PERGAMON

Progress in Surface Science 66 (2001) 3–51

Progress in SURFACE SCIENCE

www.elsevier.com/locate/progsurf

Theory of ballistic electron emission microscopy

P.L. de Andres [a,*], F.J. Garcia-Vidal [b], K. Reuter [c], F. Flores [b]

[a] *Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK*
[b] *Department de Física Teórica de la Materia Condensada (UAM), Cantoblanco, E-28049 Madrid, Spain*
[c] *Fritz-Haber Institut der MPG, Faradayweg 4-6, D-14195 Berlin, Germany*

Abstract

The present status of theories for interpreting experimental ballistic electron emission microscopy (BEEM) data is reviewed. Current formalisms may be divided into two broad classes: one-electron theories, where carriers do not exchange energy with other excitations in the solid, and scattering approaches, where such losses are considered. While the former theories have been formulated with the help of Green's functions (GFs), the latter have relied more on simulation by Monte-Carlo techniques. For the one-electron approach, we discuss why the originally suggested free propagation of carriers (e.g., ballistic electrons) does not offer a consistent interpretation of the experimental database and should be replaced instead by considering the coherent propagation of electrons interacting with the periodic potential in the metal base. Bridging towards the scattering formalisms, it is shown how GFs incorporating a complex self-energy are still a feasible approach, when only a single inelastic source of scattering (e.g., electron–electron (e–e) interaction) is operative. Within this one-electron scheme, the importance of an accurately computed transmission coefficient at the metal-semiconductor interface is stressed, when aiming to obtain *absolute* BEEM currents. Analyzing results from scattering techniques, it is argued that this coefficient should be modified to take into account the back-injection of electrons from the semiconductor into the metal. A general expression for BEEM currents is given that can be used to simulate results in real-space, reciprocal-space or energy-space (spectroscopy with BEEM). Some experimental results are discussed in relation to the theories presented. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Ballistic electron emission microscopy (BEEM); Green's functions; Keldysh' formalism; Monte-Carlo; Scattering; Transport in metallic thin films; Mean free paths

* Corresponding author. Permanent address: Instituto de Ciencia de Materiales (CSIS), Cantoblanco, E-28049 Madrid, Spain. Tel.: +34-91-334-9049; fax: +34-91-372-0623.

E-mail address: pedro@icmm.csic.es (P.L. de Andres).

0079-6816/01/$ - see front matter © 2001 Elsevier Science Ltd. All rights reserved.
PII: S 0 0 7 9 - 6 8 1 6 (0 0) 0 0 0 2 2 - 8

Nomenclature

BK	Bell–Kaiser
BZ	first Brillouin zone
BEEM	ballistic electron emission microscopy
CBM	conduction band minima
$C(V)$	capacitance characteristics
e–e	electron–electron
e–ph	electron–phonon
e–d	electron–defect
FD	Fermi–Dirac distribution for electrons
FE	free electron
GF	Green's function
IBZ	two-dimensional projection of the first Brillouin zone in the plane at the interface
IWSC	two-dimensional projection of Wigner–Seitz cell in the plane at the interface
JZ	Jones zone
LCAO	linear combination of atomic orbitals
MFP	mean free path
MC	Monte-Carlo
MS	metal-semiconductor
PL	Prietsch–Ludeke
RPA	random phase approximation
STM	scanning tunneling microscopy
SB	Schottky barrier
UHV	ultra-high vacuum
WKB	Wentzel–Kramer–Brillouin

Contents

1. Introduction 5

2. Experimental observations 10

3. One-electron models of BEEM 15
   3.1. Free electron propagation: BK and PL theories 17
   3.2. Elastic propagation in metal and band structure effects 19
   3.3. Transmission at MS interface 29

4. Metal base scattering processes 34
   4.1. E-space parametrized MC 35
   4.2. $\vec{k}$-space MC 39
   4.3. Green's functions and e–e scattering 41

5. Semiconductor scattering – BEEM currents and metal e–e mean free paths 44

6. Conclusions 49

Acknowledgements 49

References 50

1. Introduction

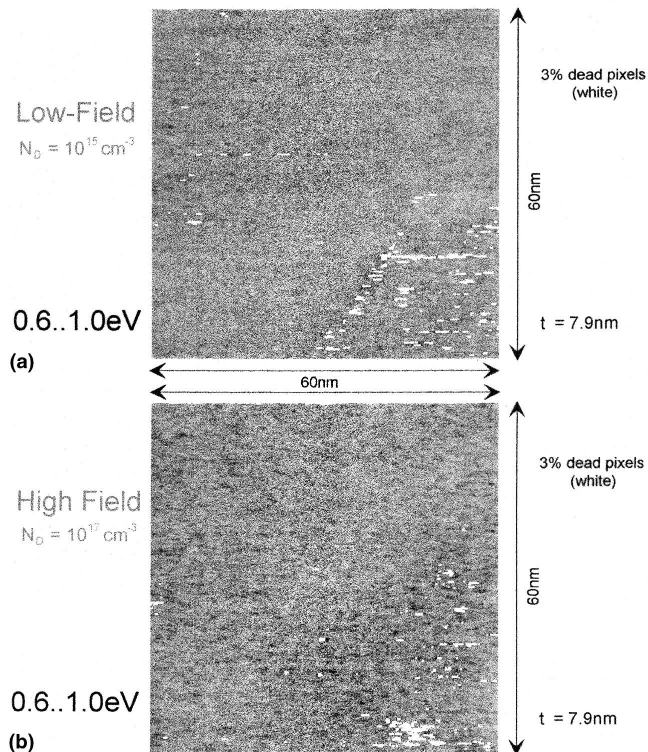
Scanning tunneling microscopy (STM) was invented in 1982 by Binnig and Rohrer [1] as a local probe to explore surfaces. Its unique capability to provide real-space images with atomic resolution clearly demonstrates that only a small surface region is involved in the operation of the instrument. This is understood in terms of the exponential decay of the tunneling current as a function of the distance, selecting the shorter paths between the STM tip (e.g., the tip point closer to the surface) and the sample underneath [2]. This local feature of the STM technique has been exploited in different ways besides the original intention of extracting atomic resolution images. Among other remarkable applications, it has made possible the manipulation of individual atoms [3], operation of an atomic switch [4], intervention in chemical reactions, etc. In this paper, we are concerned with the present status of the theory of one fruitful idea aimed to extend the power of STM to study a buried region well below the surface. The original proposal was put forward by Bell and Kaiser [5,6], only five years after the STM itself had been shown to be a practical tool, and it has now become established as a new technique in its own right [7,8]. The original purpose was to provide a direct tool to study the *Schottky barrier* (SB). Previous methods to measure the barrier at a *metal-semiconductor* (MS) interface included measurements of capacitance, $C(V)$, or intensity, $I(V)$, characteristics of the corresponding diode. These techniques depend on measuring electrical properties and need to form contacts. Therefore, they can be used on very thin films only with great difficulty. The standard way of deducing SB heights from these techniques requires a theoretical model for the $C(V)$ and $I(V)$ characteristics. According to these models, the measured values of $1/C^2$ or $\ln I$ versus V are plotted and fitted to straight lines to identify different parameters in the theoretical models, that finally can be translated to a SB value, V_{SB} [9]. Throughout this paper, atomic units are used ($e = m = (h/2\pi) = 1$), and voltages and energies will be considered dimensionally equivalent, and usually measured in eV (27.2 eV = 1 a.u.).

A different method of measuring the SB is based in photoelectric measurements. The physical principles behind this technique are not very different from those relevant to *ballistic electron emission microscopy* (BEEM). Using external tunable monochromatic light, an electron in the metalic layer is excited across the Fermi surface. If the photon energy, ω, is made greater than the SB, a current is detected in the semiconductor. As with the $C(V)$ and $I(V)$ techniques, a theoretical expression for I close to the threshold is used to get more accurate values for the barrier than

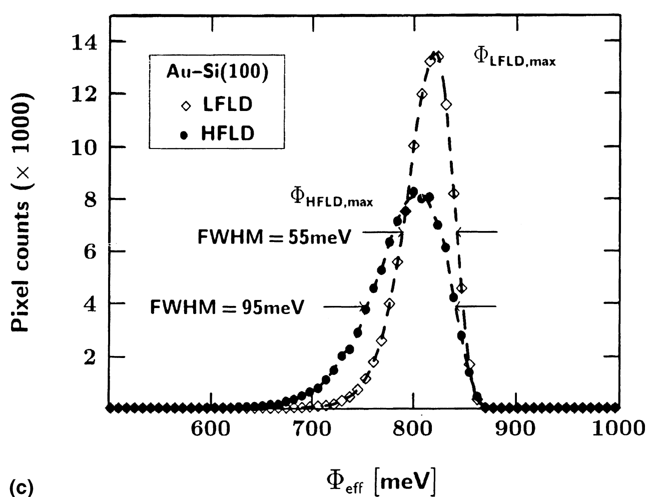
those obtained by a mere extrapolation to zero current. A popular one is an old proposal by Fowler [10]. It was originally made to get the work function of metals from photocurrents, and afterwards it was slightly modified to obtain SB values from photocurrents. The principle is simply to compute the available electron density of states, and it predicts that the intensity will depend on the photon energy quadratically, $I \propto (\omega - V_{SB})^2$. Incidentally, this is an argument similar to the one used in the *Bell–Kaiser* (BK) model, to be presented below, and it is not surprising that the same functional dependence has been found (e.g., see (3.6)). Therefore, $\sqrt{I}$ is plotted against ω, and the resulting straight line is fitted to get the barrier height. The techniques mentioned so far, have given a fairly consistent picture for SB values at different interfaces, but all of them are averaging techniques, providing values representative of regions as large as tenths of millimeters. Therefore, a considerable interest remained to introduce a new technique that could measure the barrier with smaller lateral resolution. The reasons are both practical and fundamental. If fluctuations in the barrier exist in some regions, the smaller barrier values could dominate the current transmitted over the interface, and determine the response of small semiconductor devices. Furthermore, fluctuations in the SB play an important role in the test of theoretical models set up to explain its physical origin [11,12]. Before the advent of BEEM, lateral variations in the SB height could not directly be imaged, although some experimental observations seemed to indicate their existence [13]. We should mention that, with great technical effort, the photoelectric technique could be adapted to achieve a lateral resolution in the range of some tenths of micrometer through two different modifications: (i) by focusing the photon beam to cover very small regions (scanning photoelectron spectromicroscopy), and (ii) by near-field internal photoemission spectromicroscopy [14]. However, these attempts are not comparable with BEEM capabilities, where nanometer-sized fluctuations on the SB have been measured (e.g., see Fig. 1 obtained by Palm et al. [15,16]). Furthermore, BEEM has also been recognized for its great potential to analyze other processes in the experiment, that are interesting both from a fundamental and a practical point of view. This is the case of the transport of hot electrons through metalic thin films at energies near the Fermi energy (e.g., Fig. 2 obtained by Bell [17]), which has been studied before with other techniques, but BEEM seems to have a number of advantages for obtaining better values for the *mean free paths* (MFP) parameters. From an experimental point of view, it is clear that BEEM has greater control of positions where the electron is injected and collected, allowing a better knowledge of all the multiple paths involved. From a theoretical standpoint, this problem is discussed below in more detail. These two are only but a small fraction of many other interesting results achieved with BEEM. Our primary interest is to review the theory

→

Fig. 1. Spatially resolved image of effective SB height in Au/Si(100) interface. Observed fluctuations range between 0.6 eV (black) $\leqslant V_{SB} \leqslant 1.0$ eV (white) in scanned 60 nm × 60 nm area (at a 512 × 512 pixel resolution). (a) and (b) have been taken at low doping ($N_D = 10^{15}$ cm^{-3}) and at high doping ($N_D = 10^{17}$ cm^{-3}) sample, respectively. This is estimated to correspond to interface electric fields of $\approx 10^3$ V/cm and $\approx 10^5$ V/cm, respectively. (c) Quantitative analysis showing 24 ± 10 meV change between low and high doping samples (from [15,16]).



Schottky Barrier Images Au/Si
Low-Field
N_D = 10^15 cm^-3
0.6..1.0eV
(a)
3% dead pixels
(white)
60nm
t = 7.9nm
60nm
High Field
N_D = 10^17 cm^-3
0.6..1.0eV
(b)
3% dead pixels
(white)
60nm
t = 7.9nm



Schottky Barrier Histogram Au/Si
Au–Si(100)
LFLD
HFLD
Φ_LFLD,max
Φ_HFLD,max
FWHM = 55meV
FWHM = 95meV
Pixel counts (× 1000)
14
12
10
8
6
4
2
600
700
800
900
1000
Φ_eff [meV]
(c)

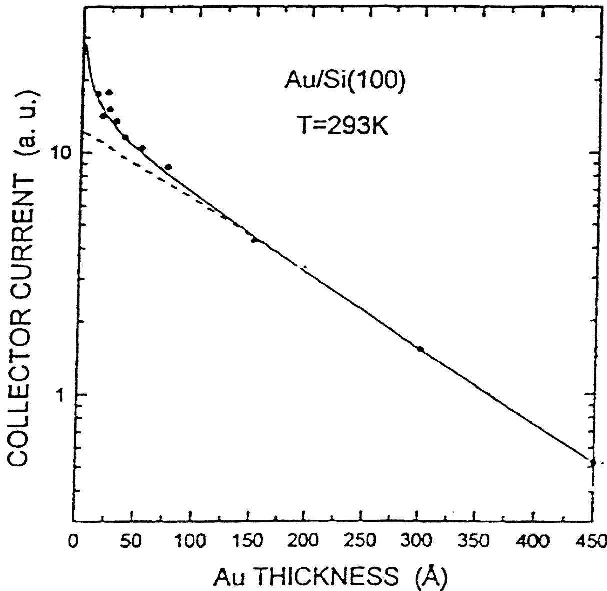


Fig. 2. Attenuation of BEEM current as function of deposited metallic film width for Au/Si(1 0 0). The deviation of experimental data from simple exponential decay at small widths indicates influence of multiple electron reflections between surface and interface. From this kind of measurement characteristic attenuation lengths can be accurately measured, and by applying adequate theoretical analysis it is, in principle, possible to determine inelastic MFPs (from [17]).

of BEEM and we shall only present experimental results that illustrate challenges for the theory. The reader interested in the experimental side is directed to any of the excellent reviews in the literature covering this subject (e.g. [7,8]).

For present purposes, it is sufficient to consider a basic experimental setup consisting of a STM tip injecting electrons onto a metalic thin film deposited on a semiconductor substrate (Fig. 3). The STM tip plays the role of a local electron gun that takes advantage of tunneling to fulfill two important conditions: (i) electrons are introduced in the metal through a very small area, and (ii) they can be put in anywhere between the Fermi energy and the vacuum level, with a very narrow energy distribution. An electrical contact is established between the tip and the metal, to measure the tunneling current, as in the standard STM experiment. This contact makes a common Fermi level for both electrodes that can be shifted by the applied voltage in the gap, V_T. A second contact is made between the tip and the semiconductor to measure the current flowing through the MS interface, which is usually referred as the BEEM current. Different interesting variations of the experiment have been tried, but, as we only want to give simple ideas for discussing different theories, we mostly consider an standard setup, roughly defined by the following parameters: the metalic film is not less than five or six layers thick, but no more than a few hundreds layers thick, typical distances from the tip to the surface are about 1 nm or less, intensities injected into the metal are around nA, energies on the order of 1 eV

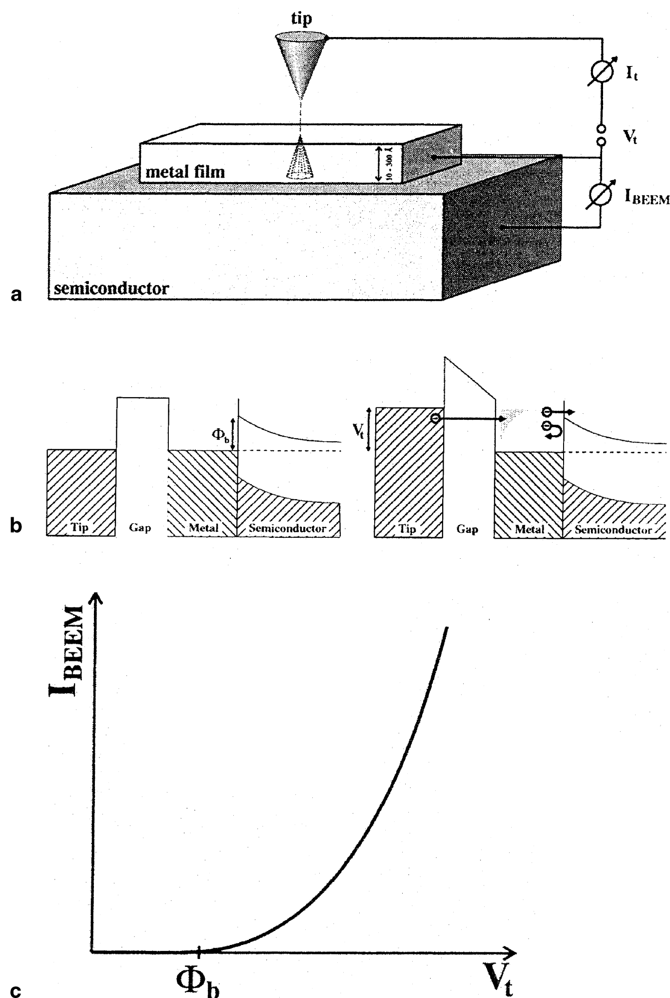


Fig. 3. (a) Schematic setup for BEEM. Current is injected from STM tip (acting as microscopic electron gun) into heterostructure formed by thin metallic film on top of semiconducting substrate. STM current is measured in the tip-metal gap, while BEEM current is measured after propagating through film and passing over SB formed at interface, in backcontact on the semiconductor. (b) Schematic corresponding 1D band-structure. (c) Schematic $I(V)$ characteristic for BEEM current, displaying typical potential law in voltage across tip-metal gap.

(e.g., $\approx V_{SB}$), and the collected BEEM current in the semiconductor backcontact is around several pA.

This review is organized as follows: we discuss first a few experimental results in Section 2, emphasizing the kind of information that has been gathered so far with the technique. In Section 3, we analyze what we call *one-electron models*, in these cases, electrons injected into the metal layer are assumed to propagate either *freely* (this is the *free electron* (FE) model, or the ballistic limit, where the technique received its

name), or *coherently*, according to the metal band structure, due to coherent interaction between an electron and the periodic crystal lattice. The important thing in these approaches is that electrons maintain a fixed energy. In a more realistic model, however, electrons suffer scattering processes which modify the previous picture. This is analyzed in Section 4, where we discuss the different approaches used to take into account the *electron–electron* (e–e), *electron–phonon* (e–ph) and other *scattering mechanisms* in the BEEM current. This has been done in the past mainly via semiclassical *Monte-Carlo* (MC) simulation, but *Green's functions* (GF) methods can also be used to tackle this problem. In Section 5, we discuss how other scattering processes inside the semiconductor can modify the BEEM current, and finally, in Section 6, we present our final comments and conclusions.

