nature physics

LETTERS

https://doi.org/10.1038/s41567-022-01529-8

Discrete symmetry breaking defines the Mott quartic fixed point

Edwin W. Huang[1,3], Gabriele La Nave[2,3] and Philip W. Phillips[1,3] ✉

Because Fermi liquids are inherently non-interacting states of matter, all electronic levels below the chemical potential are doubly occupied. Consequently, the simplest way of breaking the Fermi-liquid theory is to engineer a model in which some of those states are singly occupied, keeping time-reversal invariance intact. We show that breaking an overlooked[1] local-in-momentum space $\mathbb{Z}_2$ symmetry of a Fermi liquid does precisely this. As a result, although the Mott transition from a Fermi liquid is correctly believed to arise without breaking any continuous symmetry, a discrete symmetry is broken. This symmetry breaking serves as an organizing principle for Mott physics whether it arises from the tractable Hatsugai–Kohmoto model or the intractable Hubbard model. Through a renormalization-group analysis, we establish that both are controlled by the same fixed point. An experimental manifestation of this fixed point is the onset of particle–hole asymmetry, a widely observed[2–10] phenomenon in strongly correlated systems. Theoretically, the singly occupied region of the spectrum gives rise to a surface of zeros of the single-particle Green function, denoted as the Luttinger surface. Using *K*-homology, we show that the Bott topological invariant guarantees the stability of this surface to local perturbations. Our proof demonstrates that the strongly coupled fixed point only corresponds to those Luttinger surfaces with co-dimension $p+1$ with odd p. We conclude that both Hubbard and Hatsugai–Kohmoto models lie in the same high-temperature universality class and are controlled by a quartic fixed point with broken $\mathbb{Z}_2$ symmetry.

Symmetry is a fundamental organizing principle of nature. A case in point is the simplest example of symmetry, namely, permutations. This symmetry helps organize identical fundamental particles into two groups: fermions (odd under interchange) and bosons (even under permutation). Since the permutation group has a finite number of elements, namely, ±1, it is an example of discrete symmetry. Here we show that a group as simple as the permutation group, namely, $\mathbb{Z}_2$, controls the transition from a non-interacting collection of electrons constituting a Fermi surface to a state that strongly violates the traditional theory of metals, namely, the Mott paramagnetic state that insulates, although the band is half full. The Fermi surface retains the $\mathbb{Z}_2$ symmetry but the Mott state does not.

A manifestation of this symmetry breaking is the resultant asymmetry on particle–hole addition or removal, that is, on doping. In a non-interacting electron system, adding or subtracting an electron is a symmetrical process. However, cuprate superconductors as varied as underdoped $Bi_2Sr_2CaCu_2O_{8+\delta}$ (Bi-2212) and $Ca_{2-x}Na_xCuO_2Cl_2$ (Na-CCOC) exhibit scanning tunnelling spectra[9–12] with a distinct asymmetry in terms of particle addition and removal. The cuprates are not alone here as there are numerous electronic systems[6–8] that exhibit particle–hole asymmetry at low energies on the addition or removal of an electron. Although it is now commonplace to attribute particle–hole asymmetry to strong correlations[2,3,5], no universal operative principle has been enunciated, except for the general phenomenon of Mottness[4]. In his parting words in 2016, P. W. Anderson[13] reproached condensed-matter theorists for not facing up to this problem: 'I remain baffled by the almost universal refusal of theorists to confront this evident fact of hole–particle asymmetry head on'. We investigate this in this paper. What all cuprates have in common is that the parent material cannot be understood without considering the interactions. The minimal model thought to be relevant in this context is due to Hubbard in which electrons move on a square lattice but pay an energy cost whenever opposite-spin electrons reside on the same site. Since this model is unsolvable in any dimension other than $d=1$, it is difficult to pinpoint a clear organizing principle, other than that the interactions are important, as the root cause of the asymmetry. An added complication is that the Mott insulating state that arises from the local interactions is thought to be featureless above any temperature associated with ordering, just as is the Fermi liquid, the non-interacting limit. Consequently, appealing to some sort of symmetry breaking appears to be a non-starter.

We propose here that such an organizing principle can be unearthed by focusing on the full symmetry group of a Fermi liquid and analysing which symmetries in the Fermi liquid survive the transition to the paramagnetic Mott insulator[14]. Although it is common to use the Hubbard model to study this transition, our key point here is that the essence of the Mott transition is captured by a simpler model that breaks the fundamental local-in-momentum space $\mathbb{Z}_2$ symmetry (non-local in real space) of the Fermi-liquid state. This $\mathbb{Z}_2$ symmetry breaking serves as an organizing principle for Mott physics. We find that both local on-site Hubbard and local-in-momentum (as in the exactly solvable Hatsugai–Kohmoto (HK) model[15–17]) interactions fall into the same universality class as they both break $\mathbb{Z}_2$ symmetry. We then use *K*-theory to show that the surface of zeros that characterizes the Mott phase is stable to perturbations, thereby establishing the existence of a fixed point. Our work here is analogous to that of another study[18] on the stability of a Fermi surface.

