7

Determinant method

This chapter introduces a finite-temperature algorithm for the simulation of interacting electrons on a lattice. Because this algorithm was developed by Blankenbecler, Scalapino, and Sugar (1981; Scalapino and Sugar, 1981), it is sometimes called the *BSS algorithm*. The method uses a Hubbard-Stratonovich transformation to convert the interacting electron problem into a noninteracting one coupled to an imaginary-time-dependent auxiliary field. For this reason, it is also called the *auxiliary-field method*. We use here yet another name, the *determinant method*, which is fitting because the transformation to a problem of noninteracting electrons generates determinants as the statistical weights. The finite-temperature determinant algorithm is a general-purpose electron algorithm that enables computations of a wide variety of local observables and correlation functions. For a discussion of a zero-temperature determinant method, refer to Appendix I.

7.1 Theoretical framework

Feynman and Hibbs (1965) formulated quantum mechanics in terms of integrals over all paths in configuration space. In real time, each path contributes a phase to the integral that is determined by the classical action along the path. Two paths can interfere constructively or destructively. In the classical limit, only the stationary-phase path is important. Being characterized by many interfering paths, real-time quantum dynamics more than challenges importance sampling. Statistical mechanics, on the other hand, involves path integrals in imaginary time. Contributions to the integrals vary exponentially in magnitude but not in phase. Thus, the path integral is dominated by paths of large magnitude. The tasks of a quantum Monte Carlo method are identifying these important paths and sampling them efficiently.

In this chapter, we address the classicization of many-electron problems at finite temperatures via a Feynman path integral. The result is a method often called the *determinant method* as the weights of the paths can be expressed as determinants, hardly classical-looking weights, but ones quite suggestive of the antisymmetry of Fermion states. Sampling these weights efficiently and in a stable manner requires special techniques. We begin with a brief overview to motivate the general form of the classical representation and the weights we need to sample.

7.1.1 Hubbard-Stratonovich transformations

As noted in Chapter 5, most quantum Monte Carlo algorithms treat the Boltzmann operator $\exp(-\beta H)$ by writing its matrix elements as an imaginary-time path integral

$$\langle \psi_L | \, e^{-\beta H} \, | \psi_R \rangle = \sum_{\psi_1, \psi_2, \ldots, \psi_{N-1}} \langle \psi_L | \, e^{-\Delta\tau H} \, | \psi_{N-1} \rangle \cdots \langle \psi_2 | \, e^{-\Delta\tau H} \, | \psi_1 \rangle \, \langle \psi_1 | \, e^{-\Delta\tau H} \, | \psi_R \rangle \,, \quad (7.1)$$

where $\Delta\tau = \beta / N_\tau$. For finite-temperature statistical mechanics, the partition function

$$Z = \mathrm{Tr} \left[e^{-\beta H} \right] = \sum_{\psi} \langle \psi | \, e^{-\beta H} \, | \psi \rangle \quad (7.2)$$

thus becomes the integral over all paths that are periodic in imaginary time.

In Chapter 5, we also learned that evaluating the matrix elements of the exponentials of the Hamiltonian is aided by decomposing the Hamiltonian into its individual terms, $H = H_1 + H_2 + \cdots + H_n$, and using the Trotter approximation (5.12) so that

$$e^{-\Delta\tau H} = e^{-\Delta\tau H_1} e^{-\Delta\tau H_2} \cdots e^{-\Delta\tau H_n} + \mathcal{O}((\Delta\tau)^2). \quad (7.3)$$

The matrix elements of these exponential factors are then expressed in some convenient basis where the exponential operators are either known or easily approximated. We now add the extra requirement that the decomposition is the first step toward transforming the problem into one where the only matrix elements needed are those of a noninteracting problem. To achieve this, we single out the terms in the decomposition that describe direct interactions between the electrons.

As an illustrative example, we consider an impurity Hamiltonian where there is a Coulomb interaction acting between a pair of up and down electrons that simultaneously occupy the impurity orbital.[1] We write

[1] For a more detailed discussion of the impurity models typically studied in condensed matter physics, see Section 8.1.

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

$$H = H_1 + U\left(n_\uparrow - \tfrac{1}{2}\right)\left(n_\downarrow - \tfrac{1}{2}\right), \tag{7.4}$$

where H_1 is a Hamiltonian for some noninteracting problem, that is, it is a Hamiltonian quadratic in the electron creation and destruction operators. More deeply, this statement also implies that $H_1 = \sum_\sigma H_1^\sigma$ and that the different H_1^σ operators commute. Via the Trotter approximation (5.12),

$$e^{-\Delta\tau H} \approx e^{-\Delta\tau H_1} e^{-\Delta\tau U\left(n_\uparrow - \frac{1}{2}\right)\left(n_\downarrow - \frac{1}{2}\right)}, \tag{7.5}$$

where n_σ is the electron number operator for electron spin $\sigma = \uparrow$ or $\downarrow$.

Having separated the interaction term from the rest, we now transform its exponential into an integral over a set of auxiliary fields. If the Coulomb interaction U is repulsive, we write the argument of the exponential (7.5) as

$$\left(n_\uparrow - \tfrac{1}{2}\right)\left(n_\downarrow - \tfrac{1}{2}\right) = -\tfrac{1}{2}\left(n_\uparrow - n_\downarrow\right)^2 + \tfrac{1}{4}.$$

Then, we make the following *Hubbard-Stratonovich transformation* (Fetter and Walecka, 1971; Negele and Orland, 1988; Fulde, 1991; Enz, 1992)

$$\begin{aligned} e^{-\Delta\tau U\left(n_\uparrow - \frac{1}{2}\right)\left(n_\downarrow - \frac{1}{2}\right)} &= e^{-\frac{1}{4}\Delta\tau U} e^{\frac{1}{2}\Delta\tau U\left(n_\uparrow - n_\downarrow\right)^2} \\ &= e^{-\frac{1}{4}\Delta\tau U} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx\, e^{-\frac{1}{2}x^2 + x\sqrt{\Delta\tau U}\left(n_\uparrow - n_\downarrow\right)}, \end{aligned} \tag{7.6}$$

which introduces the auxiliary field x into the analysis. The Hubbard-Stratonovich transformation is merely the operator generalization of a shifted Gaussian integral[2]

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dx\, e^{-\frac{1}{2}(x \pm y)^2} = 1 \quad \Rightarrow \quad e^{\frac{1}{2}y^2} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dx\, e^{-\frac{1}{2}x^2 \mp xy}. \tag{7.7}$$

The result of the transformation is a Gaussian-weighted integral over the auxiliary-field variable, where the integrand is the exponential of an operator proportional to the auxiliary field and quadratic in the electron creation and destruction operators ($n_\sigma = c_\sigma^\dagger c_\sigma$). Because it is quadratic, it does not explicitly express an interaction among the electrons. In brief, the transformation has replaced the repulsive Coulomb interaction in the impurity orbital by a scalar (Bosonic) field coupled to the net spin $n_\uparrow - n_\downarrow$ of the impurity orbital. After the transformation, the effect of a positive value of U is the creation of a fluctuating spin moment at this orbital. We chose the form of the transformation such that the auxiliary field is always real. If we had not done so, we would have produced a complex-valued integrand.

[2] Since the rest of the integral is an even function in x, the choice of the sign in the argument of the one exponential is arbitrary. We choose the plus sign.

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

This result would be problematic, because we want to interpret the integrand as a probability density.

If U is attractive, we instead write

$$\left(n_{i\uparrow} - \tfrac{1}{2}\right)\left(n_{i\downarrow} - \tfrac{1}{2}\right) = \tfrac{1}{2}\left(n_{i\uparrow} + n_{i\downarrow} - 1\right)^2 - \tfrac{1}{4},$$

leading to

$$e^{-\Delta\tau U\left(n_\uparrow - \frac{1}{2}\right)\left(n_\downarrow - \frac{1}{2}\right)} = e^{+\frac{1}{4}\Delta\tau U}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} dx\, e^{-\frac{1}{2}x^2 + x\sqrt{\Delta\tau|U|}\left(n_\uparrow + n_\downarrow - 1\right)}. \tag{7.8}$$

Here, the auxiliary field couples to a charge degree of freedom. After such a transformation, the effect of a negative value of U is to create charge fluctuations in the correlated orbital.

There is an alternative form of the transformation due to Hirsch (1983), which is especially useful for lattice models that have the Coulomb interaction given as a sum of products of number operators. Hirsch's transformation introduces a *discrete auxiliary field* that in practice can be sampled more efficiently than the continuous field. His transformation is

$$e^{-\Delta\tau U\left(n_\uparrow - \frac{1}{2}\right)\left(n_\downarrow - \frac{1}{2}\right)} = \begin{cases} \frac{1}{2}e^{-\frac{1}{4}\Delta\tau U}\sum\limits_{x=\pm 1} e^{\alpha x\left(n_\uparrow - n_\downarrow\right)}, & U > 0, \\ \frac{1}{2}e^{+\frac{1}{4}\Delta\tau U}\sum\limits_{x=\pm 1} e^{\alpha x\left(n_\uparrow + n_\downarrow - 1\right)}, & U < 0, \end{cases} \tag{7.9}$$

with $\cosh\alpha = \exp(\Delta\tau|U|/2)$. Again, for positive U, the auxiliary field couples to a spin degree of freedom, while for negative U, it couples to the charge. Additional Hubbard-Stratonovich transformations are discussed in Appendix G.

Whether continuous or discrete auxiliary fields are introduced, the partition function becomes a nested sum

$$Z = \sum_C \sum_\psi w_C[\psi], \tag{7.10}$$

with $C = \{x_1, x_2, \ldots, x_{N_\tau}\}$ denoting the set of all possible auxiliary-field configurations as a function of imaginary time and

$$w_C[\psi] \equiv \langle\psi|\, e^{-\beta H}\, |\psi\rangle\,, \tag{7.11}$$

with $e^{-\beta H}$ being discretized by the Trotter approximation (7.3). The transformation breaks the interaction for each time-slice into a sum of noninteracting terms, one for each spin, and thus generates an effective noninteracting Hamiltonian. As each

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

spin piece of this Hamiltonian commutes, its exponential factors into products of exponentials, one set of products for each spin.

The states $|\psi\rangle$ in the trace are generally an occupation number basis. In an occupation number basis, each basis state is a direct product

$$\left|n_{1\uparrow}, n_{2\uparrow}, \ldots, n_{N\uparrow}\right\rangle \left|n_{1\downarrow}, n_{2\downarrow}, \ldots, n_{N\downarrow}\right\rangle$$

of a spin-up and a spin-down state, where $n_{i\sigma} = 0$ or 1, $\sum_i n_{i\sigma} = N_\sigma$, and

$$|n_{1\sigma}, n_{2\sigma}, \ldots, n_{N\sigma}\rangle = c^\dagger_{i_{N_\sigma}} \cdots c^\dagger_{i_2} c^\dagger_{i_1} |0\rangle .$$

For notational compactness, we write

$$\left|\psi_{N_\sigma}\right\rangle = |n_{1\sigma}, n_{2\sigma}, \ldots, n_{N\sigma}\rangle$$

to symbolize a basis state with N_σ electrons and use

$$\left|\psi_{N_{\text{elec}}}\right\rangle = \left|n_{1\uparrow}, n_{2\uparrow}, \ldots, n_{N\uparrow}\right\rangle \left|n_{1\downarrow}, n_{2\downarrow}, \ldots, n_{N\downarrow}\right\rangle = \left|\psi_{\text{N}\uparrow}\right\rangle \left|\psi_{\text{N}\downarrow}\right\rangle$$

to symbolize direct product states with $N_{\text{elec}} = \sum_\sigma N_\sigma$. We see that in this basis the spin states also factorize. Ultimately,

$$Z = \sum_C \sum_{N_\downarrow + N_\uparrow = N_{\text{elec}}} \prod_\sigma \sum_{\psi_{N\sigma}} w^\sigma_C[\psi_{N\sigma}]. \tag{7.12}$$

It is important to note that the spin factors in the product of the weights share the auxiliary-field configuration. This sharing is what reconstructs the interactions. Although the impurity problem discussed here is particularly simple, the determinant method expresses more general and complex Hamiltonians in the same form.

