Part I
Background

This part is dedicated to the presentation of the background material that we will use to introduce and connect the developments outlined in the Preface. The reader should beware that the presentation is not necessary standard in all points. Instead it aims to provide a unified perspective on the themes that tie this work together. More broadly I hope that it will serve as an invitation to the subjects that are active in this work and the connections that animate them. Useful resources for further study include [77, 1, 36].

1 Classical statistical mechanics

1.1 Gibbs measures

We consider statistical-mechanical models with discrete sets of sites, indexed by $i = 1, \dots, N$. Each site has a local state space X_i, and for simplicity we assume that $X_i = X$ for all i. Then the global state space is defined by $\mathcal{X} := \bigoplus_i X_i$. For now let us further assume that the local state space X is finite. Then given a Hamiltonian $H : \mathcal{X} \to \mathbb{R}$ and an inverse temperature $\beta \in (0, \infty)$, the primary object of interest is the Gibbs measure defined by the probability mass function

$$\mu_\beta(x) = \frac{1}{Z[\beta]} e^{-\beta H(x)},$$

where $Z[\beta] = \sum_{x \in \mathcal{X}} e^{-\beta H(x)}$ is the partition function, a normalization constant chosen to ensure that the right-hand side indeed defines a probability measure. Notice that in the zero-temperature limit (i.e., as $\beta \to \infty$), the Gibbs measure concentrates around the minimizer(s) of H, hence statistical mechanics at zero temperature recover the general problem of optimization.

More generally, we can lump β into our definition of H and think of $H = H_A$ itself as being parametrized by some data A. Then our Gibbs measure is likewise parametrized by A via

$$\mu_A(x) = \frac{1}{Z[A]} e^{-H_A(x)},$$

where $Z[A] = \sum_{x \in \mathcal{X}} e^{-H_A(x)}$.

A standard example is the ferromagnetic Ising model with external magnetic field, specified by a choice of graph structure for the sites (e.g., a d-dimensional lattice), local state space $X = \{-1, 1\}$, parameters $A = (\beta, \mu)$, and Hamiltonian

$$H_{\beta,\mu} = -\sum_{i \sim j} \sigma_i \sigma_j - \mu \sum_i \sigma_i,$$

where the summation over $i \sim j$ indicates summation over all pairs of indices that our adjacent in the graph. There are many other related models with local state space $X = \{-1, 1\}$, known as classical spin systems.

These considerations carry over naturally to the continuous setting, e.g., $X_i = \mathbb{R}$, which may be called 'Euclidean (alternatively, classical or statistical) lattice field theory.' Here the Gibbs measure is defined

$$d\mu_A(x) = \frac{1}{Z[A]} e^{-H_A(x)}\, dx,$$

where $Z[A] = \int_{\mathbb{R}^N} e^{-H_A(x)}\, dx$.

The major example of interest in this work is specified by taking the parameter A to be a real-symmetric $N \times N$ matrix and

$$H_A(x) = \frac{1}{2} x^T A x + U(x), \tag{1.1}$$

where $U(x)$ is thought of as a fixed 'interaction,' representing a deviation from Gaussianity, which on its own is trivial to understand. Of particular interest is the interaction form

$$U(x) = \sum_{ij} v_{ij} x_i^2 x_j^2,$$

which we call the *generalized Coulomb interaction* via its formal analogy to the Coulomb interaction of electronic structure, which is reflected in an analogy at the level of Feynman diagrams. Note that this class of models includes as a special case the lattice ϕ^4 model, specified by a diagonal kernel $v_{ij} = \lambda \delta_{ij}$.

1.2 Gibbs variational principle

The partition function, or equivalently the free energy $\Omega[A] := -\log Z[A]$, naturally encodes a great deal of information, as we shall see in our discussion of the Luttinger-Ward formalism. In the continuous setting, as we shall verify in Part III the free energy satisfies the Gibbs variational principle

$$\Omega[A] = \inf_{\mu} \left[\int H_A(x)\, d\mu(x) - S(\mu) \right],$$

where the infimum is taken over a suitable class of probability measures on $\mathbb{R}^N$ and is in fact attained by the Gibbs measure $\mu = \mu_A$. Here S is the differential entropy, defined

$$S(\mu) = -\int \log \frac{d\mu}{d\lambda}\, d\mu$$

for all μ absolutely continuous with respect to the Lebesgue measure λ (and defined $S(\mu) = -\infty$ for μ otherwise). Note that $\frac{d\mu}{d\lambda}$ is the Radon-Nikodym derivative, i.e., the probability density function for μ.

A suitable analogous variational principle is available in the discrete setting:

$$\Omega[A] = \inf_{\mu} \left[\sum_{x \in \mathcal{X}} H_A(x)\, \mu(x) - S(\mu) \right],$$

where S is the Shannon entropy, defined

$$S(\mu) = -\sum_{x \in \mathcal{X}} \mu(x) \log \mu(x),$$

with the convention $0 \log 0 = 0$.

2 Quantum statistical mechanics

What does it mean to 'quantize' a classical system? There are two aspects of the procedure: (1) upgrading the classical state space to a corresponding quantum state space and (2) choosing a Hamiltonian. The first point is straightforward, though the second point is more subtle. In elementary quantum mechanics, the procedure known as canonical quantization [36] produces a quantum Hamiltonian from the symplectic structure of a classical Hamiltonian dynamical system. However, in general classical statistical mechanics, the Hamiltonian is merely a function on states and there are no accompanying dynamics. In the setting of quantum spin systems, for example, there are many Hamiltonians of interest that are thought of as phenomenological models for interesting physics and are not derived by 'quantizing' classical Hamiltonians.

2.1 Quantum state spaces and Hamiltonians

To illustrate the first point, we describe the relation between classical and quantum spin systems. For each site i, the local quantum state space is given by $Q_i = \mathbb{C}^{X_i}$. Note that with the standard inner product, as a complex Hilbert space $Q_i \simeq L^2(X_i)$. For example, if $X_i = \{-1, 1\}$, then $Q_i \simeq \mathbb{C}^2$; this is the important case of spin-$\frac{1}{2}$. As we shall see later, bosonic systems in second quantization can be understood as spin systems in this sense with classical state space given by the nonnegative integers, i.e., $X_i = \mathbb{N}_0 := \{0, 1, 2, \ldots\}$. Moreover, there is a correspondence between fermionic systems in second quantization and quantum spin-$\frac{1}{2}$ systems via the Jordan-Wigner transformation, but this correspondence is not canonical.

Now the global quantum state space is defined as $\mathcal{Q} := \bigotimes_i Q_i \simeq \mathbb{C}^{\mathcal{X}}$, i.e., in the spin-$\frac{1}{2}$ case, we have $\mathcal{Q} \simeq \bigotimes_i \mathbb{C}^2 \simeq \mathbb{C}^{(2^N)} \simeq L^2(\{-1, 1\}^N)$. Thus each state $|\psi\rangle \in \mathcal{Q}$ can be thought of as a $\mathbb{C}$-valued function $\psi(x) = \psi(x_1, \ldots, x_N)$, i.e., a *wavefunction*. Here $x_i \in \{-1, 1\}$.

In the setting of first quantization, as discussed in section 3 below, we may think of i as an index for our *particles*, each with local classical state space $X_i = \mathbb{R}^d$, where

d is the physical dimension. With N particles, the global classical state space is $\mathcal{X} = \bigoplus_{i=1}^{N} \mathbb{R}^d = (\mathbb{R}^d)^N$, hence (ignoring the spin degree of freedom quantum particles) the global quantum state space is $\mathcal{Q} = L^2((\mathbb{R}^d)^N)$, whose elements are functions of the form $\psi(x) = \psi(x_1, \ldots, x_N)$, where $x_i \in \mathbb{R}^d$ for all i. This is the 'original' wavefunction of elementary quantum mechanics, i.e., the wavefunction appearing in the many-body Schrödinger equation.

Now a Hamiltonian in the quantum setting is a Hermitian operator $\hat{H} : \mathcal{Q} \to \mathcal{Q}$. For future reference, we let the space of Hermitian operators on a vector space V be denoted by $\mathbf{H}(V)$, so $\hat{H} \in \mathbf{H}(\mathcal{Q})$. Since $\mathcal{Q} \simeq \mathbb{C}^{\mathcal{X}}$, $\mathcal{H}(\mathcal{Q})$ may alternatively be thought of as the set of complex Hermitian matrices $(H(x, y)) \in \mathbb{C}^{\mathcal{X} \times \mathcal{X}}$. Note that restriction to diagonal $\hat{H}$ recovers the notion of a classical Hamiltonian $H : \mathcal{X} \to \mathbb{R}$.

We will discuss Hamiltonians in first quantization in section 3 below; here we discuss several examples in the quantum spin-$\frac{1}{2}$ setting. To this end, first recall the Pauli matrices:

$$\sigma^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

which, together with the identity I_2, form a basis for $\mathcal{H}(\mathbb{C}^2)$. Now let $\sigma_i^{x/y/z} \in \mathcal{H}(\bigotimes_i \mathbb{C}^2) \simeq \bigotimes_i \mathcal{H}(\mathbb{C}^2)$ be obtained by tensoring a copy of $\sigma^{x/y/z}$ for the i-th site with the identity I_2 on all the other sites, i.e., in matrix form

$$\sigma^{x/y/z}(x, y) = \sigma^{x/y/z}(x_i, y_i) \prod_{j \neq i} \delta_{x_i, y_i}.$$

(Note: the $x/y/z$ notation for the Pauli matrices is unrelated to the notation x, y for the classical state space elements.)

Given a graph structure on the site indices, we may define two model Hamiltonians of interest—the transverse-field Ising (TFI) Hamiltonian and anti-ferromagnetic Heisenberg (AFH) Hamiltonian—as follows:

$$\hat{H}_{\text{TFI}} = -h \sum_i \sigma_i^x - \sum_{\langle i,j \rangle} \sigma_i^z \sigma_j^z$$

$$\hat{H}_{\text{AFH}} = \sum_{i \sim j} \left[\sigma_i^x \sigma_j^x + \sigma_i^y \sigma_j^y + \sigma_i^z \sigma_j^z \right].$$

In the TFI Hamiltonian, h is a scalar parameter. These Hamiltonians may be used to define quantum statistical-mechanical ensembles as we shall describe presently.

2.2 Quantum Gibbs states

The quantum analog of a probability measure is a density operator, i.e., a positive semidefinite Hermitian operator $\rho : \mathcal{Q} \to \mathcal{Q}$ of unit trace. Let the space of density

operators on $\mathcal{Q}$ be denoted $\mathbf{D}(\mathcal{Q})$, so in fact $\rho \in \mathbf{D}(\mathcal{Q})$. Via diagonalization, a density operator can be thought of as a choice of orthonormal basis, *plus* a probability measure over basis elements. Hence quantum 'probability' can be thought of in fact as a *generalization* of a classical probability on $\mathcal{X}$, which is recovered in the case of a diagonal density operator. Likewise, Hermitian operators $\hat{O} \in \mathbf{H}(\mathcal{Q})$ generalize random variables $\mathcal{X} \to \mathbb{R}$, and the 'quantum expectation' is given by the trace $\mathrm{Tr}[\hat{O}\rho]$. Physically, this value is the expected value of a quantum measurement of the Hermitian operator $\hat{O}$ on a quantum system in state ρ.

A quantum Gibbs state is defined in terms of a Hamiltonian, i.e., an operator $\hat{H} \in \mathbf{H}(\mathcal{Q})$, possibly parametrized as $\hat{H}[A]$. Note that restriction to diagonal $\hat{H}$ recovers the notion of a classical Hamiltonian $H : \mathcal{X} \to \mathbb{R}$. Now the quantum Gibbs operator is defined

$$\rho[A] = \frac{1}{Z[A]} \exp(-\hat{H}[A]),$$

where 'exp' denotes the operator exponential and $Z[A] := \mathrm{Tr}\left[\exp(-\hat{H}[A])\right]$. Likewise we define the free energy $\Omega[A] = -\log Z[A]$, which (as we shall verify in Part V) satisfies the quantum Gibbs variational principle

$$\Omega[A] = \inf_{\rho \in \mathcal{D}(\mathcal{Q})} \left[\mathrm{Tr}\left(\hat{H}[A]\rho\right) - S(\rho)\right],$$

where the infimum is attained by $\rho = \rho[A]$, and S here denotes (with meaning clear from context) the von Neumann entropy

$$S(\rho) = -\mathrm{Tr}[\rho \log \rho].$$

Here 'log' is the operator logarithm. Note that the von Neumann entropy recovers the Shannon entropy in the case of diagonal ρ.

Let us explicitly focus on an inverse temperature parameter β, i.e., define

$$\rho[\beta] = \frac{1}{Z[\beta]} \exp(-\beta \hat{H}),$$

for fixed $\hat{H} \in \mathbf{H}(\mathcal{Q})$. It can be verified by diagonalization that as $\beta \to \infty$, if $\hat{H}$ has a unique (normalized) ground state (i.e., eigenvector with minimal eigenvalue) $|\Phi_0\rangle \in \mathcal{Q}$, then $\rho[\beta] \to |\Phi_0\rangle\langle\Phi_0|$. Hence quantum statistical mechanics at zero temperature recovers the problem of finding the ground state of a quantum many-body Hamiltonian, which is the quantum analog of optimization. Incidentally, any density operator of rank 1 is known as a pure state

3 First quantization and electronic structure

As mentioned earlier in section 2, the quantum state space for an N-particle system in first quantization is $L^2((\mathbb{R}^d)^N) \simeq \bigotimes_{i=1}^N L^2(\mathbb{R}^d)$, where d is the physical dimension in which the particles live and we have ignored the spin degree of freedom for

simplicity. (Note that for, e.g., spin-$\frac{1}{2}$ particles, the relevant Hilbert space is simply $L^2((\mathbb{R}^d)^N; \mathbb{C}^2)$, and our discussion can be extrapolated to this setting with minor modifications.) For notational clarity, here we shall use boldface to indicate elements $\mathbf{x} = (x_1, \ldots, x_N) \in (\mathbb{R}^d)^N$, where the $x_i \in \mathbb{R}^d$.

Consider a classical Hamiltonian dynamical system [36] specified by the Hamiltonian $H = H(\mathbf{x}, \mathbf{p})$ of position-momentum coordinates:

$$H(\mathbf{x}, \mathbf{p}) = \frac{1}{2} \sum_{i=1}^{N} |p_i|^2 + \sum_{i=1}^{N} V^{(1)}(x_i) + \frac{1}{2} \sum_{i \neq j} V^{(2)}(x_i - x_j).$$

Such a Hamiltonian specifies the classical dynamics of N particles that experience the same external potential $V^{(1)}$ and interact via the pairwise potential $V^{(2)}$, as well as kinetic energy (for which the mass of the particles is scaled to unity).

Then canonical quantization (see, e.g., [36] for a discussion of deeper principles underlying this procedure) yields the Hamiltonian $\hat{H}$

$$\hat{H} = -\frac{1}{2} \sum_{i=1}^{N} \Delta_{x_i} + \sum_{i=1}^{N} V^{(1)}(x_i) + \frac{1}{2} \sum_{i \neq j} V^{(2)}(x_i - x_j), \tag{3.1}$$

where $\Delta_{x_i} = \sum_{j=1}^{d} \frac{\partial^2}{\partial (x_i)_j^2}$ is the Laplace operator for the i-th particle slot. (Note carefully that $\sum_i \Delta_{x_i}$ is of course *not* an operator on $L^2((\mathbb{R}^d)^N)$. Hence it may seem more appropriate to look for wavefunctions in $H^1((\mathbb{R}^d)^N)$. But really the L^2 inner product is the inner product we want. This is one motivation for the notion of a 'rigged Hilbert space' [27], but we sweep such analytical difficulties under the rug, as we shall discuss first quantization mostly in passing to second quantization.)

Now a Hamiltonian $\hat{H}$ of the form (3.1) can model the physics of (relatively light) particles in the presence of the ionic potential

$$V^{(1)}(x) = -\sum_{I} \frac{Z_I}{|x - R_I|}$$

induced by fixed (relatively heavy) atomic nuclei, indexed by I, of charges Z_I and positions $R_I \in \mathbb{R}^d$, as well as a pairwise interaction specified by $V^{(2)}$; hence we have implicitly assumed the Born-Oppenheimer approximation [16], in which the positions of the atomic nuclei are fixed for the computation of the quantum state of the remaindeer of the system. Identifying our quantum particles as electrons and specifying a repulsive pairwise Coulomb interaction

$$V^{(2)}(x_1, x_2) = \frac{1}{|x_1 - x_2|},$$

we arrive at the *electronic structure* problem, modulo one important caveat, toward which we now turn.

3.1 Identical particles

The caveat is the notion of *identical particles*, which is active in the case of electrons. To motivate this requires some further background.

Here we follow Dirac's notation, i.e., denoting wavefunctions $\psi(\mathbf{x})$ and their adjoints via $|\psi\rangle$ and $\langle\psi|$, respctively, we say that a wavefunction $|\psi\rangle \in \mathcal{Q}$ is normalized if it satisfies

$$1 = \langle\psi|\psi\rangle = \int_{(\mathbb{R}^d)^N} |\psi(\mathbf{x})|^2 \, d\mathbf{x}.$$

Hence $\rho = |\psi\rangle\langle\psi|$ defines a density operator in the sense of section 2. For a set $S \in (\mathbb{R}^d)^N$, the characteristic function of this set $\chi_S(\mathbf{x})$ also defines a diagonal operator on $\mathcal{Q}$ via pointwise multiplication. Then

$$\mathrm{Tr}[\chi_S \, \rho] = \int_S |\psi(\mathbf{x})|^2 \, d\mathbf{x}.$$

Hence we interpret $|\psi(\mathbf{x})|^2$ as the probability density function for locating our N particles at positions $x_1, \ldots, x_N$, respectively.

For the purpose of this work, it can be taken as a fact of nature (though deeper justification can be made through quantum field theory; see, e.g., [36] that particles of certain species are identical, or indistinguishable, in the sense that

$$|\psi(\mathbf{x})|^2 = |\psi(\sigma \cdot \mathbf{x})|^2,$$

for all permutations $\sigma \in S_N$, where σ acts via $[\sigma \cdot \mathbf{x}]_i = x_{\sigma(i)}$.

Let us examine the consequence of such a condition. For such ψ and fixed $\mathbf{x}$ such that $\psi(\mathbf{x}) \neq 0$, it must be the case that $\psi(\sigma \cdot \mathbf{x}) = u_{\psi,\mathbf{x}}(\sigma)\psi(\mathbf{x})$ for some unique $u_{\psi,\mathbf{x}}(\sigma) \in S^1$, where $S^1 \subset \mathbb{C}$ denotes the unit circle as a subset of the complex plane. This condition defines a map $u_{\psi,\mathbf{x}} : S_N \to S^1$, evidently a group homomorphism. It can be shown that there are only two such homomorphisms: the trivial homomorphism $u_{\psi,\mathbf{x}}(\sigma) = 1$ and the signature homomorphism $u_{\psi,\mathbf{x}}(\sigma) = \mathrm{sgn}(\sigma)$, which returns ± 1 for even/odd permutations, respectively.