Relevance of the HK interaction

Part of the motivation for this work is that there seem to be two disparate ways of generating a Mott transition with no apparent relationship between them. These constitute the HK[15] and Hubbard models. Although both models contain the standard kinetic term, the HK model contains a non-standard local-in-momentum space interaction, namely,

$$H_{\rm int}^{\rm HK} = U\sum_k n_{k\uparrow}n_{k\downarrow}, \tag{1}$$

[1]Department of Physics and Institute for Condensed Matter Theory, University of Illinois, Urbana, IL, USA. [2]Department of Mathematics, University of Illinois, Urbana, IL, USA. [3]These authors contributed equally: Edwin W. Huang, Gabriele La Nave, Philip W. Phillips. ✉e-mail: dimer@illinois.edu

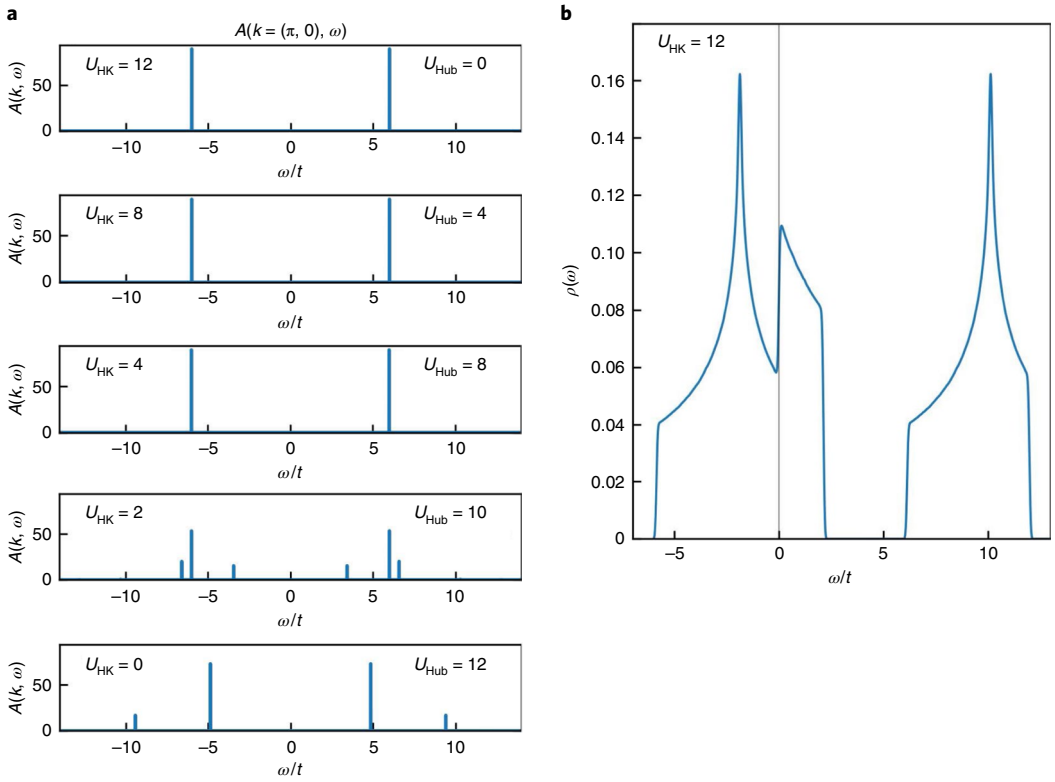

Fig. 1 | Comparison of Hubbard and HK Models. a, Spectral function of HK and Hubbard models from exact diagonalization with parameters shown. At half-filling, there is little difference between the models, showing that the HK interaction accurately models the *U*-scale physics of Hubbard. **b**, Density of states of the HK model at filling $\langle n\rangle = 0.8$, showing a strong particle–hole asymmetry at the Fermi energy.

and the Hubbard model contains the standard real-space interaction, namely,

$$H_{\rm int}^{\rm Hubb} = U\sum_i n_{i\uparrow}n_{i\downarrow} = U\sum_{k,p,q} c^\dagger_{(k-q)\uparrow}c_{k\uparrow}c^\dagger_{(p+q)\downarrow}c_{p\downarrow}, \quad (2)$$

where Fourier transform shows the explicit non-local momentum structure. Here U is the strength of the repulsion and $c^\dagger_{p\sigma}$ creates a particle in a state with momentum p and spin σ and $c_{p\sigma}$ annihilates it. Even with the kinetic energy, the former model is solvable exactly[15,16], yielding an insulating state with a hard gap for $U > W$, where W is the bandwidth. The explicit energy cost for doubly occupying the same k-state is the explicit mechanism for the Mott physics in the HK model as it leads to singly occupied states below the chemical potential. Ultimately, the same must also be true for the Hubbard model, but here only numerics[14] support a gap indicative of Mott physics. Hence, it is worth comparing both models. Note that the $q=0$, $k=p$ part of the Hubbard interaction is of the same form as the HK interaction. As will become evident from our analysis, it is this term that is the leading relevant interaction that drives Mott physics. The spectral functions for both models are roughly identical (Fig. 1). Both describe a gapped state in which the spectral weight lies at high and low energies above the gap. It is the presence of such spectral weight at high and low energies that generates the surface of zeros of the real part of the single-particle Green function, denoted as the Luttinger surface[19,20]. The surface of zeros only appears at momenta that are singly occupied[15,16,20]. There is no difference between the models in the top three panels (Fig. 1a). Differences only emerge at high energies but with reduced spectral weight. Nonetheless, the capability of the HK model to capture Mott physics is not widely appreciated. Unearthing why these two fairly apparently different models yield the same physics is the primary goal of this paper.

The first thing that must be established with the HK model is why does the interaction $H_{\rm int}^{\rm HK}$ destroy the Fermi-liquid behaviour. Two distinct arguments will be adopted here. First, we appeal to the renormalization principle[21,22] for fermions and show that $H_{\rm int}^{\rm HK}$ is a relevant perturbation. The correct starting point for the renormalization of fermions is to demand that the kinetic term in the action

$$S_0 = \int dt d^d\mathbf{p}\psi^\dagger_\sigma(\mathbf{p})(i\partial_t - (\epsilon_{\mathbf{p}} - \epsilon_{\rm F}))\psi_\sigma(\mathbf{p}) \quad (3)$$

has a zero scaling dimension under the distortion $\mathbf{p} = \mathbf{k} + s\boldsymbol{\ell}$, where $\mathbf{k}$ is along the Fermi surface, $\boldsymbol{\ell}$ is perpendicular to it and s is the scaling parameter (which will be set to 0 to preserve the Fermi surface). Here $\psi^\sigma_\sigma(p)$ creates a particle with momentum σ and momentum p, t is time, ∂_t is the derivative with respect to time, d^d represents a d-dimensional integration, ϵ_p is the particle energy and ϵ_F the Fermi energy. Expanding the dispersion relationship of an electron around the Fermi surface, namely,

