Conductance determined by transmission: probes and quantised constriction resistance

1989 J. Phys.: Condens. Matter 1 8099

(http://iopscience.iop.org/0953-8984/1/43/011)

J. Phys.: Condens. Matter 1 (1989) 8099–8110. Printed in the UK

Conductance determined by transmission: probes and quantised constriction resistance

R Landauer
IBM Research Division, T J Watson Research Center, PO Box 218, Yorktown Heights, New York 10598, USA

Received 23 February 1989

Abstract. The domain of validity and the origin of different formulae relating conductance to transmissive behaviour is discussed. The *large* variety of possible measurement probes is emphasised. The quantised conductance of constrictions, studied in recent experiments, is treated. Theoretical discussions of that have stressed the supposedly unique role of $G = (2e^2/h)\mathrm{Tr}(tt^\dagger)$ in this connection, regardless of the exact geometry or probe locations. This is shown to be incorrect, and a simple expression for the dependence of the 'quantised' conductance on the width of the wide region is given. We particularly emphasise the distinction between, on the one hand, experiments and discussions that are *strictly* limited to events at reservoirs and, on the other, those that have a concern with variations within the sample.

1. Introduction

Calculations of conductance of a sample, as a function of its transmissive behaviour, became fashionable following the work of Anderson *et al* [1]. The method was originally applied to discussions of localisation and subsequently to mesoscopic samples [2–6]. There are a number of alternative versions of this conductance formula, and valid reasons for the existence of various forms. First of all, a sample does not have a unique resistance. Its resistance depends on the way the current is introduced; how are the arriving carriers spread out both in momentum space and in real space [7, 8]. While the purely geometrical effects of the connections are obvious, the dependence on the arriving momentum distribution has received little attention, except for this author's discussions. A more widely appreciated reason for the existence of variation in conductance results from the variability in the choice of voltage used in the calculation of conductance. Where, in space, is the voltage measured? Is it measured by probes that are *invasive*, i.e. have an effect on the original transport system, or not? Are we measuring voltage or measuring electrochemical potentials? We will discuss these questions in some detail. In a final section we analyse the impressive recent work on quantised construction resistance in two-dimensional electron systems [9, 10].

2. Is $G \sim T$ or $G \sim T/(1 - T)$?

Consider the situation depicted in figure 1. The *sample* is, at least in the simplest versions of these theories, taken to be an elastic scatterer. The two reservoirs are taken to be at

0953-8984/89/438099 + 12 $02.50

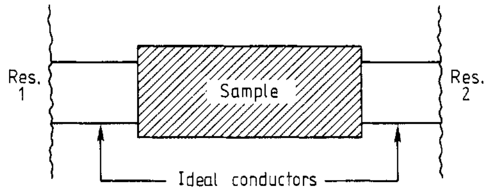

Figure 1. Standard geometry for calculation of resistance from the scattering matrix of the sample.

differing electrochemical potentials. The sample is connected to the reservoirs through ideal, i.e. ballistic, conductors. The voltage used in defining the resistance that we discuss first is that in the ideal conductors. Actually, in most cases, it makes no difference whether we use voltage or electrochemical potentials; they can differ appreciably only over regions shorter than a screening length. Furthermore, we assume that voltage or electrochemical potential can be measured by non-invasive methods that do not perturb the existing value. In that case, for a one-dimensional sample the conductance G is found to be [1, 2, 7]

$$G = (e^2/\pi\hbar)T/R \tag{2.1}$$

where T and R are the transmission and reflection probabilities of the sample. In equation (2.1), and throughout this paper, all equations are in a form allowing for a two-fold spin degeneracy, and without allowing spin to be a significant variable bearing on the scattering process. In the ideal conductor to the left of the sample there will be an energy range, defined by the respective chemical potentials μ_1 and μ_2 of the reservoirs, in which the current is carried. In that range the waves incident from the reservoir will interfere with the reflected waves, forming a standing-wave pattern. These interference effects are particularly familiar from electromigration theory [11]. The voltage that leads to equation (2.1) is that obtained in the ideal conductor after averaging over these oscillations; it does not attempt to follow the rapid variations within a Fermi wavelength.

