Lattice computation of energy moments in canonical and Gaussian quantum statistics

2000 J. Phys. A: Math. Gen. 33 4227

(http://iopscience.iop.org/0305-4470/33/23/302)

J. Phys. A: Math. Gen. **33** (2000) 4227–4247. Printed in the UK PII: S0305-4470(00)11352-6

Lattice computation of energy moments in canonical and Gaussian quantum statistics

Jani Lukkarinen

Helsinki Institute of Physics, PO Box 9, 00014 University of Helsinki, Finland

E-mail: jani.lukkarinen@helsinki.fi

Received 25 January 2000, in final form 10 April 2000

Abstract. We derive a lattice approximation for a class of equilibrium quantum statistics describing the behaviour of any combination and number of bosonic and fermionic particles with any sufficiently binding potential. We then develop an intuitive Monte Carlo algorithm which can be used for the computation of expectation values in canonical and Gaussian ensembles and give lattice observables which will converge to the energy moments in the continuum limit. The focus of the discussion is twofold: in the rigorous treatment of the continuum limit and in the physical meaning of the lattice approximation. In particular, it is shown how the concepts and intuition of classical physics can be applied in this sort of computation of quantum effects. We illustrate the use of Monte Carlo methods by computing canonical energy moments and the Gaussian density of states for charged particles in a quadratic potential.

1. Introduction

Recently, the role of the canonical Boltzmann–Gibbs ensemble as the only useful ensemble for quantum statistics has been re-evaluated. The need for re-evaluation has arisen both from experimental and theoretical considerations. Although the canonical ensemble also works surprisingly well for systems consisting of only a few particles, relativistic ion collisions have provided an arena where the corrections from the finite size of the system to the canonical results have become important [1]. A quantitative understanding of these collisions is necessary since they are central in the search for new phases of strongly interacting matter, such as a quark–gluon plasma. In addition, methods for a quantitative measurement of the properties of small quantum systems could be useful in industrial applications in the future, such as in the design of ever smaller electronic components. Similarly, the functioning of certain biological systems seems to also involve the quantum behaviour of the system—consider, for instance, the recognition of molecules by olfactory receptors [2].

The second motivation for non-canonical ensembles comes from the peculiar behaviour associated with systems with fractal-like structure. Tsallis statistics† [3] was developed for the statistical analysis of these cases and we have considered other generalized statistics of that kind in a previous work [4]. For 'ordinary' systems, however, we found that the so-called Gaussian ensemble is the simplest generalization of the canonical ensemble which could be used for the approximation of any precanonical system—these are systems for which the occupation of high-energy states decreases faster than exponentially. For a more complete description of precanonical and Gaussian ensembles, see [4, 5], respectively. Of

† A comprehensive list of references is maintained by C Tsallis at the URL http://tsallis.cat.cbpf.br/biblio.htm.

0305-4470/00/234227+21$30.00

the non-canonical ensembles, we concentrate here only on the Gaussian ensemble with the understanding that the results presented can be easily modified to apply for any precanonical ensemble.

In this work, we develop and prove the existence of a lattice approximation for a system of a fixed number of non-relativistic particles which may be fermions, bosons or a combination of both. The potential of this system can be very general, the only requirement being a fast enough increase at infinity. We do not consider the possibility of a creation or annihilation of particles—this would require the use of a quantum field theory instead of ordinary quantum mechanics and we would have to give up mathematical rigour in the transition. Formulations of the microcanonical ensemble using quantum field theory and its perturbative expansion have already been proposed by Chaichian and Senda [6] and our approach owes a lot to their work.

We begin by defining our notation and by stating assumptions we need to make on the potential in section 2. We then state a theorem which defines and proves the convergence of the lattice approximation of traces at a complex temperature and we derive a set of lattice observables which can be used for the measurement of energy moments. Section 5 presents a Monte Carlo algorithm for the generation of the canonical lattice distribution and we also try to develop an intuitive understanding of it. Section 6 contains similar derivations for the Gaussian ensemble; in particular, it contains an essentially canonical algorithm which can be used for the computation of the density of states. We conclude with an example where we test the algorithms in practice and comment on their strengths and drawbacks. We have also gathered a comparison of our methods with those presented in the literature along with suggestions for improvements in section 8. A good review of the state of the art of using Monte Carlo simulations for the path integrals of quantum fluids has been given by Ceperley [7].

2. Notation and background

Consider a system of N particles living in a d-dimensional space. In classical physics, this system is described by trajectories of these particles and the positions of the particles form the relevant degrees of freedom, an n-dimensional space with $n = Nd$. Here we shall choose the convention that the positions of particles are denoted by boldface letters with a subscript identifying the label of the particle, while the classical configuration of the system is denoted by regular italics. For example, the n-dimensional configuration x could be written in terms of the N position vectors as $x = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N)$.

The lattice regularization of the quantum effects on this system is defined on a space containing L copies of the classical configuration space. The 'lattice configurations' are thus nL-dimensional vectors and we shall put the number of the copy in parentheses after the configuration symbol. In other words, the position of the particle i in copy k in the lattice configuration x is denoted by $\boldsymbol{x}_i(k)$. The quantum results are recovered in the continuum limit, by which we mean here simply taking L to infinity in the lattice regularized formulae given below.

The following discussion applies to massive, non-relativistic particles which are bounded by a suitably strong potential and thus can be described by equilibrium ensembles. To simplify matters, we also assume that all masses have been scaled away by the transformation $\boldsymbol{x}_i \mapsto \boldsymbol{x}_i/\sqrt{m_i}$ and we use the natural units where the Planck and Boltzmann constants are equal to one.

With these choices, the Hamiltonian of the system can be written as

$$\widehat{H} = \sum_{i=1}^{N} \tfrac{1}{2}\widehat{\boldsymbol{p}}_i^{\,2} + V(\widehat{\boldsymbol{x}}_1, \ldots, \widehat{\boldsymbol{x}}_N)$$

where $\sum_{i=1}^{N} \widehat{\boldsymbol{p}}_i^{\,2} \equiv \widehat{p}^2$ denotes the n-dimensional Laplacian and $V : \mathbb{R}^n \to \mathbb{R}^n$ is the interaction potential which we assume to satisfy:

(a) V is continuous and bounded from below;
(b) $\int \mathrm{d}^n y\, \mathrm{e}^{-\beta V(y)} < \infty$ for all $\beta > 0$;
(c) V is invariant under permutations of indistinguishable particles.

For the precise definition of the Hamiltonian as a self-adjoint operator on $L^2(\mathbb{R}^n)$ see theorem X.32 in [8].

Conditions (a) and (b) are slightly more restrictive than necessary, although they are appealing in their simplicity. The first condition ensures that the Hamiltonian is well defined and the continuity assumption also simplifies the taking of the continuum limit—if the original potential is not continuous, it can be replaced by a suitable continuous approximation. The second condition requires a sufficiently fast growth of the potential at infinity and something like it is necessary for guaranteeing the convergence of the canonical traces, $\mathrm{Tr}\, \mathrm{e}^{-\beta \widehat{H}}$. It also gives the precise meaning for what we meant by a sufficiently binding potential earlier. The third requirement is more of a consistency condition than a restriction.

The classical partition function of this system is defined by

$$Z_{\mathrm{cl}}(\beta) \equiv \int \frac{\mathrm{d}^n y}{(2\pi\beta)^{n/2}}\, \mathrm{e}^{-\beta V(y)} \tag{1}$$

and clearly (b) is equivalent to the assumption that the classical partition function converges for all positive temperatures. It is also clear that if V is positive, which could be achieved by adding a constant to the potential, then $Z_{\mathrm{cl}}(\beta)$ decreases monotonically.

For the sake of simplicity, we shall now assume that there is only one kind of indistinguishable particles in the system—the generalization of our results to systems with many types of particle should be obvious. Then the physical Hilbert space of the system is not given by the full $L^2(\mathbb{R}^n)$, but by the subspace consisting of wavefunctions of either symmetric (bosons) or antisymmetric (fermions) under particle permutations. To deal with this complication, we shall need to define some simplified notation related to the particle permutations.

For each element s in the permutation group S_N, we shall define the corresponding particle permutation by $s(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_N) = (\boldsymbol{x}_{s(1)}, \ldots, \boldsymbol{x}_{s(N)})$. As defined by this formula, each particle permutation is a linear operator on $\mathbb{R}^{Nd}$ and the mapping from S_N to these operators then forms a unitary representation of the permutation group—we do not distinguish here between an element of S_N and the corresponding operator. In particular, all of these operators are orthogonal, $s^{\mathrm{T}} s = 1$, and their determinants are equal to the parity of the corresponding permutation, denoted here by $(-)^s$. With this notation, the projection operator $\widehat{P}_\pm$ onto the physical Hilbert space can be defined by

$$\widehat{P}_\pm \psi(x) = \frac{1}{N!} \sum_{s \in S_N} (\pm)^s \psi(sx)$$

where $(+)^s = +1$ and the upper sign is used for bosons and the lower for fermions. The trace of any trace-class operator (see chapter VI.6 of [9] for a definition and basic properties of the trace) $\widehat{A} \in \mathcal{T}_1(L^2(\mathbb{R}^n))$ over the physical subspace can similarly be obtained from

$$\mathrm{Tr}_{\mathrm{phys}}(\widehat{A}) = \mathrm{Tr}\left(\widehat{P}_\pm \widehat{A}\right) = \mathrm{Tr}\left(\widehat{A}\widehat{P}_\pm\right).$$

Note also that condition (c) is now equivalent to requiring that $V(sx) = V(x)$ for all permutations s.

