from (6.232) 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 w this is the same as $[w, \pi(a)] = 0$. So $w \in \pi(A)'$, and hence $\pi(A)$ is reducible by Theorem 6.22. Q.E.D.

6.11 Irreducible representations of the CAR

In determining the ground state(s) of the quantum Ising chain, we will apply Theorem 6.29 to the algebra $F = CAR(H)$, for $H = \ell^2(\mathbb{Z})$. This application relies on some knowledge of the representation theory of F [1, 5]. For the moment we leave the Hilbert space H general, equipped with a conjugation $J : H \to H$ (i.e., J antilinear, $J^2 = 1$, $J^* = J$), with associated Hilbert space $K = H \oplus H$ with conjugation $\Gamma : K \to K$ given by (6.151). An interesting class of pure states arises as follows.

Theorem 6.31 *There is a bijective correspondence between projections $P : K \to K$ that (apart form the defining properties $P^2 = P^* = P$) satisfy*

$$\Gamma P \Gamma = 1 - P, \tag{6.234}$$

and states ω_P on F that satisfy

$$\omega_P(\Phi(h)^*\Phi(h)) \;=\; \langle h, Ph\rangle \;\; \forall\, h \in K. \tag{6.235}$$

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

$$\omega_P(\Phi(h_1)\cdots\Phi(h_{2n+1})) \;=\; 0; \tag{6.236}$$

$$\omega_P(\Phi(h_1)\cdots\Phi(h_{2n})) \;=\; {\sum_{\pi\in S_{2n}}}' \epsilon(\pi) \prod_{j=1}^{n} \langle Ph_{\epsilon(2j)}, \Gamma h_{\epsilon(2j-1)}\rangle, \tag{6.237}$$

where the sum Σ' is over all permutations π of $1, \ldots, 2n$ such that $\pi(2j-1) < \pi(2j)$ and $\pi(1) < \pi(3) < \cdots < \pi(2n-1)$, and $\epsilon(\pi)$ is the sign(ature) of π.

Note that (6.235) is a special case of (6.237) because of (6.152). States like ω_P, which are entirely determined by their two-point functions, are often called *quasi-free.*[16]

Exercise 6.32 *Introduce the projection P_0 onto the first copy of H in K, given by*

$$P_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \tag{6.238}$$

1. *Using* (6.149)*, show that $\omega_{P_0}(a) = \langle 0|a|0\rangle$, where $c(f)|0\rangle = 0$ for all $f \in H$.*
2. *For a unitary operator S on K with $[S, \Gamma] = 0$, show that $\omega_{SP_0S^{-1}}(a) = \langle 0|a|0\rangle$, where this time the 'vacuum' state $|0\rangle$ is characterized by $\eta(f)|0\rangle = 0$ for all $f \in H$, in which $\eta(f)$ is the annihilation operator Bogoliubov-transformed by S, as in Exercise 6.158.*
3. *Take S as in* (6.158)*, where u and v are defined (for given dynamics) as in Exercise 6.11, so that the associated unitary operator U in* (6.182) *diagonalizes the (Fourier-transformed) single-particle Hamiltonian $\hat{h}_D^{(1)}$. Show that $SP_0S^{-1} = E_+$, the projection onto the positive-energy space for $\hat{h}_D^{(1)}$ in K.*
4. *Finally, conclude that ω_{E_+} equals the ground state ω_0 of δ, as in* (6.187) *etc.*

[16] There also exist mixed quasi-free states, in which the role of the projection P is played by some positive operator $0 \leq S \leq 1$. These occur, for example, as equilibrium states.

6.12 Proof of phase transition

To study the phase transition of the quantum Ising chain, we need to take the limit $N \to \infty$. This turns out to be problematic for the one-sides chain $\mathbb{N}$: 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 (6.114) - (6.115), the limiting dynamics has no simple form on either A or F.[17] However, the quantum Ising model on the two-sided chain $\mathbb{Z}$ can be solved [3]. This comes with a price, though: in order to define the Jordan–Wigner transformation (6.114) as an isomorphism between A and F, where

$$A = \otimes_{j\in\mathbb{Z}} M_2(\mathbb{C}); \tag{6.239}$$
$$F = CAR(\ell^2(\mathbb{Z})), \tag{6.240}$$

we would need to start the sums in the rifght-hand side of (6.114) - (6.115) at $j = -\infty$. At first sight this appears to be impossible, though, because operators like $\exp(\pi i \sum_{j=-\infty}^{i-1} \sigma_j^+ \sigma_j^-)$ 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" $\exp(\pi i \sum_{j=-\infty}^{0} \sigma_j^+ \sigma_j^-)$. This formal expression suggests the relations

$$T^2 = 1; \tag{6.241}$$
$$T^* = T; \tag{6.242}$$
$$TaT = \theta_-(a), \tag{6.243}$$