$$\epsilon(\mathbf{p}) = \epsilon_{\rm F} + \ell\frac{\partial\epsilon}{\partial\mathbf{p}} + O(\ell^2), \quad (4)$$

we find that demanding $[S_0] = 0$ requires that $\psi_\sigma(\mathbf{p}) \to s^{-1/2}\psi_\sigma(\mathbf{p})$, where the symbol in the last term, O, represents the order of magnitude. The irrelevance of a generic interaction, V, term

$$S_{\rm int} = \int dt\prod_{i=1}^4 d^{d-1}\mathbf{k}_i d\ell_i V(\mathbf{k}_1,\cdots,\mathbf{k}_4)\psi^\dagger_\sigma(\mathbf{p}_1)\psi_\sigma(\mathbf{p}_3)\psi^\dagger_{\sigma'}(\mathbf{p}_2)\psi_{\sigma'}(\mathbf{p}_4)\delta^d(\mathbf{p}_1+\mathbf{p}_2-\mathbf{p}_3-\mathbf{p}_4) \quad (5)$$

follows because $[S_{\rm int}]=1$ (arising from s^4 due to the four factors of $d\ell$, s^{-1} from $d\tau$ and $s^{-4/2}$ from the four fermion fields) and hence vanishes in the $s\to 0$ limit. This conclusion holds even if loop corrections with L loops are included as they scale as s^L, thereby vanishing for a generic interaction. Note an interaction of the form $Un_{k\sigma}n_{k'\sigma}$ has a tree-level scaling dimension of -1 and hence contributes the same order as the chemical potential. That is, it leads to the mass renormalization of Fermi-liquid theory. Loop corrections of this term once again contribute s^L and hence generate no self-energy corrections. Within this scheme, the only exception arises when electrons scatter with momenta on opposite sides of the Fermi surface. In this case, the δ function factorizes and the interaction is marginal and leads to an instability for $V<0$. Note that the form of the kinetic-energy term is irrelevant to this argument. All that is necessary is the expansion in equation (4).

In contrast, the local-in-momentum space interaction

$$\tilde{S}_{\rm int}=U\int {\rm d}t{\rm d}\ell{\rm d}^{d-1}\mathbf{k}\psi^{\dagger}_{\uparrow}(\mathbf{k})\psi_{\uparrow}(\mathbf{k})\psi^{\dagger}_{\downarrow}(\mathbf{k})\psi_{\downarrow}(\mathbf{k}) \qquad (6)$$

differs from the generic interaction in $S_{\rm int}$ in that it contains only a single integration over momentum. This interaction can be derived from a non-local-in real space interaction that preserves the centre of mass of the interacting pairs of electrons[15]. Because of the single integration over momentum in equation (6), the tree-level scaling of $\tilde{S}_{\rm int}$ takes the form s^{-2} and the interaction term is, in fact, relevant even if the electrons do not lie on the Fermi surface. Once again, loop corrections are irrelevant to this term following the argument elsewhere[22,23]. The key conclusion here is that the interaction in the HK model provides a relevant deformation of the Fermi-liquid theory. No contradiction arises from the traditional renormalization principle for fermions[21–24] as the HK interaction arises from non-local real-space interactions.

$\mathbb{Z}_2$ symmetry breaking in Mott physics

The HK interaction provides the general mechanism for the breaking of long-range real-space entanglement of a Fermi liquid, thereby providing a proxy for the Hubbard model; therefore, we investigate a little-known observation in another study[1] regarding the full symmetry group of a Fermi liquid. Their key point is that because Fermi liquids possess separately conserved currents for up and down spins, the full symmetry group for each point on the Fermi surface is O(4), the real group of rotations in 4-space. The determinant of an O(4) matrix is either $+1$ or -1, thus exhibiting the disconnected nature of this Lie group. Namely, the proper group SO(4) where the determinant is $+1$ cannot be continuously deformed into those whose determinant is -1. To understand what remains, we consider the quotient O(4)/SO(4), which is isomorphic to $\mathbb{Z}_2$. $\mathbb{Z}_2$ arises simply because there are two connected components of O(4). That is, $\pi_0({\rm O}(4))\simeq\mathbb{Z}_2$ (here $\pi_p(G)$ is the group of homotopy classes of maps of the p-dimensional sphere to G), which is equal to the group consisting of the identity I and reflection R. A reflection R through a hyperplane is represented by $R_{ij}=\delta_{ij}-n_in_j$ if $(n_0,\ldots,n_3)$ is the orthonormal vector to the hyperplane. There is, of course, a fairly distinct $\mathbb{Z}_2$ that lurks because $\pi_1({\rm SO}(4))\simeq\mathbb{Z}_2$, which tells us that there is a simply connected double cover of SO(4) called Spin(4) (the spin group) that is isomorphic to ${\rm SU}(2)\times{\rm SU}(2)$, one SU(2) for spin and the other for charge pseudospin. This gives rise to an equivalence between the spin and charge degrees of freedom in a Fermi liquid. As a result, in terms of the particle–hole spinor, $\psi^{\dagger}_{\mathbf{p}}=(c^{\dagger}_{\mathbf{p}\uparrow},c_{-\mathbf{p}\downarrow})$, we can write the Hamiltonian for a Fermi liquid as

$$H_{\rm FL}=\sum_{\mathbf{p}}\psi^{\dagger}_{\mathbf{p}}(\epsilon_{\mathbf{p}}-\epsilon_{\rm F})\tau_3\psi_{\mathbf{p}}+\cdots \qquad (7)$$

