Theory of Raman Scattering in Mott-Hubbard Systems

B. Sriram Shastri and Boris I. Shraiman

AT&T Bell Laboratories, Murray Hill, New Jersey 07974

(Received 7 May 1990)

We present a theory of Raman scattering in the Hubbard model. The scattering of light has two contributions. One gives rise to scattering by spin degrees of freedom in the insulating case where the general form of the scattering Hamiltonian is derived. The fluctuations of the "chiral" spin operator \( \sum_{\mathbf{q}} (s_{\mathbf{q}} \times s_{\mathbf{q}+\mathbf{q}}) \) are shown to contribute in the B2g scattering geometry. The other contributes in the doped case and is shown to probe the fluctuations of the "stress tensor." This quantity is not conserved, and hence its fluctuations at small \( q \) inherent in optical experiments need not be small, in striking contrast to density fluctuations in usual metals.

PACS numbers: 71.28.+d, 78.30.—j

Raman scattering (inelastic light scattering) has provided extremely important qualitative and quantitative information in high-\( T_c \) materials, in both the insulating and doped cases. Here we will be concerned with the question of describing the nonphoton contribution to Raman scattering, within a unified theoretical framework. A simplifying feature of the Mott-Hubbard system is that a single effective band, or a few bands, are involved, with the result that all couplings to the external world are readily parametrized. For example, the external electromagnetic field couples to electrons through a well-defined phase factor in the hopping matrix element—the Peierls phase. This serves below as the starting point for a unified description of both the "magnetic Raman scattering" in the insulating phase and the "electronic Raman scattering" in the doped, i.e., metallic, phase of the Mott-Hubbard system. In the former case, we reexamine the derivation and regime of validity of the effective Hamiltonian for resonant light scattering.

\[
\hat{\mathcal{H}}_r = \sum_k \epsilon(k)c_\sigma^\dagger(k)c_\sigma(k) - \frac{\epsilon}{\hbar c} \sum_k j(k) \cdot \mathbf{A}(-k) + \frac{\epsilon^2}{2 \hbar^2 e^2} \sum_{k,k'} A_\sigma(-k) \tau_{\alpha\beta}(k+k') A_\beta(-k') - i \hbar \dot{\mathbf{r}} (k) \times \mathbf{A}(-k) + \frac{\hbar}{2 m^*} \mathbf{D}(\mathbf{k}) \cdot \dot{\mathbf{r}}(k).
\]

The operators \( c_\sigma(k) \) destroy electrons with spin \( \sigma \) and momentum \( k \) defined for the square lattice in 2D. The energy is \( \epsilon(k) = -2 t [\cos(k_x) + \cos(k_y)] \), the current operator is

\[
j_{\sigma}(q) = \sum_k v_{\sigma}(k) c_\sigma^\dagger(k+q/2) c_\sigma(k-q/2),
\]

where \( v_{\sigma}(k) = \partial \epsilon(k) / \partial k_{\sigma} \), and the stress tensor \( \tau_{\alpha\beta} \) is

\[
\tau_{\alpha\beta}(q) = \sum_k \frac{\partial \epsilon(k)}{\partial k_{\alpha}} c_\sigma^\dagger(k+q/2) c_\sigma(k-q/2).
\]

Finally, \( \mathbf{A}_\sigma(q) = e_{\sigma}\mathbf{g}(a_{q} + a_{-q}) \), with \( \mathbf{g}(q) = (\hbar c^2 / V \omega_q)^{1/2} \) and with \( a_q \) destroying photons with polarization \( e \), momentum \( q \), and energy \( \omega_q = c |q| \). From this point on we switch to units where \( \hbar = c = 1 \).

The Raman-scattering cross section is determined by the transition probability rate \( R \) obtained from the "golden rule":

\[
R(q, \omega) = 2 \pi \sum_{i,j} \exp(-\beta \epsilon_i) |g(k_i) g(k_f) e_{\sigma}\mathbf{D}(\mathbf{k}) |^2 \delta(\epsilon_f - \epsilon_i - \Omega),
\]

where \( \omega_i, k_i, \) and \( e_{\sigma}^{(i)} \) denote the energy, momentum, and polarization of the initial (final) states of the photon, \( \epsilon_i(\omega) \) refer to the Hubbard eigenstates describing the "matter," \( \Omega = \omega_i - \omega_f, \) and \( q = k_i - k_f \).