3. Definition and convergence of the lattice approximation

We shall now state a theorem which can be used to obtain the lattice observables for energy moments and which proves the convergence of the continuum limit of the lattice approximation. We prove the theorem for the canonical case with a temperature that can have an imaginary part—this complex temperature trace is needed for the computation of the energy moments in precanonical ensembles as we shall see later. The proof of the theorem is given in appendix B.

Theorem 1. *Define* $\widehat{T}(z) = \exp(-z\frac{1}{2}\widehat{p}^2)\exp(-zV(\widehat{x}))$ *on the half-plane* $\operatorname{Re} z > 0$. *Then for all* $k \geqslant 0$ *and* $\operatorname{Re} z > 0$,

$$\operatorname{Tr}\left(\widehat{P}_{\pm}\widehat{H}^k \mathrm{e}^{-z\widehat{H}}\right) = \lim_{L\to\infty} (-1)^k \frac{\mathrm{d}^k}{\mathrm{d}z^k} \operatorname{Tr}\left(\widehat{P}_{\pm}\widehat{T}(z/L)^L\right).$$

All of these functions are analytic on the right half-plane and the convergence is uniform on compact subsets of the half-plane. In addition, for any fixed k *and* $0 < c \leqslant c'$, *these functions are uniformly bounded on the closed strip* $c \leqslant \operatorname{Re} z \leqslant c'$ *of the half-plane.*

The proof contains the following explicit representation for the traces on the right-hand side:

$$\operatorname{Tr}\left(\widehat{P}_{\pm}\widehat{T}(z/L)^L\right) = \frac{1}{N!}\sum_{s\in S_N} (\pm)^s \int \mathrm{d}^{nL}x\; K_L(sx(L), x; z) \tag{2}$$

where the 'lattice kernel function' is defined by

$$K_L(b, x; z) = \left(\frac{L}{2\pi z}\right)^{Ln/2} \exp\left[-\frac{1}{z}P_L(b, x) - zV_L(x)\right]$$

and the 'lattice kinetic energy' P_L and the 'lattice potential energy' V_L are

$$P_L(b, x) = \frac{L}{2}\sum_{k=1}^{L} |x(k) - x(k-1)|^2 \qquad \text{with} \quad x(0) = b$$

$$V_L(x) = \frac{1}{L}\sum_{k=1}^{L} V(x(k)).$$

This result is well known and widely used, it was presented already in the classic work by Feynman and Hibbs [10] and it looks like a simple consequence of the Trotter product formula. However, the proofs of the Trotter formula [11] only imply a strong convergence for our generators and it is necessary to do more work to prove the convergence of the traces. The convergence of the Trotter formula in the trace-norm has been proven [12] if one of the generators has a resolvent in a trace-ideal $\mathcal{T}_p$ and this can be used for proving the trace-norm convergence for systems which are contained in a bounded region $\Lambda \subset \mathbb{R}^n$, see [13] for details. However, choosing $L^2(\Lambda^N)$ as the basic Hilbert space would require considering the effect of different boundary conditions and dealing with more complicated lattice kernels. We consider the present formulation, where the necessary bounding is achieved by a fast growth of the potential, to be more practical.

The sum in (2) contains $N!$ terms and thus its computation would be impossible for even fairly small values of N. However, the result of the integral, in fact, only depends on the conjugate class of the permutation and thus the N^N increase of the number of terms to compute

can be reduced to an exponential e^N increase. To be precise, it is shown in appendix C, that

$$\mathrm{Tr}\left(\widehat{P}_\pm \widehat{T}(z/L)^L\right) = \frac{1}{N!}\sum_{\lambda_1=1}^{N}(\pm 1)^{N-\lambda_1}\sum_{\lambda_2=0}^{\lambda_1}\cdots\sum_{\lambda_N=0}^{\lambda_{N-1}}\delta\left(\sum_{l=1}^{N}\lambda_l - N\right)$$
$$\times c_{(\lambda)}\int \mathrm{d}^{nL}x\; K_L(s_{(\lambda)}x(L), x; z) \tag{3}$$

where $s_{(\lambda)}$ is a representative of the conjugate class defined by the parameters (λ_l) as explained in appendix C and $c_{(\lambda)}$ counts the number of elements in the conjugate class, as given by equation (C2). The sum goes over those decreasing sequences of N non-negative integers whose sum equals N; in the above equation, δ is the Kronecker symbol which ensures that the last condition holds.

If all the terms in the sum are relevant, then only about $N \leqslant 30$ can be handled by the above formula (for $N = 30$ there are 5604 terms in the sum). However, typically only a fraction of the terms are relevant and the sum can also be used for a larger number of particles. For positive temperatures, the estimation of which terms will be relevant can be done fairly easily by using lemma 3 given in appendix B. According to the lemma, the following extension of the Golden–Thompson–Symanzik inequality (see theorem 9.2 in [14]) is valid for all $\beta > 0$,

$$\int \mathrm{d}^{nL}x\; K_L(sx(L), x; \beta) \leqslant \int \frac{\mathrm{d}^n y}{(2\pi\beta)^{n/2}}\;\exp\left[-\frac{1}{2\beta}|sy - y|^2 - \beta V(y)\right] \tag{4}$$

where the right-hand side is much easier to evaluate than the left-hand side. This also gives one more proof of the well known result that only the identity permutation is relevant for dilute gases, for which the typical classical distance between the particles is larger than the thermal wavelength $\sqrt{2\pi\beta}$.

4. Lattice operators for energy moments

We shall next derive a set of lattice operators whose expectation values will converge in the continuum limit to the canonical expectation values of powers of the Hamiltonian. By the results of section 3,

$$\mathrm{Tr}\left(\widehat{P}_\pm \widehat{H}^k \mathrm{e}^{-z\widehat{H}}\right) = \lim_{L\to\infty}\frac{1}{N!}\sum_{s\in S_N}(\pm)^s\int \mathrm{d}^{nL}x\;(-1)^k\frac{\mathrm{d}^k}{\mathrm{d}z^k}K_L(sx(L), x; z)$$

where we have moved the differentiation inside the integral—it is part of the proof in appendix B that this is allowed.

By using the Leibniz rule and a bit of algebra we can compute the derivatives,

$$(-1)^k\frac{\mathrm{d}^k}{\mathrm{d}z^k}K_L(sx(L), x; z) = Q_k(x; L, z)K_L(sx(L), x; z)$$

where

$$Q_k(x; L, z) = \sum_{j=0}^{k}\binom{k}{j}V_L^{k-j}\sum_{i=0}^{j}\binom{j}{i}\frac{(\frac{1}{2}Ln - 1 + j)!}{(\frac{1}{2}Ln - 1 + i)!}(-P_L)^i z^{-j-i}. \tag{5}$$

We call Q_k the lattice operator for the kth energy moment. The complicated form might be a bit surprising—the lattice action, after all, is simple and it can be guessed from the expected continuum limit. We might thus expect to find a lattice Hamiltonian equally easily, the guess 'βH_L' $= \beta V_L - \frac{1}{\beta}P_L$ which goes to '$\int_0^\beta \mathrm{d}\tau\;[V(x(\tau)) - \frac{1}{2}\dot{x}^2(\tau)]$' in the continuum limit, being most natural.

In fact, the natural guess is not too far away, since by (5)

$$\beta Q_1(x; L, \beta) = \beta V_L - \frac{1}{\beta} P_L + \frac{Ln}{2} \tag{6}$$

measures the expectation value of the Hamiltonian normalized by the inverse temperature β. The important difference is that the other energy moments are *not* measured by Q_1^k. Instead, each power needs a separate 'renormalization' term, note that the extra factors all contain $\frac{1}{2}Ln$ which diverges in the continuum limit.

We have given (5) in a form from which it is apparent that it is the 'kinetic energy' part of the Hamiltonian, given by P_L, which becomes renormalized. This is very natural, since the path integration which gives the continuum results goes over continuous but non-differentiable paths and thus the derivatives should diverge in the continuum limit. In addition to (6), we have

$$Q_2(x; L, z) = Q_1^2(x; L, z) + \frac{1}{z^2}\left(\tfrac{1}{2}Ln - 2P_L\right) \tag{7}$$

which can be used for concluding that both the expectation value and the variance of the lattice operator P_L diverge as $\frac{1}{2}Lnz$ in the continuum limit.

5. Monte Carlo algorithm for the generation of the canonical lattice distribution

We saw in section 4 that the term P_L diverges proportionally to the lattice size L in the continuum limit. On the other hand, V_L has a finite continuum limit. Consider generating the lattice kernel distribution $\propto \exp(-\beta V_L - \frac{1}{\beta} P_L)$ by a straightforward Metropolis algorithm [15], i.e. make random, local changes to the lattice configuration and keep the changes with probability $\exp(-\Delta S_\beta)$. Since P_L diverges, for large L most changes will be rejected because of the large jump they cause in P_L and the hit-rate of the Metropolis algorithm is not very good. However, on the other hand, P_L is a simple quadratic function which does not depend on the system, i.e. on the potential or on the nature of the particles. We shall now present an algorithm which takes advantage of these simplifications and which behaves better for large L.