which explains the inherent SU(2) invariance of the charge sector (as proposed initially by Anderson[25] and Nambu[26]) and the existence of an infinite number of conserved currents, $n_{\mathbf{p}\sigma}$. Here τ_3 is the z component of traditional Pauli matrices. The ellipses stand for any interaction term that renormalizes to zero or terms that contribute at the same level as the chemical potential that lead to mass renormalization of the Fermi liquid[21–24]. The extra $\mathbb{Z}_2$ symmetry is evident only for the electrons precisely at the Fermi surface. In fact, although electrons at the Fermi surface have an SU(2) symmetry, those away just have a U(1). As the kinetic energy vanishes for such electrons, extra symmetries emerge. The relevant symmetry that emerges within O(4) is that the sign of only one of the spin currents can be changed without any consequence to the underlying theory. That is, at the Fermi surface, a particle–hole transformation on one species $c_{p\uparrow}\to c^{\dagger}_{p\uparrow}$ or $n_{p\uparrow}\to 1-n_{p\uparrow}$ but preserving $n_{p\downarrow}\to n_{p\downarrow}$ can be made with impunity. The remaining electrons do not enjoy this symmetry. In this sense, the $\mathbb{Z}_2$ symmetry is emergent in a Fermi liquid as it is exact only at the Fermi surface. In the presence of generic short-range interactions, the precise manifestation of this symmetry is detailed in the Supplementary Information. It is this discrete $\mathbb{Z}_2$ symmetry that a Fermi surface possesses that ultimately accounts for the inherent particle–hole symmetry at low energies. Once this symmetry is lost, the symmetry between particle and addition around the chemical potential is also lost.

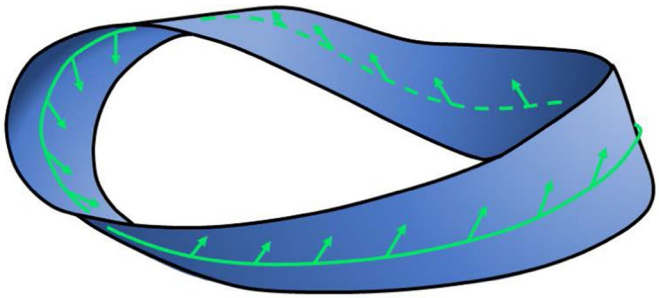

Fig. 2 | Möbius bundles. Parallel transport of the normal vector in the Möbius bundle along the yellow curve (Fermi surface) reveals that going around once leads to a reflection. A double traversal brings back the original vector, thus revealing the underlying $\mathbb{Z}_2$ symmetry. Note that this symmetry is only maintained at the point from which the parallel transport was initiated.

There is a subtlety here that points to more than O(4) defining the group structure of Fermi liquids. To establish this, we note that from the $\mathbb{Z}_2$ symmetry (of order $\frac{1}{N}$; Supplementary Information), we can view the O(4) action as giving an O(4)-bundle structure to the fermions on the Fermi surface. The $\mathbb{Z}_2$ symmetry in this context is related to orientability (a consistent orthonormal frame that remains invariant on parallel transport through a loop (Fig. 2)) of the bundle. The first step[27] is to realize that a Fermi liquid augmented by a number of trivial bands (explained below) has the same properties as the original system. We consider general free Hamiltonians as

$$H=\sum_{\sigma,\sigma'}\psi^{\dagger}_{\sigma}(\mathbf{p})A_{\sigma\sigma'}(\mathbf{p})\psi_{\sigma'}(\mathbf{p}). \qquad (8)$$

We can think of $A_{\sigma\sigma'}$ as a map from the Fermi surface (since we are interested in the $\mathbb{Z}_2$ symmetry described thus far) to a matrix group and we impose two such A values, namely, A_1 and A_2, to be equivalent when

$$A_1\sim A_2 \text{ if } A_1\oplus A_{\rm trivial}\sim_{\rm hom}A_2\oplus A_{\rm trivial}, \qquad (9)$$

where $A_{\rm trivial}$ represents the trivial system

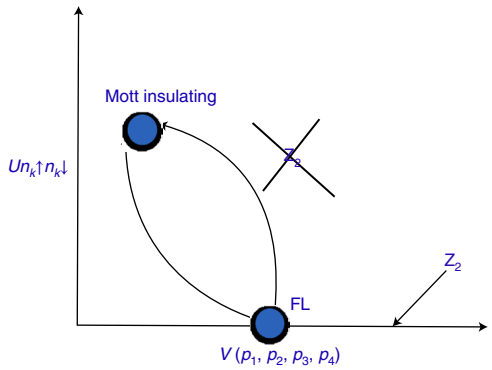

Fig. 3 | Hubbard–HK comparison. Flow diagram if the interaction in the HK model is made to have zero scaling dimension. The general 4-momentum interaction has a scaling dimension of 3 and hence is irrelevant. FL denotes the Fermi liquid that arises strictly along the x axis, which preserves a discrete momentum space $\mathbb{Z}_2$ symmetry. Any non-zero value of $\tilde{S}_{\rm int}$ destroys the Fermi liquid and consequently the $\mathbb{Z}_2$ symmetry. This scaling analysis is also supported by the exact solution of the HK model. Note that the Mott insulating behaviour persists even when the generic 4-fermion interaction $V_{\rm int}$ is turned on, implying that the Mott insulator generated from $\tilde{S}_{\rm int}$ is a stable fixed point.

$$A_{\rm trivial} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} |p|^2, \tag{10}$$

and $\sim_{\rm hom}$ means homotopically equivalent. The homotopy equivalence is reflected in being able to continuously deform the eigenvalues without changing the determinant. This equivalence class gives rise to the set of maps from the Fermi surface (which we assume to be homotopic to a sphere) to a classifying space C_q or R_q (complex or real, respectively). The only classifying spaces for which $\pi_0(G) \simeq \mathbb{Z}_2$ corresponds to either $G = {\rm O}(n)$ or $G = {\rm O}(2n)/{\rm U}(n)$ (Supplementary Information). This means that the additional group O(4)/U(2), which describes spin-polarized electrons, for example, is a possible candidate to describe Fermi liquids. However, such a group would not allow a description in terms of $H_{\rm FL}$. The types of Fermi liquid described by O(4)/U(2) is beyond the scope of this paper.

