Generalized Cooperative-Ring-Exchange Theory of the Fractional Quantum Hall Effect

D. H. Lee
IBM Thomas J. Watson Research Center, Yorktown Heights, New York 10598

G. Baskaran
Department of Physics, Princeton University, Princeton, New Jersey 08544

and

S. Kivelson
Department of Physics, State University of New York at Stony Brook, Stony Brook, New York 11794
(Received 12 January 1987)

Through the calculations of the diagonal and off-diagonal elements of the full N-particle density matrix we show that all essential features of the Laughlin variational wave functions can be derived in a generalized cooperative-ring-exchange theory of the fractional quantum Hall fluid.

PACS numbers: 71.45.Gm, 73.40.Lq

Although the Laughlin wave functions[1] are believed to contain all essential physics of the fractional quantum Hall effect,[2] the physical origin of the important correlations that are built into the wave functions remains unclear. The success of Laughlin wave functions is due, on the one hand, to their good short-range correlation[3] (good in the sense that they minimize the short-range part of the electron-electron repulsion) and, on the other hand, to their rigidity against long-wavelength density fluctuations (incompressibility). In Laughlin's theory, the analyticity constraints imposed by projection onto the lowest Landau level imply that the long-wavelength correlations are fixed by the short-distance behavior.

Recently, Kivelson *et al.*[4] adopted a totally different approach to the problem, which was later extended by Baskaran.[5] This approach focuses directly on the long-distance correlations. The idea is to study the effects of a specific class of quantum fluctuations that are particularly sensitive to small changes in the density, viz. the cooperative-ring-exchange (CRE) processes. A CRE event involves the cooperative tunneling of a set of electrons from an initial metastable configuration to a final configuration in which the "electron positions" are cyclically permuted. (Since for any state $|\Psi\rangle$ in the first Landau manifold ${}_{\rm cs}\langle\{z_i\}|\Psi\rangle = {}_{\rm ps}\langle\{z_i\}|\Psi\rangle$, where $|\{z_i\}\rangle_{\rm cs}$ is the coherent state where the electron guiding centers are localized at $\{z_i\}$ and $|\{z_i\}\rangle_{\rm ps}$ is the position eigenstate where the electrons are localized at $\{z_i\}$, we can use "electron coordinates" and "guiding-center coordinates" interchangeably.[6]) In the original CRE theory,[4,5] one studied the cooperative tunneling of electrons around hypothetical "backbone" configurations (Wigner crystal in Ref. 4 and "incompressible liquid" in Ref. 5). In the imaginary-time direction, the backbone configurations are assumed either to be completely quasistatic,[4] or to be quasistatic within imaginary-time slices of width τ_0, and completely uncorrelated from one such time slice to another.[5] Under these assumptions, it was argued that because of the constructive interference between the Aharonov-Bohm phases of the CRE processes, $\nu = 1/(2n+1)$ is more stable relative to all nearby densities. Although the CRE theory is appealing, its relevance to the real quantum Hall effect can be questioned because of the hypothesis of backbone configurations and the unrealistic treatment of their temporal evolution. Moreover, even if one ignores the problem associated with the backbone configurations, one still has to face the following question: What is the connection between the CRE theory and the Laughlin wave functions?

In this Letter we completely remove the backbone hypothesis and show that the long-wavelength properties of the Laughlin wave functions can be derived within a generalized CRE theory. In particular, we can show that the sum of the probability amplitudes of all closed-loop nearest-neighbor tunneling processes for a given electron configuration $\{z_i\}$ determine the value of the diagonal density-matrix element (DDME) corresponding to that configuration. For example, at $\nu = 1/(2n+1)$ DDME has its largest value for a perfect Wigner crystal with lattice parameter $a(\nu) = (4\pi/\sqrt{3}\nu)^{1/2}$ [in units of the magnetic length $l_0 = (\hbar c/eB)^{1/2}$], because all elementary triangles enclose the right amount of magnetic flux and hence all rings constructively interfere. For a general liquidlike electron configuration, some plaquettes enclose excessive flux (positive plaquettes), and some enclose insufficient flux (negative plaquettes). If the positive and negative plaquettes are bound closely together to form a globally "neutral" system, a dominant fraction of the tunneling rings still constructively interfere, and hence the configuration has a DDME comparable with, though somewhat less than, the Wigner crystal. On the other hand, if a liquidlike configuration is composed of uncorrelated positive and negative plaquettes, almost all ring exchanges destructively interfere, and hence the