The aim is to generate the canonical lattice distribution for x,

$$\int \mathrm{d}^{nL}x \left(\frac{L}{2\pi\beta}\right)^{Ln/2} \exp\left[-\frac{1}{\beta} P_L(sx(L), x) - \beta V_L(x)\right]$$

where s is any permutation and $\beta > 0$. Let us first separate the lattice copy containing the jump induced by the permutation and define $y = x(L)$. In the remaining $n(L-1)$ integrals we make the following change of variables from x to u:

$$x(k) = c(k) + \sum_{l=1}^{k} u(l) \qquad \text{for all} \quad 1 \leqslant k \leqslant L-1 \tag{8}$$

where $c(k) = (1 - \frac{k}{L})sy + \frac{k}{L}y$. The Jacobian of this change of variables is one. As earlier, for $k = 0$ and L we define also $x(0) = sy$ and $x(L) = y$. The inverse relation is then $u(k) = x(k) - x(k-1) - \frac{1}{L}(y - sy)$, which shows that

$$\begin{aligned} P_L(sx(L), x) &= \tfrac{1}{2}L \sum_{k=1}^{L} |x(k) - x(k-1)|^2 \\ &= \tfrac{1}{2}L \sum_{k=1}^{L-1} |u(k)|^2 + \tfrac{1}{2}L \left| \sum_{k=1}^{L-1} u(k) \right|^2 + \tfrac{1}{2}|y - sy|^2. \end{aligned}$$

A choice of other than a straight path $c(k)$ does not lead to as simple a formula, since then there would be a cross term linking y and u which is absent for a path with constant increments of c.

For reasons that will become apparent in the next section, we shall use a separate notation β' for the 'kinetic temperature'. Thus with the assumption that at this point $\beta' = \beta$, we can express the lattice integral as

$$\int \frac{\mathrm{d}^n y}{(2\pi\beta')^{n/2}} \exp\left[-\frac{1}{2\beta'}|sy - y|^2 - \beta V(y)\right] \tag{9a}$$

$$\times \int \frac{\mathrm{d}^{n(L-1)} v}{(2\pi)^{(L-1)n/2}} \mathrm{e}^{-\frac{1}{2}v^2} L^{\frac{n}{2}} \exp\left[-\tfrac{1}{2}\sum_{i=1}^{N}\left|\sum_{k=1}^{L-1} \boldsymbol{v}_i(k)\right|^2\right] \tag{9b}$$

$$\times \exp\left\{-\frac{\beta}{L}\sum_{k=1}^{L-1}[V(x(k)) - V(y)]\right\} \tag{9c}$$

where we have scaled the temperature away from the lattice fluctuations, i.e. made the change of variables $u = \sqrt{\beta'/L}\, v$. This means that in the last line we have used the definition

$$x(k) = sy + \frac{k}{L}(y - sy) + \sqrt{\frac{\beta'}{L}}\sum_{l=1}^{k} v(l) \qquad \text{for all} \quad 1 \leqslant k \leqslant L-1. \tag{10}$$

The algorithm consists of generating each of the distributions (9*a*)–(9*c*) separately.

(a) Compute the classical integral (9*a*). This is needed for the normalization of the next Metropolis step and, by (4), it can be used also for determining whether a computation of the full term is necessary.
(b) Use, for example, a Metropolis algorithm to generate the classical distribution in (9*a*) for y.
(c) Generate the distribution for v by applying a Metropolis check with normal distributed trials for each coordinate of v. This can be done independently of the generation of y and, in fact, each classical coordinate can also be generated separately. This problem is equivalent to the generation of a self-recurring one-dimensional random walk of L steps, the cumulative sums of v corresponding to the positions of the walker. Note that the constants in (9*b*) have already been chosen so that the result is a probability distribution.
(d) Apply a Metropolis check implied by (9*c*) for the generated values of y and v.

In practice, it is best to store the values of $x(k)$ and to use a separate routine for the generation of the trials for quantum fluctuations of each particle. The idea behind the algorithm is to generate a path of possible quantum fluctuations for a classically distributed particle and to test for the acceptance of the whole fluctuation path simultaneously. Since P_L does not enter this acceptance test at all, this avoids the large rejectance rate from which a local change made in just one copy suffers. However, as the proposed change is not local, the evaluation of the last Metropolis check is made more difficult and time consuming—this should be more than compensated for by the improved acceptance rate.

Let us also briefly comment on the generation of the classical distribution when s is not the identity permutation. Then s has at least one cycle of length $\ell > 1$ and assume, for simplicity, that it is composed of the first particle labels. By the same change of variables as before, i.e. by defining $\boldsymbol{y}_i = \sum_{l=1}^{i} \boldsymbol{u}_l$, we then obtain

$$\sum_{i=1}^{\ell} |\boldsymbol{y}_{s(i)} - \boldsymbol{y}_i|^2 = \sum_{i=2}^{\ell} |\boldsymbol{u}_i|^2 + \left|\sum_{i=2}^{\ell} \boldsymbol{u}_i\right|^2.$$

This is independent of $\boldsymbol{u}_1$, i.e. of one of the particle positions. On the other hand, it also implies that the distance between any two particles is of the order of $\sqrt{\beta'}$. It is thus possible to think of each ℓ-cycle as describing a cluster of ℓ particles which move together in the external potential. Similarly, especially for higher temperatures, it is best to generate the combinations $\boldsymbol{u}$ from the Gaussian distribution and use the potential only for $\boldsymbol{u}_1$ and, of course, for the Metropolis check.

6. Lattice evaluation of Gaussian traces

The Gaussian ensemble has been introduced in [5], where its connections to the canonical and the microcanonical ensemble were also illustrated. The ensemble is defined by the unnormalized density operator

$$\widehat{\rho}_\varepsilon(E) = \frac{1}{\sqrt{2\pi\,\varepsilon^2}} \exp\biggl[-\frac{1}{2\varepsilon^2}(\widehat{H}-E)^2\biggr].$$

It was proven there that for any observable $\widehat{A}$ which satisfies the condition $\mathrm{Tr}(|\widehat{A}|\mathrm{e}^{-\beta\widehat{H}}) < \infty$ for all $\beta > 0$, we have the following integral representation:

$$\mathrm{Tr}\left(\widehat{A}\widehat{\rho}_\varepsilon(E)\right) = \int_{\beta-\mathrm{i}\infty}^{\beta+\mathrm{i}\infty} \frac{\mathrm{d}z}{2\pi\,\mathrm{i}}\, \mathrm{e}^{\frac{1}{2}\varepsilon^2 z^2 + zE}\, \mathrm{Tr}\left(\widehat{A}\mathrm{e}^{-z\widehat{H}}\right) \tag{11}$$

where β is a positive constant.

For numerical computations, the proper choice of β is very important. Suppose we wish to examine the behaviour of the system near the energy scale E_0 using the energy resolution $\varepsilon < E_0$. At $E = E_0$, it was shown in [5] that the best choice for β is the unique positive solution to the equation

$$E_0 + \beta\varepsilon^2 = \overline{H} \equiv \frac{\mathrm{Tr}(\widehat{P}_\pm \widehat{H}\mathrm{e}^{-\beta\widehat{H}})}{\mathrm{Tr}(\widehat{P}_\pm \mathrm{e}^{-\beta\widehat{H}})}. \tag{12}$$

In practice, it is usually easier to begin with the temperature $1/\beta$ and then solve E_0 from equation (12). The natural energy resolution parameter is then $\varepsilon' = \beta\varepsilon$ and as the energy parameter, it will be easiest to use the scaled difference $E' = \beta(E - E_0)$. Similarly, the moments are most naturally computed as

$$\beta^k \langle (\widehat{H} - E - \beta\varepsilon^2)^k \rangle_{E,\varepsilon}^{\mathrm{gauss}} = \langle (\beta(\widehat{H} - \overline{H}) - E')^k \rangle_{E,\varepsilon}^{\mathrm{gauss}}.$$

We shall derive the Gaussian lattice expressions for this set of moments—all others can, of course, be computed from these.