From the analysis above, it is clear that any interaction of the form $n_{p\uparrow}n_{p\downarrow}$ (the interaction in $\tilde{S}_{\rm int}$) maximally breaks the momentum space $\mathbb{Z}_2$ symmetry ($n_{p\uparrow} \to (1-n_{p\uparrow}), n_{p\downarrow} \to n_{p\downarrow}$) of a Fermi surface as it transforms to $(1-n_{p\uparrow})n_{p\downarrow}$ and hence the 2-body term changes sign. Since this term is a relevant perturbation of a Fermi liquid, it is not a surprise that it breaks the $\mathbb{Z}_2$ invariance of the would-be Fermi surface. Guided by the $\mathbb{Z}_2$ symmetry and the principle of relevance, we can analyse the Hubbard interaction as well. Explicitly, equation (2) tells us that we can organize the Fourier transform as $S_{\rm int}^{\rm Hubb} = S_3 + S_2 + S_1(\tilde{S}_{\rm int}/N)$, where S_n has n independent momenta. S_1 corresponds to the $q = 0, p = k$ term, for example, which is just $\tilde{S}_{\rm int}/N$ (N is the system size) and hence has a scaling dimension of -2. As each integration over momentum carries with it a power of the scaling parameter s, S_3 and S_2 are subdominant, that is, $[S_3] = 0$ and $[S_2] = -1$, contributing the same order as the chemical potential relative to S_1 and the HK term (which has a scaling dimension of -2). Because s cannot vanish faster than $1/N$, the HK term and S_1 have identical scaling and hence both are relevant contributions to the Hubbard model. As shown by the similarity of the gaps in Fig. 1, we infer that both models are in the same high-temperature universality class. This does not mean that extra physics cannot be encoded in the Hubbard model. As in a Fermi liquid, the Landau interaction parameters can modify the susceptibilities and density of states and even make the spin and charge sectors differ in two dimensions[1]. The key point is that even in the presence of such interactions, the excitations are still governed by the full O(4) symmetry of a Fermi liquid. Likewise, in both HK and Hubbard models, the reduced symmetry as a result of the breaking of $\mathbb{Z}_2$ governs the nature of excitations and not the form of density of states. That is, the breaking of $\mathbb{Z}_2$ symmetry by $\tilde{S}_{\rm int}$ creates a new quartic fixed point (Fig. 3). The presence of a charge gap but gapless spin degrees of freedom in the half-filled state are manifestations of the breaking of the discrete $\mathbb{Z}_2$ symmetry as the spin and charge currents can no longer be freely rotated. In the doped state, it is well known[2–5] that the density of states of a doped Mott insulator (Fig. 3) lacks particle–hole symmetry as must be the case if $\mathbb{Z}_2$ symmetry in momentum space is absent. Consequently, both Mott insulating (gapped charge but gapless-charge degrees of freedom) and doped (absence of particle–hole symmetry) systems are affected by the breaking of $\mathbb{Z}_2$ symmetry. We see then that $\mathbb{Z}_2$ symmetry is a powerful organizing principle of strongly correlated Mott physics. This conclusion lends credence to the perturbative result that the gap in Weyl Mott insulator metals with HK interactions is not affected to the second order once Hubbard interactions are introduced[28]. As a result, we conclude that generically, the HK term controls the flow of a Fermi liquid to the paramagnetic Mott insulating state (Fig. 3). Hence, both HK and Hubbard interactions break the $\mathbb{Z}_2$ symmetry of a Fermi surface; as a result, the transition from a Fermi liquid to a Mott insulator involves the breaking of a discrete $\mathbb{Z}_2$ symmetry.

In the spirit of naturalness, it makes sense to scale towards the interactions and not the Fermi surface and hence away from the phase that preserves the $\mathbb{Z}_2$ symmetry. This causes a major conceptual leap as we can no longer rely on a Fermi surface. In the presence of strong interactions, it is sensible to instead choose the surface of zeros of the single-particle Green function, that is, the Luttinger surface—the locus of points in momentum space along which the single-particle Green function vanishes[20]. Such a surface demarcates the paramagnetic Mott gap in a single-band system as numerical simulations explain the two-dimensional Hubbard model[29,30]. As shown previously[31,32], such a surface has nothing to do with the particle density unlike the Fermi surface but rather sets the conditions for spectral weight transfer on the Mott scale[20]. Because such spectral weight transfer[4,14] is the defining feature of the paramagnetic Mott insulator, we are interested in a stability analysis of this surface in terms of which perturbations destroy it. As with the Fermi surface, the Luttinger surface can only be destroyed by perturbations in the perpendicular direction. For instance, in the HK model, the analogue of the expansion in equation (4) for the Luttinger surface is

$$\varepsilon(k) = \frac{U}{2} + \varepsilon(k_L) + \ell \frac{\partial \varepsilon}{\partial \mathbf{p}} + O(\ell^2), \tag{11}$$

where, once again, the only degree of freedom is a distortion perpendicular to the surface. Here k_L is the position of the surface of zeros. Hence, for the location of the Luttinger surface, U is fixed. Scaling towards the Luttinger surface is now given by equation (11). If no relevant interactions are found, then this will effectively define a strongly coupled fixed point. The requirement of $[\tilde{S}_{\rm int}] = 0$ fixes the scaling dimension of the fermion field to be $[\psi_\sigma(p)] = 0$ for the HK interaction. Under this scaling scheme, $[S_0] = 1$, implying that the kinetic term is perturbatively irrelevant in the Mott insulating state. This reinforces the naturalness of our scheme. Recall that the metallic phase exists only for $W > U$ and hence cannot be perturbatively reached from the strongly coupled fixed point. Likewise, the generic four-fermion interaction in equation (5) is also irrelevant as it scales as s^3, which vanishes when $s \to 0$. Hence, even Hubbard physics (a general 4-momenta term) cannot flow away from the

Mott insulating point. Consequently, we argue (Fig. 1a) that $\tilde{S}_{\rm int}$ constitutes a natural fixed point for Mott physics. Further, we can assess the role of pairing by considering the term

$$H_p = \frac{1}{L^d}\Delta^\dagger\Delta, \tag{12}$$

where $\Delta = \sum_k b_k = \sum_k c_{-k\downarrow}c_{k\uparrow}$. Since this term has two momenta, the action for this term, S_p, has a scaling dimension of either 2 (HK interaction) or 1 (for the generalized HK interaction), also implying that the Mott insulator is not perturbatively destroyed by pairing, consistent with our recent analysis[16] that shows that superconductivity can be obtained only in the metallic phase.