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

$$\theta_-(\sigma_j^\pm) = -\sigma_j^\pm \ (j \le 0); \tag{6.244}$$
$$\theta_-(\sigma_j^\pm) = \sigma_j^\pm \ (j > 0); \tag{6.245}$$
$$\theta_-(\sigma_j^z) = \sigma_j^z \ (j \in \mathbb{Z}); \tag{6.246}$$
$$\theta_-(\sigma_j^0) = \sigma_j^0 \ (j \in \mathbb{Z}), \tag{6.247}$$

where $\sigma^0 = 1_2$. Formally, de define an algebra extension

$$\hat{A} := A \oplus A \cdot T, \tag{6.248}$$

with elements of the type $a + bT$, $a, b \in A$, and algebraic relations given by (6.241) - (6.242). That is, we have

$$(a + bT)^* = a^* + \theta_-(b^*)T; \tag{6.249}$$
$$(a + bT) \cdot (a' + b'T) = aa' + b\theta_-(b') + (ab' + b\theta_-(a'))T. \tag{6.250}$$

[17] The reason for this is that 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 L^2, The energies ϵ_k of the fermionic quasiparticles do not define a multiplication operator on $H^2(-\pi, \pi)$.

The correct version of (6.114) - (6.115) may now be written down as

$$c_i^{\pm} = Te^{\mp\pi i\sum_{j=i}^{0}\sigma_j^+\sigma_j^-}\sigma_i^{\pm} \quad (i<1); \tag{6.251}$$
$$c_1^{\pm} = T\sigma_1^{\pm}; \tag{6.252}$$
$$c_i^{\pm} = Te^{\mp\pi i\sum_{j=1}^{i-1}\sigma_j^+\sigma_j^-}\sigma_i^{\pm} \quad (i>1), \tag{6.253}$$

with formal inverse

$$\sigma_i^{\pm} = Te^{\pm\pi i\sum_{j=i}^{0}c_j^+c_j^-}c_i^{\pm} \quad (i<1); \tag{6.254}$$
$$\sigma_1^{\pm} = Tc_1^{\pm}; \tag{6.255}$$
$$\sigma_i^{\pm} = Te^{\pm\pi i\sum_{j=1}^{i-1}\sigma_j^+\sigma_j^-}\sigma_i^{\pm} \quad (i>1), \tag{6.256}$$

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

$$\hat{F} := F \oplus F\cdot T, \tag{6.257}$$

satisfying the same rules (6.241) - (6.243), but now in terms of a "fermionic" $\mathbb{Z}_2$-action $\theta_j : F \to F$ given by extending the following action on elementary operators:

$$\theta_-(c_j^{\pm}) = -c_j^{\pm}\ (j\le 0); \tag{6.258}$$
$$\theta_-(c_j^{\pm}) = c_j^{\pm}\ (j>0); \tag{6.259}$$
$$\theta_-(c_j^{z}) = c_j^{z}\ (j\in\mathbb{Z}); \tag{6.260}$$
$$\theta_-(c_j^{0}) = c_j^{0}\ (j\in\mathbb{Z}). \tag{6.261}$$

Because of T, the Jordan–Wigner transformation fails to give an isomorphism $A\cong F$, but it does given an isomorphism $\hat{A}\cong\hat{F}$. More importantly, if, having already defined the $\mathbb{Z}_2$-action θ on F by (6.214), we define a similar $\mathbb{Z}_2$-action on A by

$$\theta(\sigma_j^{\pm}) = -\sigma_j^{\pm}\ (j\in\mathbb{Z}); \tag{6.262}$$
$$\theta(\sigma_j^{z}) = \sigma_j^{z}\ (j\in\mathbb{Z}); \tag{6.263}$$
$$\theta(\sigma_j^{0}) = \sigma_j^{0}\ (j\in\mathbb{Z}), \tag{6.264}$$

and grade A and F by the latter, i.e., by θ (see (6.209), and analogously $F = F_+\oplus F_-$), we have isomorphisms

$$A_+ \cong F_+; \tag{6.265}$$
$$A_- \cong F_-T; \tag{6.266}$$
$$A \cong F_+\oplus F_-T. \tag{6.267}$$

For given dynamics (6.159), 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 (6.265) 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 (6.265) 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 6.29 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_+ [3]. 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_P, defined according to Theorem 6.31 by a projection $P : K \to K$ that satisfies (6.234), is then $\mathbb{Z}_2$-invariant iff $[W, P] = 0$.

In our case, this simplifies to $\theta(\Phi(h)) = -\Phi(h)$, so that $W = -1$, and every projection commutes with W. In any case, we have [2, 5]:

Lemma 6.33 *Given some $\mathbb{Z}_2$-action W on K, as well as a projection $P : K \to K$ satisfying* (6.234)*, such that $[W, \Gamma] = [W, P] = 0$:*

1. *the quasi-free state ω_P of Theorem 6.31 is $\mathbb{Z}_2$-invariant (i.e., $\omega_P \circ \theta = \omega_P$);*
2. *the corresponding GNS-representation space $H_P \equiv H_{\omega_P}$ for $F = F_+ \oplus F_-$ decomposes as $H_P = H_P^+ \oplus H_P^-$, with $H_P^\pm = \overline{F_\pm \Omega_P}$. Each subspace $H_P^\pm$ is stable under $\pi_P(F_+)$, and the restriction $\pi_P^\pm$ of $\pi(F_+)$ to $H_P^\pm$ is irreducible.*

Theorem 6.29 then leads to a lemma, which also summarizes the discussion so far.