The multi-dimensional generalisation of equation (2.1) was provided by Azbel [12] and discussed in detail in [13]. That result is

$$G = \frac{2e^2}{\pi\hbar}\frac{(\Sigma_i T_i)(\Sigma_i v_i^{-1})}{\Sigma_i v_i^{-1}(1 + R_i - T_i)}. \tag{2.2}$$

Here i refers to the channel number in the ideal conductor, and v_i is the longitudinal velocity of the corresponding channel in the direction of current flow. The channels, numbered by i, can be associated with the different transverse eigenstates in the ideal conductor. Equation (2.2) assumes that the two ideal conductors are identical, but more general results are presented in [13]. Equation (2.2) has singularities arising from the reciprocal velocities, when one of the transverse channel energies is exceptionally close to the Fermi energy. These velocities also reflect the longitudinal density of states, inversely proportional to the velocity v_i. In real samples the singularity in the density of states is likely to be smeared out due to the finite length of the ideal conductor and/or to the appearance of other length scales not present in analysis of [13]. Furthermore, a little unintended residual scattering, or imperfect control of the width of the conductor, will also provide such smearing. We shall invoke such a smearing in § 5.

Recent applications of equation (2.1) were provided by [14] and [15].

Shortly after Anderson *et al* [1] drew widespread attention to equation (2.1), a number of papers appeared arguing that conductance was proportional to transmission

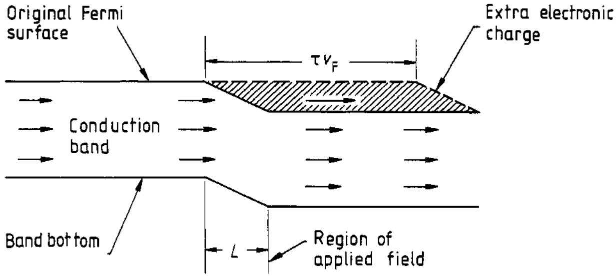

Figure 2. Energy distribution of electrons in a perfect metallic conductor, with a field applied over a region of length L (from [20]). Only electrons with a positive velocity are shown. τ is the time during which excess electrons have moved to the right.

probability, and that the expressions with more complicated denominators were incorrect. These papers, as well as the rebuttals to them, are cited in [16] and [17]. The latter [17] is, in fact, closely allied to the viewpoint of the earlier papers advocating $G \sim T$. A conductance proportional to transmission probability has been familiar for decades through the concern with tunnelling barriers [18]. For $T \ll 1$, of course, the denominator in equation (2.1) can be taken to be unity, and the distinction between the two forms becomes unimportant. The availability, however, of conductors with good transmission in which one can measure, for example, the resistance of a corner in a wire [19] has made the distinction significant.

Figure 2, [20], shows somewhat symbolically what must happen according to the papers advocating $G \sim T$, when we assume a voltage drop across a region of space that transmits perfectly and is connected to leads of the same width. Extra carriers are transmitted through the region assumed to have perfect transmission. If all the extra carriers incident from the left, in the range of bias, are transmitted, then we find $G = 2e^2/h$. This is the result implied by $G = 2e^2T/h$, for $T = 1$. By contrast, equation (2.1) yields an infinite conductance, for $T = 1$. The assumptions depicted in figure, 1 however, do not make sense. The transmitted carriers in the shaded region have a large space charge, and this simply cannot be waved aside. The equation $G = 2Te^2/h$, together with its multi-dimensional generalisations, has turned out to have a very significant range of applicability. That, however, is fortuitous; the early proposals for this expression simply ignored the space-charge questions, and did not invoke reservoirs, whose kinetics is essential for the proper applicability of the simpler result without the complicating denominators.

Eventually Imry [2], enlarging upon a discussion by Engquist and Anderson [21], pointed out that the simpler formulae did apply to the electrochemical potential difference, or the potential difference, if measured between points *way inside* the reservoirs on each side of the conductor. A reservoir is not just an extension of the sample or an extension of the ideal conductor; the physics of reservoirs has been discussed in § 7 of [8] and § 3 of [22]. We return to the physics of reservoirs later in this paper. Büttiker provided a proper treatment for the multi-dimensional conductor, with more than two attached reservoirs [23]. This has become an essential tool for interpreting the typical mesoscopic four-probe measurements, for understanding the quantum Hall effect, and for interpreting conduction in systems with inelasticity [24].

Imry's result [2] is

$$G = (e^2/\pi\hbar)\mathrm{Tr}(tt^\dagger) \tag{2.3}$$

where t is the transmission matrix specifying the transmitted wavefunctions relative to the incident waves. The trace in equation (2.3) represents a sum over the channels of an ideal conductor. The electrochemical potential difference, however, invoked in the conductance of equation (2.3) is *not* that between the ideal conductors, but is the electrochemical potential difference between the reservoirs of figure 1. Deep inside the reservoirs we have an equilibrium distribution of carriers, without current flow. It is that electrochemical potential difference which yields equation (2.3). Note, however, that *the transmission matrix of equation (2.3) characterises transmission between the ideal conductors of figure 1, and not transmission between reservoirs.* Thus, equation (2.3) cannot possibly be applicable if the interface between the reservoir and the ideal conductor adds additional scattering, not taken into account in equation (2.3). States coming out of the ideal conductor, towards the adjacent reservoir, must be able to enter that reservoir without further reflection. This requires a gently flared horn-like connection between the ideal conductor and the reservoir, or other 'matching' devices, as discussed in connection with figure 6 of [8]. Unfortunately, the rapidly growing body of theoretical literature that has appeared, attempting to interpret the experimental results of [9, 10], has not been sensitive to this distinction between ideal conductors and reservoirs, nor to the shape of the interface between these. We return to that subject, subsequently.

