# Cooperative-Ring-Exchange Theory of the Fractional Quantized Hall Effect

Steven Kivelson,[a] C. Kallin, Daniel P. Arovas,[b] and J. R. Schrieffer
*Institute for Theoretical Physics, University of California, Santa Barbara, California 93106*
(Received 3 September 1985)

A semiclassical path-integral approach is used to calculate the contribution of large-correlated-ring exchanges to the energy of a two-dimensional Wigner crystal in a strong magnetic field. This correlation energy $E_c(\nu)$ shows cusps at fractional fillings $\nu_c = n/m$ of the lowest Landau level. The uniform Wigner crystal is locally unstable for $\nu \neq \nu_c$ and the theory predicts the existence of fractionally charged quasiparticles to accommodate the extra density $\nu - \nu_c$.

PACS numbers: 71.45.Gm, 73.40.Lq

Recent experiments on the low-temperature, large-magnetic-field ($B_0$) conductivity of high-mobility electron layers[1] provide evidence of the existence of a family of novel condensed phases of the two-dimensional (2D) electron gas at "special" rational values of the dimensionless density $\nu$, where $\nu$ is the mean number of electrons in the area $2\pi l_0^2 = \phi_0/B_0$ covered by one flux quantum $\phi_0 = hc/e$. Specifically, (1) the anomalously small value of $\sigma_{xx}$ suggests[2] that there is a gap in the spectrum of current-carrying states, and hence a cusp in the correlation energy $E_c(\nu)$ at the special values of $\nu$; (2) the fractional quantization of $\sigma_{xy}$ may be evidence[3] of the existence of fractionally charged quasiparticles. We have studied this system in the absence of impurities, in the high-magnetic-field limit where the spacing $\hbar\omega_c = e\hbar B_0/m^*c$ between Landau levels as well as the Zeeman energy $\mu_B^* g B_0$ are much larger than the Coulomb repulsion between electrons $\sim V_0 = e^2/\epsilon l_0$, where $\epsilon$ is the dielectric constant. Thus, for $\nu < 1$, we assume that only spin-up electrons in the lowest Landau level (LLL) need be considered. Since the noninteracting system is highly degenerate, correlation effects are clearly very important.

Attempts to describe this system as a Wigner crystal (WC), which is the correct ground state for low densities, have been hampered by the fact that such descriptions give no evidence of cusps in the energy.[4,5] Maki and Zotos[5] used an *Ansatz* WC wave function, including Gaussian quantum fluctuations, and found densities $\nu = 1/m$ ($m$ odd) favored, but only extremely weakly as a result of the small overlap between neighboring sites. Laughlin[3] proposed a Jastrow-type trial wave function which innovatively treated pairwise electronic correlations. He found fractionally charged excitations ($q = e/m$) above the $\nu = 1/m$ ($m$ odd) ground state. At intermediate densities, Laughlin's wave function describes a liquid state with lower energy than that of the WC. Numerical studies on small numbers of electrons give support for such a state.[6] Recently, Chui, Hakim, and Ma[7] proposed a solidlike trial wave function, which they find to have a lower energy than that of Laughlin.

We have studied this problem, also beginning from a WC state, but using a systematic, semiclassical approximation.[8,9] Our results may be summarized as follows. We find exchange effects in which $L$ electrons in a ring coherently rotate to an equivalent configuration leading to contributions to $E_c(\nu)$ which can be orders of magnitude larger than pair-exchange contributions because of the reduced tunneling barrier. In addition, these contributions exhibit nonanalytic cusplike behavior for certain rational values of $\nu$.[10] Rings with large $L$ play a dominant role for two reasons. Firstly, although the contribution from any single ring decreases exponentially with $L$, there are a very large number of rings with large $L$ ($\sim K^L$ where $K$ is the connectivity). Secondly, they can make a nonanalytic contribution to $E_c(\nu)$ as a result of interference between different exchange rings. The contribution from each ring contains a phase factor $\theta = 2\pi \times B_0 A(\nu)/\phi_0$ (Bohm-Aharonov effect), where $A(\nu)$ is the enclosed area. For arbitrary $\nu$, the contributions from large rings add incoherently. However, because $A(\nu)$ is always approximately equal to an integer multiple of the area per elementary plaquette of the WC, for certain rational densities $\nu_c$ the different rings add in phase. It is this effect which makes these densities energetically favorable and which leads to cusps in $E_c(\nu)$ at $\nu = \nu_c$ when arbitrarily large rings are included.