With the present assumptions, $\widehat{A} = \widehat{P}_\pm$ satisfies the condition for the use of the integral representation (11). Therefore, for any real t and $\beta > 0$ the following holds:

$$\mathrm{Tr}\left(\widehat{P}_\pm \mathrm{e}^{t(\widehat{H}-E-\beta\varepsilon^2)}\widehat{\rho}_\varepsilon(E)\right) = \int_{\beta-\mathrm{i}\infty}^{\beta+\mathrm{i}\infty} \frac{\mathrm{d}z}{2\pi\,\mathrm{i}}\, \mathrm{e}^{\frac{1}{2}\varepsilon^2 t^2 + t\varepsilon^2(z-\beta)} \mathrm{e}^{\frac{1}{2}\varepsilon^2 z^2 + zE}\, \mathrm{Tr}\left(\widehat{P}_\pm \mathrm{e}^{-z\widehat{H}}\right). \tag{13}$$

Applying proposition 2 then shows that both sides can be analytically continued to entire functions and that for all $k \geqslant 0$,

$$\begin{aligned}\mathrm{Tr}\left(\widehat{P}_\pm(\widehat{H}-E-\beta\varepsilon^2)^k\widehat{\rho}_\varepsilon(E)\right) &= \frac{\mathrm{d}^k}{\mathrm{d}t^k}\, \mathrm{Tr}\left(\widehat{P}_\pm \mathrm{e}^{t(\widehat{H}-E-\beta\varepsilon^2)}\widehat{\rho}_\varepsilon(E)\right)\Big|_{t=0} \\ &= \int_{\beta-\mathrm{i}\infty}^{\beta+\mathrm{i}\infty} \frac{\mathrm{d}z}{2\pi\,\mathrm{i}}\, \frac{\mathrm{d}^k}{\mathrm{d}t^k}\, \mathrm{e}^{\frac{1}{2}\varepsilon^2 t^2 + t\varepsilon^2(z-\beta)}\Big|_{t=0} \mathrm{e}^{\frac{1}{2}\varepsilon^2 z^2 + zE}\, \mathrm{Tr}\left(\widehat{P}_\pm \mathrm{e}^{-z\widehat{H}}\right).\end{aligned}$$

The derivative on the right-hand side can be computed from the generating function of the Hermite polynomials H_k,

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k}\,\mathrm{e}^{\frac{1}{2}\varepsilon^2 t^2 + t\varepsilon^2(z-\beta)}\Big|_{t=0} = \big(\mathrm{i}\varepsilon/\sqrt{2}\big)^k H_k\big(-\mathrm{i}(z-\beta)\varepsilon/\sqrt{2}\big)$$
$$= \varepsilon^k \sum_{j=0}^{\lfloor k/2\rfloor} \frac{k!}{2^j j!(k-2j)!}[(z-\beta)\varepsilon]^{k-2j}.$$

Parametrizing the integral as $z = \beta(1+\mathrm{i}\alpha)$ and using the scaled variables ε' and E' we have the result

$$\beta^k \operatorname{Tr}\big(\widehat{P}_\pm(\widehat{H} - E - \beta\varepsilon^2)^k \widehat{\rho}_\varepsilon(E)\big) = \beta\varepsilon'^k \mathrm{e}^{-\frac{1}{2}\varepsilon'^2 + E' + \beta\overline{H}} \sum_{j=0}^{\lfloor k/2\rfloor} \frac{k!}{2^j j!(k-2j)!} I_{k-2j}(E', \varepsilon'; \beta)$$

where

$$I_k(E, \varepsilon; \beta) = \mathrm{i}^k \int_{-\infty}^{\infty} \frac{\mathrm{d}\alpha}{2\pi} (\varepsilon\alpha)^k \mathrm{e}^{-\frac{1}{2}\varepsilon^2\alpha^2 + \mathrm{i}\alpha(E+\beta\overline{H})} \operatorname{Tr}\big(\widehat{P}_\pm \mathrm{e}^{-\beta(1+\mathrm{i}\alpha)\widehat{H}}\big).$$

Let us next define the function

$$g(\alpha; \beta) = \mathrm{e}^{\mathrm{i}\alpha\beta\overline{H}} \operatorname{Tr}\big(\widehat{P}_\pm \mathrm{e}^{-\beta(1+\mathrm{i}\alpha)\widehat{H}}\big)$$

using which we have for even k,

$$I_k(E, \varepsilon; \beta) = (-1)^{k/2} \operatorname{Re} \int_{-\infty}^{\infty} \frac{\mathrm{d}\alpha}{2\pi} \mathrm{e}^{\mathrm{i}\alpha E} (\varepsilon\alpha)^k \mathrm{e}^{-\frac{1}{2}\varepsilon^2\alpha^2} g(\alpha; \beta) \tag{14}$$

and for odd k,

$$I_k(E, \varepsilon; \beta) = (-1)^{(k+1)/2} \operatorname{Im} \int_{-\infty}^{\infty} \frac{\mathrm{d}\alpha}{2\pi} \mathrm{e}^{\mathrm{i}\alpha E} (\varepsilon\alpha)^k \mathrm{e}^{-\frac{1}{2}\varepsilon^2\alpha^2} g(\alpha; \beta). \tag{15}$$

This shows that if we know the function g in some suitably dense set of points, we can approximate I_k for any k, ε' and E' by multiplying the known values by the weight function and then performing a discrete Fourier transform.

Thus $g(\alpha; \beta)$ contains all information even about the arbitrarily high-energy end of the spectrum and it should not come as a surprise that its computation is not easy in general. Fortunately, it has a lattice approximation and there is a relatively straightforward Monte Carlo algorithm for the evaluation of the lattice integral.

By theorem 1, the lattice approximations

$$g_L(\alpha; \beta) = \mathrm{e}^{\mathrm{i}\alpha\beta\overline{H}} \operatorname{Tr}\big(\widehat{P}_\pm \widehat{T}(\beta(1+\mathrm{i}\alpha)/L)^L\big) \tag{16}$$

converge to $g(\alpha; \beta)$ as $L \to \infty$ and they are all uniformly bounded in α. If we define $I_k(E, \varepsilon; \beta, L)$ by replacing g by g_L in (14) and (15), we can thus rely on the dominated convergence theorem and conclude that

$$\beta^k \langle(\widehat{H} - E - \beta\varepsilon^2)^k\rangle_{E,\varepsilon}^{\text{gauss}} = \lim_{L\to\infty} \varepsilon'^k \sum_{j=0}^{\lfloor k/2\rfloor} \frac{k!}{2^j j!(k-2j)!} \frac{I_{k-2j}(E', \varepsilon'; \beta, L)}{I_0(E', \varepsilon'; \beta, L)}$$

where $E' = \beta(E + \beta\varepsilon^2 - \overline{H})$ and $\varepsilon' = \beta\varepsilon$. Since the Gaussian partition function measures the density of states directly, it is of a special interest and, by the above results, it has the lattice approximation

$$\operatorname{Tr}\big(\widehat{P}_\pm \widehat{\rho}_\varepsilon(E)\big) = \lim_{L\to\infty} \mathrm{e}^{\frac{1}{2}\beta^2\varepsilon^2 + \beta E} \beta I_0(E', \varepsilon'; \beta, L).$$

6.1. Monte Carlo algorithm for the Gaussian integrals

Let us now go through the lattice algorithm for the evaluation of g_L when α and $\beta > 0$ are given. The definition (16) translates into the lattice integral

$$g_L(\alpha;\beta) = \frac{1}{N!}\sum_{s\in S_N}(\pm)^s \mathrm{e}^{\mathrm{i}\alpha\beta\overline{H}} \int \mathrm{d}^{nL}x\; K_L(sx(L),x;\beta(1+\mathrm{i}\alpha))$$

$$= \frac{1}{N!}\sum_{s\in S_N}(\pm)^s (1+\alpha^2)^{Ln/4} \int \mathrm{d}^{nL}x \left(\frac{L}{2\pi\beta'}\right)^{Ln/2} \exp\left[-\frac{1}{\beta'}P_L - \beta V_L\right]$$

$$\times \exp\left[\mathrm{i}\alpha\left(\beta\overline{H} - \beta V_L + \frac{1}{\beta'}P_L - \tfrac{1}{2}Ln\frac{\arctan\alpha}{\alpha}\right)\right] \tag{17}$$

where β' is a shorthand for the combination $\beta(1+\alpha^2)$ and $P_L = P_L(sx(L),x)$ and $V_L = V_L(x)$ as before. In addition, since the kernel of $\widehat{T}$ is obtained by analytic continuation, the arcus tangent in the above should be chosen from the primary branch, i.e. $\arctan\alpha \in (-\frac{1}{2}\pi, \frac{1}{2}\pi)$.

Now we can also finally explain why we wanted to include the term $\beta\overline{H}$ into the lattice integration. Suppose, for simplicity, that of the possible classes of permutations one dominates over the others—this is natural if the quantum behaviour of the system comes from the formation of correlated particle clusters of fixed size for the temperature in question (we saw in section 5 that when a permutation is decomposed into cycles, each cycle can naturally be identified as such a correlated cluster of particles). Then for this permutation and for small α, the last term in parentheses in (17) has an almost zero expectation value under the distribution in front of it—this follows from a comparison with the definition of Q_1 in (6). Since the α-integration is concentrated near zero, choosing $\beta\overline{H}$ as the energy origin thus gives the smallest oscillations for the integrand in (17). This also shows that for a fixed lattice size L, it is best to use the value evaluated from the same lattice for $\overline{H}$—a continuum extrapolation is not necessary or desirable.

The first term in the lattice integral in (17) gives simply a canonical distribution, which can be generated by the algorithm explained in section 5. The evaluation of the remaining exponential is then a simple computation of trigonometric functions.

Computation of g for all α would then enable the solution of the full microcanonical spectrum of the Hamiltonian. As the energy is increased, the dependence of the position of the spectral lines depends on ever finer structure of the potential. Since, eventually, a computer cannot hold the values of the potential at the accuracy needed, there must be a catch somewhere. In our case, this is most easily seen in the prefactor $(1+\alpha^2)^{Ln/4}$ which must the cancelled by the result from the lattice integral—otherwise g would diverge in the continuum limit. This means that as α is increased, the oscillations of the lattice observable increase rapidly and the maximum value of α that can be computed this way is limited by the number of Monte Carlo sweeps feasibly performed and by the resulting inaccuracy of the result from the lattice integration.

