

AN "EXACT" INTEGRAL EQUATION APPROACH TO THE INVERSE PROBLEM IN TWO-DIMENSIONAL FLUIDS

James R. ABNEY^{a,1} and John C. OWICKI^{a,b,2}

^a *Cell and Molecular Biology Division, Lawrence Berkeley Laboratory, Berkeley, CA 94720, USA*

^b *Department of Biophysics and Medical Physics, University of California, Berkeley, CA 94720, USA*

Received 14 July 1989; in final form 21 September 1989

It is shown that the Born–Green–Yvon integral equation can be used to extract interparticle pair potentials from two- and three-particle distribution functions in certain two-dimensional fluids. The approach was tested against Monte Carlo data from two pairs of closely related inverse-power-law potentials and found satisfactory over a density regime from infinite dilution to near crystal packing.

Considerable effort has been devoted to the development of a statistical-mechanical theory capable of describing the relationship between molecular forces and molecular order in fluid systems. Familiar products of this work include the Born–Green–Yvon (BGY), Percus–Yevick, and hypernetted-chain (HNC) integral equations. Initial interest centered on using these equations to characterize and predict fluid structure through the use of distribution functions, beginning with assumptions about the interparticle interactions. More recently, however, interest has also developed in the so-called "inverse problem": deducing interactions from measured structural information [1–3]^{#1}.

¹ Present address and address for correspondence: Laboratories for Cell Biology, Department of Cell Biology and Anatomy, School of Medicine, University of North Carolina, Chapel Hill, NC 27599, USA.

² Present address: Molecular Devices Corporation, 4700 Bohannon Drive, Menlo Park, CA 94025, USA.

^{#1} Ref. [1] is a seminal article suggesting that structural information can be used to deduce information on interparticle potentials. Ref. [2] is a research article which reviews more recent work on the inverse problem. Ref. [3] is a review article which discusses studies of the inverse problem in two-dimensional membrane systems. In such multi-component systems, the interaction potentials computed corresponding to the observed degrees of freedom (proteins) contain free-energy contributions from the unobserved degrees of freedom (membrane lipids and water) that have effectively been integrated out.

Both the "forward" and inverse problems typically are formulated in terms of the same integral equations. For three-dimensional systems, structural information is usually available (from scattering experiments) only in the form of the pair, or radial, distribution function $g(r)$. Consequently, most analyses of the inverse problem are based on approximate integral equations that involve only $g(r)$. Unfortunately, these equations are not accurate at high densities [2–4], although Levesque et al. [2] have markedly improved their performance using a computationally intensive predictor–corrector method, which iterates between the modified HNC equation (in their case) and computer simulations.

Another way to improve the accuracy of the analysis is to use exact rather than approximate integral equations. An example is the BGY equation, which in two dimensions states

$$-k_B T \frac{d[\ln g(r)]}{dr} = \frac{du(r)}{dr} + \int_0^\infty \frac{du(s)}{ds} K(r, s) ds, \quad (1a)$$

where

$$K(r, s) \equiv \int_0^{2\pi} \rho \frac{g^{(3)}(r, s, \theta)}{g(r)} \cos \theta s d\theta. \quad (1b)$$

Here k_B is Boltzmann's constant, T is the temperature, $u(r)$ is the pair potential, ρ is the particle num-

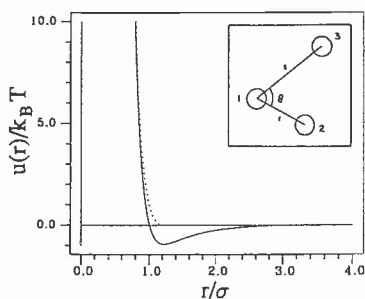


Fig. 1. Coordinate system (inset) and potentials used in the inverse scheme. The pair potential for fluid R (dotted line) was obtained from eq. (2) by truncating $u(r)$ at $r=1.2247\sigma$ and adding $k_B T$. The potential for fluid A (solid line) was obtained by truncating eq. (2) (for computational economy) at $r=3.0169\sigma$ and adding $-u(3.0169\sigma)$ to maintain continuity.

ber density, and $g^{(3)}(r, s, \theta)$ is the three-particle distribution function; see fig. 1. This equation is exact for fluids with pairwise-additive potentials if we do not invoke the standard superposition approximation, which replaces the three-particle distribution function with a product of two-particle distribution functions. Eq. (1) can be used to extract pair forces in those two-dimensional systems, such as membranes [3] and surfaces [5], for which pair and triplet information is directly accessible from molecular coordinates revealed by electron microscopy.

