

## Global Stationary Phase and the Sign Problem

André G. Moreira,<sup>1</sup> Stephan A. Baeurle,<sup>2</sup> and Glenn H. Fredrickson<sup>1</sup>

<sup>1</sup>Materials Research Laboratory, University of California, Santa Barbara, California 93106, USA

<sup>2</sup>Institut für Physikalische und Theoretische Chemie, Universität Regensburg, 93053 Regensburg, Germany

(Received 21 April 2003; published 10 October 2003)

We present a computational strategy for reducing the sign problem in the evaluation of high dimensional integrals with nonpositive definite weights whose logarithms are analytic. The method involves stochastic sampling with a positive semidefinite weight that is adaptively and optimally determined during the course of a simulation. The optimal criterion, which follows from a variational principle for analytic actions  $S(z)$ , is a global stationary phase condition that the average gradient of the phase  $\text{Im}S$  along the sampling path vanishes. Numerical results are presented from simulations of a model adapted from statistical field theories of classical fluids.

DOI: 10.1103/PhysRevLett.91.150201

PACS numbers: 02.70.-c, 05.10.-a, 82.20.Wt

A familiar problem that arises in the context of lattice gauge theory [1], quantum chemistry [2], correlated electron physics [3], and equilibrium field theories of classical fluids [4] is the evaluation of integrals of the form

$$Z = \int_{C_1} dx \exp[-S(x)], \quad (1)$$

where the path of integration  $C_1$  is the real axis and the action (or effective Hamiltonian)  $S(x)$  is *complex*. In the cases of primary interest  $x \in R^n$  is an  $n$  vector providing a discrete representation (lattice sites or spectral elements) of one or more classical or quantum fields. The dimension  $n$  is typically large, of order  $10^3$ – $10^6$ . Here we use one-dimensional notation, although the formalism is primarily intended for cases of  $n \gg 1$ .

For real  $S(x)$ , there is a variety of powerful methods available for evaluating  $Z$ , including Monte Carlo and (real) Langevin simulations [5]. However, in the case of complex  $S = S_R + iS_I$ , the integrand is not positive semidefinite, so the Monte Carlo method is not immediately applicable. Simulations can be carried out using the positive semidefinite weight  $\exp(-S_R)$ , but then an oscillatory phase factor of  $\exp(-iS_I)$  must be included in the computation of averages [6]. The rapid oscillations of this factor (i.e., the “sign problem”), which become more pronounced for large  $n$ , can dramatically slow convergence in such simulations. Alternatively, a “complex Langevin” simulation technique has been devised in which the field variables  $x$  are extended to the complex plane and a Langevin trajectory prescribed for the purpose of generating Markov chains of states [7]. Unfortunately this method is not guaranteed to converge and pathological behavior has been noted for specific models [8,9]. In the present Letter we describe a new simulation approach that is useful for reducing the sign problem in integrals of the form of Eq. (1), where  $S(z)$  is an *analytic* function of the complex  $n$  vector  $z = x + iy$ .

We begin by considering a displacement of the original integration path along the real  $x$  axis,  $C_1$ , to a new

parallel path  $C_y$  defined by  $z = x + iy$ ,  $x_j \in (-\infty, \infty)$ , in which  $y \in R^n$  is an arbitrary displacement of  $C_1$  along the imaginary axis. Note that the displacement  $y_j$  need not be uniform in  $j$  for the  $n > 1$  case. Provided  $S(z)$  is analytic in the rectangular strip bounded by  $C_1$  and  $C_y$  and  $|\exp[-S(R + iy)]| \rightarrow 0$  for  $R \rightarrow \pm\infty$ , it follows that

$$Z = \int_{C_y} dz \exp[-S(z)] = \int_{C_y} dx \exp[-S(x + iy)], \quad (2)$$

and the resulting  $Z$  is independent of the choice of  $y$ . Upon decomposing  $S$  into real and imaginary parts  $S_R(x, y) + iS_I(x, y)$ ,  $Z$  can be rewritten as