The scattering operator has both resonant and nonresonant contributions \( M(q) = M_R(q) + M_N(q) \):

\[
\langle f | M_R^{\beta}(q) | i \rangle = \langle f | \tau_{\alpha\beta}(q) | i \rangle
\]

and

\[
\langle f | M_R^{\beta}(q) | i \rangle = \sum_v \left[ \frac{\langle f | j_{\beta}(k_f) | v \rangle \langle v | j_{\alpha}(k_i) | i \rangle}{\epsilon_v - \epsilon_i - \omega_i} + \frac{\langle f | j_{\alpha}(k_f) | v \rangle \langle v | j_{\beta}(k_i) | i \rangle}{\epsilon_v - \epsilon_i + \omega_f} \right].
\]
We now proceed to examine the undoped insulating case corresponding to exactly one electron per site for large $U/t$. It is clear that the low-energy Raman scattering, at least when $\Omega \leq U$, is dominated by the resonant terms [Eq. (6)] since the stress tensor does not have low-energy matrix elements. The Hilbert space splits into bands labeled by $n$, the number of doubly occupied sites, which we denote by $|n, \alpha\rangle$. The lowest Hubbard band (LHB) $n=0$ can be parametrized by the set of $S^2(r)$ eigenvalues specifying a spin configuration $\alpha=\{|\alpha\rangle\}$, while the upper Hubbard band (UHB) with $n=1$ requires the specification of the doubly occupied site $r_d$ and the hole site $r_h$, in addition to the spins, so that $|1, \alpha\rangle=c_d^\dagger(r_d)c_h^\dagger(r_h)|\sigma\rangle$, with $|\sigma\rangle=\Pi_c c_c^\dagger|0\rangle$.

To lowest order in $t/U$, these are degenerate. The $n=0$ manifold degeneracy is lifted by the superexchange arising from "second-order" hopping "$t'"\" matrix elements giving the Heisenberg model $H_{\text{ex}}=(4t^2/U)\sum_{c,d}\{s_r^2 s_{r+\mu} - \frac{1}{2}\}$. In contrast, the degeneracy of the $n=1$ manifold (UHB) is lifted already at first order in $t$ by the kinetic energy of the hole and the "double." To leading order in $t/U$, only $n=1$ UHB states appear as intermediate states, and Eq. (6) (for $k_F \sim k_F \sim 0$) becomes

$$\langle \sigma | M_{\text{R}} | i \rangle \sim \sum_{a, r'} \langle 0, \{\sigma_f\} | j_{\mu}(r) | 1, \alpha | j_{\nu}(r') \rangle \langle 1, \alpha | \{\sigma_f\} | 0 \rangle,$$

$$\times \left\{ \frac{e^{\nu}_{a, r'}}{e_{1, a} - \omega_i} + \frac{e^{\nu}_{a, r'}}{e_{1, a} + \omega_i} \right\},$$

with the bond current operator $j_{\nu}(r)=it[c_{a,d}^\dagger(r+v) \times c(r) - \text{H.c.}]$. To the zeroth order in $t/U$, $e_{1, a} = U$, and hence can be pulled outside the intermediate-state summation. The sum can be carried out and the resulting answer for the matrix elements of $M_{\text{R}}$ written in terms of effective spin operators as in Fleury and Loudon,

$$H_{\text{LS}} = \sum_{r_d} \left( \frac{1}{U} s_{r_d} \cdot s_{r+\mu} \right) \left[ \frac{t^2}{U - \omega_i} \right] \{\sigma_f\} | \{\sigma_f\} \rangle.$$

The argument leading to Eq. (8) is the same as the one used in deriving the superexchange, with the modification that $j_{\mu} A_{\mu}$ is replaced by $H_1$, and the energy denominators do not contain $\omega_i$ [we drop the second term of Eq. (6) here and in the rest of the paper]. The latter is a crucial difference since the perturbative expansion relies on the smallness of $t/U - \omega_i$ which breaks down in the resonant regime $U \sim \omega_i$. More generally, the calculation of the scattering matrix element [Eq. (7)] involves the UHB propagator $G_{\xi}(r_{kd}, \{\sigma\}; r_{kd}, \{\sigma\}) = \langle \sigma | c_d^\dagger c_d | H_{\text{ex}} - E \rangle^{-1} | c_d^\dagger c_d | \sigma \rangle$.