We have previously used eq. (1) to calculate interparticle potentials in both biological membranes [6,7] and simulated systems [8]. In the experimental work, the goal was to deduce the pair potential between intrinsic membrane proteins. In the simulation work, the goal was to assay for the efficacy of the technique by reproducing the potential used in the simulation. We were able there to demonstrate that the BGY equation can recover an inverse-power-law potential from Monte Carlo data at intermediate densities (where the form of the distribution function still depends markedly on the potential); a similar observation was made for a three-dimensional system by Haymet and Rice [9]. However, we are fundamentally interested in using eq. (1) to characterize interactions at all relevant particle densities, which vary from infinite dilution to crystal packing in membranes. The BGY approach must therefore be tested more fully and an optimal algorithm established. Here we ask: is the BGY equation capable

of deducing pair potentials over the full range of fluid particle densities? In particular, we investigate the numerical stability of the inversion at high densities.

To address these issues, Monte Carlo simulations were run, at various densities, on two closely related potentials derived from the decomposition (following the Weeks–Chandler–Andersen prescription [10]) of a general 4–6 interaction

$$u(r) \equiv \frac{27}{4} k_B T [(\sigma/r)^6 - (\sigma/r)^4] \quad (2)$$

into purely repulsive (fluid R) and attractive-plus-repulsive (fluid A) components; see fig. 1. Here σ defines the zero crossing of the potential, while $k_B T$ gives the depth of the attractive well in fluid A. Our choice of these potentials was motivated by a desire to model the weak, long-ranged attractions and soft repulsions predicted to act between proteins in bilayer membranes [3]. For completeness, we also performed some tests using the more familiar 6–12 potential.

Two- and three-particle distribution functions corresponding to fluids R and A were obtained numerically by averaging over discrete bins the interparticle separations in configurations generated during the simulations; $g(r)$ at two reduced densities ($\rho^* \equiv \rho\sigma^2 = 0.3$ and 0.8), and further details, are given in fig. 2. At low densities, the structures of the two fluids as manifest in $g(r)$ are markedly different; both are short ranged and weakly defined. At high densities, the structures are very similar; both are long ranged, well defined, and determined largely by the common repulsive forces [10,12]. A similar convergence in structure is seen in $g^{(3)}(r, s, \theta)$ (data not shown).

The force was determined by direct inversion of the BGY equation using the distribution functions $g(r)$ and $g^{(3)}(r, s, \theta)$ computed from the simulations. At infinite dilution, $g^{(3)}(r, s, \theta)$ can be neglected and eq. (1) inverted analytically, yielding $u(r) = -k_B T \ln g(r)$. At all other densities, both distribution functions contribute to $u(r)$ and the BGY equation must be inverted numerically. This was accomplished by converting the integral equation into a system of coupled linear equations [8]. The resulting matrix problem was solved for the force using LINPACK subroutines SGECO and SGESL [13]. The potential was then obtained by numerical integration of the force using Simpson's rule [14].

