The Nernst effect and the boundaries of the Fermi liquid picture

2009 J. Phys.: Condens. Matter 21 113101

(http://iopscience.iop.org/0953-8984/21/11/113101)

J. Phys.: Condens. Matter **21** (2009) 113101 (9pp) doi:10.1088/0953-8984/21/11/113101

TOPICAL REVIEW

The Nernst effect and the boundaries of the Fermi liquid picture

Kamran Behnia

Laboratoire Photons et Matière (UPR5-CNRS), ESPCI, 10 Rue Vauquelin,
F-75005 Paris, France

Received 8 September 2008, in final form 15 January 2009
Published 9 February 2009
Online at stacks.iop.org/JPhysCM/21/113101

Abstract
Following the observation of an anomalous Nernst signal in cuprates, the Nernst effect has been explored in a variety of metals and superconductors during the past few years. This paper reviews the results obtained during this exploration, focusing on the Nernst response of normal quasi-particles as opposed to the one generated by superconducting vortices or by short-lived Cooper pairs. Contrary to what has been often assumed, the so-called Sondheimer cancelation does not imply a negligible Nernst response in a Fermi liquid. In fact, the amplitude of the Nernst response measured in various metals in the low-temperature limit is scattered over six orders of magnitude. According to the data, this amplitude is roughly set by the ratio of electron mobility to Fermi energy, in agreement with the implications of semi-classical transport theory.

(Some figures in this article are in colour only in the electronic version)

Contents

1. Introduction

The Nernst effect is the generation of a transverse electric field by a longitudinal thermal gradient in the presence of a finite magnetic field. It has attracted considerable attention over the past few years following the report on the observation by Ong's group [1] of a finite Nernst effect in the high-T_c cuprate $La_{2-x}Sr_xCuO_4$ above the critical temperature. Before this report, the Nernst effect had been studied in both conventional [2] and high-T_c [3–6] superconductors. Vortex movement caused by the application of a thermal gradient was a well-known source of a Nernst signal [7]. However, at least in the community of researchers exploring correlated electrons, general knowledge regarding other sources of the Nernst effect was fragmentary. In common metals, investigations of the Nernst effect have been out of fashion for several decades [8]. As for metals host to correlated electrons, their Nernst response is still largely unexplored.

This situation changed considerably during the first years of this century. The Nernst effect has been studied in a variety of metals and superconductors. Early measurements of the Nernst effect in the high-T_c cuprates performed in the 1990s, apart from a few exceptions [9], focused on vortex dynamics. These studies were complemented by new ones probing the anomalous Nernst response of the pseudogap state [10–17].

This has been the subject of a review by Wang *et al* [18]. Moreover, experiments were performed on other families such as organic superconductors [19, 20, 22, 21, 23] and heavy-fermion superconductors [24–26, 28, 27] as well as on a charge density wave (CDW) superconductors such as $NbSe_2$ [29]. In addition to these studies on clean superconductors, amorphous superconducting thin films were also explored for the first time [30–32].

One surprising outcome of these explorations was the detection of a sizeable Nernst signal in many cases in the absence of superconductivity or superconducting fluctuations. This was at first attributed to exotic physics, since the general scientific opinion was that the Nernst effect generated by normal quasi-particles in an ordinary metal should be negligibly small. This line of thought was supported by the fact that many cases of metals displaying a 'giant Nernst effect' displayed other anomalous transport properties. Moreover, in some cases, such as URu_2Si_2 [25] and $PrFe_4P_{12}$ [33], the puzzlingly large Nernst signal was accompanied by an exotic ground state with an unidentified order parameter.

The confusion was somewhat dissipated by the rediscovery of elemental bismuth [34]. The Nernst signal in bismuth is so large [35, 36] that it was even detectable by late 19th century technology [37]. A recent study confirmed the large magnitude of the Nernst effect in bismuth at low temperatures, which exceeds by more than one order of magnitude the Nernst response of any correlated metal [34]. Therefore, the natural question was to check if the semi-classical transport theory could account for the size of the Nernst effect in bismuth. This short review argues that the answer to this question is affirmative. The large Nernst signal in bismuth is a consequence of a large electron mobility and a small Fermi energy as implied by an equation first derived by Sondheimer in 1948 [38] and reformulated recently [39]. As we will see below, the available Nernst data for other metals are compatible with this picture. Thus, the Nernst effect roughly measures the ratio of electron mobility to Fermi energy in a given metal.

Admitting that normal quasi-particles can generate a sizeable Nernst signal does not undermine the use of the Nernst effect as a probe of superconducting fluctuations. On the contrary, knowing the order of magnitude of the purely metallic Nernst response is indispensable to disentangle the signal generated by short-lived Cooper pairs (or eventually short-lived vortices) in the normal state of a superconductor from the one expected from quasi-particles. Since a reduced electron mobility and a large Fermi energy lead to a small Nernst response, an amorphous conventional superconductor such as Nb_xSi_{1-x} is an appropriate system for the detection of the Nernst signal due to short-lived Cooper pairs. Indeed, a finite Nernst signal was observed in this system in a temperature window extending up to 30 times T_c [30, 31]. Since this signal exceeds by three orders of magnitude what is expected by the normal electrons, it is safe to assume that it is not caused by them. On the other hand, close to T_c, its magnitude is in quantitative agreement with what is theoretically expected for Gaussian fluctuations of the superconducting order parameter [40], leaving little doubt about their origin.

Figure 1. The two sign conventions for the Nernst effect: (a) the historical convention and (b) the vortex convention (see the text).

2. Nernst effect, the semi-classical picture, and sign conventions

The three conductivity tensors, $\overline{\sigma}$, $\overline{\kappa}$ and $\overline{\alpha}$ relate charge current, $\mathbf{J_e}$, and heat current, $\mathbf{J_q}$ to electric field, $\mathbf{E}$ and thermal gradient, $\nabla\mathbf{T}$ vectors:

$$\mathbf{J_e} = \overline{\sigma}\cdot\mathbf{E} - \overline{\alpha}\cdot\nabla\mathbf{T} \tag{1}$$

$$\mathbf{J_q} = T\overline{\alpha}\cdot\mathbf{E} - \overline{\kappa}\cdot\nabla\mathbf{T}. \tag{2}$$

In the absence of charge current (i.e. when $\mathbf{J_e} = 0$), the first equation yields:

$$\mathbf{E} = \overline{\sigma}^{-1}\cdot\overline{\alpha}\cdot\nabla\mathbf{T}. \tag{3}$$