Our model derives from a LLL path-integral representation for the partition function $Z = \mathrm{Tr} e^{-\beta H_N}$, with

$$H_N = \sum_{l=1}^{N} \frac{1}{2m^*}\left(\mathbf{p}_l + \frac{e}{2c} B_0 \hat{\mathbf{z}} \times \mathbf{r}_l\right)^2 + \sum_{j<k} V_2(\mathbf{r}_j - \mathbf{r}_k).$$

The single-particle Hamiltonian admits a continuous representation for LLL eigenstates, $\phi_{\mathbf{R}}(\mathbf{r}) = \langle \mathbf{r}|\mathbf{R}\rangle$:

$$\langle \mathbf{r}|\mathbf{R}\rangle = (2\pi)^{-1/2} \exp\{-\tfrac{1}{4}(\mathbf{r}-\mathbf{R})^2 + \tfrac{1}{2} i (\mathbf{r}\times\mathbf{R})\cdot\hat{\mathbf{z}}\}$$

with $H_1|\mathbf{R}\rangle = \frac{1}{2}\hbar\omega_c|\mathbf{R}\rangle$ and where we set $l_0 \equiv 1$. The resolution of the LLL projection operator, $P_0 = (1/2\pi)\int d^2R\, |\mathbf{R}\rangle\langle\mathbf{R}|$, may be used to develop a

path-integral expression[11] for $Z$:

$$Z(\nu)=\frac{1}{N!}\sum_{P\in S_N}(-1)^P\int d^{2N}r\,\langle\{\mathbf{r}_i\}|e^{-\beta H_N}|\{\mathbf{r}_{P(i)}\}\rangle,$$

where $N=\nu B_0/\phi_0$ and $|\{\mathbf{r}_i\}\rangle\equiv|\mathbf{r}_1\ldots\mathbf{r}_N\rangle$. This yields

$$Z(\nu)=\mathcal{N}\sum_{P\in S_N}(-1)^P\int\prod_{j=1}^{N}\mathcal{D}\,\mathbf{R}_j(\tau)e^{-S[R(\tau)]}, \tag{1}$$

where $\mathcal{N}$ is a normalization constant and the boundary conditions require $\mathbf{R}_j(0)=\mathbf{R}_{P(j)}(\beta)$. The action for continuous paths is[11,12]

$$S[R]=\tfrac{1}{2}\int_0^\beta d\tau\Big\{-i\sum_{j=1}^{N}(\dot{\mathbf{R}}_j\times\mathbf{R}_j)\cdot\hat{\mathbf{z}}+\sum_{j\neq k}V(\mathbf{R}_j-\mathbf{R}_k)\Big\}, \tag{2}$$

where $V$ is the matrix element of the Coulomb potential between coherent states, $V(\mathbf{R})=\frac{1}{2}\sqrt{\pi}(e^2/\epsilon)\times\exp(-\frac{1}{8}R^2)I_0(\frac{1}{8}R^2)$. We will refer to the integration variable $\tau$ as the (imaginary) time.

The partition function $Z$ is evaluated within the semiclassical approximation. This entails the finding of all paths $R^c(\tau)$ which extremize the action [$R^c$ is a vector function with $2N$ components $\mathbf{R}_j(\tau)$] and then the inclusion of quantum fluctuations by expansion of the action to second order in $R-R^c$. In this way $Z$ can be expressed as a sum over classical paths,

$$Z=\sum_c D[R^c]e^{-S[R^c]}, \tag{3}$$

where $D[R^c]$ is the fluctuation determinant. The extremal (classical) paths satisfy the equations of motion

$$i\dot{X}_j=\partial V_j/\partial Y_j,\quad i\dot{Y}_j=-\partial V_j/\partial X_j, \tag{4}$$

where $V_j\equiv\sum_{i\neq j}V(\mathbf{R}_i-\mathbf{R}_j)$. These are simply the imaginary-time $\mathbf{E}\times\mathbf{B}$ drift equations. To find solutions to Eq. (4) which satisfy the boundary conditions $\mathbf{R}_j(0)=\mathbf{R}_{P(j)}(\beta)$, we analytically continue the path integral to complex values of $X_j$ and $Y_j$.[12]