Under the reasonable assumption that $u_{\psi,\mathbf{x}}(\sigma)$ should depend continuously on $\mathbf{x}$, we arrive at two possibilities for ψ: either $\psi(\sigma \cdot \mathbf{x}) = \psi(\mathbf{x})$ for all $\mathbf{x}$, i.e., ψ is *symmetric*, or $\psi(\sigma \cdot \mathbf{x}) = \mathrm{sgn}(\sigma)\psi(\mathbf{x})$ for all $\mathbf{x}$, i.e., ψ is *antisymmetric*. The former is the case of bosons and the latter of fermions. The subspace of symmetric functions is denoted $\mathrm{Sym}^N(L^2(\mathbb{R}^d)) \subset \bigotimes_{i=1}^N L^2(\mathbb{R}^d)$, and the subspace of antisymmetric functions by $\Lambda^N(L^2(\mathbb{R}^d)) \subset \bigotimes_{i=1}^N L^2(\mathbb{R}^d)$. These are the quantum state spaces for N-particle systems of bosons, and fermions, respectively.

By contrast, certain quantum (composite) particle such as atomic nuclei can be modeled as 'boltzmannions' (note: the terminological usage is not universal), which are distinguishable and retain the full state space $\bigotimes_{i=1}^N L^2(\mathbb{R}^d)$.

3.2 Bases for boltzmannionic, fermionic, and bosonic state spaces

Given an orthonormal basis $\{\phi_p\}_{p\in\mathcal{B}}$ for $L^2(\mathbb{R}^d)$, one can construct corresponding orthogonal bases for $\bigotimes_{i=1}^N L^2(\mathbb{R}^d)$, $\Lambda^N(L^2(\mathbb{R}^d))$, and $\mathrm{Sym}^N(L^2(\mathbb{R}^d))$. Of course, technically one needs a complete orthonormal *sequence* to exhaust all of $L^2(\mathbb{R}^d)$, but in practice one may also consider truncated bases and the relevant Galerkin projections of operators. Many basis sets adapted for electronic structure have been introduced in the quantum chemisty literature [102]. We will be somewhat casual about this point in the discussion and maintain notation that is agnostic with respect to it. In other words the basis index set $\mathcal{B}$ can be either $\{1,2,3,\ldots\}$ or $\{1,\ldots,M\}$ for some finite M. Let $\mathcal{H}\subset L^2(\mathbb{R}^d)$ be the (completion of the) span of $\{\phi_p\}_{p\in\mathcal{B}}$. This is our single-particle Hilbert space, after possible truncation, and the corresponding boltzmannionic, fermionic, and bosonic strate spaces are denoted $\bigotimes_{i=1}^N \mathcal{H}$, $\Lambda^N(\mathcal{H})$, and $\mathrm{Sym}^N(\mathcal{H})$, respectively.

First, observe or recall that

$$\{\phi_{i_1}\otimes\cdots\otimes\phi_{i_N} \,:\, i_k\in\mathcal{B}\ \forall k=1,\ldots,N\}$$

is the standard induced orthonormal basis for $\bigotimes_{i=1}^N \mathcal{H}$. To construct the other bases, we must first introduce some new notation.

For $f_1,\ldots,f_N\in L^2(\mathbb{R}^d)$, define

$$\bigodot_{i=1}^N f_i := f_1\odot\cdots\odot f_N := \sum_{\sigma\in S_N} f_{\sigma(1)}\otimes\cdots\otimes f_{\sigma(N)} \in \bigodot_{i=1}^N \mathcal{H}$$

and

$$\bigwedge_{i=1}^N f_i := f_1\wedge\cdots\wedge f_N := \sum_{\sigma\in S_N} \mathrm{sgn}(\sigma)\, f_{\sigma(1)}\otimes\cdots\otimes f_{\sigma(N)} \in \bigwedge_{i=1}^N \mathcal{H}.$$

Then it is not hard to see that

$$\{\phi_{i_1}\odot\cdots\odot\phi_{i_N} \,:\, i_1\le i_2\le\cdots\le i_N,\ \ i_k\in\mathcal{B}\ \forall k=1,\ldots,N\}$$

and

$$\{\phi_{i_1}\wedge\cdots\wedge\phi_{i_N} \,:\, i_1< i_2<\cdots< i_N,\ \ i_k\in\mathcal{B}\ \forall k=1,\ldots,N\}$$

form bases for $S^N(\mathcal{H})$ and $\Lambda^N(\mathcal{H})$, respectively. (Note that $\phi_{i_1}\wedge\cdots\wedge\phi_{i_N}\equiv 0$ if $i_k=i_l$ for some $k\neq l$.)

3.3 Feynman path integral for boltzmannions

We will now describe how the Feynman path integral [34] can be used transpose a boltzmannionic quantum-statistical ensemble to the setting of classical Gibbs measures. The discussion of bases in the preceding section 3.2 will not feature here.

Consider a Hamiltonian of the form

$$\hat{H} = -\frac{1}{2}\Delta + V(\mathbf{x})$$

on the boltzmannionic state space $\mathcal{Q} = L^2((\mathbb{R}^d)^N)$, where $\Delta := \sum_{i=1}^N \Delta_{x_i}$. Hence this form recovers (3.1) as a special case. Note that by identifying $(\mathbb{R}^d)^N \simeq \mathbb{R}^K$, where $K = Nd$, then the boltzmannionic many-particle system appears as nothing but a single-particle system in higher dimension. We shall accept this simplification going forward, denoting $\mathbf{x} = (x_1, \ldots, x_K)$, where the $x_k \in \mathbb{R}$.

Consider the position operators $\mathbf{X} = (X_1, \ldots, X_K)$, which are the diagonal multipliers specified by $X_k(\mathbf{x}, \mathbf{x}) = x_k$, and the momentum operators $\mathbf{P} = (P_1, \ldots, P_K)$, defined by $P_k := -i\partial_{x_k}$. Note that $-\Delta = \sum_{k=1}^K P_k^2$, and $V = V(\mathbf{X})$ in the continuous operator calculus. We shall denote by $|\mathbf{x}\rangle$ the state of definite position $\mathbf{x} = (x_1, \ldots, x_K) \in \mathbb{R}^K$, which is the simultaneous eigenstate of the position operators $(X_1, \ldots, X_K)$ with eigenvalues $(x_1, \ldots, x_K)$. Meanwhile, let $|\mathbf{p}\rangle$ denote the state of definite momentum $\mathbf{p} = (p_1, \ldots, p_K)$, which is the simultaneous eigenstate of the position operators $(P_1, \ldots, P_K)$ with eigenvalues $(p_1, \ldots, p_K)$. As (generalized) functions we have $|\mathbf{x}\rangle = \delta_{\mathbf{x}}$ and $|\mathbf{p}\rangle = e^{-i\mathbf{p}\cdot\mathbf{x}}$. Note that $|\mathbf{x}\rangle, |\mathbf{p}\rangle \notin L^2(\mathbb{R}^K)$, and the right technical notion of 'eigenfunction' is a subtle matter. Here the theory of rigged Hilbert spaces [27] can come to the rescue, but our discussion is purely formal, and we shall elide such difficulties.

Now, as the sets $\{|\mathbf{x}\rangle\}$ and $\{|\mathbf{p}\rangle\}$ of eigenfunctions can each be formally viewed as an orthonormal basis, and we have the completeness relations

$$\mathrm{Id}_{L^2(\mathbb{R}^K)} = \int_{\mathbb{R}^K} |\mathbf{x}\rangle\langle\mathbf{x}| \; d\mathbf{x}, \quad \mathrm{Id}_{L^2(\mathbb{R}^K)} = \int_{\mathbb{R}^K} |\mathbf{p}\rangle\langle\mathbf{p}| \; d\mathbf{p},$$

and it follows that

$$\mathrm{Id}_{L^2(\mathbb{R}^K)} = \int_{\mathbb{R}^K \times \mathbb{R}^K} d\mathbf{x}\, d\mathbf{p} \; |\mathbf{x}\rangle\langle\mathbf{x}|\mathbf{p}\rangle\langle\mathbf{p}| = \int_{\mathbb{R}^K \times \mathbb{R}^K} d\mathbf{x}\, d\mathbf{p} \; e^{-i\mathbf{p}\cdot\mathbf{x}} |\mathbf{x}\rangle\langle\mathbf{p}|. \tag{3.2}$$

Now we consider the partition function

$$Z[\beta] = \mathrm{Tr}\left[e^{-\beta\hat{H}}\right],$$

which we shall expand into the path integral via insertion of the completeness relations. To wit

$$\begin{aligned} Z[\beta] &= \int d\mathbf{x}_{(0)} \; \langle \mathbf{x}_{(0)} | e^{-\beta\hat{H}} | \mathbf{x}_{(0)} \rangle \\ &= \int d\mathbf{x}_{(0)}\, d\mathbf{p}_{(0)} \; \langle \mathbf{x}_{(0)} | \mathbf{p}_{(0)} \rangle \langle \mathbf{p}_{(0)} | e^{-\beta\hat{H}} | \mathbf{x}_{(0)} \rangle \\ &= \int d\mathbf{x}_{(0)}\, d\mathbf{p}_{(0)} \; e^{-i\mathbf{p}_{(0)}\cdot\mathbf{x}_{(0)}} \langle \mathbf{p}_{(0)} | e^{-\beta\hat{H}} | \mathbf{x}_{(0)} \rangle \end{aligned}$$

$$
\begin{aligned}
&= \int d\mathbf{x}_{(0)}\, d\mathbf{p}_{(0)}\; e^{-i\mathbf{p}_{(0)}\cdot\mathbf{x}_{(0)}} \langle \mathbf{p}_{(0)}| e^{-\frac{1}{M}\beta\hat{H}} \cdots e^{-\frac{1}{M}\beta\hat{H}} |\mathbf{x}_{(0)}\rangle \\
&= \int \prod_{m=0}^{M-1} d\mathbf{x}_{(m)}\, d\mathbf{p}_{(m)}\; e^{-\sum_{m=0}^{M-1} i\mathbf{p}_{(m)}\cdot\mathbf{x}_{(m)}} \\
&\qquad \langle \mathbf{p}_{(0)}| e^{-\frac{1}{M}\beta\hat{H}} |\mathbf{x}_{(M-1)}\rangle \cdots \langle \mathbf{p}_{(1)}| e^{-\frac{1}{M}\beta\hat{H}} |\mathbf{x}_{(0)}\rangle,
\end{aligned}
$$

where in the last step we have inserted (3.2) between every pair of $e^{-\frac{1}{M}\beta\hat{H}}$ operators. Now for M large, one has formally

$$
e^{-\frac{1}{M}\beta\hat{H}} = e^{\frac{1}{M}\frac{1}{2}\sum_k P_k^2} e^{\frac{1}{M}V(\mathbf{X})} \left[\mathrm{Id} + O(M^{-2})\right],
$$

hence in the large M limit, one can replace

$$
\begin{aligned}
\langle \mathbf{p}_{(m)}| e^{-\frac{1}{M}\beta\hat{H}} |\mathbf{x}_{(m-1)}\rangle &\approx \langle \mathbf{p}_{(m)}|\, e^{-\frac{\beta}{M}\frac{1}{2}\sum_k P_k^2}\; e^{-\frac{\beta}{M}V(\mathbf{X})}\, |\mathbf{x}_{(m-1)}\rangle \\
&= e^{-\frac{\beta}{M}\left(\frac{1}{2}|\mathbf{p}_{(m)}|^2 + V(\mathbf{x}_{(m-1)})\right)} \langle \mathbf{p}_{(m)}|\mathbf{x}_{(m-1)}\rangle \\
&= e^{-\frac{\beta}{M}\left(\frac{1}{2}|\mathbf{p}_{(m)}|^2 + V(\mathbf{x}_{(m-1)})\right)} e^{i\mathbf{p}_{(m)}\cdot\mathbf{x}_{(m-1)}}
\end{aligned}
$$

where we interpret m modulo M, so

$$
\begin{aligned}
Z[\beta] &= \lim_{M\to\infty} \int \prod_{m=0}^{M-1} d\mathbf{x}_{(m)}\, d\mathbf{p}_{(m)}\; e^{-\sum_{m=0}^{M-1} i\mathbf{p}_{(m)}\cdot(\mathbf{x}_{(m)}-\mathbf{x}_{(m-1)}) - \frac{\beta}{M}\sum_{m=0}^{M-1}\left[\frac{1}{2}|\mathbf{p}_{(m)}|^2 + V(\mathbf{x}_{(m)})\right]} \\
&= \int D\mathbf{x}_{\mathrm{per}}(\,\cdot\,)\, D\mathbf{p}(\,\cdot\,)\; e^{-\int_0^\beta \left[\frac{1}{2}|\mathbf{p}(\tau)|^2 + i\mathbf{p}(\tau)\cdot\partial_\tau\mathbf{x}(\tau) + V(\mathbf{x}(\tau))\right] d\tau},
\end{aligned}
$$

where the limit is understood (for now) only formally and $D\mathbf{x}_{\mathrm{per}}(\,\cdot\,)$ is thought of as the infinite-dimensional Lebesgue measure $\prod_{\tau\in[0,\beta)} d\mathbf{x}(\tau)$ on *periodic* paths, i.e., paths satisfying $\mathbf{x}(0) = \mathbf{x}(\beta)$. Meanwhile, $D\mathbf{p}(\,\cdot\,)$ can be understood as the infinite-dimensional Lebesgue measure $\prod_{\tau\in[0,\beta]} d\mathbf{p}(\tau)$, and here via the construction the periodicity requirement is relaxed. We integrate out the $\mathbf{p}(\tau)$ path via an the formula for Gaussian integrals (formally 'extrapolated' to our infinite-dimensional setting):

$$
\int D\mathbf{p}(\,\cdot\,)\; e^{-\frac{1}{2}\int_0^\beta |\mathbf{p}(\tau)|^2\, d\tau - i\int_0^\beta \mathbf{p}(\tau)\cdot\partial_\tau\mathbf{x}(\tau)\, d\tau} = \left[\lim_{M\to\infty} (2\pi)^{M/2}\right] e^{-\frac{1}{2}\int_0^\beta |\partial_\tau\mathbf{x}(\tau)|^2\, d\tau}.
$$

The 'infinite preconstant' $\lim_{M\to\infty}(2\pi)^{M/2}$ can be ignored as a physically insignificant contribution to the partition function, or, as one prefers, it can be formally lumped into the measure $D\mathbf{x}(\,\cdot\,)$, yielding

$$
Z[\beta] = \int_{\mathbf{x}(0)=\mathbf{x}(\beta)} D\mathbf{x}(\,\cdot\,)\, e^{-\frac{1}{2}\int_0^\beta \left[|\partial_\tau\mathbf{x}(\tau)|^2 + V(\mathbf{x}(\tau))\right] d\tau},
$$

where we introduce the notation $\int D\mathbf{x}_{\mathrm{per}}(\,\cdot\,)\cdots = \int_{\mathbf{x}(0)=\mathbf{x}(\beta)} D\mathbf{x}(\,\cdot\,)\cdots$. Notice from the expression that this quantity can be viewed as a partition function for a Gibbs measure on the space of functions $[0,\beta]\to\mathbb{R}$, i.e., as an infinite-dimensional limit of the setting discussed above in section 1. In fact, this notion can be made precise via the Wiener measure; see, e.g., [34].

4 Second quantization

4.1 The Fock space

Second quantization considers an enlargement, called the Fock space, of any individual fermionic/bosonic N-particle state space. Indeed, the fermionic/bosonic Fock space can be defined as the (completion of the) direct sum of *all* fermionic/bosonic N-particle state spaces:

$$\mathcal{F}_{\mathrm{f}} := \overline{\Lambda(\mathcal{H})}, \quad \mathcal{F}_{\mathrm{b}} := \overline{\mathrm{Sym}(\mathcal{H})},$$

where $\Lambda(\mathcal{H}) := \bigoplus_{N=0}^{\infty} \Lambda^N(\mathcal{H})$ and $\mathrm{Sym}(\mathcal{H}) := \bigoplus_{N=0}^{\infty} \mathrm{Sym}^N(\mathcal{H})$ are the exterior and symmetric algebras, respectively. For simplicity we shall further denote $\mathcal{F}_{\mathrm{f}}^{(N)} := \Lambda^N(\mathcal{H})$ and $\mathcal{F}_{\mathrm{b}}^{(N)} := \mathrm{Sym}^N(\mathcal{H})$. Note that we have conflated the notions of the exterior and symmetric algebras (technically defined as quotients of the tensor algebra) with the equivalent notions, respectively, of the antisymmetric and symmetric subspaces of the tensor algebra.

Even if one is only interested in a definite particle number, a simplified picture of fermionic/bosonic quantum mechanics emerges from this transformation. Moreover, the Fock space allows one to consider states of indefinite particle number and to understand the physics of the system in terms of the creation and annihilation of particles. This perspective is fundamental to quantum field theory and can in fact ought to be viewed as more fundamental than the first-quantized perspective; however, as the names suggest, first quantization preceded second quantization historically.

Now bases for $\mathcal{F}_{\mathrm{f}}$ and $\mathcal{F}_{\mathrm{b}}$ are given, respectively, by

$$\{\phi_{i_1} \wedge \cdots \wedge \phi_{i_N} \,:\, i_1 < i_2 < \cdots < i_N,\ N = 0, 1, 2, \ldots\}$$

and

$$\{\phi_{i_1} \odot \cdots \odot \phi_{i_N} \,:\, i_1 \leq i_2 \leq \cdots \leq i_N,\ N = 0, 1, 2, \ldots\}.$$

By convention, $\Lambda^0(\mathcal{H}) = \mathrm{Sym}^0(\mathcal{H}) = \mathbb{C}$, and the basis element in the case $N = 0$ (i.e., the 'empty' wedge product) is called the vacuum state, denoted by $|-\rangle$. Note carefully the distinction between the vacuum state and the zero vector, denoted 0.

For $\mathbf{n} = (n_p)_{p \in \mathcal{B}} \in \{0, 1, 2, \ldots\}^{\mathcal{B}} = (\mathbb{N}_0)^{\mathcal{B}}$, one defines

$$|\mathbf{n}\rangle_{\mathrm{f}} := \bigwedge_p (\phi_p)^{\wedge n_p}, \quad |\mathbf{n}\rangle_{\mathrm{b}} := \bigodot_p (\phi_p)^{\odot n_p},$$

where, e.g., we denote $(\phi_p)^{\odot n_p} = \underbrace{\phi_p \odot \cdots \odot \phi_p}_{n_p \text{ times}}$, and if $n_p = 0$ then the factor is omitted. Then observe that

$$\left\{|\mathbf{n}\rangle_{\mathrm{f}} \,:\, \mathbf{n} \in \{0, 1\}^{\mathcal{B}}\right\}, \quad \left\{|\mathbf{n}\rangle_{\mathrm{b}} \,:\, \mathbf{n} \in (\mathbb{N}_0)^{\mathcal{B}}\right\}$$

are alternative representations of the same bases for $\mathcal{F}_{\mathrm{f}}$ and $\mathcal{F}_{\mathrm{b}}$, respectively. We refer to these bases as the occupation number bases, because for an element $|\mathbf{n}\rangle$, n_p indicates the number of particles occupying the p-th state ϕ_p.