By placing a term that is quadratic in the electron creation and destruction operators in the argument of the exponential, the Hubbard-Stratonovich transformation enables us to trace out the electronic degrees of freedom for a fixed configuration of auxiliary-field variables. In other words, it transforms the given problem into one we know how to solve. If we were explicitly to trace out the electron degrees of freedom, we would find

$$Z = \sum_C \prod_\sigma w^\sigma_C. \tag{7.13}$$

with the new weights, denoted as w^σ_C with no argument, expressible as determinants. It is these weights we need to specify.

While a stepping stone, this result is not what we need. The starting partition function (7.2) is in the canonical ensemble that fixes the number of electrons.

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

It is more convenient to use the grand canonical ensemble for electrons at finite temperature,

$$Z = \mathrm{Tr}\left[e^{-\beta(H-\mu N_{\mathrm{elec}})}\right] = \sum_{N_{\mathrm{elec}}} \sum_{\psi_{N_{\mathrm{elec}}}} \left\langle \psi_{N_{\mathrm{elec}}} \right| e^{-\beta(H-\mu N_{\mathrm{elec}})} \left| \psi_{N_{\mathrm{elec}}} \right\rangle, \tag{7.14}$$

where we have to sum over all possible electron numbers. Clearly, we can repeat the analysis just performed for each electron number, finding for each a determinantal weight. As we will also see, the sum of these determinants is another determinant, a determinant representation of the Fermi-Dirac partition function.

In the next subsection, we derive this grand canonical determinantal weight. Central to this development is the imaginary-time single-electron propagator. With it, we are able to construct the imaginary-time many-electron propagator, and with this propagator, we in turn can construct imaginary-time single-particle Green's functions. This construction is important: After the next subsection, we begin our presentation of the determinant method where we adopt a Metropolis sampling of auxiliary-field configurations. This sampling requires computing the ratio of two determinants to decide whether to accept or reject proposed changes in the auxiliary fields. To do the sampling efficiently, we need to use special methods that require the imaginary-time, single-particle Green's function. Having this Green's function is extra handy as it also facilitates computing many-body correlation functions. One strength of the determinant method is the ease with which a large variety of such functions can be computed.

7.1.2 Determinantal weights

From the path-integral perspective, the matrix elements of the Boltzmann factor in (7.1) are formed from a sequence of matrix multiplications where each matrix represents the imaginary-time evolution from one time step to another. The Trotter approximation (7.3) and the Hubbard-Stratonovich transformation (7.6), (7.8), or (7.9) reduce the complexity of this propagation to that of the propagation of free electron states. Central to the numerical framework of the determinant method are matrices expressing the free propagation of multi-electron states from imaginary time τ_1 to τ_2 that are built upon a matrix B whose elements $B_{ij}(\tau_2, \tau_1)$ propagate a single electron from state j at time τ_1 to a state i at time τ_2. For a given spin σ and auxiliary-field configuration, this single-particle propagator is defined as

$$B_{ij}(\tau_2, \tau_1) = \left\langle 0 \left| c_i \left[\mathcal{T} \exp\left(- \int_{\tau_1}^{\tau_2} H(\tau) d\tau \right) \right] c_j^{\dagger} \right| 0 \right\rangle, \tag{7.15}$$

where $\mathcal{T}$ is the time-ordering operator, $H(\tau)$ is the auxiliary-field-dependent noninteracting Hamiltonian ($\int d\tau H(\tau)$ is the Euclidean action), and $|0\rangle$ is the

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

vacuum state. This propagator represents the solution to the equations of motion for the electron creation and destruction operators[3]

$$c_i(\tau_2) = \sum_j B_{ij}(\tau_2, \tau_1) c_j(\tau_1),$$

$$c_i^\dagger(\tau_2) = \sum c_j^\dagger(\tau_1) B_{ji}^{-1}(\tau_2, \tau_1). \tag{7.16}$$

We note that $c_i(\tau)$ and $c_i^\dagger(\tau)$ are not Hermitian conjugates because we are working in imaginary time. This means the propagation is not unitary. As a consequence, orbitals that initially were orthonormal lose this property.

The description of an N_σ-electron system requires a multi-electron propagator

$$B(\tau_2, \tau_1) \equiv \left\langle 0 \left| c_{i_1} c_{i_2} \cdots c_{i_{N_\sigma}} \left[\mathcal{T} \exp\left(- \int_{\tau_1}^{\tau_2} H(\tau)\, d\tau \right) \right] c_{j_{N_\sigma}}^\dagger \cdots c_{j_2}^\dagger c_{j_1}^\dagger \right| 0 \right\rangle$$

$$= \det \begin{pmatrix} B_{i_1 j_1}(\tau_2, \tau_1) & B_{i_1 j_2}(\tau_2, \tau_1) & \cdots & B_{i_1 j_{N_\sigma}}(\tau_2, \tau_1) \\ B_{i_2 j_1}(\tau_2, \tau_1) & B_{i_2 j_2}(\tau_2, \tau_1) & \cdots & B_{i_2 j_{N_\sigma}}(\tau_2, \tau_1) \\ \vdots & \vdots & \ddots & \vdots \\ B_{i_{N_\sigma} j_1}(\tau_2, \tau_1) & B_{i_{N_\sigma} j_2}(\tau_2, \tau_1) & \cdots & B_{i_{N_\sigma} j_{N_\sigma}}(\tau_2, \tau_1) \end{pmatrix}, \tag{7.17}$$

which is the determinant of a matrix whose elements are single-particle propagators. We derive this identity in Appendix H. For the moment, we dwell on its physical meaning. Because the evolving electrons are independent, we should expect the many-body propagator to be a product of single-particle propagators. We inject N_σ identical particles at time τ_1 into the orbitals $j_1, j_2, \ldots, j_{N_\sigma}$, and we do not know which specific particles arrive at the orbitals $i_1, i_2, \ldots, i_{N_\sigma}$ at time τ_2. The determinant is the sum over all possible arrivals with the appropriate minus signs accounting for an even or odd number of exchanges of identical electrons.

Now we return to the expression for the many-electron partition function in the grand canonical ensemble. The Hubbard-Stratonovich transformation (7.9) adds a summation over all configurations C of auxiliary fields and together with the Trotter approximation (7.3) with $\Delta\tau = \beta/N_\tau$ generates a spin factorization

$$Z = \sum_C \prod_\sigma \sum_{N_\sigma} \sum_{\psi_{N_\sigma}} \left\langle \psi_{N_\sigma} \left| e^{-\Delta\tau c^\dagger M_{N_\tau}^\sigma c} \cdots e^{-\Delta\tau c^\dagger M_2^\sigma c} e^{-\Delta\tau c^\dagger M_1^\sigma c} \right| \psi_{N_\sigma} \right\rangle, \tag{7.18}$$

[3] The matrix B represents the exponential $\exp(-\Delta\tau H)$ where H is some noninteracting Hamiltonian. The matrix B^{-1} represents $\exp(\Delta\tau H)$. In Section 7.2.1, we discuss ways to construct B. Rather than performing a numerical inversion of B, we use the same strategy with $+\Delta\tau$ as we do with $-\Delta\tau$ to construct B^{-1}.

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

where M_i^σ is the matrix that defines the auxiliary-field-dependent noninteraction Hamiltonian for each spin (Appendix H). For the inner summations, repeated use of (7.17) yields

$$\sum_{N_\sigma}\sum_{\psi_{N_\sigma}}\left\langle\psi_{N_\sigma}\left|e^{-\Delta\tau c^\dagger M_\tau^\sigma c}\cdots e^{-\Delta\tau c^\dagger M_2^\sigma c}e^{-\Delta\tau c^\dagger M_1^\sigma c}\right|\psi_{N_\sigma}\right\rangle =$$
$$\det(1)+\det\left(B_{11}^\sigma\right)+\det\left(B_{22}^\sigma\right)+\cdots+\det\left(B_{NN}^\sigma\right)$$
$$+\det\begin{pmatrix}B_{11}^\sigma & B_{12}^\sigma\\ B_{21}^\sigma & B_{22}^\sigma\end{pmatrix}+\det\begin{pmatrix}B_{11}^\sigma & B_{13}^\sigma\\ B_{31}^\sigma & B_{33}^\sigma\end{pmatrix}$$
$$+\cdots+\det\begin{pmatrix}B_{N-1,N-1}^\sigma & B_{N-1,N}^\sigma\\ B_{N,N-1}^\sigma & B_{N,N}^\sigma\end{pmatrix}+\cdots$$
$$+\det\begin{pmatrix}B_{11}^\sigma & B_{12}^\sigma & \cdots & B_{1N}^\sigma\\ B_{21}^\sigma & B_{22}^\sigma & \cdots & B_{2N}^\sigma\\ \vdots & \vdots & \ddots & \vdots\\ B_{N1}^\sigma & B_{N2}^\sigma & \cdots & B_{NN}^\sigma\end{pmatrix}.$$

In the above, $B_{ij}^\sigma = B_{ij}^\sigma(\beta,0)$ (Appendix H). As is easily verified by direct expansion, the right-hand side equals

$$\det\left[I+B^\sigma(\beta,0)\right]=\det\begin{pmatrix}1+B_{11}^\sigma & B_{12}^\sigma & \cdots & B_{1N}^\sigma\\ B_{21}^\sigma & 1+B_{22}^\sigma & \cdots & B_{2N}^\sigma\\ \vdots & \vdots & \ddots & \vdots\\ B_{N1}^\sigma & B_{N2}^\sigma & \cdots & 1+B_{NN}^\sigma\end{pmatrix},$$

so the partition function (7.18) simplifies to

$$Z=\sum_C\prod_\sigma w_C^\sigma, \tag{7.19}$$

with

$$w_C^\sigma=\det\left[I+B^\sigma(\beta,0)\right]. \tag{7.20}$$

As promised, we have shown that the weights are determinants. We note that

$$B^\sigma(\beta,0)\equiv B^\sigma(\beta,\beta-\Delta\tau)\cdots B^\sigma(2\Delta\tau,\Delta\tau)B^\sigma(\Delta\tau,0) \tag{7.21}$$

is a string of matrix multiplications (Appendix H).