The path with the smallest action is the triangular WC. This path with pairwise exchange and Gaussian (phonon) fluctuations about it makes the leading-order contribution to $Z$, $Z_0\equiv D_0e^{-S_0}=\exp[-\beta\times NE_0(\nu)]$, where $E_0$ is the energy per site of the static WC as computed by Maki and Zotos.[5] $E_0(\nu)$ is a smooth, monotonic function of $\nu$ for $\nu<\frac{1}{2}$. Since the shear modulus is negative for $\nu>\nu_+\simeq0.45$,[5] we restrict our analysis to $\nu<\nu_+$. (For $\nu>1-\nu_+$, the same considerations apply to the hole lattice.) So as not to have prohibitively large action, the important classical paths must resemble the WC at most points in space and time. Therefore, we consider only classical paths whose initial configuration is the static WC. In addition, as discussed earlier, we focus on those paths which are most likely to lead to structure in $E_c(\nu)$ because of their systematic dependence on $\nu$; that is, we consider paths, such as those illustrated in Fig. 1, which consist of a cyclic, coherent superposition of nearest-neighbor exchanges.

For convenience, we factor out the leading-order contribution that is common to all classical paths by writing Eq. (3) as

$$Z=Z_0\sum_c\tilde{Z}_c=Z_0\sum_c\tilde{D}[R_c]e^{-\tilde{S}[R_c]}, \tag{5}$$

where $\tilde{D}=D/D_0$ and $\tilde{S}=S-S_0$. Let us consider the contribution of a single large exchange ring to $Z$. The real part of the classical action is approximately proportional to the number of electrons in the ring $L$ and the imaginary part is $\theta=\pm2\pi(\phi/\phi_0)+\pi(L-1)$, where $\phi$ is the enclosed flux, the $\pm$ refers to positive or negative sense of rotation, and the term $\pi(L-1)$ reflects the Fermi statistics. If the classical path exactly followed the straight-line segments between sites (as shown in Fig. 1) then $\theta=\pm\pi(\nu^{-1}-1)N_A+\pi(\mathrm{mod}2\pi)$, where $N_A$ is the number of enclosed plaquettes of area $\pi l_0^2/\nu$, and we have used the fact that $N_A$ is even (odd) when $L$ is even (odd). Similarly, the real part of the action would be $\alpha_0(\nu)L$, where $\alpha_0$ is independent of path.[13] We have argued that the net effect of deviations from such linear paths is to renormalize $\alpha_0$.[8] The fluctuation determinant is $\tilde{D}[R_c]=-\tau_0^{-1}d\tau\exp[-\Delta\alpha L+O(\ln L)]$, where $\tau_0$ is the cooperative tunneling time.[14] Therefore, we approximate the contribution of a large-ring exchange as

$$\Delta\tilde{Z}_c(\nu)=\tau_0^{-1}d\tau\exp[-\alpha(\nu)L+ihN_A+O(\ln L)],$$

where $\alpha=\alpha_0+\Delta\alpha$ and $h=\pi(\nu^{-1}-1)$.

Whether or not large-ring exchanges contribute significantly to $Z(\nu)$ is determined by the numerical value of $\alpha(\nu)$. We have estimated $\alpha(\nu)$ for the simple case in which a single line of electrons exchange,

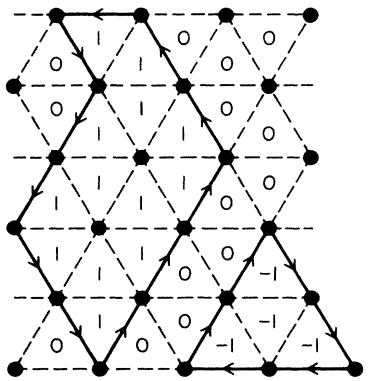


FIG. 1. Examples of exchange paths and their representation in terms of a configuration of the dual lattice.