configuration has a vanishing DDME relative to that of the Wigner crystal. If ν deviates slightly from $1/(2n+1)$, the Wigner crystal with lattice constant $a(\nu)$ has a negligible DDME because the elementary plaquette has the wrong area, causing destructive interference between rings of all sizes. Instead, configurations where the excess (or the deficit) density is accumulated as localized defects such that all rings enclosing these defects have their enclosed magnetic flux increased (or decreased) by an integer multiple of the flux quantum have significant DDME. What we are able to show is that the DDME calculated according to the generalized CRE theory agrees with that predicted by the Laughlin wave functions. Moreover, by summing over all probability amplitudes of tunneling paths connecting different configurations, we can even calculate the off-diagonal density-matrix elements (ODME) and hence obtain phase differences between eigenwave functions evaluated for these configurations.

Our approach is based on an approximate evaluation of the density matrix

$$\rho(\{z_i'\},\{z_i\};\tau_0)\equiv {}_{\rm AS}\langle\{z_i'\}\,|\,e^{-\tau_0\mathcal{H}}\,|\,\{z_i\}\rangle_{\rm AS},$$

where $|\,\rangle_{\rm AS}$ denotes a properly antisymmetrized state, $\{z_i\}$ and $\{z_i'\}$ are the initial and final electron configurations, and τ_0 is the cooperative tunneling time. The hope is to demonstrate that for $\nu=1/(2n+1)$ (unless otherwise mentioned we shall always assume the "magic" filling factors through the paper),

$$\rho(\{z_i'\},\{z_i\};\tau_0)=\sum_\alpha e^{-\tau_0 E_\alpha}\Phi_\alpha(\{z_i'\})\Phi_\alpha^*(\{z_i\})$$
$$\simeq e^{-\tau_0 E_0}\left[\Psi_0(\{z_i'\})\Psi_0^*(\{z_i\})+\sum_{m=1}^{\infty}\prod_{j=1}^{m}\sum_{R_j\bar{R}_j}e^{-2m\Delta\tau_0}\Psi(\{z_i'\};\{R_j,\bar{R}_j\})\Psi^*(\{z_i\};\{R_j,\bar{R}_j\})\right], \tag{1}$$

where E_α and $\Phi_\alpha(\{z_i\})$ are exact eigenenergies and eigenwave functions, respectively. $\Psi_0(\{z_i\})$ is the Laughlin ground-state wave function, Δ is the average of the quasiparticle and quasihole gaps, $\Psi(\{z_i\};\{R_j,\bar{R}_j\})$ are the Laughlin quasiparticle-quasihole wave functions, where the quasiparticles are localized at $\{R_j\}$ and the quasiholes are localized at $\{\bar{R}_j\}$. Here we have made an approximation, viz., that at temperature τ_0^{-1}, the dominant contributions to the density matrix come from the ground state and states composed of a gas of noninteracting Laughlin quasiparticles and quasiholes.

To calculate the left-hand side of (1) we write down its path-integral representation,[4]

$$\rho(\{z_i'\},\{z_i\};\tau_0)=\frac{1}{\mathcal{N}}\sum_P(-1)^P\int_{\substack{z_i(0)=z_i,\\ z_i(\tau_0)=z'_{P_i}}}D[z(\tau)]\exp\left\{-\int_0^{\tau_0}\mathcal{L}(z(\tau))d\tau\right\}, \tag{2}$$

where $\mathcal{N}$ is the normalization constant, P is a permutation, and

$$\mathcal{L}(z(\tau))\equiv\tfrac{1}{4}\sum_i\left(\frac{dz_i^*}{dt}z_i-\frac{dz_i}{dt}z_i^*\right)+\tfrac{1}{2}\sum_{i\neq j}V(z_i-z_j).$$

To make the computation tractable, we make the semiclassical approximation.[4] Terms in (2) have the final electron configuration differing from the initial one by some permutation P. The saddle-point paths under such a boundary condition represent cooperative ring exchanges. Because we are restricted to a time interval of order τ_0, the CRE tunneling time, we do not have to worry about the temporal repetitions of the ring exchanges. This simplifies the calculation considerably, because now we need only sum over all possible spatial CRE's which occur in a time interval τ_0. Since the CRE processes that involve electrons tunneling through a distance larger (smaller) than $a(\nu)$ have longer (shorter) tunneling time, our calculations are most accurate for those configurations where the distribution function of the nearest-neighbor distance is sharply peaked at $a(\nu)$.