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

7.1.3 Single-particle Green's function

The standard many-body physics definition of the matrix elements of the single-particle equal-time Green's function (7.22) is[4]

$$G^{\sigma}_{ij}(\tau,\tau) \equiv \left\langle c_{i\sigma}(\tau)c^{\dagger}_{j\sigma}(\tau)\right\rangle = \delta_{ji} - \left\langle c^{\dagger}_{j\sigma}(\tau)c_{i\sigma}(\tau)\right\rangle, \tag{7.22}$$

where the expectation values are over the grand canonical ensemble. To connect this definition with (7.17), we start by coupling the action to the operator $c^{\dagger}_{j\sigma}c_{i\sigma}$ for an instant of imaginary time

$$\mathcal{T}\exp\left(-\int_0^{\beta} H(\tau)d\tau\right) \rightarrow \mathcal{T}\exp\left(-\int_{\tau}^{\beta} H(\tau)d\tau\right)\exp(hc^{\dagger}_{j\sigma}c_{i\sigma})\mathcal{T}\exp\left(-\int_0^{\tau} H(\tau)d\tau\right).$$

With the insertion of this interaction, (7.15) and (7.20) yield a modified weight,

$$w'_C = \det\left[I + B^{\sigma}(\beta,\tau)e^{hQ}B^{\sigma}(\tau,0)\right]\det\left[I + B^{-\sigma}(\beta,\tau)B^{-\sigma}(\tau,0)\right],$$

where Q is a null matrix except for $Q_{ji} = 1$. From this expression it follows that

$$\left\langle c^{\dagger}_{j\sigma}(\tau)c_{i\sigma}(\tau)\right\rangle = \left.\frac{\partial}{\partial h}\ln w'_C\right|_{h=0}.$$

We now use the identity

$$\ln\det A = \mathrm{Tr}\ln A, \tag{7.23}$$

and the cyclic property of a trace, $\mathrm{Tr}ABC = \mathrm{Tr}BCA = \mathrm{Tr}CAB$, to find

$$\begin{aligned}
\left.\frac{\partial}{\partial h}\ln w'_C\right|_{h=0} &= \mathrm{Tr}\frac{\partial}{\partial h}\ln\left[I + B^{\sigma}(\beta,\tau)e^{hQ}B^{\sigma}(\tau,0)\right]\Big|_{h=0} \\
&= \mathrm{Tr}\left[B^{\sigma}(\tau,0)\left(I + B^{\sigma}(\beta,\tau)B^{\sigma}(\tau,0)\right)^{-1}B^{\sigma}(\beta,\tau)Q\right] \\
&= \left[B^{\sigma}(\tau,0)\left(I + B^{\sigma}(\beta,\tau)B^{\sigma}(\tau,0)\right)^{-1}B^{\sigma}(\beta,\tau)\right]_{ji} \\
&= \left[\left(I + B^{\sigma}(\beta,\tau)^{-1}B^{\sigma}(\tau,0)^{-1}\right)^{-1}\right]_{ji} \\
&= \delta_{ji} - \left[\left(I + B^{\sigma}(\tau,0)B^{\sigma}(\beta,\tau)\right)^{-1}\right]_{ji}.
\end{aligned}$$

[4] Note the transposition of indices and the choice of sign.

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

Thus,

$$G^{\sigma}(\tau,\tau) = I - [I + B^{\sigma}(\tau,0)B^{\sigma}(\beta,\tau)]^{-1}. \tag{7.24}$$

This Green's function plays a central role in the determinant method, as it is used for two different key purposes. The first use is for sampling of the auxiliary fields. The variable $x_i(\tau)$ for orbital i and imaginary time τ couples only to $n_{\uparrow}(\tau)$ and $n_{\downarrow}(\tau)$, information embedded in $G^{\sigma}_{ii}(\tau,\tau) = 1 - \langle n_{i\sigma}(\tau)\rangle$. The second is for measuring expectation values of operators. The noninteracting character of the simulation enables Wick's theorem (Negele and Orland, 1988) to reduce such expectation values to sums of products of Green's functions.

7.2 Finite temperature algorithm

The methods in this chapter are most conveniently applied to interacting electron models such as

$$H = -\sum_{ij\sigma} T^{\sigma}_{ij} c^{\dagger}_{i\sigma} c_{j\sigma} + \sum_{ij} V_{ij} n_i n_j \tag{7.25}$$

that have the Coulomb interaction expressed as sums of number operator products (density-density interaction).[5] In (7.25), the spin-dependent symmetric matrix elements $T^{\sigma}_{ij} = T^{\sigma}_{ji}$ are hopping integrals, and the $c^{\dagger}_{i\sigma}$ and $c_{i\sigma}$ are the creation and destruction operators for an electron of spin $\sigma = \uparrow$ and $\downarrow$ on site i. The Coulomb interaction is parameterized by the spin-independent symmetric matrix $V_{ij} = V_{ji}$, and the on-site charge density is $n_i = n_{i\uparrow} + n_{i\downarrow}$. This Hamiltonian is quite generic and includes such prominent models as the Hubbard and extended Hubbard models, the single-impurity and lattice Anderson models, and the Falicov-Kimball model.

As a working example, we adopt the *repulsive Hubbard model*,

$$\begin{aligned} H &= H_1 + H_2 + H_3 \\ &= -\sum_{ij\sigma} T_{ij} c^{\dagger}_{i\sigma} c_{j\sigma} + U \sum_{i} \left(n_{i\uparrow} - \tfrac{1}{2}\right)\left(n_{i\downarrow} - \tfrac{1}{2}\right) - \mu \sum_{i} n_i, \end{aligned} \tag{7.26}$$

where $U > 0$ is the strength of the on-site Coulomb interaction, μ is the chemical potential, and the hopping matrix is spin-independent. While the hopping integrals in the Hubbard model are customarily restricted to nearest neighbors with a value of t, there is no a priori need to make this assumption. In what follows, we use N to designate the number of lattice sites, work in an occupation number basis, and assume periodic boundary conditions.

[5] The use of more general Hubbard-Stratonovich transformations (Appendix G) extends the applicability of the methods to other classes of Hamiltonians, although they cannot in general prevent a sign problem.

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

7.2.1 Matrix representation

Using the Trotter approximation (7.3), we first break the time evolution operator for a single time step into two factors

$$e^{-\Delta\tau H} \approx e^{-\Delta\tau(H_2+H_3)}e^{-\Delta\tau H_1}. \tag{7.27}$$

Both pieces of H in the first factor on the right are diagonal in the occupation number representation and thus commute. Consequently, this factor can be written as

$$e^{-\Delta\tau(H_2+H_3)} = \prod_i e^{-\Delta\tau U\left(n_{i\uparrow}-\frac{1}{2}\right)\left(n_{i\downarrow}-\frac{1}{2}\right)} \prod_i e^{-\Delta\tau\mu n_{i\uparrow}} e^{-\Delta\tau\mu n_{i\downarrow}}. \tag{7.28}$$

To the first product on the right, we now apply the discrete Hubbard-Stratonovich transformation (7.9) at each lattice site i. This transformation adds a summation over the auxiliary fields, but places a quadratic number operator into the arguments of the exponentials and allows us to factor the exponential for each site into commuting up- and down-spin terms.

For a given configuration of the auxiliary-field variables, we have

$$B^\sigma(\tau',\tau) = B^\sigma(\tau',\tau'-\Delta\tau)\cdots B^\sigma(\tau+2\Delta\tau,\tau+\Delta\tau)B^\sigma(\tau+\Delta\tau,\tau), \tag{7.29}$$

where

$$B^\sigma(\tau,\tau-\Delta\tau) = e^{\Delta\tau\mu}A^\sigma(\tau)\exp(\Delta\tau T) \tag{7.30}$$

and

$$A^\sigma(\tau) = \begin{pmatrix} e^{\sigma\alpha x_1(\tau)} & & & 0 \\ & e^{\sigma\alpha x_2(\tau)} & & \\ & & \ddots & \\ 0 & & & e^{\sigma\alpha x_{N_\tau}(\tau)} \end{pmatrix}. \tag{7.31}$$

The exponent arguments are $\sigma\alpha x = \pm\alpha x$, depending on whether $\sigma = \uparrow$ or $\downarrow$. The parameter α is found from $\cosh(\alpha) = \exp(\frac{1}{2}\Delta\tau U)$. The matrix A^σ (7.31) is diagonal because the auxiliary fields couple only to local electron degrees of freedom. In the imaginary-time propagation, the effect of $\exp(\Delta\tau T)$ is diffusive, $A^\sigma(\tau)$ acts like an external scattering potential, and $\exp(\Delta\tau\mu)$ is an attenuation factor. We recall that the Trotter approximation (7.27) makes B^σ an asymmetric matrix.

We can compute $\exp(\Delta\tau T)$ by finding the eigenvalues λ_i and eigenvectors v_i of T. The eigenvectors form the columns of the similarity transformation matrix $V = (v_1, v_2, \ldots, v_N)$ that diagonalizes T. With this transformation,

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

$$\exp(\Delta\tau T) = V \begin{pmatrix} e^{\Delta\tau\lambda_1} & & & 0 \\ & e^{\Delta\tau\lambda_2} & & \\ & & \ddots & \\ 0 & & & e^{\Delta\tau\lambda_N} \end{pmatrix} V^T. \tag{7.32}$$

Alternately, we can approximate $\exp(\Delta\tau T)$ via a checkerboard decomposition (Section 5.2.4),

$$\exp(\Delta\tau T) = \exp\left(\Delta\tau \sum_{ij} T^{(ij)}\right) = \prod_{ij} \exp\left(\Delta\tau T_{ij}^{(ij)}\right),$$

where the $T^{(ij)}$ are sparse matrices with only $T_{ij}^{(ij)} = T_{ji}^{(ij)} = T_{ij}$ nonzero,

$$T^{(ij)} = \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & 0 & \cdots & T_{ij} & \cdots & 0 \\ \vdots & & \vdots & \ddots & \vdots & & \vdots \\ 0 & \cdots & T_{ij} & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \end{pmatrix},$$

and thus

$$\exp(\Delta\tau T) = \prod_{ij} \begin{pmatrix} 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & & \vdots & & \vdots \\ 0 & \cdots & \cosh\left(\Delta\tau T_{ij}\right) & \cdots & \sinh\left(\Delta\tau T_{ij}\right) & \cdots & 0 \\ \vdots & & \vdots & & \vdots & & \vdots \\ 0 & \cdots & \sinh\left(\Delta\tau T_{ij}\right) & \cdots & \cosh\left(\Delta\tau T_{ij}\right) & \cdots & 0 \\ \vdots & & \vdots & & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{pmatrix}.$$

Each $\exp(\Delta\tau T^{(ij)})$ is a sparse matrix with only the ii, ij, ji, and jj elements differing from those of the identity matrix. This construction of an approximate exponential of the hopping matrix is particularly useful if T itself is sparse, for example, when the hopping is only between nearest neighbors. In such cases, the sparsity reduces the N^3 operations in a matrix-matrix multiplication to $N_{\text{bonds}}N^2$ where N_{bonds} is the number of sites j connected to site i. In most Hubbard-like models, $N_{\text{bonds}} \ll N$. The gain in the reduction of computation time per step, however, must be balanced by the need to use smaller $\Delta\tau$ to maintain the accuracy of the Trotter approximation and hence the need to take more imaginary-time steps.

