THE SIMULATION OF QUANTUM SYSTEMS

D. M. Ceperley

National Center for Supercomputing Application
Department of Physics
University of Illinois at Urbana-Champaign

I am deeply honored that the committee has chosen me to receive the Feenberg Medal. I was quite taken aback when I was informed of the award, I suppose because I always felt that I am privileged enough to to be involved in the discovery of new scientific knowledge, knowledge that will soon be of significant practical importance. Of course there are many benefits to being a physicist, not the least of which is the opportunity to travel to Schloss Seggau and rendezvous with the many-body community. I would like to thank the Professor Schachinger for his hospitality, and for organizing such a wonderful meeting.

I will take this opportunity to share my thoughts on how far we have gone and how far we may hope to go with the simulation of quantum systems. Of course, much is really unknown, as I hope to convey. The situation is analogous to that of an explorer: the result of his long voyage is a map with a few points on a coastline and a couple of rivers. It would be nice if there were some natives around that could tell us the lay of the land, but barring that, my colleagues and I have the pleasure of doing the exploration ourselves. Of course, the use of statistical methods to “solve” problems in many-body physics is only interesting because there exist fast computers. The field has been and will be driven by the rapid advance of computer technology. Most indications are that this progress will continue at the current rate for at least another decade. In fact, one could argue that the power of computers today has far outpaced the sophistication of simulation methods, so even if the progress in hardware were to stop soon, algorithmic advances, proceeding at a slower rate, will ensure that the art and practice of quantum simulations will thrive for many years. This difference in rate is understandable if one considers the relative number of individuals in theoretical many-body physics versus computer science.

One of the first, and most famous use of computers to solve a previously intractable problem in statistical mechanics was by Metropolis, Rosenbluth and Teller\textsuperscript{1} who simulated a 2d system of hard disks and in so doing introduced the ubiquitous Metropolis Monte Carlo algorithm. This was soon followed by the development of the molecular dynamics techniques by Alder and Wainright\textsuperscript{2} and for more realistic systems by Rahman.\textsuperscript{3} A recent advance was that of Car and Parrinello\textsuperscript{4} who showed how to solve
the electronic structure problem by local density functional methods simultaneous with the molecular dynamics evolution of the classical degrees of freedom, thus eliminating the need for parameterized potentials. There were some early suggestions that the simulation methods could be used to solve the quantum many-body problem (including unpublished remarks by Fermi) but it was Kalos who really showed how it could be done. The paper on the ground state of hard core bosons with Levesque and Verlet\(^5\) was the first major simulation of a quantum many-body system.

**THE QUANTUM PROGRAM**

My goal is to learn how to compute the properties of a microscopic quantum many-body system on a modest computer in a fully routine way. One should just be able to set up the numbers of particles, their interactions, perhaps the temperature and then directly determine the thermodynamic properties, the correlation functions and so forth. By routine, I mean that the calculation should not have to require a Ph.D. in physics to set up the input or to understand the results.

Because I want this simulation to be a black box, it is important that the calculations come with real error bars that are determined internally. I doubt that we will soon be able to calculate exact properties of a quantum many-boson system to arbitrary accuracy like one can with simple mathematical functions. After all, one cannot even do exact calculations in classical statistical mechanics and there is no reason that quantum mechanics should be any easier (at least at non-zero temperature). It would be a much easier task to make a program that guesses the answers (guessing in the sense that there is no way of estimating the error of an individual calculation), maybe based on an intelligent use of prior experience or experimental measurements, but I find that the design of such a semi-empirical program a less interesting and not a mathematically well defined task. It is particularly difficult to make a semi-empirical program which is universal, *i.e.* can accurately treat a wide variety of physical systems from liquid helium, to superconductors, to the electronic states of molecules. So one of the requirements is that the program be “exact” within its own computed errors.

There has been a great deal of progress in the last few decades in learning how to simulate systems at the quantum single body level, mainly with local density functional theory. For electronic structure calculations, one assumes that the nuclei are classical and the electrons are weakly correlated. There are codes that are fast approaching the goals that I have set, where commercial software is available on lots of computers whose use does not require an expert that knows all the details of the program. One of the recent issues is to find Local Density Function (LDA) algorithms which are order(\(N\)) in the number of electrons, \(N\). This means the computer time per step scales linearly with the number of electrons for sufficiently large systems. Thus systems with thousands of electrons will be amenable to those methods on today’s most powerful computers. But, although the mean-field descriptions may be quite satisfactory for ground state geometries, for more complex properties, and for highly correlated systems, there are some real problems with those methods and no one knows if one can practically extend the local density approximation.

There is another important characteristic of the program which I would like to emphasize. It should have the ability to simulate experimentally relevant systems. By that I mean that we do not want to simulate just simplified models. This more or less rules out methods restricted to lattice models since it requires a lot of expertise to know if a lattice model is applicable to a particular physical system. One must be able to treat thermal effects, systems with many different types of particles and large units cells and
to calculate relevant properties such as transport coefficients. Of course, a key problem in condensed matter physics is the calculation of the electronic properties of materials. It is this aspiration which requires that we be able to calculate properties of systems with thousands of atoms and tens of thousands of electrons. This places a stringent condition on the type of algorithm which can be used. Experimental relevance also places a demand on the accuracy of the program. Clearly for many applications, one will need to compute energies to an accuracy of much better than room temperature, for example better than 0.01eV.