Therefore, the Nernst signal, N, which is the transverse electric field, E_y, generated by a longitudinal thermal gradient, $\nabla_x T$, would be:

$$N = \frac{E_y}{\nabla_x T} = \frac{\alpha_{xy}\sigma_{xx} - \alpha_{xx}\sigma_{xy}}{\sigma_{xx}^2 + \sigma_{xy}^2}. \tag{4}$$

The sign convention for the Nernst effect has been a source of confusion, since two different sign conventions have been used (see figure 1). In the first convention, a positive Nernst signal corresponds to an electric field along the y-axis, when the thermal gradient is along the x-axis and the magnetic field along the z-axis. This older convention is the one used in thermoelectric literature [41] and textbooks [42] unrelated to the Nernst signal of the vortices. According to this convention, the Nernst signal in bismuth is negative [41, 36]. However, a second sign convention has been widely used during the last few years. According to it, the Nernst signal expected by the vortices moving from hot to cold is taken as positive [18]. As seen in figure 1, this is opposite to the first convention. The scientific literature on the superconducting Nernst signal, apart from a few exceptions [43], has used this latter convention. The existence of two opposite sign conventions led to a confusion in the case of $CeCoIn_5$ [24, 27, 28]. The Nernst signal in the normal state of this system is negative according to the first (or the historical) convention, but positive according to the second (or the vortex) convention. This feature, not correctly grasped in the first communication on the observation of the Nernst effect in this system [24] was subsequently corrected [27, 28]. In this text we are going to use the more popular vortex

convention, according to which the Nernst signal in bismuth (which is negative in the historical convention) would be positive.

The solution of the Boltzmann equation leads to the following link between the electric and the thermoelectric conductivity tensors:

$$\overline{\alpha} = -\frac{\pi^2}{3}\frac{k_B^2 T}{e}\frac{\partial \sigma}{\partial \epsilon}\bigg|_{\epsilon=\epsilon_F}. \tag{5}$$

In other words, the thermoelectric response is a measure of the variation in conductivity caused by an infinitesimal shift in the chemical potential.

Combining equations (4) and (5) with the definition of the Hall angle ($\tan\theta_H = \frac{\sigma_{xy}}{\sigma_{xx}}$) yields:

$$N = -\frac{\pi^2}{3}\frac{k_B^2 T}{e}\frac{\partial \tan\theta_H}{\partial \epsilon}\bigg|_{\epsilon=\epsilon_F}. \tag{6}$$

Such an expression directly linking the Nernst effect to the energy derivative of the Hall angle was first put forward by Oganesyan and Ussishkin [39]. Now, in a one-band picture, the Hall angle is equal to:

$$\tan\theta_H = \omega_c \tau = \frac{eB\tau}{m^*}. \tag{7}$$

Here ω_c is cyclotron frequency, τ is the scattering time, and m^* is the effective mass. Therefore, assuming that the scattering time is the only energy-dependent component of the Hall angle, an alternative expression for equation (6) would be:

$$\nu = N/B = -\frac{\pi^2}{3}\frac{k_B^2 T}{m^*}\frac{\partial \tau}{\partial \epsilon}\bigg|_{\epsilon=\epsilon_F}. \tag{8}$$

This is the expression which appears in Sondheimer's monograph of 1948 [38]. Note the equivalence between equations (6) and (8). A superficial reading of equation (8) would erroneously conclude that the Nernst effect inversely scales with the effective mass. But this is misleading, since any change in the effective mass would have consequences on the Fermi energy. For this reason among others, equation (6) is more transparent. It states that *if an infinitesimal shift in the chemical potential leaves the Hall angle unchanged, then the Nernst response of the system is nil*. This statement is no more than one formulation of what has been dubbed Sondheimer cancelation by Wang and co-workers [10].

There is no well-established experimental procedure to determine $\frac{\partial \tan\theta_H}{\partial \epsilon}|_{\epsilon=\epsilon_F}$. The simplest approximation is to assume that the Hall angle does not depend on energy in the vicinity of the Fermi energy. This assumption would lead to a strictly zero Nernst response.

3. Two routes towards a finite Nernst signal

There are two distinct roads to the finite Nernst response observed in real metals. The first is the presence of multiple bands and the second is an energy-dependent Hall angle.

3.1. Ambipolarity

If $\frac{\partial \tan\theta_H}{\partial \epsilon}|_{\epsilon=\epsilon_F} = 0$, then equation (5) implies the equality:

$$\frac{\alpha_{xy}}{\alpha_{xx}} = \frac{\sigma_{xy}}{\sigma_{xx}}. \tag{9}$$

In a one-band metal, this would lead to a vanishing Nernst response. However, in a two-band metal with both electron-like and hole-like carriers, a finite Nernst response can arise even in these conditions. Indeed, in this case, equation (4) should be replaced by:

$$N = \frac{(\alpha_{xy}^+ + \alpha_{xy}^-)(\sigma_{xx}^+ + \sigma_{xx}^-) - (\alpha_{xx}^+ + \alpha_{xx}^-)(\sigma_{xy}^+ + \sigma_{xy}^-)}{(\sigma_{xx}^+ + \sigma_{xx}^-)^2 + (\sigma_{xy}^+ + \sigma_{xy}^-)^2}. \tag{10}$$

The superscripts + and − refer to hole-like and electron-like bands. Now, since the signs of σ_{xy} and α_{xx} depend on the sign of the carriers, the validity of equation (9) for each band does not lead to a vanishing numerator in equation (10).

3.2. Energy-dependent mobility

Even in a one-band picture, the Hall angle can be energy-dependent. In this case, in a first approximation, $\frac{\partial \tan\theta_H}{\partial \epsilon}|_{\epsilon=\epsilon_F}$ can be replaced by $\frac{\tan\theta_H}{\epsilon_F}$. Note that this is equivalent to assuming that $\tan\theta_H$ is a linear function of energy in the vicinity of the Fermi energy. If the energy dependence is smooth but stronger or weaker than linear, then the two expressions should differ by a constant of the order of unity. The derivative vanishes only in the specific case of the energy dependence presenting an extremum at the chemical potential. This would be equivalent to a perfect electron–hole symmetry, a feature which is often assumed but never demonstrated to occur in real metals.

Moreover, it is preferable to substitute the Hall angle by what it physically measures, that is the carrier mobility, μ. The latter can be expressed as:

$$\tan\theta_H/B = \mu = \frac{e\tau}{m^*} = \frac{e\hbar k_F}{\ell_e}. \tag{11}$$

Here k_F is the Fermi wavevector and ℓ_e is the carrier mean-free-path. This is particularly useful in the case of multi-band metals. Since the sign of the Hall angle is different for hole-like and electron-like carriers, the overall Hall angle of an ambipolar metal can be substantially reduced compared to the Hall angle (and the mobility) of each band.