HK as a fixed point

In this section, we explain how the renormalization-group flow works for non-local theories and propose a K-theory stability analysis for the underlying fixed point. The notion of renormalizabilty is, in general, ill-posed as normally stated, as one generally neglects to mention the space of operators within which a theory is renormalizable. More explicitly, consider a certain theory described by a classically local action $S(\phi_i)$ of some (not necessarily scalar) fields $\phi_1,\ldots,\phi_n$. One fixes a certain energy scale Λ and integrates fields whose energy is higher than Λ so that the effective action S_Λ can be obtained. This is done by integrating the fields whose frequency $\omega>\Lambda$, thus splitting the fields into high ($\phi_{\rm H}$) and low ($\phi_{\rm L}$) frequencies $\phi=\phi_{\rm L}+\phi_{\rm H}$ and then integrating

$$\int D\phi e^{iS(\phi)} = \int D\phi_{\rm L} e^{iS_\Lambda(\phi_{\rm L})}, \tag{13}$$

where $S_\Lambda(\phi_{\rm L}) = -i\log\left(\int D\phi_{\rm H} e^{iS(\phi_{\rm L},\phi_{\rm H})}\right)$. If S_* is a fixed point, one can write

$$S_\Lambda = S_* + \int d^dx \sum_i g_i\mathcal{O}_i \tag{14}$$

for some local operators $\mathcal{O}_i$ (they are local, despite the integration of high-frequency fields, because we focus on fields with $\omega<\Lambda$). The core of renormalization is in the observation that there is a dimension (of operators) $D=D(d,H_0)$ (where H_0 is the Hamiltonian of free energy), above which the operators are irrelevant, and the number of local operators $\mathcal{O}_i$ whose dimension is less than (or equal) to D is finite (this is because classically, local operators are polynomials in the fields ϕ and their derivative $\frac{\partial^I\phi}{\partial x^I}$, having used the multi-index notation). Since there are finitely many of these, one can make sense of such theories. The point we want to make is that this makes sense only because we restrict ourselves to a class of operators allowed (in this case, the classically local operators $\mathcal{O}_i$). But this argument can be generalized to a non-local theory in real space whose Fourier transform is, of course, local. Hence, non-locality in real space poses no real hurdle to the renormalization program. Here locality in momentum space is the standard notion of locality in which position is replaced by momentum.

For the sake of simplicity, we explain this procedure for the case in which the Hamiltonian is $H=H_0+H_1$, where H_0 arises from a classically local operator (such as kinetic energy) and H_1 is non-local in real space, but its Fourier transform is local as in the example of $\tilde{S}_{\rm int}$ in the HK model. We now simply allow operators $\mathcal{O}_i$ whose Fourier transform is local and can be written as a combination of fundamental operators; that is, operators that are either classically local in position space or that are polynomials in the operator components of H_1. Since the degree of these polynomials has to be bounded for the dimension of the operators to be bounded, there are only finitely many of this latter type as well.

Stability of the Mott fixed point is tantamount to showing that the defining feature of Mott physics[20]—the Luttinger surface (defined in equation (11)) although not necessarily related to the particle density[31,32]—is stable to perturbations, for example, the non-HK terms in the Hubbard model. Recall the surface of zeros is possible as long as the spectral weight bifurcates (Fig. 1). To this end, we show that the Luttinger surface, which can be exactly established for the HK model[16], under perturbations of the Hamiltonian is determined by Bott periodicity[33,34] and ultimately K-theory, similar to a Fermi surface, as shown elsewhere[18]. For our purposes, the importance of the renormalization group, besides the existence of the fixed point, is that for small values of parameters g_i, the Green function continuously changes by applying perturbation theory to $Z(g_i) = \int D\phi_{\rm L} e^{iS_\Lambda(\phi_{\rm L})} = \int D\phi_{\rm L} e^{i\left(S_* + \int d^dx \sum_i g_i\mathcal{O}_i\right)}$. Consider the Green function

$$G(k,\omega) = \langle\psi(0,0)\psi^\dagger(k,\omega)\rangle = \frac{1}{\omega-\varepsilon(k)+\Sigma(k,\omega)} \tag{15}$$

for some Hamiltonian that vanishes along a surface of zeros—the Luttinger surface. In a d+1-dimensional (k,ω) space, we will regard the Luttinger surface Ω to have dimension $d-p$ and hence its co-dimension is $p+1$. Here Σ is the exact self-energy. The precise equation denoting the zero surface (equation (11)) is determined by the locus of $(k_L,\omega=0)$ points at which Σ diverges. We assume the fields $\psi(k,\omega)$ represent complex fermions consisting of N components. We consider a point k_L in momentum space that is an element of the Luttinger locus, that is, $\Omega:=\{\det G=0\}$ and consider a p-sphere of radius ϵ centred at point $k_\perp$ in the normal direction (that is, $k_\perp$ is in the normal bundle ν_Ω to Ω and we take a fibre of the ϵ-tubular neighbourhood of Ω identified via the exponential map with the ϵ-sphere bundle $S_\Omega(\epsilon)=\{(k_L,k_\perp)\in\nu_\Omega : |k_L-k_\perp|=\epsilon\}$). A perturbation that preserves the Luttinger surface moves the zero of G along $k_\perp$. If not, it moves it elsewhere in which case the Luttinger surface is destroyed. We investigate the topology to show that the latter cannot be obtained. At points in $S_\Omega(\epsilon)$, the complex $N\times N$ matrix G is non-degenerate since, by definition, the locus of points in the momentum space on which it is degenerate is Ω. Therefore, we obtain a continuous analytic map