Lemma 6.34 *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.*

We 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 (6.265) *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 $\pi_+^T \cong \pi_0^+$.*

1. *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.*
2. *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

$$\hat{F} = F \oplus FT = F_+ \oplus F_- \oplus F_+ T \oplus F_- T, \tag{6.268}$$

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

$$\hat{H}_0 = \overline{F_+ \hat{\Omega}_0} \oplus \overline{F_- \hat{\Omega}_0} \oplus \overline{F_+ T \hat{\Omega}_0} \oplus \overline{F_- T \hat{\Omega}_0}, \tag{6.269}$$

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 $\pi_0^{\pm}(F_+)$;
- $\hat{\pi}_-^T(F_+)$ is just our earlier $\pi_-^T(F_+)$;
- $\hat{\pi}_+^T(F_+)$ has not been encountered before.

To understand the latter, we rewrite (6.269) as

$$\hat{H}_0 = H_0 \oplus \hat{H}_0^T; \tag{6.270}$$

$$H_0 = \overline{F_+\hat{\Omega}_0} \oplus \overline{F_-\hat{\Omega}_0} \cong \overline{F_+\Omega_0} \oplus \overline{F_-\Omega_0}; \tag{6.271}$$

$$\hat{H}_0^T = \overline{F_+T\hat{\Omega}_0} \oplus \overline{F_-T\hat{\Omega}_0}, \tag{6.272}$$

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

$$\hat{\omega}_0^T(a) := \langle \hat{\pi}(T)\hat{\Omega}_0, \hat{\pi}(a)\hat{\pi}(T)\hat{\Omega}_0 \rangle_{\hat{H}_0} = \langle \hat{\Omega}_0, \hat{\pi}(\theta_-(a))\hat{\Omega}_0 \rangle_{\hat{H}_0}, \tag{6.273}$$

where the second equality follows from (6.243). Comparing H_0 and $\hat{H}_0$, for all $b \in F$ (and hence especially for $b = \theta_-(a)$) we simply have

$$\langle \hat{\Omega}_0, \hat{\pi}(b)\hat{\Omega}_0 \rangle_{\hat{H}_0} = \hat{\omega}_0(b) = \omega_0^F(b), \tag{6.274}$$

so that $\hat{\omega}_0^T = \omega_0^F \circ \theta_- \equiv \theta_-^* \omega_0^F$. Consequently, if we decompose the GNS-representation space $H_{\theta_-^*\omega_0^F}$ of $\pi_{\theta_-^*\omega_0^F}(F)$ as $H_{\theta_-^*\omega_0^F} = H^+_{\theta_-^*\omega_0^F} \oplus H^-_{\theta_-^*\omega_0^F}$, then $\hat{\pi}_+^T(F_+)$ is (equivalent to) the restriction of $\pi_{\theta_-^*\omega_0^F}(F_+)$ to $H^+_{\theta_-^*\omega_0^F}$. Therefore, the representation $\hat{\pi}(F)$ restricted to $\hat{H}_0^T$ is (equivalent to) the GNS-representation $\pi_{\theta_-^*\omega_0^F}(F)$, so that in turn $\hat{\pi}_+^T(F_+)$ is (equivalent to) $\pi_{\theta_-^*\omega_0^F}(F_+)$, restricted to $H^+_{\theta_-^*\omega_0}$. Hence, further to (6.270) - (6.272),

$$\hat{\pi}(F) \cong \pi_{\omega_0^F}(F) \oplus \pi_{\theta_-^*\omega_0^F}(F). \tag{6.275}$$

The point is that for the quantum Ising chain Hamiltonian (6.166) we have:

Lemma 6.35 *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 the fundamental Theorem 6.37 below. The third follows from Lemma 6.34 and the previous claims.

Exercise 6.36 *Prove the second claim by repeatedly applying Theorem 6.29 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 6.31 and Exercise 6.32, and thence realizing that

$$\pi_{\omega_0^F} = \pi_{E_+}; \tag{6.276}$$

$$\pi_{\theta_-^*\omega_0^F} = \pi_{W_-E_+W_-}. \tag{6.277}$$

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 6.33; 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, we invoke a basic result on the CAR-algebra [1, 2, 5].