7. An example: charged particles in a quadratic potential

We now want to check how well the algorithm works in practice. We shall consider particles living in a three-dimensional world and bound by a harmonic potential. After we have checked that the algorithm works for this analytically solvable case, we shall add a Coulomb interaction

between the particles. In other words, we shall consider the potential

$$V(x) = \sum_{i=1}^{N} \tfrac{1}{2}\omega^2 \boldsymbol{x}_i^2 + \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} \frac{1}{2} \frac{q^2}{|\boldsymbol{x}_i - \boldsymbol{x}_j|}$$

where N gives the number of particles, ω is the binding frequency and q denotes the charge of one particle.

The scaling properties of the potential part of the action are now simple,

$$\beta V_L(\sqrt{\beta}x) = \beta^2 V_Q(x) + \sqrt{\beta} V_C(x)$$

where

$$V_Q(x) = \frac{1}{L} \sum_{k=1}^{L} \sum_{i=1}^{N} \frac{1}{2} \omega^2 \boldsymbol{x}_i^2(k)$$

and

$$V_C(x) = \frac{1}{L} \sum_{k=1}^{L} \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} \frac{1}{2} \frac{q^2}{|\boldsymbol{x}_i(k) - \boldsymbol{x}_j(k)|}.$$

By first scaling the temperature β away from the kinetic term, and by subsequent differentiation we can thus derive a second set of observables for the computation of the energy moments. The energy expectation value can now be measured either by using Q_1 in (6) or by using the observable

$$\bar{Q}_1 = 2V_Q + \tfrac{1}{2} V_C \tag{18}$$

which has been rescaled back to the original lattice variables. Similarly, the second moment is given by (7) or by

$$\bar{Q}_2 = \bar{Q}_1^2 + \frac{1}{\beta}(2V_Q - \tfrac{1}{4} V_C)$$

and the other observables $\bar{Q}_k$ could be computed equally easily. Note that, again, simply taking the second power of the observable giving the energy expectation value does not yield the correct result.

The advantage of using both of these two kinds of observables comes from their very different dependence on the lattice kinetic energy. For large L the second set of observables is better, since it does not contain any cancellation of two large numbers as is necessary for the first set of observables. For all values of L, both methods should nevertheless yield mutually consistent results and we, in fact, used this for fine-tuning and checking of the Monte Carlo algorithm.

We have presented the results of the Monte Carlo computation of the energy moments of a system of two distinguishable three-dimensional particles at two different temperatures in table 1. The results were computed by performing 3×10^6 sweeps on an AlphaStation 1000 computer which took a few hours of computer time. The error estimates for the moments were computed using the bootstrap method and the number of sweeps was enough to obtain a clean signal for all of the measured six moments.

In both cases, we used a lattice with $L = 8$ steps and it is clear that the results are very close to the continuum limit. In fact, the main difference with the classical $L = 1$ results comes already at the addition of one classical copy for the quantum fluctuations, i.e. already at $L = 2$. As expected, the convergence to the continuum limit is also faster for smaller values of β.

Table 1. First six moments $\langle \beta^k (\widehat{H} - \langle \widehat{H} \rangle)^k \rangle$ of a two-particle system in a three-dimensional space. Only the $s = \mathrm{id}$ term has been considered, i.e. the particles have been assumed to be distinguishable, and the binding frequency is $\omega = 1$. The results were obtained from an eight-step lattice using either Q_k as defined in equation (5) or $\bar{Q}_k$ as defined in section 7. For the non-charged case, the known continuum value is also given.

β	q	Obs.	$\langle \beta \widehat{H} \rangle$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$k = 6$
1.0	0	Q	6.483(3)	5.59(2)	11.7(2)	131(2)	778(25)	7 100(290)
		$\bar{Q}$	6.481(4)	5.49(3)	11.4(3)	123(5)	790(100)	8 800(2500)
		c.l.	6.4919	5.524	11.95	127.6	804.4	7 664
	2	Q	8.735(3)	5.03(2)	10.8(2)	110(2)	647(23)	6 150(430)
		$\bar{Q}$	8.665(4)	4.69(3)	9.8(3)	98(7)	730(170)	9 400(4000)
0.1	0	Q	6.000(3)	5.99(2)	11.8(2)	147(2)	832(27)	8 930(440)
		$\bar{Q}$	6.000(2)	5.970(12)	11.87(15)	143(2)	886(40)	9 340(790)
		c.l.	6.0050	5.995	12.00	143.8	863.4	8 629
	5	Q	6.583(3)	5.82(2)	11.6(2)	139(2)	852(26)	8 690(370)
		$\bar{Q}$	6.585(2)	5.79(2)	11.8(2)	138(3)	907(65)	10 300(1500)

We have included the $\beta = 1$ case as a warning about the possible systematical errors induced by the Monte Carlo algorithm. It is clear that the two lattice observables at $\beta = 1$ do not yield results which are as mutually consistent as those at $\beta = 0.1$. This is most clearly seen from the energy expectation value of the charged case when the two results are not consistent within their statistical errors when $\beta = 1$, but they *are* consistent for $\beta = 0.1$, even though we have increased the charge for the second case.

The source of these systematic errors is in the *a priori* distribution which is used for obtain samples to the Metropolis check. If the *a priori* distribution does not concentrate on the correct region of the configuration space, the resulting probability distribution in this region will be coarser than suggested by the number of sweeps used and the results, likewise, less accurate.

For instance, the differences in the non-charged case are caused by the use of quantum fluctuations as the *a priori* distribution although we are already in the low-temperature region $\beta \to \infty$ where the lowest energy state dominates. This could be remedied by also using sweeps with the usual Monte Carlo sampling where the binding potential $\mathrm{e}^{-\beta V_L}$ serves as the *a priori* distribution. The discrepancy in the low-temperature charged case is more problematic, since it is also caused by the singularity of the Coulomb distribution. The cure in the second case would be to replace the Coulomb distribution with some smooth approximation of it, but the effect of this change would then have to be analysed separately. Of course, it is always possible simply to increase the number of sweeps until the final probability distribution is smooth enough, but this might not always be feasible.

We also measured the Gaussian partition function, i.e. the density of states, for this system. The measurements were performed by computing $g_L(\alpha; \beta)$ for $|\alpha| \leqslant 1.6$ at intervals of 0.05 on a four-step lattice using $\beta = 0.1$. Samples were then taken from the measured values with the assumption of normally distributed errors and a discrete Fourier transformation with the correct weight was performed for each of these samples. We have given the stable results of this analysis in figure 1 for two different values for the energy resolution, $\varepsilon' = 1.5$ and 0.1. It is interesting that we were also able to obtain accurate results for the second case, for which the effect of the Gaussian weight is almost negligible.

Otherwise, the results behave exactly as expected. The most accurate results are obtained near $E' = 0$, which corresponds to $E \approx 40$ in the first case and to $E \approx 60$ in the second case. The accuracy deteriorates as the energy difference increases, faster in the case of smaller energy resolution. The addition of the Coulomb term also seems to only affect the density of states at the low-energy spectrum.

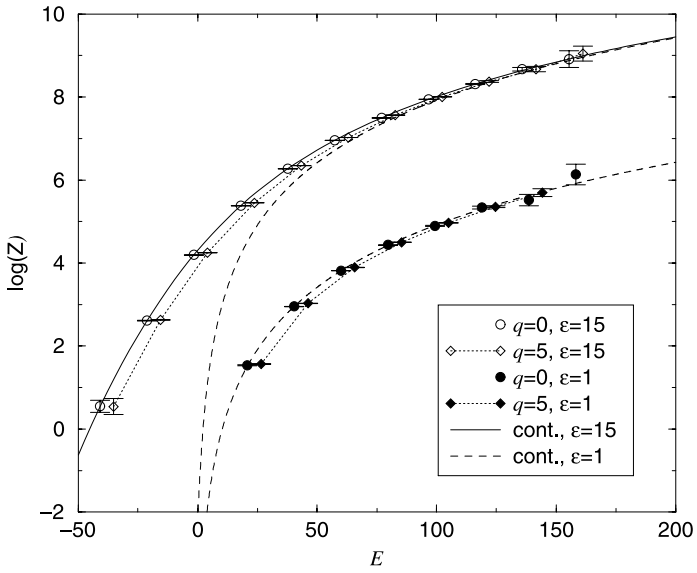

Figure 1. The base-ten logarithm of the Gaussian partition function obtained using a four-step lattice with $\omega = 1$ and $\beta = 0.1$ and the two values of the energy resolution shown. For the sake of clarity, the results for $\varepsilon = 1$ are plotted as three orders of magnitude smaller than they were measured. The full and broken curves represent the continuum results for $q = 0$; the broken curve has been plotted twice, both at its correct and reduced location.

The non-charged results were also compared with their continuum counterparts, which were computed using the known partition function

$$\mathrm{Tr}\left(\mathrm{e}^{-z\widehat{H}}\right)_{q=0} = \frac{1}{(2\sinh(z\omega/2))^{Nd}}.$$

If we remember that we used only an $L = 4$ lattice, the results presented in figure 1 are surprisingly accurate. When we later checked the difference between the classical $L = 1$ results and the full quantum results at $\beta = 0.1$ for the values of α used, the differences were almost negligible. This seems to imply that, at least for the harmonic potential, the classical density of states does give a highly accurate measurement of the high energy density of states of the quantum system.

In contrast, we were not able to measure the discrete structure of the energy spectrum by measuring g_L since this would have needed impractically high values for α. We also checked that for detecting the discrete spectrum, the use of large values of L is essential. Thus the present method is suitable only for measurement of quantities which do not depend directly on the discrete energy spectrum, but only on the density of states.