These two simplifications lead us to the following expression for the magnitude of the Nernst coefficient:

$$\nu = \frac{\pi^2}{3}\frac{k_B}{e}\frac{K_B T}{\epsilon_F}\mu. \tag{12}$$

Since this expression uses fundamental constants and two system-dependent parameters, it is easy to use it in order to confront the measured value of the Nernst coefficient with the expectations of the semi-classical transport theory. It was first used in the context of the investigation into the source of the Nernst signal in URu_2Si_2 [25] and $PrFe_4P_{12}$ [33].

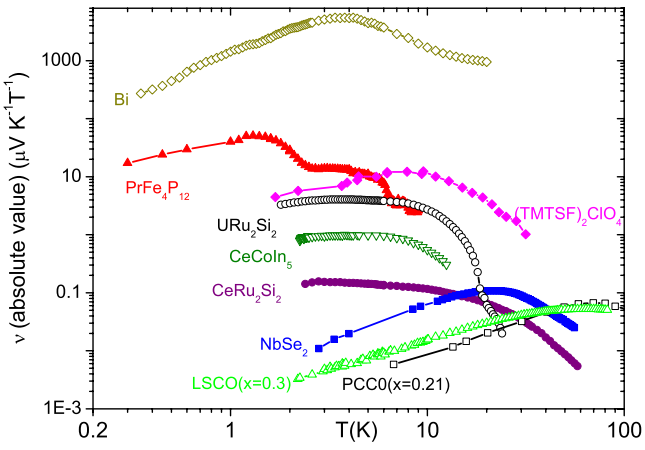

Figure 2. The magnitude of the Nernst coefficient in a number of metals.

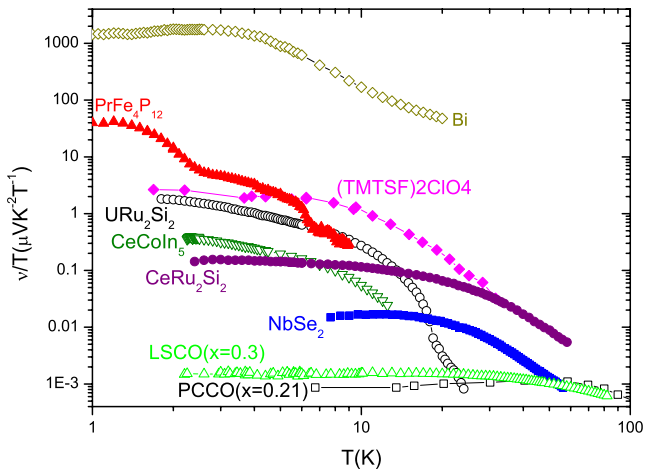

Figure 3. The magnitude of the Nernst coefficient divided by temperature.

Note that the expression proposed by Oganesyan and Ussishkin for a compensated two-band metal (that is equation A3 in [39]):

$$B\nu = \frac{2\pi^2}{3}\frac{k_B}{e}\frac{K_B T\tau}{\hbar}\frac{1}{(k_F\ell_B)^2} \tag{13}$$

is identical to equation (10). The strict equivalency between the two expressions becomes explicit if one replaces the magnetic length, ℓ_B and the scattering time, τ by their corresponding values (that is $\ell_B^2 = \frac{\hbar}{eB}$ and $\tau = \frac{\ell_e m^*}{k_F}$).

The only difference between the two expressions is the visibility of the physical parameters which enhance the Nernst response. According to equation (12), the necessary ingredients for an enhanced Nernst signal is a large electronic mobility and a small Fermi energy. Equation (13) yields the same message by tracing the source of the Nernst signal to a long scattering time and a small wavevector. In the following, we are going to use equation (12), because of the simplicity of distinguishing three scales: first a universal scale ($\frac{\pi^2}{3}\frac{k_B}{e} = 283.7\ \mu\text{V K}^{-1}$), second the inverse of the Fermi energy in kelvins, and finally the mobility in T^{-1} (or in $\text{m}^2\ \text{V}^{-1}\ \text{s}^{-1}$). As we shall see below, the order of magnitude of the available Nernst data is in reasonable agreement with the expectations of the semi-classical theory.

3.3. Nernst, Seebeck, and Hall coefficients

When the temperature is much lower than the Fermi temperature, the Seebeck coefficient of a metallic system, S, is expected to become T-linear with a slope linked to the Fermi temperature through the following simple expression:

$$S = \frac{\pi^2}{2}\frac{k_B}{e}\frac{T}{T_F}. \tag{14}$$

This expression is strictly valid only in the case of a free electron gas. It is very similar to the one linking the electronic specific heat, γ, of a free electron gas to its Fermi temperature:

$$\gamma = \frac{\pi^2}{2}\frac{k_B}{T_F}n. \tag{15}$$

Here n is the (molar) carrier density.

Interestingly, the link between γ and S/T survives even in the presence of strong electronic interaction. An examination of the available thermopower and specific heat data in various families of correlated metals suggests that, at low enough temperatures, the dimensionless ratio of the Seebeck coefficient to the specific heat remains of the order of unity [44].

If in equation (12), one replaces the Fermi energy with the Seebeck coefficient (using equation (14)) and the mobility with the Hall angle one finds

$$\nu B = \tfrac{2}{3} S \tan\theta_H. \tag{16}$$

Therefore, equation (12) is another formulation of a fundamental link between the Nernst and Seebeck coefficients through the Hall angle. However, this equivalency between the two equations only holds in the case of one-band systems. In a multi-band system, the measured Hall angle can be significantly lower than the mobility. In bismuth, for example, in the presence of a magnetic field, $\sigma_{xx} \gg \sigma_{xy}$, but $N \gg S$. Therefore, equation (16) fails. However, as we shall see below, equation (12) holds.

4. Short review of experimental data

In this section we put under scrutiny the available Nernst data. The study should focus on cases associated with a Nernst effect generated by normal electrons (as opposed to the signal linked to superconductivity). The temperature dependence of the absolute value of the Nernst coefficient of the systems considered in this review is presented in figure 2. Theoretically, the Nernst coefficient should become T-linear at low enough temperature. It is the order of magnitude of this T-linear coefficient, which should be confronted to equation (10) (or (11)). With this in mind, figure 3 presents the Nernst

data as a plot of ν/T versus temperature. In order to see how the magnitude of ν/T at low temperatures conforms to the expectations of equation (10), one needs to extract the magnitude of the Fermi energy and the electronic mobility in each system. Let us briefly consider them. A list of extracted parameters is given in table 1.