so that $X_j(\beta)=X_j(0)+a_\nu$, $Y_j(\beta)=Y_j(0)$, where $a_\nu=(4\pi/\sqrt{3}\nu)^{1/2}$ is the lattice constant of the WC. We assume that only the electrons in this one line move in the background of the static potential of all other electrons. Thus we *overestimate* $\alpha$; however, we believe that the relaxation corrections are small. We find that the extremal path corresponds to rigid motion of the line [$\mathbf{R}_j(\tau)-\mathbf{R}_j(0)=\mathbf{R}(\tau)$ for $j$ on the line] and that $\alpha(\frac{1}{3})\sim 0.81$ (as a function of the density, $\alpha$ is approximately proportional to $1/\nu$). As we shall see, this indicates that large-ring exchanges are important at densities of experimental interest.

A general path in Eq. (5) contains many exchange rings and for small $\alpha$, they are sufficiently dense that the exchange events do not form a dilute gas. Therefore, we include what we believe to be the most important interactions between exchange events that overlap in space and time. The time interval $\beta$ is divided into slices of width $\tau_0$ and, because the classical paths are exponentially localized in time, we ignore interactions between exchange events which occur in different time slices; i.e., $\tilde{Z}_c\simeq[Z_{\rm slice}]^{\beta/\tau_0}$, where $Z_{\rm slice}$ is the trace over all exchanges in a given time slice. The exchange paths in a given time slice are enumerated in terms of integer-valued spin variables $S_\lambda$; $S_\lambda$ is defined to be the number of clockwise minus the number of counterclockwise exchange rings that encircle the plaquette $\lambda$. Hence, $\lambda$ labels a site on the dual (honeycomb) lattice. We associate with each spin configuration an energy

$$H_{\rm DG}=\alpha\sum_{\langle\lambda,\gamma\rangle}(S_\lambda-S_\gamma)^2+ih\sum_\lambda S_\lambda, \tag{6}$$

where $\langle\lambda,\gamma\rangle$ denotes nearest-neighbor sites. Then we make the approximation $Z_{\rm slice}\simeq{\rm Tr}\exp(-H_{\rm DG})$ which is exact for all configurations of isolated rings and includes a repulsion between rings that share one or more nearest-neighbor bonds. Equation (6) is the Hamiltonian of the discrete Gaussian (DG) model in an imaginary field, where $\alpha^{-1}$ plays the role of temperature. This model is known to have a phase transition at a critical value of $\alpha=\alpha_c(h)$.[15] For $h=2\pi m$ [$\nu=1/(2m+1)$], $\alpha_c$ takes on its maximum value, which we estimate to be $\alpha_c\simeq 1.1$. For $\alpha(\nu)>\alpha_c$, the system behaves like a classical WC, while for $\alpha(\nu)<\alpha_c$, the system is highly quantum mechanical and arbitrarily large exchange rings dominate the behavior of the system.

To study the $\alpha<\alpha_c$ phase it is useful to exploit the exact equivalence of the DG model and the Coulomb gas (CG)[15]:

$$H_{\rm CG}=\frac{2\pi^2}{\alpha}\sum_{\lambda\gamma}\left(q_\lambda-\frac{h}{2\pi}\right)G_{\lambda\gamma}\left(q_\gamma-\frac{h}{2\pi}\right),$$

where $G_{\lambda\gamma}\sim\ln|R_\lambda-R_\gamma|$ is the (honeycomb) lattice Green's function and $q_\lambda$ is an integer charge. The small-$\alpha$ phase can be analyzed by a study of the ground-state properties of the CG. The $h=2\pi m$ ground state of $H_{\rm CG}$ has $q_\lambda=m$ and zero energy. The ground state for $|h-2\pi m|\equiv\delta h\ll 2\pi$ has a fraction $\delta h/2\pi$ of sites with charge $1-\delta h/2\pi$, which themselves form a Wigner lattice. The remaining sites have charge $-\delta h/2\pi$. Thus the free energy $F_{\rm CG}$ of the CG at small $\alpha$ is proportional to $|\delta h|\ln|2\pi/\delta h|$, which for our problem implies $E_c(\nu)\sim|\delta\nu|\ln|\delta\nu|$ for $|\delta\nu|\ll 1/(2m+1)$. Since $\partial E_c/\partial\nu$ diverges as $\delta\nu\to 0$, the uniform WC state is thermodynamically unstable in the open neighborhood of $\nu=1/(2m+1)$. This motivates the need for quasiparticles discussed below.