Theorem 6.37 *Let P and P' be projections on K that satisfy* (6.234)*. Then:*

1. *$\pi_P(F) \cong \pi_{P'}(F)$ iff $P - P' \in B_2(K)$;*
2. *$\pi_P^+(F_+) \cong \pi_{P'}^+(F_+)$ iff $P - P' \in B_2(K)$ and $\dim(PK \cap (1-P')K)$ is even.*

Here $B_2(K)$ is the class of Hilbert–Schmidt operators on K, which consists of all compact operators a on K for which $\sum_i \mu_i < \infty$, where the μ_i are the eigenvalues of a^*a. 'Morally', the first condition indicates that P and P' should differ only by a finite-dimensional part, but formally this statement is only literally equivalent to the states condition if $[P, P'] = 0$. In any cae, 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 6.34 and 6.35 and Theorem 6.37, we finally obtain:

Theorem 6.38 *The unique $\mathbb{Z}_2$-invariant ground state ω_0 of the Hamiltonian* (6.166) *is pure (and hence forms the unique ground state) iff both of the following hold:*

$$E_+ - W_-E_+W_- \in B_2(K); \tag{6.278}$$

$$\dim(E_+K \cap (1 - W_-E_+W_-)K) \text{ is even.} \tag{6.279}$$

This is true for all λ with $|\lambda| \geq 1$. If , on the other hand, $|\lambda| < 1$, 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 $\omega_0^\pm \circ \theta = \omega_0^\mp$.

The computations establishing this final result are formidable, and we refer to [3].

6.13 Classical Ising model in $d = 2$

6.13.1 Motivation from quantum theory

In general, a *quantum* lattice model in dimension d at inverse temperature $\beta = (k_B T)^{-1}$, defined on $\Lambda \subset \mathbb{Z}^d \subset \mathbb{R}^d$, can be mapped into a *classical* lattice model in dimension $d+1$ on $\Lambda \times \Lambda_0$, where the additional dimension Λ_0 is contained in the interval $[0, \beta]$ (this includes the limit $\beta \to \infty$, in which case thermal averages give ground state averages). This mapping becomes exact in the limit where Λ_0 approaches $[0, \beta]$, in that the lattice spacing in Λ_0 is taken to be a and one takes the double limit $a \to 0$, $\Lambda_0 \to \mathbb{N}$, with $a|\Lambda_0| = \beta$. For example, take the quantum Ising chain (6.124), where $d = 1$ and $\Lambda = \underline{N} = \{1, \ldots, N\}$ i.e.,

$$h_N = -J \sum_{i=1}^{N} \sigma_i^z \sigma_{i+1}^z - \Gamma \sum_{i=1}^{N} \sigma_i^x. \tag{6.280}$$

Here we take periodic boundary conditions, so that $s_{N+1} \equiv s_1$. Using the Trotter product formula for (noncommuting) matrices A and B, i.e.,

$$e^{A+B} = \lim_{L \to \infty} \left(e^{A/L} e^{B/L} \right)^L, \tag{6.281}$$

in terms of $a = \beta/L$ (so that $L = |\Lambda_0|$) we obtain

$$\operatorname{Tr} e^{-\beta h_N} = \lim_{L \to \infty, a \to 0, aL = \beta} \operatorname{Tr} \left(\left(e^{a\Gamma \sum_i \sigma_i^x} e^{aJ \sum_i \sigma_i^z \sigma_{i+1}^z} \right)^L \right), \tag{6.282}$$

where the trace is taken over the Hilbert space $\otimes^N \mathbb{C}^2$. In each copy of $\mathbb{C}^2$ we use basis states $e_+ = (1, 0)$ and $e_- = (0, 1)$, so that we can speak of e_m for $m = \{\pm 1\} \equiv \underline{2}$, with corresponding basis vectors $e_s = \otimes_{i=1}^N e_{s_i}$ of $\otimes^N \mathbb{C}^2$, defined by $s \in \underline{2}^{\underline{N}}$, or $s : \underline{N} \to \underline{2}$, or $s = (s_i)_{i=1}^N$ (cf. §1.2.1). Writing $e_s \equiv |s\rangle = |s_1 \cdots s_N\rangle$, we have

$$\operatorname{Tr}(T^L) = \sum_{s(1) \in \underline{2}^{\underline{N}}, \ldots, s(L) \in \underline{2}^{\underline{N}}} \langle s(1)|T|s(2)\rangle \langle s(2)|T|s(3)\rangle \cdots \langle s(L)|T|s(1)\rangle, \tag{6.283}$$

where each $|s(j)\rangle$ runs over all basis vectors of $\otimes^N \mathbb{C}^2$. Using the identity

$$\langle m | e^{K^* \sigma^x} | m' \rangle = \sqrt{\sinh(K^*) \cosh(K^*)}\, e^{K m m'}, \tag{6.284}$$

where K and K^* (which will be $K^* = ha$) are related by

$$K^* = \operatorname{arctanh}(e^{-2K}) \;\Leftrightarrow\; K = -\tfrac{1}{2} \ln \left(\frac{e^{K^*} - e^{-K^*}}{e^{K^*} + e^{-K^*}} \right), \tag{6.285}$$

the expression (6.282) becomes

$$\operatorname{Tr} e^{-\beta h_N} = \lim_{L \to \infty, a \to 0, aL = \beta} (\sinh(ha) \cosh(ha))^{L/2} \sum_{s \in \underline{2}^{\underline{N} \times \underline{L}}} e^{h_{N,L}^{I_{d=2}}(s)}, \tag{6.286}$$

$$h_{N,L}^{I_{d=2}}(s) = \sum_{i=1}^{N} \sum_{j=1}^{L} \left(K(a) s_{i,j} s_{i,j+1} + J(a) s_{i,j} s_{i+1,j} \right), \tag{6.287}$$

where $K(a)$ is given as K in (6.285) with $K^* = ha$, and $J(a) = Ja$. This is a certain limit of the partition function of the classical Ising model in $d = 2$, with anisotropic couplings (that is, $K(a) \neq J(a)$). This motivates a general study of this model.