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

7.2.2 Metropolis sampling

Having identified the statistical weight (7.20) of the auxiliary fields, we now want to sample them. Here, we adopt the Metropolis algorithm[6] (Section 2.5.1), and at each imaginary time, we go from lattice site to lattice site proposing a flip of the Ising-like discrete auxiliary-field variable x. We accept or reject the proposal depending on whether the ratio $\mathcal{R} = w_{C'}/w_C$ is greater than some random number ζ selected from the uniform distribution. The proposal affects only the time evolution through the matrix A defined in (7.31). For a flip from x to x' at time τ at site i,

$$A^\sigma(\tau) \to \begin{pmatrix} e^{\sigma\alpha x_1(\tau)} & & & & 0 \\ & \ddots & & & \\ & & e^{\sigma\alpha x_i'(\tau)} & & \\ & & & \ddots & \\ 0 & & & & e^{\sigma\alpha x_N(\tau)} \end{pmatrix} \equiv \left[I + \Delta^\sigma(i,\tau)\right] A^\sigma(\tau),$$

where the matrix $\Delta^\sigma(i,\tau)$ is null except for the $\Delta^\sigma_{ii}(i,\tau)$ element, which is

$$\Delta^\sigma_{ii}(i,\tau) = \exp[\sigma\alpha(x_i'(\tau) - x_i(\tau))] - 1.$$

The accompanying change in the propagator is

$$B^\sigma(\beta,0) = B^\sigma(\beta,\tau)B^\sigma(\tau,0) \to B^\sigma(\beta,\tau)\left[I + \Delta^\sigma(i,\tau)\right]B^\sigma(\tau,0)$$

so that

$$\mathcal{R}^\sigma \equiv \frac{w^\sigma_{C'}}{w^\sigma_C} = \frac{\det\left[I + B^\sigma(\beta,\tau)\left[I + \Delta^\sigma(i,\tau)\right]B^\sigma(\tau,0)\right]}{\det\left[I + B^\sigma(\beta,\tau)B^\sigma(\tau,0)\right]}. \tag{7.33}$$

This expression perhaps seems a bit daunting to evaluate, but with the use of two standard determinant relations, the expression simplifies considerably. One relation is

$$\det A / \det B = \det B^{-1}A = \det AB^{-1}, \tag{7.34}$$

and the other is

$$\det(I + AB) = \det(I + BA). \tag{7.35}$$

The algebra is straightforward. First, we write (7.33) as

$$\mathcal{R}^\sigma = \frac{\det\left[I + B^\sigma(\beta,0) + B^\sigma(\beta,\tau)\Delta^\sigma(i,\tau)B^\sigma(\tau,0)\right]}{\det\left[I + B^\sigma(\beta,0)\right]}.$$

[6] The heat-bath algorithm works just as well. The resulting modifications to the discussion are minor.

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

Then, we use the relation (7.34) to rewrite this equation as

$$\mathcal{R}^\sigma = \det\left[I + (I + B^\sigma(\beta,0))^{-1} B^\sigma(\beta,\tau)\Delta^\sigma(i,\tau)B^\sigma(\tau,0)\right].$$

Next, we use (7.35) to write

$$\mathcal{R}^\sigma = \det\left[I + \Delta^\sigma(i,\tau)B^\sigma(\tau,0)(I + B^\sigma(\beta,0))^{-1} B^\sigma(\beta,\tau)\right].$$

Finally, using

$$A(I + BA)^{-1}B = (I + (AB)^{-1})^{-1} = I - (I + AB)^{-1}$$

with $A \equiv B^\sigma(\tau,0)$ and $B \equiv B^\sigma(\beta,\tau)$, we obtain

$$\mathcal{R}^\sigma = \det\left[I + \Delta^\sigma(i,\tau)\,(I - G^\sigma(\tau,\tau))\right], \tag{7.36}$$

where

$$G^\sigma(\tau,\tau) = \left[I + B^\sigma(\tau,0)B^\sigma(\beta,\tau)\right]^{-1}. \tag{7.37}$$

$G^\sigma(\tau,\tau)$ is the equal-time single-particle Green's function (7.24). Because of the sparseness of Δ^σ,

$$\mathcal{R} = \mathcal{R}^\uparrow\mathcal{R}^\downarrow = \prod_\sigma \left[1 + \Delta^\sigma_{ii}(i,\tau)\left(1 - G^\sigma_{ii}(\tau,\tau)\right)\right], \tag{7.38}$$

a scalar, as opposed to a matrix, expression that is exceedingly simple to calculate, provided $G^\sigma_{ii}(\tau,\tau)$ is known. Computing G^σ from (7.29) and (7.37) for every accepted flip of $x_i(\tau)$ requires a significant computational effort. Matrix algebra, however, provides a technique to reduce this effort by a factor of N. We now proceed to derive this technique.

Under a proposed flip in the sign of $x_i(\tau)$,

$$\begin{aligned} G^\sigma(\tau,\tau) = (I + B^\sigma)^{-1} &\rightarrow \left[I + (I + \Delta^\sigma(\tau))\,B^\sigma\right]^{-1} \\ &= \left[(I + B^\sigma) + \Delta^\sigma(\tau)B^\sigma\,(I + B^\sigma)^{-1}\,(I + B^\sigma)\right]^{-1} \\ &= (I + B^\sigma)^{-1}\left[I + \Delta^\sigma(\tau)B^\sigma\,(I + B^\sigma)^{-1}\right]^{-1} \\ &= G^\sigma\left[I + \Delta^\sigma(\tau)\,(I - G^\sigma)\right]^{-1}, \end{aligned} \tag{7.39}$$

where we wrote B^σ as shorthand for $B^\sigma(\tau,0)B^\sigma(\beta,\tau)$. The equations of motion (7.16) for the electron operators yield

$$G^\sigma(\tau',\tau) = B^\sigma(\tau',\tau)G^\sigma(\tau,\tau) = G^\sigma(\tau',\tau')B^\sigma(\tau',\tau), \tag{7.40}$$

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

which are a pair of relations consistent with the propagator character of B^σ. Similarly, for $\tau' > \tau$, we advance the equal-time Green's function to a later time via

$$G^\sigma(\tau', \tau') = B^\sigma(\tau', \tau) G^\sigma(\tau, \tau) B^\sigma(\tau', \tau)^{-1}. \tag{7.41}$$

We still must show how to update $G^\sigma(\tau, \tau)$ efficiently if the Metropolis proposal is accepted.

Analytically,[7]

$$\left[I + \Delta^\sigma(\tau)\,(I - G^\sigma)\right]^{-1} = I - \frac{1}{\mathcal{R}^\sigma}\Delta^\sigma(\tau)\,(I - G^\sigma)\,, \tag{7.42}$$

where $\mathcal{R}^\sigma$ ($\sigma = \uparrow, \downarrow$) is the ratio of the weights before and after the flip. In a compact notation,

$$\mathcal{R}^\sigma = 1 + \Delta^\sigma_{ii}\left(1 - G^\sigma_{ii}\right). \tag{7.43}$$

Hence, under the proposed flip,

$$G^\sigma \rightarrow G^\sigma - \frac{1}{\mathcal{R}^\sigma} G^\sigma\,\Delta^\sigma(\tau)\,(I - G^\sigma)\,. \tag{7.44}$$

In terms of matrix elements,

$$G^\sigma_{jk} \rightarrow G^\sigma_{jk} - \frac{1}{\mathcal{R}^\sigma} G^\sigma_{ji}\,\Delta^\sigma_{ii}(\tau)\,(I - G^\sigma)_{ik}\,. \tag{7.45}$$

The significance of (7.44) becomes clearer: If we know G^σ_{ij} and accept the flip, we are able to compute the Green's function for the new configuration of auxiliary fields with a number of arithmetic operations of the order of the number of elements of G^σ, which is N^2. This number is a factor of N less than the $\mathcal{O}(N^3)$ operations associated with a matrix inversion. The gain is significant particularly because the updating of G^σ sits in the inner loop of the algorithm.

7.2.3 The algorithm

We now discuss the algorithm and several issues related to it. The overall Monte Carlo algorithm structure is the same as that discussed in Section 3.9. In the initialization, we not only set the auxiliary-field variables x_i to values of ± 1 for each site i at each time τ, but we also build the matrices $B^\sigma(\tau, \tau - \Delta\tau)$ for all times and the two

[7] We can prove (7.42) as follows:

$$\begin{aligned}(I + \Delta(I - G))(I - \mathcal{R}^{-1}\Delta(I - G)) &= I + \frac{\mathcal{R} - 1}{\mathcal{R}}\Delta(I - G) - \frac{1}{\mathcal{R}}\Delta(I - G)\Delta(I - G) \\ &= I + \frac{\mathcal{R} - 1}{\mathcal{R}}\Delta(I - G) - \frac{1}{\mathcal{R}}\Delta_{ii}(1 - G_{ii})\Delta(I - G) = I,\end{aligned}$$

where we have used the sparseness of Δ and (7.43).

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

spin components of the Green's function matrix for the initial time. We can build this Green's function using (7.37), by performing all the matrix multiplications and then computing the required matrix inversion. The steps seem quite direct, but as we discuss below, they have a hidden complexity.

Algorithm 23 details a Monte Carlo sweep. To facilitate the discussion of the sweep procedure, we use the more convenient notations $G_i \equiv G(\tau_i, \tau_i)$ and $B_i^\sigma \equiv B^\sigma(\tau_i, \tau_{i-1})$ for the i-th imaginary-time step $\tau_i = i\Delta\tau$. With these notations,

Algorithm 23 Finite-temperature determinant Monte Carlo algorithm.

Input: The $N_\tau \times N$ auxiliary field variables $x_{i,j}$, the $N \times N$ matrices B_i^σ, and the $N \times N$ matrices G_1^σ.

```
for i = 1 to N_τ do
  if mod(i, n_stab)=0 then
    for σ =↑ to ↓ do
      Make G_i^σ from (7.46) ;
    end for
  end if
  for j = 1 to N do
    Compute R^σ(x_ij) from (7.43) ;
    Draw ζ uniform randomly in [0, 1] ;
    if R^σ > ζ then
      x_ij ← −x_ij ;
      for σ =↑ to ↓ do
        Update G_i^σ via (7.45) ;
      end for
    end if
  end for
  if i < N_τ then
    for σ =↑ to ↓ do
      G_{i+1}^σ ← B_{i+1}^σ G_i^σ [B_{i+1}^σ]^{-1} ;
    end for
  else
    for σ =↑ to ↓ do
      G_1^σ ← B_1^σ G_{N_τ}^σ [B_1^σ]^{-1} ;
    end for
  end if
end for
return G_i^σ for i = 1, 2, ..., N_τ.
```

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

$$G_i^\sigma = \left[I + B_i^\sigma \cdots B_1^\sigma B_{N_\tau}^\sigma \cdots B_{i+1}^\sigma\right]^{-1}. \tag{7.46}$$

To update the auxiliary fields, we use the Metropolis algorithm and visit each lattice site at each imaginary time. If the proposed change is accepted, we update the Green's function using (7.44). After visiting all lattice sites, we advance the equal-time Green's function to the next time using (7.41) and then revisit the lattice sites. One sweep is finished after we complete the sampling for all times. We note that (7.37) implies $G^\sigma(\beta,\beta) = G^\sigma(0,0)$ and that (7.41) rolls $G^\sigma(\beta,\beta)$ over to $G^\sigma(\Delta\tau,\Delta\tau)$, as it must because of the periodic boundary conditions in imaginary time. When a proposed change is accepted, it changes the matrix A_i^σ in the definition (7.30) of the matrix B_i^σ. In advancing G^σ to the next time, we use the B^σ with the new auxiliary-field configuration.