The intermediate-state summation in Eq. (7) can be written in terms of the off-diagonal elements of $G_{\xi}$ where $E = U - \omega$, and since holes and doubles form extended states in the UHB, we expect $G_{\xi}$ to have a continuous spectrum with bandwidth of $-t$. However, because of the interaction of the carriers with the background spins, the proper calculation is an intricate Brinkman-Rice kind of a problem which will be considered elsewhere.

The lowest-order scattering Hamiltonian, as in Eq. (8), is clearly the leading term in a moment expansion $(H_{\text{ex}} - E)^{-1} = -E \sum_m \langle m | E^{-1} | H_{\text{ex}} \rangle^m$, while the general terms of the expansion contain both nonlocal and multispin exchanges of a rich variety. Here we present the next few terms in the moment expansion, which are obtained by combining Eqs. (7) and (9) and some tedious algebra.

Defining $\Delta_{\mu, \nu, \rho, \sigma}^\pm = \langle g(k_{\mu}) g(k_{\nu}) \rangle^{-1} [A_{\mu}^+, A_{\nu}^+ \pm A_{\rho}^+, A_{\sigma}^+]$, $\mathcal{P}_{\Delta} = \sum_{\sigma, \tau} \delta_{\sigma, \tau} + \sum_{\sigma, \tau} (s_{\rho} s_{\sigma} \pm s_{\tau} s_{\tau}) (s_{\rho} s_{\sigma} \pm s_{\tau} s_{\tau})$, and $\Delta = t^{-1} (U - \omega_i)$, we find that

$$\langle \sigma | M_{\text{R}} | i \rangle = \langle \sigma | (O_1 A_{\pm, y} y + O_2 A_{\pm, x} x, y + O_3 A_{\pm, y} x, y + O_4 A_{\pm, y, x} x) | \{\sigma\} \rangle,$$

with the operators

$$O_1 = t \Delta [N/4 - \frac{1}{2} (P_x + P_y)] + 2t \Delta^3 [N/4 + P_x + P_y + P_x + P_y + P_x + y + P_x + y + 2(Q_{x, y, x} + Q_{y, x, x} + Q_{x, y, y} - Q_{y, y, y})],$$

$$O_2 = t \Delta \{ \frac{1}{2} - 4 \Delta^2 \} (P_x - P_x) - 2t \Delta^3 [P_x - P_y],$$

$$O_3 = -t \Delta^3 [N/4 + 4P_x + 4P_y + 4P_x - y],$$

and

$$O_4 = 8t \Delta^3 \sum_{r} e_{\mu, \nu, r} (s_{r+\mu} \times s_{r+\nu}),$$

where $\mu = \pm x, \pm y$ and $e_{\mu, \nu, r} = -e_{\nu, \mu} = -e_{-\mu, -\nu}$. The operators in Eqs. (11)–(14) may be regarded as new terms in the effective light-scattering Hamiltonian. Several comments are in order. The terms involving $O_1$ and $O_2$ appear already in the earlier work of Fleury and Loudon, however, one sees that $[O_1, H_{\text{ex}}]\neq 0$, and so the $A_{1g}$ scattering does not vanish as it would if $O_1$ were truncated at the lowest order. The previously overlooked terms $O_3$ and $O_4$ contribute, respectively, in the $A_{1g}, B_{2g}$ and in the $B_{1g}, B_{2g}$ geometries. Their precise contribution is calculable with standard techniques for the antiferromagnet.