6.13.2 Hamiltonian, partition function, and transfer matrix

The variables of the classing Ising model in $d = 2$, defined on a rectangular lattice $\Lambda = \underline{N} \times \underline{L}$, are $s \in \underline{2}^{\underline{N}\times\underline{L}}$, but it is useful to treat each direction differently and regard s as a function $s : \underline{N} \to \underline{2}^{\underline{L}}$. In other words, we use the canonical isomorphism

$$\underline{2}^{\underline{N}\times\underline{L}} \cong \left(\underline{2}^{\underline{L}}\right)^{\underline{N}}, \tag{6.288}$$

given by $s(i,j) \equiv s_{i,j} \leftrightarrow s_i(j)$. So from now on, s will be a function $\underline{N} \to \underline{2}^{\underline{L}}$, and instead of $s(i)$ we write $s_i \in \underline{2}^{\underline{L}}$. In terms of these, the Hamiltonian is given by

$$h_{N,L}^{I_{d=2}}(s) = \sum_{i=1}^{N} \left(H_{N,L}^{(0)}(s_i) + H_{N,L}^{(I)}(s_i, s_{i+1}) \right); \tag{6.289}$$

$$H_{N,L}^{(0)}(s_i) = -J_2 \sum_{j=1}^{L} s_i(j) s_i(j+1); \tag{6.290}$$

$$H_{N,L}^{(I)}(s_i, s_k) = -J_1 \sum_{j=1}^{L} s_i(j) s_k(j). \tag{6.291}$$

Here the idea is that $H_{N,L}^{(0)}$ incorporates the vertical nearest-neighbour interactions, which take place within a single column, whereas $H_{N,L}^{(I)}$ takes care if the horizontal nearest-neighbour interactions, which take place between different columns. In this way, the model already looks like a one-dimensional Ising model, in which, however, the s_i take values in $\underline{2}^{\underline{L}}$ rather than in $\underline{2}$. As a case in point, let us recall that the transfer matrix T of the classical *one*-dimensional Ising model (with periodic boundary conditions), with Hamiltonian

$$h_N^{I_{d=1}} = -J \sum_{i=1}^{N} s_i s_{i+1}, \tag{6.292}$$

is given (in terms of the usual coupling constant $K = \beta J$) by

$$T = \begin{pmatrix} e^K & e^{-K} \\ e^{-K} & e^K \end{pmatrix} \;\Leftrightarrow\; T(s_i, s_k) = e^{K s_i s_k}, \tag{6.293}$$

where we label the matrix entries by $s_i, s_k = (\pm 1, \pm 1)$ instead of $i, j = (1, 2)$. Since

$$e^{-\beta h_N^{I_{d=1}}} = \prod_{i=1}^{N} T(s_i, s_{i+1}), \tag{6.294}$$

the partition function is now given by

$$Z_N \equiv \sum_{s \in \underline{2}^{\underline{N}}} e^{-\beta h_N^{I_{d=1}}} = \mathrm{Tr}\,(T^N), \tag{6.295}$$

so that the free energy density is

$$f_\beta \equiv -\frac{1}{\beta} \lim_{N\to\infty} \frac{1}{N} Z_N = -\frac{1}{\beta} \ln(\lambda^+), \tag{6.296}$$

where λ^+ is the largest eigenvalue of T (which, by the Perron–Frobenius Theorem of linear algebra, is real and nondegenerate; it comes out as $\lambda^+ = e^K + e^{-K}$).

We return to the two-dimensional classical Ising model, and define the $2^L \times 2^L$ transfer matrix by specifying its matrix elements as

$$T_L(s_i, s_k) = \exp -\beta \left(\tfrac{1}{2} H^{(0)}_{N,L}(s_i) + H^{(I)}_{N,L}(s_i, s_k) + \tfrac{1}{2} H^{(0)}_{N,L}(s_k) \right), \tag{6.297}$$

where the labelling is by $s_i, s_k \in \underline{2^L}$ rather than by $i, j \in \{1, \ldots, 2^L\}$. This yields

$$e^{-\beta h_N^{I_{d=2}}} = \prod_{i=1}^{N} T_L(s_i, s_{i+1}); \tag{6.298}$$

$$Z_{N,L} \equiv \sum_{s \in \left(\underline{2^L}\right)^{\underline{N}}} e^{-\beta h_N^{I_{d=2}}} = \mathrm{Tr}\,(T_L^N), \tag{6.299}$$

so that (cf. the one-dimensional case) the free energy density may be computed as

$$f_\beta \equiv -\frac{1}{\beta} \lim_{N,L\to\infty} \frac{1}{N \cdot L} Z_{N,L} = -\frac{1}{\beta} \lim_{L\to\infty} \frac{1}{L} \ln(\lambda_L^+), \tag{6.300}$$

where λ_L^+ is the (necessarily real and nondegenerate) largest eigenvalue of T_L. So we need to compute the spectrum of T_L and select the largest eigenvalue, as $L \to \infty$.