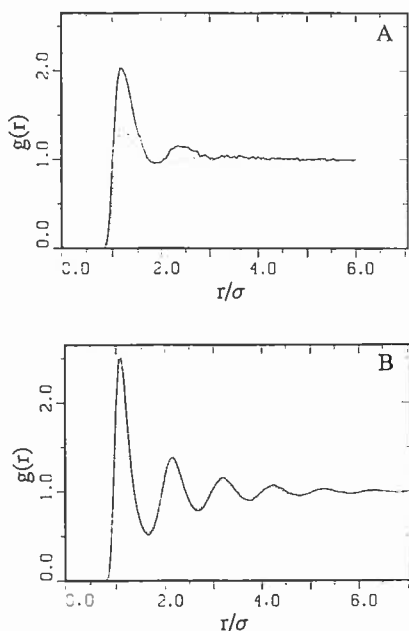


Fig. 2. Radial distribution functions for fluids R (dotted line) and A (solid line) at $\rho^* = 0.3$ (A) and 0.8 (B). We analyzed 256 particles in a square cell using periodic boundary conditions and the standard Metropolis et al. algorithm [11]. Results were averaged over bins of width 0.05σ every 10 cycles for a total of 2000 (A) to 5000 (B) cycles. A cycle corresponds to one sequentially attempted perturbation of each particle in the system. The upper limit on data acquisition was chosen to carry $g(r)$ to (near) its asymptotic value. It is reasonable to assume, without a priori knowledge, that the range of the potential is comparable to or shorter than the range of the structure revealed in the distribution functions.

The most effective numerical algorithm was the simplest: equal weighting of terms in the sum representing the integral and a five-point cubic-quartic fit for the derivative (which was computed after taking the logarithm). This approach flexibly handled the sometimes rapid changes in $u(r)$ and $g(r)$. We avoided smoothing, which effectively changes the number of particles associated with each bin and which can thus introduce artifacts into the pair potential. Specifically, a depletion of particles might appear as an effective repulsion, while an addition of particles might appear as an effective attraction. In contrast, variations in the unsmoothed distribution functions correspond to legitimate statistical fluctuations and probably contribute only noise.

Calculated values of the pair potentials for fluids

R and A are shown in fig. 3 corresponding to data in fig. 2. The agreement between measured and analytical values was good at all densities analyzed ($\rho^* = 0.0, 0.3, 0.5, 0.7$, and 0.8). Slightly poorer performance was obtained at the highest densities. This was selected in the condition of the matrix (as reported by LINPACK), which diminished monotonically by two orders of magnitude as ρ^* increased from 0.3 to 0.8, and in the difficulty of correctly computing $d[\ln g(r)]/dr$ as $g(r)$ varied more rapidly. However, even when the distribution functions were virtually identical, the forces were accurately extracted by the BGY equation. Note that the validity of the results in fig. 3 was verified by repeating the analysis on a second, independent set of simulations.

We also performed less exhaustive tests on repulsive and attractive-plus-repulsive fluids derived from a Weeks-Chandler-Andersen decomposition of the

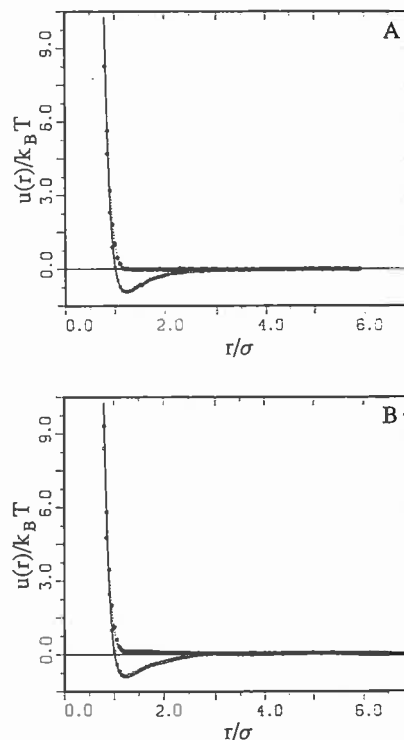


Fig. 3. Pair potentials for fluids R (filled circles) and A (open circles) derived from BGY inversion of Monte Carlo data. Results shown correspond to the data in fig. 2. Analytical potentials for fluids R (dotted line) and A (solid line) are plotted for reference.

6–12 potential, again setting the well depth in the starting potential equal to $k_B T$. Simulations were thus run in the two-dimensional hypercritical fluid phase [15]; we investigated systems at relatively low density ($\rho^*=0.3$) and near the fluid–solid transition ($\rho^*=0.8$). Results were comparable to those shown in fig. 3, though slightly poorer at the higher density.