Since computers will continue to get more powerful, it is clear that the complexity issue i.e. how the algorithm performs as the system size gets larger, is crucial. Unless we have methods that scale with some low power of the number of electrons, the ability to simulate “real world” systems will always be limited. As I mentioned above, the focus of most research today in the condensed matter community is on approximate methods, such as local density functional theory. The hot issue is whether proposed order(N) methods really work and how to implement them in an efficient manner. So those methods will be on the workstation of the typical condensed matter physicist in 5 to 10 years. One will be able to simulate thousands of electrons, but with some unknown, systematic error coming from the local density approximation. The question that I want to address in the remainder of the talk is to what extent one could develop exact methods for general use.

I suppose it is unnecessary to convince this audience that the numerical solution of the quantum many-body problem is a worthwhile goal but often one hears the objection that just calculating numbers does not really lead to any understanding. Let me list some of the reasons why numerical methods are an indispensable part of science. It seems that science progresses very slowly when there are no high quality experimental or numerical data to tell if a proposed theory or approximation is any good. This has partly to do with the sociology of science, one needs to arbitrate between various theories. Computation can sometimes calculate a quantity which is simply not possible to measure. For example, one can find out what happens to liquid helium if the particles obey Boltzmann rather than Bose or Fermi statistics. I think if a quantum simulation tool were around, statistical physics would move to a higher level, the issues would become more sophisticated and fundamental. Many of the theoretical approaches that are worked on today would become obsolete, but I suspect just as many new ones would open up. Finally, can we really say we understand something if we can’t compute the consequences? Another justification is that physics needs to be able to provide reliable predictions of quantum systems to be useful to the practical scientists. An engineer may only be interested in designing a useful gadget, not in the details of the many-body physics. This is not really theory, but is the side of science that pays the bills and provides one of the principle long-term justifications.

Leaving aside the above considerations, here I will focus only on what has always been my motivation, the intellectual, the mathematical and the technical challenge of the quantum many-body problem. To what extent is it possible to exactly calculate the microscopic properties of quantum many-body systems? I think that this is one of the major challenges and opportunities in all of theoretical physics. As I will describe, there is plenty left to be done here, only the coastline of the continent has been sighted. I believe that in the process of discovering how to numerically solve the quantum many-body problem, it is inevitable that our understanding of quantum many-body physics will become much deeper.
THE COMPLEXITY OF QUANTUM MECHANICS

Let me now formulate the quantum many body problem as a problem of complexity. We ask how much computer time, $T_O(\epsilon)$, it takes to calculate some property, $O$, of a particular quantum system to a specified accuracy, $\epsilon$. The error, $\epsilon$, is the total error, including both systematic and statistical errors. The actual computer time is not very well defined since it depends on which computer you have, how clever the programmer is and so forth. The dependence of the computer time on the accuracy needed and on the number of particles, $N$, the complexity, is a more general question and goes to the heart of the problem. There are general methods that can numerically solve the 1-body, 2-body, 3-body or even 4-body quantum problem. Some of those methods have been around since the dawn of quantum mechanics. But one sees that those methods are not being used to solve the 5-body or 6-body problem, even with the enormous growth in computer power. For example, J. Reading, elsewhere in this volume, gives an example of ion-atom scattering and how difficult it is to calculate the cross section for a single proton scattering from a lithium atom.

I define an explicit quantum method as one where the full many-body wave function is tabulated in some basis. The simplest basis is a Cartesian grid with $d$ grid points in each dimension. Since the many-body wave function is an arbitrary function of $3N$ variables, the total number of grid points is $d^{3N}$. Hence the size of computer memory and computer time will be proportional to $d^{3N}$. Now if we use the trapezoidal rule to do the needed integrals we find that the error goes as $\epsilon \propto d^{-2}$. Then the computer time needed to do an integral goes as: $T(\epsilon, N) \propto \epsilon^{-3N/2}$. Explicit methods are exponential in the number of particles and thus rapidly become intractable.

Of course, a lot of effort has gone into finding better ways to represent the wave function and to do the multidimensional integrals. For example, we can use Simpson’s rule to do integrals, if the integrand is smooth. What one finds is that the computer time now goes as:

$$T(\epsilon, N) \propto \epsilon^{-3N/l},$$

where $l$ is a smoothness parameter. If the trapezoidal rule is used, $l = 2$. For Simpson’s rule, $l = 4$, and so forth. But $l$ can never be made too big since that requires more and more smoothness in the wave function. So the use of higher order schemes will only reduce the coefficient of $N$ in the exponent by a limited amount. Other basis functions, the common example in quantum chemistry is the Linear Combination of Atomic Orbitals (LCAO), will end up with a similar scaling. While they can rather efficiently solve the one-electron equations, expansions of the many body wave functions necessary to correlate the electrons, the Configuration Interaction method, leads to an exponentially large number of determinants. So the coefficient of $N$ in the exponent depends on the problem, the basis set and the algorithm, but if the basis is complete, one ends up with an exponential increase in time.