2. Experimental observations

Experiments with BEEM permit the direct imaging of SB height fluctuations, as has been shown by Palm et al. [15,16] on the Au/Si interface (Fig. 1), or the identification of different barriers on inhomogeneous films [18,19]. Therefore, it is not surprising to find that many experimental groups have applied BEEM to the precise measurement of averaged SB values on many different MS interfaces. As commented above, it is beyond the present scope to discuss such a vast literature in detail, and the interested reader is referred to the excellent accounts given by Prietsch [7] or Bell and Kaiser [8]. Most of these studies proceed by measuring the $I(V)$ characteristics and by considering the V_{SB} as a best-fit parameter. In this context, it is interesting to understand which theoretical law is better suited for this least-squared fit determination of the SB. As we shall see below, this point can be discussed from a theoretical point of view, and it is well illustrated by the empirical search performed by Cuberes et al. [20] for the best fitting power law, $I \propto V_T^n$, on the Au/Si interface (Fig. 4).

There is a clear consensus on the BEEM community around the existence of a nanometric lateral resolution even after the electron beams have been propagated through relatively thick metallic layers (e.g., nearly 50 references on this point are given in Section 5.3 of Prietsch's review [7]). The logical hypothesis is that it must be a consequence of sharp electronic beams impinging on the interface, but it is not clear how this could happen within the ballistic electron scenario. An elegant experiment proving this particular point comes from Milliken et al. [21]: a BEEM scan is performed on Au/Si, but part of the interface is covered with an insulating layer blocking the transmission of electrons. The growth properties of this layer are well known, and allow the design of two sharply defined regions. An accurate value for the lateral resolution can be obtained by measuring the rising of the BEEM current, when the tip crosses from the blocked to the allowed region, which is seen in Fig. 5, where an abrupt change in BEEM current is found in a region about 15 Å wide. This can be measured independently of the silicon substrate orientation, even after the electronic beam has been propagated through a gold metallic film of width $w > 100$ Å. Different examples of very high lateral sensitivity have been quantitatively reported in the literature, but probably one of the most recent and impressive cases is

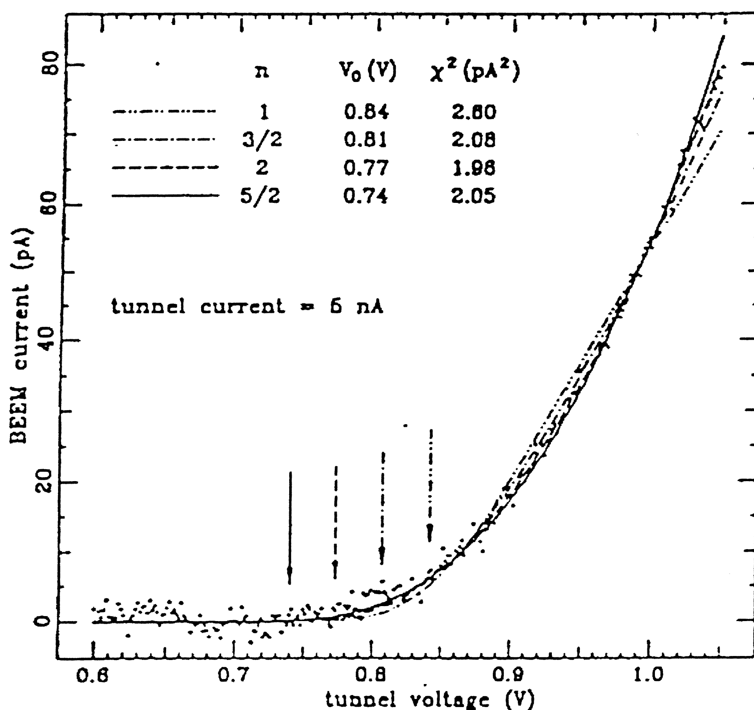


Fig. 4. Experimental $I(V)$ curve for 40 Å thick Au film evaporated on n-Si(1 1 1)-(7 × 7) under UHV conditions. Threshold position (i.e., the SB height) and its corresponding χ^2 deviation have been computed for different power laws thermally-broadened from best-fit procedure. A small variation in SB values with different theoretical assumptions is shown (from [20]).

that of von Känel's and co-workers [22,23]. Working on 20 or 30 Å $CoSi_2$ films deposited on Si(1 1 1), images were found showing quasi-atomic resolution of interface dislocations and point-defects (see Fig. 6). These pictures clearly imply that a very narrow beam must be scanning the interface. A high lateral spatial resolution, compatible with the reported experiments, has been theoretically predicted for both systems, Au/Si [24], and $CoSi_2$/Si [25].

Experiments performed on silicides deposited on silicon enable work with interfaces of very good quality, and metallic films with low densities of defects. In combination with *ultra-high vacuum* (UHV) and low-temperature conditions, these systems have provided high quality data, where interesting points on BEEM have been learnt. In particular, Sirringhaus et al. [22,26] observed atomic resolution on a BEEM image, reflecting a reconstruction forming part of the surface (Fig. 7). This unexpected correlation between the STM image and the BEEM image was explained by Reuter et al. [25], by considering the effect of surface states on the injected tunneling electrons: these effects offer the interesting possibility of realizing surface state spectroscopy by combining STM and BEEM images, as shown recently by Weilmeier et al. [27].

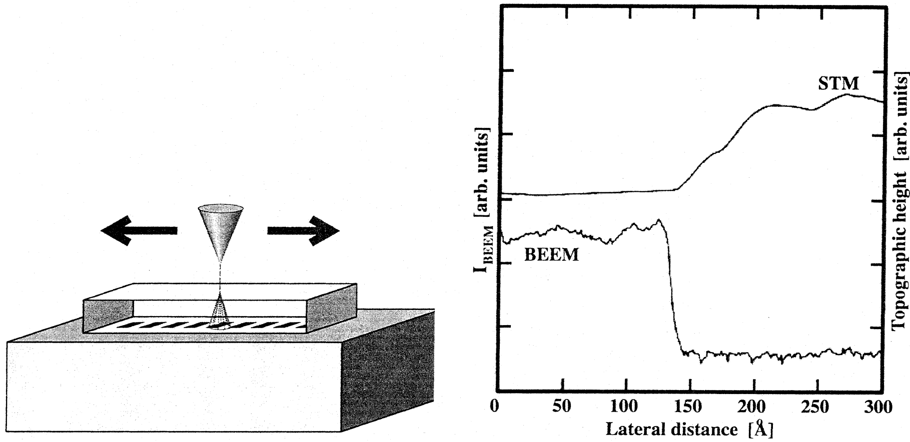


Fig. 5. Scans over patterned Au/Si(1 1 1) interface (across a buried oxide layer of SiO_2 patch). The border of patch can be extracted from onset of STM topographic (upper) line, while sharp step riser in corresponding (lower) BEEM line indicates high spatial focalization of beam scanning the interface. After propagation of electrons across ≈100 Å metalic film, spatial resolution of ≈15 Å is deduced from experiment both for Au/Si(1 1 1) and Au/Si(1 0 0) – not shown – interfaces, indicating that transport is not simply ballistic (from [21]).

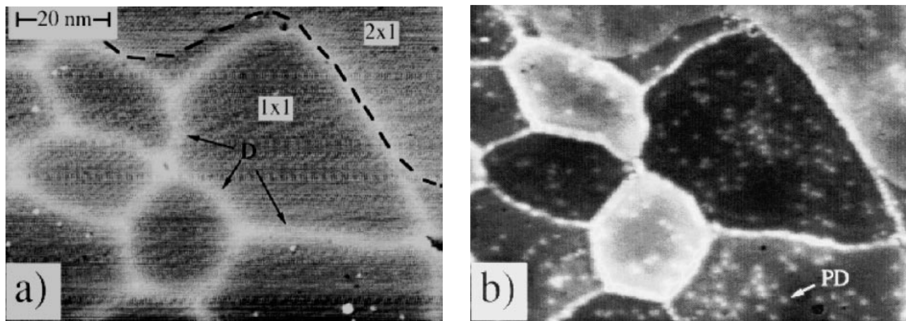


Fig. 6. STM (a) and BEEM (b) images from 28 Å $CoSi_2$ film deposited on Si(1 1 1). STM parameters are $I_t = 10$ nA and $V = E_F + 1.6$ eV. Quasi-atomic resolution of dislocations (D) and point defects (PD) is still seen in the BEEM image, suggesting existence of narrowly focused beams in metal layer (from [23]).

Information about the $\vec{k}_{\parallel}$ distribution at the interface can be obtained by choosing different semiconductor substrates, or different faces for the same semiconductor, while tunneling and propagation conditions through the metalic layer are kept fixed. By changing the semiconductor projected interface band structure, the experiment selects different regions in the 2D *Brillouin zone* (BZ), and the resulting changes observed in the BEEM current can be used to understand the $\vec{k}_{\parallel}$ distribution incoming from the metal side into the semiconductor. This point has been demonstrated by Schowalter and Lee [28], by performing experiments on Au films

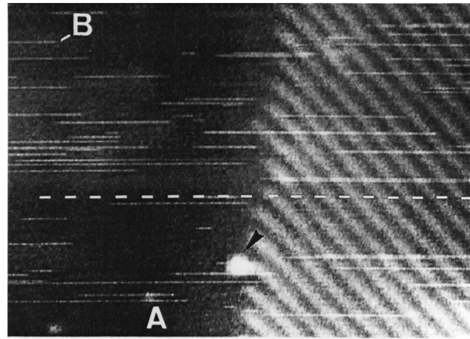


Fig. 7. BEEM image for $CoSi_2$/Si(1 1 1) interface. Metal film is ≈ 30 Å thick, and good resolution is achieved either in region below 2 × 1 reconstruction of surface (right), where anticorrugation w.r.t. STM image is obtained, or in region below unreconstructed patch of surface (left), where point defects can be also imaged (A and B) (from [26]).

deposited either on Si(1 0 0) or on Si(1 1 1) surfaces (Fig. 8 schematically illustrates the projected *conduction band minima* (CBM) on both orientations). The surprising finding from these experiments was a BEEM current quite insensitive to the differences in the projected semiconductor band structure at the interface. Similar results have also been found by other groups (e.g., see Ref. [8] in [28], or compare data for a 70 Å thick Au/Si(1 0 0) in [29] to a corresponding 75 Å thick Au/Si(1 1 1) in [17]). Moreover, Weilmeier et al. [27] have shown that the BEEM current for Au(1 1 1)/Si(1 1 1) is consistently around 1.7 higher than for Au(1 1 1)/Si(1 0 0), at a fixed bias of $V_T = 1.2$ eV.

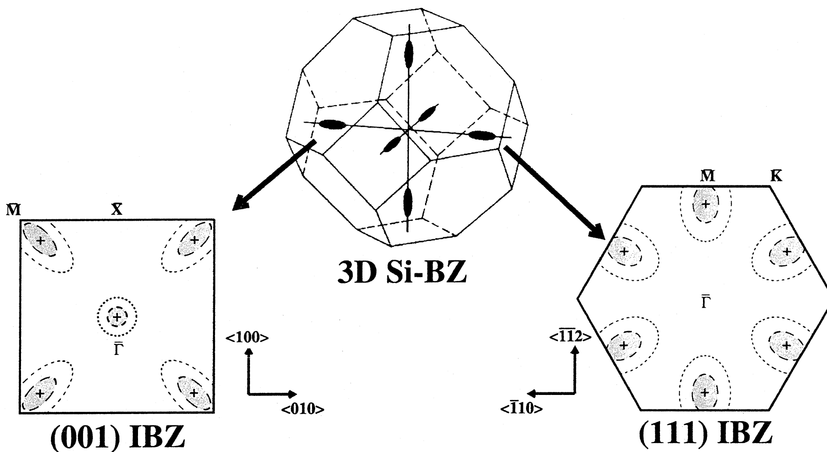


Fig. 8. Constant energy surface of Si conduction band (middle panel) and its respective projections onto (1 0 0) and (1 1 1) orientation (left and right hand panel): shown are high-symmetry points and fundamental directions inside interface BZ, CBM locations (+) and open-state pockets at 0.2 eV (dashed line) and 0.6 eV (dotted line) above minimum.

Although it might seem a very indirect way of probing the actual $\vec{k}_{\parallel}$ distribution impinging in the interface, these results have originated a lot of discussion, because they do not fit naturally with the pure ballistic model. In such a model, the $\vec{k}_{\parallel}$ distribution is fixed at the tunneling step, and it is not modified by propagation through the metal layer. The *Wentzel–Kramers–Brillouin* (WKB)-like planar tunneling theories predict the formation of a narrow cone propagating in the forward direction, dictated by the exponential dependence with energy and the corresponding higher difficulty to make tunneling for high $k_{\parallel}$ components (e.g., see (3.1)). This single point can already be probed by substituting the Si(1 0 0) surface by the Si(1 1 1), as the former possess a suitable CBM near the $\overline{\Gamma}$ point, and can transmit current distributions peaked around that point, whereas this possibility is precluded for the later, as the silicon CBM in the (1 1 1) orientation are located at about 40° off the normal direction, giving a smaller overlap with the initial distribution than for the (1 0 0) case. The lack of observation of important differences between both orientations, points to the opposite picture to the one initially assumed: the $\vec{k}_{\parallel}$ distribution must be somehow homogeneously scattered on the 2D BZ. This cannot be explained by a mere ballistic model, and calls for including some interaction, although the constraint of keeping the good lateral resolution, found in different experiments, has also to be kept in mind.

In spite of these comments, many theories described below assume, quite generally, that energy and $\vec{k}_{\parallel}$, the two good quantum numbers for the ideal model, are conserved at the interface. However, the necessity to explain a broad $\vec{k}_{\parallel}$ distribution and the suspected bad quality of interfaces such as Au/Si, have prompted several groups to consider an important relaxation of the momentum conservation law, keeping the energy conservation as the only restriction. In pioneering work, Schowalter and Lee introduced this simple modification to the ballistic BK theory by assuming that the distribution in $\vec{k}_{\parallel}$-space must be uniform, due to defect scattering processes inside the metal base [28] and/or randomization of the $\vec{k}_{\parallel}$, due to imperfections at the interface [28,29]. This automatically produces similar $I(V)$ curves for both (1 1 1) and (1 0 0) orientations, but apparently clashes with the measured spatial resolution mentioned above. We discuss later a theoretical framework that can explain concomitantly both experimental observations. This is consistent with recent data by Dähne-Prietsch et al. [30]. They have shown, by measuring the increase in the BEEM current induced by locating of the tip at terrace edges of gold films, a predominant conservation of lateral momentum at the MS interface.

It is interesting to notice that, irrespective of the $\vec{k}_{\parallel}$ issue, the quadratic functional dependence with V_T near the threshold, predicted by BK theory and discussed in the next section, continues to be valid, even if the momentum conservation condition is lost. This results from the approximation $I_T(E,\vec{k}_{\parallel}) \approx \rho(E_F)$, previously only acceptable around $\vec{k}_{\parallel} = 0$, that is now extended to the whole 2D BZ, and only a different normalization is needed (exactly for the same argument, the $\frac{5}{2}$ law, to be discussed below, is kept unchanged).

Finally, we mention that BEEM presents a good potential to investigate attenuation lengths on metalic thin films. The measured BEEM current cannot be theoretically explained, except for very thin films or energies very close to the threshold,

without introducing attenuation. Therefore, inelastic effects related to this attenuation in the transport through the metalic film, can be studied with good accuracy and detail. This application needs of some theoretical description for the inelastic scattering channels, but even very simple theories can already show interesting effects. As an example, Bell's work [17] has immediately allowed the identification of the importance of multiple reflections between the two surfaces of the metallic film for small widths (Fig. 2).

3. One-electron models of BEEM

The goal of the simple models, initially proposed for BEEM, was to set up the minimum theoretical description needed to obtain a reliable determination of the SB from the experimental data. Such a goal allows drastic simplification in the theory, in particular, it is possible to consider the transport of electrons as ballistic. Simple as this is, it carries a practical importance, because it yields valuable information on a technologically relevant magnitude: the SB. In close analogy to early photoemission theories, it is assumed that the BEEM current can be computed as the result of four independent, but linked steps [7,31]:

(a) tunneling from the STM-tip to the metalic layer,
(b) transport through the metal,
(c) transmission at the MS interface,
(d) Different effects in the semiconductor affecting the measured BEEM current.

Accordingly, a simple theoretical description has been proposed based on the effective 1D band scheme sketched in Fig. 1, which enables the following approximate dynamical factors to be discussed:

(a) *The injected tunneling current* as derived from a one-dimensional model. In a first approximation, this can be achieved by computing tunneling currents between two metalic planes. The WKB approximation can be used to derive analytical expressions, but, for our purpose, we only need to write down the dependence of the tunneling current with the tip-sample separation, d, and, more significatively for the BEEM problem, with the electron energy, E

$$I_T \propto \mathrm{e}^{-2kd}. \tag{3.1}$$

Within such a rough model, it is not worth bothering too much with details. Therefore, k can be estimated from a square barrier assuming a FE-like energy dispersion, $k = \sqrt{2W - (2E - k_{\parallel}^2)}$ ($W \approx 4$–5 eV, a typical value for the work function of many metals).

(b) *Transport in metal base via FE model*. As widths for metal films have, in some experiments, been smaller than the typical attenuation lengths measured for those materials, it has been thought acceptable to describe electrons as ballistic carriers. This is the origin of the name given to the technique and its main implication is to keep for the carriers an energy and $\vec{k}_{\parallel}$-distribution determined at the tunneling step. Whenever attenuation plays a role, this model can be easily extended by giving the

carrier a fixed probability of losing energy, independent of its past story. This results in an exponential decay:

$$A(E) \propto e^{-l/\lambda} \tag{3.2}$$

with a typical relaxation length, λ, which, in general, can be a function of E. Two cases are of interest here: $\lambda \gg l$, where we recover the ballistic model, and $\lambda \ll l$, where the current is truly controlled by the exponential decay. We shall see that the region in between is rich in details, and cannot be fully accounted for by such a simple model.

(c) *Electron transmission at MS interface*. The simplest approach is to assume that it takes a constant value, independent of energy, to be absorbed in several constants affecting only the prefactor of the qualitative theory. A further step would be to consider the quantum-mechanical expression by modelling the SB by a simple step (*Prietsch–Ludeke* (PL) model [32])

$$T(E, \vec{k}_{\parallel}) = \frac{4k_{\mathrm{M}}k_{\mathrm{S}}}{(k_{\mathrm{M}} + k_{\mathrm{S}})^2}, \tag{3.3}$$

where k_{M} and k_{S} represent the wave vector perpendicular to the interface in the metal and the semiconductor, respectively. In this model, one keeps in mind that energies are measured in the metal with respect to the bottom of the band, and wave vectors with respect to the Γ point, while, in the semiconductor, they are referred to the bottom of the conduction band, V_{SB} (not necessarily located at Γ, as, e.g., in Si). As in the tunneling process, it is not more difficult to consider other shapes for the barrier (e.g., a triangular one). However, at this level of sophistication it is not necessary, and it is not worth to care about band bending, or the image potential in the depletion layer inside the semiconductor.