From studies of Josephson junction arrays in a transverse field,[16] $F_{\rm CG}$ is believed to have cusps of the form $|\delta h|\ln|2\pi/\delta h|$ at all rational $h/2\pi$ for $\alpha(\nu)<\alpha_c(h(\nu))$.[17] From our estimate of $\alpha(\nu)$ and from Monte Carlo calculations of Shih and Stroud,[16] we find that this inequality is most likely to be satisfied by $\nu=\frac{1}{3},\frac{1}{5},\frac{2}{5},\frac{2}{7},\frac{1}{4},\frac{3}{7}$, and $\frac{4}{9}$, although some of these phases may be unstable with respect to competing phases.[18] Since the $\nu=\frac{1}{3}$ uniform-density state is much more stable than any other special density, one might also consider a hierarchy of states formed by starting with a WC of quasiparticles and repeating our analysis for ring exchange of quasiparticles with fractional charge and fractional statistics.[19,20] In this case, one obtains a sequence of stable densities in agreement with previous hierarchical analyses.[20,21] We will discuss the relative stability of these states elsewhere.[8]

Our model is thus incompressible since it is rigid with respect to uniform dilations. However, there are quasiparticle (qp) excitations of charge $Q^*$ which correspond to local dilations of the Wigner crystal by an amount $\delta A$. As a result of the deformation of the lattice, all rings that enclose the qp acquire an extra phase $\Delta\theta=2\pi B_0\delta A/\phi_0$ relative to those that do not. One can see immediately from the CG representation that this will cost a logarithmically divergent energy unless $\Delta\theta=2\pi\times$integer, and hence the elementary excitations have charge $Q^*=\pm\nu e$. We have constructed an approximate qp with a core size $\sim l_0$ and estimate its creation energy $E_{\rm qp}(\nu)\sim 0.5\nu^2e^2/\epsilon l_0$. When $\nu$ is near a favored density $\nu_c$, the system can reduce its energy by forming a nonuniform state with density (averaged over a plaquette) equal to $\nu_c$ everywhere except in the core of the qp. The cusps in the energy are thus $E(\nu)\sim\nu_c^{-1}|\nu-\nu_c|E_{\rm qp}(\nu_c)$ for $\nu$ near $\nu_c$. (We have not evaluated the creation energies with sufficient accuracy to distinguish between quasiparticle and quasihole energies.)

Of great interest is the magnetophonon (mp) spectrum. In the sparse-ring phase, the mp dispersion resembles the standard $\omega_k\sim k^{3/2}$ form.[5] In the dense-ring phase, however, nonlinear interactions between density fluctuations strongly alter the mp

spectrum.[8] This effect and the issue of dissipation are presently under consideration.

We conclude with some brief comments on the consistency of our approach. In the small-$\alpha$ phase, the gas of exchange loops is quite dense. Hence, the time-slice path decomposition is not well defined. However, the major role of the time-slice approximation is only to provide an ultraviolet cutoff corresponding to a repulsion between overlapping exchange loops. One particular feature of the dense-ring phase is the proliferation of intersecting rings in the dual spin model. Some of these configurations correspond to high-action paths in real space (e.g., crossings), and should be discouraged by the inclusion of additional short-ranged spin-spin interactions.[22] We expect that such terms should lead to a renormalization of $\alpha$, and that they will not change the universality class of the spin model.

An interesting open question concerns the existence of long-ranged charge-density-wave (CDW) order. At finite temperature, the magnetophonons will destroy any long-ranged order. At zero temperature, none of the terms that we have computed explicitly destroy the CDW. However, we have yet to establish fully whether or not such order is actually present in this limit, and the relation between our theory and that of Laughlin remains unclear. Nevertheless, it seems likely that our results do not depend essentially on the answer to this question, since the vanishing compressibility at rational $\nu$ derives from the coherent addition of many large exchange loops, an effect that could persist in the absence of CDW order.