For the DDME $\{z_i\}=\{z_i'\}$, we represent each CRE as a directed closed-loop diagram. The graph rules are as follows: (1) Draw the lattice (in general nonperiodic) defined by the initial (=final) electron configuration $\{z_i\}$; (2) draw all possible single-bond directed closed loops; (3) associate a weight $\exp(-\alpha_{ij})$ to every bond $\langle ij\rangle$ on the loop; and (4) associate a phase factor $e^{\pm i\theta}$ to every loop. Here α_{ij} is the effective tunneling barrier between sites i and j, $\theta=2\pi(\phi/\phi_0)+\pi S$ (the first term is the Aharonov-Bohm phase factor and the second term comes from the Fermi-Dirac statistics), and the sign is determined by the sense of the loop. Also ϕ is the magnetic flux enclosed by the loop, $\phi_0=he/c$ is the flux quantum, and S is the number of bonds in the loop. (We have followed Ref. 4 and postulated the existence of a factor of -1 from the fluctuation determinant of every loop.[7]) We identify the graph rules described above with those of the loop expansion of a two-dimensional statistical-mechanical model defined as follows:

$$\rho(\{z_i\},\{z_i\};\tau_0)\equiv Z_{XY}(\{z_i\})e^{-F_0(\{z_i\},\{z_i\})}, \tag{3}$$
$$Z_{XY}=\int\prod_i\frac{d\theta_i}{2\pi}\prod_{ij}[1+2e^{-\alpha_{ij}}\cos(\theta_i-\theta_j+\Delta_{ij})],$$

where $F_0(\{z_i\},\{z_i\})$ is the contribution from processes other than the CRE (basically, this is the electrostatic

energy of the configuration $\{z_i\}$), θ_i is an angle variable defined at z_i, and $\Delta_{ij} \equiv (2\pi/\phi_0)\int_j^i \mathbf{A}\cdot d\mathbf{l} + \pi$, where $\mathbf{A}$ is the vector potential and the integral is performed along the saddle-point path. For a range of values for the parameter α_{ij}, the model defined in (3) exhibits a line of fixed points and is in the same universality class as the fully frustrated XY models.[8] Since tunneling between nonnearest-neighbor sites is suppressed by a large α_{ij}, we restrict ourselves to nearest-neighbor interactions only. By performing standard duality transformations[9] one can show that

$$Z_{XY}(\{z_i\}) \propto \sum_{\{m_R\}} \delta\left[\sum_R (m_R + f_R)\right] \exp\left[-\mu \sum_R (m_R+f_R)^2 - K_{\text{eff}} \sum_{R\neq R'} (m_R+f_R)\ln|R-R'|(m_{R'}+f_{R'})\right], \tag{4}$$

where the $\{R\}$ span the dual lattice of $\{z_i\}$, the $\{m_R\}$ are integers, and f_R for the plaquette P_R centered at R is defined as

$$f_R \equiv (2\pi)^{-1} \sum_{\langle ij\rangle \in P_R} \Delta_{ij}.$$

Also, μ (the chemical potential of m_R's) and K_{eff} (the long-wavelength spin-stiffness constant) are both functions of $\{\alpha_{ij}\}$. Since the effect of m_R is to shift f_R, and hence the magnetic flux, of P_R by a quantized unit, we associate the sum over $\{m_R\}$ with the sum over different quasiparticle and quasihole configurations in (1). This identification is valid so long as there is a well developed gap in the spectrum, so that only the ground state and few-quasiparticle states contribute to ρ. In the Coulomb gas, this situation pertains for K_{eff} greater than the critical value for the Kosterlitz-Thouless transition. For K_{eff} less than the critical value, the sudden proliferation of free vortices signals the transition to a compressible (gapless) state. As the result, we find that apart from some short-range terms [terms such as $F_0(\{z_i\},\{z_i\})$, whose value falls off faster than $\ln|z_i - z_j|$ as the particles are brought apart],

$$\begin{aligned} \ln|\Psi_0(\{z_i\})|^2 &\propto -\mu\sum_R (f_R - m_R^0)^2 + K_{\text{eff}} \sum_{R\neq R'} (f_R - m_R^0)\ln|R-R'|(f_{R'} - m_{R'}^0), \\ \ln|\Psi(\{z_i\};\{R_\alpha,\bar{R}_\alpha\})|^2 &\propto -\mu\sum_R (f_R + m_R - m_R^0)^2 + K_{\text{eff}} \sum_{R\neq R'} (f_R + m_R - m_R^0)\ln|R-R'|(f_{R'} + m_{R'} - m_{R'}^0). \end{aligned} \tag{5}$$