As a Fermion algorithm, the finite-temperature determinant method generally has a sign problem (Sections 5.4 and 11.1). A sign problem arises when w_C^σ is negative while $w_C^{-\sigma}$ is positive. From (7.43), we see that for a given x a w_C^σ is negative when a $1 - G_{ii}^\sigma$ is. Since $1 - G_{ii}^\sigma = n_{i\sigma}$, this situation is clearly unphysical. Gubernatis and Zhang (1994) investigated whether the source of this situation was numerical imprecision in computing the Green's function. They concluded that while the lack of precision can cause a sign problem, imprecision was not the dominant source of the problem. They noted that in some cases the topology of the Hubbard-Stratonvich fields is directly connected to the topology of the Fermion fields and suggested as the source of the sign problem topological defects (zero modes) that develop in the Hubbard-Stratonovich configurations. We also comment that in some cases, for example, low electron fillings, the sign problem is not debilitating.

Algorithm 23 works well if we can perform all arithmetic operations with infinite numerical precision. With finite-precision arithmetic, the repeated updating of the Green's function and the stepping forward in time build up numerical errors. To combat this loss of accuracy, as indicated in Algorithm 23, we use (7.46) to rebuild the Green's functions periodically from the current values of the auxiliary fields. One strategy is to determine the period n_{stab} empirically: Fortunately when numerical accuracy breaks down, it does so rather suddenly. Consequently, we can try several values of n_{stab} until the breakdown occurs and then pick a value below this number.

The most serious loss of accuracy occurs in the inversion of the matrix in (7.46). As the temperature is lowered, this matrix becomes very ill-conditioned, and inverting it with a standard matrix-inversion routine produces a result that may have lost all numerical accuracy. In Section 7.4, we describe techniques, called *matrix product stabilization methods*, that enable this inversion. In brief, these methods accumulate the products of the B-matrices in a factored form in such a way as

to reduce the loss of information when the identity I and the product of the B's are added. The factored form is easy to invert, but the need to use these methods requires more coding than is implied in Algorithm 23. The stabilization methods add only a small cost to the total computation. *In the algorithm, where we say* "Make G_i^σ from (7.46)" *we assume that this step is done with the matrix product stabilization methods.* The description of these techniques is technical, yet essential. Because of their technical nature, we postpone their discussion until later in the chapter (Section 7.4). Using them is also important for accurate estimates of observables, especially those that require unequal-time Green's functions. We now discuss the basics of observable estimation.

7.2.4 Measurements

In analytic calculations of noninteracting models, Wick's theorem (Fetter and Walecka, 1971; Negele and Orland, 1988; Fulde, 1991; Enz, 1992) converts expectation values of products of operators into sums of products of all possible pair-wise contractions of the creation and destruction operators. The expectation values of these contractions are elements of the (noninteracting) Green's function. Thus, Wick's theorem expresses results of measurements in terms of the central quantity in the simulation.

As an example of the use of Wick's theorem in the determinant method, we consider the *anti ferromagnetic structure factor*,

$$S^{zz}(\pi,\pi) = \frac{1}{N}\sum_{ij}(-1)^{i-j}\left\langle\left(n_{i\uparrow}-n_{i\downarrow}\right)\left(n_{j\uparrow}-n_{j\downarrow}\right)\right\rangle,$$

where N is the number of lattice sites, $n_{i\sigma}$ is the number of electrons on site i with spin σ, and $(-1)^{i-j}$ is $+1$ if the sites i and j are on the same sublattice and -1 if they are not. The Monte Carlo simulation provides an estimate of $S^{zz}(\pi.\pi)$ by averaging over M independent samples of the auxiliary-field variables so that

$$\begin{aligned} S^{zz}(\pi,\pi) = \frac{1}{M}\sum_{\{C:C_1,C_2,\ldots,C_M\}}\frac{1}{N}\sum_{ij}(-1)^{i-j} \\ \times\left[\left\langle n_{i\uparrow}n_{j\uparrow}\right\rangle_C - \left\langle n_{i\downarrow}n_{j\downarrow}\right\rangle_C - \left\langle n_{i\downarrow}n_{j\uparrow}\right\rangle_C + \left\langle n_{i\uparrow}n_{j\downarrow}\right\rangle_C\right], \end{aligned}$$

where the quantum mechanical expectation values are for a specific configuration C of the auxiliary-field variables. Because the Hubbard-Stratonovich transformation factorizes the problem in the electron spin, contractions between different spin pairs, such as $\langle c_\sigma^\dagger c_{-\sigma}\rangle_C$, are zero. Hence,

$$\left\langle n_{i,\sigma}n_{j,-\sigma}\right\rangle_C = \left\langle n_{i,\sigma}\right\rangle_C\left\langle n_{j,-\sigma}\right\rangle_C = \left(1-G_{ii}^\sigma\right)\left(1-G_{jj}^{-\sigma}\right).$$

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

Like-spin terms produce more nonzero contractions

$$\begin{aligned}\left\langle n_{i\sigma} n_{j\sigma} \right\rangle_C &= \left\langle c^\dagger_{i\sigma} c_{i\sigma} c^\dagger_{j\sigma} c_{j\sigma} \right\rangle_C \\ &= \left\langle c^\dagger_{i\sigma} c_{i\sigma} \right\rangle_C \left\langle c^\dagger_{j\sigma} c_{j\sigma} \right\rangle_C + \left\langle c^\dagger_{i\sigma} c_{j\sigma} \right\rangle_C \left\langle c_{i\sigma} c^\dagger_{j\sigma} \right\rangle_C \\ &= \left(1 - G^\sigma_{ii}\right)\left(1 - G^\sigma_{jj}\right) + \left(\delta_{ij} - G^\sigma_{ji}\right) G^\sigma_{ij}.\end{aligned}$$

We note that the $i = j$ case reduces to $(1 - G^\sigma_{ii})$, as is necessary because $\langle n_{i\sigma} n_{i\sigma} \rangle_C = \langle n_{i\sigma} \rangle_C$.

When measurements are made, we may exploit the symmetries of the Hamiltonian to gather at once many possible estimates of a quantity of interest. Doing this is especially practical for the current algorithms because the measurement calculations typically execute in considerably less computer time than a Monte Carlo sweep. To estimate the kinetic energy for a translationally invariant system, for example, we may average the measurements of $\langle c^\dagger_{i,\sigma}(\tau) c_{i+\delta,\sigma}(\tau) \rangle_C$ over all lattice positions i, symmetry-related sites $i + \delta$, electron spins σ, and imaginary times τ. While individual measurements do not obey the Hamiltonian symmetries, averages over many configurations will. Averaging over as many equivalent measurements as possible reduces the variance of the final estimate.

Features in the simulation might break a symmetry of the Hamiltonian artificially. For example, in the Hubbard model, the *staggered magnetization*

$$M^\alpha(\pi, \pi) = \sum_i (-1)^i \left(c^\dagger_{i\uparrow} \;\; c^\dagger_{i\downarrow} \right) \sigma^\alpha \begin{pmatrix} c_{i\uparrow} \\ c_{i\downarrow} \end{pmatrix}, \quad \alpha = x, y, z,$$

where σ^α is a Pauli spin matrix, is rotationally invariant. Because it couples the auxiliary field to the z-component of the local spin $n_{i\uparrow} - n_{i\downarrow}$, the Hubbard-Stratonovich transformation breaks this symmetry, making the longitudinal estimator

$$\frac{1}{N} \left\langle M^z(\pi, \pi)^2 \right\rangle$$

much noisier than the transverse one (Hirsch, 1987)

$$\frac{1}{N} \left\langle M^x(\pi, \pi)^2 + M^y(\pi, \pi)^2 \right\rangle.$$

Using the latter is often more advisable.

7.3 Hirsch-Fye algorithm

In this section, we revisit the updating equation for the Green's function from a general point of view. The more general viewpoint does not lead to a more efficient algorithm for lattice models, but it does lead to an interesting variant of the updating

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

called the *Hirsch-Fye algorithm*. This algorithm, which is particularly useful for the simulation of impurity models, provides a different perspective on the updating just presented for nonimpurity problems.

Updating the Green's function $G^\sigma(x_1, x_2, \ldots, x_i, \ldots, x_N)$ for a set of values of Hubbard-Stratonovich fields at a given imaginary-time τ via (7.45) generates the Green's function $G^\sigma(x_1, x_2, \ldots, x_i', \ldots, x_N)$ with just one field changed at the same imaginary time. The procedure exactly solves the problem of changing the field at one site at a given time. We could start with the noninteracting Green's function and use the same update equation with $\Delta_{11}^\sigma(i,\tau) = \exp[\sigma\alpha x_1(\tau)] - 1$ to produce $G^\sigma(x_1)$ for an interaction activated at one site and then use this Green's function to produce $G^\sigma(x_1, x_2)$ for the Green's function with the interaction activated at two sites, and so on, until the interaction is active at all sites. The final Green's function would inherit the imaginary time associated with the Hubbard-Stratonovich fields. To update the Hubbard-Stratonovich fields for all sites requires $\mathcal{O}(N^3)$ floating-point operations, that is, $\mathcal{O}(N^2)$ operations for N sites. To do so for all times requires $\mathcal{O}(N_\tau N^3)$ operations. In Section 7.2.2, we also discussed procedures that scale as $\mathcal{O}(N^3)$ for moving the equal-time Green's function from one imaginary time to another and for generating unequal-time Green's functions from equal-time ones.

In the general case, the path integral for the partition function for discrete Hubbard-Stratonovich fields has the form (Blankenbecler et al., 1981; Negele and Orland, 1988)

$$Z = \sum_C \det X(C) \tag{7.47}$$

where $X(C) = X^\sigma(C)X^{-\sigma}(C)$ is the matrix representation of the noninteracting electron dynamics and its coupling to the auxiliary field. It is of order $N_\tau \times N$ but remarkably sparse:

$$X^\sigma = \begin{pmatrix} I & 0 & \cdots & \cdots & 0 & B_{N_\tau}^\sigma \\ -B_1^\sigma & I & \ddots & \ddots & 0 & 0 \\ 0 & -B_2^\sigma & I & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & I & 0 \\ 0 & 0 & 0 & \cdots & -B_{N_\tau-1}^\sigma & I \end{pmatrix}, \tag{7.48}$$

that is, it is an $N_\tau \times N_\tau$ block matrix where each block is an $N \times N$ matrix. The Green's function $\tilde{G}^\sigma$ for all sites and times is the inverse of $X^\sigma(x)$ (Negele and Orland, 1988). While these statements are not obvious, with a little algebra they can be made palatable. For example, starting with (7.47), choosing conveniently various permutations of the block rows and columns, and using the following formulas for

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

the *block inversion* of a matrix

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} + A^{-1}BS^{-1}CA^{-1} & -A^{-1}BS^{-1} \\ -S^{-1}CA^{-1} & S^{-1} \end{pmatrix} = \begin{pmatrix} T^{-1} & -T^{-1}BD^{-1} \\ -D^{-1}CT^{-1} & D^{-1} + D^{-1}CT^{-1}BD^{-1} \end{pmatrix}, \tag{7.49}$$

where

$$S = D - CA^{-1}B, \quad T = A - BD^{-1}C,$$

we can generate all the equal-time (7.24) and unequal-time (see (7.56) below) expressions for the Green's function. Hence, taking advantage of the structure and sparseness of $X^{\sigma}(x)$ reduces the $(N_\tau N)^3$ scaling of the cost of inverting (7.48) to the more efficient $N_\tau N^3$ scaling of the algorithm previously derived.