8. Prologue: improvements and a brief comparison with existing methods

A number of references came to our attention after the Monte Carlo simulations were finished and in this section we shall suggest some improvements motivated by them. We first present a better algorithm for the generation of the self-recurrent random walk inspired by the work done mostly by Freeman and Doll, see, e.g., [16]. The second subsection contains a discussion and references to other possible improvements.

8.1. Generation of the random walk by Fourier components

Suppose we make, instead of (8), a change of variables defined by a discrete Fourier sine transform,

$$a(k) = \frac{2}{L}\sum_{j=1}^{L}[x(j) - c(j)]\sin\left(\pi\frac{k}{L}j\right) \qquad \text{for all} \quad 1 \leqslant k \leqslant L-1.$$

The following 'orthogonality' relations, which can be found from [17] or computed directly by expressing sines as exponentials, are valid for the discrete sine transform,

$$\sum_{j=1}^{L}\sin\left(\pi\frac{k}{L}j\right)\sin\left(\pi\frac{k'}{L}j\right) = \frac{L}{2}\delta_{kk'}$$

$$\sum_{j=1}^{L}\sin\left(\pi\frac{k}{L}j\right)\sin\left(\pi\frac{k'}{L}(j-1)\right) = \begin{cases} \dfrac{L}{2}\cos\left(\pi\dfrac{k}{L}\right) & \text{for } k = k' \\ \dfrac{1-(-1)^{k+k'}}{2}\dfrac{\sin\left(\pi\frac{k}{L}\right)\sin\left(\pi\frac{k'}{L}\right)}{\cos\left(\pi\frac{k}{L}\right)-\cos\left(\pi\frac{k'}{L}\right)} & \text{for } k \neq k' \end{cases}$$

where we have assumed that both k and k' belong to $\{1, 2, \ldots, L-1\}$. Using these formulae it is easy to prove that the linear transformation which defines the change of variables is always invertible with an inverse

$$x(j) = c(j) + \sum_{k=1}^{L-1} a(k)\sin\left(\pi\frac{j}{L}k\right) \qquad \text{for all} \quad 0 \leqslant k \leqslant L \tag{19}$$

and that with the choice $c(k) = (1 - \frac{k}{L})sy + \frac{k}{L}y$ the kinetic term becomes simply

$$P_L = \tfrac{1}{2}|y - sy|^2 + \sum_{k=1}^{L-1}|a(k)|^2 L^2 \sin^2\left(\frac{\pi k}{2L}\right).$$

Since equation (9*b*) integrates to unity, we can then also deduce the Jacobian of this change of variables and conclude that (9*b*) can be replaced by

$$\int \frac{\mathrm{d}^{n(L-1)}a}{(2\pi)^{n(L-1)/2}}\prod_{k=1}^{L-1}\frac{1}{\sigma_k^n}\exp\left[-\tfrac{1}{2}\sum_{k=1}^{L-1}|a(k)|^2/\sigma_k^2\right] \tag{20}$$

where

$$\sigma_k^2 = \frac{\beta'}{2L^2\sin^2(\pi k/2L)}. \tag{21}$$

Therefore, we can replace the Monte Carlo step suggested in (c) in section 5 by a generation of the path $x(k)$ via the coefficients $a(k)$ in equation (19). Since all $a(k)$ have a normal distribution, their generation is fast and straightforward. Note also that the discrete Fourier transformation need not slow down the algorithm, since all necessary trigonometric functions can be computed before entering the Monte Carlo loop.

8.2. Discussion about the Fourier method and further improvements

Using a Fourier transformation of the path to express quantum statistical path integrals was suggested already by Feynman [10]. Later, Freeman and Doll [16] used a form with a cut-off for the number of Fourier modes in a Monte Carlo simulation of energies of quantum mechanical systems and they found a performance better than in a conventional Metropolis Monte Carlo simulation. However, the convergence of energy expectation values was found to be non-monotonic for this 'primitive Fourier method' and a relatively high number of Fourier modes, $k_{\max} \approx 50$, was necessary before the continuum value was approached [18, 19].

In our simulations, we did not see any non-monotonicity of the convergence of the lattice energy expectation values and, in addition, we found that a relatively coarse lattice was sufficient for the computation of nearly continuum results. Emboldened by these results we now suggest that using the version of the Fourier transformation given in the previous section we should be able to solve the problems related to the approach to the continuum limit that are manifest in the primitive Fourier method.

In other words, if the 'Matsubara frequencies' used in the primitive Fourier method are replaced by those derived in the previous section, i.e. if we replace

$$\sigma_k^2 = \frac{2\beta'}{(k\pi)^2}$$

by

$$\sigma_k^2 = \frac{\beta'}{2L^2 \sin^2(\pi k/2L)}$$

then the Fourier method should converge for smaller values of $k_{\max} = L - 1$. Similarly, the integrals needed in the previous energy estimators should probably be replaced by their discretized lattice versions. The idea of using the frequencies defined by equation (21) is not new and in [20] Coalson arrived essentially at the same conclusions as we do here. Nevertheless, apart from a few examples [21], we are not aware that it has have been widely used.

Since we did not use as complicated a system as in the references, it would now be necessary and interesting to repeat the test of the convergence and the comparison of computing times for the different methods that were used in [18, 22], respectively. For this comparison, we would like to point out that the differentiation which we used to produce the second set of observables, $\bar{Q}_k$ in section 7, would for a generic potential lead to what is essentially the virial estimator defined in [16]. Note, however, that our method also yields similar estimators for all other energy moments, not just for the first one. In addition, both of our lattice observables always give the same result on the same lattice, whereas the previous ones agree only in the continuum limit.

We can now also explain why the Fourier method is better than the conventional lattice Monte Carlo method: using the Fourier coefficients as the *a priori* distribution gets rid of the kinetic term which we suggested was the culprit for the slow convergence of the conventional Monte Carlo algorithm for large L. We can also explain why typically no estimators are given for energy moments beyond the first one: since each power of the kinetic term requires a separate renormalization, guessing the correct lattice observables from continuum expressions is difficult and the effect of using incorrect constants can easily be drastic. In addition, observables for the higher moments cannot be reduced to 'coordinate-space observables' (i.e. observables living in only one time slice) as is the case for the virial estimator of the energy for distinguishable particles.

Since the changes we suggest here make the Fourier lattice integrals *equivalent* to the discrete time-step integrals, the methods used for improving the performance of the earlier

methods should apply equally easily here. One could, for example, employ classical cluster algorithms [23] or use smeared or effective potentials [24]. One interesting application which we did not discuss here at all is in the computation of real-time correlation functions (for this see, e.g., [25, 26]).

Although the algorithm for the computation of Gaussian traces is not practical for resolving the discreteness of the energy spectrum, the lattice formulae derived are valid for arbitrarily fine energy precision. By following the steps which where used in [27], we could derive a lattice formulation with a complicated oscillating kernel which could be used to also solve the low-energy discrete spectrum. How high in energy and for how complicated systems this kernel can be used in practice needs testing.

9. Conclusions

We have presented a lattice Monte Carlo algorithm for the computation of energy moments in the canonical Boltzmann–Gibbs ensemble for a system of non-relativistic quantum mechanical particles. In fact, none of our results would have changed if we were to add a bounded classical observable $A(x)$ to the trace. Thus the present method can also be used for the evaluation of microcanonical corrections to the canonical expectation values of such observables as is explained in [5]. The overall performance of the algorithm, especially for high temperatures, was very good. In addition to the algorithm, we presented rigorous limits which can be used for a quantitative estimation of when and how the indistinguishability of particles will begin to affect the results.

We also proposed a lattice formulation which can be used for the evaluation of Gaussian expectation values and the Gaussian partition function of these systems. The algorithm we gave for the computation of the lattice integrals was shown to yield correct information about the density of states even for quite coarse lattices. If knowledge about the discrete energy spectrum is required, then using analytical methods or a modification of the algorithm seems to be necessary. The present method has the advantage that other non-canonical ensembles in addition to the Gaussian one, as well as many different energy resolutions, can be analysed from the same lattice data.

In both cases, we saw that the classical ensembles work surprisingly well and we noted how the intuitive understanding of the classical case can help in the computation of the full quantum results. The main use of the Gaussian lattice methods would probably be in the detection of phase transitions. Although it is necessary to increase the number of particles to infinity before a clean signal for a phase transition appears, such effects should be visible for a finite number of particles as well.

Acknowledgments

I am indebted for the motivation and advice provided by M Chaichian. I also wish to thank A Kupiainen and C Montonen for their comments on the manuscript and J Doll and D Freeman for comments and references.

Appendix A. Analytic integrals

The following result is a straightforward consequence of textbook results, but since it is central to all the results presented in the text, we give a brief derivation for it. The main content of the theorem is that an analytically parametrized integral is an analytic function of the parameter

if the integrand is uniformly dominated by an integrable function in every compact subset of the parameter space and that in this case it is also always possible to take differentiation of the parameter inside the integral.