6.13.3 Computing the largest eigenvalue

This computation is originally due to [9] and may also be found in detail in [5, 13]:

1. Use the isomorphism $\mathbb{C}^{2^L} \cong \otimes^L \mathbb{C}^2$ to rewrite T_L as an operator on $\otimes^L \mathbb{C}^2$ (specifically, in terms of the Pauli matrices σ_j^μ ($j = 1, \ldots, L$, $\mu = 1, 2, 3$);
2. Using the isomorphism $\otimes^L \mathbb{C}^2 \cong F_-(\mathbb{C}^L)$, rewrite T_L in terms of the fermionic operators $c_j^\pm$ (using the Jordan–Wigner transformation);
3. Diagonalize the latter using a Fourier- and a Bogoliubov-transformation.

The first step is implemented using $T_L(s_i, s_k) = \langle e_{s_i} | T_L | e_{s_k} \rangle$. Corresponding to the three terms in (6.297), and with some foresight, it is convenient to factorize T_L as

$$T_L = (2 \sinh(2K_1))^{L/2} S_L^{1/2} V_L S_L^{1/2}, \tag{6.301}$$

upon which the first step gives (exercise!)

$$V_L = e^{K_1^* \sum_{j=1}^L \sigma_j^x}; \tag{6.302}$$

$$S_L = e^{K_2 \sum_{j=1}^L \sigma_j^z \sigma_{j+1}^z}. \tag{6.303}$$

Secondly, the Jordan–Wigner transformation yields (exercise!)

$$V_L = e^{-2K_1^* \sum_{j=1}^L (c_j^* c_j - \frac{1}{2})}; \tag{6.304}$$

$$S_L = e^{-K_2 \sum_{j=1}^{L-1} (c_{j+1}^* + c_{j+1})(c_j^* - c_j) - (-1)^N (c_1^* + c_1)(c_L^* - c_L)}, \tag{6.305}$$

where the fermion number operator N is given by

$$N = \sum_{j=1}^{L} c_j^* c_j. \tag{6.306}$$

Defining $W = -N$, we have $W^* = W$ and $W^2 = 1$, so that the fermionic Fock space decomposes as $F_-(\mathbb{C}^L) = F_-(\mathbb{C}^L)^+ \oplus F_-(\mathbb{C}^L)^-$, where $F_-(\mathbb{C}^L)^{+/-}$ has an odd/even number of occupied states. Since $[S_L, W] = 0$, also $S_L = S_L^+ \oplus S_L^-$, where

$$S_L^{\pm} = e^{-K_2 \sum_{j=1}^{L} (c_{j+1}^* + c_{j+1})(c_j^* - c_j)}, \tag{6.307}$$

where $c_{L+1} = \pm c_L$. For the plus sign, the Fourier transform is just given by

$$\hat{c}(k) = \frac{1}{\sqrt{L}} \sum_{j \in \mathbb{Z}_L} e^{-ijk} c_j, ; \tag{6.308}$$

$$c_j = \frac{1}{\sqrt{L}} \sum_{k \in \hat{\mathbb{Z}}_L} e^{ijk} \hat{c}(k) \tag{6.309}$$

where $\mathbb{Z}_L = \{1, \ldots, L\}$ (at least initially, see below) is now seen as an abelian group under addition modulo L, with dual group,[18]

$$\hat{\mathbb{Z}}_L = \{k = 2\pi n/L, n \in \mathbb{Z}_L\}. \tag{6.310}$$

In the limit $L \to \infty$ this yields $k \in [0, 2\pi]$, but (assuming L is even) it is more convenient to shift $\mathbb{Z}_L$ by $-\frac{1}{2}L$, so as to have

$$\mathbb{Z}_L = \{-\tfrac{1}{2}L + 1, \ldots, 0, \ldots, \tfrac{1}{2}L\}. \tag{6.311}$$

This time, k takes the values $\{-\pi + 2\pi/L, \ldots, 0, \ldots, \pi - 2\pi/L, \pi\}$, where each step has size $2\pi/L$, with limit $[-\pi, \pi]$ as $L \to \infty$.