$$\Omega_\epsilon : S_\Omega(\epsilon) \to {\rm GL}(N,\mathbb{C}). \tag{16}$$

Here ${\rm GL}(N,\mathbb{C})$ is the group of invertible complex matrices with $N\times N$ entries. Fixing a point $k_L\in\Omega$, the relevant set is then $S_\Omega(\epsilon)_{k_L} = \{k_\perp : |k_L-k_\perp|=\epsilon\}$ (which, in the language of fibre bundles, is the fibre of $S_\Omega(\epsilon)$ at Ω_L) and this set $S_\Omega(\epsilon)_{k_L}$ is an S^p sphere and the map Ω at fixed k_L is

$$\Omega'_\epsilon : S^p \to {\rm GL}(N,\mathbb{C}). \tag{17}$$

Any deformation $H_{\rm M}+gH_2$ of the Hamiltonian $H_{\rm M}$ (here we think of this as the Hamiltonian of the fixed point that exhibits a Mott insulating nature) will deform this map Ω continuously, thus preserving its homotopy class. Now, the main observation is that if the homotopy class of Ω'_ϵ in the pth fundamental group $\pi_p({\rm GL}(N,\mathbb{C}))$ is non-zero, the Luttinger surface Ω must be stable under small deformations. In fact, if the image of S^p via Ω'_ϵ was the trivial class, then the map would be homotopic (continuously deformable) to a constant map (that is, mapping the whole of S^p to a constant invertible matrix). But this would mean that the map Ω'_ϵ could be extended to a map from the solid ball $B_\epsilon(k_L)=\{(k_L,k_\perp)\in\nu_\Omega : |k_L-k_\perp|=\epsilon\}$ centred at k_L and of radius ϵ to ${\rm GL}(N,\mathbb{C})$. This is impossible because $\mathcal{G}$ is degenerate at k_L as per the definition of a Luttinger surface. As a result, stability follows. Higher fundamental groups are notoriously complicated to calculate; fortunately, for classical groups, via the use of Morse theory, Bott[33] was able to prove that they are periodic (and the period depends on the group)

$$\pi_k\left(\varinjlim \mathrm{GL}(N,\mathbb{C})\right) = \pi_{k+2}\left(\varinjlim \mathrm{GL}(N,\mathbb{C})\right); \tag{18}$$

particularly in the so-called stable regime or when N is sufficiently large compared with p ($N > \frac{p}{2}$ suffices),

$$\pi_k\left(\varinjlim \mathrm{GL}(N,\mathbb{C})\right) = \begin{cases} 0 & \text{if } k \text{ is even} \\ \mathbb{Z} & \text{if } k \text{ is odd} \end{cases}. \tag{19}$$

We have thus established the fact that Luttinger surfaces of co-dimension $p+1$ in momentum–energy space are stable for odd p and unstable for even p, similar to the case of the Fermi surface.

Our work shows that models exhibiting a Luttinger surface, that is, a surface of zeros, ultimately have a rigorous stability condition based in K-theory. It would be erroneous to associate the winding number $\pi_k(\varinjlim \mathrm{GL}(N,\mathbb{C}))$ with the charge density because for the Fermi surface, the winding number counts the multiplicity of poles; because each pole has a quasiparticle interpretation, the winding is equivalent to knowing the charge. For the zero surface[31,32], no quasiparticle interpretation of zeros is obtained. Hence, their multiplicity as indicated by the non-trivial winding number $\pi_k(\varinjlim \mathrm{GL}(N,\mathbb{C}))$ has no physical significance. This ultimately sheds light on why deviations from the Luttinger count[35] with the charge density have been so numerous[31,32]. The existence of our stability condition implies that the details of the underlying Hamiltonian are irrelevant. The only quantity of relevance is the Luttinger surface. Consequently, our analysis puts all the models with Luttinger surfaces under the same umbrella as they are controlled by a fixed point whose stability is ultimately controlled by K-theory and lack the $\mathbb{Z}_2$ symmetry of a Fermi surface. The superconducting transition found earlier[16] should then be a generic feature of this fixed point. It is from the breaking of the discrete $\mathbb{Z}_2$ symmetry that the particle–hole asymmetry (Fig. 3) naturally arises, thereby leading to a direct response to Anderson's reproach.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41567-022-01529-8.

Received: 23 June 2021; Accepted: 27 January 2022;
Published online: 21 March 2022