$$Z = Z_y \int_{C_y} dx P_y(x) \exp[-iS_I(x, y)], \quad (3)$$

where  $Z_y \equiv \int_{C_y} dx \exp[-S_R(x, y)]$  and  $P_y(x)$  is a normalized, positive semidefinite, probability distribution for a random variable  $x$  at the fixed value of  $y$ :

$$P_y(x) = \exp[-S_R(x, y)]/Z_y. \quad (4)$$

It follows that the average of an analytic observable  $f(x)$  can be evaluated alternatively from the formulas

$$\begin{aligned} \langle f(x) \rangle &= Z^{-1} \int_{C_1} dx \exp[-S(x)] f(x) \\ &= \frac{\langle \exp[-iS_I(x, y)] f(x + iy) \rangle_y}{\langle \exp[-iS_I(x, y)] \rangle_y}, \end{aligned} \quad (5)$$

where  $\langle h(x) \rangle_y \equiv \int_{C_y} dx P_y(x) h(x)$  denotes an average with probability weight  $P_y(x)$ .

It is the second expression in Eq. (5) that is of interest in the present Letter. A poor choice of  $y$  will lead to significant oscillations in the phase factor  $\exp[-iS_I(x, y)]$  as  $x$  is stochastically varied along the sampling path  $C_y$  in a simulation. This would drive both numerator and denominator in Eq. (5) to zero and dramatically slow or prevent convergence of average quantities of interest. One approach to alleviate this difficulty would be to choose  $y = y^*$ , where  $y^*$  is the imaginary

component of a saddle point  $z^*$  defined by  $S'(z^*) = 0$ . The deformed integration path  $C_{y^*}$  would then be a line passing through the saddle point parallel to the real axis. If this path happened to be a constant phase (steepest ascent) path locally around the saddle point, then the phase oscillations would be reduced on trajectories that remain close to the saddle point [10]. In general, however, path  $C_{y^*}$  will not be a constant phase path, even in the close vicinity of  $z^*$ . A local analysis about each saddle point, costing  $O(n^2)$  in computational effort, can be used to identify proper constant phase paths. However, in typical problems where field fluctuations are strong, significant weight is given to trajectories that are *not localized* around saddle points.

The essence of our method is a *global* strategy for selecting an optimal displacement  $y$ , denoted  $\bar{y}$ . To this end, we introduce a “generating” function (functional)

$$G(y) = \ln \int_{C_y} dx \exp[-S_R(x, y)]. \quad (6)$$

Invoking the Cauchy-Riemann (CR) equations, it is straightforward to show that the first derivative of  $G(y)$  is given by

$$\frac{\partial G(y)}{\partial y_j} = \left\langle \frac{\partial}{\partial x_j} S_I(x, y) \right\rangle_y, \quad (7)$$

where  $x_j$  and  $y_j$  are components of the real  $n$  vectors  $x$  and  $y$ . The second derivative follows from repeated application of the CR equations and an integration by parts

$$\begin{aligned} \frac{\partial^2 G(y)}{\partial y_j \partial y_k} = & \left\langle \left[ \frac{\partial}{\partial x_j} S_I - \left\langle \frac{\partial}{\partial x_j} S_I \right\rangle_y \right] \left[ \frac{\partial}{\partial x_k} S_I - \left\langle \frac{\partial}{\partial x_k} S_I \right\rangle_y \right] \right\rangle_y \\ & + \left\langle \left[ \frac{\partial}{\partial x_j} S_R \right] \left[ \frac{\partial}{\partial x_k} S_R \right] \right\rangle_y, \end{aligned} \quad (8)$$

which is the sum of two positive definite forms. It follows that  $G(y)$  is manifestly a *convex* function for any  $y$ .

We now claim that the “optimal” choice  $y = \bar{y}$  is such that

$$\left. \frac{\partial G(y)}{\partial y_j} \right|_{\bar{y}} = \left\langle \frac{\partial}{\partial x_j} S_I(x, \bar{y}) \right\rangle_{\bar{y}} = 0. \quad (9)$$

Evidently such a point would be a *local minimum* of  $G(y)$ . Moreover, it implies that  $S_I$  has *vanishing gradients on average* along the sampling path  $C_{\bar{y}}$ . This condition can be viewed as a *global*, rather than a local [10], stationary phase criterion and would seem to be an excellent way to minimize the effect of phase fluctuations. Since  $G(y)$  has a unique minimum, it follows that  $\bar{y}_j$  is homogeneous in  $j$  for bulk systems with translationally invariant actions. The method evidently produces nontrivial inhomogeneous  $\bar{y}$  when applied to field theories in bounded geometries.