3. Probes

There is now a prevailing presumption in this field that only electrochemical potentials, rather than electrostatic potentials, can be measured, and that they have to be measured by reservoirs connected through leads that are similar to the conducting sample and perturb the conduction process. This is the type of configuration analysed in [23]. The prevailing assumption is best characterised by a quotation from Stone [25]: '. . . it is unreasonable to suppose that the voltage drop across some region of a small sample can be measured without the presence of the voltage probes strongly affecting the results of the measurement.' This *presumption* for the needs of strongly perturbing leads is remarkable; normally in physics we look for minimally invasive methods of measurement. There are alternative voltage measurement methods available, and we summarise these below.

First we emphasise that we are not limited to measuring electrochemical potentials; voltages can be measured. This is clear in the case of contact potential differences. Two connected dissimilar metals, in thermal equilibrium, have no electrochemical potential difference. But they do have an electrostatic potential difference, which can be measured, e.g. through phase-locked techniques, with oscillating capacitive probes. Capacitive probes have also been used in more modern contexts [26–28]. Additionally, capacitive probes have been used in AC transport measurements [29]. We could measure transport voltage differences by measuring the alternating current between two connected probes, both capacitively coupled to the sample, in the presence of AC transport through the sample. What is the back effect of these probes on the transport? We note that the voltage drop within a conductor, at sufficiently low frequencies, is determined by the requirement for continuity of current. Externally applied *electrostatic* fields (i.e. through capacitive coupling) only have an effect on transport through the field effect, i.e. through the modulation of carrier density at the conductor surface. Furthermore, electrostatic fields that are themselves proportional to the transport current, as can be the case for capacitive measurement probes, only have an effect on the transport field distribution *within the conductor*, which is proportional to the square of the current flow. They will not affect the first-order (linear) behaviour.

Electric fields can, in principle, be measured or characterised in a great many other ways. In addition to the traditional hypothetical test charge, there is nuclear quadrupole resonance or the deflection (outside the sample) of electron beams. Electromigration is a complex and indirect way of probing electric fields, but unlikely to occur at very low temperatures [11]. The internal voltage distribution, in the sample, has an effect on the non-linear behaviour. See, for example, the discussion in [30]. Experiments on non-linearity [31–33] have been done, and many more are likely to come.

Can capacitive experiments have high spatial resolution? Consider a scanning tunnelling microscope (STM) probe, used capacitively. The problem with that, if used in the most obvious way, is that it it not a high-resolution probe. The capacitance will not just be determined by the end of the tip where tunnelling occurs, but will be dominated by the fringing fields. There is a way out, however, suggested by the work of van Bentum *et al* [34]. They observe the Coulomb blockade associated with small tunnelling capacitances, in an STM geometry. Why? Because only that part of the capacitance counts which can deliver charge to the tip, within the time taken by the electron to tunnel. To evaluate that capacitance requires not only an understanding of traversal time in tunnelling, but also of the charge propagation process. Does it only involve surface plasmons, or also electromagnetic wave propagation effects? The results of [34], however, do show that the effective capacitance is limited, and well below the static capacitance. The Coulomb blockade is, of course, a tunnelling effect. The onset of tunnelling is, however, determined by capacitive effects.

The above schemes will only work for metallic samples with an accessible surface. For the semiconductor field-effect geometry there may be a more practical scheme. There we can use the non-linear capacitance of measurement gates to measure the voltage in the underlying channel, or ballistic conductor. The potential disturbance within that conductor can be kept within a range, and be made smooth enough, so as to cause no additional scattering. Or, we can tie two otherwise floating gates together and measure the current between them. The conductance of equations (2.1) and (2.2) leads to a key point: *When there is no intervening barrier, there should be no voltage drop. A perfect conductor has no resistance!* The configuration consisting of two successive measurement gates over the same ballistic channel should be particularly simple. More generally, at two different positions along an ideal conductor (far enough away from reservoirs or obstacles so that evanescent modes are not present), all quantities of interest such as carrier populations, or potentials, *must* be the same. *Non-invasive methods of measurement must yield the same potential at the two points.* Only if we insist on counting the carriers moving to the right at the left end, and the carriers moving to the left at the right end, do we find the electrochemical potential difference found in [17]. But there is no reason for such a procedure.