If we look at the progress in a field, one can quickly see whether the algorithms are stuck with the problem of exponential complexity. The floating point performance of the most powerful computers has been growing at an exponential rate over the last 50 years, with the speed doubling roughly every 2 years. Recently the rate has been even faster. If a quantum algorithm is explicit, then the system size amenable to calculation will grow only linearly with time. For example, exact diagonalization has proved an effective tool to understand lattice spin models. The spin 1/2 Heisenberg model is very economical model since each spin can only be in one of two states. Lattice symmetry
further reduces the size of the matrices which must be diagonalized. In 1978, Oitmaa and Betts\textsuperscript{6} diagonalized a system with 18 spins. A few years ago in 1990, lattices with 36 spins were diagonalized by several groups. Hence it appears that roughly one spin can be added each year. Of course one should not be so pessimistic. We heard in this conference about the work of Roder and Silver who have a new method to calculate thermodynamic properties of spin systems, which although it uses an explicit representation, also has a random element and can calculate densities of states. Lanczos methods are usually limited to low lying states. Exact diagonalization may well be able to understand much of the physics of simple lattice models, especially if the algorithms are further developed, but those lattice models are very far from real materials.

In the rest of the talk I will only discuss simulation methods since they do not have exponential scaling. By simulation, I mean methods such as Monte Carlo or molecular dynamics, where phase space is sampled rather than explicitly integrated over. One can justify this type of approach by saying that since quantum mechanics is probabilistic, one might as well sample the square of the wave function as explicitly integrate over it. Progress, over the years in the simulation of classical systems, has been dramatic. Rahman\textsuperscript{3} considered on the order of 800 argon atoms in 1964. By 1984, there were molecular dynamics simulations of systems containing 161,000 atoms interacting with realistic forces.\textsuperscript{7}

In any type of simulation method, the error decreases inversely as the square root of the number of independently sampled points, while the computer time for each sampled point is linear in the size of the system, assuming that the particles interact with short-ranged forces. This "slow" convergence of the error is the curse of simulation methods but for many dimensional problems it is vastly superior to that of explicit methods. (There have been mathematical demonstrations\textsuperscript{8} to the effect that there are better ways of placing the points than at random, the so-called "quasi-random" lattice points, which will give a convergence like $T^{-1}$ instead of $T^{-1/2}$. In practice, this method has not achieved a change in the exponent.) Putting these rates together, we find that the computer time goes as:

$$T(\epsilon, N) \propto N^{\alpha} \epsilon^{-2}. \quad (2)$$

The exponent, $\alpha$, depends on the quantity being computed. For a local property such as the energy per atom, and for a local potential energy function then $\alpha = 0$. This is the best possible situation, we will call it "classical scaling." Even for long-ranged Coulomb interactions, there exists an order($N$) algorithm, the "fast multipole algorithm." Most classical systems really have $\alpha \approx 1$ if you take into account how the autocorrelation time depends on the size of the system. Even if $\alpha$ is greater than one, for example near a phase transition, or in a frustrated system, the scaling of simulation methods is enormously better than that of explicit methods, since the doubling time for the systems size is only $\alpha$ times the doubling time for computational resources. This is one reason why simulations methods continue to have an enormous impact in the field of classical statistical mechanics. While I do not know if there is a proof that simulation methods are needed for classical systems, I do not know of any competitors either.

Since quantum mechanics reduces to classical mechanics in certain limits, and since simulation is the only known general exact method in classical mechanics, we expect that quantum mechanics will require simulation. (This argument is actually only valid at finite temperature.) In addition to being economical for large $N$, another nice feature of simulation is its simplicity. This simplicity will be crucial in building a robust "black-box" capable of treating a wide variety of quantum systems. So now the question arises whether it is possible to do quantum mechanical calculations with a simulation method.

Feynman\textsuperscript{9} has already answered this question, arguing that it is impossible to sim-
ulate quantum mechanics on a local classical computer. (By the way Feynman made the suggestion that perhaps we should try to build a quantum mechanical Turing machine, a programmable device whose state evolves according to the laws of quantum mechanics, not classical Boolean logic.) Of course, you should really read Feynman’s article, but let me try to give a very brief synopsis. First he makes the complexity argument. For a large system, simply specifying the initial conditions is too difficult. It will take an inordinate amount of memory not to speak of the operations needed for time evolution. But theoretically, any initial conditions are possible, therefore explicit methods are out. At best, one will be able to do a probabilistic simulation. That is not too bad, since quantum measurements have a random character anyway. Let us now assume that the simulation algorithm is “local.” Carve up the physical system into regions of space and assign the computation of the particles in each spatial region to a given processor. The processor for each spatial region can only communicate with neighboring regions and evolves in time according to its current state and the state of its neighbors. This arrangement is equivalent to a local hidden variables theory and from Bell’s theorem,\(^\text{10}\) the computer can never completely reproduce the results of quantum mechanics. Hence Feynman concluded, simulations of quantum systems on local classical computers are impossible.