One may note that $O_3$ is odd under time reversal, and the corresponding term only contributes inasmuch as $e_{\mu} \neq e_{\nu}$. It is precisely the chiral spin operator that has been suggested to have a nonzero ground-state expectation in some recent theories of high-$T_c$ systems. If that were the case, it would contribute to $\varepsilon_{xy}(\Omega = 0)$ as can be seen from the elastic limit of Eq. (6). In Raman scattering, we have shown that the fluctuations of this operator contribute in the $B_{2g}$ geometry.
We now turn to the doped case where the nonresonant terms [Eq. (5)] give nontrivial dominant contributions since the kinetic energy, or stress tensor, has matrix element between states in the LHB, i.e., with low-energy transfers. The simplest case corresponds to the geometry $e_i = -e_f = x$ for which the scattering intensity is proportional to

$$I_{xx}(q, \Omega) = \sum_{\tau \neq \rho} \exp(-\beta e_i) \langle \sigma | \tau_{xx}(q) | \sigma \rangle^2 \delta(e_f - e_i - \Omega).$$

(15)

The case of parabolic bands is obtained by specializing to $-\epsilon(k) = k^2$, which by Eq. (3) replaces $\tau_{x,\rho}(q)$ by $\delta_{x,\rho}(q)$. In this case, the scattering vanishes as $q \to 0$, the integrated intensity vanishing as $q^n$, with $n = 1$ if the long range of the Coulomb interaction is neglected, and with $n = 2$ otherwise. This is the inescapable consequence of particle-number conservation, i.e., $|\rho, H\rangle \to 0$ as $q \to 0$. However, for Mott-Hubbard systems, this suppression does not apply since $\tau_{xx}(q) \neq \rho(q)$ does not commute with the full Hamiltonian, and can cause scattering between the eigenstates of $H$. The effect of nonparabolicity of the bands (i.e., $\tau_{xx}(q) \neq \rho(q)$) was discussed by Wolff, who found that the random-phase approximation that $I_{xx}(q, \Omega) = 0$ for $\Omega \leq q_v f$, the latter constraint being a consequence of the limited phase space available for scattering the quasiparticle. However, for Mott-Hubbard systems, the scattering is actually dominated by the incoherent part of the carrier spectral function which defeats the phase-space limitation, resulting in the nonvanishing of $I_{xx}(q, \Omega)$, over a broad range of $\Omega \sim t$.

We can relate $I_{xx}$ to a “stress susceptibility” by the fluctuation-dissipation theorem and write $I_{xx}(q, \Omega) = -i \chi^{(c)}(q, \Omega) / \Omega - \exp(-\beta \Omega)$ and express

$$\chi^{(c)}(q, \Omega) = N^{-1} \sum_{k, k'} \gamma^{(c)}(k) \gamma^{(c)}(k') M_{k, k'}(q, \Omega),$$

(16)

where the bare vertex $\gamma^{(c)}(k) = \cos(k_x)$. The function $M$ is given as the sum of the “bubble” and the vertex contributions,

$$M_{k, k'}(q, \Omega) = \delta_{k, k'} A_{k}(q, \Omega) + N^{-1} B_{k, k'}(q, \Omega).$$

The term $A_{k}$ corresponds to the bubble diagram with fully renormalized propagators and can be formally evaluated in terms of the single-particle spectral functions. The contribution arising from the coherent part of these vanishes as $q \to 0$ at zero temperature. The incoherent part is, in general, nonzero at all wave vectors but vanishes as $\Omega \to 0$ as $\Omega^2$, with $\gamma \geq 1$. The same function $M_{k, k'}$ determines the density and current susceptibilities $\chi^{(d)}(q, \Omega)$ and $\chi^{(j)}(q, \Omega)$, which are obtainable from Eq. (16), by replacing the stress vertex $\gamma^{(c)}$ by unity and the current vertex $\nu_{c}(k)$, respectively.

The conservation of particle number leads to the Ward-Takahashi identities relating the vertex correction term $B_{k, k'}$ and the bubble term $A_{k}$ through $\langle B_{k, k'}(0, \Omega) \rangle_{k, k'} = -\langle A_{k}(0, \Omega) \rangle_{k}$ for all $\Omega$, whereby $\Im \chi^{(d)}(0, \Omega) = 0$ for $\Omega \neq 0$. The implication is that the Raman cross section in the doped case, if due to density fluctuations, should vanish with $q$, the photon momentum transfer which is usually very small compared to, say, the Fermi momentum. In actual experiments on high-$T_c$ materials, the typical $q$ is roughly the inverse skin depth $\lambda_s \approx 1000 \AA$. We then expect that for nonparabolic Mott-Hubbard systems, the resulting integrated intensity should be larger than that in a typical free-electron metal by factors of order $(r_s/a_0)^m$, where $a_0 (\approx 5 \AA)$ is a lattice constant, or typical interparticle spacing, and $m$ is either 1 or 2.