Probes, of course, can also measure the carrier population, e.g. by attempts to have a small reservoir whose electrochemical potential adjusts to make the net current between the sample and the reservoir vanish [23]. We will not discuss such probes in detail comparable to that just devoted to capacitive probes. We do, however, stress that the probes can be very loosely coupled, as in the case of STM voltage measurements [35, 36]. Probes coupled to the electron population in the sample do not need to be invasive; they do not need to have an appreciable effect on the underlying transport process. Even though the probes are not invasive, that does not guarantee that the measurements will be easy to interpret. The probes can still couple to the electron motion in the sample in a way that depends delicately on the multiple interference between scattering centres in the sample and in the probes [23, 37]. On the other hand, we do not need to sanctify the currently and readily available experimental situation as that which is inevitable. We repeat from [22]:

> Note that if we are concerned with measurements on an ideal conductor, inserted between a sample reservoir, then we can (in principle, probably not in reality) use a whole array of loosely identical probes to achieve a variety of measurements. We can, for example, use the equivalent of a phased array to measure the carriers present in a particular 'channel', moving in a specified direction (toward reservoir or toward sample). Alternatively, we can use a random array to eliminate the effect of interference oscillations, and thereby measure an average carrier population in the ideal conductor.

4. Central issue

We now emphasise our most central point. We can, on the one hand, focus on experiments where we are only concerned with events at reservoirs. Thus, for example, the current comes from one reservoir and flows to another. Voltages are measured at additional reservoirs. In that case the formalism of [23] applies. Alternatively, we may have a concern with the voltage at some point away from a reservoir. Once we correct voltage-probe measurements for a series resistance, as is done in [9] and [10], we are asking a question of the latter type, regardless of the presumed simplicity of the correction. (Actually, [10] presents raw data in its figures; the series-resistance correction arises in the supplementary discussion.) Spatial variation of the self-consistent voltage distribution must be understood through some variation of the methods employed in my papers [22, 38]. It *cannot* be done without recourse to Poisson's equation, space-charge neutrality, a frequency- and wavenumber-dependent dielectric constant for the electron gas, or some other explicit way of coping with screening. Nothing of that sort turns up in [17] and [25]. Papers that ignore Poisson's equation, or papers that go to the even further extent of explicitly stating that it is not necessary to pay attention to such questions [17, 39], may be correct within their declared purpose, but cannot possibly have any bearing on the many questions related to the spatial features of the voltage drop within the system.

If we want to understand the voltage division in a sample that includes elastic scattering and is attached to leads that include *inelastic* scattering, then equation (2.3) and its multi-terminal generalisations are useless. We face a problem that has not been analysed in detail. But it is susceptible to the methods reviewed in [22]. We follow the carriers out of their reservoirs, and through the sample, and then find a voltage division through the subsequent self-consistent screening of these carriers. The voltage drop across the central 'sample' with purely elastic scattering will not correspond to any of the explicit results displayed in § 2, because the carrier distribution incident on the sample is, in part, determined by the kinetics in the leads. The kind of difficulty is emphasised in [7] and in my other papers.

5. Conduction through narrow apertures

Recent papers [9, 10] have drawn attention to the dramatic nature of the conductance steps encountered as the width of a narrow channel is varied. This effect could have been anticipated on the basis of a number of theoretical discussions, but nevertheless came unexpectedly, For example, [13] states:

> For a system where the Fermi energy E_F can be changed, such as in metal–oxide–semiconductor device, E_F may be made to cross a particular transverse level and to switch on or off its contribution to the conductance. This will lead to sharp changes in the conductance near these crossings.

But the accompanying description in [13] is too flawed to make this a serious candidate as an anticipation. The sentences following equation (4.19) of [2] come closer. The clearest anticipation came from those concerned with an STM geometry, particularly in a conference paper by Garcia in the summer of 1987, which displayed the repeated steps in the conductance [40]. After the experiments on two-dimensional electron gas (2DEG) geometries were reported [9, 10], a great many additional calculations appeared. We only relate to papers that analyse the conductance as a function of transmission, calculated for a geometry that includes a narrow region separating two wider ballistic conductors. Our comments bear little or no relationship to other theoretical discussions [41–43].