Now any element $|\psi\rangle \in \mathcal{F}_{\mathrm{f/b}}$ can be written $\sum_{p\in\mathcal{B}} \psi(\mathbf{n})\,|\mathbf{n}\rangle$, hence can equivalently be viewed as a function $\mathbf{n} \mapsto \psi(\mathbf{n})$, with $\psi \in L^2\left(\{0,1\}^{\mathcal{B}}\right)$ or $\psi \in L^2\left((\mathbb{N}_0)^{\mathcal{B}}\right)$. Recall that in first quantization, $|\psi(\mathbf{x})|^2$ indicates the likelihood of finding particles at positions $(x_1, \ldots, x_N)$. Second quantization turns this conceit on its head; for second-quantized wavefunction ψ, $|\psi(\mathbf{n})|^2$ indicates the probability of finding, for each $p \in \mathcal{B}$, n_p particles in state ϕ_p. Hence the the basis functions ϕ_p are the 'sites' of our model, as viewed through the lens of statistical mechanics.

Hence $\mathcal{F}_{\mathrm{f}} \simeq (\mathbb{C}^2)^{\mathcal{B}}$ and $\mathcal{F}_{\mathrm{b}} \simeq \left(\mathbb{C}^{\mathbb{N}_0}\right)^{\mathcal{B}}$ via the correspondence(s) $|\mathbf{n}\rangle \leftrightarrow e_{n_1} \otimes e_{n_2} \otimes \cdots$, where the $e_k \in \mathbb{C}^2$ are the (zero-indexed) standard unit vectors. Hence fermionic and bosonic ensembles can be viewed as quantum spin systems in the sense of section 2. In the case of fermions, as we shall see, this isomorphism is not canonical in that it depends on the numerical ordering of the basis functions ϕ_p. We shall examine further difficulties in section 4.4 below.

4.2 The creation and annihilation operators

All operators on the Fock space can be written in terms of the so-called creation and annihilation operators, denoted (respsectively) by $c_p^\dagger$ and c_p in the fermionic case and by $b_p^\dagger$ and b_p in the bosonic case. When the context is clear it is also common to use $a_p^\dagger$ and a_p for either case.

Now we define the fermionic creation operator $c_p^\dagger$ via its action on a basis for $\mathcal{F}_{\mathrm{f}}$:

$$c_p^\dagger\left[\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}\right] = \phi_p \wedge \phi_{i_1} \wedge \cdots \wedge \phi_{i_N}.$$

For the $N = 0$ case we understand this to mean $c_p^\dagger|-\rangle = \phi_p$.

Meanwhile, the annihilation operator c_p can be defined as the formal adjoint of $c_p^\dagger$, and it can be shown without difficulty that

$$c_p\left[\phi_p \wedge \phi_{i_1} \wedge \cdots \wedge \phi_{i_N}\right] = \begin{cases} \phi_{i_1} \wedge \cdots \wedge \phi_{i_N}, & p \neq i_k \ \forall k = 1, \ldots, N \\ 0, & \text{otherwise.} \end{cases}$$

Moreover $c_p|-\rangle = 0$ for all p.

Meanwhile for $i_1, \ldots, i_N \in \mathcal{B}$ and $n_p := |\{k \,:\, i_k = p\}|$, we define

$$b_p^\dagger\left[\phi_{i_1} \odot \cdots \odot \phi_{i_N}\right] = \sqrt{n_p + 1}\ \phi_p \odot \phi_{i_1} \odot \cdots \odot \phi_{i_N}$$

and extend by linearity. The formal adjoint b_p satisfies

$$b_p\left[\phi_p \odot \phi_{i_1} \odot \cdots \odot \phi_{i_N}\right] = \sqrt{n_p + 1}\ \phi_{i_1} \odot \cdots \odot \phi_{i_N},$$

where still $n_p = |\{k \,:\, i_k = p\}|$, and $b_p|-\rangle = 0$ for all p.

It is not hard to verify that

$$\{c_p^\dagger, c_q^\dagger\} = \{c_p, c_q\} = 0, \quad \{c_p, c_q^\dagger\} = \delta_{ij},$$

where $\{\,\cdot\,,\,\cdot\,\}$ denotes the anticommutator, and

$$[b_p^\dagger, b_q^\dagger] = [b_p, b_q] = 0, \quad [b_p, b_q^\dagger] = \delta_{ij},$$

where $[\,\cdot\,,\,\cdot\,]$ denote the commutator. These are the canonical (anti)commutation relations, known for short as the CAR/CCR.

More abstractly, the set of operators $\mathrm{End}(\mathcal{F}_{\mathrm{f/b}})$ can be viewed as the star-algebra $\mathcal{A}_{\mathrm{f/b}}$ generated by the $a_p^\dagger$, subject to the CAR/CCR. (A star-algebra is just an algebra with a star (or adjoint) operation satisfying certain predictable axioms; see, e.g., [18]) Meanwhile, roughly speaking, the Fock space can be thought of abstractly as the orbit of a vacuum state $|-\rangle$ under an action of $\mathcal{A}_{\mathrm{f/b}}$, with equivalences defined via the CCR/CAR and the relations $a_p|-\rangle = 0$.

Indeed, note that in both the fermionic and bosonic cases, we can write

$$|\mathbf{n}\rangle = (a_1^\dagger)^{n_1}(a_2^\dagger)^{n_2}\cdots|-\rangle.$$

Hence all objects of interest to us (i.e., Hamiltonians and wavefunctions) can be understood purely in terms of the algebra $\mathcal{A}_{\mathrm{f/b}}$, together with the vacuums state, satisfying certain algebraic relations.

To conclude this section, we define the number operators $\hat{n}_p = a_p^\dagger a_p$ and the total number operator $\hat{N} = \sum_p \hat{n}_p$. Observe that $|\mathbf{n}\rangle$ is an eigenstate of $\hat{n}_p$ with eigenvalue n_p for all p. Moreover, the N-particle subspaces $\mathcal{F}_{\mathrm{f/b}}^{(N)}$ are precisely the N-eigenspaces of $\hat{N}$.

4.3 Second-quantized operators

Observe that an operator of the form (3.1) has the essential structure

$$\hat{H} = \sum_{k=1}^{N} \hat{O}_k^{(1)} + \sum_{k\neq l}^{N} \hat{O}_{kl}^{(2)},$$

where $\hat{O}_k^{(1)}$ is a one-body operator on $\mathcal{Q}$ obtained by tensoring a copy of some operator $\hat{O}^{(1)}$ on $\mathcal{H}$ for site i with copies of the identity for all other tensor factors $1,\ldots,N$, and $\hat{O}_{kl}^{(2)}$ is a two-body operator on $\mathcal{Q}$ obtained by tensoring a copy of some operator $\hat{O}^{(2)}$ on $\mathcal{H}\otimes\mathcal{H}$ for sites k,l with copies of the identity for all other tensor factor $1,\ldots,N$.

We will show how to write such operators (which preserve the (anti)symmetry of the wavefunction) in terms of the creation and annihilation operators.

Lemma 1. *After restriction to $\mathcal{F}_{\mathrm{f/b}}^{(N)}$, we have*

$$\sum_{k=1}^{N} \hat{O}_k^{(1)} = \sum_{p,q\in\mathcal{B}} O_{pq} a_p^\dagger a_q,$$

where $O_{pq} = \langle \phi_p | \hat{O}^{(1)} | \phi_q \rangle$.

Proof. Though the fermionic and bosonic cases are very similar, it is less confusing to treat them separately. Let us consider the case of fermions first.

We check the claimed operator equality by checking on the basis element $\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}$, and we begin by applying the first-quantized operator $\sum_{k=1}^N \hat{O}_k^{(1)}$ as follows, where for simplicity we write $\hat{O} = \hat{O}^{(1)}$:

$$\begin{aligned}
&\left(\sum_k \hat{O}_k^{(1)}\right) [\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}] \\
&\quad = \sum_k \sum_{\sigma\in S_N} \operatorname{sgn}(\sigma)\, \hat{O}_k^{(1)} \left[\phi_{i_{\sigma(1)}} \otimes \cdots \otimes \phi_{i_{\sigma(N)}}\right] \\
&\quad = \sum_{\sigma\in S_N} \sum_k \operatorname{sgn}(\sigma) \left[\phi_{i_{\sigma(1)}} \otimes \cdots \otimes \underbrace{(\hat{O}\phi_{i_{\sigma(k)}})}_{k\text{-th slot}} \otimes \cdots \otimes \phi_{i_{\sigma(N)}}\right] \\
&\quad \overset{(\star)}{=} \sum_k \sum_{\sigma\in S_N} \operatorname{sgn}(\sigma) \left[\phi_{i_{\sigma(1)}} \otimes \cdots \otimes \underbrace{(\hat{O}\phi_{i_k})}_{[\sigma^{-1}(k)]\text{-th slot}} \otimes \cdots \otimes \phi_{i_{\sigma(m)}}\right] \\
&\quad = \sum_{k=1}^{N} \phi_{i_1} \wedge \cdots \wedge (\hat{O}\phi_{i_k}) \wedge \cdots \wedge \phi_{i_N}.
\end{aligned}$$

In the step $(\star)$ we changed the inner summation variable k according to $k \mapsto \sigma^{-1}(k)$ and then exchanged the sums.

Now we will show that we obtain the same expression by applying $\sum_{p,q\in\mathcal{B}} O_{pq}^{(1)} a_p^\dagger a_q$. First observe that according to the definition of O_{pq}, we have $\hat{O}\phi_q = \sum_p O_{pq}\phi_p$. Then compute:

$$\begin{aligned}
\sum_{pq} O_{pq} a_p^\dagger a_q \,[\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}] &= \sum_p \sum_{k=1}^{N} O_{pi_k} a_p^\dagger a_{i_k} \,[\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}] \\
&= \sum_p \sum_{k=1}^{N} (-1)^{k-1} O_{pi_k} a_p^\dagger \,[\underbrace{\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}}_{\text{factor } k \text{ omitted}}] \\
&= \sum_{k=1}^{N} (-1)^{k-1} \sum_p O_{pi_k} \,[\phi_p \wedge \underbrace{\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}}_{\text{factor } k \text{ omitted}}] \\
&= \sum_{k=1}^{N} (-1)^{k-1} \,[(\hat{O}\phi_{i_k}) \wedge \underbrace{\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}}_{\text{factor } k \text{ omitted}}]
\end{aligned}$$

$$= \sum_{k=1}^{N} \phi_{i_1} \wedge \cdots \wedge (\hat{O}\phi_{i_k}) \wedge \cdots \wedge \phi_{i_N}.$$

This completes the proof for fermions. For bosons, the proof of the identity

$$\left(\sum_k \hat{O}_k^{(1)}\right) [\phi_{i_1} \odot \cdots \odot \phi_{i_N}] = \sum_{k=1}^{N} \phi_{i_1} \odot \cdots \odot (\hat{O}\phi_{i_k}) \odot \cdots \odot \phi_{i_N}$$

is identical to the proof of the first identity above, up to the removal of $\mathrm{sgn}(\sigma)$ from the computation. For the second identity, fix $i_1, \ldots, i_N$ and define $n_q := |\{k : i_k = q\}|$ for all q. Then compute:

$$\begin{aligned}
\sum_{pq} O_{pq} a_p^\dagger a_q \, [\phi_{i_1} \odot \cdots \odot \phi_{i_N}] &= \sum_p \sum_{k=1}^{N} \frac{1}{n_{i_k}} O_{pi_k} a_p^\dagger a_{i_k} \, [\phi_{i_1} \odot \cdots \odot \phi_{i_N}] \\
&= \sum_p \sum_{l=1}^{N} \frac{1}{\sqrt{n_{i_k}}} O_{pi_k} a_p^\dagger \, [\underbrace{\phi_{i_1} \odot \cdots \odot \phi_{i_N}}_{\text{factor } k \text{ omitted}}] \\
&= \sum_{l=1}^{N} \sum_p O_{pi_k} \, [\phi_p \odot \underbrace{\phi_{i_1} \odot \cdots \odot \phi_{i_N}}_{\text{factor } k \text{ omitted}}] \\
&= \sum_{l=1}^{N} (\hat{O}\phi_{i_k}) \odot \underbrace{\phi_{i_1} \odot \cdots \odot \phi_{i_N}}_{\text{factor } k \text{ omitted}} \\
&= \sum_{l=1}^{N} \phi_{i_1} \odot \cdots \odot (\hat{O}\phi_{i_k}) \odot \cdots \odot \phi_{i_N}.
\end{aligned}$$

□

Lemma 2. *After restriction to $\Lambda^N(\mathcal{H})$ or $\mathrm{Sym}^N(\mathcal{H})$ according to the whether the case is fermionic or bosonic, we have*

$$\sum_{k \neq l}^{N} \hat{O}_{kl}^{(2)} = \sum_{p,q,r,s \in \mathcal{B}} O_{pq,rs} a_p^\dagger a_q^\dagger a_s a_r,$$

where $O_{pq,rs} = \langle \phi_p \phi_q | \hat{O}^{(2)} | \phi_r \phi_s \rangle$. (Here, e.g., $|\phi_r \phi_s\rangle$ denotes $\phi_r \otimes \phi_s$.)

Proof. We shall only give the proof for fermions; the bosonic proof follows by making similar changes as those made in Lemma 1 above. First note that we can assume without loss of generality that $\hat{O}^{(2)} = \hat{A} \otimes \hat{B}$, so $O_{pq,rs} = A_{pr} B_{qs}$, where $A_{pr} = \langle \phi_p | \hat{A} | \phi_q \rangle$ and $B_{qs} = \langle \phi_q | \hat{B} | \phi_s \rangle$. (The general case follows from linear combination of such operators.)

As in the proof of Lemma 1, fix a basis element $\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}$ and compute

$$\begin{aligned}
&\sum_{k\neq l}^{N} \hat{O}^{(2)}_{kl} \left[\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}\right] \\
&\quad = \sum_{k} \sum_{\sigma \in S_N} \operatorname{sgn}(\sigma)\, \hat{O}^{(2)}_{kl} \left[\phi_{i_{\sigma(1)}} \otimes \cdots \otimes \phi_{i_{\sigma(N)}}\right] \\
&\quad = \sum_{\sigma \in S_N} \sum_{k\neq l} \operatorname{sgn}(\sigma) \left[\phi_{i_{\sigma(1)}} \otimes \cdots \otimes \underbrace{(\hat{A}\phi_{i_{\sigma(k)}})}_{k\text{-th slot}} \otimes \cdots \otimes \underbrace{(\hat{B}\phi_{i_{\sigma(l)}})}_{l\text{-th slot}} \otimes \cdots \otimes \phi_{i_{\sigma(N)}}\right] \\
&\quad \overset{(\star)}{=} \sum_{k\neq l} \sum_{\sigma \in S_N} \operatorname{sgn}(\sigma) \left[\phi_{i_{\sigma(1)}} \otimes \cdots \otimes \underbrace{(\hat{A}\phi_{i_k})}_{\sigma^{-1}(k)\text{-th slot}} \otimes \cdots \otimes \underbrace{(\hat{B}\phi_{i_l})}_{\sigma^{-1}(l)\text{-th slot}} \otimes \cdots \otimes \phi_{i_{\sigma(N)}}\right] \\
&\quad = \sum_{k\neq l} \phi_{i_1} \wedge \cdots (\hat{A}\phi_{i_k}) \wedge \cdots \wedge (\hat{B}\phi_{i_l}) \wedge \cdots \wedge \phi_{i_N}.
\end{aligned}$$

In the step $(\star)$ we changed the inner summation variable $k \neq l$ according to $(k,l) \mapsto (\sigma^{-1}(k), \sigma^{-1}(l))$ then exchanged the sums.

Finally, we apply the second-quantized operator to the fixed basis element $\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}$. For visual clarity, we first introduce some auxiliary notation. For $k \neq l$, let ε_{kl} be the sign of the permutation that permutes i_k, i_l to the first two spots of $(i_1, \ldots, i_N)$ without changing the rest of the ordering. Hence $\varepsilon_{kl} = (-1)^{k-1}(-1)^{l-1}$ if $k > l$ and $\varepsilon_{kl} = -(-1)^{k-1}(-1)^{l-1}$ if $k < l$. Then compute

$$\begin{aligned}
&\sum_{pqrs} O_{pq,rs} a_p^\dagger a_q^\dagger a_s a_r \left[\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}\right] \\
&\quad = \sum_{pqr} \sum_{k=1}^{N} (-1)^{k-1} O_{pq,i_k s}\, a_p^\dagger a_q^\dagger a_s \left[\underbrace{\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}}_{\text{factor } k \text{ omitted}}\right] \\
&\quad = \sum_{pq} \sum_{k\neq l} \varepsilon_{kl}\, O_{pq,i_k i_l}\, a_p^\dagger a_q^\dagger \left[\underbrace{\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}}_{\text{factors } k,l \text{ omitted}}\right] \\
&\quad = \sum_{pq} \sum_{k\neq l} \varepsilon_{kl}\, O_{pq,i_k i_l} \left[\phi_p \wedge \phi_q \wedge \underbrace{\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}}_{\text{factors } k,l \text{ omitted}}\right] \\
&\quad = \sum_{k\neq l} \varepsilon_{kl} \sum_{pq} A_{pi_k} B_{qi_l} \left[\phi_p \wedge \phi_q \wedge \underbrace{\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}}_{\text{factors } k,l \text{ omitted}}\right] \\
&\quad = \sum_{k\neq l} \varepsilon_{kl} \left[(\hat{A}\phi_{i_k}) \wedge (\hat{B}\phi_{i_l}) \wedge \underbrace{\phi_{i_1} \wedge \cdots \wedge \phi_{i_N}}_{\text{factors } k,l \text{ omitted}}\right] \\
&\quad = \sum_{k\neq l} \phi_{i_1} \wedge \cdots (\hat{A}\phi_{i_k}) \wedge \cdots \wedge (\hat{B}\phi_{i_l}) \wedge \cdots \wedge \phi_{i_N}
\end{aligned}$$

Note that throughout the computation, the sign factor isn't really doing anything but hanging out and waiting to help us anticommute things back to the middle of the wedge product. $\square$

More generally, one may consider m-body operators for $m \leq N$ with notation analogous to the above. Then the following general result should be apparent from the proofs of Lemmas 1 and 2.

Lemma 3. *After restriction to* $\Lambda^N(\mathcal{H})$ *or* $\mathrm{Sym}^N(\mathcal{H})$ *according to the whether the case is fermionic or bosonic, we have*

$$\sum_{k_1,\ldots,k_m \text{ distinct}} \hat{O}^{(m)}_{k_1\cdots k_m} = \sum_{p_1,\ldots,p_m,q_1,\ldots,q_m} O_{p_1\cdots p_m,q_1\cdots q_m}\, a^\dagger_{p_1}\cdots a^\dagger_{p_m} a_{q_m}\cdots a_{q_1},$$

where $O_{p_1\cdots p_m,q_1\cdots q_m} = \langle \phi_{p_1}\cdots\phi_{p_m}|\hat{O}^{(m)}|\phi_{q_1}\cdots\phi_{q_m}\rangle$.

4.4 The Jordan-Wigner transformation

In this section we will first focus on the fermionic case. We have already seen how $\mathcal{F}_\mathrm{f}$ can be put into correspondence with $(\mathbb{C}^2)^{\mathcal{B}}$ via the $|\mathbf{n}\rangle \leftrightarrow e_{n_1}\otimes e_{n_2}\otimes\cdots$, which puts the occupation number basis for $\mathcal{F}_\mathrm{f}$ in correspondence with the standard basis of $(\mathbb{C}^2)^{\mathcal{B}}$. (Recall that here the $e_k \in \mathbb{C}^2$ are the *zero-indexed* standard unit vectors.)