Proposition 2. *Let Ω be an open subset of the complex plane and let X be a measure space which is σ-finite with respect to a positive measure μ. Consider a function $F : X \times \Omega \to \mathbb{C}$, which is measurable in the product space and which defines an analytic function $z \mapsto F(x, z)$ for almost every (with respect to μ) x. If for every compact subset K of Ω there is a function $g_K \in L^1(\mu)$ such that $|F(x, z)| \leqslant g_K(x)$ for almost every x and for every $z \in K$, then the function $f : z \mapsto \int \mathrm{d}\mu(x)\, F(x, z)$ is well defined and analytic in Ω. Then also for all $z \in \Omega$ and positive integers k, $\frac{\mathrm{d}^k}{\mathrm{d}z^k} f(z) = \int \mathrm{d}\mu(x)\, \frac{\mathrm{d}^k}{\mathrm{d}z^k} F(x, z)$.*

Proof. f is clearly well defined for all $z \in \Omega$. By the dominated convergence theorem it is also continuous and then an application of Morera's theorem shows that it is analytic. Using Cauchy's integral formula for the derivatives of f, followed by an application of Fubini's theorem then leads to the given expression for the derivatives. For more details see [28]. □

When using the proposition here, the σ-finiteness is obvious since we use it only for the counting measure on the positive integers and for Lebesque measures on Euclidean spaces. The measurability of F in the product space is also trivial since the functions we are dealing with will be continuous.

Appendix B. Proof of the lattice convergence theorem

We prove here how theorem 1 follows from assumptions (a) and (b) given in section 2. The proof consist of four parts: we first show that the $k = 0$ trace is an analytic function which works as a generating function for the other powers. We then derive an integral representation for the lattice traces and use this to prove their analyticity. Next we show that the lattice traces are uniformly bounded on compact subsets with a bound given in terms of the classical partition function. Since the traces converge for positive values of the parameter, theorem 1 follows from an application of the Vitali convergence theorem (see problem 33 in chapter I in [9]) and properties of analytic functions.

However, let us begin by defining the following two functions of a lattice configuration $x = (x(1), \ldots, x(L))$:

$$P_L(b, x) = \tfrac{1}{2} L \sum_{k=1}^{L} |x(k) - x(k-1)|^2 \qquad \text{with} \quad x(0) = b$$

and

$$V_L(x) = \frac{1}{L} \sum_{k=1}^{L} V(x(k)).$$

In the proof and in practical computations the following bound for their exponentials will become useful.

Lemma 3. *For any permutation $s \in S_N$ and for all $L \geqslant 1$ and $\beta > 0$,*

$$\int \mathrm{d}^{nL}x \left(\frac{L}{2\pi\beta}\right)^{Ln/2} \exp\left[-\frac{1}{\beta} P_L(sx(L), x) - \beta V_L(x)\right] \leqslant \int \frac{\mathrm{d}^n y}{(2\pi\beta)^{n/2}} \exp\left[-\frac{1}{2\beta}|sy - y|^2 - \beta V(y)\right] \leqslant Z_{\mathrm{cl}}(\beta). \tag{B1}$$

Proof. By the relation between arithmetic and geometric means, we always have $\exp\bigl[-\frac{1}{L}\sum_{k=1}^{L}\beta V(x(k))\bigr] \leqslant \frac{1}{L}\sum_{k=1}^{L}\exp[-\beta V(x(k))]$. Applying this and the rules of Gaussian integrals then yields

$$\begin{aligned}\int \mathrm{d}^{nL}x \left(\frac{L}{2\pi\beta}\right)^{Ln/2} &\exp\left[-\frac{1}{\beta}P_L(sx(L),x)-\beta V_L(x)\right] \\ &\leqslant \frac{1}{L}\int \mathrm{d}^n y\, \mathrm{e}^{-\beta V(y)}\left(\frac{1}{2\pi\beta}\right)^{n/2}\exp\left[-\frac{1}{2\beta}|sy-y|^2\right] \\ &+\frac{1}{L}\sum_{k=1}^{L-1}\int \mathrm{d}^n y\,\mathrm{d}^n x\,\mathrm{e}^{-\beta V(y)}\left(\frac{L}{2\pi\beta k}\right)^{n/2}\exp\left[-\frac{L}{2\beta k}|sx-y|^2\right] \\ &\times\left(\frac{L}{2\pi\beta(L-k)}\right)^{n/2}\exp\left[-\frac{L}{2\beta(L-k)}|y-x|^2\right].\end{aligned}$$

However, s is an orthogonal transformation and thus $|sx-y|^2 = |x-s^{\mathrm{T}}y|^2$. Substituting this into the previous formula and computing the resulting Gaussian integral over x will then show that, in fact, all the terms in the sum are equal and the result is precisely the middle term in equation (B1). This proves the first inequality. The second inequality is obvious from the definition of Z_{cl}, equation (1). □

Let us begin the proof of the theorem by defining for all $\mathrm{Re}\, z > 0$

$$g(z) = \mathrm{Tr}(\widehat{P}_{\pm}\mathrm{e}^{-z\widehat{H}}).$$

It has been proven in [29] that, for the potentials we are interested in, the operator $\mathrm{e}^{-z\widehat{H}}$ is of trace class on the right half-plane. Since $\widehat{P}_{\pm}$ is bounded, the trace converges and can be given in terms of the eigenvectors ψ_j and the corresponding eigenvalues E_j of the Hamiltonian as the (denumerable) sum

$$g(z) = \sum_j \langle\psi_j|\widehat{P}_{\pm}\psi_j\rangle \mathrm{e}^{-zE_j}.$$

Suppose next that $\mathrm{Re}\, z \geqslant c > 0$. Since then $|\langle\psi_j|\widehat{P}_{\pm}\psi_j\rangle\mathrm{e}^{-zE_j}| \leqslant \mathrm{e}^{-cE_j}$ and $\sum_j \mathrm{e}^{-cE_j} = \mathrm{Tr}\,\mathrm{e}^{-c\widehat{H}} < \infty$, we can apply proposition 2 and conclude that g is analytic on the right half-plane and that

$$g^{(k)}(z) = \sum_j \langle\psi_j|\widehat{P}_{\pm}\psi_j\rangle(-E_j)^k\mathrm{e}^{-zE_j} = (-1)^k\,\mathrm{Tr}\left(\widehat{P}_{\pm}\widehat{H}^k\mathrm{e}^{-z\widehat{H}}\right). \tag{B2}$$

The operator defining the lattice trace, $\widehat{T}(z)$, is for all $\mathrm{Re}\, z > 0$ a Hilbert–Schmidt operator with a kernel

$$K^{(T)}(a,b) = \frac{1}{(2\pi z)^{n/2}}\exp\left(-\frac{1}{2z}|a-b|^2 - zV(b)\right).$$

Therefore, $\widehat{P}_{\pm}\widehat{T}(z)$ is also Hilbert–Schmidt and it has a kernel

$$\frac{1}{N!}\sum_{s\in S_N}(\pm)^s K^{(T)}(sa,b).$$

Thus for any $L \geqslant 2$ the function $f_L(z) \equiv \mathrm{Tr}\left(\widehat{P}_{\pm}\widehat{T}(z/L)^L\right)$ is well defined and it has an integral representation

$$f_L(z) = \frac{1}{N!}\sum_{s\in S_N}(\pm)^s\int \mathrm{d}^{nL}x\left(\frac{L}{2\pi z}\right)^{Ln/2}\exp\left[-\frac{1}{z}P_L(sx(L),x)-zV_L(x)\right]. \tag{B3}$$

We shall use equation (B3) to also define $f_1(z)$.

Suppose next that $\mathrm{Re}\, z \geqslant c > 0$. Then the integrand in (B3) is bounded by $(L/2\pi c)^{Ln/2} \exp(-cV_L(x))$ which is integrable by assumption. Thus the conditions for proposition 2 are satisfied and each $f_L(z)$ is an analytic function on the right half-plane and all of its derivatives can be computed by a differentiation inside the integral.

We shall now first finish proving the theorem for the $k = 0$ case, from which the $k > 0$ case follows quite easily. The uniform boundedness on compact subsets will follow from the bound

$$|f_L(z)| \leqslant \mathrm{e}^{-b_L \mathrm{Re}\, z} Z_{\mathrm{cl}}(a_L \mathrm{Re}\, z) \qquad \text{where} \quad a_L = \begin{cases} 1 & \text{for } L \text{ even} \\ (L-1)/L & \text{for } L \text{ odd} \end{cases} \tag{B4}$$

and $b_L = (L \bmod 2) V_{\min}/L$ with $V_{\min} = \inf_x V(x)$. Note that for even L we have the simple bound $|f_L(z)| \leqslant Z_{\mathrm{cl}}(\mathrm{Re}\, z)$.

By lemma IV.1 of [29], for any bounded operator $\widehat{A}$ and Hilbert–Schmidt operator $\widehat{T}$, $|\mathrm{Tr}(\widehat{A}\widehat{T}^{2\ell})| \leqslant \|\widehat{A}\| \mathrm{Tr}(\widehat{T}^{\dagger}\widehat{T})^{\ell}$. Since $\|\widehat{P}_{\pm}\| = 1$ and $\|\widehat{T}(z)\| \leqslant \mathrm{e}^{-\mathrm{Re}\, z V_{\min}}$, we thus have

$$|f_L(z)| \leqslant \mathrm{e}^{-b_L \mathrm{Re}\, z} \mathrm{Tr}(\widehat{T}(a_L \mathrm{Re}\, z/\ell)^{\ell}) \qquad \text{where} \quad \ell = \lfloor L/2 \rfloor.$$

Choosing $s = \mathrm{id}$ in the left-hand side of equation (B3) yields an integral representation for the trace in the above equation. By lemma 3, then $\mathrm{Tr}(\widehat{T}(\beta/\ell)^{\ell}) \leqslant Z_{\mathrm{cl}}(\beta)$ and we have finished proving the bound (B4).