At the time this is being written the author is aware of six efforts that explain the quantised resistances observed in [9] and [10] via a transmission probability calculation of the type alluded to above [44–49]. Undoubtedly there will be more entries, and possibly some of the six that are in preprint form will change on their way into print. As a result we minimise attention to the individual nature of [44–49] and emphasise remarks that apply to several of them. These papers [44–49] calculate the transmissive behaviour of a narrow constriction using a variety of geometries, and a variety of analytical procedures. None of these papers consider the actual details of the potential variation in the electron channel, as caused by the controlling gate structure [50]. We have no arguments, at all, with quantum-mechanical details of [44–49] and their calculation of transmission matrices. All of our subsequent comments relate to the use of this information in the calculation of a conductance. References [44–49] invoke equation (2.3), or its equivalent. Considered as a collection, they appear like a clear vote in favour of equation (2.3), relative to equation (2.2). Indeed, [17] and [45] make this particularly explicit. For example, [17], after an illusion to [9], states: '. . . Therefore we conclude that Equation (4), including the two-probe and one-channel limits, is the physically relevant Landauer formula for all the present experiments.' Equation (4) in this citation refers to the generalisation of equation (2.3) to more than two connections [23]. None of the six cited theoretical contributions seem to be sensitive to the point made in our § 4. Additionally, none of them seem to be sensitive to the very elementary fact that we have taken up in § 2, that equation (2.2) gives the voltage between reservoirs, not between the ballistic conductors. While we have lumped the six citations, we should actually distinguish between those [44–46] which explicitly allow for ballistic conductors with a well defined width, as contrasted to those [47–49] which explicitly or implicitly assume that the constricted region occurs between ballistic conductors of unlimited width. The former [44–46] are the primary focus of our comments. The latter [47–49], which do not attempt to model the experiments in equal detail, are not actually in error. The 'error' is, admittedly, a very minor quantitative issue. It is the qualitative issue, the implicit suggestion (and more than implicit in [17] and [45]) that equation (2.3) is 'justified' by the constriction experiments, that is our concern.

We adopt the terminology of [45] and refer to the wide ballistic regions as W, and W for the actual width of that region. Similarly N and N refer to the narrow region, and

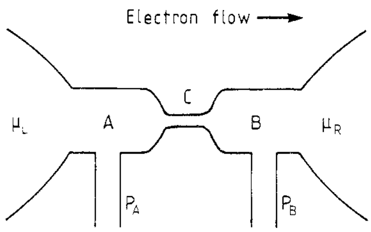

Figure 3. Two ballistic conductors with probes P_A and P_B separated by a narrow constriction C. Current connects to reservoirs at far left and right.

WNW describes the overall configuration of [9] and [10]. The actual geometry of [9] and [10] differ somewhat. Both papers, however, invoke corrections for presumed series resistances, and, therefore, our comments in § 4 apply. Figure 3 idealises the situation, but is closer to that described in [10] than that of [9].

Our discussion is aimed at a general point, applicable to the case, for example, where the W and N regions have an interface with sharp corners. But, for simplicity, we first invoke the specifics of an adibatic transition between the W and N regions. We assume that the narrow region ends in a tapered horn, and that all modes in the narrow section can pass on out to the wider region without reflection. We invoke a totally adiabatic horn which flares out gradually, until it covers the whole width of the wide area. In that case the lowest transverse mode in the wide section (no nodes in the transverse direction) maps onto the lowest mode in the constricted region. We also assume that the constriction is long enough so that evanescent modes excited by arriving waves at one end have a negligible transmission probability. We invoke equation (2.1), repeated here for convenience

$$G = \frac{2e^2}{\pi\hbar} \frac{(\Sigma_i T_i)(\Sigma_i v_i^{-1})}{\Sigma_i v_i^{-1}(1 + R_i - T_i)}. \tag{5.1}$$

The channels in the wide conductor are either reflected totally by the constriction, or transmitted totally. Then $\Sigma_i T_i$ counts the number of transmitted channels, i.e. the number of channels present in the narrow constriction. Call that number N_N. The other factor in the numerator, $\Sigma_i v_i^{-1}$, will be evaluated in detail. In the denominator only the channels that are reflected contribute. We have

$$\Sigma_i v_i^{-1}(1 + R_i - T_i) = 2\,(\Sigma_A v_i^{-1} - \Sigma_T v_i^{-1}) \tag{5.2}$$

where A represents summation over all channels and T over transmitted channels. The first term on the RHS in equation (5.2) is the same as the one that occurs in the numerator of equation (5.1). What about $\Sigma_T v_i^{-1}$? Let us assume $W \gg N$. Then only the channels with the lowest transverse energy (in the region) will be transformed adiabatically, via the 'horn', into the transmitted channels in the N region. These are channels which, in the W region, have motion almost parallel to the direction of current; v_i will be close to v_F. Therefore, we approximate $\Sigma_T v_i^{-1}$ by $N_N v_F^{-1}$, where v_F is the Fermi velocity. N_N is the number of channels (not counting spin) in the constriction.