It remains to discuss how to incorporate this optimal choice of sampling path into a simulation algorithm. We

propose the following “optimal path sampling” (OPS) algorithm:

- (i) Initialize vectors  $x$  and  $y = y^{(k)}$  with  $k = 0$ .
- (ii) Carry out a stochastic simulation in  $x$  at fixed  $y^{(k)}$  to generate a Markov chain of  $x$  states of length  $M$ .  $P_{y^{(k)}}(x)$  should be used as a statistical weight for importance sampling. The simulation method could be the METROPOLIS Monte Carlo algorithm, its “smart” or “hybrid” variants [11], or a real Langevin technique.
- (iii) Evaluate  $G(y^{(k)})$  and  $\partial G(y^{(k)})/\partial y^{(k)}$  by averaging over the  $x$  configurations accumulated in the  $M$ -state simulation. Update  $y$  to approach  $\bar{y}$  by making a steepest descent step

$$y^{(k+1)} = y^{(k)} - \lambda \frac{\partial G(y^{(k)})}{\partial y^{(k)}},$$

where  $\lambda$  is an adjustable relaxation parameter. Alternatively, the accumulated information on  $G(y)$  could be used to carry out approximate line minimizations, which would permit conjugate gradient updates from  $y^{(k)}$  to  $y^{(k+1)}$ .

- (iv) Repeat steps (ii) and (iii) for  $k = 1, 2, \dots$  until the sequence of  $y^{(k)}$  converges to within some prescribed tolerance to  $\bar{y}$ . The simulation has now *equilibrated*.

- (v) Carry out a long stochastic simulation (“production run”) with statistical weight  $P_{\bar{y}}(x)$ .

- (vi) Compute averages over the simulated states according to Eq. (5) with  $y = \bar{y}$ .

Evidently, the parameters  $M$  and  $\lambda$  can be adjusted to accelerate the “equilibration” period.

Our OPS method has some similarities to (and was inspired by) the complex Langevin (CL) simulation technique. In that approach, one generates a Markov chain of states in the complex plane by integrating the Langevin equations [7]

$$\frac{\partial x}{\partial t} = -\text{Re} \frac{dS}{dz} + \eta(t), \quad (10)$$

$$\frac{\partial y}{\partial t} = -\text{Im} \frac{dS}{dz}, \quad (11)$$

where  $\eta(t)$  is a real Gaussian white noise with  $\langle \eta(t) \rangle = 0$  and  $\langle \eta_j(t) \eta_k(t') \rangle = 2\delta(t - t') \delta_{jk}$ . Ensemble averages  $\langle f(x) \rangle$  are computed as time averages of  $f(x + iy)$  over the chain of states. Under conditions where the CL method converges, we have observed that  $y$  drifts to a nearly constant value that is not associated with any saddle point  $y^*$ . Equation (11) reduces approximately in this case to  $\langle \text{Im} dS/dz \rangle_y = 0$ , which is equivalent to the condition (9). The OPS technique is also distinct from the “stationary phase Monte Carlo” methods, which apply filtering and sparse sampling methods to suppress phase oscillations [2,12]. These methods are effective but apparently have no variational basis.

Before providing a numerical example of the OPS method, it is illustrative to see how our global stationary

phase criterion works in a simple one-dimensional example

$$\text{Ai}(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dx \exp[i(x^3/3 + tx)], \quad (12)$$

which is a representation of the Airy function. In this case  $S(x, t) = -i(x^3/3 + tx)$  and Eq. (9) leads to  $\bar{y}^2 - \langle x^2 \rangle_{\bar{y}} - t = 0$ . This equation has a single root, corresponding to the minimum of  $G(y)$ , that yields  $\bar{y}(t)$ . For example,  $\bar{y}(1) = 1.19149$ . Of particular interest is the effect of the optimal displacement on phase oscillations. In Fig. 1 we plot  $\text{Re}\{\exp[-iS_I(x, y, t)]\}$  vs  $x$  at  $t = 1$  for  $y = 0$  (no shift) and  $y = \bar{y}$  (optimal). Clearly the optimal shift dramatically suppresses phase oscillations over the interval  $-2 \lesssim x \lesssim 2$ . The global stationary phase criterion has no effect outside this interval, because  $P_{\bar{y}}(x)$  decays supraexponentially there as  $\sim \exp(-x^2\bar{y})$  and so no statistical weight is given to  $|x| \gtrsim 2$ .