We next need the result that for all $\beta > 0$,

$$\lim_{L\to\infty} f_L(\beta) = \mathrm{Tr}\left(\widehat{P}_{\pm} \mathrm{e}^{-\beta \widehat{H}}\right).$$

The proof of this result can be proved essentially identically to the proof of theorem III.4 in [29], which contains the above result for the case $\widehat{P} = \widehat{1}$. We do not reproduce the proof here.

We have now shown that the sequence f_L consists of functions analytic on the right half-plane, which converge on the real axis and which are uniformly bounded on every compact subset of the half-plane. By the Vitali convergence theorem, the sequence then converges on the whole half-plane, the convergence is uniform on compact subsets and the limit function is analytic. Since the limit function is equal to g on the real axis, and g is analytic, it follows that the limit function is given by g on the whole half-plane. Also, since the convergence is uniform on compact subsets, then all derivatives converge as expected, $f_L^{(k)}(z) \to g^{(k)}(z)$. By equation (B2) we have now completed the proof of the first half of theorem 1.

Suppose then that $0 < c \leqslant c'$ and z satisfies $c \leqslant \mathrm{Re}\, z \leqslant c'$. By the Cauchy estimates for derivatives, then for all $0 < t < 1$,

$$|f_L^{(k)}(z)| \leqslant \frac{k!}{t^k c^k} \sup_{(1-t)c \leqslant \mathrm{Re}\, w \leqslant c' + tc} |f_L(w)|$$

and it is thus enough to prove the uniform boundedness of $f_L(z)$. However, since Z_{cl} is now a continuous function, this is an obvious consequence of the inequality (B4). Especially, if V is positive then Z_{cl} is decreasing and we have for all even L, $0 < t < 1$ and $\mathrm{Re}\, z \geqslant c$,

$$|f_L^{(k)}(z)| \leqslant \frac{k!}{t^k c^k} Z_{\mathrm{cl}}((1-t)c).$$

Appendix C. Simplification of the permutation sum

Let r be a permutation and make a change of variables from x to y to the lattice integral (2) as defined by $x(k) = ry(k)$ for all $k = 1, \ldots, L$. The Jacobian of this transformation is one and, by assumption, $V_L(x) = V_L(y)$. The transformation r is orthogonal, and thus $|x(k) - x(k-1)|^2 = |y(k) - y(k-1)|^2$ for all k with the exception of $k = 1$ for which we have $|x(1) - sx(L)|^2 = |y(1) - r^{\mathrm{T}}sry(L)|^2$. Therefore, the lattice kernel satisfies $K_L(sx(L), x; z) = K_L((r^{-1}sr)y(L), y; z)$.

Since r was arbitrary, we can now conclude that the result of the lattice integral depends only on the conjugate class of the permutation. We shall thus need a way of parametrizing the conjugate classes and of choosing an easy representative element from each class. The solution is well known and we present the results as they are given in [30].

Every permutation can be decomposed into independent cyclic permutations, cycles, and its conjugate class is determined by the number ν_l of cycles of length l. Conversely, a collection of numbers ν_l, $l = 1, \ldots, N$, defines a conjugate class provided it satisfies the consistency condition $\sum_{l=1}^{N} l\nu_l = N$. For each class, we can then choose a representative element as the permutation which performs the cyclic permutations for consecutive elements and puts the longest cycles first, e.g. if $N = 7$ and $\nu_3 = 1$, $\nu_2 = 1$, $\nu_1 = 2$, the representative permutation would be $(1, 2, 3, 4, 5, 6, 7) \to (3, 1, 2, 5, 4, 6, 7)$. The number of distinct permutations in each class can be computed from ν_l by the formula

$$\frac{N!}{\prod_{l=1}^{N}(l^{\nu_l}\nu_l!)} = \prod_{l=1}^{N}\left[\binom{N - \sum_{j=1}^{l-1} j\nu_j}{l\nu_l}\frac{(l\nu_l)!}{l^{\nu_l}\nu_l!}\right].$$

Each cycle with an even number of elements has an odd parity and a cycle with an odd number of elements has an even parity. Thus the parity of the permutations in the class is given by the formula

$$(-1)^{\sum_{l=1}^{N}(l-1)\nu_l} = (-1)^{N-\sum_{l=1}^{N}\nu_l}. \tag{C1}$$

For our purposes, there exists a better alternative parametrization. From ν_l define the new parameters λ_l, $l = 1, \ldots, N$, via

$$\lambda_l = \sum_{j=l}^{N} \nu_j.$$

The inverse relation is clearly given by $\nu_l = \lambda_l - \lambda_{l+1}$ for $l = 1, \ldots, N$, where it is understood that $\lambda_l = 0$ whenever $l > N$. The consistency condition for λ_l is now simply

$$\sum_{l=1}^{N} \lambda_l = N \qquad \text{and} \qquad \lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_N \geqslant 0$$

and thus the generation of the conjugate classes reduces to the generation of the partitions of N into N ordered non-negative integers. After this is done, we can compute ν_l and choose the representative permutation, denoted from now on by $s_{(\lambda)}$, as was explained in the previous paragraphs. In addition, the number of permutations in a class is given by

$$c_{(\lambda)} \equiv \frac{N!}{\prod_{l=1}^{N}[l^{\lambda_l - \lambda_{l+1}}(\lambda_l - \lambda_{l+1})!]} \tag{C2}$$

and, by (C1), the parity of the class is given by $(-1)^{N-\lambda_1}$.

References

[1] Redlich K, Cleymans J, Satz H and Suhonen E 1994 *Nucl. Phys.* A **566** 391c
Braun-Munzinger P, Stachel J, Wessels J P and Xu N 1996 *Phys. Lett.* B **365** 1
[2] Wright R H 1982 *The Sense of Smell* (Boca Raton, FL: Chemical Rubber Company)
Turin L 1996 *Chem. Senses* **21** 773
[3] Tsallis C 1988 *J. Stat. Phys.* **52** 479
Tsallis C, Mendes R S and Plastino A R 1998 *Physica* A **261** 534
[4] Lukkarinen J 1999 On the use of non-canonical quantum statistics *Preprint* cond-mat/9911027
[5] Lukkarinen J 1999 *J. Phys. A: Math. Gen.* **32** 287
[6] Chaichian M and Senda I 1993 *Nucl. Phys.* B **396** 737
Chaichian M, Satz H and Senda I 1994 *Phys. Rev.* D **49** 1566
[7] Ceperley D M 1995 *Rev. Mod. Phys.* **67** 279
[8] Reed M and Simon B 1975 *Methods of Modern Mathematical Physics II: Fourier Analysis, Self-Adjointness* (New York: Academic)
[9] Reed M and Simon B 1972 *Methods of Modern Mathematical Physics I: Functional Analysis* (New York: Academic)
[10] Feynman R P and Hibbs A R 1965 *Quantum Mechanics and Path Integrals* (New York: McGraw-Hill)
[11] Trotter H F 1959 *Proc. Am. Math. Soc.* **10** 545
Kato T and Masuda K 1978 *J. Math. Soc. Japan* **30** 169
[12] Neidhardt H and Zagrebnov V A 1990 *Commun. Math. Phys.* **131** 333
[13] Zagrebnov V A 1988 *J. Math. Phys.* **29** 888
[14] Simon B 1979 *Functional Integration and Quantum Physics* (New York: Academic)
[15] Montvay I and Münster G 1994 *Quantum Fields on a Lattice* (Cambridge: Cambridge University Press) section 7.2
[16] Freeman D L and Doll J D 1984 *J. Chem. Phys.* **80** 5709
[17] Hansen E R 1975 *A Table of Series and Products* (Englewood Cliffs, NJ: Prentice-Hall) section 31
[18] Eleftheriou M, Doll J D, Curotto E and Freeman D L 1999 *J. Chem. Phys.* **110** 6657
[19] Doll J D and Freeman D L 1999 *J. Chem. Phys.* **111** 7685
[20] Coalson R D 1986 *J. Chem. Phys.* **85** 926
[21] Runge K J and Chester G V 1988 *Phys. Rev.* B **38** 135
[22] Chakravarty C, Gordillo M C and Ceperley D M 1998 *J. Chem. Phys.* **109** 2123
[23] Chandler D and Wolynes P G 1981 *J. Chem. Phys.* **74** 4078
[24] Doll J D and Myers L E 1979 *J. Chem. Phys.* **71** 2880
Doll J D, Coalson R D and Freeman D L 1985 *Phys. Rev. Lett.* **55** 1
[25] Thirumalai D and Berne B J 1983 *J. Chem. Phys.* **79** 5029
[26] Beck T L, Doll J D and Freeman D L 1989 *J. Chem. Phys.* **90** 3181
[27] Freeman D L and Doll J D 1994 *J. Chem. Phys.* **101** 848
[28] Rudin W 1987 *Real and Complex Analysis* 3rd edn (New York: McGraw-Hill)
[29] Lukkarinen J 1998 *J. Math. Phys.* **39** 3670
[30] Hamermesh M 1962 *Group Theory* (London: Addison-Wesley)