(d) *Other semiconductor processes.* Finally, one has to incorporate a number of physical processes in the semiconductor that differ in importance, depending on the energy. As we shall see below, the main use of ballistic or near-ballistic models such as the ones described here is to characterize the threshold region. Therefore, as far as the semiconductor is concerned, we completely neglect effects, such as impact ionization, which are important at higher voltages [31], and we comment only on a process relevant for the threshold region: the *backscattering* of electrons towards the metal by interaction with phonons in the semiconductor. This can be introduced into the theory, without excessive complications, by simply correcting the transmission coefficient at the interface to get the charge effectively injected in the semiconductor, after taking into account the back-injection into the metal. This factor, $S(E)$, was first proposed by Lee and Schowalter [33], and later computed in a more refined scheme by Hohenester [34] (see below). The main effect, besides making the interface more reflecting, is to smooth considerably the energy dependence of the bare quantal transmission coefficient.

Different approaches have tried to improve, in one way or another, the description of the different step processes mentioned above. We shall concentrate now our discussion on presenting theories that have been proposed for describing the

BEEM current within *one-electron models*. In this approach, electrons are assumed to propagate either ballistically in the metal base, or according to the band structure of the metal crystal. The implications of both assumptions is discussed below. It is also important for determining the BEEM currents, to calculate with a certain degree of accuracy the electron transmission coefficients across the MS interface. As this task is a typical one-electron problem, we discuss also in this section how different theories have calculated this coefficient, assuming $\vec{k}_{\parallel}$-conservation at the interface.

3.1. Free electron propagation: BK and PL theories

The essence of the first theory proposed for BEEM, the so-called BK theory or *phase-space* model [5,6], is to compute the flux of every ballistic electron, which is allowed to enter the semiconductor. Therefore, the current is obtained as the integral of the flux distribution at the interface, viz.,

$$
\begin{aligned}
I_{\mathrm{B}}(V_T) &= \int v_{\perp}[F(E-V_T)-F(E)]\theta(E-V_{\mathrm{SB}})\frac{\mathrm{d}^3\vec{k}}{(2\pi)^3} \\
&= \frac{1}{(2\pi)^3}\int_0^{\infty} v_{\perp}[F(E-V_T)-F(E)]\theta(E-V_{\mathrm{SB}})\mathrm{d}k_{\perp}\int_{\mathrm{IBZ}}\mathrm{d}^2k_{\parallel}. \qquad (3.4)
\end{aligned}
$$

The original BK model assumes that ballistic propagation takes place through a FE metal, $\vec{k}_{\parallel}$ and E are conserved at the interface, and the intervening dynamical factors are approximated by: $T(E)=\theta(E-V_{\mathrm{SB}})$, $S(E)=1$. $v_{\perp}$ represents the perpendicular component to the surface of the electrons group velocity, and $F(E)$ is the *Fermi–Dirac* (FD) distribution at temperature T (Fig. 3 illustrates the schematics of this 1D model). As the good quantum numbers of the problem are energy, E, and parallel momentum, $\vec{k}_{\parallel}$, it is convenient to substitute $k_{\perp}$ for E. Using that $v_{\perp}=\partial E/\partial k_{\perp}$, we find:

$$
I_{\mathrm{B}}(V_T) = \frac{1}{(2\pi)^3}\int_0^{\infty}[F(E-V_T)-F(E)]\theta(E-V_{\mathrm{SB}})\mathrm{d}E\int_{\mathrm{IBZ}}\mathrm{d}^2k_{\parallel}. \qquad (3.5)
$$

The available phase space is determined by the kinematic condition that electrons should carry enough energy in the perpendicular direction to overcome the barrier: $\frac{1}{2}k_{\perp}^2 \geqslant V_{\mathrm{SB}}$, which defines a maximum $k_{\parallel}$ value for electrons that can be injected into the semiconductor: $k_{\parallel} \leqslant k_{\parallel}^{\max} = \sqrt{2(E-V_{\mathrm{SB}})}$. The integration in (3.5) can be easily performed at $T=0$ K (to simplify the calculations, cylindrical symmetry is assumed)

$$
\begin{aligned}
I_{\mathrm{B}}(V_T) &= \frac{1}{(2\pi)^2}\int_0^{V_T}\theta(E-V_{\mathrm{SB}})\,\mathrm{d}E\int_0^{k_{\parallel}^{\max}}k_{\parallel}\mathrm{d}k_{\parallel} = \frac{1}{(2\pi)^2}\int_{V_{\mathrm{SB}}}^{V_T}(E-V_{\mathrm{SB}})\,\mathrm{d}E \\
&= \frac{1}{2(2\pi)^2}(V_T-V_{\mathrm{SB}})^2, \quad V_{\mathrm{SB}} \leqslant V_T, \qquad (3.6)
\end{aligned}
$$

where V_T is the voltage across the STM junction. For the sake of brevity and clarity, the original argument has been particularized to $T = 0$ K. However, for energies only a few $k_B T$ away from the threshold, the proper FD distribution should be used. A simple approach to this problem is still feasible for obtaining analytical results. For normal metalic densities ($r_s < 7$), and not too high temperatures ($T < 1000$ K), it is possible to introduce the following approximation to the FD distribution:

$$F(E) = \begin{cases} 1 - \frac{1}{2}\mathrm{e}^{(E-\mu)/k_B T} & \text{if } E \leqslant \mu, \\ \frac{1}{2}\mathrm{e}^{-(E-\mu)/k_B T} & \text{if } E \geqslant \mu, \end{cases} \tag{3.7}$$

in the whole range of energies $-\infty < E < \infty$. Using (3.7) in (3.5), the BEEM current at a finite temperature is obtained (sometimes referred in the literature as *thermally broadened* BEEM current)

$$I_B(V_T, T) = \frac{1}{2(2\pi)^2}\left[(V_T - V_{SB})^2 + (k_B T)^2\left(2 - \exp\left(-\frac{V_{SB}}{k_B T}\right) - \exp\left(-\frac{V_T - V_{SB}}{k_B T}\right)\right)\right], \quad V_{SB} \leqslant V_T, \tag{3.8}$$

which shows how thermal effects can change the BEEM-current near the threshold ($(V_T - V_{SB}) < 4k_B T$). It is interesting to note that, due to the tail of the FD distribution for electrons in the tip, a BEEM current $\approx (k_B T)^2$ is obtained, even for $V_T = V_{SB}$.

Eqs. (3.6) or (3.8) result from neglecting any energy dependence for the tunneling process, the transport through the metal, the transmission across the interface, and the transport through the semiconductor. PL have tried to improve this model by incorporating the energy dependence of the transmission coefficient at the interface [7,32]. Close to the threshold, $T(E)$ in (3.3) can be approximated by the perpendicular momentum in the semiconductor, $k_{\perp,S}$, so (3.6) becomes

$$I_B(V_T) \propto \int_{V_{SB}}^{V_T} \mathrm{d}E \int_0^{k_\parallel^{\max}} k_\parallel \sqrt{2(E - V_{SB}) - k_\parallel^2}\, \mathrm{d}k_\parallel \tag{3.9}$$

At this stage, we are more interested in the basics of the different physical models, so we are not concerned with the details, such as allowing different effective masses for the carriers in the metal and the semiconductor, etc. The integrals in (3.9) are again straightforward, and the energy dependence predicted by this model is

$$I_B(V_T) \propto \int_{V_{SB}}^{V_T} (E - V_{SB})\sqrt{E - V_{SB}}\, \mathrm{d}E \propto (V_T - V_{SB})^{5/2}. \tag{3.10}$$

This confronts us with a different energy dependence from the one predicted by the BK model, and the question of which one would lead to a better determination of the SB. PL have analyzed this point and have shown that the two power laws result in differences in the SB values of 0.04–0.08 eV for five different metals deposited on GaP(1 1 0) (systematically, the 5/2-law giving lower values, because of the thermal broadening). In the particular case of a 4 nm thick Au film evaporated on Si(1 1 1) in UHV, it has been shown by Cuberes et al. [20] that the SB determination changes by

0.03 eV, depending on the power law used (Fig. 4). A slightly better agreement is reported for $n = 2$, but it also slightly depends on the interval of voltages used for the fit. It has been suggested that, eventually, theory or experiment will elucidate which law is better for obtaining SB values. The derivation outlined above, indicates that, while the $n = 2$ law is valid when $V \rightarrow V_{SB}$, the $n = \frac{5}{2}$ one takes care of a dynamical effect that is important in the vicinity of the threshold region. However, it has to be kept in mind that the electrons, backscattered from the semiconductor by interaction with phonons, introduce a factor in the transmission coefficient (see discussion below) that tends to cancel the strong square-root dependence [33,34], giving a further chance to the constant transmission coefficient model. Therefore, in general, n lies between these two values, and slightly depends on the particular system under consideration.

There have been some attempts to improve the BK model to include elastic scattering at the MS interface [35]. This model starts by calculating the eigenstates of an ideal interface Hamiltonian and then allows interfacial elastic scattering to redistribute the electrons that reach MS interface. Transport inside the metal base is assumed to be ballistic, as in the BK model. The strength of the interface scattering is then fitted to the experimental data of Au/Si(1 0 0) and Au/Si(1 1 1) interfaces. The authors found that this scattering strength must be quite large, in order to reduce the ballistic current for Si(1 0 0) and increase current for Si(1 1 1). However, concluding from this calculations that the electrons parallel momentum is not conserved at the MS interface in the transmission process must be taken with certain caveats, as the propagation of electrons inside the metal base is computed using the simple ballistic approach (see below).

3.2. Elastic propagation in metal and band structure effects

The FE propagation model, assumed above, can only describe a BEEM experiment, if the metalic film exhibits a complete lack of crystalline order. Otherwise, the elastic interaction of the electrons with the ordered lattice will result in a strong interaction that, among other effects, can produce gaps for some specific k-space regions. Consequently, the electron propagation departs drastically from the ballistic assumed free flight, at least for these directions. In particular, many transition metals (like Au, Pd, etc.) exhibit such a gap for [1 1 1] directions, which turn out to be very relevant for a significative number of BEEM experiments. To include these effects, a GF approach, based in Keldysh's formalism, has been set up, which is free of adjustable parameters for an elastic scenario, and enjoys the advantage of being able to describe, on the same footing, several parts of the problem, namely, tunneling and propagation in the metallic film.

The system hamiltonian is split into three independent terms: the tip, $\hat{H}_T$, the metallic sample, $\hat{H}_M$, and their interaction, $\hat{H}_I$. It is assumed that solutions for $\hat{H}_T$ and $\hat{H}_M$ are known, and the goal is to construct a solution of the whole system when $\hat{H}_I \neq 0$. The different hamiltonians are written in a *linear combination of atomic orbitals* (LCAO) basis:

$$\hat{H}_{\rm T} = \sum \epsilon_\alpha \hat{n}_\alpha + \sum \hat{T}_{\alpha\beta} \hat{c}^\dagger_\alpha \hat{c}_\beta, \tag{3.11}$$

$$\hat{H}_{\rm M} = \sum \epsilon_i \hat{n}_i + \sum \hat{T}_{ij} \hat{c}^\dagger_i \hat{c}_i, \tag{3.12}$$

$$\hat{H}_{\rm I} = \sum \hat{T}_{\alpha i} \hat{c}^\dagger_\alpha \hat{c}_i, \tag{3.13}$$

where Greek indices (α, β) label sites on the tip, Latin indices (i, j) are referred to sites in the metal, and the orbitals forming the basis are not explicitly shown, but they are denoted by the circumflex symbol, ˆ, on the operators. A useful approximation to the formal basis is given by the Slater and Koster tight-binding approach [36], where the tight-binding parameters $\hat{T}_{ij}$ are obtained by a fitting to ab initio band structure calculations.

Currents between two nodes in the system can be obtained via the work of Caroli et al. [37], that we reproduce for a 3D lattice. An auxiliary surface S surrounding an atomic position i is drawn so two different regions of space can be distinguished: *in* and *out*. The current between point i (in) and any other generic point j (out) can be generally expressed by means of creation and annihilation operators, whose net effect is to destroy an electron on j at time t and to create it on i at a later time $t + 0^+$, minus the opposite process:

$$\hat{J}_{i,j} = A_{j,i} \hat{c}^\dagger_j (t + 0^+) \hat{c}_i(t) - A_{i,j} \hat{c}^\dagger_i (t + 0^+) \hat{c}_j(t). \tag{3.14}$$

The Heisenberg picture is assumed here, and the small positive quantity, 0^+, is useful to ensure the analytical properties of the various GF introduced later.

To fix the arbitrary constants $A_{i,j}$ and $A_{j,i}$, under stationary conditions, we assume that the electron charge around any site i, $\hat{\rho}_i = \hat{c}^\dagger_i \hat{c}_i$, must be conserved

$$\frac{\partial \hat{\rho}_i}{\partial t} = \sum_j \hat{J}_{i,j}. \tag{3.15}$$

Since the Heisenberg picture is used, and the density current in a stationary situation is independent of time in the Schrödinger picture, one can take advantage of the Heisenberg's *equation of motion* for that operator, i.e.,

$$\frac{\partial \hat{\rho}_i}{\partial t} = -i\left[\hat{\rho}_i, \hat{H}\right] = -i\left[\hat{c}^\dagger_i \hat{c}_i, \sum_{j,k} T_{j,k} \hat{c}^\dagger_j \hat{c}_k\right] = -i\left\{\sum_j T_{i,j} \hat{c}^\dagger_i \hat{c}_j - T_{j,i} \hat{c}^\dagger_j \hat{c}_i\right\} \tag{3.16}$$

and, by comparing with (3.15), the unknown constants $A_{i,j}$ are identified: $A_{i,j} = iT_{i,j}$.

A measurable magnitude is obtained from the expectation value of the operator $\hat{J}_{i,j}$ with respect to the ground state of the decoupled system, which is related to the definition of the Keldysh GF $G^{+-}_{i,i+1}(0^+)$ [38–40].

$$\begin{aligned} J_{i,j} = \langle 0|\hat{J}_{i,j}|0\rangle &= i\Big(T_{i,j}\langle 0|\hat{c}^\dagger_i c_j|0\rangle - T_{j,i}\langle 0|\hat{c}^\dagger_j c_i|0\rangle\Big) \\ &= T_{i,j} G^{+-}_{i,j}(0^+) - T_{j,i} G^{+-}_{j,i}(0^+). \end{aligned} \tag{3.17}$$

Under stationary conditions, it is more convenient to Fourier transform and work in the energy domain (while the time domain is better suited to study transitory regimes)

$$i\langle 0|\hat{c}_i^\dagger \hat{c}_{i+1}|0\rangle = G_{i,i+1}^{+-}(t+0^+) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} G_{i,i+1}^{+-}(E)\,\mathrm{d}E. \tag{3.18}$$

We have arrived at the starting point for computing currents for both BEEM and STM problems in this formalism

$$J_{i,j} = \frac{2\mathrm{e}}{h}\int \mathrm{Tr}\left\{\hat{T}_{i,j}\left[\hat{G}_{i,j}^{+-}(E) - \hat{G}_{j,i}^{+-}(E)\right]\right\}\mathrm{d}E. \tag{3.19}$$

The original work of Keldysh shows how to construct expressions to compute a number of propagators for non-equilibrium problems, including the $\hat{G}_{j,i}^{+-}(E)$ used in the above equation. For the purposes of computing BEEM or STM currents, only two rules need to be remembered [41]:

- $\hat{G}_{j,i}^{+-}(E)$ is related to standard retarded and advanced GF of the coupled system and to the $\hat{g}_{j,i}^{+-}(E)$ of the uncoupled one by

$$\hat{G}_{j,i}^{+-}(E) = \left(\hat{1} + \hat{G}^R(E)\hat{T}\right)\hat{g}^{+-}\left(\hat{1} + \hat{T}\hat{G}^A(E)\right). \tag{3.20}$$

- $\hat{G}_{j,i}^{R,A}(E)$ can be computed from the GF of the two decoupled parts of the system by applying *Dyson's equation*, namely,

$$\hat{G}_{j,i}^{R,A}(E) = \hat{g}^{R,A} + \hat{g}^{R,A}\hat{T}\hat{G}^{R,A}. \tag{3.21}$$

These rules fully determine the objects of interest, $G_{j,i}^{+-}(E)$, in terms of GF of two non-interacting systems and their interaction given by the hopping matrix $\hat{T}_{ij}$.

3.2.1. BEEM current

Eq. (3.19) was applied to compute tunneling currents before being used for BEEM. Martin-Rodero et al. [42] have used it to develop a non-perturbative STM formalism, which is useful when the interaction between the tip and the sample is not weak and the standard Tersoff and Hamann approach breaks down (e.g., short tunneling distances or non-zero voltages). It is outside our scope to discuss STM currents, and, for applications of (3.19) to this problem, we refer the interested reader to the literature (e.g., see [43]). We give, however, an expression for the STM current conveying the main physical ingredients, and valid for the large tip-sample distances typically found in BEEM experiments

$$J_T(V_T) = \frac{8\pi^2 e}{h} Re\int_{-\infty}^{+\infty} \mathrm{Tr}\sum_{m\alpha\beta n}\left[\hat{T}_{m\alpha}\hat{\rho}_{\alpha\beta}(E-V_T)\hat{T}_{\beta n}\hat{\rho}_{mn}(E)\right][F(E-V_T) - F(E)]\,\mathrm{d}E. \tag{3.22}$$

We stress that one of the assets of the Keldysh GF formalism for BEEM is to describe, on the same footing, the tunneling step and the electron propagation through the metal up to the MS interface. We notice that the electronic-structure calculations used to obtain the necessary GF, and based on decimation techniques [44], can also be extended to the problem of the transmission coefficient with little extra effort, taking care of all the physical processes where the metal plays an important role.

BEEM currents have been also analyzed by making use of (3.19) as a function of the unperturbed metal and tip GF [24,45,46]. The result is a closed expression that accurately describes the physics of elastic electronic transport inside the metal by taking into account band-structure effects. Given the two alternative modes of performing BEEM experiments, current distributions in real and reciprocal space are needed, depending on the operating mode chosen. In this formalism, both expressions share the same structure, being built, respectively, with real-space functions or their corresponding Fourier transforms:

$$J_{ij}(\vec{r}_\parallel, E) = \frac{8\pi^2 e}{h} \mathrm{Im\ Tr} \sum_{m\alpha\beta n} \left[\hat{T}_{\beta n} \hat{g}^A_{ni}(\vec{r}_\parallel, E) \hat{T}_{ij} \hat{g}^R_{jm}(\vec{r}_\parallel, E) \hat{T}_{m\alpha} \hat{\rho}_{\alpha\beta}(E) \right] \tag{3.23}$$

and

$$J_{ij}(\vec{k}_\parallel, E) = \frac{8\pi^2 e}{h} \mathrm{Im\ Tr} \sum_{m\alpha\beta n} \left[\hat{T}_{\beta n} \hat{g}^A_{ni}(\vec{k}_\parallel, E) \hat{T}_{ij}(\vec{k}_\parallel) \hat{g}^R_{jm}(\vec{k}_\parallel, E) \hat{T}_{m\alpha} \hat{\rho}_{\alpha\beta}(\vec{k}_\parallel, E) \right], \tag{3.24}$$

where (3.23) gives the current between two *nodes* in the periodic lattice, i, j, while (3.24) refers to two *planes* in the metal, i, j. Recall that Greek indices α and β refer to the tip, while Latin indices refer to the metal surface. It is interesting to note that both expressions are not the Fourier-transform of each other, as perhaps expected intuitively. Instead, they are obtained by Fourier-transforming each of the objects in the expression. In practice this is performed by approximating the integrals by a summation in the 2D BZ

$$g^R_{j1}(\vec{r}_\parallel, E) = \sum_{\vec{k}_\parallel} w_{\vec{k}_\parallel}\, g^R_{j1}(\vec{k}_\parallel, E) \mathrm{e}^{i\vec{k}_\parallel \vec{r}_\parallel}, \tag{3.25}$$

which usually is performed with the help of a set of special points, $\vec{k}_\parallel$, and their associated weights, $w_{\vec{k}_\parallel}$ [47,48].