One verifies that under this isomorphism $c^\dagger_p$ (abusing notation slightly by overloading the notation for $c^\dagger_p$) can be written

$$c^\dagger_p = \underbrace{\sigma^z\otimes\cdots\otimes\sigma^z}_{p-1\,\text{factors}} \otimes \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \otimes I_2 \otimes I_2 \otimes \cdots,$$

from which the corresponding formula for c_p is apparent. This transformation from fermionic creation and annihilation operators to quantum spin-$\frac{1}{2}$ operators (or vice versa) is known as the Jordan-Wigner transformation.

Observe that, due to the that a reordering of the index p *does not* commute with the corresponding reording of the tensor factors in $c^\dagger_p$. Hence our representation of fermionic operators in terms of quantum spin operators depends importantly on the choice of ordering of the basis. Moreover, a one-body Hermitian operator such as $a^\dagger_p a_q + a^\dagger_q a_p$, as might appear in a second-quantized Hamiltonian, acts nontrivially on *all spins* between the indices p and q, inclusive. In particular, any physical locality of the fermionic Hamiltonian may be destroyed by Jordan-Wigner transformation.

By contrast the bosonic creation operator can be written as an operator on $(\mathbb{C}^{\mathbb{N}_0})^{\mathcal{B}}$ via

$$b^\dagger_p = \underbrace{\mathrm{Id}\otimes\cdots\otimes\mathrm{Id}}_{p-1\,\text{factors}} \otimes \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & \sqrt{2} & 0 & 0 & \cdots \\ 0 & 0 & \sqrt{3} & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \otimes \mathrm{Id} \otimes \mathrm{Id} \otimes \cdots.$$

It is evidently more natural to view bosonic operators in this way, which is independent of the basis ordering and which preserves physical locality.

4.5 Canonical transformations and noninteracting problems

We now discuss how a change of the orthonormal basis for $\mathcal{H}$ in first quantization may be understood as a transformation (namely, a canonical transformation) of the creation and annihilation operators in second quantization. Consider a basis $\{\tilde{\phi}_p\}$ for $\mathcal{H}$, written in terms of the original basis $\{\phi_p\}$ via

$$\tilde{\phi}_p = \sum_q U_{pq}\phi_q,$$

where $U = (U_{pq})$ is unitary. Consider the fermionic case for concreteness (the bosonic case is almost identical), and recall that for $\Phi \in \mathcal{F}_{\mathrm{f}} = \Lambda(\mathcal{H})$, the action of the creation operator $c_p^\dagger$ is given by

$$c_p^\dagger \Phi = \phi_p \wedge \Phi.$$

Now let $\tilde{c}_p^\dagger$ denote the set of creation operators induced by the basis $\{\tilde{\phi}_p\}$. Then

$$\tilde{c}_p^\dagger \Phi = \tilde{\phi}_p \wedge \Phi = \left(\sum_q U_{pq}\phi_q\right) \wedge \Phi = \sum_q U_{pq} c_p^\dagger \Phi$$

for all Φ, hence $\tilde{c}_p^\dagger = \sum_q U_{pq} c_p^\dagger$. By similar reasoning for bosons and conjugation, we obtain the general formulas

$$\tilde{a}_p^\dagger = \sum_q U_{pq} a_p^\dagger, \quad a_p = \sum_q \overline{U}_{pq} a_q$$

for canonical transformation. The canonical transformation can be thought of in terms of the CCR/CAR alone, without any reference to a first-quantized setting. (Indeed, this is a more fundamental point of view, physically.)

The canonical transformation allows us to completely solve so-called *noninteracting* systems, specified by Hamiltonians of the form

$$\hat{H} = \sum_{pq} A_{pq} a_p^\dagger a_q$$

because after a suitable canonical transformation, we can assume that A is diagonal, i.e., we can assume $\hat{H} = \sum_p u_p \hat{n}_p$, so the states decouple (as the $\hat{n}_p$ commute). Such systems are derived from first-quantized Hamiltonians that lack any many-body terms. In the context of many-body physics, noninteracting systems may be thought of as 'trivial' and can often be viewed as a building block or point of departure for methods designed for many-body systems.

4.6 Second-quantized model Hamiltonians

Second quantization allows us to consider—in addition to Hamiltonians derived from first quantization via the choice of an orbtial basis—various model problems that may capture physical phenomenology of interest.

Of particular note is the fermionic Hubbard model, whose states we enumerate via the orbital-spin index (i, σ), where $i = 1, \ldots, M$, $\sigma = \uparrow, \downarrow$.

$$\hat{H} = -t \sum_{ij\sigma} A_{ij} a^\dagger_{i\sigma} a_{j\sigma} + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow},$$

where A_{ij} is the adjacency matrix of a graph with vertex set $\{1, \ldots, M\}$, e.g., a square lattice.

More generally, one can consider a 'generalized Coulomb model' of the form

$$\hat{H} = \sum_{ij\sigma} h_{ij} a^\dagger_{i\sigma} a_{j\sigma} + \sum_{ij\sigma\tau} U_{ij} \hat{n}_{i\sigma} \hat{n}_{j\tau},$$

which includes in particular the Hubbard model and variants with longer-range interactions. In fact, via certain choices of orbital bases such as the recently introduced Gausslets [105], electronic structure problems in the continuum can be mapped to second-quantized Hamiltonians of this form.

Recall that in electronic structure, the most general Hamiltonian of interest (arising from an arbitrary choice of orbital basis) can be written

$$\hat{H} = \sum_{pq} A_{pq} a^\dagger_p a_q + \sum_{pqrs} U_{pq,rs} a^\dagger_p a^\dagger_q a_s a_r. \tag{4.1}$$

5 Fermionic and bosonic statistical mechanics

In this section we adopt the notation $\zeta = +1, -1$ to indicate the bosonic and fermionic cases, respectively. Moreover, we consider Fock spaces with finitely many states $d = |\mathcal{B}|$. We indicate these parameters in the notation via $\mathcal{F}_{\zeta,d}$. Moreover, we let $\mathcal{F}^{(N)}_{\zeta,d}$ indicate the N-particle sector of the Fock space, i.e., the N-eigenspace of the total number operator $\hat{N}$.

5.1 The zero-temperature ensemble

At zero temperature, typically one first fixes a particle number N, and attention is restricted to the N-particle subspace. Let $\left|\Psi^{(N)}_0\right\rangle \in \mathcal{F}^{(N)}_{\zeta,d}$ be the *N-particle ground state* of $\hat{H}$, i.e., the minimal normalized eigenvector of $\hat{H}$ considered as an operator on the N-particle subspace. The role of the density operator is assumed by the orthogonal projector $\left|\Psi^{(N)}_0\right\rangle\left\langle\Psi^{(N)}_0\right|$ onto the ground state $\left|\Psi^{(N)}_0\right\rangle$, i.e., the statistical

average of a linear operator $\hat{A}$ (with respect to the N-particle *canonical ensemble*) is given by

$$\langle \hat{A} \rangle_N = \big\langle \Psi_0^{(N)} \big| \hat{A} \big| \Psi_0^{(N)} \big\rangle.$$

5.2 The finite-temperature ensemble

At inverse temperature $\beta \in (0, \infty)$, the *partition function* is defined by

$$Z := \mathrm{Tr}[e^{-\beta(\hat{H}-\mu\hat{N})}].$$

where 'Tr' indicates the Fock space trace. Here $\mu \in \mathbb{R}$ is the *chemical potential*, but before commenting on its role, some further elaboration on the trace is owed in the bosonic case, in which the Fock space is infinite-dimensional.

By assumption, $\hat{H}$ conserves particle number, i.e., it maps $\mathcal{F}_{\zeta,d}^{(N)}$ to itself for all N. Thus $e^{-\beta(\hat{H}-\mu\hat{N})}$ does as well and can be viewed as a positive-definite operator on each $\mathcal{F}_{\zeta,d}^{(N)}$. The trace can then be constructed as

$$\mathrm{Tr}[e^{-\beta(\hat{H}-\mu\hat{N})}] = \sum_{N=0}^{\infty} \mathrm{Tr}_N[e^{-\beta(\hat{H}-\mu\hat{N})}] = \sum_{N=0}^{\infty} e^{\beta\mu N} \mathrm{Tr}_N[e^{-\beta\hat{H}}],$$

where 'Tr_N' indicates the trace on the N-particle subspace. Since each of the summands is positive, $\mathrm{Tr}[e^{-\beta(\hat{H}-\mu\hat{N})}] \in (0, +\infty]$ is well-defined.

More generally, the trace is defined for all operators in the trace class of $\mathcal{F}_{\zeta,d}$, i.e., the set of bounded linear operators $\hat{O} : \mathcal{F}_{\zeta,d} \to \mathcal{F}_{\zeta,d}$ for which

$$\sum_{\mathbf{n} \in \mathcal{N}_\zeta^d} \langle \mathbf{n} | \, (\hat{O}^\dagger \hat{O})^{1/2} \, | \mathbf{n} \rangle < +\infty,$$

in which case

$$\mathrm{Tr}[\hat{O}] = \sum_{\mathbf{n} \in \mathcal{N}_\zeta^d} \langle \mathbf{n} | \hat{O} | \mathbf{n} \rangle.$$

See, e.g., [89] for further details on trace class operators.

Now since the partition function can be viewed as a normalization factor, the scenario $Z = +\infty$ is to be avoided. It is now that we turn to the chemical potential. We can view Z as defined above as a function of μ. Evidently $\mu \mapsto Z(\mu)$ is non-decreasing.

First we want to rule out the case that $Z \equiv +\infty$. Unfortunately, this case cannot be ruled out without further assumptions! To see why, suppose that $d = 1$ (so write $a = a_1$), and let $\hat{H} = -a^\dagger a - a^\dagger a^\dagger a a = -a^\dagger a a^\dagger a = -\hat{N}^2$. Then

$$\mathrm{Tr}[e^{-\beta(\hat{H}-\mu\hat{N})}] = \sum_{N=0}^{\infty} e^{\beta(N^2+\mu N)} \mathrm{Tr}_N \left[\mathrm{Id}_{\mathcal{F}_{\zeta,d}^{(N)}} \right] = \sum_{N=0}^{\infty} e^{\beta(N^2+\mu N)} \binom{N+d-1}{d-1} = +\infty,$$

for *all* $\mu \in \mathbb{R}$.

We conclude that such a choice of $\hat{H}$ is *unphysical*, and to rule out such pathologies, we adopt the following:

Assumption 4. *We assume, in the case of bosons, that there exist some positive integer N_0 and some $\mu \in \mathbb{R}$ such that $\hat{H} - \mu\hat{N} \succeq 0$ as an operator on all N-particle subspaces for $N \geq N_0$. (It is equivalent to require that there exist N_0, μ such that $\hat{U} - \mu\hat{N} \succeq 0$ on all N-particle subspaces for $N \geq N_0$.)*

This condition is satisfied in particular if $\hat{U}$ is a two-body interaction as in (4.1) such that $\tilde{U}_{ik,jl} := U_{kj,il}$, interpreted as a $d^2 \times d^2$ matrix, is positive semidefinite. Indeed, in this case, $\hat{U}$ is equal to (up to a single-body term)

$$\frac{1}{2}\sum_{ijkl} \tilde{U}_{ik,jl} \left[a_i^\dagger a_k\right]^\dagger \left[a_j^\dagger a_l\right] \succeq 0.$$

If the $U_{ij,kl}$ are derived from a real-space two-body potential v that is a positive semidefinite function (i.e., has nonnegative Fourier transform), then it follows from Lemma 2 that the matrix $(\tilde{U}_{ik,jl})$ is positive definite as desired. Note that it is possible for v to be positive definite but take negative values at long ranges, i.e., v can act attractively at long range.

Now that we have argued that Assumption 4 is natural, let us see how it guarantees that the set $\operatorname{dom} Z := \{\mu : Z(\mu) < +\infty\}$ is non-empty. Indeed, choose μ' negative enough such that $\hat{H} - \mu'\hat{N} \succeq 0$ as an operator on all N-particle subspaces, and let $\mu = \mu' - \delta$, where $\delta > 0$. Then

$$\operatorname{Tr}[e^{-\beta(\hat{H}-\mu\hat{N})}] \leq \sum_{N=0}^{\infty} \operatorname{Tr}_N[e^{-\beta\delta\hat{N}}] = \sum_{N=0}^{\infty} e^{-\beta\delta N}\binom{N+d-1}{d-1}.$$

Now the binomial coefficient in the last expression is $O(N^{d-1})$ as $N \to +\infty$, so the sum converges.

We will always assume in the finite-temperature setting that $\mu \in \operatorname{int}\operatorname{dom} Z$. It can be shown that if $\hat{U} = 0$, then $\operatorname{dom} Z = \{\mu : h \succ \mu\, I_d\}$. Moreover, if there exist $N_0, \delta > 0$ such that $\hat{U} \succeq \delta\hat{N}^2$ on all N-particle subspaces for $N \geq N_0$ (which holds in particular if $\hat{U}$ is is a two-body interaction as in (4.1) where the $d^2 \times d^2$ matrix $\tilde{U}_{ki,jl} := U_{ij,kl}$ is *positive definite*), then $\operatorname{dom} Z = \mathbb{R}$.

Notice that if $\operatorname{dom} Z$ is open, then since Z is increasing we can write $\operatorname{dom} Z = (-\infty, \mu_c)$ for some μ_c possibly infinite. If $\mu_c < +\infty$, then by Fatou's lemma we have that $\liminf_{\mu\to\mu_c^-} Z(\mu) \geq Z(\mu_c) = +\infty$, so $Z(\mu) \to +\infty$ as $\mu \to \mu_c^-$. (And in any case it follows from the definition of Z that $Z(\mu) \to +\infty$ as $\mu \to +\infty$, so we can write more compactly that $Z(\mu) \to +\infty$ as $\mu \to \mu_c$, no matter whether μ_c is finite or infinite.)

The *grand canonical ensemble* is defined by the *density operator*

$$\rho := Z^{-1}e^{-\beta(\hat{H}-\mu\hat{N})},$$

and the statistical average of an operator $\hat{A}$ with respect to the grand canonical ensemble is denoted

$$\langle \hat{A} \rangle_{\beta,\mu} = \operatorname{Tr}[\hat{A}\rho]$$

whenever $\hat{A}\rho$ is in the trace class. Note that if $\hat{A}$ conserves particle number then

$$\operatorname{Tr}[\hat{A}\rho] = \sum_{N=0}^{\infty} \operatorname{Tr}_N[\hat{A}\rho] = Z^{-1} \sum_{N=0}^{\infty} e^{\beta\mu N} \operatorname{Tr}_N[\hat{A}e^{-\beta\hat{H}}]$$

is defined under the condition that the sum is absolutely convergent, which holds in particular if the norm of $\hat{A}$ as an operator on the N-particle subspace grows only polynomially with N, via the assumption that $\mu \in \operatorname{int}\operatorname{dom} Z$. When the context is clear we simply write $\langle \cdot \rangle$.

Of particular interest is the *expected particle number*

$$\langle \hat{N} \rangle = \frac{\sum_{N=0}^{\infty} N e^{\beta\mu N} \operatorname{Tr}_N[e^{-\beta\hat{H}}]}{\sum_{N=0}^{\infty} e^{\beta\mu N} \operatorname{Tr}_N[e^{-\beta\hat{H}}]}.$$

Observe that $\langle \hat{N} \rangle_{\beta,\mu} \to 0$ as $\mu \to -\infty$. Also note that if $\operatorname{dom} Z = \mathbb{R}$, then $\langle \hat{N} \rangle_{\beta,\mu} \to +\infty$. Defining the *free energy* $\Omega(\mu) := \beta^{-1} \log Z(\mu)$, we see that $\langle \hat{N} \rangle_{\beta,\mu} = \Omega'(\mu)$.

It is not hard to check that Ω is (strictly) convex, i.e., $\langle \hat{N} \rangle_{\beta,\mu}$ is increasing in μ for $\mu \in \operatorname{int}\operatorname{dom} Z$. Recall that if $\operatorname{dom} Z = (0, \mu_c)$, then $Z(\mu) \to +\infty$ as $\mu \to \mu_c$, hence $\Omega(\mu) \to +\infty$ as $\mu \to \mu_c$. If $\mu_c < +\infty$, it follows that $\Omega'(\mu) \to +\infty$ as $\mu \to \mu_c^-$. (Otherwise, since Ω' is increasing, it approaches a finite limit $\mu \to \mu_c^-$. But in this case it would follow from the fundamental theorem of calculus that Ω approaches a finite limit as well: contradiction.) In summary we have established that if $\operatorname{dom} Z$ is open, then $Z(\mu) \to +\infty$ as $\mu \to \mu_c$, no matter whether μ_c is finite or infinite. It follows that in this case $\mu \mapsto \langle \hat{N}_{\beta,\mu} \rangle$ is a bijection from $\operatorname{dom} Z = (-\infty, \mu_c)$ to $(0, +\infty)$. Thus one can *select* the chemical potential μ to yield a predetermined expected particle number.

6 The coherent state path integral

There is a path integral expansion of the partition function in second quantization that is similar in spirit to the original Feynman path integral of section 3.3. It is simplest to treat the bosonic case first because the fermionic path integral formalism requires the introduction of new abstractions.

6.1 The bosonic coherent state path integral

We let $\mathbf{b} = (b_p)_{p \in \mathcal{B}}$ denote the vector of annihilation operators and likewise use bold notation throughout to denote vectors indexed by the state index set $\mathcal{B}$. (Note: we

also retain some bold notation from section 3.3.) We say that a Hamiltonian $\hat{H}$ is written in normal-ordered form if it is a polynomial $\hat{H} = H(\mathbf{b}^\dagger, \mathbf{b})$ of the creation and annihilation operators such that, for each monomial term, all creation operators appear to the left of all annihilation operators, e.g., $\sum_{ij} A_{ij} b_i^\dagger b_j$. Without loss of generality we shall assume that our second-quantized Hamiltonian $\hat{H}$ is of this form. For an operator $\hat{O}$, viewed symbolically as a polynomial of creation and annihilation operators, we write $:\hat{O}:$ for the normal-ordered symbolic operator obtained by formally commuting creation and annihilation operators, e.g., $:bb^\dagger: = b^\dagger b \neq bb^\dagger$.

For inspiration we recapitulate the essential points of the derivation of the Feynman path integral. Recall that we considered a Hamiltonian of the form $\hat{H} = H_{\text{kin}}(\mathbf{P}) + V(\mathbf{X})$, where $\mathbf{P}$ and $\mathbf{X}$ were the momentum and position operators, respectively. Then for momentum and position eigenstates $|\mathbf{p}\rangle$ and $|\mathbf{x}\rangle$ we have

$$\langle \mathbf{p}|\hat{H}|\mathbf{x}\rangle = \left(\overline{H_{\text{kin}}(\mathbf{p})} + V(\mathbf{x})\right) \langle \mathbf{p}|\mathbf{x}\rangle.$$

Using this observation, together with resolutions of the identity in terms of the momentum and position eigenstates, we derived the path integral.