For the minus sign in (6.307), the Fourier transform is defined by the same formulae, but to achieve antiperiodicity $c_{L+1} = -c_L$, the set $\hat{\mathbb{Z}}_L \equiv \hat{\mathbb{Z}}_L^+$ is replaced by

$$\hat{\mathbb{Z}}_L^- = \{k = (2n-1)\pi/L, n \in \mathbb{Z}_L\}, \tag{6.312}$$

so that $k \in \{-\pi + \pi/L, \ldots, \pi - \pi/L\}$, again in steps of $2\pi/L$; note that $k = 0$ does not occur. We now isolate the positive parts $\hat{\mathbb{N}}_L^{\pm} = \hat{\mathbb{Z}}_L^{\pm} \cap \mathbb{N}$ of $\hat{\mathbb{Z}}_L^{\pm}$ (where $\mathbb{N}$ includes 0, so that $0 \in \hat{\mathbb{Z}}_L^+$), and for each k in either $\hat{\mathbb{N}}_L^+ \backslash \{0, \pi\}$ or $\hat{\mathbb{N}}_L^-$ we define

$$S_L(k) = \exp\{2K_2[\cos(k)(\hat{c}^*(k)\hat{c}(k) + \hat{c}^*(-k)\hat{c}(-k) - 1) \tag{6.313}$$

$$+ i \sin(k)(\hat{c}^*(-k)\hat{c}^*(k) + \hat{c}(k)\hat{c}(-k))]\}; \tag{6.314}$$

$$V_L(k) = \exp\{-2K_1^*(\hat{c}^*(k)\hat{c}(k) + \hat{c}^*(-k)\hat{c}(-k) - 1)\}. \tag{6.315}$$

[18]The dual group $\hat{G}$ of a (locally compact) abelian group G consists of the characters on G, i.e., the one-dimensional representations $\chi : G \to \mathbb{T}$, where $\mathbb{T} \subset \mathbb{C}$ is the unit circle. As the name suggest, the dual group is itself a group under $(\chi \cdot \chi')(g) = \chi(g)\chi'(g)$. In our case, $G = \mathbb{Z}_L$ and each $k \in \hat{\mathbb{Z}}_L$ labels a character $\chi_k(j) = \exp(ijk)$. In our case, $\hat{\mathbb{Z}}_L$ is just a reparametrization of $\mathbb{Z}_L$ itself. The Schur orthogonality relations for compact groups then immediately yield useful relations like $(1/L)\sum_{j=1}^{L} \exp(ijk) = \delta_{k,0}$ and $(1/L)\sum_{k=\pi+2\pi/L}^{\pi} \exp(ijk) = \delta_{j,0}$. These, in turn, give the inversion formula (6.309).

For $k = 0$ and $k = \pi$, however, in order to avoid double counting below, the factors 2 above are to be omitted, so that at $k = 0$ and $k = \pi$ we have

$$S_L(k) = \exp\{K_2 \cdot \cos(k)(2\hat{c}^*(k)\hat{c}(k) - 1)\}; \tag{6.316}$$
$$V_L(k) = \exp\{-K_1^*(2\hat{c}^*(k)\hat{c}(k) - 1)\}. \tag{6.317}$$

The point, then, is this: since also $[T_L, W] = 0$, we have $T_L = T_L^+ \oplus T_L^-$, with

$$T_L^\pm = (2\sinh(2K_1))^{L/2} \prod_{k \in \hat{\mathbb{N}}_L^\pm} T_L(k); \tag{6.318}$$
$$T_L(k) = S_L(k)^{1/2} V_L(k) S_L(k)^{1/2}, \tag{6.319}$$

compare (6.301). Within the product in (6.318), all factors $T_L(k)$ mutually commute, so that the set $\sigma(T_L)$ of eigenvalues of T_L also decomposes and then factorizes as

$$\sigma(T_L) = \sigma(T_L^+) \cup \sigma(T_L^-); \tag{6.320}$$
$$\sigma(T_L^\pm) = (2\sinh(2K_1))^{L/2} \prod_{k \in \hat{\mathbb{N}}_L^\pm} \sigma(T_L(k)). \tag{6.321}$$

To compute $\sigma(T_L^\pm)$, the operators (6.319) need to be diagonalized. As we see from (6.316) - (6.317), for fixed k the operator $T_L(k)$ is just a 4×4 matrix for all k except $k = 0, \pi$, for which it is merely 2×2 (namely, each pair of fermion operator $\hat{c}^\pm(k)$ acts on a two-dimensional space, so that operators involving $\hat{c}^\pm(k)$ as well as $\hat{c}^\pm(-k)$ act on a four-dimensional space). A lengthy but straightforward computation, which may be accomplished either through the Bogoliubov-transformation mentioned in step 3 above [9, 5], or directly [13] (using a few tricks), yields

$$\prod_{k \in \hat{\mathbb{N}}_L^{+/-}} \sigma(T_L(k)) = \left\{ \exp - \left(\sum_{k \in \hat{\mathbb{Z}}_L^{+/-}} \varepsilon_k (n_k - \tfrac{1}{2}) \right) \mid n_k \in \{0, 1\}, \sum_{k \in \hat{\mathbb{Z}}_L^{+/-}} n_k \text{ odd/even} \right\}, \tag{6.322}$$

where, for $K_1 = K_2 = K$ (do the general case yourself!), the 'energies' are given by

$$\varepsilon_k = \operatorname{arccosh}(\cosh(2(K^* - K)) + 1 - \cos(k)), \tag{6.323}$$

where the sign ambiguity is resolved by the rules [5, 13]

- $\varepsilon_0 = 2(K^* - K)$ (which may have either sign, determined by the value of K);
- $\varepsilon_k > 0$ for all $k \neq 0$ (so that $\varepsilon_{-k} = \varepsilon_k$).