Eqs. (3.23) and (3.24) define an approach to BEEM calculations, where the main task is to compute the *equilibrium* GF, $\hat{g}^{R,A}_{ij}$, for an isolated metalic layer. Besides reducing the whole calculation to finding familiar GFs, this formulation has a number of advantages: (i) up to this point, it is free of adjustable parameters; (ii) for a practical calculation, a well-suited minimal basis can be obtained via a Slater and Koster LCAO parametrization, which is no more than a convenient way to quickly recover a set of previously calculated bands by other ab initio methods; (iii) it can include inelastic effects by incorporating a self-energy to the GF, which can be done to different degrees of sophistication: from a self-consistent calculation to a phenomenological level, where only an energy-dependent optical potential is fitted to the data. Below we discuss, in more detail, how this works in practice.

Eq. (3.24) can be used to compute I(V) characteristics (while (3.23) is useful to analyze BEEM real-space images). Once the current distribution in the metal part of the interface has been computed as a function of E and $\vec{k}_\parallel$, it is only a matter of applying a transmission coefficient, $T(E, \vec{k}_\parallel)$, and integrating over all the energies to obtain the injected charge in the semiconductor

$$I(V_T) = \int_{-\infty}^{\infty} \mathrm{d}E \int \int_{\mathrm{IBZ}} \mathrm{d}\vec{k}_\parallel J_{n-1,n}\left(E,\vec{k}_\parallel\right) T\left(E,\vec{k}_\parallel\right) S(E,\vec{k}_\parallel)[F(E-V_T)-F(E)], \tag{3.26}$$

$J_{n-1,n}$ being the current impinging upon the last metalic layer, n, forming the interface in contact with the first semiconductor layer. $\vec{k}_\parallel$ takes values on the metalic BZ projected on the interface, and $S(E,\vec{k}_\parallel)$ is a factor correcting the transmission coefficient T to take into account the probability for an electron in the semiconductor to be scattered back into the metal (e.g., by interaction with phonons).

3.2.2. Au/Si interface

Using (3.23) and a *semiclassical approximation* for the GFs it is possible to understand the lateral spatial resolution observed on Au(1 1 1) films, as the result of the elastic interaction of propagating electrons with the ordered lattice [24]. These calculations have been extended to a full quantal approach based on *decimation techniques* [45,46,49,50] to study coherent phenomena. Within this formalism, the 2D current distribution, at a given plane parallel to the surface, can be calculated. For Au(1 1 1), this current distribution is displayed in Fig. 9 for the second, 10th and 25th

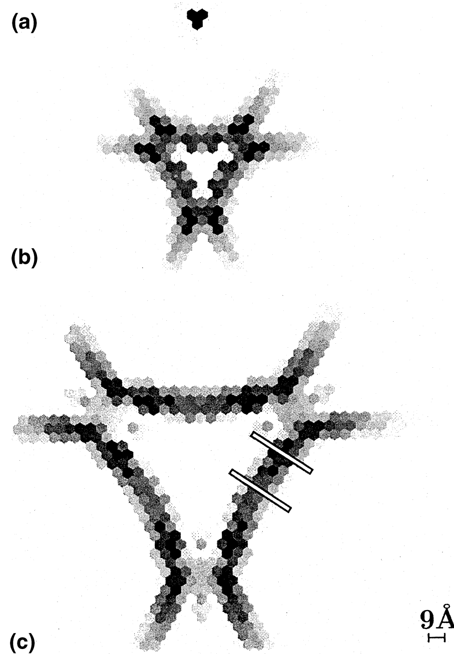


Fig. 9. Real-space current distributions for Au(1 1 1) after injection into one atom in first layer (located in center of drawn layer) at $E = E_F + 1.00$ eV. Distributions drawn are in the (a) second layer (2.35 Å down the surface), (b) 10th layer (21.19 Å, and (c) 25th layer (56.51Å). Each dot represents one atom in corresponding layer and greyscale indicates amount of current passing through atom: black for maximum current to white for zero current. White lines mark positions of profile cuts displayed in Fig. 11.

atomic layers. As seen in Fig. 9(a), in the first stages of propagation the distribution is basically governed by the crystal geometry and, consequently, the maximum current is mainly concentrated in the center of the region (just where the tunneling active atom is located). However, an interesting change then occurs, which is related to the gradual formation of a Bloch wave inside the crystal: propagation becomes only possible in directions consistent with the metal band structure. It is well-known that the Fermi surface of Au is nearly spherical with an influence of the crystal potential most apparent in the $\langle 1\,1\,1\rangle$ directions, where *necks* open up, and approach the border of the BZ. For energies ≈ 0.8 eV above the Fermi level, further smaller necks also develop in the $\langle 100\rangle$ directions, yielding a constant energy surface, as shown in Fig. 10 ($E = E_F + 1.0$ eV). Hence, as open necks symbolize lack of propagating states, electron transport is totally suppressed both along $\langle 1\,1\,1\rangle$ and $\langle 1\,0\,0\rangle$ directions, as a consequence of the interaction of the electron with the crystal lattice. Because Au layers in Au/Si systems tend to grow in the $\langle 1\,1\,1\rangle$ direction [27,51,52], the electron beam has then to make way sidewards, which can be clearly seen in Fig. 9b, showing an opening of the injected beam with very little current remaining in the central plane atoms, corresponding to the $\langle 111\rangle$ forward direction perpendicular through the film. A second effect of the metallic potential is the formation of narrowly focused *Kossel-like lines* that can also be understood by looking at the constant energy surface displayed in Fig. 10. Semiclassically, the direction of propagation of an electron with momentum $\vec{k}_0$ is dictated by the vector perpendicular to the constant energy surface at $\vec{k}_0$. Hence, electron propagation is strongly enhanced in directions corresponding to the rather flat parts of the constant energy surface. Since all electrons with wavevectors inside these extended areas travel in the

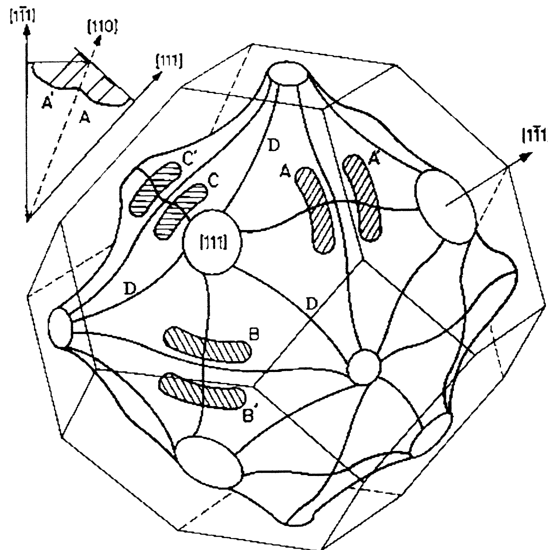


Fig. 10. Schematic shape of Au constant energy surface at $E = E_F + 1.0$ eV. Hatched areas mark some of the regions contributing most to projected 2D current distribution in $\langle 1\,1\,1\rangle$ direction.

same real-space direction, collimated beams could result. In the constant energy surface of Au, one can identify some flat parts, as indicated in Fig. 10. The closer the angle of the corresponding $\vec{k}$-vector comes to the $\langle 1\,1\,1\rangle$-direction, the more populated with electrons the states will be, which is why the area A, instead of the equivalent A', will dominate the current transport. Furthermore, regions *B* and *C* are symmetry-related to A, giving the pattern a three-fold nature, which is just the pattern obtained in $\vec{k}$-space for the semiclassical limit displayed in Fig. 12(c). It is worth stressing how different this *hollow* tetrahedron pattern is from the one obtained with a FE model, where the $\vec{k}$-spreading of the initial tunneling injection leads to a *solid* cone of roughly 20° angular width.

The formation of these narrowly focused Kossel-like lines, as caused by the Au band structure, can account for the nanometric resolution reported in the Milliken experiment: as seen in Fig. 9, the width of these lines is typically 3–4 atomic distances, which is better appreciated in Fig. 11, where cuts perpendicular through these lines at two different locations are performed and the intensity along these cuts is depicted. The derived width of 10 Å is in good agreement with the experimentally obtained value of 15 Å, which is a value very different from the FE result, that, after propagation through a 100–150 Å thick Au-film would predict $\approx$80 Å beam size.

We now focus our attention on how the current distribution in the metal matches the available states in the semiconductor. Thus, we calculate the 2D interface current distributions inside the Au interface BZ according to (3.24). Figs. 12(a) and (b) display the reciprocal-space pattern obtained in the purely elastic limit, where the pattern does not change during the passage through an infinite number of layers, because there are no processes enforcing $\vec{k}_{\|}$-pickups of the electrons after injection. As a direct reflection of the Au(1 1 1) band gap, which projects onto $\bar{\Gamma}$ in the center of the hexagonal interface BZ, the distribution possess a ring-like structure with a six-fold symmetry. It is again worthwhile to check how different this ensemble is from the FE prediction, which would inject the main current with $\vec{k}_{\|} = 0$ just in the central part. With the help of this reciprocal-space current distribution, we can now address

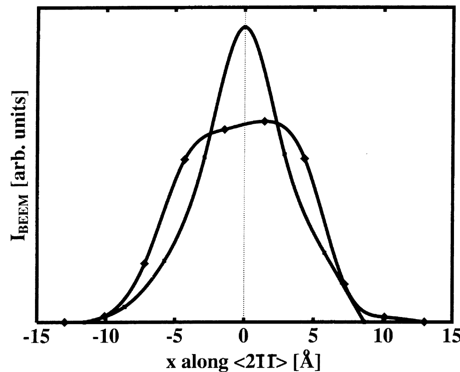


Fig. 11. Intensity profile along two perpendicular cuts through one of Kossel-like lines (locations of cuts are marked in Fig. 9 by white lines).

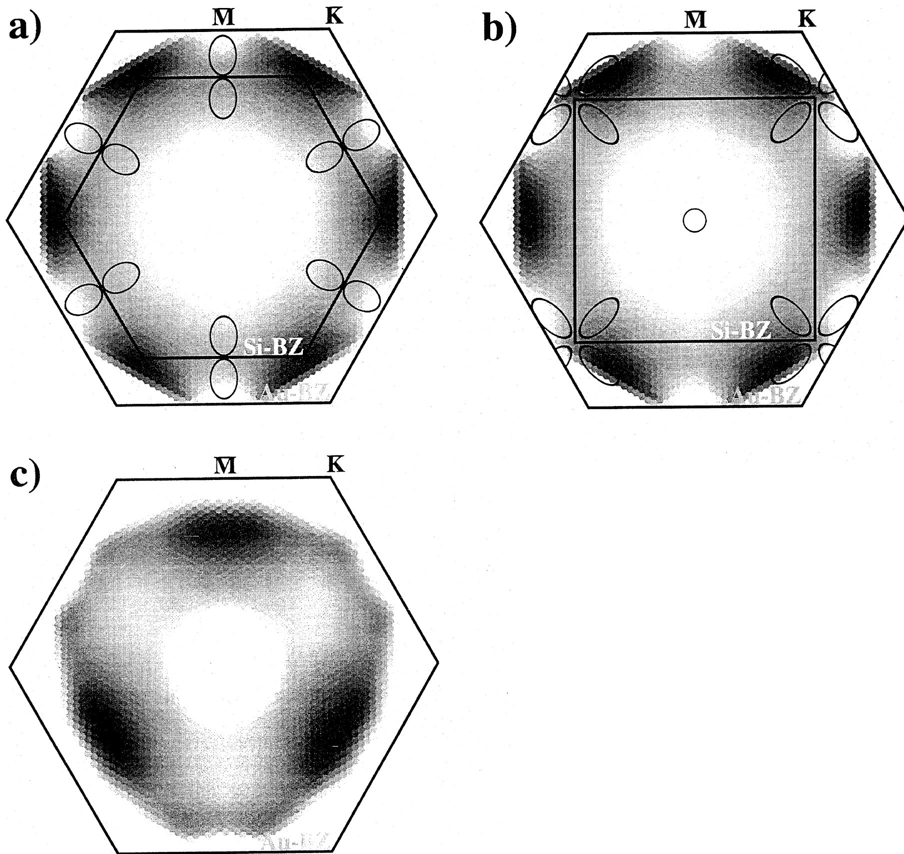


Fig. 12. Reciprocal space BEEM current distribution in 2D Au interface BZ at $E = E_F + 1.00$ eV (black represents high current, and white zero current: quadratic relation between intensities and grey scale is used to stress variations in intensities). For imaginary part of energy small enough ($\eta = 0.001$ eV) large coherent spatial region is found, where the current distribution is six-fold symmetric, as it is expected from quantal considerations. (a) and (b) shows such result, where projected ellipsoids from Si CBM have been plotted for (1 1 1) Si orientation in (a) and for (1 0 0) Si orientation in (b). Remapping of these ellipsoids by reciprocal vectors from the Si BZ is also shown. Notice that only regions inside these ellipses allow matching between states in the metal and the semiconductor, i.e., injection of electrons in the semiconductor. High enough imaginary part added to energy ($\eta = 0.1$ eV) results in loss of coherence even in regions close to surface, and three-fold symmetry, akin to semiclassical calculations, is found as shown in (c). Electrons only can be injected into the semiconductor where suitable projected conduction band minima allow matching between states in the metal and the semiconductor. These regions are visualized in (a) for the 1 1 1 orientation and in (b) for the 100 one by using an effective mass approximation for bands in the semiconductor and by drawing the different ellipses where corresponding states exist.

the explanation of why $I(V)$-curves on Si(0 0 1) and Si(1 1 1) substrates are so alike, despite their very distinct electronic structure. The matching between the current distribution in the metal with the available states in the two orientations of Si is illustrated also in Fig. 12(a) for the (1 1 1) orientation and in Fig. 12(b) for the (0 0 1)

one. The ring-like distribution, in the zeroth order, can be described as azimuthally symmetric, making it highly plausible why the BEEM spectroscopic curves in the two different Si orientations show similar onset and magnitude. This $\vec{k}_{\parallel}$ distribution makes it unnecessary to introduce, as a further hypothesis, the lack of $\vec{k}_{\parallel}$ conservation at the interface. In fact, the resulting distribution, generated by elastic scattering with the periodic lattice, is such that this issue is no longer a crucial one for theory.

The change in symmetry of the current distribution, when inelastic effects are included in Au, also has important consequences in the collector current, and, is discussed further below in Section 4. As shown in Fig. 12(c), if the current distribution is three-fold (i.e., the thickness of the metallic film is larger than the coherence region defined by η), the matching of this distribution with the available Si-states is better for the (1 1 1) orientation of the semiconductor than for the (1 0 0) one. Consequently, a larger BEEM current is obtained for the (1 1 1) orientation [53], in agreement with the available experimental evidence [27].

3.2.3. $CoSi_2$/Si(1 1 1) interface

As in the Au/Si case, we concentrate on the purely elastic-limit current at one particular energy to obtain an estimate of the beam evolution within the silicide base. In Fig. 13(a), the real-space current distribution, after propagation through a 30 Å $CoSi_2$(1 1 1) film (typical thickness in the experiments), is displayed. Contrary to the case of Au, the absence of a projected band gap in the normal film direction causes the maximum current to be concentrated in the region close to $\vec{k}_{\parallel} = 0$. However, in this case, an effect somewhat opposite to the sideward deflection in Au(1 1 1) is observed in form of a strong *focalization* along the $\langle 111 \rangle$ direction itself. The FWHM of 8.9 Å, extracted from the intensity profile in Fig. 13(b), is remarkably similar to

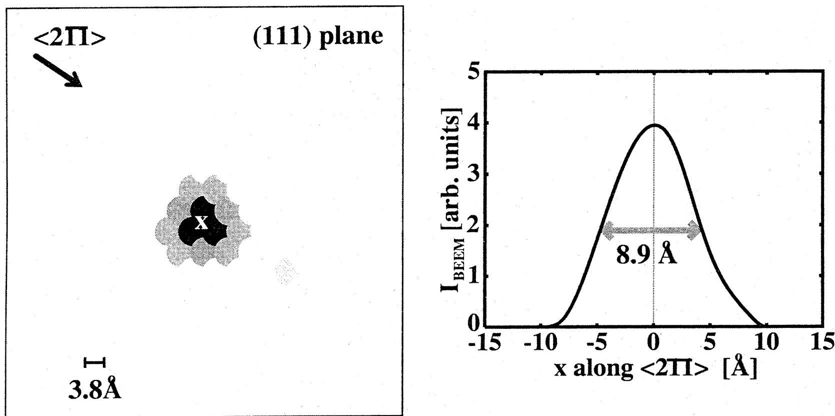


Fig. 13. (a) Real-space BEEM current distribution after propagation through 30 Å $CoSi_2$(1 1 1)–(1 × 1)Si-rich terminated film. Injection from tip at $E = E_F + 1.50$ eV occurred in center of shown plane (white X). Each dot represents one atom in Si-type layer and linear greyscale indicates the current intensity at each site: black maximum to white zero. (b) Cut through focused beam in $\langle 2\,\bar{1}\,\bar{1} \rangle$ direction, from which FWHM of 8.9 Å can be derived.

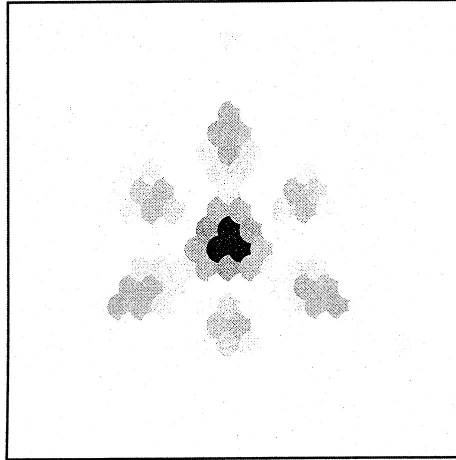


Fig. 14. Same as Fig. 13(a), but after propagation through hypothetical 70 Å film. Area covered is identical to Fig 13(a). Notice theory predicts that, even after propagation by a thick layer, the beam is still mainly focused around forward direction.

the experimental value indicated by the $\approx$10 Å width of the individual point defects. It is interesting to see if this focalization effect is also found with thicker films. In Fig. 14, we show the real-space current distribution for a hypothetical 70 Å thick film. The degree of focalization exhibited in this figure is rather impressive, because of the large number of layers the electrons traversed. The result shows that narrow collimation should be an intrinsic feature of $CoSi_2$(1 1 1) oriented films, which might be exploited in new applications.