Loosely following such a recipe, we are inspired to consider *eigenstates of the annihilation operators*, which will be known as the coherent states. Let $|\mathbf{z}\rangle$ be such an eigenstate for $\mathbf{z} = (z_p)_{p\in\mathcal{B}} \in \mathbb{C}^{\mathcal{B}}$, satisfying $b_p|\mathbf{z}\rangle = z_p|\mathbf{z}\rangle$. Then for normal-ordered $\hat{H} = H(\mathbf{b}^\dagger, \mathbf{b})$, we have

$$\langle w|\hat{H}|z\rangle = H(\overline{w}, z)\langle w|z\rangle.$$

Then if we can construct coherent states and determine a resolution of the identity in terms of them, we will be in good shape.

6.1.1 Bosonic coherent states

Consider the case of a bosonic system of a single state, i.e., $|\mathcal{B}| = 1$. Then we want to find

$$|z\rangle = \sum_{n=0}^{\infty} \lambda_n |n\rangle$$

such that $b|z\rangle = z|z\rangle$, i.e.,

$$\sum_{n=0}^{\infty} z\lambda_n |n\rangle = \sum_{n=1}^{\infty} \lambda_n \sqrt{n}|n-1\rangle = \sum_{n=0}^{\infty} \lambda_{n+1}\sqrt{n+1}|n\rangle.$$

By equating corresponding terms we conclude that we must have

$$\lambda_{n+1} = \frac{z}{\sqrt{n+1}}\lambda_n,$$

and choosing $\lambda_0 = 1$, we obtain $\lambda_n = \frac{z^n}{\sqrt{n!}}$. Therefore

$$|z\rangle = \sum_{n=0}^{\infty} \frac{z^n}{\sqrt{n!}}|n\rangle = \sum_{n=0}^{\infty} \frac{z^n}{n!}(b^\dagger)^n|-\rangle,$$

or, more succinctly,

$$|z\rangle = e^{zb^\dagger}|-\rangle.$$

More generally, we may derive

$$|\mathbf{z}\rangle = e^{\mathbf{z}\cdot\mathbf{b}^\dagger}|-\rangle = \sum_{\mathbf{n}\in(\mathbb{N}_0)^{\mathcal{B}}} \frac{\mathbf{z}^{\mathbf{n}}}{\sqrt{\mathbf{n}!}}|\mathbf{n}\rangle,$$

where we interpret $\mathbf{z}^{\mathbf{n}} = \prod_p z_p^{n_p}$ and $\mathbf{n}! = \prod_p n_p!$. This is the general formula for the bosonic coherent state, indexed by $\mathbf{z} \in \mathbb{C}^{\mathcal{B}}$.

Now for $\mathbf{w}, \mathbf{z} \in \mathbb{C}$,

$$\langle \mathbf{w}|\mathbf{z}\rangle = \sum_{\mathbf{n}} \frac{\overline{\mathbf{w}}^{\mathbf{n}}\mathbf{z}^n}{\mathbf{n}!} = \prod_{p\in\mathcal{B}}\sum_{n=0}^{\infty} \frac{(\overline{w_p}z_p)^n}{n!} = \prod_{p\in\mathcal{B}} e^{w_p z_p} = e^{\mathbf{w}^*\mathbf{z}}.$$

In particular, the coherent states are *not* orthogonal.

Nonetheless, we can write a resolution of the identity in terms of the coherent states. To wit, we have in the case $|\mathcal{B}| = 1$ that

$$\mathrm{Id} = \int \frac{d\overline{z}\,dz}{2\pi i}\; e^{-|z|^2}|z\rangle\langle z|, \tag{6.1}$$

as we shall confirm below. We write the integration in this particular way to lay some conceptual groundwork for the fermionic case and to clarify certain analogies. For the reader unfamiliar with such notation, we shall record presently the relevant details.

6.1.2 Complex coordinates

To interpet the integration we view $\frac{1}{2i}d\overline{z}\,dz = \frac{1}{2i}d\overline{z}\wedge dz$ as a differential form. Writing $z = x + iy$ for $x, y \in \mathbb{R}$, we see that

$$\frac{1}{2i}d\overline{z}\wedge dz = \frac{1}{2i}(dx - idy)\wedge(dx + idy) = dx\wedge dy,$$

i.e., the integration measure is the standard Lebesgue measure on $\mathbb{R}^2$, identified with $\mathbb{C}$ via the decomposition into real and imaginary parts. Now for a smooth (not necessarily analytic) function $f : \mathbb{C} \to \mathbb{C}$, $f_z = \partial_z f = \frac{\partial f}{\partial z}$ and $f_{\overline{z}} = \partial_{\overline{z}} f = \frac{\partial f}{\partial \overline{z}}$ are defined by the formula

$$df = f_z\,dz + f_{\overline{z}}\,d\overline{z}.$$

By substituting the formulas $dz = dx + idy$ and $d\overline{z} = dx - idy$, one obtains the concrete identities

$$f_z = \frac{1}{2}(f_x - if_y), \quad f_{\overline{z}} = \frac{1}{2}(f_x + if_y),$$

where f_x and f_y are the standard partial derivatives. By construction

$$\int f_z\,d\overline{z}\,dz = 0, \quad \int f_{\overline{z}}\,d\overline{z}\,dz = 0 \tag{6.2}$$

for any f (with derivatives of sufficient decay, e.g., of the Schwartz class). Indeed, to see this, note, e.g., that $f_z\, d\overline{z}\, dz = -df \wedge d\overline{z} = -d(f\, d\overline{z})$, hence the claim follows by Stokes' theorem.

It is natural to consider $f(z) = g(z, \overline{z})$, where $g : \mathbb{C}^2 \to \mathbb{C}$ is analytic. (In fact it is not hard to check that any real-analytic function $\mathbb{R}^2 \to \mathbb{C}$ can be written this way.) For example, choosing $g(z, w) = zw$ yields $f(z) = |z|^2$. In this case $f_z(z) = g_z(z, \overline{z})$ and $f_{\overline{z}}(z) = g_w(z, \overline{z})$. Roughly speaking, we can think of $z, \overline{z}$ algebraically as independent variables and compute the derivatives f_z and $f_{\overline{z}}$ via the application of the usual symbolic rules to any given formula for f. In our example $f(z) = |z|^2$, this means that $f_z = \overline{z}$ and $f_{\overline{z}} = z$. To confirm this claim, one observes (by writing difference quotients) that $f_x(z) = g_z(z, \overline{z}) + g_w(z, \overline{z})$ and $f_y(z) = i(g_z(z, \overline{z}) - g_w(z, \overline{z}))$, hence $f_z(z) = g_z(z, \overline{z})$ and $f_{\overline{z}}(z) = g_w(z, \overline{z})$, as desired.

6.1.3 The resolution of identity

Consider $f(z) = e^{-|z|^2}$, which can be written $f(z) = g(z, \overline{z})$, where $g(z, w) = e^{-zw}$. By inductively applying (6.2) to derivatives $\partial_z^i \partial_{\overline{z}}^j f$, one can show that

$$\int z^m \overline{z}^n\, e^{-|z|^2}\, d\overline{z}\, dz = 0, \quad \text{if } m \neq n.$$

One also has the elementary identity

$$\int e^{-|z|^2}\, \frac{d\overline{z}\, dz}{2\pi i} = \frac{1}{\pi} \int e^{-(x^2+y^2)}\, dx\, dy = 1.$$

This identity is the motivation for the normalization of the measure $\frac{d\overline{z}\, dz}{2\pi i}$ and shall be directly analogized later on in the fermionic setting. Now by using (6.2) and induction once again, one derives that in turn

$$\int |z|^{2m} e^{-|z|^2}\, \frac{d\overline{z}\, dz}{2\pi i} = m!$$

Hence in summary

$$\frac{1}{m!} \int z^m \overline{z}^n\, e^{-|z|^2}\, \frac{d\overline{z}\, dz}{2\pi i} = \delta_{mn}. \tag{6.3}$$

Via polynomial approximation, (6.3) tells us how to integrate arbitrary functions against the measure $e^{-|z|^2}\, \frac{d\overline{z}\, dz}{2\pi i}$; hence (6.3) can be thought of as an *algebraic* specification of this measure. It is this sense that can be extrapolated to the fermionic setting.

Now we have the tools needed to verify the resolution of the identity (6.1), for which it suffices to apply $\langle -|b^m$ from the left, as follows:

$$\langle -|b^m \int \frac{dz\, d\overline{z}}{2\pi i}\, e^{-|z|^2} |z\rangle\langle z| \;\; = \;\; \int \frac{d\overline{z}\, dz}{2\pi i}\, e^{-|z|^2} z^m \langle -|z\rangle\langle z|$$

$$
\begin{aligned}
&= \int \frac{d\overline{z}\, dz}{2\pi i}\; e^{-|z|^2} z^m \langle z| \\
&= \int \frac{d\overline{z}\, dz}{2\pi i}\; e^{-|z|^2} z^m \langle -| \sum_{n=0}^{\infty} \frac{\overline{z}^n}{n!} b^n \\
&= \langle -| \sum_{n=0}^{\infty} b^n \frac{1}{m!} \int z^m \overline{z}^n\, e^{-|z|^2}\, \frac{d\overline{z}\, dz}{2\pi i} \\
&= \langle -| b^m
\end{aligned}
$$

Note the essential features of the derivation: the eigenfunction property of the coherent state $|z\rangle$, the normalization $\langle -|z\rangle = 1$, and the integration identity (6.3). The reader should keep this features in mind for the fermionic setting.

It is straightforward to likewise verify the more general resolution of identity for arbitrary $|\mathcal{B}| \geq 1$:

$$
\mathrm{Id} = \int \left[\prod_{p \in \mathcal{B}} \frac{d\overline{z}_p\, dz_p}{2\pi i} \right]\; e^{-|\mathbf{z}|^2} |\mathbf{z}\rangle\langle \mathbf{z}| = \int d(\overline{\mathbf{z}}, \mathbf{z})\; e^{-|\mathbf{z}|^2} |\mathbf{z}\rangle\langle \mathbf{z}|, \tag{6.4}
$$

where we introduce the formal notation $d(\overline{\mathbf{z}}, \mathbf{z}) := \prod_{p \in \mathcal{B}} \frac{d\overline{z}_p\, dz_p}{2\pi i}$. The relevant integration identity is simply the product measure version of (6.3):

$$
\frac{1}{\prod_p m_p!} \int \left(\prod_{p \in \mathcal{B}} z^{m_p} \overline{z}^{n_p} \right)\; e^{-|z|^2}\, d(\overline{\mathbf{z}}, \mathbf{z}) = \prod_{p \in \mathcal{B}} \delta_{m_p n_p}. \tag{6.5}
$$

6.1.4 Path integral

As suggested above, we expand the partition function (temporarily lumping the chemical potential contribution into the Hamiltonian $\hat{H}$) as:

$$
\begin{aligned}
Z &= \mathrm{Tr}\left[e^{-\beta \hat{H}} \right] \\
&= \mathrm{Tr}\left[e^{-\beta \hat{H}}\, \mathrm{Id} \right] \\
&= \int d\big(\overline{\mathbf{z}}_{(0)}, \mathbf{z}_{(0)}\big)\, e^{-|\mathbf{z}_{(0)}|^2} \langle \mathbf{z}_{(0)} | e^{-\beta \hat{H}} | \mathbf{z}_{(0)} \rangle \\
&= \int d\big(\overline{\mathbf{z}}_{(0)}, \mathbf{z}_{(0)}\big)\, e^{-|\mathbf{z}_{(0)}|^2} \langle \mathbf{z}_{(0)} | e^{-\frac{1}{M}\beta \hat{H}} \cdots e^{-\frac{1}{M}\beta \hat{H}} | \mathbf{z}_{(0)} \rangle \\
&= \int \left[\prod_{m=0}^{M-1} d\big(\overline{\mathbf{z}}_{(m)}, \mathbf{z}_{(m)}\big) \right]\; e^{-\sum_{m=0}^{M-1} |\mathbf{z}_{(m)}|^2} \langle \mathbf{z}_{(0)} | e^{-\frac{1}{M}\beta \hat{H}} | \mathbf{z}_{(M-1)} \rangle \cdots \langle \mathbf{z}_{(1)} | e^{-\frac{1}{M}\beta \hat{H}} | \mathbf{z}_{(0)} \rangle.
\end{aligned}
$$

Now for M large, one hopes that $e^{-\frac{1}{M}\beta \hat{H}} = :e^{-\frac{1}{M}\beta \hat{H}}: + O(M^{-2})$, allowing us to substitute

$$\begin{aligned}\langle \mathbf{z}_{(m+1)}|e^{-\frac{1}{M}\beta\hat{H}}|\mathbf{z}_{(m)}\rangle &\approx \langle \mathbf{z}_{(m+1)}| :e^{-\frac{1}{M}\beta\hat{H}}: |\mathbf{z}_{(m)}\rangle \\ &= e^{-\frac{1}{M}\beta H(\overline{\mathbf{z}}_{(m+1)},\mathbf{z}_{(m)})}\langle \mathbf{z}_{(m+1)}|\mathbf{z}_{(m)}\rangle \\ &= e^{-\frac{1}{M}\beta H(\overline{\mathbf{z}}_{(m+1)},\mathbf{z}_{(m)})}e^{\mathbf{z}^*_{(m+1)}\mathbf{z}_{(m)}}.\end{aligned}$$

Indeed $e^{-\frac{1}{M}\beta\hat{H}} = 1 - \frac{1}{M}\beta\hat{H} + O(M^{-2})$, and by assumption $:\hat{H}: = \hat{H}$, so indeed our hope is justified (though a complete justification of the path integral would require significant further argument). Proceeding, we compute (interpreting the index m modulo M):

$$\begin{aligned}Z &= \lim_{M\to\infty}\int\left[\prod_{m=0}^{M-1} d(\overline{\mathbf{z}}_{(m)},\mathbf{z}_{(m)})\right] e^{-\sum_{m=0}^{M-1}\left[|\mathbf{z}_{(m+1)}|^2 - \mathbf{z}^*_{(m+1)}\mathbf{z}_{(m)}\right] - \frac{\beta}{M}\sum_{m=0}^{M-1} H(\overline{\mathbf{z}}_{(m+1)},\mathbf{z}_{(m)})} \\ &= \lim_{M\to\infty}\int\left[\prod_{m=0}^{M-1} d(\overline{\mathbf{z}}_{(m)},\mathbf{z}_{(m)})\right] e^{-\sum_{m=0}^{M-1}\left[\mathbf{z}^*_{(m+1)}(\mathbf{z}_{(m+1)} - \mathbf{z}_{(m)})\right] - \frac{\beta}{M}\sum_{m=0}^{M-1} H(\overline{\mathbf{z}}_{(m+1)},\mathbf{z}_{(m)})} \\ \text{“}=\text{”}\quad &\int D_{\text{per}}\left[\overline{\mathbf{z}}(\,\cdot\,),\mathbf{z}(\,\cdot\,)\right]\, e^{-\int_0^\beta [\mathbf{z}(\tau)^*\partial_\tau \mathbf{z}(\tau) + H(\overline{\mathbf{z}}(\tau),\mathbf{z}(\tau))]\,d\tau}\end{aligned}$$

where $D_{\text{per}}\left[\overline{\mathbf{z}}(\,\cdot\,),\mathbf{z}(\,\cdot\,)\right]$ is formally the infinite-dimensional Lebesgue measure (properly normalized) on periodic paths $\mathbf{z} : [0,\beta) \to \mathbb{C}$. Here the “ = ” indicates that the expression in the last line of the display is only formal and ought to be more rigorously understood as a limit as $M \to \infty$. Nonetheless, the formal perspective offers significant insight!

Then by replacing $\hat{H} \leftarrow \hat{H} - \mu\hat{N}$ and noting that $\hat{N}(\overline{\mathbf{z}},\mathbf{z}) = |\mathbf{z}|^2$, we obtain the path integral formulation of the partition function

$$Z = \int D_{\text{per}}\left[\overline{\mathbf{z}}(\,\cdot\,),\mathbf{z}(\,\cdot\,)\right]\, e^{-S(\overline{\mathbf{z}},\mathbf{z})},$$

where the action S is defined by

$$S(\overline{\mathbf{z}},\mathbf{z}) := \int_0^\beta \left[\mathbf{z}(\tau)^*(\partial_\tau - \mu)\mathbf{z}(\tau) + H(\overline{\mathbf{z}}(\tau),\mathbf{z}(\tau))\right]\,d\tau$$

If we write $\hat{H}$ as a sum of a noninteracting part $\hat{H}_0$ and an interaction $\hat{U} = U(\mathbf{b}^\dagger,\mathbf{b})$, i.e.,

$$\hat{H} = \hat{H}_0 + \hat{U} = \sum_{p,q} h_{pq} b_p^\dagger b_q + \hat{U},$$

then we can write

$$S(\overline{\mathbf{z}},\mathbf{z}) = S_0(\overline{\mathbf{z}},\mathbf{z}) + S_{\text{int}}(\overline{\mathbf{z}},\mathbf{z}),$$

where

$$S_0(\overline{\mathbf{z}},\mathbf{z}) := \int_0^\beta \mathbf{z}(\tau)^*(\partial_\tau + h - \mu)\mathbf{z}(\tau)\,d\tau, \quad S_{\text{int}}(\overline{\mathbf{z}},\mathbf{z}) = \int_0^\beta U(\overline{\mathbf{z}}(\tau),\mathbf{z}(\tau)).$$

In particular, for a general two-body interaction

$$\hat{U} = \sum_{p,q,r,s} U_{pqrs} a_p^\dagger a_q^\dagger a_s a_r, \tag{6.6}$$

we have

$$S_{\text{int}}(\overline{\mathbf{z}}, \mathbf{z}) = \sum_{pqrs} U_{pqrs} \int_0^\beta \overline{z}_p(\tau)\overline{z}_q(\tau) z_s(\tau) z_r(\tau)\, d\tau,$$

and for the generalized Coulomb interaction $\hat{U} = \sum_{pq} v_{pq} \hat{n}_p \hat{n}_q$ (with $v_{pp} = 0$), which corresponds to the choice $U_{pqrs} = v_{pq}\delta_{pr}\delta_{qs}$, we have

$$S_{\text{int}}(\overline{\mathbf{z}}, \mathbf{z}) = \sum_{pq} v_{pq} \int_0^\beta |z_p(\tau)|^2 |z_q(\tau)|^2\, d\tau.$$

Observe, at this point, the formal similarity of the path integral to the Euclidean field theory presented in section 1. There is, however, a crucial difference. The contribution of the term $\int_0^\beta \mathbf{z}(\tau)^* \partial_\tau \mathbf{z}(\tau)\, d\tau$ to the action includes an imaginary part, hence the path integral cannot be interpreted as a Gibbs measure, even in an infinite-dimensional sense. (This scenario should be contrasted with that of the Feynman path integral of section 3.3.)