4.1. $NbSe_2$

A study of the Nernst effect in $NbSe_2$ was reported by Bel and co-workers [29] who attributed the existence of a finite Nernst signal to ambipolarity. However, the magnitude of the Nernst coefficient was not put under analysis. $NbSe_2$ is not a strongly correlated system and has a conventional carrier density of about one carrier per formula unit. Unsurprisingly, its Nernst response is smaller than correlated metals with high-mobility electrons. We use the thermoelectric and Hall data to estimate the electronic mobility and the Fermi temperature. The Fermi temperature can be estimated by taking the slope of the Seebeck coefficient at low temperature ($S/T = 0.3\ \mu\mathrm{V\ K^{-2}}$ [29]) and equation (14). The electronic mobility is estimated using the low-temperature magnitude of the Hall angle of the sample studied in [29].

4.2. Ce-based heavy fermions

The observation of a large Nernst coefficient (of the order of $\mu\mathrm{V\ K^{-1}\ T^{-1}}$) in $CeCoIn_5$ by Bel and co-workers [24] was unexpected. This result was confirmed and extended by subsequent studies by Onose *et al* [27] and Izawa *et al* [28]. The latter study focused on the low-temperature region and found that the Nernst response is particularly enhanced in the vicinity of the field-induced quantum critical point (QCP), which occurs at 5 T [45, 46]. We will return to this study in section 4.3 on quantum criticality. Here, let us consider the Nernst response in the zero-field limit, which was probed down to the onset of superconductivity ($T \geqslant 2.2$ K). Since the transport properties of the system such as thermopower and Hall coefficient are markedly non-Fermi liquid in the temperature window extending from T_c to 20 K [47], they cannot be used to estimate the Fermi temperature and the mobility. One crude estimate of the Fermi temperature is yielded by the magnitude of the electronic specific heat, γ, equation (15) and assuming a carrier density of 1/f.u. (i.e. the system is close to half filling), the magnitude of γ in $CeCoIn_5$ ($0.6\ \mathrm{J\ K^{-2}\ mol^{-2}}$) [45] implies a Fermi temperature of 60 K.

As for electronic mobility, the value given in table 1 is extracted from the Dingle temperature ($k_B T_D = \hbar/\tau$) and the effective masses given by de Haas–van Alphen measurements [48].

Measurements on $CeRu_2Si_2$ revealed a Nernst signal with a magnitude somewhat smaller than what was found in $CeCoIn_5$ [26]. Here, the zero-field state is a Fermi liquid and the Fermi temperature estimated by the slope of thermopower ($S/T = 2.4\ \mu\mathrm{V\ K^{-2}}$ [49]) or by the magnitude of electronic specific heat ($\gamma = 0.35\ \mathrm{J\ mol^{-1}\ K^{-2}}$ [50]) are comparable. The electronic mobility was estimated using the Hall angle data [51].

Table 1. The magnitude of the Nernst coefficient divided by temperature at low temperature, together with estimations of the electronic mobility and the Fermi energy in various metals. The fourth column yields the expected magnitude of ν/T according to equation (10).

System	ν/T ($\mu\mathrm{V\ K^{-2}\ T^{-1}}$)	μ ($\mathrm{T^{-1}}$)	ϵ_F ($\mathrm{K^{-1}}$)	$\frac{\pi^2}{3}\frac{k_B}{e}\frac{\mu}{\epsilon_F}$ ($\mu\mathrm{V\ K^{-2}\ T^{-1}}$)
Bi	750	420	130	914
$CeRu_2Si_2$	0.16	0.2	180	0.25
$CeCoIn_5$	0.5	0.3	60	1.4
URu_2Si_2	1.8	0.08	25	0.9
$PrFe_4P_{12}$	57	0.85	8	30
$(TMTSF)_2ClO_4$	2.6	0.75	110	1.9
$La_{1.7}Sr_{0.3}CuO_4$	0.0015	0.01	5900	4.8×10^{-4}
$Pr_{1.79}Ce_{0.21}CuO_4$	9×10^{-4}	0.005	4300	3.3×10^{-4}
$NbSe_2$	0.015	0.09	1400	0.018

4.3. Heavy-electron metals with unidentified orders

Among heavy-fermion metals, the magnitude of the Nernst coefficient in URu_2Si_2 [25] and in Pr-based skutterudite $PrFe_4P_{12}$ [33] becomes particularly large when they order. Interestingly, in both these systems the order parameter appears to be exotic and remains unidentified. However, the large magnitude of the Nernst coefficient in both cases can be traced back to the semi-metallic nature of the ordered state implying (in the language of equation (11)) a small Fermi wavevector and a long scattering time [34].

In both cases, de Haas–van Alphen (dHvA) studies have detected only small pockets of Fermi surface. In the case of URu_2Si_2, three frequencies have been detected. The largest corresponds to a Fermi surface whose volume in only 0.02 of the Brillouin zone [52]. In the case of $PrFe_4P_{12}$, only one frequency is detected, corresponding to 0.0015 of the Brillouin zone [53]. In both cases, the mass and the volume of the pockets detected do not sum up to the magnitude of the measured electronic specific heat. This suggests that one or several massive low-mobility pockets have not been detected yet.

In URu_2Si_2 [25], the slope of the Seebeck coefficient continues to increase down to the superconducting transition temperature ($T_c = 1.5$ K), making it difficult to extract the zero-temperature value needed to estimate the Fermi temperature. A more reliable process would be to use $\gamma = 0.065\ \mathrm{J\ mol^{-1}\ K^{-2}}$ and a carrier density of 0.04/f.u. for this compensated system [54]. This estimate, ($T_F = 25$ K) is comparable to the Fermi temperature of the β-band ($T_F = 22$ K), which has the lowest Fermi temperature among the three detected frequencies.

In $PrFe_4P_{12}$ [33], the situation is more straightforward: the slope of thermopower ($S/T = 56\ \mu\mathrm{V\ K^{-2}}$) yields a rather low Fermi temperature ($T_F = 8$ K). This value is very close to the width of the Kondo resonance (8.7 K) found in the specific heat data [55].

The mobilities of table 1 are based on the measurements of the Hall angle on the crystals which were used in the Nernst studies. Let us note that in the new generation of ultraclean URu_2Si_2 single crystals [54] electronic mobility is at least one order of magnitude larger. Therefore, according to our current

understanding, the Nernst response of these crystals should also be enhanced by an order of magnitude.