Here $m_R^0 = \frac{1}{2}S_R(1+\nu^{-1}) - \nu^{-1}$, with S_R being the number of sides of the plaquette centered at R; $m_R = -1$ if $R = R_\alpha$ and $m_R = 1$ if $R = \bar{R}_\alpha$. Since $f_R = \phi_R/\phi_0 + \frac{1}{2}S_R$ (ϕ_R is the magnetic flux piercing through the plaquette P_R), $f_R - m_R^0 = \nu^{-1}[\nu\phi_R/\phi_0 - (\frac{1}{2}S_R - 1)]$. Because $\nu\phi_R/\phi_0$ is the total neutralizing-background charge contained in the plaquette P_R, and $\frac{1}{2}S_R - 1$ is the average number of electrons enclosed in P_R, we recognize that (5) is nothing but the Laughlin wave functions

$$\begin{aligned} \ln|\Psi_0(\{z_i\})|^2 &\propto -\nu\int d^2r\, d^2r'\, \delta\rho(r)\ln|r-r'|\delta\rho(r'), \\ \ln|\Psi(\{z_i\};\{R_\alpha,\bar{R}_\alpha\})|^2 &\propto -\nu\int d^2r\, d^2r' [\delta\rho(r) + \rho_{\text{qp}}(r)]\ln|r-r'|[\delta\rho'(r') + \rho_{\text{qp}}(r')], \end{aligned} \tag{6}$$

with their short-wavelength (shorter than the mean interparticle spacing) plasma charge-density fluctuations $\delta\rho(r)$ integrated over. [In (6), $\rho_{\text{qp}}(r)$ is the plasma charge density produced by quasiparticles at $\{R_\alpha\}$ and quasiholes at $\{\bar{R}_\alpha\}$.] As the result of this integration, the plasma temperature is renormalized, and the renormalized value $(K_{\text{eff}})^{-1}$ is, in general, different from the corresponding "bare" value ν^{-1} given by the Laughlin wave functions.[10]

For the ODDM, the initial and final electron coordinates in (1) are no longer the same. For simplicity, let us assume $z_1 \neq z_1'$ and $z_i = z_i'$ for all $i \neq 1$. The graph rules for summing up the CRE's are the same as for the DDME, except that (a) the lattice is now defined by $\{z_1', z_1, z_2, \ldots, z_N\}$, (b) the allowed graphs now consist of open paths from z_1 to z_1' and closed loops everywhere else, and (c) with the open path we associate a phase factor $[\exp(i2\pi/\phi_0)\int \mathbf{A}\cdot d\mathbf{l} + i\pi S]$, where S is the number of bonds (excluding the bond between z_1 and z_1'). These new rules are precisely the ones used to calculate the two-point correlation functions for the model defined by (3). Therefore,

$$\begin{aligned} \rho(\{z_i'\},\{z_i\};\tau_0) &\equiv Z'_{XY}(z_1', z_1, z_2, \ldots, z_N) e^{-F_0(\{z_i'\},\{z_i\})}, \\ Z'_{XY}(z_1', z_1, z_2, \ldots, z_N) &= \int \frac{d\theta_{1'}}{2\pi} \prod_{i=1}^{N} \frac{d\theta_i}{2\pi} e^{i(\theta_1 - \theta_{1'})} \prod_{ij} [1 + 2e^{-\alpha_{ij}} \cos(\theta_i - \theta_j - \Delta_{ij})], \end{aligned} \tag{7}$$

where the product is over $i, j = 1', 1, \ldots, N$ and Δ_{ij} is defined as before except that $\Delta_{1'1}$ is defined without the additional π from the Fermi statistics. Following standard duality transformation, we can show that apart from some short-range

terms

$$Z'_{XY} \propto \sum_{\{m_R\}} \delta\left(\sum_R m_R + f_R\right) \exp\left[-\mu \sum_R (m_R+f_R)^2 - K_{\text{eff}} \sum_{R\neq R'} (m_R+f_R)\ln|R-R'|(m_{R'}+f_{R'})\right] \times \exp\left[-i\sum_R (m_R + \tfrac{1}{2}S_R)[\Theta(z'_1 - z_R) - \Theta(z_1 - z_R)]\right], \quad (8)$$

where S_R is the number of sides of the plaquette centered around R, $z_R \equiv R_x + iR_y$, and