I want now to describe what we can do. I will come back to ways around Feynman’s argument at the end.

THE PATH INTEGRAL METHOD

In 1953, Feynman\(^\text{11}\) applied his newly invented path integrals to superfluid liquid \(^4\text{He}.\) Feynman did his path integral calculations analytically, so he had to make some approximations. However, once quantum mechanics has been formulated in terms of imaginary time path integrals, it is amenable to direct simulation. Luckily, all thermodynamic properties can be calculated in imaginary time since they are traces of the thermal density matrix:

\[
\rho(R, R'; \beta) = \sum_\alpha e^{-\beta E_\alpha} \phi_\alpha^*(R) \phi_\alpha(R')
\]

(3)

where \(R = \{r_1, r_2 \ldots r_N\}\) is a point in \(3N\) dimensional configuration space, \(\beta = 1/k_B T\) is the inverse temperature and \(\mathcal{H} = V + T\) is a non-relativistic Hamiltonian, with eigenvalues, \(E_\alpha\), and eigenfunctions, \(\phi_\alpha(R)\). I am neglecting relativistic effects everywhere. The non-relativistic problem is already hard enough.

The discrete PI expression then follows:

\[
\rho(R, R_0; \beta) = \frac{1}{N!} \sum_P (\pm)^P \int dR_1 \ldots dR_{M-1} \exp \left[ - \sum_{k=1}^M S(R_k, R_{k-1}; \beta/M) \right]
\]

(4)

where the path, \(\{R_1, R_2 \ldots R_{M-1}\}\), is to be integrated over. We have gotten rid of quantum mechanics by introducing many more classical degrees of freedom. The “action”, \(S(R, R'; \tau)\) in the “primitive approximation” is given by:

\[
S(R, R'; \tau) = \frac{(R - R')^2}{4\alpha \tau} + \frac{\tau}{2} [V(R) + V(R')] + \frac{3N}{2} \ln(4\pi \lambda \tau)
\]

(5)

where \(\lambda = \hbar^2/2m\). In practice, much better approximations are used for the action. The number of points on the path, \(M\), is to be taken to infinity although one can have
Figure 1. A typical path integral contribution for a system of 6 \(^4\)He atoms at a
temperature of 0.75K. The dark circles are the positions of the atoms at the beginning of
imaginary time. The number density is 0.0432 Å\(^{-2}\). The dashed square represents the
periodic boundary conditions. Three of the atoms are involved in an exchange which winds
around the boundary in the x direction thus showing that the system is superfluid.

good and controlled approximations for small \(M \approx 20\). The approximation of finite \(M\)
will not affect the discussion of complexity since systematic errors converge sufficiently
quickly as \(M\) becomes large.

The boundary condition at the end of the path is: \(PR_M = R\). Here \(P\) is a permuta-
tion operator. The upper, \((+)^P\) sign is for bose statistics, the lower, \((-)^P\) sign is for fermi
statistics. Bose statistics enter in a very physical way: to calculate any observable, one
must sum over all possible relabelings of the particles between the beginning and the
end of the path. It is this relabeling, or macroscopic exchange which is responsible for
the phenomena of superfluidity in a rather direct way. In a simulation the permutation,
\(P\), is sampled.

What Feynman has achieved is to map the quantum statistical mechanics of any
Bose system onto classical statistical mechanics, the equilibrium properties of a ficti-
tious system of \(N\) ring “polymers.” The important feature is that the integrand of the
path integral expression is non-negative. Figure 1 shows a typical path of 6 \(^4\)He atoms
from a 2d simulation. The paths are closed because we are using the simulation to
calculate the trace of the density matrix.

We have developed the numerical methods\(^{12,13}\) for sampling paths and permuta-
tions by generalizing the Metropolis rejection algorithm and shown that it is practical
to obtain good agreement with the experimentally measured properties of liquid \(^4\)He
above, below and at the lambda transition. In addition to such static quantities as the
energy, specific heat and pair correlation function, we found a few new ways path inte-
grals can help understand superfluidity. We showed\(^{14}\) the superfluid density in periodic
boundary conditions is exactly proportional to the mean squared winding number of
the paths. (A path with a nonzero winding number makes a non-trivial loop around
the periodic boundary conditions. An example is shown in the figure.) Hence there is
“dynamical” content to the imaginary time paths. The condensate fraction, the frac-
tion of \(^4\)He atoms in the zero momentum state, which is the order parameter of the
superfluid state, is directly related to the delocalization of two ends of a cut path, one
Table 1. Dictionary of the Quantum-Classical Isomorphism