6.1.5 Path integral in frequency space

Since the action in the path integral is time-translation-invariant and our paths are periodic, it makes sense to consider our paths in the frequency space. To begin we define frequency representations $\hat{\overline{\mathbf{z}}}(\omega_n)$ and $\hat{\mathbf{z}}(\omega_n)$ of the periodic complex paths on the Matsubara frequencies $\omega_n = 2n\pi/\beta$ (where $n \in \mathbb{Z}$):

$$\mathbf{w}(\omega_n) = \frac{1}{\sqrt{\beta}} \int_0^\beta \mathbf{z}(\tau) e^{i\omega_n \tau}\, d\tau,$$

so

$$\overline{\mathbf{w}}(\omega_n) = \frac{1}{\sqrt{\beta}} \int_0^\beta \overline{\mathbf{z}}(\tau) e^{-i\omega_n \tau}\, d\tau,$$

and we have

$$\overline{\mathbf{z}}(\tau) = \frac{1}{\sqrt{\beta}} \sum_n \overline{\mathbf{w}}(\omega_n) e^{i\omega_n \tau}, \quad \mathbf{z}(\tau) = \frac{1}{\sqrt{\beta}} \sum_n \mathbf{w}(\omega_n) e^{-i\omega_n \tau}.$$

Then we convert our action to the frequency representation by computing

$$\int_0^\beta \mathbf{z}^*(\tau) \partial_\tau \mathbf{z}(\tau)\, d\tau \;=\; \frac{1}{\beta} \sum_{nm} -i\omega_m \mathbf{w}^*(\omega_n)\mathbf{w}(\omega_m) \underbrace{\int_0^\beta e^{i(\omega_n - \omega_m)\tau}\, d\tau}_{=\beta\delta_{nm}}$$

$$= \sum_n (-i\omega_n)\mathbf{w}^*(\omega_n)\mathbf{w}(\omega_n),$$

and

$$\int_0^\beta \mathbf{z}^*(\tau)(h-\mu)\mathbf{z}(\tau)\,d\tau = \sum_n \mathbf{w}^*(\omega_n)(h-\mu)\mathbf{w}(\omega_n).$$

Finally, for the two-body interaction (6.6) we compute

$$\begin{aligned} S_{\text{int}}(\overline{\mathbf{z}},\mathbf{z}) &:= \int_0^\beta \sum_{ijkl} U_{ijkl}\overline{z}_i(\tau)\overline{z}_j(\tau)\xi_l(\tau)\xi_k(\tau)\,d\tau \\ &= \frac{1}{\beta^2}\sum_{ijkl} U_{ijkl} \sum_{mnpq} \overline{w}_i(\omega_m)\overline{w}_j(\omega_n)w_l(\omega_q)w_k(\omega_p) \int_0^\beta e^{i(\omega_m+\omega_n-\omega_p-\omega_q)\tau}\,d\tau \\ &= \frac{1}{\beta}\sum_{ijkl}\sum_{mnpq} U_{ijkl}\,\delta_{m+n,p+q}\,\overline{w}_i(\omega_m)\overline{w}_j(\omega_n)w_l(\omega_q)w_k(\omega_p) \\ &=: \hat{S}_{\text{int}}(\overline{\mathbf{w}},\mathbf{w})\,. \end{aligned}$$

Since the transformations $\mathbf{z} \mapsto \hat{\mathbf{z}}$ is a unitary change of variables, we have that

$$Z = \int \hat{D}\left[\overline{\mathbf{w}}(\,\cdot\,),\mathbf{w}(\,\cdot\,)\right]\; e^{-\hat{S}(\overline{\mathbf{w}},\mathbf{w})}$$

where $\hat{D}\left[\overline{\mathbf{w}}(\,\cdot\,),\mathbf{w}(\,\cdot\,)\right]$ is understood as the infinite-dimensional Grassmann Lebesgue measure $\prod_{n\in\mathbb{Z}} d(\overline{\mathbf{w}},\mathbf{w})$, and

$$\hat{S}(\overline{\mathbf{w}},\mathbf{w}) = \hat{S}_0(\overline{\mathbf{w}},\mathbf{w}) + \hat{S}_{\text{int}}(\overline{\mathbf{w}},\mathbf{w})$$

with

$$\hat{S}_0(\overline{\mathbf{w}},\mathbf{w}) := \sum_n \mathbf{w}^*(\omega_n)(-i\omega_n + h - \mu)\mathbf{w}(\omega_n).$$

6.2 The fermionic coherent state path integral

When we try to mimic the derivation of the bosonic coherent states we immediately run into a difficulty. Indeed, consider the case of a single-state fermionic system, i.e., $|\mathcal{B}| = 1$, and suppose that $|\mathbf{z}\rangle$ is an eigenstate of the annihilation operators c_p with corresponding eigenvalues z_p. Then $c_pc_q|\mathbf{z}\rangle = z_qz_p|\mathbf{z}\rangle$, but also $c_pc_q|\mathbf{z}\rangle = -c_qc_p|\mathbf{z}\rangle = -z_pz_q|\mathbf{z}\rangle$, so $z_pz_q = -z_qz_p$. In particular, it follows that $z_p = 0$ for all p, hence apparently any coherent state is in the null space of all of the annihilation operators. But the only state satisfying this property is the vacuum state! Clearly this won't do.

6.2.1 Grassmann numbers

To find our coherent states, we have to expand the space of numbers in which we look for eigenvalues. In particular, following the above reasoning, we want our eigenvalues ξ_i to satisfy $\xi_i\xi_j = -\xi_j\xi_i$, i.e., we want them to anticommute. This motivates the introduction of the algebra $\mathcal{G} = \mathcal{G}(\mathcal{B})$ of Grassmann numbers (also known as supernumbers [30]), which can be identified with the exterior algebra $\Lambda(\mathbb{C}^{\mathcal{B}})$ via the isomorphism

$$e_{p_1} \wedge \cdots \wedge e_{p_m} \mapsto \xi_{p_1} \cdots \xi_{p_m},$$

with the additional stipulation that $\Lambda^0(\mathbb{C}^{\mathcal{B}}) \simeq \mathbb{C}$ corresponds to the complex part of a Grassmann number, sometimes referred to as the 'body.' More concretely, a Grassmann number z can be written uniquely as

$$z = z_B + z_S = \sum_m \sum_{p_1<\cdots<p_m} c_{p_1\cdots p_m}\xi_{p_1} \cdots \xi_{p_m},$$

where z_B is the complex part or body, and z_S is the rest, i.e., the 'soul.'

6.2.2 Fermionic coherent states

In fact, we will always consider an extension $\mathcal{G}^*$ of this algebra that allows us to take adjoints, and the extended algebra (itself a Grassmann algebra) will in fact be a star-algebra. Concretely, the extension is achieved by considering the Grassmann algebra generated by an enlarged set $\{\xi_p, \overline{\xi}_p\}_{p\in\mathcal{B}}$ of anticommuting symbols; hence our algebra will be isomorphic to $\Lambda(\mathbb{C}^{\mathcal{B}\sqcup\mathcal{B}})$. Moreover, the adjoint '$*$' is defined via $(c\,\xi_{p_1} \cdots \xi_{p_m})^* = \overline{c}\,\overline{\xi}_{p_m} \cdots \overline{\xi}_{p_1}$. (Note that the notation for $\mathcal{G}$ and $\mathcal{G}_*$ is nonstandard.)

Moreover, we can consider Grassmann numbers as multipliers on Fock space operators as in the expression $\xi_i c_i^\dagger$. Mathematically we are considering $\xi_i c_i^\dagger$ as an element of the extended star-algebra $\mathrm{End}_{\mathcal{G}^*}(\mathcal{F}) = \mathcal{G}^* \otimes_{\mathbb{C}} \mathrm{End}_{\mathbb{C}}(\mathcal{F})$, where multiplication is defined, for $z \in \mathcal{G}^*$ and by $(z_1 \otimes \hat{O}_1)(z_2 \otimes \hat{O}_2) = z_1 z_2 \otimes \hat{O}_1\hat{O}_2$, or, for short, $(z_1\hat{O}_1)(z_2\hat{O}_2) = z_1 z_2 \hat{O}_1\hat{O}_2$. Moreover, the adjoint is defined by $(z\hat{O})^\dagger = z^*\hat{O}^\dagger$. Likewise we can extend the Fock space via the Grassmann algebra as $\mathcal{G}^* \otimes_{\mathbb{C}} \mathcal{F}$ to consider elements such as $\xi_p|\mathbf{n}\rangle$, and we can extend the dual space to consider elements such as $\langle\mathbf{n}|\overline{\xi}_p$.

For the most part, such technicalities need not be emphasized. Nonetheless, we have given some indication of the mathematical structures in order to reassure the readers that the Grassmann numbers and all accompanying manipulations can in fact be backed by honest mathematical definition.

Finally, by analogy with the bosonic case, we define the fermionic coherent state

$$|\boldsymbol{\xi}\rangle = e^{\boldsymbol{\xi}\cdot\mathbf{c}^\dagger}|-\rangle.$$

It is important to note that unlike complex numbers, the Grassmann generators ξ_i should not be thought of as variables with indeterminate numerical value; rather they

are mere symbols constrained to satisfy certain algebraic relations. In the development of the theory of bosonic coherent states, we attempted to emphasize the role of $z_i, \overline{z}_i$ as mere symbols and of integration as a recipe for assigning numerical values to algebraic expressions in these symbols. To transfer our developments to the fermionic setting, we will likewise only need a recipe for 'integrating' (or assigning numerical value to) elements of the Grassmann algebra.

Now observe that the via the anticommutation of both the Grassmann generators and the fermionic creation operators we have that that $\xi_i c_i^\dagger$ all commute within $\mathrm{End}_{\mathcal{G}^*}(\mathcal{F})$, much like the analogous bosonic operators $z_i b_i^\dagger$, where $z_i \in \mathbb{C}$. Hence

$$e^{\boldsymbol{\xi}\cdot\mathbf{c}^\dagger} = e^{\sum_p \xi_p c_p^\dagger} = \prod_p e^{\xi_p c_p^\dagger},$$

where the order in the product of the latter expression can be arbitrary. Now $e^{\xi_p c_p^\dagger} = 1 + \xi_p c_p^\dagger$; note that the Taylor series series terminates abruptly because $\xi_p^2 = 0$.

Now let us verify that $|\boldsymbol{\xi}\rangle \in \mathcal{G}^* \otimes_{\mathbb{C}} \mathcal{F}$ are eigenfunctions for the annihilation operators c_p with eigenvalues $\xi_p \in \mathcal{G}^*$:

$$\begin{aligned}
c_p|\boldsymbol{\xi}\rangle &= c_p(1+\xi_p c_p^\dagger)\prod_{q\neq p} e^{\xi_q c_q^\dagger}|-\rangle \\
&= (c_p + \xi_p c_p c_p^\dagger)\prod_{q\neq p} e^{\xi_q c_q^\dagger}|-\rangle \\
&= (c_p + \xi_p(1-c_p^\dagger c_p))\prod_{q\neq p} e^{\xi_q c_q^\dagger}|-\rangle \\
&= \xi_p \prod_{q\neq p} e^{\xi_q c_q^\dagger}|-\rangle + (1-c_p^\dagger)c_p \prod_{q\neq p} e^{\xi_q c_q^\dagger}|-\rangle.
\end{aligned}$$

Now notice that $\prod_{q\neq p} e^{\xi_q c_q^\dagger}|-\rangle$ is in the zero-eigenspace of $\hat{n}_p$, hence $c_p \prod_{q\neq p} e^{\xi_q c_q^\dagger}|-\rangle = 0$, and the second term in the last display is zero. Moreover $\xi_p = \xi_p(1+\xi_p c_p^\dagger)$, so we have derived

$$c_p|\boldsymbol{\xi}\rangle = \xi_p|\boldsymbol{\xi}\rangle,$$

as desired. Note that the adjoint coherent state is given by

$$\langle\boldsymbol{\xi}| = |\boldsymbol{\xi}\rangle^* = \langle -|e^{\overline{\boldsymbol{\xi}}\cdot c_p}.$$

6.2.3 Grassmann integration

Now in order to formulate a resolution of identity, we need an integration formula.[2] First let us focus on the case of $|\mathcal{B}| = 1$, i.e., the case of $\mathcal{G}^* = \langle\overline{\xi}, \xi\rangle$, where we use angle

[2]We will only consider integration on algebras with adjoint symbols in order to emphasize the analogy with the bosonic case, but in fact Grassmann integration can also be defined without difficulty on any Grassmann algebra. However, the evenness that accompanies the adjoint structure makes some aspects of the theory more elegant because even elements of the Grassmann algebra are commute with all elements of the algebra.

brackets to indicate the Grassmann algebra generated by the anticommuting symbols therein contained. The motivation here will be to analogize (6.5) from the bosonic setting. Since $\xi^m = \overline{\xi}^n = 0$ for $m, n \geq 2$, it is sufficient to *define* an integration functional $\mathcal{I} : \mathcal{G}^* \to \mathbb{C}$ via

$$\delta_{mn} = \mathcal{I}\left[\xi^m \overline{\xi}^n \, e^{-\overline{\xi}\xi}\right] =: \int \xi^m \overline{\xi}^n \, e^{-\overline{\xi}\xi} \, d(\overline{\xi}, \xi)$$

for $m, n \in \{0, 1\}$. Note that

$$e^{-\overline{\xi}\xi} = 1 - \overline{\xi}\xi = 1 + \xi\overline{\xi}$$

$$\xi e^{-\overline{\xi}\xi} = \xi(1 - \overline{\xi}\xi) = \xi$$

$$\overline{\xi} e^{-\overline{\xi}\xi} = \overline{\xi}(1 - \overline{\xi}\xi) = \overline{\xi}$$

$$\xi\overline{\xi} e^{-\overline{\xi}\xi} = \xi\overline{\xi}(1 - \overline{\xi}\xi) = \xi\overline{\xi},$$

so it follows that one could otherwise define $\mathcal{I}$ via

$$\mathcal{I}[1] = \mathcal{I}[\xi] = \mathcal{I}[\overline{\xi}] = 0, \quad \mathcal{I}[\xi\overline{\xi}] = 1.$$

To define an integration on a more general algebra $\mathcal{G}^*(\mathcal{B}) = \langle \{\overline{\xi}_p, \xi_p\}_{p \in \mathcal{B}} \rangle$, we seek to analogize the bosonic integration formula via

$$\int \left(\prod_{p \in \mathcal{B}} \xi^{m_p} \overline{\xi}^{n_p} \right) e^{-\boldsymbol{\xi}^* \boldsymbol{\xi}} \, d(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) := \mathcal{I}\left[\left(\prod_{p \in \mathcal{B}} \xi^{m_p} \overline{\xi}^{n_p} \right) e^{-\boldsymbol{\xi}^* \boldsymbol{\xi}} \right] := \prod_{p \in \mathcal{B}} \delta_{m_p n_p} \tag{6.7}$$

for $m_p, n_p \in \{0, 1\}$. Sometimes we will alternatively write $d(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) = \prod_{p \in \mathcal{B}} d(\overline{\xi}_p, \xi_p)$ to denote the multivariate Grassmann integration 'measure.' We will sometimes write $d(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi})$ immediately after the integration sign $\int$, as in $\int d(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) \cdots$, but the meaning is unchanged. Note carefullly that the definition makes sense regardless of the ordering of $p \in \mathcal{B}$ in the product because whenever $m_p + n_p$ is odd for some p, the result of the integration is defined to be zero; meanwhile, whenever $m_p + n_p$ is even, $\xi^{m_p} \overline{\xi}^{n_p}$ commutes with all elements of $\mathcal{G}^*$. One verifies that our definition of $\mathcal{I} : \mathcal{G}^*(\mathcal{B}) \to \mathbb{C}$ is equivalent to the definition

$$\mathcal{I}\left[\prod_{p \in \mathcal{B}} \xi_p \overline{\xi}_p \right] = 1, \quad \mathcal{I}\left[\text{any other monomial}\right] = 0.$$

We may also consider partial integration $\mathcal{I}_{\mathcal{S}} : \mathcal{G}^*(\mathcal{B}) \to \mathcal{G}^*(\mathcal{B} \backslash \mathcal{S})$, defined by

$$\mathcal{I}_{\mathcal{S}}\left[f\left((\overline{\xi}_p, \xi_p)_{p \in \mathcal{S}} \right) g\left((\overline{\xi}_p, \xi_p)_{p \notin \mathcal{S}} \right) \right] = \mathcal{I}\left[f\left((\overline{\xi}_p, \xi_p)_{p \in \mathcal{S}} \right) \right] g\left((\overline{\xi}_p, \xi_p)_{p \notin \mathcal{S}} \right)$$

for polynomials f, g. The left-hand side may alternatively be denoted by

$$\int f\left((\overline{\xi}_p, \xi_p)_{p\in\mathcal{S}}\right) g\left((\overline{\xi}_p, \xi_p)_{p\notin\mathcal{S}}\right) \left[\prod_{p\in\mathcal{S}} d(\overline{\xi}_p, \xi_p)\right].$$

There is in fact another perspective on the definition of the Grassmann integration; we may view it as an attempt to analogize (6.2). To pursue such an analogy, we need to define a suitable notion of Grassmann differentiation, i.e., linear operators $\partial_{\xi_p}, \partial_{\overline{\xi}_p} : \mathcal{G}^* \to \mathcal{G}^*$. These operators are determined entirely by the formulas

$$\partial_{\xi_p}\left[\xi_p f\left(\overline{\boldsymbol{\xi}}, \{\xi_p\}_{p\neq q}\right)\right] = f\left(\overline{\boldsymbol{\xi}}, \{\xi_p\}_{p\neq q}\right), \quad \partial_{\xi_p}\left[f\left(\overline{\boldsymbol{\xi}}, \{\xi_p\}_{p\neq q}\right)\right] = 0$$

$$\partial_{\overline{\xi}_p}\left[\overline{\xi}_p f\left(\{\overline{\xi}_p\}_{p\neq q}, \boldsymbol{\xi}\right)\right] = f\left(\{\overline{\xi}_p\}_{p\neq q}, \boldsymbol{\xi}\right), \quad \partial_{\overline{\xi}_p}\left[f\left(\{\overline{\xi}_p\}_{p\neq q}, \boldsymbol{\xi}\right)\right] = 0.$$

Note carefully, e.g., that $\partial_{\xi_1}(\xi_2\xi_1) = -\partial_{\xi_1}(\xi_1\xi_2) = -\xi_2$.

Then we can think of the stipulation that $\mathcal{I}\left[\left(\prod_{p\in\mathcal{B}} \xi_p\overline{\xi}_p\right) e^{-\boldsymbol{\xi}^*\boldsymbol{\xi}}\right] = 1$, or equivalently that $\mathcal{I}\left[\prod_{p\in\mathcal{B}} \xi_p\overline{\xi}_p\right] = 1$, as a kind of arbitrary 'normalization' of the Grassmann measure, just as in the bosonic case. Meanwhile, integration of arbitrary polynomials can then by defined via the stipulation, analogous to (6.2), that

$$\int \partial_{\xi_p} f \; d(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) = \int \partial_{\overline{\xi}_p} f \; d(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) = 0$$

for all polynomials f. Since every monomial in $\mathcal{G}^*(\mathcal{B})$ besides $\prod_{p\in\mathcal{B}} \xi_p\overline{\xi}_p$ can be written as a derivative, the integration rule introduced above follows.