Now let us face $\Sigma_A v_i^{-1}$. We will approximate the sum through an integral. This not only makes the calculation easier, it also provides the smearing discussed in § 2 and needed to eliminate the anomalous effects that result when one of the transverse energies is exceptionally close to the Fermi energy. The result is $Wm/2\hbar$, where m is the carrier mass and W the width of the wide region. Avishai [51] has kindly checked the validity of this approximation through a computer simulation. Our result can, equivalently, be written as $(\pi/2)N_W v_F^{-1}$, where N_W is the number of channels (not counting spin) in the wide region. Thus equation (5.1) becomes

$$G = \frac{2e^2}{\pi\hbar} N_N \frac{(\pi/2)N_W v_F^{-1}}{2[(\pi/2)N_W v_F^{-1} - N_N v_F^{-1}]} \tag{5.3}$$

or

$$G = \frac{e^2}{\pi\hbar} N_N \left(\frac{N_W}{N_W - (2/\pi)N_N}\right). \tag{5.4}$$

The $(e^2/\pi\hbar)N_N$ prefactor in equation (5.4) is the widely publicised result, based on $G \sim \mathrm{Tr}(tt^\dagger)$. The remaining correction is close to 1, if $N_W \gg N_N$, but is larger than unity. *This larger conductance is real*; it represents the fact that the W regions are not perfect reservoirs. The carriers, there, are not in equilibrium, but to a small extent reflect the electrochemical potential of the far-away reservoir, on the other side. The larger W is in comparison to N, the closer the final factor in equation (4) is to unity. Thus W/N is a measure of the quality of the W regions as approximate reservoirs.

Note that equation (5.4), which measures the quality of the W region as a reservoir for the current in the N region, did not need to invoke elastic or inelastic scattering. As we go further into a reservoir, the electronic motion must approach that characteristic of thermal equilibrium appropriate to the carriers emerging from the reservoir into the conducting sample. This can be achieved through simple geometrical dilution. As we go further into the reservoir, a higher proportion of the carriers present at a particular point originated far inside of the reservoir, rather than the sample. This author has, in the past, perhaps exaggerated the need for inelastic processes in the reservoir. Clearly we want the carriers coming out of the reservoir to be incoherent with those entering it. But if the reservoir is large enough, then for a long time that incoherence can be provided simply by the fact that the states coming out of the reservoir are not the states that entered from the sample. It is the geometrical dilution of incoming carriers that constitutes the most essential aspect of a reservoir.

Let us now continue with an alternative formulation of the notions just presented, but with wider applicability. We pay for the greater applicability through an expression for our result that has physical meaning, but is not readily calculated. Equation (2.3) gives a conductance for the situation in figure 3 if voltage is measured between reservoirs in figure 3. For voltage probes,as shown in figure 3, the conductance can be more complex than the experiments indicate [23]. Here we calculate the potential difference between the ballistic regions without attention to probes, using the concepts of [13] but without actually invoking equation (2.2). As stated, such an expression assumes spatial averaging to remove interference effects. We treat the diffusion of non-interacting carriers; allowing later for self-consistent screening will not change the drop in electrochemical potential. In lead A in figure 3 the electrochemical potential, reflecting carrier density there, will be below μ_L deep inside the left-hand reservoir. That reservoir has an equilibrium distribution, given by μ_L; this also characterises carriers arriving at A from the left. The carriers arriving at A from the right will, in part, have come through the constriction and their population will depend on μ_R. The current j, through A, is due to an imbalance between left-moving and right-moving carriers, and has the magnitude $n_d W e v_F \langle \cos\theta \rangle$. Here, W is the width of lead A, v_F the Fermi velocity, n_d the deficit in carrier density of the carriers from the right, relative to those from the left, and θ the angle of the carriers relative to the direction of current flow. $\langle\cos\theta\rangle$ denotes an average for the mix of channels in the flow. The electrochemical potential in lead A is

$$\mu_A = \mu_L - (d\mu/dn)n_d = \mu_L - (j/Wev_F\langle\cos\theta\rangle)\, d\mu/dn \tag{5.5}$$

with a similar expression for μ_B. For simplicity, we use the same W, $v_F\langle\cos\theta\rangle$ and $d\mu/dn$ at B as at A. Using $d\mu/dn = hv_F/2k$, where k is the Fermi wavenumber, and $kW/\pi = N_W$ as the number of transverse channels, we find

$$\Delta R = j^{-1}[(\mu_L - \mu_R) - (\mu_A - \mu_B)] = h/\pi e^2 N_W\langle\cos\theta\rangle. \tag{5.6}$$