Related to (7.49), let us also note the useful relation

$$\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det A \det \left(D - CA^{-1}B\right) = \det D \det \left(A - BD^{-1}C\right). \tag{7.50}$$

Using this relation, it is easy to show the equivalence of $\det X(C)$ with (7.19), (7.20), and (7.21).

In the bigger space of dimension $N_\tau \times N$, what is the updating procedure for the total Green's function? To derive it, we start with the definition of the M^{σ}-matrix and use the definition (7.30) of the B-matrices to define

$$B_i^{\sigma} \equiv e^{-\Delta\tau V_i^{\sigma}} e^{\Delta\tau T},$$

where we absorbed the chemical potential factor into the hopping matrix T, and where from (7.31) we have the $N \times N$ diagonal matrix

$$e^{-\Delta\tau V_i^{\sigma}} = \begin{pmatrix} e^{\sigma\alpha x_1(\tau_i)} & & & 0 \\ & e^{\sigma\alpha x_2(\tau_i)} & & \\ & & \ddots & \\ 0 & & & e^{\sigma\alpha x_N(\tau_i)} \end{pmatrix}.$$

This definition separates the auxiliary-field-dependent terms from the terms with no such dependence. Next, we define the $N_\tau \times N_\tau$ block diagonal matrix

$$e^{-V^{\sigma}} = \begin{pmatrix} e^{-\Delta\tau V_1^{\sigma}} & & & \\ & e^{-\Delta\tau V_2^{\sigma}} & & \\ & & \ddots & \\ & & & e^{-\Delta\tau V_{N_\tau}^{\sigma}} \end{pmatrix}$$

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

and then introduce still another definition, the matrix $\hat{G}^\sigma \equiv \bar{G}^\sigma e^{V^\sigma}$, that is,

$$\hat{G}^\sigma = \begin{pmatrix} e^{-\Delta\tau V_1^\sigma} & & & & e^{\Delta\tau T} \\ -e^{\Delta\tau T} & e^{-\Delta\tau V_2^\sigma} & & & \\ & -e^{\Delta\tau T} & \ddots & & \\ & & \ddots & \ddots & \\ & & & -e^{\Delta\tau T} & e^{-\Delta\tau V_{N_\tau}^\sigma} \end{pmatrix}^{-1}.$$

From this last expression and the operator identity

$$I/(A+B) = I/A - \big(I/A\big)\, B\, \big[I/(A+B)\big],$$

it is straightforward to show that if $\exp(-V^\sigma)$ is updated to $\exp(-V'^\sigma)$, the updated Green's function $\hat{G}'^\sigma$ satisfies the Dyson equation

$$\hat{G}'^\sigma = \hat{G}^\sigma + \hat{G}^\sigma \left(e^{-V'^\sigma} - e^{-V^\sigma}\right) \hat{G}'^\sigma.$$

Finally, the substitution of the definition of $\hat{G}^\sigma$ into this equation delivers

$$\bar{G}'^\sigma = \bar{G}^\sigma + \bar{G}^\sigma \left(e^{-(V'^\sigma - V^\sigma)} - I\right)\big(I - \bar{G}'^\sigma\big) \tag{7.51}$$

for the Green's function $\bar{G}^\sigma$ of interest. This equation generalizes (7.39). Instead of a procedure for updating all sites at a given time, the new Dyson equation updates all sites at all times, a $\mathcal{O}((N_\tau N)^3)$ operation.

The generality of (7.51) allows us to define a procedure to update a given site for all times. This capability is especially useful for an impurity problem, that is, for a problem where one site is physically different from the rest; see, for example, (7.4). From (7.51), we can easily convince ourselves that if the interaction is localized at a particular site, say, 0, then the general update equation reduces to an $N_\tau \times N_\tau$ matrix equation just for that site's Green's function for all times:

$$\bar{G}'^\sigma_{00} = \bar{G}^\sigma_{00} + \bar{G}^\sigma_{00} \left[e^{-(V'^\sigma - V^\sigma)} - I\right]_{00} \big(I - \bar{G}'^\sigma_{00}\big).$$

Following similar algebra as previously presented reduces this updating equation to one with the same form as (7.45). The updates here are done one τ value at a time. As the matrix is $N_\tau \times N_\tau$, the computation time needed to update all times scales as N_τ^3.

This impurity updating procedure is the *Hirsch-Fye algorithm*. It was first developed for impurity problems (Hirsch and Fye, 1986) and then adapted for lattice problems. The algorithm easily generalizes to a cluster of N_c impurity sites leading to an algorithm scaling as $(N_c N_\tau)^3$. At a certain cluster size, the "all sites at a given

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

time" procedure becomes more efficient. Before the development of the continuous-time impurity solvers (Chapter 8), the Hirsch-Fye algorithm was widely used to solve single-site and cluster impurity problems.

There are several important points about the Hirsch-Fye algorithm: Although only the impurity Green's function is needed to execute the Monte Carlo updating, once this Green's function is known, the general update equation tells us how to use the impurity Green's function and the Green's function in the absence of the impurity to compute the Green's function with spatial matrix elements between the impurity and the other sites in the lattice (Gubernatis et al., 1987). This capability is the spatial analog of finding the unequal-time Green's function for all sites for one imaginary time.

At the beginning of the section, we noted that for a lattice of interacting sites, it was possible to activate the interactions one by one if the noninteracting Green's function is known. In the impurity problem something similar, but better, is possible. To create the noninteracting impurity Green's function, we need the Green's function for the problem without interactions on the impurity site. Once we have the impurity Green's function, we can activate the Hubbard-Stratonovich fields. Then, we do not need the other (spatially nonlocal) Green's function. For the Green's function without the impurity, we can use the one for an infinite lattice, if available, or calculate it for a very large lattice. In short, we are able to simulate the impurity problem without finite-size effects. In the next chapter, we discuss methods to eliminate the error caused by discretizing imaginary time.

7.4 Matrix product stabilization

The matrix methods described in this section are an essential part of the determinant method, but they have little to do with many-electron physics per se. They are principally used to calculate the equal-time Green's function. In this calculation, the product of the B-matrices is accumulated and added to the identity in a way that adequate precision remains after the matrix sum is inverted.[8] The problem is that the matrix needing inversion is very ill-conditioned.

The *condition number* κ of a matrix M is $\kappa = \|M\| \, \|M^{-1}\|$, where $\|M\|$ is some matrix norm (Golub and Loan, 1989; Meyer, 2000). The largest absolute value of the eigenvalues of the matrix is one such norm. The condition number measures how close a matrix is to being singular, that is, to having one or more zero eigenvalues. If $\kappa = 10^d$, then when a matrix is inverted, for example, using Gaussian elimination with partial pivoting, d represents roughly the number of

[8] Because the Hirsch-Fye algorithm does not accumulate the product of these matrices, it does not need these methods.

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

decimal places lost (Golub and Loan, 1989). We can get a feel of the difficulties in the current applications by considering the $U = 0$ case and restricting the hopping to nearest neighbors. For a square lattice, the eigenvalues of the nearest-neighbor hopping matrix T are well known to be $\lambda(k_x, k_y) = -2t[\cos(k_x) + \cos(k_y)]$ $(0 \leq k_{x,y} \leq 2\pi)$, so that $-4t \leq \lambda(k_x, k_y) \leq 4t$. With $\mu = 0$,

$$B(\beta, 0) = \left[I + e^{\beta T}\right]^{-1} \tag{7.52}$$

is a matrix whose condition number is roughly $\exp(8\beta t)$. When the temperature is about $t/5$, a small fraction of the electron's band width of $8t$, this condition number implies a loss of 17 decimal places of precision, a number just beyond what is representable in double-precision arithmetic.

The basic difficulty in thermodynamic studies of many-electron systems at low temperatures is that the Boltzmann distribution $e^{-\beta(H_1 - \mu N)}$ gives the lowest-energy states, relative to the chemical potential, exponentially large weights. The Pauli exclusion principle, however, prevents these states from becoming macroscopically occupied and forces them to be singly occupied up to some "Fermi energy." The matrix elements of a diagonalized $B(\beta, 0)$ are just 1 minus the Fermi-Dirac distribution,

$$[1 + \exp(-\beta(\lambda - \mu))]^{-1} = 1 - [1 + \exp(\beta(\lambda - \mu))]^{-1},$$

for each eigenvalue λ. Properly adding 1 to $\exp[-\beta(\lambda - \mu)]$ is the first step toward ensuring a proper Fermi-Dirac distribution of the electrons. The procedure for inverting $I + B(\beta, 0)$, in effect, is about properly adding the identity I to the Boltzmann distribution $\exp[-\beta(H - \mu N)]$ so that the inversion of the matrix sum preserves the Fermi-Dirac character of the electrons. The single particle propagators B^σ for any configuration of the auxiliary fields attenuate the "high-energy" scales while amplifying the "low-energy" scales. Accordingly, the extraction of information about "intermediate" energies near the chemical potential is a very difficult numerical task, but one that must be done accurately as it is these energy states that control the finite-temperature physics. Without the procedures to be described, the Fermionic character of the simulation gets lost.

The numerical limitation in inverting a matrix embodying very diverse numerical scales is quite different from that for inverting a single number. For a single number, double-precision arithmetic typically allows the scale to range between 10^{-307} and 10^{+309}. The strategy devised (Loh et al., 1989; Loh and Gubernatis, 1992; Loh et al., 2005) for inverting our ill-conditioned matrix is stratifying the matrix multiplications with a representation of the B matrices that places these numerical scales into diagonal matrices. We can accurately invert them as doing so just requires inverting the individual numbers. The procedures also accumulate the matrix multiplications in a way that avoids mixing the scales. Only a final set of matrix multiplications

mixes the scales and generates a loss of accuracy. The remaining accuracy is enough that the measurement error is at least as small as the statistical error.

The procedure employed decomposes the matrices into the form UDV where the diagonal matrix D contains the diverse scales and U and V are reasonably well conditioned and readily invertible. The following illustrates the character of the factored form and the problem we are trying to avoid:

$$UDV = \begin{pmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{pmatrix} \begin{pmatrix} X & & & \\ & x & & \\ & & x & \\ & & & x \end{pmatrix} \begin{pmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{pmatrix}$$
$$= \begin{pmatrix} X & X & X & X \\ X & X & X & X \\ X & X & X & X \\ X & X & X & X \end{pmatrix},$$

where we are using the size of the character x to represent different numerical scales. The matrices U and V have unit scales. The three matrices in the first line of the above equation represent what we want to achieve in the factorization. The second line illustrates what happens upon multiplication – the largest numerical scale dominates and small-scale information is lost.