In order to shed light on the physics underling this fascinating focalization phenomenon, Fig. 15 provides the constant-energy surface sheet corresponding to the eight-hole band, which exhibits the highest group velocity and, therefore, is the most influential on transport issues. From the figure, one can identify three flat terraces on the bulges of the star-shaped sheet. The normal to these plain terraces points almost exactly in the $\langle 1\,1\,1\rangle$ direction. But, since the perpendicular vector to the location on such a surface defines the direction of the group velocity of electrons travelling with corresponding momentum, a focused propagation of the latter, perpendicularly through (1 1 1) oriented samples, naturally ensues, as seen in Figs. 13 and 14.

As already mentioned, there is a puzzling observation of a surface 2×1 reconstruction on a BEEM image recorded for the interface $CoSi_2(1\,1\,1)/Si(1\,1\,1)$ (Fig. 16(a)). This effect must be related to tunneling, though it is not simple to explain, because the injected tunneling current is held constant in the experiment. The first explanation was given in terms of microscopic variations of the energy distribution of tunneling electrons, caused by atomic-scale variations in the effective tunneling barrier. Moreover, variations in the $\vec{k}_{\parallel}$-distribution were suspected to be partially responsible for this effect. However, on the basis of tunneling planar theory, they were estimated to be less important than the corresponding energetic variations [26]. In order to analyze this effect, Reuter et al. [25] computed the current impinging

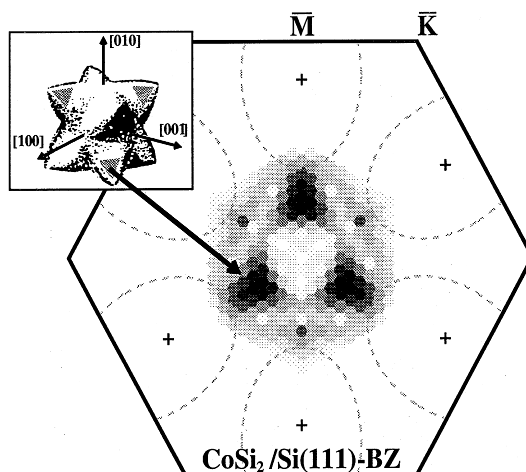


Fig. 15. Reciprocal-space BEEM current distribution within 2D interface BZ after propagation through 30 Å $CoSi_2$ (1 1 1)–(1 × 1)Si-rich terminated film ($E = E_F + 1.50$ eV, linear greyscale ranges from black maximum to white zero). Also shown are *Si* CBM ellipses (dashed grey lines). The inset contains constant energy surface sheet mainly responsible for current propagation: shaded planar terraces point in ⟨1 1 1⟩ direction and correspond to dark patches of 2D current distribution.

on the interface, as a function of the STM tip position (Fig. 16b). Results for the BEEM current are shown in Fig. 16 (all the effects related to transmission over the SB and/or back-injection of electrons have been neglected, because they are assumed to be rather insensitive to the position of the tip). This result reproduces two different experimental aspects: (a) The BEEM current is out of phase with the topographic corrugation, and (b) the relative variation in the BEEM current from the high to low chain is approximately 20%, which is in accord with the experimental value of about 10%, taking into account that a quantitative comparison was not intended and that features, such as the transmission through the interface have not been included in the calculation. The theory also explains the reason for the anticorrugation: injection of electrons into surface states causes the current to decay faster for the high chain than for the lower one.

3.3. Transmission at MS interface

As already mentioned, the first and most simple approach to the transmission over the SB at the MS interface, $T(E)$, is to fix its value at a constant. A better approach is to solve a 1D model for a step-like barrier, which already incorporates the square-root energy dependence. It is easy to understand that trying to make the 1D barrier more realistic is not worth the effort invested. However, there are reasons to extend the calculation beyond 1D models, because a correct description of the BEEM current requires a good description of the transmission over the barrier, which is necessary to avoid associating features, originating at the transmission (e.g., some characteristic energy dependence), to any of the other steps, where they would

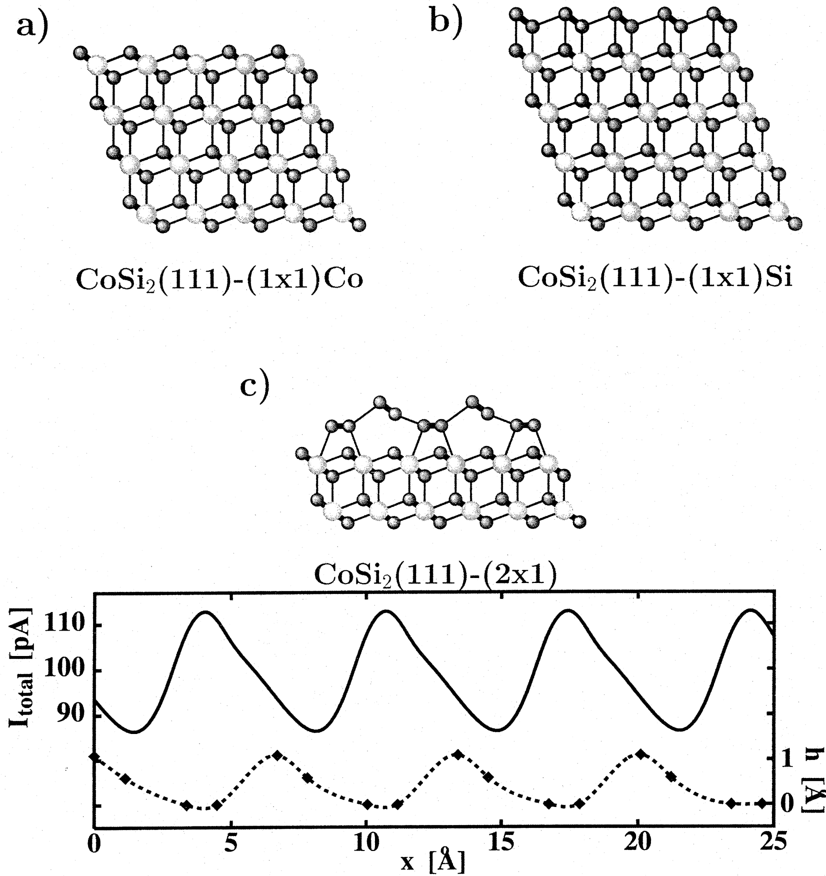


Fig. 16. Schematic representation of (1 × 1) (a) Co and (b) Si-rich phases and (c) (2 × 1) reconstruction on $CoSi_2(1\,1\,1)$. In lower panel, it is shown computed BEEM current through interface (solid line, left y-axis) compared with surface topography along considered STM scan line (dotted line, right y-axis).

be difficult to fit or would result on unphysical interpretations. On the other hand, BEEM presents us with another opportunity to test theories claiming to describe electronic processes at surfaces and interfaces, so it is reasonable to use this technique to test them thoroughly.

3.3.1. Approximate matching methods

First of all, we discuss an approximate method, which is used for studying the transmission coefficient for the Au/Si interfaces. The problem with these interfaces is their mismatch, whereby the transmission coefficient requires a full calculation with complicated conditions on the $k_{\|}$-component. To avoid this problem, the metal band structure is approximated by a FE band and the transmission coefficient, for a given $\vec{k}_{\|}$ and E, is calculated assuming that the incoming wavefunction is a plane-wave. Likewise, for the semiconductor, it is appropriate to introduce an approximate

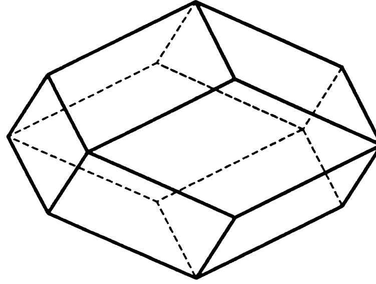


Fig. 17. 3D JZ used to compute approximated transmission coefficient at interface.

model, that, keeping the main physics of the problem, allows a rather simple calculation of the transmission coefficient. This is provided by the so-called first Jones-zone approach of the Si band structure (Fig. 17) which we could regard as the equivalent, for the semiconductor, of the free electron model for the metal.

Solving the matching problem in the *Jones–Zone* (JZ) approximation has the advantage of providing an analytical solution. As an example, we first solve the 1D situation, which incorporates most of the features found later in the 3D case. This is achieved by analyzing a two-band model with a narrow gap [54]. The wavefunction in the metal is assumed to be a combination of two plane waves travelling towards $+\infty$ and $-\infty$ with $k_\perp = \sqrt{2E - k_\parallel^2}$

$$\phi_{\mathrm{M}} = \mathrm{e}^{-\mathrm{i}zk_\perp} + r\mathrm{e}^{\mathrm{i}zk_\perp}. \tag{3.27}$$

Inside the semiconductor, the transmitted wavefunction is computed, by using a nearly FE model, where a $2|v|$ gap is opened, i.e.,

$$\phi_{\mathrm{S}} = t\left[\mathrm{e}^{\mathrm{i}q+(g/2)z} + s(\epsilon)\mathrm{e}^{\mathrm{i}q-(g/2)z}\right], \quad s(\epsilon) = \frac{\epsilon + \sqrt{\epsilon^2 - v^2}}{v}, \tag{3.28}$$

the wavevectors being measured with respect to the BZ edge, $q = k - \frac{1}{2}g$, and energies are referred to the midgap point, $\epsilon = E - \frac{1}{2}g^2$. The reflection, $R(E) = |r|^2$, and transmission coefficients, $T(E) = 1 - R$, are obtained by matching values and slopes at the interface. Since the main interest is to match states with $\vec{k}_\parallel$ near the $\overline{M}$ or $\overline{K}$ points, we particularize now to the BZ edge and write the reflection coefficient in a form that facilitates the generalization to the 2D interface problem

$$r = \frac{((1+\gamma)/(1-\gamma)) - ((s-1)/(s+1))}{((1+\gamma/1-\gamma)) + ((s-1)/(s+1))}, \qquad \gamma = \frac{1 - (g/2k_\perp)}{1 + (g/2k_\perp)}. \tag{3.29}$$

In the near threshold region, as $v \to 0$, this transmission coefficient behaves like that of the FE model and changes abruptly with a square-root law (measured with respect to the semiconductor conduction band minimum).

The next step is to improve the band structure description to account for 3D features, which, in the metal, has been done within the FE model by splitting the energy between the interface's perpendicular and parallel directions. In the semiconductor, an intermediate level of sophistication is achieved via de JZ

approximation, which improves the model significantly. This zone determines a minimum number of reciprocal $\vec{g}$ vectors needed to obtain a reasonable description of the semiconductor band-structure, while the matching still can be performed analytically for high-symmetry $\vec{g}$ points. It is assumed that the transmission coefficient does not change very quickly with $\vec{k}_{\parallel}$, and the equations can be extended from high-symmetry points to its near vicinity.

Without going into further details, let us only mention that for a Si(1 1 1) surface, we obtain the reflection coefficient by matching at the $\overline{M}$ point, obtaining an effective value for γ that can be inserted in (3.29)

$$\gamma = \gamma_1\gamma_2 = \frac{(1-(2|g_{111}|/3k_{\perp}))(1-(5|g_{111}|/6k_{\perp}))}{(1+(2|g_{111}|/3k_{\perp}))(1+(5|g_{111}|/6k_{\perp}))}. \tag{3.30}$$

This approach yields transmission coefficients about 10% larger than (3.3), resulting in around 30% more intensity injected at $V = 1.5$ eV (Fig. 18).

A similar analysis for the Si(1 0 0) orientation, now matching at the $\overline{K}$ point, gives the following value for γ:

$$\gamma = \gamma_1\gamma_2 = \left\{1 - \frac{(|g_{100}|/2) + q_{\parallel}}{\sqrt{2\left[\epsilon - \frac{1}{2}((|g_{100}|/2) - q_{\parallel})^2\right]}}\right\} \times \left\{1 - \frac{(|g_{100}|/2) - q_{\parallel}}{\sqrt{2\left[\epsilon - \frac{1}{2}((|g_{100}|/2) + q_{\parallel})^2\right]}}\right\}. \tag{3.31}$$

In (3.31) and (3.32), g_{111} and g_{100} are particular g-vectors associated with the Si structure. Eqs. (3.29) to (3.31) can be used as the basis for calculating the reflection coefficients for the Au/Si(1 1 1) and Au/Si(1 0 0) interfaces. For example, in (3.26), T is provided by this approach, and the purely elastic BEEM-current is given by this equation, calculating $J_{n-1,n}$ via the above GF method, and taking $S = 1$.

3.3.2. Variational principle method

The calculation of the transmission coefficient can be performed in a more convenient way, when the metal and the semiconductor structures are well matched, as is the case of some silicides-silicon. Generally speaking, the first-principles problem of finding the transmission at the interface reduces to one of finding the wave-functions for a composite of two materials. By assuming perfect 2D periodicity in the direction parallel to the interface, the task is effectively reduced to a 1D matching problem for each $\vec{k}_{\parallel}$. In principle, there are ways to find the wave-functions for such a composite system given either the GFs functions [54] or the Bloch states of both parts. A simpler example arises when one of the two materials is the vacuum, where one computes the reflectivity of a surface [55]. Stiles and Hamann [56] proposed a different approach, feasible for the ideal interface, where the variational principle is used to solve the Schrödinger equation with appropriate boundary conditions.

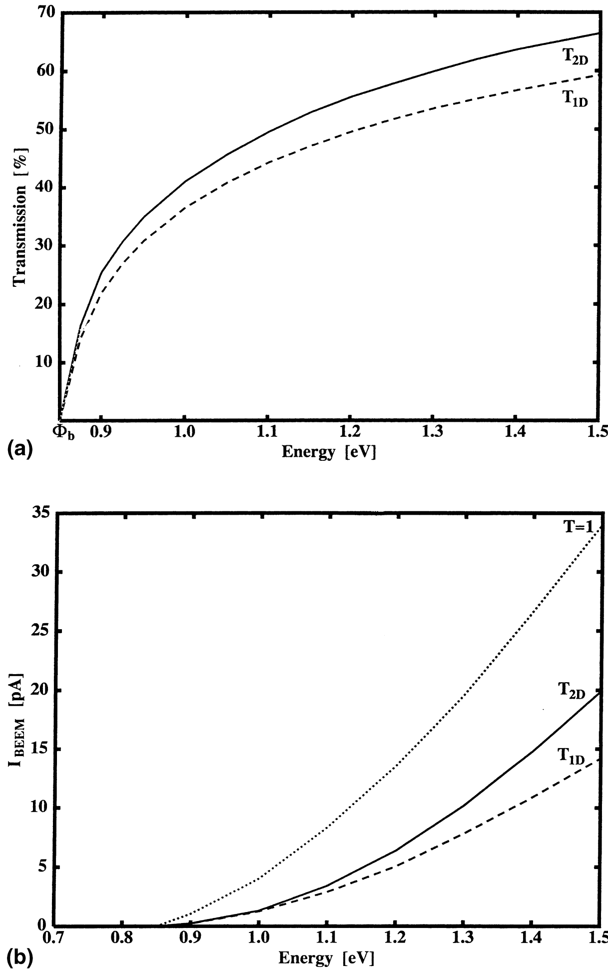


Fig. 18. (a) Energy dependence of transmission coefficients computed either in JZ approximation (solid line) or in simple 1D step barrier approximation (dashed line). Parameters entering into computation correspond to the Au/Si(1 1 1) system, and $\vec{k}_{\parallel}$ dependence has been averaged out over ellipsoids, where electrons are allowed to be injected in semiconductor. (b) Effect of different transmission coefficients on $I(V)$ characteristic (same parameters as (a)).

Boundary conditions can be specified, as in a transfer-matrix method, by giving the wavefunction and its normal derivative at some layer, $l-1$. The Schrödinger equation can be used to propagate the solution to the next layer, l, defining a transfer matrix, $\hat{\Gamma}_{l-1,l}$ that solves the problem. Moreover, Wachutka [57] has shown that by using boundary conditions involving the layer behind, $l-1$, and the layer in front, $l+1$, better numerical stability is achieved.

Such a microscopic model provides new insight that cannot be obtained from simpler approaches (e.g., 1D step models based on effective masses). Stiles and

Hamman have studied the transmission through $NiSi_2(1\,1\,1)/Si(1\,1\,1)$ and $CoSi_2(1\,1\,1)/Si(1\,1\,1)$ interfaces, which have the great advantage that the lattice parameter mismatch is less than 1%, so very good interfaces can be achieved. Two different interfaces are known to form on the $NiSi_2$, depending on growth conditions: either it grows with the same crystallographic orientation as Si (type A) or it grows rotated 180° (type B) [58,59]. Fig. 19 shows a striking result: both interfaces are strongly reflecting, but transmission for interface B is about 2–3 times lower than for interface A. Although the determination of the SB can be quite insensitive to effects that do not change the threshold position, but merely the absolute value of the BEEM current, such a strong variation in transmission coefficients can influence the interpretation of BEEM data. In particular, it can affect the determination of e–e inelastic MFP on thin metalic films.

An interesting effect, appearing in $CoSi_2(1\,1\,1)/Si(1\,1\,1)$, is the apparent existence of a delayed onset. The SB value of $V_{SB} = 0.66$ derived for this interface from photoresponse measurements, cannot be found in BEEM according to theory, because the analysis of the $CoSi_2(1\,1\,1)$ projected band structure shows that there are no states overlapping with allowed CBM in Si(1 1 1) [60–62]. This effect has been also found in independent calculations [25], and may be used as an indicator of the importance of $\vec{k}_{\parallel}$-conservation at the interface. In fact, experimental data measured ex situ on this system by Kaiser et al. [63] reproduced the theoretical expectation for a delayed onset at about 0.85 eV, a considerable larger energy than the real SB value. However, further measurements by Sirringhaus et al. [64], resulted in an onset agreeing well with the accepted SB value (Fig. 20). They interpreted this observation as an indication of how small errors in the band-structure calculation can affect the theoretical determination of the SB, and pointed out that allowing a small change of $\approx$0.6 eV in the $CoSi_2$ bands could explain the lower observed onset. A different, and also plausible, explanation might be based on a relaxation of the $\vec{k}_{\parallel}$-conservation condition, which deforms the $\vec{k}_{\parallel}$, distribution so that some overlap with semiconductor states is achieved. In fact, it would possible to speculate, from the experimental data presented in Fig. 20, that the open circles values are consistent with the theoretically predicted onset at about $\approx$0.8 eV, as deduced from the extrapolated slope at high energies. This allows the effect of non-conserving $\vec{k}_{\parallel}$ processes to be estimated by comparing with the data taken on top of a dislocation (crosses), where some amount of momentum redistribution is expected. Indeed, this indicates that, for energies away from the onset by more than 0.2 or 0.3 eV, the non-conserving processes amounts to less than 10% of the current.