References

1. Anderson, P. W. & Haldane, F. D. M. The symmetries of fermion fluids at low dimensions. *J. Stat. Phys.* **103**, 425–428 (2001).
2. Chen, C. T. et al. Electronic states in $La_{2-x}Sr_xCuO_{4+\delta}$ probed by soft-X-ray absorption. *Phys. Rev. Lett.* **66**, 104 (1991).
3. Eskes, H., Meinders, M. B. J. & Sawatzky, G. A. Anomalous transfer of spectral weight in doped strongly correlated systems. *Phys. Rev. Lett.* **67**, 1035 (1991).
4. Phillips, P. Colloquium: identifying the propagating charge modes in doped Mott insulators. *Rev. Mod. Phys.* **82**, 1719 (2010).
5. Kohno, M. Mott transition in the two-dimensional Hubbard model. *Phys. Rev. Lett.* **108**, 076401 (2012).
6. Yu, Y. et al. High-temperature superconductivity in monolayer $Bi_2Sr_2CaCu_2O_{8+\delta}$. *Nature* **575**, 156–163 (2019).
7. Silva, J. B. et al. Particle-hole asymmetry in the two-impurity Kondo model. *Phys. Rev. Lett.* **76**, 275 (1996).
8. Ghosh, S. S., Quan, Y. & Pickett, W. E. Strong particle-hole asymmetry in a 200 kelvin superconductor. *Phys. Rev. B* **100**, 094521 (2019).
9. Randeria, M., Sensarma, R., Trivedi, N. & Zhang, F.-C. Particle-hole asymmetry in doped Mott insulators: implications for tunneling and photoemission spectroscopies. *Phys. Rev. Lett.* **95**, 137001 (2005).
10. Anderson, P. & Ong, N. Theory of asymmetric tunneling in the cuprate superconductors. *J. Phys. Chem. Solids* **67**, 1–5 (2006).
11. Hanaguri, T. et al. A 'checkerboard' electronic crystal state in lightly hole-doped $Ca_{2-x}Na_xCuO_2Cl_2$. *Nature* **430**, 1001–1005 (2004).
12. Miller, T. L., Zhang, W., Eisaki, H. & Lanzara, A. Particle-hole asymmetry in the cuprate pseudogap measured with time-resolved spectroscopy. *Phys. Rev. Lett.* **118**, 097001 (2017).
13. Anderson, P. W. Last words on the cuprates. Preprint at https://arxiv.org/abs/1612.03919 (2016).
14. Georges, A., Kotliar, G., Krauth, W. & Rozenberg, M. J. Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions. *Rev. Mod. Phys.* **68**, 13 (1996).
15. Hatsugai, Y. & Kohmoto, M. Exactly solvable model of correlated lattice electrons in any dimensions. *J. Phys. Soc. Jpn* **61**, 2056–2069 (1992).
16. Phillips, P. W., Yeo, L. & Huang, E. W. Exact theory for superconductivity in a doped Mott insulator. *Nat. Phys.* **16**, 1175–1180 (2020).
17. Baskaran, G. An exactly solvable fermion model: spinons, holons and a non-Fermi liquid phase. *Mod. Phys. Lett. B* **05**, 643–649 (1991).
18. Hořava, P. Stability of Fermi surfaces and *K theory*. *Phys. Rev. Lett.* **95**, 016405 (2005).
19. Setty, C. Superconductivity from Luttinger surfaces: emergent Sachdev-Ye-Kitaev physics with infinite-body interactions. *Phys. Rev. B* **103**, 014501 (2021).
20. Dzyaloshinskii, I. Some consequences of the Luttinger theorem: the Luttinger surfaces in non-Fermi liquids and Mott insulators. *Phys. Rev. B* **68**, 085113 (2003).
21. Polchinski, J. Effective field theory and the Fermi surface. Preprint at https://arxiv.org/abs/hep-th/9210046 (1992).
22. Weinberg, S. Preface to volume II. in *The Quantum Theory of Fields* Vol. 2 XVII–XIX (Cambridge Univ. Press, 1996).
23. Weinberg, S. Effective action and renormalization group flow of anisotropic superconductors. *Nucl. Phys. B* **413**, 567–578 (1994).
24. Shankar, R. Renormalization-group approach to interacting fermions. *Rev. Mod. Phys.* **66**, 129 (1994).
25. Anderson, P. W. Random-phase approximation in the theory of superconductivity. *Phys. Rev.* **112**, 1900 (1958).
26. Nambu, Y. Quasi-particles and gauge invariance in the theory of superconductivity. *Phys. Rev.* **117**, 648 (1960).
27. Kitaev, A. In *Advances in Theoretical Physics: Landau Memorial Conference* (eds. Lebedev, V. and Feigel'Man, M. **1134**, 22–30 (American Institute of Physics Conference Series, 2009).
28. Morimoto, T. & Nagaosa, N. Weyl Mott insulator. *Sci. Rep.* **6**, 19853 (2016).
29. Stanescu, T. D. & Kotliar, G. Fermi arcs and hidden zeros of the Green function in the pseudogap state. *Phys. Rev. B* **74**, 125110 (2006).
30. Sakai, S., Motome, Y. & Imada, M. Evolution of electronic structure of doped Mott insulators: reconstruction of poles and zeros of Green's function. *Phys. Rev. Lett.* **102**, 056404 (2009).
31. Rosch, A. Breakdown of Luttinger's theorem in two-orbital Mott insulators. *Eur. Phys. J. B* **59**, 495–502 (2007).
32. Dave, K. B., Phillips, P. W. & Kane, C. L. Absence of Luttinger's theorem due to zeros in the single-particle Green function. *Phys. Rev. Lett.* **110**, 090403 (2013).
33. Bott, R. An application of the Morse theory to the topology of lie-groups. *Bull. Soc. Math. Fr.* **84**, 251–281 (1956).
34. Atiyah, M., Bott, R. & Shapiro, A. Clifford modules. *Topology* **3**, 3–38 (1964).
35. Luttinger, J. M. Fermi surface and some simple equilibrium properties of a system of interacting fermions. *Phys. Rev.* **119**, 1153 (1960).

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgements

P.W.P. thanks D. Gross for a probing question, asked during a KITP online seminar, regarding the renormalization of non-local interactions that ultimately inspired this work, DMR-2111379 for partial funding of this project and M. Kaplan-Hartnett and Janice Benner for assistance with Fig. 2. E.W.H. was supported by the Gordon and Betty Moore Foundation EPiQS Initiative through the grants GBMF 4305 and GBMF 8691.

Author contributions

Edwin W. Huang (numerics and conception), Gabriele La Nave (conception), Philip W. Phillips (conception and writing of manuscript). Edwin W. Huang, Gabriele La Nave and Philip W. Phillips contributed equally.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41567-022-01529-8.

Correspondence and requests for materials should be addressed to Philip W. Phillips.

Peer review information *Nature Physics* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.