We are able at present to give only a very crude estimate of the Raman intensity. It is readily seen that the intensity integrated over frequencies reduces to

$$\int_{0}^{\Omega_s} d \Omega \ I_{xx}(0, \Omega) = \langle P_G \tau_{xx} P_G \tau_{xx} P_G \rangle - \langle P_G \tau_{xx} P_G \rangle^2,$$

where $P_G$ projects out doubly occupied sites, and the cutoff frequency $\Omega_s$ is assumed to be in the range $t \leq \Omega \leq U$, so that only LHB states can contribute (crudest $\Omega_s \sim$ insulating optical gap), and should scale like $t^2$ times the hole density $\delta$. This intensity should be seen in the $B_{1g}$ and $A_{1g}$ geometries with $\frac{1}{2} (\tau_{xx} \mp \tau_{xy})$ replacing $\tau_{xx}$ in Eq. (15). The various moments can be calculated as higher commutators, and expressed as correlation functions. In contrast to the density-fluctuation picture of Raman scattering, which leads to well-known features including a particle-hole continuum with a well-defined cutoff $q_f$ and a sharp plasma mode arising from collective density fluctuations, in Mott-Hubbard systems the intermediate states probed need not necessarily have any sharp structure since the stress tensor is not expected to create well-defined elementary excitations. Thus we expect a broad continuum with a bandwidth of order $t$, with an “anomalously” large intensity compared to free-electron metals, scaling like the hole density near half filling. This description is qualitatively consistent with the experiments where the largest energy transfer is $\approx 1 \text{ eV}$.

An interesting cross-check is provided by considering the optical conductivity $\Re \sigma_{xx}(0, \Omega) = \Im \chi^{(j)}(0, \Omega)/\Omega$. In the approximation where only the bubble contributions to $\Re \sigma_{xx}$ are retained, these are essentially identical since $\langle \cos^2(k_x) \rangle = 1 - (\sin^2(k_x))$. The optical experiments on $\text{YBa}_2\text{Cu}_3\text{O}_7$ do seem to bear out, albeit crudely, this pseudoidentity,

$$I_{xx}(0, \Omega) \sim \Omega / \left[ 1 - \exp(-\beta \Omega) \right] \Re \sigma_{xx}(0, \Omega),$$

for small enough $\Omega$, although at higher frequencies there seem to be significant departures. In the limit $\Omega \leq kT$, the above reduces to $I_{xx} \sim kT / \Re \sigma_{xx}$; together with a temperature-independent $I_{xx}(0, 0)$, it implies a linear resistivity, which is a ubiquitous feature in the high-$T_c$ materials.

For this pseudoidentity and also the Ward-Takahashi identity to hold, we would have to argue that the vertex corrections are small in the non-$s$-wave channels of the function $B_{k, k'}$. Recent work on the Hubbard model shows that this scenario is realized in the limit of high dimensions. We should also stress that these considera-
tions apply only to small momentum transfers in the singlet particle-hole channel; the triplet-channel spin susceptibility for $q \sim [\pi, \pi]$ is expected to have substantial structure in view of the oxygen and copper planar NMR relaxation rates.\textsuperscript{20}

We summarize by returning to recent experiments\textsuperscript{21} in the insulating case where it is seen that the integrated intensity in the $B_{1g}$, $B_{2g}$, and $A_{1g}$ geometries are all of the same order of magnitude, and, in fact, as the laser frequency $\omega_L$ is changed, the $A_{1g}$ and $B_{2g}$ intensities can be larger than that in $B_{1g}$. Within $B_{1g}$ geometry, a reasonable understanding of the scattering has been reached by Singh et al.\textsuperscript{22} The fact that the scattering intensity in $A_{1g}$ and $B_{2g}$ geometries exceeds that in $B_{1g}$ is not easy to reconcile with the fact that the second-neighbor exchange $J_2$ is smaller than $J_1$ by at least an order of magnitude. Within our framework, however, it is clear that the ratio of the prefactors of the scattering operators that contribute to $A_{1g}$ and $B_{1g}$ [see Eqs. (11) and (14)] is different from $J_2/J_1$, and, in fact, it involves a different power of $t/(U - \omega_L)$ when such an expansion is possible. In general, when $U - \omega_L$, the forms of the spin operator that can contribute involve arbitrarily long strings of spins; the upper-Hubbard-band propagator is transmitted into the scattering operator. Experiments have been done with laser frequencies between 2.4 and 2.8 eV, which corresponds fairly closely to the effective $U$ in the 2D cuprates.\textsuperscript{23} Hence we believe that the experiments are probing the most interesting and difficult region of Raman scattering in the Mott-Hubbard system.