4. Metal base scattering processes

Going beyond the one-electron model implies introducing in the propagation of the electron in the metal base or the semiconductor different electron scattering mechanisms, such as e–e and e–ph interactions. In this section, we discuss how different scattering processes in the metal base can be taken into account in the calculation of the BEEM current.

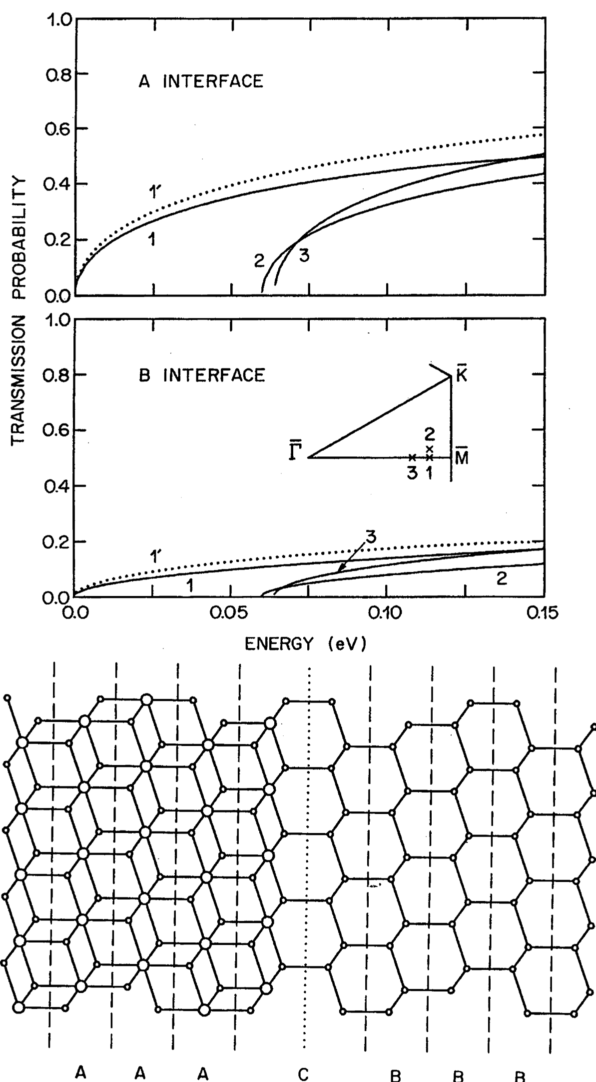


Fig. 19. First-principles calculation for transmission coefficient through $NiSi_2(1\,1\,1)/Si(1\,1\,1)$ type A and B interfaces (schematically represented in low panel). Three selected $\vec{k}_{\|}$ vectors on irreducible wedge of 2D BZ are shown to illustrate $\vec{k}_{\|}$ dependence and energy dependence of the transmission coefficient. Note important difference found for transmission on two interfaces that are otherwise similar. This is a marked effect that only true microscopic model can account for (from [60,61]).

4.1. E-space parametrized MC

The necessity of more sophisticated models for describing the different scattering effects which are believed to be relevant in real systems, directed several groups to perform MC simulation. The MC approach is an efficient way to obtain numerical

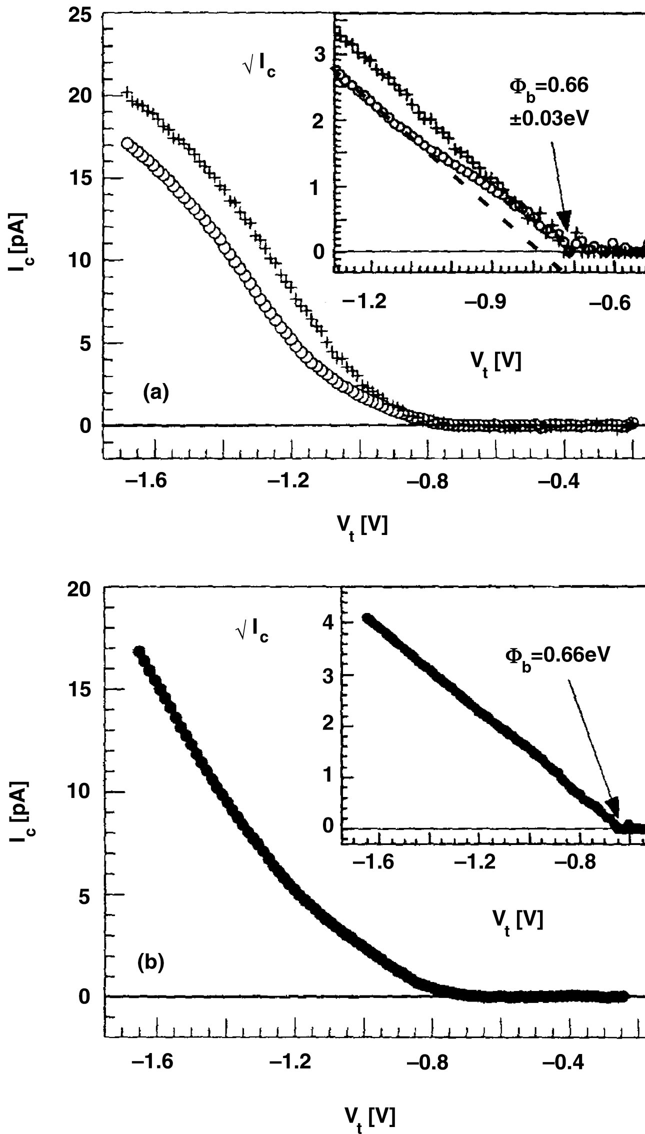


Fig. 20. Onset of $I(V)$ spectra of 35 Å $CoSi_2$/*n*-Si(1 1 1) film. In the top panel, it is observed that neither measurement directly on top of dislocation (crosses), nor in adjacent dislocation free region (circles), show any sign of delayed onset. This becomes even more apparent in inset, where $I_B^{1/2}$ has same slope for energies about 0.2–0.3 eV higher than threshold. Note, however, how dislocations free curve (circles) can be linearly extrapolated (dashed line) to yield the theoretically predicted onset at $\approx$0.8 eV (from [64]).

solutions to the semiclassical *Boltzmann transport equation*, which is very appropriate, when the main interest is to discuss ensemble averages. Moreover, it is useful to keep in mind that this method requires an underlying physical model, so that

probabilities can be asigned to different scattering mechanisms. The first objective behind the use of MC in BEEM was to include attenuation to extend the energy range, where the pure ballistic approach sketched in Section 3 is valid. A direct estimation for the region of validity of ballistic theories can be obtained by comparing experimental values with theoretical results obtained without taking into account losses. Recent realistic calculations, incorporating a fair description of all the four steps involved in the theory give, at most, a validity range of $V_T \leqslant V_{SB} + 0.1$ eV [65] when compared with experimental $I(V)$ curves measured on favourable conditions, where losses are expected to be less important (small metalic width and low temperature). Therefore, it is clear that the experimental data cannot be represented by a pure ballistic theory always, and it is important to introduce several scattering mechanisms for producing attenuation. Indeed, this becomes crucial for thicker films, higher temperatures, or lower metalic densities.

If only one inelastic scattering channel is considered as responsible for a given energy loss, and, consequently, for the attenuation of the BEEM current, it can be conveniently modelled within a relaxation-time formalism. As a practical example, let us choose the e–e interaction as the physical mechanism responsible for attenuation. This is very convenient, because it is known that the average loss is as large as half the kinetic energy [66]. Therefore, for $E < 2V_{SB}$, once the electron has undergone a single inelastic event, it is likely it will be lacking the necessary energy to overcome the barrier, which is why the BEEM current becomes attenuated. In a relaxation-time approach, the hypothesis is made that $\kappa \mathrm{d}l$ gives the probability that an electron will lose energy, while travelling in the interval $[l, l + \mathrm{d}l]$. Therefore, the probability that the carrier keeps its energy after travelling a total distance $l + \mathrm{d}l$ is simply

$$P(l + \mathrm{d}l) = P(l)(1 - \kappa\, \mathrm{d}l), \tag{4.1}$$

which is easily integrated to give the standard exponential attenuation law

$$P(l) \propto \mathrm{e}^{-\kappa l}, \quad \kappa = \frac{1}{\lambda}. \tag{4.2}$$

We observe that, within this trivial MC simulation, the exponential decay is a result of two key assumptions: first, that the probability to undergo an e–e interaction is independent of the carrier's past history, and second, that the loss is large enough to discount the electron from the measured current. This would not be the case for a quasi-elastic e–ph interaction, or any elastic process we wish to consider, and in these cases the redistribution of current among available states in $\vec{k}$-space is the crucial feature.

Therefore, the problem arises of obtaining, simultaneously, an elastic or quasi-elastic (λ_{el}) and a inelastic (λ_{in}) MFP from a single-measured attenuation length (λ_{at}), when both processes are assumed to be equally important. Ludeke and Bauer provided a simple 1D model to deal with such a situation [67]. They assumed that the scattering cross-section for elastic scattering is isotropic, so that the current is equally redistributed amongst the forward and backward directions. After such an event, the carrier is either scattered inelastically (and removed) or scattered elastically again, thus producing a series of events that one needs to sum to convergence. The

probability for a carrier to reach the MS interface results from adding the probability to travel the metal layer without attenuation, $\exp(-l/\lambda_{at})$, plus the probability of being scattered elastically, $(1-\exp(-l/\lambda_{el}))$, times the probability of travelling an extra distance $\overline{n}\lambda_{el}$, before reaching the interface, $\exp(-\overline{n}\lambda_{el}/\lambda_{in})$. The average number of elastic scatterings, when the scattering cross-section is isotropic, can be estimated from the distance travelled on the path to the interface, and the mean distance between elastic scaterers: $\overline{n} \approx (2l/\lambda_{el})^2$. Combining this, Ludeke and Bauer [67] proposed the following functional dependence on the metal width, l:

$$P(l) = \exp\left(-\frac{l}{\lambda_{at}}\right) + \left(1-\exp\left(-\frac{l}{\lambda_{el}}\right)\right)\exp\left(-\frac{4l^2}{\lambda_{el}\lambda_{at}}\right). \tag{4.3}$$

Two limiting cases are of interest: $\mathrm{d}(\ln P)/\mathrm{d}l$ approaches $(1/\lambda_{el})-(1/\lambda_{at})$ for $w \to 0$ and $-1/\lambda_{at}$ for $w \to \infty$. It is not surprising that, for thick layers, the exponential decay of the current involves λ_{at}, since it was assumed at the beginning, but even this simple calculation shows that, for thin films, the elastic scattering is mixed with the attenuation length in some way, which depends crucially on the hypothesis used to set up the model. In this example, a rule to combine the different MFP to obtain the total attenuation length has to be introduced. A popular one is *Matthiessen's rule*

$$\frac{1}{\lambda_{at}} = \frac{1}{\lambda_{in}} + \frac{1}{\lambda_{el}},$$

which is justified when the different processes contribute independently to the current attenuation. For the particular case of mixing elastic and inelastic processes, it is not clear whether this is the case or not, and other rules exist, e.g., Fermi age theory [68]. None of them is fully justified from a theoretical point of view [69], so results based on this approach should be interpreted with care.

From these simple considerations, it can be understood how the inelastic e–e interaction inside the metal film produces a typical exponential attenuation, which can be experimentally measured. However, it is useful to keep in mind the two limiting situations: for thin layers, multiple reflections between the surface and the interface must be taken into account and might result in deviations from the pure exponential law. For wide films, e–ph and static defects might play a dominant role. It is also clear that mixing elastic and inelastic mechanisms is not a simple task, and, in general, there is not a precise rule to obtain one single attenuation length, λ_{at}, by mixing several processes, e.g., $\lambda_{e\text{–}e}$, $\lambda_{e\text{–}ph}$, $\lambda_{e\text{–}d}$, etc. Therefore, several authors have tried to obtain the relaxation times associated to different scattering channels simultaneously, by running MC simulations where the probability of scattering is taken as a parameter to be fitted to experiment. In these models, the different scattering probabilities being fitted to experiments only depend on energy, which is referred to as *E-space MC simulation* in the literature. Results within this approach have been shown to depend significatively on all the details of the model, and it is possible to find in the literature MFPs for the same material differing by an order of magnitude [18,19,70], or proposals to adopt a smoother energy dependence for the e–e MFP than the one derived from the quasi-particle concept and the Landau

theory of the Fermi liquid [17]. These dilemmas are also related to the intrinsic difficulty of minimizing a function depending on many variables, as a consequence of correlations amongst the different parameters, or even because of the existence of local minima that are very hard to distinguish from the real global one.

4.2. $\vec{k}$-space MC

A step forward applying MC techniques is achieved by setting up a physical model by fixing the various scattering mechanisms and avoiding the direct fitting of several parametrized MFPs to agree with the experiment. This must be done in reciprocal space, as scattering cross-sections are computed as a function of $\vec{k}$ and E. Such an approach was first proposed by Hohenester and collaborators [71]. The starting point is also a semiclassical Boltzman picture for the system: between scattering events, ballistic electrons (i.e., particles moving freely) are characterized by a momentum $\vec{k}$ and a position $\vec{r}$. The method is completely equivalent to finding a numerical solution to the Boltzmann equation by means of the ensemble MC technique [73,74].

The main capability of MC simulation is to describe scattering processes inside the metallic film or the semiconductor, but it is an inferior method for treating the elastic interaction with the periodic lattice (e.g., it cannot describe the existence of gaps). Therefore, Hohenester et al. proposed to use, as the initial current distribution, the elastic component computed by the ab initio GFs (Section 3.2). This is a practical way to suplement MC techniques with effects due to the tunneling step and the band-structure in the metal, both problems being hard to describe by MC alone.

Let us now consider an electron that is scattered from a state $\vec{k}$ into a state $\vec{k}'$ by a perturbing potential, V. The probability that an electron in state $\vec{k}$ makes a transition to state $\vec{k}'$ is

$$S(\vec{k}',\vec{k}) = 2\pi|\langle\vec{k}'|V|k\rangle|^2\delta\left[E_{\vec{k}'} - E_{\vec{k}}\right], \tag{4.4}$$

and the change in the distribution function entering in the Boltzmann equation is given by the number of electrons scattered into the state $\vec{k}$ minus the number of electrons scattered out of that same state

$$\frac{\partial f}{\partial t} = \int \left\{S(\vec{k}',\vec{k})[1 - F_{\vec{k}}] - S(\vec{k},\vec{k}')\left[1 - F_{\vec{k}'}\right]\right\} \mathrm{d}\vec{k}'. \tag{4.5}$$

A number of different scattering channels can be easily taken into account with this approach.

4.2.1. Electron–electron scattering

The e–e scattering probability for a free electron gas is given by [75]

$$P(\vec{k}) = 2\sum_{\vec{q}} \frac{4\pi}{q^2}\left[1 - F_{\vec{k}-\vec{q}}\right]\mathrm{Im}\left[-\frac{1}{\epsilon(\vec{q},\omega_{\vec{k},\vec{q}})}\right], \tag{4.6}$$

where $\omega_{\vec{k},\vec{q}} = E_{\vec{k}+\vec{q}} - E_{\vec{k}}$ is the energy of the excitation $\vec{q}$, and the dielectric function $\epsilon(\vec{q}, w)$ is computed by using different approximations (e.g., the *random phase approximation* (RPA)). This scattering probability includes most of the relevant energy dependence, even for models more realistic than the electron gas on a constant background potential, and has been shown that theoretical refinements to include band-structure effects, exchange, etc, can be added accurately by simply scaling the scattering cross-section by a constant factor.

4.2.2. Electron–phonon scattering

Building on simple, but well established, formalisms for the e–ph interaction in metals [69,76], Hohenester et al. [72] suggested that the related scattering probability could be found from a deformation potential model: for small displacements, a Taylor expansion is used to find the relevant matrix element, $\langle\vec{k}|\nabla v|\vec{k}\rangle$, which can be approximated by

$$(\vec{k} - \vec{k}')D \int_{|r|<r_s} \mathrm{d}\vec{r} \psi^*_{\vec{k}'} \psi_{\vec{k}} = \frac{(\vec{k} - \vec{k}')}{|\vec{k} - \vec{k}'|} \frac{2E_F}{r_s} j_1(|\vec{k} - \vec{k}'|r_s), \tag{4.7}$$

where $j_1(x)$ is a spherical Bessel function and the deformation potential, D, has been replaced by $(2/3)E_F$ (e.g., see [77]). The total scattering rate is

$$P(\vec{k}) = \sum_{\vec{q}} \left[1 - F_{\vec{k}-\vec{q}}\right] \frac{\pi q D^2}{\rho c} j_1^2(qr_s) \begin{cases} n_{\vec{q}} \delta(E_{\vec{k}+\vec{q}} - E_{\vec{k}} - w_{\vec{q}}), & \text{absorption}, \\ (1 + n_{\vec{q}})\delta(E_{\vec{k}-\vec{q}} - E_{\vec{k}} - w_{\vec{q}}), & \text{emission}, \end{cases} \tag{4.8}$$

where ρ is the metal density, c is the sound velocity, r_s the Wigner–Seitz sphere radius, n_q the Bose–Einstein occupation number for phonons at a given temperature T, and the dispersion relation is assumed for longitudinal modes described by a *Debye model*: $\omega_q = cq$. Optical phonons can also be introduced by using similar approximations.

4.2.3. Electron–defect scattering

A number of static defects, affecting the momentum distribution of carriers, can be considered by modeling their scattering potential: point imperfections (impurities, vacancies, interstitials), line defects (dislocations), and planar imperfections (stacking faults, grain boundaries). Usually, the mass of these defects is four or five orders of magnitude larger than the electron's, so that the process can be considered as purely elastic, which makes it easier to find an explicit expression for the collision term in the Boltzmann equation

$$\frac{\partial f}{\partial t} = -\frac{f_1(\vec{k})}{\tau}, \tag{4.9}$$

where f_1 is the deviation of the equilibrium distribution, and a relaxation time τ is introduced, which depends on f_1 and the probability of scattering $S(\vec{k}', \vec{k})$ [77]. Such an approach is useful, when f_1 is independent of the static potential, thus creating a

departure from the equilibrium distribution. As an example, consider a spherically symmetric point defect, so that $S(\vec{k}', \vec{k})$ becomes a function of the angle θ between $\vec{k}'$, and $\vec{k}$, and the constant magnitude $|\vec{k}|$ only. The angular dependence of the scattering probability can be exemplified by further assuming that the corresponding constant energy surface of the actual band structure is spherical, to obtain

$$\frac{1}{\tau} = 2\pi\rho_{\mathrm{d}}c \int (1 - \cos\theta)\sigma(\theta)\, \mathrm{d}(\cos\theta) \tag{4.10}$$

where $\sigma(\cos\theta)$ is the differential cross-section, and ρ_{d} is the density of the scatterers. It is observed that large-angle scattering becomes important, because it is very effective in restoring a spherically symmetric distribution.