Here ΔR denotes the resistance between reservoirs minus that between A and B. Let the conductance between reservoirs be denoted by G_R. The conductance between A and B is $G_{AB} = G_R(1 + G_R\Delta R)$, if ΔR is small compared to $1/G_R$. With equation (5.6)

$$G_{AB}/G_R = 1 + 2\,\mathrm{Tr}(tt^\dagger)/\pi N_W\langle\cos\theta\rangle. \tag{5.7}$$

In a quantised plateau $\mathrm{Tr}(tt^\dagger) = N_N$, where N_N is the channel count in the constriction. Thus, G_{AB}/G_R in equation (5.7) becomes $1 + (2N_N/\pi N_W)\langle\cos\theta\rangle^{-1}$. We see that the smallness of $(2N_N/\pi N_W)\langle\cos\theta\rangle^{-1}$ measures the quality of the ballistic regions as reservoirs. If the constriction ends in slowly tapered horns, then the transverse wavenumbers are reduced in passage out through the horn, and $\langle\cos\theta\rangle$ is close to 1. In that case G/G_R becomes $1 + (2N_N/\pi N_W)$, in agreement with equation (5.4). Thus, the conductances evaluated in [44–46] need the correction in equation (5.7). The peculiarities of invasive probes may cause that to be incorrect, but surely the potential *in the ballistic regions* will be more relevant than that in the reservoirs. The correction described by equation (5.7) may be unimportant quantitatively. We emphasise it to demonstrate the greater applicability of the views expressed in [13] relative to that of equation (2.3).

Note that for the density of states in equation (5.5) we used the simple Fermi–Thomas expression $hv_F/2k$, rather than one more delicately adjusted to the particular geometry at hand. This provides the same smearing discussed earlier. The precise validity of this approximation depends on the exact method of voltage or electrochemical potential measurement. It will also, of course, depend on the degree to which further residual scattering is really absent, and on the degree with which the transverse dimension W is really controlled.

6. Conclusions

The equation $G = (2e^2/h)\,\mathrm{Tr}(tt^\dagger)$ and its multi-terminal generalisation apply when we are *strictly* concerned with current sources and voltage measurements at reservoirs. Elastic or inelastic scattering along a conducting lead does not turn it into a better approximation of a reservoir than is allowed by the purely geometrical spreading. If we are dealing with the voltage distribution within a sample, e.g. by allowing for a series-resistance correction from the voltages measured at probes, then we must utilise theories that invoke Poisson's equation in their formulation. A theory of this sort [13] not only can be used to explain quantised constriction resistances, but is essential to allow us to understand the effects of the width of the wide regions on the measured results. Voltage probes, along a sample, need not be invasive and can be purely capacitive.

Acknowledgments

Interaction with most of the authors of the experimental and theoretical papers on quantised constriction resistance helped us to develop our argument, and is appreciated. These interactions have not necessarily produced complete agreement. I am particularly indebted to two groups (van Wees *et al* [9] and Wharam *et al* [10]) for clarification about their geometry and their use of series-resistance corrections.

Note added in proof. See [52] for a closely related viewpoint.