$$\Theta(z-z') \equiv \tan^{-1}[(y-y')/(x-x')]$$

is the polar angle of $z-z'$. Following the analogy with (1) and (3), we again interpret the sum over $\{m_R\}$ in (8) to be the sum over different quasiparticle and quasihole configurations so that,

$$\begin{aligned} &\mathrm{Im}\{\ln[\Psi_0(z'_1)\Psi_0^*(z_1)]\} = -i\sum_R (\tfrac{1}{2}S_R - m_R^0)[\Theta(z'_1 - z_R) - \Theta(z_1 - z_R)], \\ &\mathrm{Im}(\ln[\Psi(z'_1;\{R_a,\bar{R}_a\})\Psi^*(z_1;\{R_a,\bar{R}_a\})]) = -i\textstyle\sum_R (m_R - m_R^0 + \frac{1}{2}S_R)[\Theta(z'_1 - z_R) - \Theta(z_1 - z_R)]. \end{aligned} \quad (9)$$

Since $m_R^0 - \frac{1}{2}S_R = \nu^{-1}(\frac{1}{2}S_R - 1)$, and $\frac{1}{2}S_R - 1$ is the average number of electrons per elementary plaquette, (9) implies that the phase difference between $\Psi_0(z'_1, z_2, \ldots, z_N)$ and $\Psi_0(z_1, z_2, \ldots, z_N)$ is equal to the phase change of the wave function of a charged particle that moves from z_1 to z'_1 in the magnetic field produced by solenoids located at $\{z_i; i=2, \ldots, N\}$. The flux going through each solenoid is ν^{-1} flux quanta. Similarly, the change in phase of the excited-state wave functions is the same as that of a charged particle that moves in the magnetic field produced by solenoids located at $z_2, \ldots, z_N$ and $R_a, \bar{R}_a$. These features are again consistent with the properties of the Laughlin wave functions. A word of warning should be given here. Whereas for some fixed configurations, e.g., the Wigner crystal, the model defined in (3) exhibits quasi long-range order in spin-spin correlation, the reduced one-body density matrix does not. This is because to calculate the reduced density matrix we have to calculate the spin-spin correlation function upon integration over the coordinates of the $N-1$ particles. Because of the phase oscillation, the "annealing" destroys the quasi long-range order.

In summary, we have shown that the long-distance behaviors of the DDME and ODME are consequences of interference effects. This result therefore serves as an independent test for the validity of any proposed wave functions for the fractional quantum Hall liquid.

We thank Dr. C. Kallin for useful discussions and acknowledge the hospitality of the Brookhaven National Laboratory where part of this work was done. One of us (S.K.) is supported in part by National Science Foundation Grant No. DMR18051 and an Alfred P. Sloan fellowship. Another of us (G.B.) is supported in part by National Science Foundation Grant No. DMR 8020236.

[1]R. B. Laughlin, Phys. Rev. Lett. **50**, 1395 (1983).

[2]D. C. Tsui, H. L. Störmer, and A. C. Gossard, Phys. Rev. Lett. **48**, 1599 (1982).

[3]S. Trugman and S. Kivelson, Phys. Rev. B **31**, 5280 (1985); F. D. M. Haldane, in *The Quantum Hall Effect,* edited by R. E. Prange and S. M. Gurvin (Springer-Verlag, New York, 1986), Chap. 8.

[4]S. Kivelson, C. Kallin, D. P. Arovas, and J. R. Schrieffer, Phys. Rev. Lett. **56**, 873 (1986).

[5]G. Baskaran, Phys. Rev. Lett. **56**, 2716 (1986).

[6]D. H. Lee, unpublished.

[7]Note that with the -1 per loop from the fluctuation determinant, it is not clear whether our theory predicts Laughlin-type wave functions for the boson analog of the fractional quantum Hall effect.

[8]B. Nienhuis, Phys. Rev. Lett. **49**, 1062 (1982); S. Teitel and C. Jayaprakash, Phys. Rev. B **27**, 598 (1983).

[9]J. V. Jose, L. P. Kadanoff, S. Kirkpatrick, and D. R. Nelson, Phys. Rev. B **16**, 1217 (1977).

[10]Note that because of the analyticity of the wave functions, the magnetic charge (defined as the integer multiple of 2π through which the phases of the Laughlin wave functions change as one electron is brought around another) cannot renormalize; however, the electric charge (defined as the power-law exponent which governs the decay of the modulus of wave functions as two electrons are taken apart) does, upon coarse graining.