When several scattering mechanisms act together, the simulation has to choose between them according to their respective probabilities [73,74]. The standard algorithm is as follows: if P_i represent these probabilities for $i = 1, N$ channels, they are first added together, $P = \sum_{i=1,N} P_i$, and a random number uniformly distributed between 0 and 1, r, is used to compare $r \times P$ to the different partial sums, $S_j = \sum_{i=1,j} P_i$. The mechanism j is selected, when $S_{j-1} < r \times P < S_j$, which ensures that channel j is chosen with a probability proportional to P_j.

After some scattering occurs, the metal band structure acts to restore an electron distribution compatible with the corresponding band structure. This elastic interaction, produced by the periodicity of the lattice, has been shown to be very effective, and propagation through only three or four layers is needed to completely re-establish a distribution compatible with the bulk band structure. Therefore, the *characteristic length*, where the distribution conforms again to the metalic band structure, is considerably shorter than the shorter MFP, due to any scattering mechanism. This makes it reasonable to introduce these effects by using a *rejection technique* [73,74]: any scattering event resulting in propagation along forbidden directions is taken to be a *self-scattering* event, and is ignored.

Using these ideas, Hohenester et al. analyzed a number of problems associated with scattering mechanisms that are hard to study with a GF formalism. As an example, the following question is: after the formation of highly focused beams by elastic interaction with the periodic potential, how effective are inelastic events as a defocusing mechanism? This calculation is very difficult in a GF formalism, but MC techniques can provide an answer: Fig. 21 shows that, after propagation of 100 Å in a Au film, the elastic beam is still more important than the electrons defocused by inelastic (e–e and e–ph) interactions [72].

4.3. Green's functions and e–e scattering

Inelastic effects, associated with the e–e interaction, can be incorporated into the GF formalism (Section 3.2) simply by adding to the energies a complex *optical potential*, $\eta(E)$, describing the attenuation of the waves

$$\hat{G}^R(E) = \frac{1}{E - \hat{H} + \mathrm{i}\eta}. \tag{4.11}$$

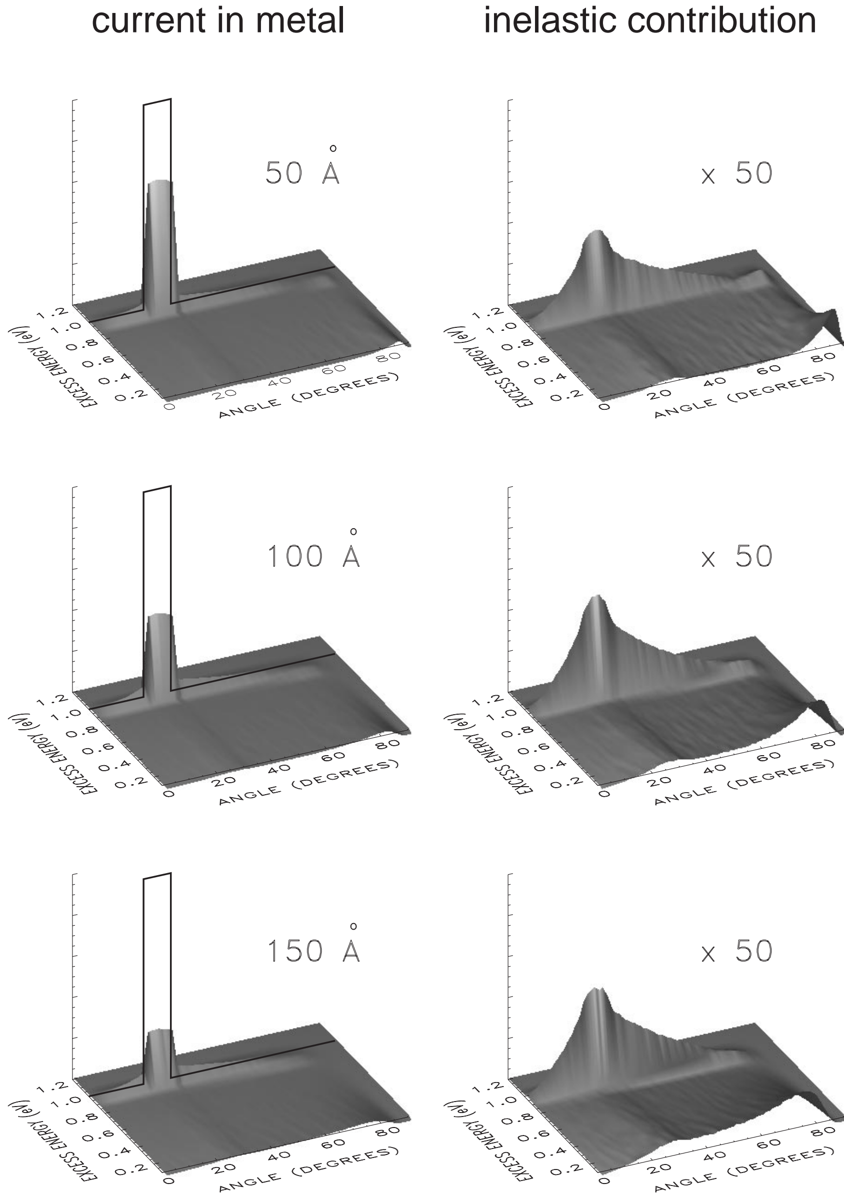


Fig. 21. MC simulation for defocusing effects on elastic beams by inelastic scattering with electrons and phonons. Focused beams (continuous line) are assumed to carry current in approximate directions 20° degrees away from normal at $E = E_F + 1.5$ eV. Upper, middle and lower panels show the elastic (focalized) and inelastic (defocused) currents, at 50, 100 and 150 Å from surface. Even at 150 Å, focused elastic beams still carries significant part of total BEEM current (from [72]).

When the limit $\eta \to 0$ is taken, this amounts to the usual mathematical trick that is introduced to ensure convergence on an infinite semicircle in the upper or lower complex plane, thus allowing to work with analytical GFs. Since it is inherent in GF methods to use $\eta \to 0^+$, it is useful to talk about a *coherence region*, defined roughly by $\lambda_c \approx (\sqrt{2E}/\eta)$, where quantum effects are important, while outside it the loss of coherence is appreciable and the scenario becomes semiclassical. From a physical point of view, it is clear that inelastic processes must introduce some loss of coherence. This effect has already been seen in Fig. 12, where a full quantal calculation is shown for small and large η values. The effect is that the initial six-fold current distribution is progressively transformed into a three-fold one as η is increased. This interesting effect reveals the building up of a *semiclassical scenario*, which is better discussed in reciprocal space, since the most prominent feature is the change in symmetry of the current distribution. A close examination of (3.24) shows that the decisive objects for symmetry are the retarded and advanced GFs. These objects can be computed, by using a semiclassical approximation due to Koster [78]. This approach is based on a stationary phase condition, which identifies a single $\vec{k}_0$, that happens to be the most relevant for electrons propagating in a given direction in real space, $\vec{R}_j - \vec{R}_1$ (hence, its semiclassical character)

$$\hat{g}^R_{j1}(E) = \sum_{\vec{k}} \frac{e^{i\vec{k}(\vec{R}_j - \vec{R}_1)} \phi^*_{\vec{k}}(j)\phi_{\vec{k}}(1)}{E - E_{\vec{k}}}$$

$$\approx -\frac{1}{\sqrt{(\partial^2\epsilon/\partial k_1^2)(\partial^2\epsilon/\partial k_2^2)}} \exp\left(i\frac{\vec{k}_0(\vec{R}_j - \vec{R}_1)}{\mid \vec{R}_j - \vec{R}_1 \mid} \right) \phi^*_{\vec{k}_0}(j)\phi_{\vec{k}_0}(1), \tag{4.12}$$

where $\partial^2\epsilon/\partial k_1^2$, $\partial^2\epsilon/\partial k_2^2$ are the two principal curvatures associated with the constant energy surface and the eigenvectors $\phi^*_{\vec{k}_0}(j)$, $\phi_{\vec{k}_0}(1)$ are obtained by diagonalizing the bulk Hamiltonian, $\hat{H}$. This expression admits a simple geometrical rule for identifying $\vec{k}_0$: it is the reciprocal space vector linking the origin to a point in the constant energy surface, where the normal to the surface is parallel to the propagation direction in real space. Therefore, it is possible to see that the GF symmetry is related to the symmetry of the constant energy surface (e.g., the Fermi surface), which in turn, is related to the symmetry of the reciprocal lattice, that must belong to the same space group as the real lattice. Thus, for example, an fcc film terminated by $\langle 1\,1\,1 \rangle$ planes has a three-fold rotational symmetry for the current distribution in reciprocal space (the current distribution in real space must also be three-fold symmetric). However, a different picture emerges for the quantal case: the Hamiltonian for the bulk material introduces an extra inversion symmetry, so vectors with $\pm\vec{k}$ contribute equally to the wave function. In the particular case we are considering, it is possible to see that, by projecting with equal weights on an fcc (1 1 1) face, Bloch waves travelling towards or away from the surface ($-\vec{k}$ and $+\vec{k}$, respectively), an extra mirror plane arises, so the symmetry becomes six-fold instead of three-fold, as shown in Figs. 12. Therefore, if an experiment that is sensitive to the symmetry of the currents in reciprocal space can be designed, that symmetry can be used to define an order parameter for studying the transition between the quantal and classical worlds.

Such an interesting goal is linked with quantum coherence studies [79] that can be relevant for several fields, such as, quantum computation, the single quantum transistor, quantum cryptography, transport in quantum dots, etc. In our particular example, the transition from six- to three-fold symmetry is driven by the loss of coherence, which can be controlled experimentally by turning on the inelastic interactions that continuously depend on parameters defining the system, such as, temperature, electronic density, film width, etc. Furthermore, external parameters breaking suddenly some symmetry can be varied to distinguish between $+\vec{k}$ and $-\vec{k}$ (e.g., a magnetic field can be turn on and off for breaking the temporal inversion).

5. Semiconductor scattering – BEEM currents and metal e–e mean free paths

After the electron has been injected into the semiconductor, a number of physical processes can affect the measured collector current. These have not yet been well investigated from a theoretical point of view. The two main points discussed in the literature are: the scattering of electrons by phonons, creating a back-injection current into the metal, and impact ionization, both of which have been studied by MC techniques.

Scattering of electrons by optical phonons, as a source of back-injected current, was first proposed by Schowalter and Lee [28] as a factor for modifying the effective transmission coefficient. They estimated an energy independent value $T(E) \times S(E) \approx 0.2$. Bauer et al. [18,19] also resorted to estimations for the energy independent value of $S(E) \approx 0.5$. Hohenester [34] studied, in more detail, the effect of electron scattering by acoustic and optical phonons, using a *deformation potential approximation*, which is a well-known approach for semiconductors, but whose accuracy is not well established for metals. The interesting feature found by Hohenester, in agreement with previous studies, is that the percentage of electrons backscattered into the metal depends strongly on energy (Fig. 22), especially in the near-threshold region, which suggest a contribution affecting the power law for the $I(V)$ characteristics. Effects associated with the depletion field in the semiconductor have also been investigated by Hohenester, by increasing the doping concentration from $1 \times 10^{15}\mathrm{cm}^{-3}$ to $5 \times 10^{15}\mathrm{cm}^{-3}$. Variations found of $S(E)$, however, are less than 10% for this five-fold increase in doping, indicating that this factor is less influential in explaining these experiments (Fig. 22).

The structure of the $I(V)$ functions for energies away from the threshold is rich, and deserves some attention. The parametrized models, set up to account for the main features, have been of great value in identifying the important physical mechanisms in that part of the spectrum. Although those models are still primitive, and depend on a heavy parametrization of the problem, useful information can be drawn, if fits are interpreted with some care. The relevant process at these energies is most likely impact ionization (Fig. 23). At energies larger than the band gap of the semiconductor, an electron can create an electron–hole pair between the valence and conduction bands. The pair can be subsequently separated by the electric field in the depletion region, and an additional electron is contributed to the BEEM current.

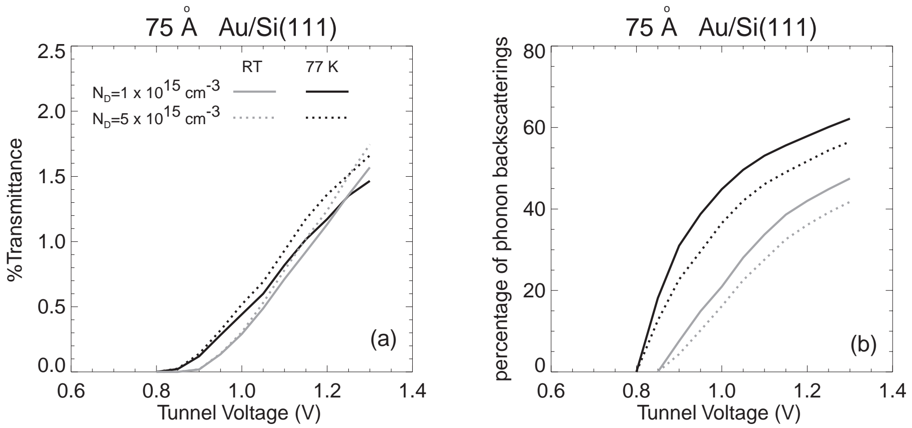


Fig. 22. At two temperatures (RT and 77 K), the influence of depletion field on interface transmitance is shown in left panel. On the right panel, the percentage of electrons backscattered from semiconductor to metal after e–ph interaction is given (from [34]).

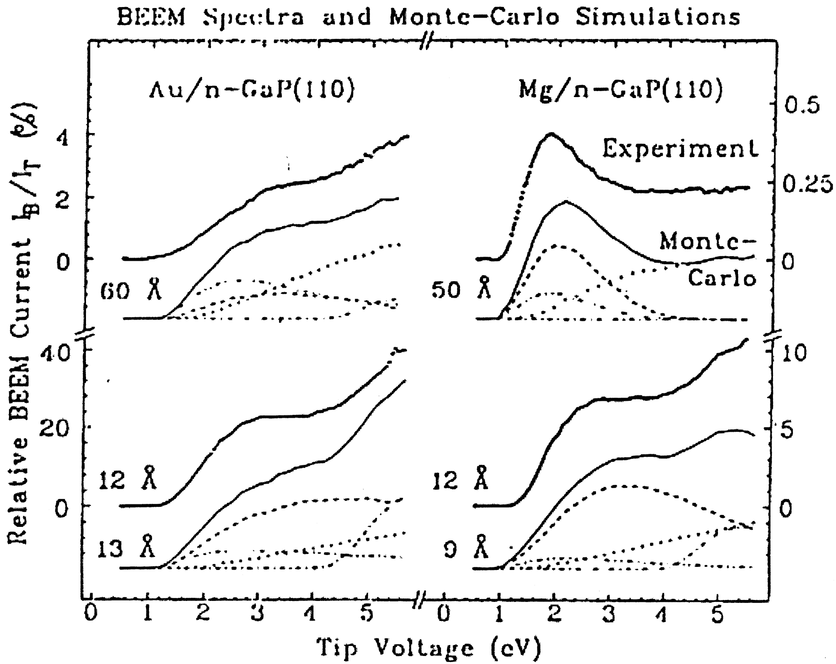


Fig. 23. I–V curves for Au and Mg films of varying thickness deposited on GaP(1 1 0) (thick solid lines). Using MC simulations, Bauer et al. identified different possible physical mechanisms contributing to intensity (thin solid lines). Notice onset of impact ionization (dash-dotted lines) at about 3.0 eV above threshold (from [18,19]).

This induces an increase of the BEEM current, which can even become larger than the initially injected tunneling current [31]. The signature of impact ionization is often a second onset at energies around the impact ionization threshold (taking slightly larger values than the semiconductor band gap). Different theoretical approaches have been tried to model the experimental data. The simplest, but still used, is due to Keldysh [80], who took a parabolic-band approximation for the semiconductor to derive the main energy dependence of the relaxation time

$$\frac{1}{\tau_I(E)} = \frac{P_0}{\tau_{\rm op}(E_I)}\left[\frac{E-E_I}{E_I}\right]^2, \tag{5.1}$$

where $\tau_{\rm op}^{-1}$ gives the scattering rate of electrons by optical phonons, and P_0 and the threshold E_I are treated as parameters to be fitted to the experiment. Bauer et al. [18,19] applied this expression to E-space MC calculations, which have permitted an unambiguous identification of the impact ionization at 3–4 eV above the SB threshold.