On the other hand, if a matrix M has its numerical scales stratified into columns, we can factorize it in a stable way, preserving its small-scale information:

$$U^{-1}MV^{-1} = \begin{pmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{pmatrix} \begin{pmatrix} X & x & x & x \\ X & x & x & x \\ X & x & x & x \\ X & x & x & x \end{pmatrix} \begin{pmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{pmatrix}$$
$$= \begin{pmatrix} X & & & \\ & x & & \\ & & x & \\ & & & x \end{pmatrix} = D.$$

Multiplication on the left of M by U^{-1} only combines the elements in a given column and thus causes no loss of information. The multiplication on the right by V^{-1} combines columns of different scales but does not overwrite any small-scale information *as long as large-scale columns are first scaled down* before they are added into the columns of smaller scale. Using $MV^{-1} = (V^{-T}M^T)^T$ reduces the mixing.

To compute the product (7.29) of many matrices in a stable manner, we separate the many scales throughout the calculation. As an illustration, we first imagine we

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

have $B(\tau,0) = UDV$. If we want this propagator at $\tau + \tau_0$, where τ_0 is the size of a few imaginary time steps,[9] we can extend the propagator with standard matrix multiplication methods. We write

$$\begin{aligned} B(\tau+\tau_0,0) &= B(\tau+\tau_0,\tau)UDV = [B(\tau+\tau_0,\tau)UD]\,V \\ &= \left[B(\tau+\tau_0,\tau)U \begin{pmatrix} X & & & \\ & x & & \\ & & x & \\ & & & x \end{pmatrix} \right] V = \begin{pmatrix} X & x & x & x \\ X & x & x & x \\ X & x & x & x \\ X & x & x & x \end{pmatrix} V \\ &= \left(U'D'V'\right)V = U'D'\left(V'V\right), \end{aligned} \tag{7.53}$$

thus obtaining the UDV factorization of $B^\sigma(\tau+\tau_0,0)$. What we did was to decompose the stratified matrix $B^\sigma(\tau+\tau_0)UD$ into $U'D'V'$. The V matrices must be sufficiently well conditioned so that we can multiply many of them together in a stable way. Repeating this procedure as often as necessary enables the computation of a long string of products of the B matrices (7.21) by multiplying from the left.

After the matrix products are assembled in factored form, we compute the inverse $[I+UDV]^{-1}$ by the following steps:

$$\begin{aligned} G = [I+UDV]^{-1} &= V^{-1}\left[U^{-1}V^{-1}+D\right]^{-1}U^{-1} \\ &= V^{-1}\left[U'D'V'\right]^{-1}U^{-1} = \left(V'V\right)^{-1}\left(D'\right)^{-1}\left(UU'\right)^{-1}. \end{aligned}$$

We note that after forming the sum $U^{-1}V^{-1}+D$, we decompose it into $U'D'V'$ and perform the inverses of the factors separately.

An alternative to the above is performing the factorization of the two partial products

$$\begin{aligned} B^\sigma(\tau,0) &= U_1D_1V_1, \\ B^\sigma(\beta,\tau) &= V_2D_2U_2, \end{aligned}$$

where in the second equation we reverse the UDV factors because we intend to build up the partial product by multiplications from the left. We now have

$$\begin{aligned} G^\sigma(\tau,\tau) &= [I+U_1D_1V_1V_2D_2U_2]^{-1} \\ &= U_2^{-1}\left[U_1^{-1}U_2^{-1}+D_1V_1V_2D_2\right]^{-1}U_1^{-1}. \end{aligned} \tag{7.54}$$

[9] We do not apply the factorization at each time step but rather with a period τ_0 in imaginary time.

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

Schematically, the piece to be inverted is

$$\begin{pmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{pmatrix} + \begin{pmatrix} X & & & \\ & x & & \\ & & x & \\ & & & x \end{pmatrix} \begin{pmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{pmatrix} \begin{pmatrix} X & & & \\ & x & & \\ & & x & \\ & & & x \end{pmatrix}$$
$$= \begin{pmatrix} x & x & x & x \\ x & x & x & x \\ x & x & x & x \\ x & x & x & x \end{pmatrix} + \begin{pmatrix} XX & Xx & X_x & X_x \\ xX & xx & x_x & x_x \\ {}_xX & {}_xx & {}_{xx} & {}_{xx} \\ {}_xX & {}_xx & {}_{xx} & {}_{xx} \end{pmatrix}.$$

Because we kept the diagonal matrices on the outside, the small-scale cut-off occurs only in the last step when we add together elements of different scales. We invert the resulting matrix by first doing another *UDV* factorization.

How do we create the *UDV* factorization? A number of standard numerical matrix methods work, including the modified Gram-Schmidt, *QR*, and the singular value decomposition methods (Golub and Loan, 1989; Meyer, 2000). A UDV^T factorization is the standard result of the singular value decomposition method. The standard modified Gram-Schmidt and *QR* algorithms are easily rewritten to produce this result. The modified Gram-Schmidt method, though slightly less stable numerically, is noticeably faster and usually more than adequate. How do we invert the various matrices? Diagonal matrices are trivial to invert. In fact, so are U and V. The modified Gram-Schmidt method produces a column orthonormal U and an upper-unit triangular V. The first we invert by transposition, and the second by a back-substitution solution of the linear system of equations $VV^{-1} = I$ for the unknown V^{-1}, which is also a triangular matrix.

The Gram-Schmidt orthogonalization method is the following: Given the set of vectors $a_1, a_2, \ldots, a_M$, which may be the columns of a matrix, find a set of vectors $v_1, v_2, \ldots, v_M$ that are orthonormal and span the same space. The idea is quite simple: Take a_1 and divide it by its length $\|a_1\| = \sqrt{a_1^T a_1}$ to get v_1. Then subtract the component of a_2 in the direction of v_1. The remainder is orthogonal to v_1. Next, divide the remainder by its length to get v_2. Repeat the process by subtracting the component of a_3 in the directions of v_1 and v_2. Dividing the remainder by its length gives v_3. By continuing the process, we obtain the desired set of orthonormal vectors. Unfortunately, this process is numerically unstable if any of the a_i are nearly linearly dependent. Fortunately, the fix is simple: The computation is reformulated so that the a_i are never used once the v_i are obtained. The algorithm is given in Algorithm 24. In a more computer-language-like fashion, Algorithm 25 describes the orthonormalization and the matrix factorization in pseudocode.

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

Algorithm 24 Modified Gram-Schmidt method.

Input: M vectors $a_1, a_2, \ldots, a_M$.
  for $i = 1$ to M **do**
    $a_i \leftarrow a_i / \|a_i\|$;
    for $j = i + 1$ to M **do**
      $a_j \leftarrow a_j - (a_i^T a_j) a_i$;
    end for
  end for
  return M orthonormal vectors $a_1, a_2, \ldots, a_M$.

Algorithm 25 Modified Gram-Schmidt method: *UDV* factorization.

Input: $N \times N$ array U.
  for $k = 1$ to N **do**
    $d(i) = 0$;
    for $i = 1$ to N **do**
      $d(k) \leftarrow d(k) + U(i,k)U(i,k)$;
    end for
    $d(k) \leftarrow \sqrt{d(k)}$;
    for $i = 1$ to N **do**
      $U(i,k) \leftarrow U(i,k)/d(k)$;
    end for
    $V(k,k) = 1$;
    for $j = k + 1$ to N **do**
      $V(k,j) = 0$;
      for $i = 1$ to N **do**
        $V(k,j) \leftarrow V(k,j) + U(i,k)U(i,j)$;
      end for
      for $i = 1$ to N **do**
        $U(i,j) \leftarrow U(i,j) - V(k,j)U(i,k)$;
      end for
      $V(k,j) \leftarrow V(k,j)/d(k)$;
    end for
  end for
  return the $N \times N$ orthonormal matrix U, the N diagonal elements d of a $N \times N$ diagonal matrix D, and the $N \times N$ unit upper triangular matrix V.

While the procedure for updating the equal-time Green's function enables convenient Monte Carlo sampling, the goal of any sampling remains the estimation of various physical observables. The determinant method permits the computation of

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

these quantities as a function of imaginary time. For many such quantities, knowing just equal-time Green's functions suffices. For others, such as *susceptibilities*,

$$\chi\,(i\omega_n) = \int_0^\beta \left\langle A^\dagger(\tau)A(0)\right\rangle e^{i\omega_n\tau}\,d\tau. \tag{7.55}$$

Wick's theorem, generates contractions of the electron operator pairs at different imaginary times. Thus, for the measurement part of the simulation, we often need the matrix elements of unequal-time Green's functions.

The unequal-time Green's functions are

$$G^\sigma(\tau',\tau) = B^\sigma(\tau',\tau)\left[I + B^\sigma(\tau,0)B^\sigma(\beta,\tau)\right]^{-1}. \tag{7.56}$$

For small changes $\tau' \to \tau''$, multiplication by the single-particle propagator produces the appropriate Green's functions

$$G^\sigma(\tau'',\tau) = B^\sigma(\tau'',\tau')G^\sigma(\tau',\tau).$$

As with the equal-time Green's function, this procedure propagates with adequate accuracy for some time τ_0. To extend this range, we write

$$G^\sigma(\tau,0) = B^\sigma(\tau,0)\left[I + B^\sigma(\beta,0)\right]^{-1} = \left[B^\sigma(\tau,0)^{-1} + B^\sigma(\beta,\tau)\right]^{-1}$$

and accumulate the two propagators in factored form,

$$G^\sigma(\tau,0) = \left[V_1^{-1}D_1^{-1}U_1^{-1} + V_2D_2U_2\right]^{-1}, \tag{7.57}$$

where $U_1D_1V_1$ is the factorization of $B^\sigma(\tau,0)$. Then, we isolate the most ill-conditioned diagonal matrix, for example, D_2, and write

$$G^\sigma(\tau,0) = U_1\left[D_1^{-1} + V_1V_2D_2U_2U_1\right]^{-1}V_1.$$

The matrix inverse of the sum is done by Gaussian elimination with partial pivoting.

A more stable procedure, related to (7.54), is to reexpress (7.57) as

$$G^\sigma(\tau,0) = U_2^{-1}\left[D_1^{-1}U_1^{-1}V_2^{-1} + V_1V_2D_2\right]^{-1}V_1,$$

then factor each diagonal matrix D as $D^{\max}D^{\min} = D^{\min}D^{\max}$ and rewrite (7.4) as

$$G^\sigma(\tau,0) = U_2^{-1}\left(D_2^{\max}\right)^{-1}\Big[\left(D_1^{\max}\right)^{-1}U_1^{-1}V_2^{-1}\left(D_2^{\max}\right)^{-1} + \left(D_1^{\min}\right)^{-1}V_1V_2\left(D_2^{\min}\right)^{-1}\Big]^{-1}\left(D_1^{\min}\right)^{-1}V_1.$$

$D^{\max}$ is D with its nonzero elements less than 1 replaced by 1, and $D^{\min}$ is D with elements greater than 1 replaced by 1. The matrix inverse of the sum is done by Gaussian elimination with partial pivoting.

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

7.5 Comments

The birth of supercomputers gave life to the determinant method. In the mid-1980s, it pushed the fastest and biggest computers to their limits. While processor speeds, the size of internal and external memory, and the access rate of these memories have all improved dramatically, the range of applications of these quantum Monte Carlo methods has remained somewhat static because for many interesting applications, its true computation time scales exponentially with the problem size due to the sign problem (Loh et al., 1990). Today, for systems lacking a sign problem, or ones where it is minor, the execution cost of this method is relatively cheap.