References

[1] Anderson P W, Thouless D J, Abrahams E and Fisher D S 1980 *Phys. Rev.* B **22** 3519
[2] Imry Y 1986 *Directions in Condensed Matter Physics* ed. G Grinstein and G Mazenko (Singapore: World Scientific) p 101
[3] Aronov A G and Sharvin Yu V 1987 *Rev. Mod. Phys.* **59** 755
[4] Webb R A and Washburn S 1988 *Phys. Today* **41** (12) 46
[5] Al'tshuler L and Lee P A 1988 *Phys. Today* **41** (12) 36
[6] Datta S 1989 *Quantum Phenomena* (Reading, Mass: Addison-Wesley)
[7] Landauer R 1985 *Localisation, Interaction and Transport Phenomena* vol. 61, ed. B Kramer, G Bergmann and Y Bruynseraede (Heidelberg: Springer) p 38
[8] Landauer R 1987 *Z. Phys.* B **68** 217
[9] van Wees B J, van Houten H, Beenakker C W J, Williamson J G, Kouwenhoven L P, van der Marel D and Foxon C T 1988 *Phys. Rev. Lett.* **60** 848
[10] Wharam D A, Thornton T J, Newbury R, Pepper M, Ahmed H,, Frost J E F, Hasko D G, Peacock D C, Ritchie D A and Jones G A C 1988 *J. Phys. C: Solid State Phys.* **21** L209
[11] Sorbello R S 1989 *Phys. Rev.* B **39** 4984
[12] Azbel M Ya 1981 *J. Phys. C: Solid State Phys.* **14** L225
[13] Büttiker M, Imry Y, Landauer R and Pinhas S 1985 *Phys. Rev.* B **31** 6207
[14] Gasparian V M, Altshuler B L, Aronov A G and Kasamanian Z A 1988 *Phys. Lett.* **132A** 201
[15] Heinrichs J, Jayannavar A M and Kumar N 1988 *Solid State Commun.* **67** 1103
[16] Landauer R 1981 *Phys. Lett.* **85A** 91
[17] Stone A D and Szafer A 1988 *IBM J. Res. Dev.* **32** 384
[18] Sommerfeld A and Bethe H 1933 *Handbuch der Physik* vol. 24, ed. A Smekal (Heidelberg: Springer) p 446
[19] Timp G, Barangar H U, de Vegvar P, Cunningham J E, Howard R E, Behringer R and Mankiewich P M 1988 *Phys. Rev. Lett.* **60** 2081
[20] Landauer R 1981 *Phys. Lett.* **85A** 91
[21] Engquist H-L and Anderson P W 1981 *Phys. Rev.* B **24** 1151
[22] Landauer R 1988 *IBM J. Res. Dev.* **32** 306
[23] Büttiker M 1988 *IBM J. Res. Dev.* **32** 317
[24] Datta S and McLennan M 1988 *Preprint*
[25] Stone A D 1988 *Physics and Technology of Submicron Structures* ed. H Heinrich, G Bauer and F Kuchar (Heidelberg: Springer) p 108
[26] Zeller R T, Fang F F, Goldberg B B, Wright S L and Stiles P J 1986 *Phys. Rev.* B **33** 1529
[27] Dabiran A M, Zeller R T, Fang F F, Wright S L and Stiles P J 1988 *Surf. Sci.* **196** 712
[28] Fang F F, Nocera J J, Luo J and Smith T P III 1988 *Proc. 19th Int. Conf. Physics of Semiconductors* ed. W Zawadski (Warsaw: Polish Academy of Sciences) p 169
[29] Smith T P and Stiles P J 1984 *Solid State Commun.* **52** 941
[30] Landauer R 1987 *Nonlinearity in Condensed Matter* ed. A R Bishop, D K Campbell, P Kumar and S E Trullinger (Heidelberg: Springer) p 2 (see § 4)
[31] Webb R A, Washburn S and Umbach C P 1988 *Phys. Rev.* B **37** 8455
[32] de Vegvar P, Timp G, Mankiewich P M, Cunningham J C, Behringer R and Howard R E *Phys. Rev.* B **38** 4326
[33] Kaplan S B and Hartstein A 1988 *IBM J. Res. Dev.* **32** 347

[34] van Bentum P J M, van de Leemput L E C, Smokers R T M and van Kempen H 1989 *Phys. Scr.* **T25** 122
[35] Kirtley J R, Washburn S and Brady M J 1988 *IBM J Res. Dev.* **32** 414
[36] Kirtley J, Washburn S and Brady M J 1988 *Phys. Rev. Lett.* **60** 1546
[37] Baranger H 1988 Presentation at *Workshop on Quantum Electrical Engineering (Wayzata, Minn.) 1988*
[38] Landauer R 1957 *IBM J. Res. Dev.* **1** 223
[39] Kane C L, Serota R A and Lee P A 1988 *Phys. Rev.* B **37** 6701
[40] Garcia N 1987 Presentation at *STM Workshop, ITCP (Trieste) 1987* unpublished
[41] Johnston R and Schweitzer L 1988 *J. Phys. C: Solid State Phys.* **21** L861
[42] Widom A and Tao R 1988 *J. Phys. C: Solid State Phys.* **21** L1061
[43] Kramer B 1988 *Physics and Technology of Submicron Structures* ed. H Heinrich, G Bauer and F Kuchar (Heidelberg: Springer) p 138
[44] Escapa L and Garcia N 1989 *J. Phys.: Condens. Matter.* **1** 2125
Garcia N and Escapa L 1989 *Appl. Phys. Lett.* **54** 1418
[45] Szafer A and Stone A D 1989 *Phys. Rev. Lett.* **62** 300
[46] Avishai Y and Band Y B 1988 *Preprint*
[47] Glazman L I, Lesovick G B, Khmel'nitskii D E and Shekhter R E 1988 *Pis. Zh. Eksp. Teor. Fiz.* **48** 218 (Engl. transl. 1988 *JETP Lett.* **48** 238)
[48] Haanappel E G and van der Marel D 1989 *Phys. Rev.* B **39** 5484
[49] Kirczenow G 1988 *Solid State Commun.* **68** 715
[50] Kumar A, Laux S E and Stern F 1989 *Appl. Phys. Lett.* **54** 1270
[51] Avishai Y 1988 private communication
[52] Payne M C 1989 *J. Phys.: Condens. Matter* **1** 4931