<table>
<thead>
<tr>
<th>Quantum</th>
<th>Classical Analogue</th>
</tr>
</thead>
<tbody>
<tr>
<td>bose condensation</td>
<td>delocalization of cut ends</td>
</tr>
<tr>
<td>boson statistics</td>
<td>joining of polymers</td>
</tr>
<tr>
<td>exchange frequency</td>
<td>free energy to link polymers</td>
</tr>
<tr>
<td>free energy</td>
<td>free energy</td>
</tr>
<tr>
<td>imaginary velocity</td>
<td>bond vector</td>
</tr>
<tr>
<td>kinetic energy</td>
<td>negative spring energy</td>
</tr>
<tr>
<td>momentum distribution</td>
<td>FT of end-end distribution</td>
</tr>
<tr>
<td>particle</td>
<td>ring polymer</td>
</tr>
<tr>
<td>potential energy</td>
<td>iso-‘time’ potential</td>
</tr>
<tr>
<td>superfluid density</td>
<td>mean squared winding number</td>
</tr>
<tr>
<td>superfluid state</td>
<td>macroscopic polymer</td>
</tr>
<tr>
<td>temperature</td>
<td>polymer length</td>
</tr>
</tbody>
</table>

that is not periodic in imaginary time. There is a translation of terms and concepts as we go from the quantum system into the classical system. Some examples are shown in Table I. The numerical calculations have borne out Feynman’s hypothesis: the essential features of superfluidity, the phase transition, the absence of viscosity and bose condensation, are all a result of macroscopic quantum exchange.

In this short article, I have only briefly introduced the path integral Monte Carlo method of simulating quantum systems. There is another technique for computing properties of the bosonic ground state introduced by Kalos\(^5\) known as Green’s Function Monte Carlo. Conceptually it is very close to PIMC, the main differences being that the paths do not close on themselves, instead computer time is along the imaginary time axis, and a trial function is introduced to guide the resulting random walk. Diffusion Monte Carlo\(^15\) is a variant of this technique. Path Integral Monte Carlo has a number of advantages over GFMC in the construction of a “black box” program, namely one can take into account thermal effects, there is no importance function to worry about, and that all expectation values can be computed without systematic errors, not just the energy. However, the methods are closely related and it is for reasons of brevity that I have discussed only PIMC.

Now, let us return to the question of how much computer time it takes to do a simulation of superfluid helium. Undoubtedly, quantum simulations are slower than those of the same-sized classical system. In fact, the first simulations of Kalos in 1970 at zero temperature, and the ones that I have been involved with at finite temperature are comparable in physical size to classical simulations of the 1960’s. Nonetheless, the complexity is the same for superfluid \(^4\)He and a classical liquid. After one has expanded an \(N\) atom system into an imaginary time path with \(M\) steps, one has a classical system of \(NM\) particles. One can calculate local quantities like the energy in a computer time independent of the number of particles and of the temperature because the action is local in space and in imaginary time. It is only the prefactor and the lack of efficient implementations which has limited simulations of large bosonic systems. The crucial property of the integrand that makes the simulations do-able and even straightforward is that the integrand is strictly non-negative. We will show below why this is crucial.

This is the end of the good news. Let me touch on an unsolved but very important problem for bosonic superfluids on which we have been recently working. A break-
Figure 2. Shown is the density-density response as a function of imaginary time for $k = 0.76 \text{Å}^{-1}$ (solid line), 2.01 $\text{Å}^{-1}$ (short dashes) and 2.52 $\text{Å}^{-1}$ (long dashes) in liquid $^4\text{He}$ at a temperature of 1.2K as computed with PIMC.

through in understanding liquid $^4\text{He}$ came when Landau guessed that the excitation spectrum was unusual. He postulated the familiar picture of phonons and rotons. Experimentally, one can measure the density-density correlation function, $S_k(\omega)$, with neutron scattering and directly see this dispersion relation. This is a quantity which we would like to be able to calculate and interpret since the excitations are central to understanding superfluid helium. Feynman\textsuperscript{16} introduced a variational way of calculating the phonon-roton spectrum. See the article in this volume by Vitiello and Reatto concerning recent improvements. The variational wave function approach is the accepted microscopic way of calculating the phonon-roton spectrum but it leaves a number of questions unanswered and does not satisfy our criteria since the errors are not controlled. Variational approaches are guaranteed to find a single excitation energy, not the entire excitation spectrum, and not the lifetime of the excitation. It is not clear how superfluidity comes into play in the variational approach, and extending the approach to non-zero temperatures brings in more approximations. There are theoretical and experimental conjectures\textsuperscript{19} that the excitation spectrum is more complicated because of the relationship between the spectrum and the superfluid order parameter.

The imaginary-time density-density response function defined as:

$$F_k(t) = \frac{1}{Z} Tr \left( \rho_{-k} e^{-\tau \hat{H}} \rho_k e^{-(\beta-t) \hat{H}} \right)$$

is straightforward to calculate with path integrals. Here $\rho_k = \sum_i e^{i k \rho_i}$ is the fourier transform of the density. This function is shown in Fig. 2 for various wavelengths. It is a function which decays exponentially up to $\beta/2$ and then rises again because of the periodicity in imaginary time (that part is not shown) and is related to $S_k(\omega)$ by a Laplace transform:

$$F_k(t) = \int_{-\infty}^{\infty} d\omega e^{-\tau \omega} S_k(\omega)$$

One can see little in the way of structure in $F_k(t)$, it is a featureless exponential decay. But mathematically, $F_k(t)$ and $S_k(\omega)$ are equivalent and it is possible to do inverse Laplace transforms analytically. But the presence of any statistical or systematic errors, rules out a direct inversion. The numerical inversion of a Laplace transform is a classic ill-conditioned problem.
Figure 3. A comparison of the phonon-roton dispersion energies of liquid $^4\text{He}$ calculated with PIMC plus MaxEnt inversion (filled circles) to those determined by neutron scattering and thermodynamical measurements (solid curve).