6.2.4 The resolution of identity

Due to the eigenfunction property of the coherent state $|\boldsymbol{\xi}\rangle$, the normalization $\langle -|\boldsymbol{\xi}\rangle = 1$, and the integration identity (6.7) (analogous to (6.5)), the proof of our resolution of identity will be analogous to the proof in the bosonic case. The resolution of identity is written

$$\mathrm{Id}_{\mathcal{F}} = \int d(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) \; e^{-\boldsymbol{\xi}^*\boldsymbol{\xi}} |\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}|, \tag{6.8}$$

and we prove by applying an arbitrary occupation number basis element $\langle -|c_{p_m}\cdots c_{p_1}$ from the left, as

$$\begin{aligned}
\langle -|c_{p_m}\cdots c_{p_1} \int d(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) \; e^{-\boldsymbol{\xi}^*\boldsymbol{\xi}} |\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}| &= \int d(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) \; e^{-\boldsymbol{\xi}^*\boldsymbol{\xi}} \xi_{p_1}\cdots\xi_{p_m}\langle -|\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}| \\
&= \int d(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) \; e^{-\boldsymbol{\xi}^*\boldsymbol{\xi}} \xi_{p_1}\cdots\xi_{p_m}\langle -|\prod_{p\in\mathcal{B}}(1+\overline{\xi}_p c_p) \\
&= \langle -|\int \xi_{p_1}\cdots\xi_{p_m}\prod_{p\in\mathcal{B}}(1+\overline{\xi}_p c_p)\, e^{-\boldsymbol{\xi}^*\boldsymbol{\xi}} \; d(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}),
\end{aligned}$$

where we have used the fact that $e^{-\boldsymbol{\xi}^*\boldsymbol{\xi}}$ commutes with all of $\mathcal{G}^*$. Then note that upon expanding the product, by (6.7) the only term that survives is $\overline{\xi}_{p_m}\cdots\overline{\xi}_{p_1}c_{p_m}\cdots c_{p_1}$, hence

$$\begin{aligned}
&\langle -|c_{p_m}\cdots c_{p_1}\int d(\overline{\boldsymbol{\xi}},\boldsymbol{\xi})\; e^{-\boldsymbol{\xi}^*\boldsymbol{\xi}}|\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}| \\
&\quad = \langle -|c_{p_m}\cdots c_{p_1}\left[\int \xi_{p_1}\cdots\xi_{p_m}\overline{\xi}_{p_m}\cdots\overline{\xi}_{p_1}\, e^{-\boldsymbol{\xi}^*\boldsymbol{\xi}}\; d(\overline{\boldsymbol{\xi}},\boldsymbol{\xi})\right] \\
&\quad = \langle -|c_{p_m}\cdots c_{p_1},
\end{aligned}$$

as was to be shown. In the last step we used (6.7), together with the fact that

$$\xi_{p_1}\cdots\xi_{p_m}\overline{\xi}_{p_m}\cdots\overline{\xi}_{p_1} = (\xi_{p_1}\overline{\xi}_{p_1})\cdots(\xi_{p_m}\overline{\xi}_{p_m}),$$

which follows from grouping the factor $\xi_{p_m}\overline{\xi}_{p_m}$, which commutes with the entire algebra, and moving it all the way to the right, then repeating for $\xi_{p_{m-1}}\overline{\xi}_{p_{m-1}}$, etc.

In order to use the resolution of identity to compute traces, it is useful to derive the following identity:

$$\langle\mathbf{m}|\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}|\mathbf{n}\rangle = \langle -\boldsymbol{\xi}|\mathbf{n}\rangle\langle\mathbf{m}|\boldsymbol{\xi}\rangle, \quad \text{when} \quad \sum_p (m_p - n_p) \equiv 0 \mod 2. \tag{6.9}$$

Here we interpret $|-\boldsymbol{\xi}\rangle = e^{(-\boldsymbol{\xi})\cdot\mathbf{c}^\dagger}|-\rangle$, and $\langle -\boldsymbol{\xi}| = |-\boldsymbol{\xi}\rangle^* = \langle -|e^{(-\overline{\boldsymbol{\xi}})\cdot\mathbf{c}}$. Note that $c_p|-\boldsymbol{\xi}\rangle = -\xi_p|-\boldsymbol{\xi}\rangle$ for all p.

Now to prove the claim, write $|\mathbf{m}\rangle = c^\dagger_{p_1}\cdots c^\dagger_{p_M}|-\rangle$ and $|\mathbf{n}\rangle = c^\dagger_{q_1}\cdots c^\dagger_{q_N}|-\rangle$, where $M-N$ is even, and compute

$$\begin{aligned}
\langle\mathbf{m}|\boldsymbol{\xi}\rangle &= \langle -|c_{p_M}\cdots c_{p_1}|\boldsymbol{\xi}\rangle = \xi_{p_1}\cdots\xi_{p_M}\langle -|\boldsymbol{\xi}\rangle = \xi_{p_1}\cdots\xi_{p_M} \\
\langle\boldsymbol{\xi}|\mathbf{n}\rangle &= \langle\boldsymbol{\xi}|c^\dagger_{q_1}\cdots c^\dagger_{q_N}|-\rangle = \overline{\xi}_{q_N}\cdots\overline{\xi}_{q_1} \\
\langle -\boldsymbol{\xi}|\mathbf{n}\rangle &= \langle -\boldsymbol{\xi}|c^\dagger_{q_1}\cdots c^\dagger_{q_N}|-\rangle = (-\overline{\xi}_{q_N})\cdots(-\overline{\xi}_{q_1}) = (-1)^N\overline{\xi}_{p_1}\cdots\overline{\xi}_{p_m}.
\end{aligned}$$

Then

$$\begin{aligned}
\langle\mathbf{m}|\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}|\mathbf{n}\rangle &= \xi_{p_1}\cdots\xi_{p_M}\overline{\xi}_{q_N}\cdots\overline{\xi}_{q_1} \\
&= (-1)^{MN}\overline{\xi}_{q_N}\cdots\overline{\xi}_{q_1}\xi_{p_1}\cdots\xi_{p_M} \\
&= (-1)^{MN}(-1)^N\langle -\boldsymbol{\xi}|\mathbf{n}\rangle\langle\mathbf{m}|\boldsymbol{\xi}\rangle.
\end{aligned}$$

But

$$(-1)^{MN} = (-1)^{(N+M-N)N} = (-1)^{N^2}(-1)^{(M-N)N} = (-1)^N,$$

where we have used the facts that $(-1)^{N^2} = (-1)^N$ and that $M-N$ is even (so $(M-N)N$ is even as well). The claim (6.9) follows.

From the identity (6.9), together with our resolution of identity (6.7), we may derive a formula for $\mathrm{Tr}(\hat{O})$ for operators $\hat{O} = O(\mathbf{c}^\dagger, \mathbf{c})$, where O is an *even* polynomial (as is required of physical fermionic operators):

$$\mathrm{Tr}(\hat{O}) = \int d(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi})\ e^{-\boldsymbol{\xi}^*\boldsymbol{\xi}} \langle -\boldsymbol{\xi}|\hat{O}|\boldsymbol{\xi}\rangle. \tag{6.10}$$

To derive the identity, we expand as

$$\begin{aligned}
\mathrm{Tr}(\hat{O}) &= \mathrm{Tr}(\hat{O}\,\mathrm{Id}_{\mathcal{F}}) \\
&= \int d(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi})\ e^{-\boldsymbol{\xi}^*\boldsymbol{\xi}} \mathrm{Tr}\left(\hat{O}|\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}|\right) \\
&= \int d(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi})\ e^{-\boldsymbol{\xi}^*\boldsymbol{\xi}} \sum_{\mathbf{n}} \langle \mathbf{n}|\hat{O}|\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}|\mathbf{n}\rangle.
\end{aligned}$$

Now write $\hat{O} = \sum_{\mathbf{m}',\mathbf{m}} O_{\mathbf{m}'\mathbf{m}} |\mathbf{m}'\rangle\langle\mathbf{m}|$, where $O_{\mathbf{m}'\mathbf{m}} = 0$ whenever $\sum_p (m'_p - m_p)$ is odd. Then inserting this expression we obtain

$$\begin{aligned}
\mathrm{Tr}(\hat{O}) &= \int d(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi})\ e^{-\boldsymbol{\xi}^*\boldsymbol{\xi}} \sum_{\mathbf{n},\mathbf{m}} O_{\mathbf{n}\mathbf{m}} \langle \mathbf{m}|\boldsymbol{\xi}\rangle\langle\boldsymbol{\xi}|\mathbf{n}\rangle \\
&= \int d(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi})\ e^{-\boldsymbol{\xi}^*\boldsymbol{\xi}} \sum_{\mathbf{n},\mathbf{m}} O_{\mathbf{n}\mathbf{m}} \langle -\boldsymbol{\xi}|\mathbf{n}\rangle\langle\mathbf{m}|\boldsymbol{\xi}\rangle \\
&= \int d(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi})\ e^{-\boldsymbol{\xi}^*\boldsymbol{\xi}} \langle -\boldsymbol{\xi}|\hat{O}|\boldsymbol{\xi}\rangle,
\end{aligned}$$

as was to be shown. (In the second equality of the last display, we used (6.9), together with the fact that $O_{\mathbf{n}\mathbf{m}} = 0$ whenever $\sum_p (n_p - m_p)$ is odd.)

6.2.5 Path integral

Expand the partition function (again temporarily lumping the chemical potential contribution into the Hamiltonian $\hat{H}$) via the trace identity (6.10):

$$\begin{aligned}
Z &= \mathrm{Tr}\left[e^{-\beta\hat{H}}\right] \\
&= \int d(\overline{\boldsymbol{\xi}}_{(0)}, \boldsymbol{\xi}_{(0)})\, e^{-\boldsymbol{\xi}^*_{(0)}\boldsymbol{\xi}_{(0)}} \langle -\boldsymbol{\xi}_{(0)}|e^{-\beta\hat{H}}|\boldsymbol{\xi}_{(0)}\rangle \\
&= \int d(\overline{\boldsymbol{\xi}}_{(0)}, \boldsymbol{\xi}_{(0)})\, e^{-\boldsymbol{\xi}^*_0\boldsymbol{\xi}_0} \langle -\boldsymbol{\xi}_{(0)}|e^{-\frac{1}{M}\beta\hat{H}} \cdots e^{-\frac{1}{M}\beta\hat{H}}|\boldsymbol{\xi}_{(0)}\rangle \\
&= \int \left[\prod_{m=0}^{M-1} d(\overline{\boldsymbol{\xi}}_{(m)}, \boldsymbol{\xi}_{(m)})\right] e^{-\sum_{m=0}^{M-1} \boldsymbol{\xi}^*_{(m)}\boldsymbol{\xi}_{(m)}} \\
&\qquad \langle -\boldsymbol{\xi}_{(0)}|e^{-\frac{1}{M}\beta\hat{H}}|\boldsymbol{\xi}_{(M-1)}\rangle \cdots \langle \boldsymbol{\xi}_{(1)}|e^{-\frac{1}{M}\beta\hat{H}}|\boldsymbol{\xi}_{(0)}\rangle.
\end{aligned}$$

Observe that in the last line, the integration takes place in the enlarged Grassmann algebra

$$\mathcal{G}_M^* := \left\langle \{\overline{\xi}_{(m),p}, \xi_{(m),p}\}_{p\in\mathcal{B},\, m=0,\ldots,M-1} \right\rangle .$$

Evidently, in this enlarged Grassmann algebra we shall have to compute the overlaps $\langle \boldsymbol{\xi}_{(m)} | \boldsymbol{\xi}_{(m-1)} \rangle$.

More generally, we compute the overlap $\langle \boldsymbol{\theta} | \boldsymbol{\xi} \rangle$ within $\left\langle \{\overline{\xi}_{,p}, \xi_{,p}, \overline{\theta}_p, \theta_p\}_{p\in\mathcal{B}} \right\rangle$:

$$\langle \boldsymbol{\theta} | \boldsymbol{\xi} \rangle = e^{\boldsymbol{\theta}^* \boldsymbol{\xi}}, \tag{6.11}$$

analogously to the bosonic case. To verifty this identity, first rewrite

$$|\boldsymbol{\xi}\rangle = \prod_p (1 + \xi_p c_p^\dagger)|-\rangle = \sum_{\mathcal{S}\subset\mathcal{B}} \prod_{p\in\mathcal{S}} (\xi_p c_p^\dagger)|-\rangle.$$

Note that the ordering of p within the product does not matter. Similarly,

$$\langle \boldsymbol{\theta}| = \sum_{\mathcal{S}\subset\mathcal{B}} \langle -| \prod_{p\in\mathcal{S}} (\overline{\theta}_p c_p),$$

from which it follows that

$$\begin{aligned}
\langle \boldsymbol{\theta} | \boldsymbol{\xi} \rangle &= \sum_{\mathcal{S}\subset\mathcal{B}} \prod_{p\in\mathcal{S}} (\overline{\theta}_p \xi_p) \\
&= \prod_{p\in\mathcal{B}} (1 + \overline{\theta}_p \xi_p) \\
&= \prod_{p\in\mathcal{B}} e^{\overline{\theta}_p \xi_p} \\
&= e^{\boldsymbol{\theta}^* \boldsymbol{\xi}},
\end{aligned}$$

as was to be shown.

Now for M large, we again make use of $e^{-\frac{1}{M}\beta\hat{H}} = :e^{-\frac{1}{M}\beta\hat{H}}: + O(M^{-2})$, allowing us to substitute

$$\begin{aligned}
\langle \boldsymbol{\xi}_{(m+1)} | e^{-\frac{1}{M}\beta\hat{H}} | \boldsymbol{\xi}_{(m)} \rangle &\approx \langle \boldsymbol{\xi}_{(m+1)} | :e^{-\frac{1}{M}\beta\hat{H}}: | \boldsymbol{\xi}_{(m)} \rangle \\
&= e^{-\frac{1}{M}\beta H(\overline{\boldsymbol{\xi}}_{(m+1)}, \boldsymbol{\xi}_{(m)})} \langle \boldsymbol{\xi}_{(m+1)} | \boldsymbol{\xi}_{(m)} \rangle \\
&= e^{-\frac{1}{M}\beta H(\overline{\boldsymbol{\xi}}_{(m+1)}, \boldsymbol{\xi}_{(m)})} e^{\boldsymbol{\xi}^*_{(m+1)} \boldsymbol{\xi}_{(m)}}.
\end{aligned}$$

Proceeding, we compute, adopting the convention $\boldsymbol{\xi}_{(M)} = -\boldsymbol{\xi}_{(0)}$:

$$Z = \lim_{M\to\infty} \int \left[\prod_{m=0}^{M-1} d(\overline{\boldsymbol{\xi}}_{(m)}, \boldsymbol{\xi}_{(m)}) \right] e^{-\sum_{m=0}^{M-1} \left[\boldsymbol{\xi}^*_{(m+1)} (\boldsymbol{\xi}_{(m+1)} - \boldsymbol{\xi}_{(m)}) \right] - \frac{\beta}{M} \sum_{m=0}^{M-1} H(\overline{\boldsymbol{\xi}}_{(m+1)}, \boldsymbol{\xi}_{(m)})}$$

$$= \int D_{\text{a-per}}\left[\overline{\boldsymbol{\xi}}(\,\cdot\,), \boldsymbol{\xi}(\,\cdot\,)\right]\; e^{-\int_0^\beta \left[\boldsymbol{\xi}(\tau)^* \partial_\tau \boldsymbol{\xi}(\tau) + H(\overline{\boldsymbol{\xi}}(\tau), \boldsymbol{\xi}(\tau))\right] d\tau}$$

where $D_{\text{a-per}}\left[\overline{\boldsymbol{\xi}}(\,\cdot\,), \boldsymbol{\xi}(\,\cdot\,)\right]$ is formally the infinite-dimensional Lebesgue measure (properly normalized) on *antiperiodic* 'Grassmann paths' $\boldsymbol{\xi}(\tau)$. Again " $=$ " indicates that the expression in the last line of the display is only formal and ought to be more rigorously understood as a limit as $M \to \infty$. The notion even of a Grassmann path is shakily interpreted at best; by contrast to a complex path, it cannot be interpreted as an (anti)periodic function on $[0, \beta)$. Rather, it's meaning is only symbolic.

Then by replacing $\hat{H} \leftarrow \hat{H} - \mu \hat{N}$ and noting that $\hat{N}(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) = \boldsymbol{\xi}^* \boldsymbol{\xi}$, we obtain the path integral formulation of the partition function

$$Z = \int D_{\text{a-per}}\left[\overline{\boldsymbol{\xi}}(\,\cdot\,), \boldsymbol{\xi}(\,\cdot\,)\right]\; e^{-S(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi})},$$

where the action S is defined by

$$S(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) := \int_0^\beta \left[\boldsymbol{\xi}(\tau)^* (\partial_\tau - \mu) \boldsymbol{\xi}(\tau) + H(\overline{\boldsymbol{\xi}}(\tau), \boldsymbol{\xi}(\tau))\right]\, d\tau$$

If we write $\hat{H}$ as a sum of a noninteracting part $\hat{H}_0$ and an interaction $\hat{U} = U(\mathbf{c}^\dagger, \mathbf{c})$, i.e.,

$$\hat{H} = \hat{H}_0 + \hat{U} = \sum_{p,q} h_{pq} c_p^\dagger c_q + \hat{U},$$

then we can write

$$S(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) = S_0(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) + S_{\text{int}}(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}),$$

where

$$S_0(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) := \int_0^\beta \boldsymbol{\xi}(\tau)^* (\partial_\tau + h - \mu) \boldsymbol{\xi}(\tau)\, d\tau, \quad S_{\text{int}}(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) = \int_0^\beta U(\overline{\boldsymbol{\xi}}(\tau), \boldsymbol{\xi}(\tau)).$$

Again one can observe the formal similarity of the path integral to the Euclidean field theory presented in section 1. However, the analogy is even more restricted here for obvious reasons.

This concludes our discussion of the coherent-state path integral. The use of this construction in this dissertation is limited to section 7.1 below, where we use it to motivate the connection between Green's functions in the Euclidean lattice field theory (which will be key in Parts II, III, IV, and VI) and Green's functions in fermionic and bosonic statistical mechanics (which will be key in Parts VI and VII).

7 Green's functions

7.1 Motivation via functional derivatives

Before we proceed with standard definitions to many-body Green's functions, we first offer some motivating discussion from a more general perspective. In this section we will consider $d = |\mathcal{B}| < \infty$.

In the setting of Euclidean field theory, our notion of the 'Green's function' associated to a Gibbs meausure $d\mu(x) = e^{-H(x)}\,dx$ is simply the two-point correlator

$$G = \int_{\mathbb{R}^N} xx^\top d\mu(x).$$

Note that for the choice $H_A(x)$ of (1.1), by defining the free energy

$$\Omega[A] := -\log Z[A] = \int_{\mathbb{R}^N} e^{-\frac{1}{2}x^T Ax - U(x)}\,dx,$$

we can write $G = G[A]$ as a gradient via $G[A] = -\nabla_A \Omega[A]$, where we define $\nabla_A := \left(\frac{\partial}{\partial A_{ij}} + \frac{\partial}{\partial A_{ji}}\right)$. We comment that the self-energy is defined as the difference $\Sigma = A - G^{-1}$ so that $\Sigma = 0$ if $U \equiv 0$ (i.e., in the noninteracting case). Further detail is provided in Part III, which views this relation as the foundation of the so-called Luttinger-Ward formalism.