We have now computed the entire spectrum of T_L, but we are really just interested in the largest eigenvalue λ_L^+ as $L \to \infty$, see (6.300). There are two cases:

1. If $\varepsilon_0 > 0$, i.e., if $K^* > K$, then the supremum $\tilde{\lambda}_L^+$ of the set in (6.322) is obtained at $n_k = 0$ for all k, which implies that $\sum_k n_k = 0$ is even, and hence $\ln(\tilde{\lambda}_L^+) = \frac{1}{2}\sum_{k\in\hat{\mathbb{Z}}_L^-} \varepsilon_k$ comes from $\sigma(T_L^-)$. Then

$$\lim_{L\to\infty} \frac{1}{L} \ln(\tilde{\lambda}_L^+) = \tfrac{1}{2} \lim_{L\to\infty} \frac{1}{L} \sum_{k\in\hat{\mathbb{Z}}_L^-} \varepsilon_k = \int_{-\pi}^{\pi} \frac{dk}{4\pi}\, \varepsilon_k. \tag{6.324}$$

2. If $\varepsilon_0 < 0$, i.e., if $K^* < K$, the supremum lies at $n_0 = 1$, $n_k = 0$ for all $k \neq 0$, in which case $\sum_k n_k = 1$ is odd, and hence $\ln(\tilde{\lambda}_L^+) = \frac{1}{2}(|\varepsilon_0| + \sum_{k\in\hat{\mathbb{Z}}_L^+\setminus\{0\}} \varepsilon_k)$ comes from $\sigma(T_L^+)$. However, because $\lim_{k\to 0} \varepsilon_k = |\varepsilon_0|$ (see [5, 13]), the limiting expression (6.324) is exactly the same for this case.

So in both cases we obtain Onsager's famous expression for the free energy of the classical Ising model in $d = 2$ for $K_1 = K_2$ and vanishing external field, viz.

$$f_\beta(K) = -\frac{1}{2\beta} \left(\ln(\sinh(2K)) + \int_{-\pi}^{\pi} \frac{dk}{2\pi}\, \varepsilon_k(K) \right). \tag{6.325}$$

This gives a phase transition at $K = K^*$, at which point $f_\beta(K)''$ diverges. So the first case $K^* > K$ corresponds to $T > T_c$, with a unique thermal equilibrium state, whereas in the second case $K^* < K$, ot $T < T_c$, there are two thermal equilibrium states (this conclusion follows from a detailed analysis of the type given earlier for the quantum model in $d = 1$, though now classically, see [5]). As a first indication of this phase structure, note that in the first case $\sup \sigma(T_L^-)$ as computed above continues to be above $\sup \sigma(T_L^+)$ even as $L \to \infty$, whereas in the second case these suprema asymptotically coincide. Identifying the limits of the corresponding eigenvectors of T_L with equilibrium states then yields the claim. See [5] for a detailed argument.

Exercise 6.39 *1. Verify all computations in this chapter.*

2. Extend the results to the case $K_1 \neq K_2$.

3. Extend the results so as to include an external field.

References

[1] H. Araki, Bogoliubov automorphisms and Fock representations of canonical anticommutation relations, *Contemporary Mathematics* 62, 23–165 (1987).

[2] H. Araki and D.E. Evans, On a C*-algebra approach to phase transition in the two-dimensional Ising model, *Commun. Math. Phys.* 91, 489–503 (1983).

[3] H. Araki and T. Matsui, Ground states of the XY-model, *Commun. Math. Phys.* 101, 213–245 (1985).

[4] O. Bratteli and D.W. Robinson, *Operator Algebras and Quantum Statistical Mechanics, Vol.* II*: Equilibrium States, Models in Statistical Mechanics* (Springer, 1981).

[5] D.E. Evans and Y. Kawahigashi, *Quantum Symmetries on Operator Algebras* (Oxford University Press, 1998).

[6] A. van Enter, R. Fernandez, A. Sokal, *J. Stat. Phys.* 72, 879–1167 (1993).

[7] H.O. Georgii, *Gibbs Measures and Phase Transitions* (De Gruyter, 1988).

[8] E. Lieb, T. Schultz, D. Mattis, Two soluble models of an antiferromagnetic chain, *Ann. Phys.* 16, 407–466 (1961).

[9] E. Lieb, T. Schultz, D. Mattis, Two-dimensional Ising model as a soluble problem of many fermions, *Rev. Mod. Phys.* 36, 856–871 (1964).

[10] P. Pfeuty, The one-dimensional Ising model with a transverse field, *Ann. Phys.* 47, 79–90 (1970).

[11] G. L. Sewell, *Quantum Theory of Collective Phenomena* (Oxford University Press, 1986).

[12] G. L. Sewell, *Quantum Mechanics and its Emergent Macrophysics* (Princeton University Press, 2002).

[13] B. Simon, *The Statistical Mechanics of Lattice Gases, Vol. I* (Princeton University Press, 1993).