While for finite-temperature Fermion systems a good way to reduce the scaling of the sign problem is still lacking, there are ways to reduce the computational cost when there is no sign problem. Reducing the number of matrix-matrix multiplications and matrix inversions is the most direct route. In Section 7.2.1, we saw how using a sparse matrix (checkerboard) representation for the exponential of the hopping matrix can reduce the computational scaling from N^3 to N^2. Even the use of a matrix stabilization factorization producing triangular as opposed to dense matrices helps. Sometimes we can also achieve significant gains by computing, storing, and craftily using strings of partial matrix products rather than repeatedly recomputing the entire string. We illustrate one such approach for the zero-temperature determinant method described in Appendix I.

A single spin-flip algorithm, such as the Metropolis or heat-bath algorithm, causes the determinant method to share the same broken ergodicity problems as the Monte Carlo simulation of a classical spin model in a strong external magnetic field. The Hubbard-Stratonovich transformation (7.9) in effect couples a local fluctuating quantum spin to an imaginary-time-dependent magnetic field that depends on the magnitude of U. As such, when U becomes large, sampling ergodically becomes a problem. Furthermore, the Trotter approximation restricts the use of this method to the weak and intermediate coupling regimes. Global spin moves (Scalettar et al., 1991) promote ergodicity and increase the magnitude of U, which can be simulated. In the past, to keep the cost of the simulation manageable, $\Delta\tau$'s of $0.25t$ or larger were used where t is the magnitude of the nearest-neighbor hopping. The recently developed continuous-time version of the finite-temperature determinant method (Iazzi and Troyer, 2014) is discussed in Appendix M.[10] These algorithms do not use a Trotter decomposition, but instead sample configurations from a weak-coupling perturbation expansion in continuous imaginary time, using techniques very similar to those that we introduce for impurity models in the following chapter.

The advantages of global or multiple-site updating methods of auxiliary and Bosonic fields have led to other algorithms, most notably the hybrid quantum Monte

[10] For the zero-temperature implementation, see Wang et al. (2015).

Carlo method (Scalettar et al., 1987). It updates at once all sites at all times. This method employs a continuous Hubbard-Stratonovich transformation, and it is conceivable that this method could be increasingly used for simulations of complex electron lattice models requiring multiple on-site or longer-ranged Coulomb interactions, cases when the discrete auxiliary fields become numerous. In these situations, the use of the continuous Hubbard-Stratonovich transformation is potentially more advantageous (Appendix G).

Briefly, the hybrid method uses continuous Hubbard-Stratonovich fields and reformulates (7.47) as

$$Z = \int \mathcal{D}x\, \mathcal{D}\phi e^{-\left[S(x)+\phi^T O^{-1}\phi\right]}, \tag{7.58}$$

where $\exp[-S(x)]$ is the weight of the fields and the matrix $O = X(x)^T X(x)$. Next, fictitious momentum fields p, canonical to the auxiliary fields x, are introduced and a Gaussian integral over the p is added to the path integral without changing the physical content of the partition function,[11]

$$Z = \int \mathcal{D}p\, \mathcal{D}x\, \mathcal{D}\phi e^{-H_{\text{eff}}(p,x,\phi)}.$$

Here

$$H_{\text{eff}}(p,x,\phi) = \sum_i \frac{p_i^2}{2m} + V_{\text{eff}}(x,\phi), \quad V_{\text{eff}}(x,\phi) = S(x) + \phi^T O^{-1}(x)\phi.$$

The simulation strategy is to sample all p and x fields, keeping the ϕ fields fixed, and then to sample all ϕ fields, keeping the p and x fields fixed. When the ϕ fields are fixed, the hybrid Monte Carlo method (Liu, 2001) is used. In this method, molecular dynamics evolves the system of particles described by H_{eff} in a fictitious time to generate a proposed change of all the p and x fields. Normally, molecular dynamics (Allen and Tildesley, 1987) conserves energy so the Metropolis algorithm would always accept the change. The intent here, however, is not to conserve energy well but rather to use a rough molecular dynamics simulation with a time step a little too large. The step size is chosen by adjusting it and the pseudo-mass m to cause an energy nonconservation so the Metropolis acceptance rate is around 80 to 90%. We recall that the Metropolis algorithm does not specify how the proposal is to be made. However it is made, the Metropolis algorithm ensures we are sampling the auxiliary fields with a probability $\exp(-V_{\text{eff}})$. Sampling the ϕ fields from $\exp(-\phi^T O^{-1}\phi)$ with $O = X^T X$ is done by sampling fields R^σ from a multidimensional Gaussian distribution $\exp(-\frac{1}{2}R^{\sigma T}R^\sigma)$ to find the sample of ϕ^σ via $\phi^\sigma = X^\sigma R^\sigma$ (Appendix G.1).

[11] Recall the factorization of the Gaussian integration of momentum of the partition function in classical statistical mechanics.

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

An intermediate step of the molecular dynamics is the computation of V_{eff}. It requires computing

$$Y^\sigma = \left[O^\sigma\right]^{-1}\phi^\sigma = \left[X^\sigma\right]^{-1}R^\sigma,$$

which is done by solving the large sparse linear system of equations $X^\sigma Y^\sigma = R^\sigma$. We note that the elements of $[X^\sigma]^{-1}$ are the imaginary-time elements of the Green's function. For example,

$$[X_i^\sigma]^{-1} = G_i^\sigma = [I + B_i^\sigma \cdots B_1^\sigma B_{N_\tau}^\sigma \cdots B_{i+1}^\sigma]^{-1}.$$

Accordingly, both solving the linear system defined by X^σ and computing the inverse of $I + B_i^\sigma \cdots B_{N_\tau}^\sigma B_1^\sigma \cdots B_{i-1}^\sigma$ involve matrices that are very ill-conditioned. For this reason, preconditioned, not the standard, conjugate-gradient methods (Golub and Loan, 1989; Meyer, 2000) are used to solve the system of equations. Giving more details will take us too far from our central focus.

In closing, we comment that embedded in our discussion of the hybrid quantum Monte Carlo method is a basic strategy for simulating systems where the electrons are coupled to Bosonic fields such as phonons or magnons: We fix the Bosonic fields and then use an appropriate Monte Carlo method to sample the scalar variables representing the electron degrees of freedom. Then, we fix these scalar variables and use an appropriate simulation method to sample the Bosonic degrees of freedom. The two sampling methods need not be the same. Combining Bosonic fields with the scalar auxiliary field is also possible with the determinant method.

Suggested reading

E. Y. Loh, Jr. and J. E. Gubernatis, "Stable simulations of models of interacting electrons in condensed-matter physics," in *Electronic Phase Transitions*, ed. W. Hanke and Yu. V. Kopaev (Amsterdam: North-Holland, 1992), chapter 4.

A. Muramatsu, "Quantum Monte Carlo for lattice Fermions," in *Quantum Monte Carlo Methods in Physics and Chemistry*, ed. M. P. Nightingale and C. J. Umrigar, NATO Science Series C, vol. 525 (Dordrecht: Kluwer Academics, 1999), chapter 13.

E. Y. Loh, Jr., J. E. Gubernatis, R. T. Scalettar, S. R. White, D. J. Scalapino, and R. L. Sugar, "Numerical stability and the sign problem in the determinant Monte Carlo method," *Int. J. Mod. Phys. C* **16**, 1319 (2005).

Exercises

7.1 For the Hamiltonian

$$H = U(n_\uparrow - \tfrac{1}{2})(n_\downarrow - \tfrac{1}{2}), \tag{7.59}$$

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

where $U > 0$, analytically evaluate the grand canonical partition function

$$Z = \mathrm{Tr}\, e^{-\beta(H-\mu N)} \tag{7.60}$$

1. By summing over all the basis states $|00\rangle$, $|01\rangle$, $|10\rangle$, and $|11\rangle$.
2. By making the continuous Hubbard-Stratonovich transformation (7.6), tracing out the electron degrees of freedom, and performing the remaining Gaussian integral over the auxiliary field.
3. By making the discrete Hubbard-Stratonovich transformation (7.9), tracing out the electron degrees of freedom, and performing the remaining sum over the auxiliary field.

7.2 What unequal-time Green's functions are needed to calculate the susceptibility (7.55) when $A = n_\uparrow - n_\downarrow$?

7.3 On a bipartite lattice, with no magnetic field, the *particle-hole transformation* is

$$c_{i\uparrow} \to d_{i\uparrow},$$
$$c^\dagger_{i\downarrow} \to \begin{cases} +d^\dagger_{i\downarrow}, & i \in A, \\ -d^\dagger_{i\downarrow}, & i \in B. \end{cases}$$

A *bipartite lattice* is one composed of two sublattices (A and B) where the nearest-neighbor sites of one lattice belong to the other.

1. Show that under this transformation, the Hubbard Hamiltonian transforms as

$$H(t, U, N) \to H(t, -U, N) + UN_\uparrow,$$

where $N = N_\uparrow + N_\downarrow$ is the total number of electrons. This relation allows us to make statements about the repulsive Hubbard model by studying the attractive model and vice versa. Note that the N_σ are conserved.
2. Similarly, show that under this transformation the total charge Q and the z-component S^z of the total spin transform as $Q \to S^z + 1$ and $S^z \to Q - 1$. These relations connect the behavior of particles and spins within a repulsive or attractive model and between these models.
3. Show that under the particle-hole transformation, the Hubbard Hamiltonian also transforms as $H(N) = H(2L - N) + U(N - L)$.
4. What does the last relation say about the symmetry of the ground state and the excited states for the square lattice at half-filling? Are they the same? For a square lattice, half-filling corresponds to one electron per site and the number of up and down electrons being equal.

7.4 Verify the validity of the discrete Hubbard-Stratonovich transformations (7.9).

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press

7.5 Prove the three determinant relations (7.34), (7.35), and (7.23). Note the first two require the matrices to be square; the last one does not.

7.6 The operator product $n_\uparrow n_\downarrow$ may be written as the sum of two squares $\frac{1}{4}[(n_\uparrow + n_\downarrow)^2 - (n_\uparrow - n_\downarrow)^2]$. What discrete Hubbard-Stratonvich transformation follows from this expression for the interaction?

7.7 Show that the diagonal and off-diagonal elements of the inverse of M^σ (7.48) are (7.37) and (7.40).

7.8 For the noninteracting problem, (7.32) represents a convenient and obvious choice of a *UDV* factorization of $\exp(\Delta\tau T)$. For

$$T = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix},$$

compute analytically its eigenvalues and eigenvectors plus

$$[I + \exp(\beta T)]^{-1}.$$

Then compare this exact result with its numerical computation via the Trotter breakup

$$[I + \exp(\Delta\tau T)\exp(\Delta\tau T)\cdots\exp(\Delta\tau T)]^{-1}$$

with and without the procedure defined by (7.53) for different values of β and $\Delta\tau$. From the point of view of matrix analysis and linear algebra, what are the differences between this computation and the one found in a quantum Monte Carlo simulation?

7.9 Repeat the above using the modified Gram-Schmidt factorization instead of the eigenvalue diagonalization.

https://doi.org/10.1017/CBO9780511902581.008 Published online by Cambridge University Press