The most striking prediction of this theory concerns the $q$ dependence of the inelastic Raman scattering: We are predicting that the integrated intensity does not vanish as $q \rightarrow 0$, simply because the conservation laws do not force this vanishing. This prediction should be amenable to experimental test, and is in our opinion a crucial test of the applicability of the standard dielectric function theory of Raman scattering\textsuperscript{11,12} to the high-$T_c$ systems. It also follows that the semiconductor $n$-type InSb and others having strong nonparabolicity could also display a nonvanishing intensity as $q \rightarrow 0$, the effects being proportional to $1 - z_k$, and largest when the electron density is low (i.e., the effective $r_2 \geq 3$). The other important result shown here is the fact that the $B_{2g}$ geometry is particularly favorable for studying quasilastic time-reversal violation, both in the insulating phase [Eq. (14)] and in the doped phase.\textsuperscript{24} It is outside the scope of this work to give an estimate of the relevant correlation functions, and we hope that our result stimulates concrete calculations of these that can be tested against experiment.\textsuperscript{24}

We are pleased to acknowledge stimulating discussions with S. L. Cooper, P. Fleury, K. Lyons, and P. Sulewski, and helpful comments by G. Kotliar and P. A. Lee.


\textsuperscript{6}In the insulating case, this is a good approximation since $\epsilon_{\sigma} - \epsilon_{\alpha} \sim U$, in the doped case, however, the intermediate states corresponding to the LHB have $\epsilon_{\sigma} - \epsilon_{\alpha} \ll \omega_L \sim U$, whereby there is a tendency for the two terms in Eq. (6) to cancel out, particularly in the diagonal components $M_{\alpha \beta}^{\alpha \beta}$. We are dropping this term for clarity of presentation, although we expect it to be nonzero in the doped case (see Refs. 10 and 24).


\textsuperscript{8}B. S. Shastry and B. I. Shraiman (to be published).


\textsuperscript{10}The resonant term is expected to continue to contribute in this phase on general grounds. Its structure, however, is expected to depart significantly from the insulating phase in view of the fact that the current operator has matrix elements in the lower Hubbard band, as discussed later in connection with the optical conductivity. The intensity is expected to be smaller than in the insulating phase for $\omega_L \sim U$ (see Refs. 6 and 24).

\textsuperscript{11}D. Pines and P. Nozieres, Theory of Quantum Liquids (Benjamin, Reading, MA, 1966).


\textsuperscript{14}W. Hayes and R. Loudon, Scattering of Light by Crystals (Wiley, New York, 1978).

\textsuperscript{15}P. A. Wolff, Phys. Rev. 171, 436 (1968).

\textsuperscript{16}Numerical evaluation of this function should be feasible using exact diagonalization methods for small clusters.


\textsuperscript{18}In order to make a more complete comparison one needs to have optical conductivity data without the contributions for chains in the 1-2-3 materials.


\textsuperscript{24}We observe that in the doped case, with $\omega_L \sim \epsilon_{\alpha}$, and with $\epsilon_{\alpha} \sim \epsilon_{\alpha} - i$, the two terms in Eq. (6) reduce to essentially $M_{\alpha \beta}^{\alpha \beta} = \langle f_{\alpha f_{\alpha}} P_{\alpha f_{\alpha}} \rangle \omega_L$, which would tend to $\int \Im \sigma^{\alpha \alpha}(\omega) - \sigma^{\alpha \alpha}(\omega) \omega d\omega$ if $|f_{\alpha}| = |f_{\beta}|$.\textsuperscript{3}