We note that other groups have applied the BGY equation to the inverse problem in three-dimensional fluids [9,16–19], albeit less successfully. It seems likely that the difficulties they encountered arose from errors that are inherited when the superposition approximation is used in the description of three-dimensional systems; further evidence to support this conclusion is given by Haymet and Rice [9]. Kumaravadivel et al. [18] suggested that some of these problems may be reduced using a k -space solution of the BGY equation; we have not attempted this form of solution because our data are obtained directly in r -space (unlike the structure factor $S[k]$) and because the transform into k -space is sensitive to values of the distribution functions at larger r for which it is difficult to obtain good data.

In conclusion, the BGY equation provides a quick and relatively inexpensive method for determining pair potentials at all densities in systems such as membranes and surfaces for which pair and triplet distribution functions can be measured. At low densities, it can extract the force from relatively weak structural information. At high densities, where other integral equations fail, it can discern differences in the potentials from very subtle differences in the pair and triplet distribution functions. Finally, under circumstances in which greater accuracy is desired than that provided by a simple inversion of the BGY equation, improvements in the potential may be possible at greater cost using an iterative technique such as that mentioned previously [2].

We would like to acknowledge stimulating conversations with Karl Runge, Beth Scalettar, and Richard Scalettar. This work was funded by the National Institutes of Health grant AI-22860 (JCO) and

an appointment to the Alexander Hollaender Distinguished Postdoctoral Fellowship Program, supported by the United States Department of Energy, Office of Health and Environmental Research, and administered by Oak Ridge Associated Universities (JRA).

References

- [1] M.D. Johnson, P. Hutchinson and N.H. March, Proc. Roy. Soc. A 282 (1964) 283.
- [2] D. Levesque, J.J. Weis and L. Reatto, Phys. Rev. Letters 54 (1985) 451.
- [3] J.R. Abney and J.C. Owicki, in: Progress in protein–lipid interactions, eds. A. Watts and J.J.H.M. de Pont (Elsevier, Amsterdam, 1985) p.1.
- [4] W.G. Madden and S.A. Rice, J. Chem. Phys. 72 (1980) 4208.
- [5] M. Stenberg and H. Nygren, Phys. Rev. Letters 59 (1987) 1164.
- [6] J. Braun, J.R. Abney and J.C. Owicki, Nature 310 (1984) 316.
- [7] J.R. Abney, J. Braun and J.C. Owicki, Biophys. J. 52 (1987) 441.
- [8] J. Braun, J.R. Abney and J.C. Owicki, Biophys. J. 52 (1987) 427.
- [9] a.D.J. Haymet and S.A. Rice, J. Chem. Phys. 76 (1982) 661.
- [10] D. Chandler, J.D. Weeks and H.C. Andersen, Science 220 (1983) 787.
- [11] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller and E. Teller, J. Chem. Phys. 27 (1953) 720.
- [12] B. Widom, Science 157 (1967) 375.
- [13] J.J. Dongarra, C.B. Moler, J.R. Bunch and G.W. Stewart, LINPACK: users' guide (SIAM, Philadelphia, 1979).
- [14] W.H. Press, B.P. Flannery, S.A. Teukolsky and W.T. Vetterling, Numerical recipes: the art of scientific computing (Cambridge Univ. Press, Cambridge, 1986).
- [15] J.M. Phillips, L.W. Bruch and R.D. Murphy, J. Chem. Phys. 75 (1981) 5097.
- [16] W.S. Howells and J.E. Enderby, J. Phys. C 5 (1972) 1277.
- [17] N.K. Ailawadi, P.K. Banerjee and A. Choudry, J. Chem. Phys. 60 (1974) 2571.
- [18] R. Kumaravadivel, R. Evans and D.A. Greenwood, J. Phys. F 4 (1974) 1839.
- [19] W. Schommers, Phys. Rev. A 28 (1983) 3599.