It has been proposed$^{17}$ that Bayesian, or maximum entropy techniques could aid in this inversion by making the problem better conditioned. In addition to using the Monte Carlo data for the inverse Laplace transform, one uses prior information about what are acceptable functions, $S_k(\omega)$. In some lattice models, the maximum entropy method has given very satisfactory inversions. We have done such a maximum entropy reconstruction using the imaginary time data from a moderately short simulation (1 month on a SPARC workstation). The results are compared to the very clean experimental measurements by neutron scattering$^{18}$ in Fig. 4. You can see that we are far from a satisfactory calculation of this quantity with path integrals. Although the method reproduces the overall shape of the response function, it does not resolve the two peak structure and does not tell us how narrow the phonon-roton line is. One can get better results by assuming a better default model (here we used a very broad Lorentzian model) but that begs the question. Ideally, one would like a method which uses only rigorous theoretical input and nothing derived from experiment or approximate theories. If we take the position of the peak height as an estimate of the excitation energy, one gets the dispersion curve shown in Fig. 3. The agreement on the excitation energy with experiment is better than 10%, but there is no reliable way of assigning error bars to the inversion procedure. I am unaware of any work on the complexity, $T(\epsilon, N)$, when Bayesian methods are used to improve the inversion. There are intriguing hints of other methods for calculating this spectrum in a different way, using the information about macroscopic exchanges present in the simulations, but I will not go into the details here.

Another related problem is the calculation of the structure and energy of a single vortex in liquid helium. This is another characteristic excitation of superfluid helium and is traditionally calculated with variational techniques. Elsewhere, Vitiello and Reatto discuss improved calculations with shadow wave functions. One would like to see vortices come naturally from path integrals and to be able to calculate their properties in an approximation-free way, as we can for quantities like the momentum distribution or the superfluid density.
FERMIONS AND THE SIGN PROBLEM

Unfortunately, bosons are relatively rare. Most of the world is made of fermions. If quantum simulations are to a big scientific impact, we have to simulate electronic systems. Probably most of you have heard about the “fermion sign problem.” Let me briefly explain the problem. It arises from a simple minded extension of the Feynman path integral method from Bose statistics to Fermi statistics. One has to take the minus sign in Eq. (4) whenever the relabeling involves an odd permutation of fermions. So now the integrand for any observable is partly positive and partly negative and it cannot be interpreted as a probability distribution. Before, we simply sampled a path with probability proportional to the integrand. For an integrand which is which is both positive and negative, the best function to sample is the absolute value of the integrand. That means one samples the paths from the *bosonic* distribution! Then the fermion minus sign is taken into account as a weight. Any physical observable for fermions is computed as the difference between the observable averaged over the positive paths, and averaged over the negative paths. One quickly gets into difficulty as either the temperature gets low or the number of particles gets large. One finds that the computer time needed to compute some quantity to an accuracy, $\epsilon$, is:

$$T(\epsilon, N) \propto \frac{e^{2\beta \Delta F}}{e^2}$$

where $\Delta F$ is the difference between the free energy of the fermion system and the free energy of the bose system which was actually sampled. If fermi statistics are important, the difference will be large and, in any case, proportional to the number of particles, $N$. Hence the computer time depends exponentially on both the number of particles and the inverse temperature, just as for the explicit quantum methods. That the computer time blows up is not too surprising, since it is totally unphysical to sample with bose statistics. After all, bosons will undergo a superfluid transition but fermions will not. The minus signs have to take account of this huge difference.
As trivial as the problem may seem, no “exact” method has been found that does not scale exponentially. Remember, it is not really the sign that is the issue, but how it scales in \( N \). There is a simple thermodynamic argument which shows that once minus signs are allowed into any calculation, one eventually has an exponential growth in computer time for large enough \( N \). Suppose we have a system which has a probability \( p \) of having a positive contribution and a probability \( q = 1 - p \) of having a negative contribution. Now the efficiency, or signal-to-noise ratio, is simply the integral divided by the total number of samplings: \( (p - q)/(p + q) \). Now put \( N \) of these systems together. Using the binomial theorem, the signal-to-noise ratio is now: \( (p - q)^N/(p + q)^N = e^{-cN} \) where \( c = -\ln(1 - 2q) \approx 2q \). Thus no matter how small \( q \) may be, one gets exponentially scaling. The same argument applies if one goes a factor \( N \) times lower in temperature. The general result is that once minus signs are allowed into the integrand, the computer time becomes proportional to \( e^{\beta N} \Delta \varepsilon \). Monte Carlo is great at addition but cannot subtract very well. That is the minus sign problem.