The GF formalism of Section 3.2 has also been used for the calculation of BEEM-current using (3.26), and the fitting of some transport parameter, such as the e–e MFP in metals. In particular, more quantitative information on inelastic processes, and their associated MFPs, can be obtained by combining BEEM experiment and theory. As mentioned before, these are technologically important processes, and BEEM seems most adequate for accurately characterizing them on real metalic thin films. Reuter et al. [65] analyzed the $I(V)$ curves for Au/Si(1 1 1), which were measured by Bell [17] as a function of the film width and temperature. Data measured for thin films and low temperatures, where the inelastic e–e interaction is likely to be dominant, can be understood in terms of the standard RPA approximation for a Fermi liquid. In particular, following the work of Luttinger [81], Quinn [66] and others, a simple expression can be obtained for the MFP related to e–e interaction in an electron gas

$$\lambda_{\rm e\text{–}e}(E) = \lambda_0^{\rm RPA}(r_{\rm s})\frac{(E/E_F)}{((E/E_F)-1)^2}, \quad \lambda_0^{\rm RPA}(r_{\rm s}) \approx 4(1+r_{\rm s}), \tag{5.2}$$

which is valid in the interval $1 \leqslant r_{\rm s} \leqslant 6$ ($\lambda_0^{\rm RPA}$ is in atomic units). In this simple equation, it is important to realize that further improvements to the theory are more likely to affect the prefactor, λ_0, which is strongly linked to the homogeneous electron–gas model, than the functional dependence on energy, derived from general requirements, such as the existence of a Fermi surface and the validity of the quasiparticle Landau model for the Fermi liquid. In fact, experimental measurements [82] and theoretical improvements over the electron-gas model [83] seem to indicate that $\lambda_0^{\rm RPA}$ should be corrected by a constant factor, $\lambda_0 = \alpha\lambda_0^{\rm RPA}$, to obtain good agreement between theory and experiment for $\lambda_{\rm e\text{–}e}$. A good theoretical model should be able to provide an estimation of α, which is not yet perfectly well known, but is generally accepted to be something like twice or thrice the RPA value [83]. Reuter et al. [65], therefore, considered α as a single free parameter, and found an

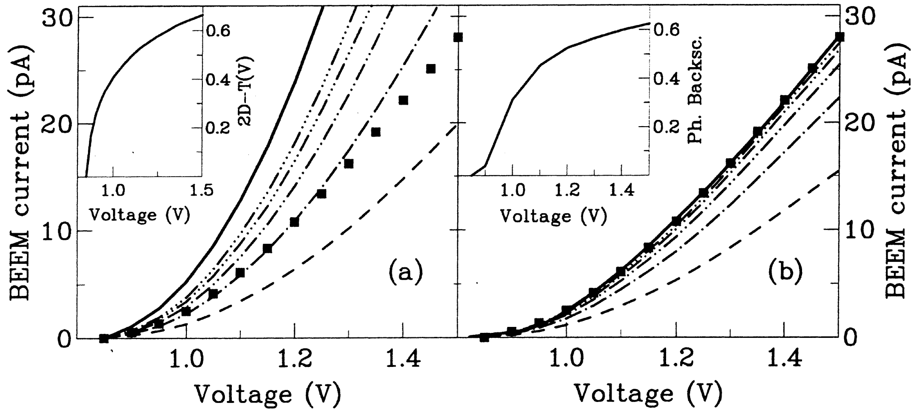


Fig. 24. $I(V)$-curves for 75 Å Au/Si(1 1 1) at 77 K. Experimental data (solid squares both in (a) and (b)) are taken from [17]. (a) Ballistic results ($\lambda_0 = \infty$) for direct injection (dashed), with one reflection at metal surface (dashed one-dotted), two reflections (dashed two-dotted), three reflections (dashed three-dotted), four reflections (dashed four-dotted), and to infinite order (solid thick line). Different reflections are calculated using reflection coefficient obtained by matching wave functions across (1 1 1) interface, $1 - T(E)$ (see inset in Fig. 1(a)), and probability for an electron to be back-injected into the metal, due to electron–phonon scattering in semiconductor, $1 - P(E)$ (see inset in Fig. 1(b)). (b) Full GF result ($\lambda_0 = 30.4$ Å) decomposed in multiple reflections as in (a). Infinite order result (solid thick line in (b)) is to be compared with experiment.

excellent agreement between theory and experiment for $\alpha \approx 2$ (Fig. 24(b)). The fit is of a striking quality, especially when considering the contrast between fitting only one free parameter and the interplay of multiple factors in the four steps, all influencing the total BEEM spectrum in a complicated way. This is typified, for example, by including the separate contributions from the lowest multiple reflections in the figure. The crucial importance of including the electrons multiple reflections in the final BEEM current in thin Au films was first proposed by Bell [17]. Fig. 24 clearly shows that a considerable increase in the BEEM current is due up to the third passage, thus explaining why some initial attempts to simulate such thin film data had to come up with much larger values of $\lambda_{e\text{–}e}$, when considering only direct injection. This illustrates how delicately the extraction of electron lifetimes from BEEM suffers from deficiencies in the underlying model. Consequently, it is important to check the weaknesses of any approach to ascertain the likely error bars. There are two hypotheses in GF theory that are still under discussion and cannot be unambiguously confirmed either by experiments or first-principle arguments: specular reflections and the degree of $\vec{k}_{\parallel}$-conservation during transmission. It must be admitted that disorder at both the rough surface and the non-epitactic interface can create diffuse reflections at the metal surface [17,84] and/or $\vec{k}_{\parallel}$-violation at the interface [35,67], but it is difficult to include arbitrary or selective $\vec{k}_{\parallel}$-pickups in a GF formalism. On the other hand, as already mentioned, the new distributions of currents make it much less important to introduce this kind of randomization at the interface. Therefore, to understand better its role on the theory, we have performed

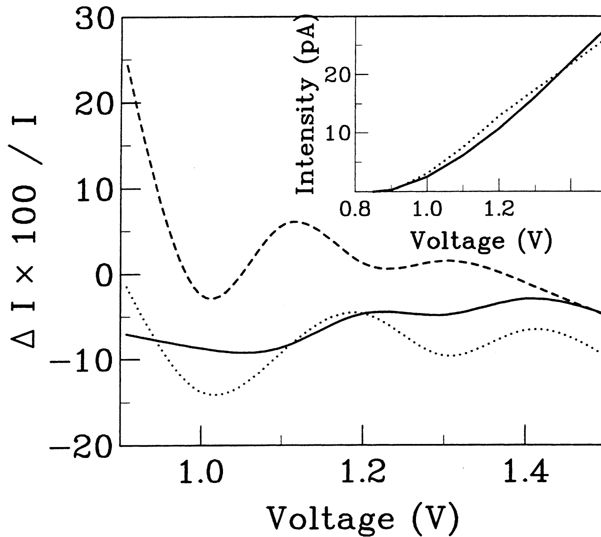


Fig. 25. Agreement between GF method (solid line) and MC simulation (dotted) is demonstrated in inset by two separate calculations on a 75 Å wide Au layer, where specular reflections are assumed at surface and interface. Same $\lambda_{e\text{-}e}(E)$ is used in both calculations, as defined in (5.2). In figure, relative change in intensity between MC specular model and three following models is given: diffuse reflections at surface (solid line), randomization of $k_{\parallel}$ at interface (dotted line), and elastic scattering with defects, $\lambda_{e\text{-}d} \approx 100$ Å (dashed line).

additionally semiclassical MC simulations, using the framework described in Section 4.2. The first benchmark is to test whether both schemes (GF formalism and k-space MC simulations) are compatible under equal conditions. This comparison is shown in Fig. 25, where the agreement between both theories is rather good, in spite of the approximations made in the MC simulation to mimic band structure effects. We now analyze the effect on the BEEM spectra of the different quasi-elastic scattering processes at the metal surface and/or the MS interface, which are not included in our estimation of $\lambda_{e\text{-}e}$. The solid line in Fig. 25 corresponds to the relative change in the final BEEM current, when diffuse reflections (instead of specular ones) are assumed in the metal surface. The dotted line represents the effect of non-$\vec{k}_{\parallel}$ conserving interface scatterings, which is treated by allowing the pickup of an arbitrary $\vec{k}_{\parallel}$ with a 40%-probability. From these case studies, it is clear that none of these quasi-elastic scatterings has a major influence on such thin film spectra, being typically less than ±10% on the BEEM current. It is worth stressing that this small change stems from the relatively extended ring-like current distribution in k-space, due to the band structure of Au. A shuffling of the electron ensemble at the film boundaries does not introduce a drastic change in the interface BZ pattern, so the spectra is not greatly changed. Hence, we can conclude that the results obtained with the k-space MC technique reinforce previous GF results, and allow to estimate a conservative worst-case error in the determined e–e mean free path of ±20%.

6. Conclusions

An overview of different theories has been presented, with a balanced emphasis on both their physical elements and their ability to explain experimental evidence. Significant experimental facts are found relevant for direct space (e.g., nanometric spatial resolution, even for relatively thick metal layers), for reciprocal space (e.g., broad $\vec{k}_{\parallel}$-distributions, suggested by results on different semiconductor orientations), and E-space (e.g., the spectroscopic energy-dependence of BEEM intensities, when the tip-sample voltage is varied).

It is concluded that most of the experimental results can be readily interpreted by using available theoretical tools. All the approaches have been broadly classified according to whether they allow carriers to exchange energy with other particles in the solid (Section 4, scattering processes) or not (Section 3, one-electron approaches).

For one-electron theories, we see that the mere free propagation of carriers is not a good approximation, except in the close vicinity of the SB. This is at variance with the earlier expectations of on the technique and, in particular, with its now generally accepted ballistic name. Taking into account the coherent propagation of electrons in the metal base naturally explains the experimental observed resolution and the similar currents measured for different orientations of the semiconductor substrate.

For theories allowing energy-exchange processes, the traditional E-space MC approach is slowly moving away from the earlier heavily parametrized simulations to a more fundamental k-space approach. This has the further advantage of making it possible to incorporate into MC calculations, in an approximate way, the band-structure obtained from one-electron theories by using a rejection technique.

GF techniques and the MC approach are shown to coincide, when only e–e scattering needs to be considered (i.e., at low temperatures and small film widths). Under this condition, an energy-dependent optical potential, calculated from Fermi liquid theory for an electron gas, can be introduced into the GF, reproducing well the measured attenuation and in perfect agreement with the MC result. If the model has to include concurrently several different types of interaction (e.g., e–e, e–ph and e–d), or other imperfections in the system (e.g., diffuse reflections at the surface or interface) semi-classical k-space MC still remains the more reasonable approach.

The final perspective is very promising, since BEEM has evolved from a method for determining local SBs to a technique that can be confidently used to determine accurately electron MFPs in metals, and local density of states associated with surface bands and impurities at the MS interface. All this development has been partially due to a deeper theoretical understanding of the physics underlying BEEM and, based on this trend, we can expect further developments in this area in the near future.

Acknowledgements

This work has been funded by the Spanish CICYT (PB97-0028 and PB98-0524). We acknowledge useful conversations with Dr. Ulrich Hohenester and Prof. Peter Kocevar on Monte-Carlo techniques. Pedro de Andres would like to thank the BBV

Foundation for financial support during a sabbatical in the University of Cambridge, and is grateful to Prof. David King and the Surface Science Group in the Department of Chemistry for their kind hospitality.

References

[1] G. Binnig, H. Rohrer, Ch. Gerber, E. Weibel, Phys. Rev. Lett. 49 (1982) 57.
[2] J. Tersoff, D.R. Hamman, Phys. Rev. Lett. 50 (1983) 25.
[3] D.M. Eigler, E.K. Schweizer, Nature 344 (1990) 524.
[4] D.M. Eigler, C.P. Lutz, W.E. Rudge, Nature 352 (1991) 600.
[5] W.J. Kaiser, L.D. Bell, Phys. Rev. Lett. 60 (1988) 1406.
[6] L.D. Bell, W.J. Kaiser, Phys. Rev. Lett. 61 (1988) 2368.
[7] M. Dähne-Prietsch, Phys. Rep. 253 (1995) 164.
[8] L.D. Bell, W.J. Kaiser, Ann. Rev. Mater. Sci. 26 (1996) 189.
[9] M.S. Tyagi, Introduction to Semiconductor Materials and Devices, Wiley, New York, 1991.
[10] R.H. Fowler, Phys. Rev. 38 (1931) 45.
[11] W.E. Spicer, I. Lindau, P. Skeath, C.Y. Su, J. Vac. Sci. Technol. 17 (1980) 1019.
[12] F. Flores, C. Tejedor, J. Phys. C 20 (1987) 145.
[13] A. Tanabe, K. Konuma, N. Teranishi, S. Tohyama, K. Masubuchi, J. Appl. Phys. 69 (1990) 850.
[14] G. Margaritondo, Prog. Surf. Sci. 56 (1997) 311.
[15] H. Palm, M. Arbes, M. Schulz, Phys. Rev. Lett. 71 (1993) 2224.
[16] H. Palm, M. Arbes, M. Schulz, Appl. Phys. A 56 (1993) 1.
[17] L.D. Bell, Phys. Rev. Lett. 77 (1996) 3893.
[18] A. Bauer, M.T. Cuberes, M. Prietsch, G. Kaindl, Phys. Rev. Lett. 71 (1993) 149.
[19] A. Bauer, M.T. Cuberes, M. Prietsch, G. Kaindl, J. Vac. Sci. Technol. B 11 (1993) 1584.
[20] M.T. Cuberes, A. Bauer, H.J. Wen, D. Vandré, M. Prietsch, G. Kaindl, J. Vac. Sci. Technol. B 12 (1994) 2422.
[21] A.M. Milliken, S.J. Manion, W.J. Kaiser, L.D. Bell, M.H. Hecht, Phys. Rev. B 46 (1992) 12826.
[22] H. Sirringhaus, E.Y. Lee, H. von Känel, Phys. Rev. Lett. 73 (1994) 577.
[23] T. Meyer, H. Sirringhaus, H. von Känel, Thin Solid Films 318 (1998) 195.
[24] F.J. Garcia-Vidal, P.L. de Andres, F. Flores, Phys. Rev. Lett. 76 (1996) 807.
[25] K. Reuter, F.J. García-Vidal, P.L. de Andres, F. Flores, K. Heinz, Phys. Rev. Lett. 81 (1998) 4963.
[26] H. Sirringhaus, E.Y. Lee, H. von Känel, Surf. Sci. 331-333 (1995) 1277.
[27] M.K. Weilmeier, W.H. Rippard, R.A. Buhrman, Phys. Rev. B 59 (1999) R2521.
[28] L.J. Schowalter, E.Y. Lee, Phys. Rev. B 43 (1991) 9308.
[29] E.Y. Lee, L.J. Schowalter, Phys. Rev. B 45 (1992) 6325.
[30] T. Kalka, M. Dähne-Prietsch, G. Kaindl, in: M. Scheffler, R. Zimmermann (Eds.), The Physics of Semiconductors, World Scientific, Singapore, 1996.
[31] R. Ludeke, Phys. Rev. Lett. 70 (1993) 214.
[32] M. Prietsch, R. Ludeke, Phys. Rev. Lett. 66 (1991) 2511.
[33] E.Y. Lee, L.J. Schowalter, J. Appl. Phys. 70 (1991) 2156.
[34] U. Hohenester, Private communication.
[35] D.L. Smith, E.Y. Lee, V. Narayanamurti, Phys. Rev. Lett. 80 (1998) 2433.
[36] J.C. Slater, G.F. Koster, Phys. Rev. 94 (1954) 1498.
[37] C. Caroli, R. Combescot, P. Nozieres, D. Saint-James, J. Phys. C 4 (1971) 916.
[38] L.V. Keldysh, Zh. Eskp. Teor. Phys. 47 (1964) 1515.
[39] L.V. Keldysh, Sov. Phys. JETP 20 (1965) 1018.
[40] E.M. Lifshitz, L.P. Pitaevskii, Physical Kinetics, Pergamon Press, Oxford, 1993.
[41] C. Caroli, R. Combescot, P. Nozieres, D. Saint-James, J. Phys. C 5 (1972) 21.
[42] A. Martin-Rodero, F. Flores, N.H. March, Phys. Rev. B 38 (1988) 10047.
[43] F. Flores, P.L. de Andres, F.J. Garcia-Vidal, L. Jurczyszyn, N. Mingo, R. Perez, Prog. Surf. Sci. 48 (1992) 27.

[44] F. Guinea, C. Tejedor, F. Flores, E. Louis, Phys. Rev. B 28 (1983) 4397.
[45] K. Reuter, P.L. de Andres, F.J. García-Vidal, D. Šestović, F. Flores, K. Heinz, Phys. Rev. B 58 (1998) 14036.
[46] K. Reuter, P.L. de Andres, F.J. García-Vidal, D. Šestović, F. Flores, K. Heinz, Comp. Phys. Comm. 127 (2000) 327.
[47] D.J. Chadi, M.L. Cohen, Phys. Rev. B 8 (1973) 5747.
[48] R. Ramirez, M. Böhm, Int. J. Quant. Chem. 19 (1988) 571.
[49] P.L. de Andres, K. Reuter, F.J. Garcia-Vidal, D. Sestovic, F. Flores, Appl. Surf. Sci. 123 (1998) 199 (cond-mat/9710091).
[50] P.L. de Andres, F.J. Garcia-Vidal, D. Sestovic, F. Flores, Phys. Scr. T 66 (1996) 277.
[51] K. Oura, T. Hanawa, Surf. Sci. 82 (1979) 202.
[52] A.K. Green, E. Bauer, J. Appl. Phys. 47 (1976) 1284.
[53] P.F. de Pablos, F.J. Garcia-Vidal, F. Flores, P.L. de Andres Surf. Sci. (in press).
[54] F. Garcia-Moliner, F. Flores, Introduction to the Theory of Solid Surfaces, Cambridge University Press, Cambridge, 1980.
[55] J.B. Pendry, Low-Energy Electron Diffraction, Academic Press, London, 1974.
[56] M.D. Stiles, D.R. Hamann, Phys. Rev. B 38 (1988) 2021.
[57] G. Wachutka, Phys. Rev. B 34 (1986) 8512.
[58] E. Vlieg, A.E.M.J. Fischer, J.F. van der Veen, B.N. Dev, G. Materlik, Surf. Sci. 178 (1986) 36.
[59] E.J. vanLoenen, J.W.M. Frenken, J.F. van der Veen, S. Valeri, Phys. Rev. Lett. 54 (1985) 827.
[60] M.D. Stiles, D.R. Hamann, Phys. Rev. B 40 (1989) R1349.
[61] M.D. Stiles, D.R. Hamann, Phys. Rev. Lett. 66 (1991) 3179.
[62] M.D. Stiles, D.R. Hamann, J. Vac. Sci. Tech. B 9 (1991) 2394.
[63] W.J. Kaiser, M.H. Hecht, R.W. Fathauer, L.D. Bell, Phys. Rev. B 44 (1991) 6546.
[64] H. Sirringhaus, T. Meyer, E.Y. Lee, H. von Känel, Phys. Rev. B 53 (1996) 15944.
[65] K. Reuter, U. Hohenester, P.L. de Andres, F.J. García-Vidal, F. Flores, K. Heinz, Phys. Rev. B 61 (2000) 4522.
[66] J.J. Quinn, Phys. Rev. 126 (1962) 1453.
[67] R. Ludeke, A. Bauer, Phys. Rev. Lett. 71 (1993) 1760.
[68] E. Fermi, Nuclear Physics, University of Chicago Press, Chicago, 1950, p. 187.
[69] J.M. Ziman, Electrons and Phonons, Oxford University Press, Oxford, 1960.
[70] C. Manke, Y. Bodschwinna, M. Schulz, Appl. Surf. Sci. 117 (1997) 321.
[71] U. Hohenester, P. Kocevar, P.L. de Andres, F. Flores, Phys. Stat. Sol. B 204 (1997) 397 (cond-mat/9710151).
[72] U. Hohenester, P. Kocevar, P.L. de Andres, F. Flores, in: E.T. Cullis (Ed.), Proceedings of the 10th Conference on Microscopy of Sem. Mat., vol. 157, Oxford, 1997, p. 615 (cond-mat/9710147).
[73] C. Jacoboni, L. Reggiani, Rev. Mod. Phys. 55 (1983) 645.
[74] C. Jacoboni, P. Lugli, The Monte-Carlo Method for Semiconductor Device Simulation, Springer, Berlin, 1989.
[75] D. Pines, Elementary Excitations in Solids, Benjamin, New York, 1964.
[76] F.J. Blatt, Physics of Electronic Conduction in Solids, McGraw-Hill, New York, 1968.
[77] F.J. Blatt, Solid State Phys. 4 (1957) 199.
[78] G.F. Koster, Phys. Rev. 95 (1954) 1436.
[79] W.H. Zurek, Phys. Today 44 (1991) 36.
[80] L.V. Keldysh, Sov. Phys. JETP 10 (1960) 509.
[81] J.M. Luttinger, Phys. Rev. 121 (1961) 942.
[82] H. Petek, S. Ogawa, Prog. Surf. Sci. 56 (1997) 239.
[83] I. Campillo, J.M. Pitarke, A. Rubio, E. Zarate, P.M. Echenique, Phys. Rev. Lett. (1999) 2230.
[84] R. Menegozzi, P.G. Reinhard, M. Schulz, Surf. Sci. 411 (1998) L810.