4.4. Bechgaard salts

The expression 'giant Nernst effect' was first employed by Wu *et al* following the observation of a very large resonant Nernst response in $(TMTSF)_2PF_6$ [19]. A similar feature, but less pronounced, was observed in the angular dependence of the Nernst effect in $(TMTSF)_2ClO_4$ [20]. In Bechgaard salts when the field is oriented along some peculiar orientations called 'magic angles', all transport properties are anomalous and the large resonant Nernst response is a fascinating problem which should be addressed in this context [21]. The magnitude of the Nernst coefficient at zero-field on the other hand, may be easier to understand. Nam *et al* [22] studied the Nernst response of $(TMTSF)_2ClO_4$ in the low field limit and found that the Nernst coefficient becomes large below the anion-ordering temperature. Their data are included in figures 1 and 2. The slope of the thermopower in the data reported by Nam *et al* ($\sim 4\ \mu V\ K^{-2}$) [22] can be used to estimate the Fermi temperature in $(TMTSF)_2ClO_4$. The mobility can be estimated using the scattering time ($\tau = 4.3 \times 10^{-12}$ s) deduced from angular magnetoresistance studies by Danner *et al* [56].

4.5. Elemental bismuth

The Nernst effect in elemental bismuth for temperatures exceeding 4.2 K was studied decades ago by three different groups [35, 36, 41]. The magnitude of the Nernst response in this semi-metal easily dwarfs all other cases of 'giant' Nernst effect. The result was rediscovered, confirmed, and extended to lower temperatures recently [34]. The Fermi surface in bismuth is well known, the small Fermi temperature for the electron pocket (27 meV) and hole pocket (11 meV) has been known for a long time [57, 58]. It is also known that mobility in bismuth is very large and can exceed $10^7\ cm^2\ V^{-1}\ s^{-1}$ [58]. There is, therefore, no surprise that bismuth, in which the Nernst effect was originally discovered, is still the metal with the largest known Nernst coefficient. The values given in table 1 correspond to the mobility of the crystal used in [34] and the Fermi temperature of the hole pocket (which is more mobile than the electron pocket as indicated by their respective Dingle temperatures [59]). Note that, since bismuth is a compensated system, the Hall angle is much lower than the mobility of either holes or electrons. Therefore, equation (16) fails.

4.6. Overdoped cuprates

It was the discovery of an anomalous Nernst effect in hole-doped cuprates [1] which started a tremendous interest in the physics of the Nernst effect. It is well known that the existence of a robust superconducting ground state is a major obstacle to probe the properties of the normal ground state in the hole-doped cuprates. This is not the case of the electron-doped ones. Studies of the Nernst effect in the electron cuprates [60, 43, 61–63] have detected a sizeable normal-state Nernst response and attributed it to the existence of two bands.

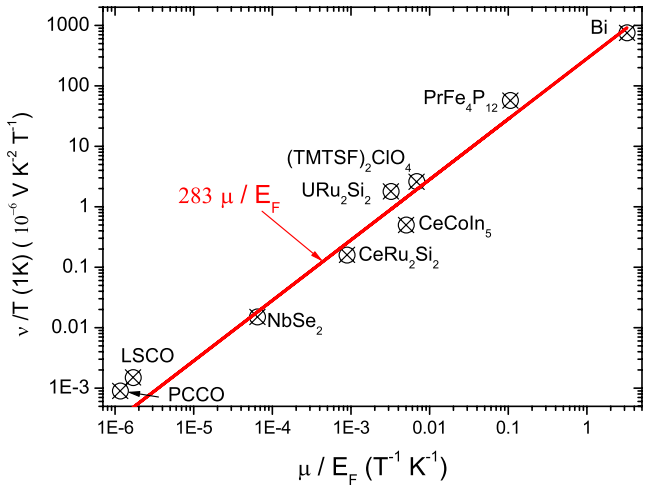

Figure 4. The low-temperature slope of the Nernst coefficient as a function of the ratio of electron mobility to the Fermi energy. The values are those listed in table 1.

The data included in figures 3 and 4 were those reported for PCCO ($x = 0.21$) by Li and Greene [63]. At this doping level, the system is overdoped and its is expected to display a Fermi liquid behavior. The Fermi temperature and the mobility given in the table are extracted from the reported data for the slope of thermopower [64] and the Hall angle [65].

In the case of hole-doped cuprates, the data presented are unpublished results obtained in my group on $La_{1.7}Sr_{0.3}CuO_4$ [66]. At this doping level, superconductivity is totally destroyed and no trace of it could be found down to 0.1 K [67]. The system becomes a Fermi liquid and its resistivity displays a purely T^2 behavior [67]. The sign of the Nernst signal is negative in the vortex convention. The Fermi energy given in table 1 is calculated using the magnitude of the electronic specific heat ($\gamma = 6.9\ mJ\ mol^{-1}\ K^{-2}$ [67]). The mobility is extracted from the slope of the B-square magnetoresistance [67].

Let us note that the Nernst responses of $La_{1.7}Sr_{0.3}CuO_4$ and in $Pr_{1.79}Ce_{0.21}CuO_4$ are comparable in magnitude but opposite in sign. Since the hole-doped cuprate is believed to have a single large Fermi surface, no ambipolar Nernst effect is expected there. Its sizeable Nernst response suggests that a detectable Nernst signal could exist even in the absence of ambipolarity. In both cases, the low-temperature Nernst signal is three to four times larger than what is expected according to the simple picture. This discrepancy appears to be much larger than the uncertainty on the magnitude of either the mobility or the Fermi energy and appears significant. It is tempting to link it to a non-trivial energy dependence of the relaxation time. Interestingly, both the magnitude and the temperature dependence of the Hall number in $La_{1.7}Sr_{0.3}CuO_4$ differ from what is expected in the isotropic Drude–Boltzmann picture, as recently pointed out by Narduzzo *et al* [68]. They have shown that this departure can be explained by invoking a strong in-plane anisotropic scattering. It would be interesting to explore the consequences of such anisotropy for the Nernst response.