RESTRICTED PATH INTEGRALS

What then can be done to simulate fermion systems? What I am going to discuss now, the fixed-node or restricted path integral method, is not an exact solution, but one where an uncontrolled approximation is made. But we learn that a mapping from fermion systems to a “classical” distribution (one with a non-negative integrand) exists, but unfortunately to find the mapping, requires that we determine the sign of the many-fermion density matrix before we begin. However, much progress can be made with approximate mappings. Practically, the method appears to be more accurate than any other method, and it scales reasonably well. Even if we cannot predict properties of real systems with 100% reliability, maybe we will be able to determine properties of generic fermions.

The restricted Path Integral method is based on the following exact identity\(^{20}\) for the fermion density matrix:

\[
\rho(R, R_0; \beta) = \frac{1}{N!} \sum_{P} (-1)^P \int dR_1 \ldots dR_{M-1} \exp \left[ -\sum_{k=1}^{M} \left( S(R_k, R_{k-1}; \frac{\beta}{M}) + S_N(R_k, R_0; \frac{k\beta}{M}) \right) \right] (9)
\]

As before \( PR_M = R \) and \( S \) is the bosonic action as before. All we have done is to include an additional “nodal” action, \( S_N(R, R_0; t) \), given in the limit \( M \rightarrow \infty \) by:

\[
S_N(R, R_0; t) = \begin{cases} 
0 & \text{if } \rho(R, R_0; t) > 0 \\
\infty & \text{if } \rho(R, R_0; t) \leq 0.
\end{cases}
\] (10)

The nodal action restricts the paths to regions where the fermion density matrix is strictly positive. We have accomplished our main goal, namely that for calculations involving the trace of the fermion density matrix, where \( R = R_0 \), all negative paths with odd permutations are killed by the nodal action so the integrand is non-negative.

But the unknown fermion density matrix appears on both the left and right hand side of the identity since one needs it for the definition of \( S_N \). But we only need the sign of the fermion density matrix for the nodal action, or equivalently its nodes, to have an exact “classical” path integral expression. The nodes are a \( 3N - 1 \) dimensional hypersurface which bisect the full configurations space into symmetrical pieces.\(^{20}\) Generally one doesn’t know the nodal locations except for certain limits. The exception
is in one dimension, where the nodes correspond to two fermions being at the same place. Knowledge of the nodes implies a solution to the complexity problem if one can determine the sign of the density matrix quickly enough.

To do practical calculations in the absence of exact nodal information, one makes a good ansatz for the sign of the density matrix needed on the right hand side. Then the simulation of the fermion system proceeds exactly as with the boson system except the sum over permutations is only over even permutations and the paths are restricted to those with a positive trial density matrix with respect to $R_0$. One might worry about how one can throw away so much of the integrand and still be left with an exact result. The answer is that the flux of positive paths at any spot on the nodal surface exactly cancels the flux of negative paths because the gradient of the density matrix is continuous across the node. Hence there is no contribution to the integral of paths crossing the nodes.

A number of recent calculations show that even with free-particle nodes, the simplest nodal approximation, the method gives accurate properties for strongly correlated fermions.\textsuperscript{21} In a recent study\textsuperscript{22} we were able to see a plasma phase transition in hot liquid hydrogen, starting out with the “elementary particles”, 32 electron and 32 protons. In the case of liquid $^3\text{He}$ and the 2d electron gas, it is known how to construct quite accurate nodal surfaces\textsuperscript{23} by using another idea that Feynman introduced, that of backflow. The calculations are somewhat more involved and have not yet been implemented at finite temperature. Even though the restricted path integral method is approximate, I am optimistic that it will lead to an accurate way of simulating strongly correlated fermion systems and possibly a different way of understanding fermi liquids and superconductors. One of the beautiful aspects of path integrals is that it can unify different aspects of many-body theory, so bose and fermi systems can be understood and calculated within the same method.

Let us now address the complexity issue for restricted path integrals. The difference with the bosonic calculation is that at each step one has to check that the trial density matrix is positive. Let us take as an example non-interacting fermions. The free particle density matrix for $N$ fermions is a determinant of the $N \times N$ matrix of single particle density matrices:

$$\rho(R,R_0;\tau) \propto \text{det} \left[ e^{\frac{-|r-r_0|^2}{4\lambda^2}} \right].$$

(11)

In the worst case, it takes $N^3$ operations to evaluate the determinant after all the particles have been moved. If a single particle is moved it takes only $N$ operation to find the new determinant in terms of the old inverse matrix and $N^2$ operations to update the inverse matrix if the move is accepted. This time can be greatly reduced since the single particle density matrix is localized. They decay to zero in a distance equal to the thermal de Broglie wavelength: $\lambda_D = \sqrt{\frac{6\lambda}{K_B T}}$ where $T$ is the temperature. Then the number of non-zero elements in any row of the matrix will be on the order of $\lambda_D^3 \rho$ where $\rho$ is the number density. As long as the temperature is not so low that $\lambda_D$ extends over the entire simulation cell, the time it takes to find the new determinant will be independent of $N$. Essentially it goes as: $\lambda_D^6 \propto T^{-3}$. Then one finds that the complexity is:

$$T(\epsilon, N) \propto \begin{cases} 
\epsilon^{-2}T^{-3} & \text{if } \lambda_D < L \\
\epsilon^{-2}N^2 & \text{if } \lambda_D > L
\end{cases}.$$  

(12)

For localized electrons, say in an insulator, the single particle density matrices will have a finite range even at zero temperature, they become localized Wannier functions.