As a numerical test of the OPS method, we have carried out simulations of the model

$$S(x) = \sum_{j=1}^n [\alpha x_j^2 + (x_{j+1} - x_j)^2 - \chi \exp(-ix_j)], \quad (13)$$

which can be viewed as a lattice field theory for the one-dimensional classical Yukawa fluid in the grand canonical ensemble ( $\alpha$  is a measure of interaction strength and  $\chi$  is the activity). For the case of  $n > 1$ , periodic boundary conditions are applied. The model has a saddle point  $z_j^* = iy_j^*$  that lies on the imaginary axis and is homogeneous in the index  $j$  (as well as an inhomogeneous “1D crystal-like” saddle point). Its location is given by the solution of  $\chi \exp(y_j^*) + 2\alpha y_j^* = 0$ . The optimal displaced path  $\bar{y}_j$  is homogeneous in  $j$  and is given by the solution of  $\chi \exp(\bar{y}_j) \langle \cos x_j \rangle_{\bar{y}} + 2\alpha \bar{y}_j = 0$ . We see that  $y^*$  and  $\bar{y}$  are coincident under conditions ( $\alpha \gg 1$ ) where the random variable  $x$  fluctuates closely about the saddle point  $x^* = 0$ .

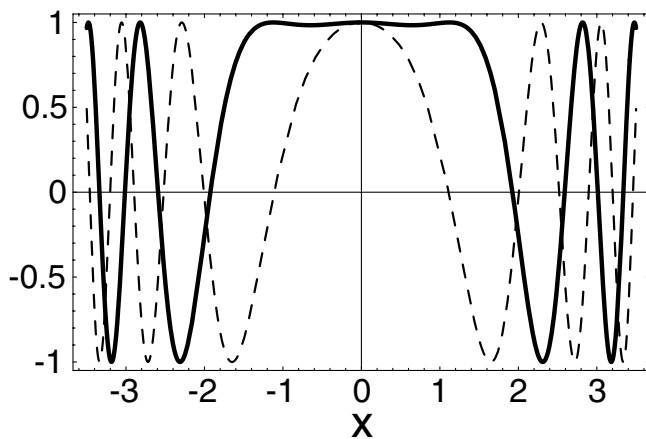


FIG. 1. Variation of the phase factor of the Airy integrand  $\text{Re}\{\exp[-iS_I(x, y, 1)]\}$  with  $x$  for  $y = 0$  (dashed line) and  $y = \bar{y} = 1.19149$  (solid line).

In the strongly fluctuating regime ( $\alpha \ll 1$ ),  $\langle \cos x_j \rangle_{\bar{y}}$  will be dramatically reduced, resulting in a large shift of  $\bar{y}$  away from  $y^*$ . These expectations are borne out in numerical simulations of the model.

We have carried out a conventional METROPOLIS Monte Carlo (MC) simulation [i.e., Eq. (5) with  $y = 0$ ], OPS, and CL simulations of the model with action Eq. (13). The results were obtained from runs with a total of  $10^7$  MC cycles or Langevin steps, a time step of 0.001 in the case of CL, and parameters  $M = 1000$ ,  $\lambda = 0.05$  for OPS. In Fig. 2 we compare the results obtained from OPS and CL simulations with  $n = 1$  and  $\chi = 1$ . The top panel (a) shows  $\langle z \rangle$  as a function of  $\alpha$ , while the bottom panel (b) displays the real and imaginary parts of the “sign”  $\langle \exp(-iS_I) \rangle$ . In contrast to OPS, CL fails to converge, or converges very slowly, for  $\alpha \leq 0.15$ . The conventional MC method also converges, but the average sign is approximately 0.8, as opposed to  $\sim 1$  shown by the OPS.