For now, let us analogize this construction to the setting of the coherent state path integral. Note that the discussion will be only informal, with rigorous definitions to follow later. For concreteness, we will stick to the fermionic case. There, the 'quadratic part' of the action (i.e., the analogy of $\frac{1}{2}x^T Ax$ in the Euclidean setting) is

$$S_0(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) = \int_0^\beta \boldsymbol{\xi}(\tau)^*(\partial_\tau + h - \mu)\boldsymbol{\xi}(\tau)\,d\tau.$$

We can extend this particular action to the broadest possible parametric class of quadratic actions as

$$S_0[A](\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) := \int_0^\beta \boldsymbol{\xi}(\tau)^*\partial_\tau \boldsymbol{\xi}(\tau)\,d\tau + \int_0^\beta \int_0^\beta \boldsymbol{\xi}(\tau')^* A(\tau', \tau)\boldsymbol{\xi}(\tau)\,d\tau\,d\tau',$$

so that the action map $(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) \mapsto S_0[A](\overline{\boldsymbol{\xi}}, \boldsymbol{\xi})$ is itself a functional of the Hermitian-operator-valued kernel $(\tau', \tau) \mapsto A(\tau', \tau)$.

Then by considering the partition function $Z = Z[A]$ as a functional of the kernel A and defining $\Omega[A] = \log Z[A]$, we may in turn define a Green's function via

$$\begin{aligned} G(\tau, \tau')[A] : &= \frac{\delta \Omega}{\delta A(\tau', \tau)}[A] \\ &= \frac{-1}{Z[A]} \int D_{\text{a-per}}\left[\overline{\boldsymbol{\xi}}(\,\cdot\,), \boldsymbol{\xi}(\,\cdot\,)\right] \boldsymbol{\xi}(\tau)\boldsymbol{\xi}(\tau')^*\, e^{-S_0[A](\overline{\boldsymbol{\xi}}, \boldsymbol{\xi}) - S_{\text{int}}(\overline{\boldsymbol{\xi}}, \boldsymbol{\xi})}. \end{aligned}$$

By evaluating at a kernel of the form $A(\tau', \tau) = (h - \mu)\delta(\tau' - \tau)$ and *reversing the steps* of the derivation of the path integral, we find that

$$G_{ij}(\tau, \tau') = \frac{-1}{Z}\text{Tr}\left[\mathcal{T}\left\{a_i(\tau)a_j^\dagger(\tau')\right\} e^{-\beta(\hat{H} - \mu\hat{N})}\right],$$

where $\mathcal{T}\left\{a_i(\tau)a_j^\dagger(\tau')\right\}$ indicates the imaginary-time-ordering operator, formally defined by

$$\mathcal{T}\left\{a_i(\tau)a_j^\dagger(\tau')\right\} = \begin{cases} a_i(\tau)a_j^\dagger(\tau'), & \tau' < \tau \\ -a_j^\dagger(\tau')a_i(\tau), & \tau' \geq \tau. \end{cases}$$

In fact this matches the definition of the *Matsubara Green's function* to be given below.

7.2 Green's functions and the self-energy at zero temperature

For $t \in \mathbb{R}$, we denote the annihilation and creation operators in the Heisenberg representation by

$$a_i(t) := e^{i\hat{H}t}a_i e^{-i\hat{H}t}, \quad a_i^\dagger(t) := e^{i\hat{H}t}a_i^\dagger e^{-i\hat{H}t}.$$

Then for a zero-temperature ensemble with N particles, the *time-ordered, single-body, real-time Green's function* (which we call the *Green's function* for short) is a function $G : \mathbb{R} \times \mathbb{R} \to \mathbb{C}^{d\times d}$ defined by

$$G_{ij}(t,t') = -i\left\langle \Psi_0^{(N)} \right| \mathcal{T}\left\{a_i(t)a_i^\dagger(t')\right\} \left| \Psi_0^{(N)} \right\rangle,$$

where $\mathcal{T}$ is the *time-ordering operator*, formally defined by

$$\mathcal{T}\left\{a_i(t)a_i^\dagger(t')\right\} = \begin{cases} a_i(t)a_j^\dagger(t'), & t' < t \\ \zeta a_j^\dagger(t')a_i(t), & t' \geq t. \end{cases}$$

Note that $\mathcal{T}$ is not really an operator and it is interpreted merely via the *symbolic* content of its argument.

We can write

$$G(t,t') = G^+(t,t') + G^-(t,t'),$$

where

$$\begin{aligned} iG^+(t,t') &:= \left\langle \Psi_0^{(N)} \middle| a_i(t)a_j^\dagger(t') \middle| \Psi_0^{(N)} \right\rangle \theta(t-t'), \\ iG^-(t,t') &:= \left\langle \Psi_0^{(N)} \middle| a_j^\dagger(t')a_i(t) \middle| \Psi_0^{(N)} \right\rangle (1-\theta(t-t')), \end{aligned}$$

with

$$\theta(s) := \begin{cases} 1, & s > 0 \\ 0, & s \leq 0. \end{cases}$$

It is easy to show that $G(t,t')$, $G^+(t,t')$, and $G^-(t,t')$ depend only on $t-t'$, so we can define $G(t) := G(t,0)$, $G^+(t) := G^+(t,0)$, and $G^-(t) := G^-(t,0)$ and consider these objects without any loss of information. It is then equivalent to consider the Fourier transforms

$$G(\omega) := \int_{\mathbb{R}} G(t)e^{i\omega t - \eta|t|}\,dt$$

and likewise $G^+(\omega)$ and $G^-(\omega)$ defined similarly, so

$$G(\omega) = G^+(\omega) + G^-(\omega).$$

Here η is interpreted as a positive, infinitesimally small quantity needed to ensure the convergence of the relevant integrals, and $G(\omega)$, $G^+(\omega)$, and $G^-(\omega)$ are not really functions, but rather distributions on $\mathbb{R}$ defined via the limit $\eta \to 0^+$.

One can show that

$$G^+_{ij}(\omega) = \big\langle \Psi_0^{(N)} \big| a_i \frac{1}{\omega - (\hat{H} - E_0^{(N)}) + i\eta} a_j^\dagger \big| \Psi_0^{(N)} \big\rangle$$

and

$$G^-_{ij}(\omega) = -\zeta \big\langle \Psi_0^{(N)} \big| a_j^\dagger \frac{1}{\omega + (\hat{H} - E_0^{(N)}) - i\eta} a_i \big| \Psi_0^{(N)} \big\rangle,$$

where $E_0^{(N)}$ is the energy of the N-particle ground state, i.e., $\hat{H}\big|\Psi_0^{(N)}\big\rangle = E_0\big|\Psi_0^{(N)}\big\rangle$.

Now we can think of $G^\pm$ as the restriction to the real axis of the rational function $G^\pm : \mathbb{C} \to \mathbb{C}^{d\times d}$ defined by

$$\begin{aligned} G^+_{ij}(z) &:= \big\langle \Psi_0^{(N)} \big| a_i \frac{1}{z - (\hat{H} - E_0^{(N)})} a_j^\dagger \big| \Psi_0^{(N)} \big\rangle \\ G^-_{ij}(z) &:= -\zeta \big\langle \Psi_0^{(N)} \big| a_j^\dagger \frac{1}{z + (\hat{H} - E_0^{(N)})} a_i \big| \Psi_0^{(N)} \big\rangle, \end{aligned}$$

and we can define $G(z) := G^+(z) + G^-(z)$ accordingly to be rational on $\mathbb{C}$.

Note that here we have left out the $\pm i\eta$ in the denominators, which specified whether poles should be viewed as being infinitesimally above or below the real axis. This erases the distinction between the time-ordered Green's function and the advanced and retarded Green's functions, which we do not define here, though see [77] for details. In fact the distinction does not matter for our sparsity results, which applies equally well in all of these cases.

The self-energy is the rational function $\Sigma : \mathbb{C} \to \mathbb{C}^{d\times d}$ defined by

$$\Sigma(z) := z - h - G(z)^{-1}.$$

7.3 Green's functions and the self-energy at finite temperature

As above, for $t \in \mathbb{R}$, we denote the annihilation and creation operators in the Heisenberg representation by

$$a_i(t) := e^{i\hat{H}t} a_i e^{-i\hat{H}t}, \quad a_i^\dagger(t) := e^{i\hat{H}t} a_i^\dagger e^{-i\hat{H}t}.$$

Then at finite inverse temperature $\beta \in (0, \infty)$ and chemical potential $\mu \in \operatorname{int}\operatorname{dom} Z$, the *time-ordered, single-body, real-time Green's function* (which we call the *Green's*

function for short when the context is clear) is a function $G : \mathbb{R} \times \mathbb{R} \to \mathbb{C}^{d\times d}$ defined by

$$G_{ij}(t,t') = -i \left\langle \mathcal{T}\{a_i(t)a_i^\dagger(t')\} \right\rangle_{\beta,\mu}.$$

We can write

$$G(t,t') = G^+(t,t') + G^-(t,t'),$$

where

$$iG^+(t,t') = \frac{1}{Z}\mathrm{Tr}\left[a_i(t)a_j^\dagger(t')e^{-\beta(\hat{H}-\mu\hat{N})}\right]\theta(t-t'),$$
$$iG^-(t,t') = \frac{\zeta}{Z}\mathrm{Tr}\left[a_j^\dagger(t')a_i(t)e^{-\beta(\hat{H}-\mu\hat{N})}\right](1-\theta(t-t')),$$

with

$$\theta(s) := \begin{cases} 1, & s > 0 \\ 0, & s \le 0. \end{cases}$$

as above.

Once again it is easy to show that $G(t,t')$, $G^+(t,t')$, and $G^-(t,t')$ depend only on $t-t'$, so we can define $G(t) := G(t,0)$, $G^+(t) := G^+(t,0)$, and $G^-(t) := G^-(t,0)$ and consider these objects without any loss of information. It is then equivalent to consider the Fourier transforms

$$G(\omega) := \int_{\mathbb{R}} G(t)e^{i\omega t - \eta|t|}\,dt$$

and likewise $G^+(\omega)$ and $G^-(\omega)$ defined similarly, so

$$G(z) = G^+(\omega) + G^-(\omega).$$

Now since $\hat{H}$ preserves particle number, we can safely diagonalize $\hat{H}$ as an operator on each of the N-particle subspaces separately. Then the spectrum of $\hat{H}$ consists of the union of its spectra on the N-particle subspaces. It follows from Assumption 4 that $\hat{H} - \mu\hat{N}$ has a ground state, i.e., that its spectrum is bounded from below, for $\mu \in \operatorname{int}\operatorname{dom} Z$. Let $m = 0, 1, \ldots,$ (terminating at $m = 2^d$ in the case of fermions) index the spectrum of $\hat{H}$, and let $|\Psi_m\rangle$ denote the m-th eigenstate. Let N_m be the particle number of $|\Psi_m\rangle$ (which is an eigenstate of $\hat{N}$), and let E_m be defined by $\hat{H}|\Psi_m\rangle = E_m|\Psi_m\rangle$.

One can show that

$$G_{ij}^+(\omega) = \frac{1}{Z}\sum_m e^{-\beta(E_m-\mu N_m)}\big\langle\Psi_m\big|a_i\frac{1}{\omega-(\hat{H}-E_m)+i\eta}a_j^\dagger\big|\Psi_m\big\rangle$$

and

$$G_{ij}^-(\omega) = \frac{-\zeta}{Z}\sum_m e^{-\beta(E_m-\mu N_m)}\big\langle\Psi_m\big|a_j^\dagger\frac{1}{\omega+(\hat{H}-E_m)-i\eta}a_i\big|\Psi_m\big\rangle.$$

Recall that

$$Z = \sum_m e^{-\beta(E_m - \mu N_m)}.$$

Now we can think of $G^{\pm}$ as the restriction to the real axis of the rational function $G^{\pm} : \mathbb{C} \to \mathbb{C}^{d\times d}$ defined by

$$G^+_{ij}(z) := \frac{1}{Z}\sum_m e^{-\beta(E_m-\mu N_m)}\big\langle \Psi_m \big| a_i \frac{1}{z - (\hat{H} - E_m)} a_j^\dagger \big| \Psi_m \big\rangle$$
$$G^-_{ij}(z) := \frac{-\zeta}{Z}\sum_m e^{-\beta(E_m-\mu N_m)}\big\langle \Psi_m \big| a_j^\dagger \frac{1}{z + (\hat{H} - E_m)} a_i \big| \Psi_m \big\rangle,$$

and we can define $G(z) := G^+(z) + G^-(z)$ accordingly to be rational on $\mathbb{C}$. Once again we have ignored the infinitesimal η in this definition; the same comments made in Appendix 5.1 apply here.

The self-energy is the rational function $\Sigma : \mathbb{C} \to \mathbb{C}^{d\times d}$ defined by

$$\Sigma(z) := z - h - G(z)^{-1}.$$

7.4 Matsubara Green's functions and self-energy

For $\tau \in \mathbb{R}$, we define (abusing notation)

$$a_i(\tau) := e^{(\hat{H}-\mu\hat{N})\tau} a_i e^{-(\hat{H}-\mu\hat{N})\tau}, \quad a_i^\dagger(\tau) := e^{(\hat{H}-\mu\hat{N})\tau} a_i^\dagger e^{-(\hat{H}-\mu\hat{N})\tau}.$$

Although we have overloaded the notation, the distinction between, e.g., $a_i(\tau)$ and $a_i(t)$ should be clear from context. Note carefully that $a_i^\dagger(\tau)$ is *not* the adjoint of $a_i(\tau)$. This is merely a notation. The operators $a_i(\tau)$ and $a_i^\dagger(\tau)$ can be thought of as the imaginary-time Heisenberg representation of the annihilation and creation operators. Although the analogy with the real-time Heisenberg representation is broken by considering $\hat{H} - \mu\hat{N}$ in place of $\hat{H}$, our convention is indeed the more widely used due to its naturality in the context of the imaginary-time path integral.

Then at finite inverse temperature $\beta \in (0, \infty)$ and chemical potential $\mu \in \operatorname{int}\operatorname{dom} Z$, the *time-ordered, single-body, imaginary-time Green's function* (which we call the *Matsubara Green's function* for clarity) is a function $G^{\mathrm{M}} : [0, \beta]^2 \to \mathbb{C}^{d\times d}$ defined by

$$G^{\mathrm{M}}_{ij}(\tau, \tau') = -\big\langle\, \mathcal{T}\big\{ a_i(\tau) a_i^\dagger(\tau') \big\}\, \big\rangle_{\beta,\mu},$$

where $\mathcal{T}$ here indicates the imaginary-time-ordering operator, formally defined by

$$\mathcal{T}\big\{ a_i(\tau) a_i^\dagger(\tau') \big\} = \begin{cases} a_i(\tau) a_j^\dagger(\tau'), & \tau' < \tau \\ \zeta a_j^\dagger(\tau') a_i(\tau), & \tau' \geq \tau. \end{cases}$$

We can write

$$G^{\mathrm{M}}(\tau, \tau') = G^{\mathrm{M},+}(\tau, \tau') + G^{\mathrm{M},-}(\tau, \tau'),$$

where

$$\begin{aligned}-G_{ij}^{\mathrm{M},+}(\tau,\tau') &= \frac{1}{Z}\mathrm{Tr}\left[a_i(\tau)a_j^\dagger(\tau')e^{-\beta(\hat{H}-\mu\hat{N})}\right]\theta(\tau-\tau'),\\ &= \frac{1}{Z}\mathrm{Tr}\left[a_i e^{-(\hat{H}-\mu\hat{N})(\tau-\tau')}a_j^\dagger e^{(\tau-\tau'-\beta)(\hat{H}-\mu\hat{N})}\right]\theta(\tau-\tau'),\end{aligned}$$

and

$$\begin{aligned}-G_{ij}^{\mathrm{M},-}(\tau,\tau') &= \frac{\zeta}{Z}\mathrm{Tr}\left[a_j^\dagger(\tau')a_i(\tau)e^{-\beta(\hat{H}-\mu\hat{N})}\right](1-\theta(\tau-\tau')).\\ &= \frac{\zeta}{Z}\mathrm{Tr}\left[a_j^\dagger e^{-(\hat{H}-\mu\hat{N})(\tau'-\tau)}a_i e^{(\tau'-\tau-\beta)(\hat{H}-\mu\hat{N})}\right](1-\theta(\tau-\tau')).\end{aligned}$$

Once again it is easy to show that $G^{\mathrm{M}}(\tau,\tau')$, $G^{\mathrm{M},+}(\tau,\tau')$, and $G^{\mathrm{M},-}(\tau,\tau')$ depend only on $\tau-\tau'$. Then the full information of $G^{\mathrm{M}}(\tau,\tau')$ can be recovered from

$$G^{\mathrm{M}}(\tau) := \begin{cases} G^{\mathrm{M},+}(\tau,0), & \tau > 0\\ G^{\mathrm{M},-}(0,-\tau), & \tau \le 0,\end{cases}$$

defined for $\tau \in (-\beta,\beta)$. Now for $\tau \in (0,\beta)$, we can compute via the above formulas:

$$G^{\mathrm{M}}(\tau-\beta) = G^{\mathrm{M},-}(0,\beta-\tau) = \zeta G^{\mathrm{M},+}(0,\tau) = \zeta G^{\mathrm{M}}(\tau).$$

Therefore, by considering $G^{\mathrm{M}}(\tau) = G^{\mathrm{M},+}(\tau,0)$ only on $(0,\beta)$, i.e.,

$$G^{\mathrm{M}}(\tau) = \frac{-1}{Z}\mathrm{Tr}\left[a_i e^{-\tau(\hat{H}-\mu\hat{N})}a_j^\dagger e^{(\tau-\beta)(\hat{H}-\mu\hat{N})}\right], \quad \tau\in(0,\beta)$$

and extending by β-(anti)periodicity, we can recover the full information of the Matsubara Green's function.

It is then equivalent to consider the frequency-space representation of at the Matsubara frequencies

$$\omega_n = \begin{cases} 2n\pi/\beta, & \zeta = +1\\ (2n+1)\pi/\beta, & \zeta = -1\end{cases}$$

for $n \in \mathbb{Z}$, defined via

$$G^{\mathrm{M}}(i\omega_n) := \int_0^\beta G^{\mathrm{M}}(\tau)e^{i\omega_n\tau}\,d\tau,$$

so

$$G^{\mathrm{M}}(\tau) = \frac{1}{\beta}\sum_n G^{\mathrm{M}}(i\omega_n)e^{-i\omega_n\tau}.$$

One can show that

$$G^{\mathrm{M}}(i\omega_n) = G(i\omega_n+\mu),$$

where G is the rational function $\mathbb{C}\to\mathbb{C}^{d\times d}$ defined in the preceding subsection.

The Matsubara self-energy is defined by

$$\Sigma^{\mathrm{M}}(i\omega_n) = i\omega_n - (h-\mu) - G^{\mathrm{M}}(i\omega_n)^{-1} = \Sigma(i\omega_n + \mu),$$

where Σ is the rational function $\mathbb{C} \to \mathbb{C}^{d\times d}$ defined in the preceding subsection. Thus to study the Matsubara Green's function and self-energy it suffices to study the rational functions G and Σ defined earlier.

Finally, we comment that in the imaginary-time representation, we can write

$$-\partial_\tau G^{\mathrm{M}}(\tau,\tau') - (h-\mu)G^{\mathrm{M}}(\tau,\tau') - \int_0^\beta \Sigma^{\mathrm{M}}(\tau,\tau'')G^{\mathrm{M}}(\tau'',\tau')\,d\tau'' = I_d\delta(\tau-\tau').$$