5. Overall picture

Figure 4 displays the magnitude of the low-temperature Nernst coefficient divided by temperature as a function of the ratio of mobility divided by the Fermi energy. The low-temperature Nernst coefficient in bismuth is 10^6 times larger than in $Pr_{1.79}Ce_{0.21}CuO_4$ and in between lie the scattered data for other metals. Most of these metals are multi-band systems with different pockets of Fermi surface and the Fermi energy and the mobility varies among the bands. Some (such as $(TMTSF)_2ClO_4$) are very anisotropic with a hierarchy of energy scales, making the notion of a single Fermi energy questionable. Given all these uncertainties on the precise magnitude of $\frac{\mu}{\epsilon_F}$, it is remarkable that, as seen in the figure, the data points scatter around the red line expressing equation (10) ($\frac{\nu}{T} = 283\frac{\mu}{\epsilon_F}$).

The main message here is that the order of magnitude of the Nernst response in the zero-temperature regime is in agreement with the expectations of the semi-classical theory even when the ratio of mobility to Fermi energy is changed by six orders of magnitude. Note also that in several cases the discrepancy points to the unsatisfactory determination of the Fermi energy. In $CeCoIn_5$, for example, there is a large uncertainty on the magnitude of the Fermi energy. In $La_{1.7}Sr_{0.3}CuO_4$, the Fermi energy extracted from the slope of thermopower or the coefficient A, the prefactor of inelastic resistivity ($\rho = \rho_0 + AT^2$) would yield a smaller Fermi energy, reducing the discrepancy seen in figure 4. At this stage, it is fair to conclude that the ratio of mobility to Fermi energy is an adequate measure for the expected order of magnitude of the Nernst response of a Fermi liquid in the zero-temperature limit.

6. Nernst effect and quantum criticality

The Fermi energy, broadly taken as the main energy scale of the Fermi liquid vanishes in the vicinity of the quantum critical point (QCP). If the Nernst response is inversely proportional to the Fermi energy, it should be enhanced in the vicinity of a QCP. This has been confirmed by Izawa and co-workers [28], who studied the thermoelectric response of $CeCoIn_5$ in the vicinity of the field-induced QCP at 5.2 T. The case is well-documented, so that the relationship between the field dependence of ν/T and ϵ_F can be checked by looking at other probes of the Fermi energy, which are the prefactor of T^2 resistivity [46] and, γ, the electronic specific heat [45]. Moreover, the direct observation of the Fermi surfaces with de Haas–van Alphen studies allows us to quantitatively link the drop in the Fermi energy and the mass enhancement of the heaviest detected band [48]. The field dependences of three quantities which track the Fermi energy, that is ν/T, $A^{1/2}$, and γ are almost identical [28]. The available data for the Nernst coefficient and resistivity suggests that the Fermi energy increases by a factor of five between 16 and 6 T. (Note that there are no low-temperature specific heat data for $B > 9$ T).

It is interesting to complement these data with the field dependence of the zero-temperature Hall coefficient and the effective mass. The field dependence of the Hall number

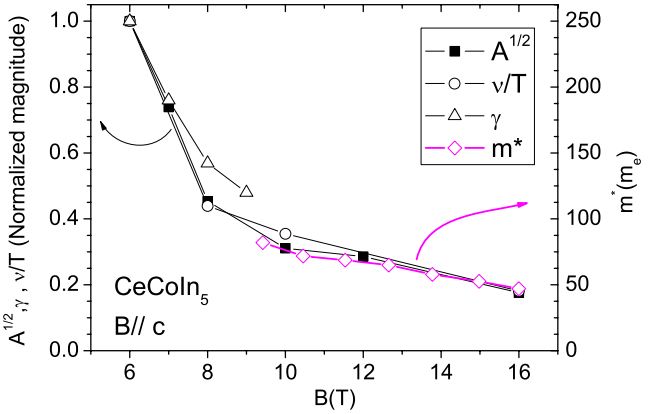

Figure 5. Quantum criticality in $CeCoIn_5$. The field variation of the Nernst coefficient [28], the prefactor of the T^2-resistivity [46], the electronic specific heat [45], and the effective mass of the β_1-band detected by dHvA measurements [48]. Apart from the effective mass, the magnitude of the other quantities is normalized by their value at 6 T, for the sake of comparison.

close to QCP is small (less than 10% between 6 and 8 T) [69]. Moreover, its magnitude is comparable with the Hall coefficient in the parent La-based compound [47]. This suggests that the volume of the large Fermi surface in $CeCoIn_5$ does not change near the QCP. On the other hand, the variation of the effective mass of the β_1-band (as seen by de Haas–Van Alphen studies) between 9 and 16 T is in quantitative agreement with the decrease in the Fermi energy implied by the field variation of ν/T and $A^{1/2}$ (see figure 5). The overall picture drawn by these sets of data is thus the following: the QCP in $CeCoIn_5$ is associated with an exploding mass without any notable change in the volume of the Fermi surface. This should present an important constraint for theoretical scenarios.

In the context of a possible link between quantum criticality and an enhanced Nernst signal, let us note that Li and Greene have reported a maximum in the doping dependence of the Nernst signal in $Pr_{1-x}Ce_xCuO_4$ close to $x = 0.15$ [64]. Since, this is close to the critical doping level as evidenced by Hall measurements [65] and backed up by thermopower data [63], it is tempting to speculate on the possible link between quantum criticality and this maximum. The absence of data for the critical doping make a definite judgment difficult (see figure 6). Note, however, that if there is a QCP, it is not driven by the same mechanism as in $CeCoIn_5$, since it is accompanied with a drastic change in the Fermi surface topology, as suggested by the abrupt change in the Hall number. Independent of microscopic details, the link between a maximum in the Nernst response and a drastic change in the Hall number is naturally explained by equation (6). At the critical doping, the Hall angle is extremely sensitive to any shift in chemical potential and therefore a maximum in the Nernst response is expected.

7. Conclusions and open questions

The main conclusion of this short review is that the so-called Sondheimer cancelation does not imply a zero Nernst signal

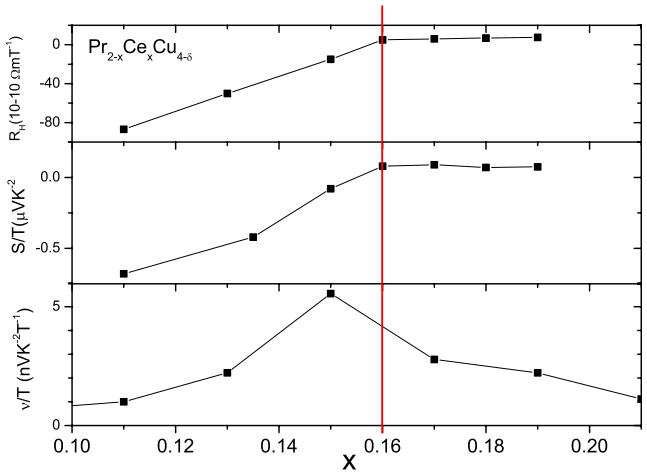

Figure 6. Quantum criticality in $Pr_{1-x}Ce_xCuO_4$. The doping dependence of the Hall coefficient, measured in the presence of a field exceeding the upper critical field and at 0.35 K [65], the low-temperature slope of the Seebeck coefficient [63], and the Nernst coefficient [64]. The existence of a QCP at $x = 0.16$ was deduced from the abrupt change in the Hall number and the thermoelectric response. Nernst coefficient peaks at $x = 0.15$, but the absence of a data point at $x = 0.16$ makes it difficult to definitely link this peak to quantum criticality.