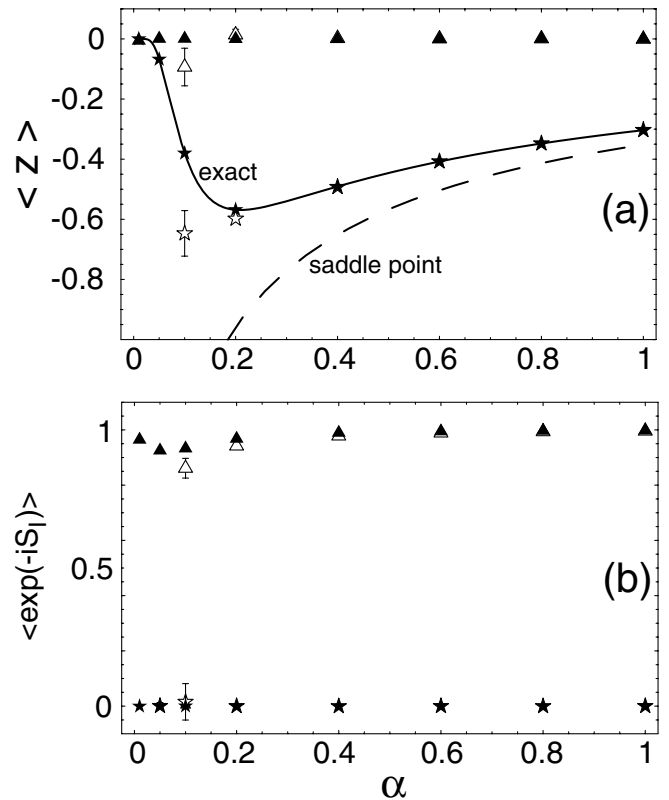


FIG. 2. Comparison between OPS and CL simulations. (a) The average of  $z = x + iy$  as a function of the parameter  $\alpha$  for the model of Eq. (13) with  $n = 1$  and  $\chi = 1$ . Open and filled symbols are results, respectively, from CL and OPS. Stars denote the average of the imaginary part  $y$ , and triangles the average of the real part  $x$ . The full line is the exact solution, the dashed line the corresponding saddle point. (b) The average sign  $\exp(-iS_I)$  for the same parameters as in (a). The convention for the symbols is the same as in (a). Error bars are comparable to the symbol sizes if not explicitly shown.