So except for the case where one is essentially in the ground state, restricted paths have achieved classical scaling. At very low temperatures, one has an $O(N^2)$ method.
The complexity of the restricted Path Integral algorithm is rather similar to that found in single particle calculations. If the electrons are localized, one gets classical complexity. Things are worse if the electrons are delocalized as in a metal and one is interested in low temperature details. Many of the order($N$) methods being developed for LDA-DF calculations can be applied directly to the restricted path integral algorithm. More realistic backflow nodes appear to introduce an extra power of $N$ in the scaling.

Since the fermion minus-sign problem is unsolved there are many open questions. Can better density matrix nodes be found? Can these be effectively determined with variational principles or self-consistently? What features of the nodal surfaces are important to determine accurate properties? For example, which features correspond to superconductivity? As more and more accurate nodes are developed, does that bring back the exponential scaling? How do we compute properties other than those involving the trace of the density matrix? Off-diagonal matrix elements bring back some of the minus signs, so that a direct calculation of the superfluid density as was done for bosons, appears to be much more difficult for superconductors.

THE LIMITS OF QUANTUM SIMULATIONS

Now let us return to Feynman’s argument against quantum simulations on local classical computers and try to understand how we have managed to do exact simulations of boson systems. There are at least two important ways that we have violated Feynman’s assumptions. First, we have formulated the problem in terms of complexity, not locality. It is possible to make a simulation with complexity $N \ln(N)$ which is non-local. In fact, it is done all the time. In band structure calculations, plane waves are often used as basis states. In order to calculate the electrostatic potential energy, one has to go into a position representation with a Fast-Fourier Transform (FFT) which takes on the order of $N \ln(N)$ operations. Switching back and forth between k-space and r-space is the epitome of a non-local algorithm. It may well be that one needs to do part of a quantum simulation in a non-local representation and part in a local representation. The bosonic imaginary time path integrals are local in space and in imaginary time. Does this count as being local? Presumably Bell’s theorem inequalities are important at very low temperatures since thermal fluctuations will wipe out phase coherence. In that limit, the imaginary time path will extend over large spatial regions and the path of an entire particle will be non-local. In addition, for fermion restricted path integrals, the nodal action is non-local in space and imaginary time.

This brings up a more significant difference. The path integrals we have been discussing are all in imaginary time. As far as I know, Bell’s theorem is applicable to real measurements, i.e. where systems evolve in real time from the past into the future. In the path integral technique, what is in the computer’s memory is the entire imaginary time trajectory. There is no causality since we simultaneously treat all times. In fact, imaginary time is periodic. Hence the implication of Feynman’s argument seems to be that we really cannot simulate quantum dynamics on a local classical computer. In fact, simulation of quantum dynamics is known as being computationally much more difficult than imaginary time dynamics. Actual calculations involve only a few particles and very short propagations.

For boson and boltzmannon statistics, the computation of thermodynamic properties seems well understood and manageable. We can do routine calculations of the thermodynamic type, traces of the density matrix, or things which are expressible in terms of this. It is this last loophole which makes me somewhat optimistic. Not only
can we get things like energy, pressure, compressibility, but it turns out there many of
other quantities which can be expressed as integrals over the thermal density matrix.
One example is the exchange frequencies of atoms in a quantum crystal.\textsuperscript{24} I don’t think
it is really known how far we can push this. How many experimental measurements can
be related to imaginary time matrix elements? What is required is more than writing an
expression for a given quantity in terms of the density matrix. One really has to show
how the computer time scales with the number of particles and to demonstrate that
the calculation is practical. Tractable algorithms to treat fermions or excited states
of bosons are still to be invented. Simulation of true quantum molecular dynamics
(\textit{i.e.} scattering junk on junk) seems very far away. It is likely that the full range
of microscopic many-body theory will be needed to perfect the simulation algorithms,
from good mean-field techniques, to renormalization group methods to perturbation
theory. Much exploration remains to be done.

In closing, I want to acknowledge my colleagues that have contributed so much to
my work over the years, in particular to Geoffrey Chester, Mal Kalos, Berni Alder,
Roy Pollock, Bernard Bernu and Richard Martin. Of course I owe a special debt to
my wife, Peri, to my parents and to my children, Natalie, Martin and Elizabeth. This
research has been supported by the department of physics at the University of Illinois
at Urbana-Champaign, by the National Center for Supercomputing Applications,
by the National Science Foundation (NSF-DMR 91-17822 and CARM grant) and by the
Office of Naval Research (ONR-N00014-92-J-1320).

REFERENCES

   Mechanical Theory of Transport Properties,” I. Prigogine, ed., Interscience, New