in a Fermi liquid. The order of magnitude of the Nernst signal should become larger as the mobility increases and as the Fermi energy decreases. This statement is backed by an examination of available data on different families of remarkable metals. It appears that this correlation between the magnitude of the Nernst effect and the Fermi energy remains valid even near a quantum critical point making the Nernst effect a powerful probe of quantum criticality.

It is important to underline the limits of this simple picture. For example, the large quantum oscillations of the Nernst effect at low Landau levels observed in bismuth [70, 71] remain unexplainable in this approach. Explaining why the quantum oscillations in the Nernst response are much larger than quantum oscillations of conductivity (the Shubnikov–de Haas effect) remains a challenge to the theory.

The Nernst effect has proved to be a very powerful probe of superconducting fluctuations, a subject not addressed by this review. While, there is a satisfactory experimental confirmation [30] of the theory for the Nernst signal of Gaussian fluctuations [40], the issue of the Nernst effect generated by short-lived vortices (or superconducting phase fluctuations) in superconductors with small phase stiffness remains unsettled.

In the underdoped cuprates, even if one assumes that they are Fermi liquids, the analysis of the Nernst data is particularly complicated due to the possible role of three distinct sources of the Nernst signal: normal quasi-particles, short-lived Cooper pairs (amplitude fluctuations), and short-lived vortices (phase fluctuations). The recent detection of a small Fermi surface pocket in underdoped cuprates [72] dramatically modifies our picture of the normal state of these materials. Due to its small Fermi energy, the size of the Nernst signal generated by this pocket could be comparable with what is expected due to superconducting fluctuations.

Acknowledgments

I would like to thank H Aubin, R Bel, C Capan, J Flouquet, K Izawa, H Jin, I Sheikin, Y Nakajima, Y Matsuda, M-A Méasson, A Pourret, C Proust, and P Spathis, for their precious collaboration during the experimental investigation of the Nernst effect in metals and superconductors in the past few years. I am grateful to R Greene for his critical reading of the manuscript. I would also like to acknowledge the hospitality of the University of Campinas, Brazil, where this review was partly written. This work was supported by Agence Nationale de la Recherche in France and FAPESP in Brazil.

References

[1] Xu Z A, Ong N P, Wang Y, Kakeshita T and Uchida S 2000 *Nature* **406** 486
[2] Huebener R P and Seher A 1969 *Phys. Rev.* **181** 710
[3] Palstra T T M, Batlogg B, Schneemeyer L F and Waszczak J V 1990 *Phys. Rev. Lett.* **64** 3090
[4] Hagen S J, Lobb C J, Greene R L, Forrester M G and Talvacchio J 1990 *Phys. Rev.* B **42** 6777
[5] Hagen S J, Xu X Q, Peng J L, Li Z Y, Jiang W and Greene R L 1991 *Physica* C **185–189** 1275
[6] Ri H-C, Gross R, Gollnik F, Beck A, Huebener R P, Wagner P and Adrian H 1994 *Phys. Rev.* B **50** 3312
[7] Huebner R P 1979 *Magnetic Flux Structures in Superconductors* (Berlin: Springer)
[8] Delves R T 1965 *Rep. Prog. Phys.* **28** 249
[9] Clayhold J A, Linnen A W, Chen F and Chu C W 1994 *Phys. Rev.* B **50** 4252
[10] Wang Y, Xu Z A, Kakeshita T, Uchida S, Ono S, Ando Y and Ong N P 2001 *Phys. Rev.* B **64** 224519
[11] Wang Y, Ong N P, Xu Z A, Kakeshita T, Uchida S, Bonn D A, Liang R and Hardy W N 2002 *Phys. Rev. Lett.* **88** 257003
[12] Wang Y, Ono S, Onose Y, Gu G, Ando Y, Tokura Y, Uchida S and Ong N P 2003 *Science* **299** 86
[13] Capan C, Behnia K, Hinderer J, Jansen A G M, Lang W, Marcenat C, Marin C and Flouquet J 2002 *Phys. Rev. Lett.* **88** 056601
[14] Capan C, Behnia K, Li Z Z, Raffy H and Marin C 2003 *Phys. Rev.* B **67** 100507
[15] Capan C and Behnia K 2005 *Phys. Rev. Lett.* **95** 259703
[16] Wen H H *et al* 2003 *Europhys. Lett.* **63** 583
[17] Rullier-Albenque F, Tourbot R, Alloul H, Lejay P, Colson D and Forget A 2006 *Phys. Rev. Lett.* **96** 067002
[18] Wang Y, Li L and Ong N P 2006 *Phys. Rev.* B **73** 024510
[19] Wu W, Lee I J and Chaikin P M 2003 *Phys. Rev. Lett.* **91** 056601
[20] Choi E S, Brooks J S, Kang H, Jo Y J and Kang W 2005 *Phys. Rev. Lett.* **95** 187001
[21] Wu W, Ong N P and Chaikin P M 2005 *Phys. Rev.* B **72** 235116
[22] Nam M S, Ardavan A, Wu W and Chaikin P M 2006 *Phys. Rev.* B **74** 073105
[23] Nam M-S, Ardavan A, Blundell S J and Schlueter J A 2007 *Nature* **449** 584
[24] Bel R, Behnia K, Nakajima Y, Izawa K, Matsuda Y, Shishido H, Settai R and Onuki Y 2004 *Phys. Rev. Lett.* **92** 217002
[25] Bel R, Jin H, Behnia K, Flouquet J and Lejay P 2004 *Phys. Rev.* B **70** 220501(R)