We are grateful to A. J. Berlinsky, V. J. Emery, F. D. M. Haldane, T. C. Halsey, D. Stroud, and E. Tosatti for useful conversations. This work was supported in part by the National Science Foundation through Grants No. DMR 82-16285 and No. DMR 83-18051. One of us (D.P.A.) acknowledges support by an AT&T Bell Laboratories Scholarship, the work of another of us (C.K.) is partially supported by a Canadian Natural Sciences and Engineering Research Council Postdoctoral Fellowship, and one of us (S.K.) is the recipient of an Alfred P. Sloan Fellowship.

[a]Permanent address: Department of Physics, State University of New York at Stony Brook, Stony Brook, N.Y. 11794.

[b]Permanent address: Department of Physics, University of California, Santa Barbara, Cal. 93106.

[1]See, for example, D. C. Tsui, H. L. Stormer, and A. C. Gossard, Phys. Rev. Lett. **48**, 1559 (1982); H. L. Stormer, A. M. Chang, D. C. Tsui, J. C. M. Hwang, A. C. Gossard, and W. Wiegmann, Phys. Rev. Lett. **50**, 1953 (1983).

[2]R. B. Laughlin, Phys. Rev. B **23**, 5632 (1981); B. I. Halperin, Helv. Phys. Acta **56**, 75 (1983).

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

[4]D. Yoshioka and P. A. Lee, Phys. Rev. B **27**, 4986 (1983), and references therein.

[5]K. Maki and X. Zotos, Phys. Rev. B **28**, 4349 (1983).

[6]F. D. M. Haldane and E. H. Rezayi, Phys. Rev. Lett. **54**, 237 (1985); D. Yoshioka, B. I. Halperin, and P. A. Lee, Phys. Rev. Lett. **50**, 1219 (1983).

[7]S. T. Chui, T. M. Hakim, and K. B. Ma, to be published; see also S. T. Chui, Phys. Rev. B **32**, 1436 (1985).

[8]S. Kivelson, C. Kallin, D. P. Arovas and J. R. Schrieffer, to be published.

[9]This approximation is justified by the fact that the quantum parameter is small for $\nu < \frac{1}{2}$. See Ref. 8

[10]The possibility that large-ring exchanges might lead to a phase transition was discussed in a quite different context by R. P. Feynman, Phys. Rev. **91**, 1291 (1953). See also S. Chakravarty and D. B. Stein, Phys. Rev. Lett. **49**, 582 (1982).

[11]See, for example, L. Schulman, *Techniques and Applications of Path Integration* (Wiley, New York, 1981), Chap. 27.

[12]When discontinuous paths are important, one must use the discrete version of Eq. (1). See Ref. 11 and J. R. Klauder, in *Path Integrals and Their Applications in Quantum Statistical and Solid State Physics,* edited by G. Papadopoulos and J. T. Devreese (Plenum, New York, 1977), p. 5.

[13]The effective interaction between particles on the exchange path falls off sufficiently rapidly that $\alpha_0$ is independent of path. See Ref. 8.

[14]The fluctuation determinant has been evaluated approximately. It contains a factor $-1$ per ring-exchange event. The $\tau_0^{-1}$ term arises due to the zero mode associated with time translation.

[15]S. T. Chui and J. D. Weeks, Phys. Rev. B **14**, 4978 (1976).

[16]W. Y. Shih and D. Stroud, Phys. Rev. B **32**, 158 (1985), and references therein.

[17]T. C. Halsey, unpublished.

[18]Note that the plateau at $\nu = \frac{1}{4}$ does not appear in the usual hierarchy scheme. In this context, we note that a plateau in $\sigma_{xy}$ at $\nu = \frac{3}{4}$ as well as a broad minimum in $\sigma_{xx}$ at $\nu = \frac{1}{2}$ have been reported by G. Ebert *et al.*, J. Phys. C **17**, L775 (1984).

[19]D. P. Arovas, J. R. Schrieffer, and F. Wilczek, Phys. Rev. Lett. **53**, 722 (1984).

[20]B. I. Halperin, Phys. Rev. Lett. **52**, 1583, 2390(E) (1984).

[21]F. D. M. Haldane, Phys. Rev. Lett. **51**, 605 (1983); R. B. Laughlin, Surf. Sci. **142**, 163 (1984).

[22]At the level of the spin model, several unphysical paths have already been discarded or suppressed. See Ref. 8.