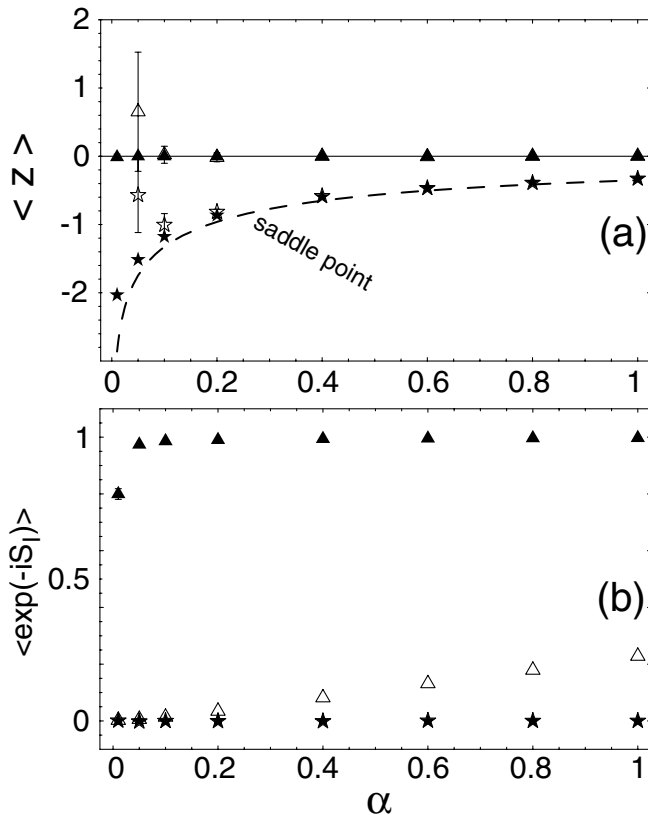


FIG. 3. Comparison between OPS and the conventional MC method. (a) The average of  $z = x + iy$  as a function of the parameter  $\alpha$  for the model of Eq. (13) with  $n = 10$  and  $\chi = 1$ . Open and filled symbols are results, respectively, from the MC method and OPS. Stars, triangles, and the dashed curve as in Fig. 2. (b) The average sign  $\exp(-iS_I)$  for the same parameters as in (a). At small  $\alpha$  the real part of the sign in the MC method rapidly approaches zero, and the averages fail to converge. Error bars are comparable to the symbol sizes if not explicitly shown.

It is often observed [3] that the sign in conventional MC simulations decreases exponentially with  $n$ , causing a breakdown of the method. This is illustrated for the present model in Fig. 3 with parameters  $n = 10$  and  $\chi = 1$ . The conventional MC method fails to converge for  $\alpha \lesssim 0.1$  in contrast to OPS. Moreover, the real part of the sign is strongly suppressed in the MC results, even at large values of  $\alpha$ . The sign problem is evidently strongly suppressed, if not eliminated entirely for this model in OPS.

The OPS method is applicable to any field theory with an action  $S(z)$  that is analytic throughout a domain of  $z$  relevant to numerical simulations. This includes

the important cases of classical fluids in the grand canonical ensemble and path integral formulations of time-dependent quantum chemical problems. Other situations including fluids in the canonical ensemble, strongly correlated electrons, and lattice gauge theories are characterized by analytic  $\exp(-S)$ , but with zeros along the real axis and hence logarithmic singularities in  $S$ . It is unclear at present whether extensions of the OPS technique will prove fruitful for such systems. Finally, we note that the displaced paths considered here were parallel to the real axis. Generalization of the method to optimize both the displacement and shape of the path could be even more powerful.

In summary, we have identified a variational principle that permits a global stationary phase analysis of integrals of arbitrary dimension in which the logarithm of the integrand is an analytic function. We expect that this technique will have important implications for analytical and numerical investigations of field theories in the complex plane.

This work was supported by the NSF under the MRSEC program Grant No. DMR00-80034. We are grateful to H. Metiu, C. Garcia-Cervera, R. Sugar, J. S. Langer, M. P. A. Fisher, and D. Scalapino for helpful discussions.

- 
- [1] I. Montvay and G. Münster, *Quantum Fields on the Lattice* (Cambridge University Press, Cambridge, 1994).
  - [2] V. S. Filinov, Nucl. Phys. **B271**, 717 (1986); J. D. Doll and D. L. Freedman, Adv. Chem. Phys. **73**, 289 (1988); N. Makri and W. H. Miller, Chem. Phys. Lett. **139**, 10 (1987).
  - [3] E. Y. Loh, Jr. *et al.*, Phys. Rev. B **41**, 9301 (1990).
  - [4] G. H. Fredrickson, V. Ganesan, and F. Drolet, Macromolecules **35**, 16 (2002); S. A. Baeurle, Phys. Rev. Lett. **89**, 080602 (2002).
  - [5] D. P. Landau and K. Binder, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, New York, 2000).
  - [6] H. Q. Lin and J. E. Hirsch, Phys. Rev. B **34**, 1964 (1986).
  - [7] G. Parisi, Phys. Lett. **131B**, 393 (1983); J. R. Klauder, Phys. Rev. A **29**, 2036 (1984).
  - [8] S. Lee, Nucl. Phys. **B413**, 827 (1994).
  - [9] W. J. Schoenmaker, Phys. Rev. D **36**, 1859 (1987).
  - [10] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill Publishing Company, New York, 1978).
  - [11] A. D. Kennedy, Parallel Comput. **25**, 1311 (1999); P. J. Rossky and J. D. Doll, J. Chem. Phys. **69**, 4628 (1978).
  - [12] D. Sabo, J. D. Doll, and D. L. Freedman, J. Chem. Phys. **116**, 3509 (2002).