[26] Sheikin I, Jin H, Bel R, Behnia K, Proust C, Matsuda Y, Aoki D and Onuki Y 2006 *Phys. Rev. Lett.* **96** 077207
[27] Onose Y, Li L, Petrovic C and Ong N P 2007 *Europhys. Lett.* **79** 17006
[28] Izawa K, Behnia K, Matsuda Y, Shishido H, Settai R, Onuki Y and Flouquet J 2007 *Phys. Rev. Lett.* **99** 147005
[29] Bel R, Behnia K and Berger H 2003 *Phys. Rev. Lett.* **91** 066602
[30] Pourret A, Aubin H, Lesueur J, Marrache-Kikuchi C A, Berg L, Dumoulin L and Behnia K 2006 *Nat. Phys.* **2** 683
[31] Pourret A, Aubin H, Lesueur J, Marrache-Kikuchi C A, Berg L, Dumoulin L and Behnia K 2007 *Phys. Rev.* B **76** 214504
[32] Spathis P, Aubin H, Pourret A and Behnia K 2008 *Eur. Phys. Lett.* **83** 57005
[33] Pourret A, Behnia K, Kikuchi D, Aoki Y, Sugawara H and Sato H 2006 *Phys. Rev. Lett.* **96** 176402
[34] Behnia K, Méasson M-A and Kopelevich Y 2007 *Phys. Rev. Lett.* **98** 076603
[35] Korenblit I Ya, Kusnetsov M E and Shalyt S S 1969 *JETP* **29** 4
[36] Mangez J H, Issi J P and Heremans J 1976 *Phys. Rev.* B **14** 4381
[37] Ettingshausen A V and Nernst W 1886 *Wied. Ann.* **29** 343
[38] Sondheimer E H 1948 *Proc. R. Soc.* A **193** 484
[39] Oganesyan V and Ussishkin I 2004 *Phys. Rev.* B **70** 054503
[40] Ussishkin I, Sondhi S L and Huse D A 2002 *Phys. Rev. Lett.* **89** 287001
[41] Sugihara K 1969 *J. Phys. Soc. Japan* **27** 362
[42] Nolas G S, Sharp J and Goldsmid H J 2001 *Thermoelectrics* (Berlin: Springer)
[43] Gollnik F and Naito M 1998 *Phys. Rev.* B **58** 11734
[44] Behnia K, Jaccard D and Flouquet J 2004 *J. Phys.: Condens. Matter* **16** 5187
[45] Bianchi A, Movshovich R, Vekhter I, Pagliuso P G and Sarrao J L 2003 *Phys. Rev. Lett.* **91** 257001
[46] Paglione J, Tanatar M A, Hawthorn D G, Boaknin E, Hill R W, Ronning F, Sutherland M, Taillefer L, Petrovic C and Canfield P C 2003 *Phys. Rev. Lett.* **91** 246405
[47] Nakajima Y, Izawa K, Matsuda Y, Behnia K, Kontani H, Hedo M, Uwatoko Y, Matsumoto T, Shishido H, Settai R and Onuki Y 2006 *J. Phys. Soc. Japan* **95** 023705
[48] Settai R, Shishido H, Ikeda S, Murakawa Y, Nakashima M, Aoki D, Haga Y, Harima H and Onuki Y 2001 *J. Phys.: Condens. Matter* **13** L627
[49] Amato A, Jaccard D, Sierro J, Lapierre F, Haen P, Lejay P and Flouquet J 1988 *J. Magn. Magn. Mater.* **76/77** 263
[50] Lacerda A, de Visser A, Haen P, Lejay P and Flouquet J 1989 *Phys. Rev.* B **40** 8759
[51] Bel R 2004 *PhD Thesis* University Paris VI, unpublished
[52] Ohkuni H *et al* 1999 *Phil. Mag.* B **79** 1045
[53] Sugawara H, Matsuda T D, Abe K, Aoki Y, Sato H, Nojiri S, Inada Y, Settai R and Onuki Y 2002 *Phys. Rev.* B **66** 134411
[54] Kasahara Y, Iwasawa T, Shishido H, Shibauchi T, Behnia K, Haga Y, Matsuda T D, Onuki Y, Sigrist M and Matsuda Y 2007 *Phys. Rev. Lett.* **99** 116402
[55] Aoki Y, Namiki T, Matsuda T D, Abe K, Sugawara H and Sato H 2002 *Phys. Rev.* B **65** 064446
[56] Danner G M, Kang W and Chaikin P M 1994 *Phys. Rev. Lett.* **72** 3714
[57] Smith G E, Baraff G A and Rowell J M 1964 *Phys. Rev.* **135** A1118
[58] Hartman R 1969 *Phys. Rev.* **181** 1070
[59] Bhargava R N 1967 *Phys. Rev.* **156** 785
[60] Jiang Wu, Mao S N, Xi X X, Jiang Xiuguang, Peng J L, Venkatesan T, Lobb C J and Greene R L 1994 *Phys. Rev. Lett.* **73** 1291
[61] Fournier P, Jiang X, Jiang W, Mao S N, Venkatesan T, Lobb C J and Greene R L 1997 *Phys. Rev.* B **56** 14149
[62] Balci H, Hill C P, Qazilbash M M and Greene R L 2003 *Phys. Rev.* B **68** 054520
[63] Li P and Greene R L 2007 *Phys. Rev.* B **76** 174512
[64] Li P, Behnia K and Greene R L 2007 *Phys. Rev.* B **75** 020506(R)
[65] Dagan Y, Qazilbash M M, Hill C P, Kulkarni V N and Greene R L 2004 *Phys. Rev. Lett.* **92** 167001
[66] Jin H, Narduzzo A, Nohara M, Takagi H, Hussey N E and Behnia K 2009 unpublished
[67] Nakamae S, Behnia K, Mangkorntong N, Nohara M, Takagi H, Yates S J C and Hussey N E 2003 *Phys. Rev.* B **68** 100502
[68] Narduzzo A, Albert G, French M M J, Mangkorntong N, Nohara M, Takagi H and Hussey N E 2008 *Phys. Rev.* B **77** 220502
[69] Singh S, Capan C, Nicklas M, Rams M, Gladun A, Lee H, DiTusa J F, Fisk Z, Steglich F and Wirth S 2007 *Phys. Rev. Lett.* **98** 057001
[70] Behnia K, Méasson M-A and Kopelevich Y 2007 *Phys. Rev. Lett.* **98** 166602
[71] Behnia K, Balocas L and Kopelevich Y 2007 *Science* **317** 1729
[72] LeBoeuf D, Doiron-Leyraud N, Levallois J, Daou R, Bonnemaison J-B, Hussey N E, Balicas L, Ramshaw B J, Liang R, Bonn D A, Hardy W N, Adachi S, Proust C and Taillefer L 2007 *Nature* **450** 533