Field emission theory for an enhanced surface potential: a model for carbon field emitters

T C Choy, A H Harker and A M Stoneham

Department of Physics and Astronomy, University College London, Gower Street,
London WC1E 6BT, UK

E-mail: a.harker@ucl.ac.uk

Abstract
We propose a non-JWKB-based theory of electron field emission for carbon
field emitters in which, for electrons with energy in the vicinity of the order of ϑ
to the Fermi level, the effective $(1/x)$ surface potential is strongly enhanced. The
model grossly violates the WKB validity criteria and necessitates an analytic
treatment of the one-dimensional Schrödinger equation, which we first obtain.
We determine ϑ (which is field-dependent) from the wavefunction matching
point close to the surface. For reasonable values of the surface parameters—
work function $\varphi \approx$ 2–5 eV, electron affinity $\chi \approx 2\varphi$ and an empirical
electron loss factor $\sigma \approx 10^{-3}$ (and with no other adjustable parameters)—the
theory provides an intriguing agreement with experimental data from carbon
epoxy graphite composite (PFE) and certain graphitized carbon nanotube field
emitters. We speculate on the surface potential enhancement, which can be
interpreted as a massive (field-induced) dielectric effect of dynamic origin.
This can be related via time-dependent perturbation theory to second-order non-
linear polarizability enhancements at ultraviolet $\sim$3000 Å wavelengths near the
tunnelling region. Finally some exact mathematical results are included in the
appendix for future reference.

1. Introduction

The classic theories of electron field emission from a cold metal surface due to Fowler and
Nordheim (FN) [1] and later with image modifications due to Nordheim [2] were one of
the earliest applications of wave mechanics after its foundation in the early 1920s. FN
were originally motivated by their conviction of Sommerfeld's (Fermi–Dirac-based) theory of
electrons in metals and thus their need to reject early interpretations of experiments by others
such as Millikan and Lauritsen [3], who suggested that the emitted electrons, 'thermions', are
to be distinguished from the conduction electrons in the metal[1]. The Nordheim theory [2]

---

[1] In the modern context, especially in non-metals for which tunnelling almost certainly comes from surface states
which are distinguished from the bulk, such ideas should probably not be scorned.

actually consists of two parts: a full image potential, zero field ($F = 0$) but analytically exact solution of the one-dimensional Schrödinger equation (based on modern day Whittaker functions) and (for reasons of intractability at that time) an approximate image potential + finite field ($F \neq 0$) theory using the Jeffreys' method [4]. The latter is the now familiar modern-day WKB approximation (henceforth we shall call this the JWKB approximation) [5–7]. This approximate theory is also widely known as the Fowler–Nordheim (FN) theory by many later workers, which is indeed a source of confusion. The FN–JWKB theory, with later corrections for Nordheim's erroneous reduction of elliptic integrals [8, 9], has become the landmark theory for field emission studies from metal surfaces for the past 75 years, while the original exact full image–zero field Nordheim theory has somehow disappeared into obscurity [10, 11].

With increased interest in field emission processes, especially for use in printable field emitting (PFE) display screens [12], and the difficulty that the FN–JWKB theory is now challenged in order-of-magnitude terms by the high current densities (of up to 0.1 A cm^{-2}) observed in such systems, it is perhaps worthwhile re-examining some of its basic assumptions. These are:

(1) Electrons are fermions and are governed by Fermi–Dirac statistics.
(2) They are thus characterized by well-defined quantities in the bulk, namely the Fermi energy (E_F at zero temperature $T = 0$ K) or chemical potential (μ at finite temperatures $T \neq 0$ K), a work function φ and an electron affinity χ.[2]
(3) The *effective* surface potential, which is *assumed* to have a classical form outside the surface:

$$V(x) = \begin{cases} -\mu & \text{for } x < 0 \text{—region 1 (inside);} \\ \varphi - eFx - \dfrac{e^2}{16\pi\varepsilon_0 x} & \text{for } x > 0 \text{—region 2 (outside).} \end{cases} \quad (1)$$

Henceforth we shall take the zero of energy at the Fermi level, all energy being positive above the Fermi level and negative below, with $\mu = (\chi - \varphi)$. With minor exceptions, such as the electric field F being in V μm^{-1} and the current density J being in A cm^{-2}, we shall use SI units and hence $\varepsilon_0 = 8.854\,18 \times 10^{-12}$ F m^{-1} is the permittivity of free space and $e = 1.602\,18 \times 10^{-19}$ C is the elementary charge.
(4) The JWKB approximation for the evaluation of the tunnelling probability $\mathcal{T}(W)$, including an integration over *all* energy levels W for the emission current with this probability.

In this paper we shall not critically examine each of these assumptions which would also entail an examination of the criteria, regimes and conditions when each of them may fail. Obviously, there are also regimes and conditions in which the theory has been rather successful [9–11]. Our purpose is merely to innovate on reasonable grounds in a modern context, perhaps incorporating necessary modifications, such as abandoning the JWKB approximation as we shall see. Hence we shall state from the start that we shall not argue against (1) and indeed retain a Fermi–Dirac-based supply function, see equations (6) and (43) below. Our theory is based on a proposed modification to assumption (3), in which the $1/x$ potential is significantly enhanced. The latter leads to a gross violation of the WKB validity criteria (to be discussed later) and hence we are forced to reject assumption (4). Thus we are compelled to consider an analytic solution of the one-dimensional Schrödinger equation close to the surface. As well as exotic metal–vacuum interfaces, this theory may have further applications for graphite–vacuum, graphite–insulator–vacuum, certain graphitized carbon-nanotube–vacuum [13] and amorphous diamond-like carbon (DLC) Schottky junctions [14].

---

[2] The temperature dependence and band-bending modifications of these parameters are more realistic considerations which we will not pursue here.

Since our approach can be generalized to other cases of interfaces between dissimilar media, it may motivate a revisit of the systems discussed earlier by Mott and Gurney [15] and Herring [16]. These other cases might include vacuum breakdown and electrical forming processes [17]. In section 2, we shall discuss the basic features of our model, including a rejection of the JWKB approximation in our case in view of the strong enhancement. In section 3 we establish the analytic treatment of the one-dimensional Schrödinger equation including its generalization to include the full image plus finite ($F \neq 0$) field problem, the mathematical details of which we then postpone to the appendix B. While no exact closed form analytical treatment is yet available, unlike the closed form Whittaker function analysis historically first obtained by Nordheim [2] and later studied by MacColl [18], nevertheless we have decided to include these results for future reference. Our solution is exact in the sense of an infinite series in terms of Airy-type functions, even though for an enhanced potential (i.e. large $\bar{\beta}$, see later) the series is divergent. Nevertheless modern mathematical methods developed over the last two decades of re-summation, such as Borel–Laplace transform asymptotic theory [19], means that further progress towards a closed form solution is likely to be on the horizon. Moreover we have found, as in our case, that in the limit of an infinitely large surface potential, an exact solution is available in terms of Hankel functions of order one, which we shall be using extensively. Hence re-summation of the exact finite field solution can be guided by this limiting result, which could also be extended by perturbation analysis. Unfortunately we have achieved no success in terms of a series expansion for the exact finite field problem in terms of Whittaker functions, which is probably why this problem has not received much attention for the last 75 years [10, 11].

In section 5 we are particularly interested in applying our theory to epoxy-graphite PFE and graphitized carbon nanotube systems. We are intrigued not only by its removal of many orders-of-magnitude discrepancies between traditional FN–JWKB-based theories and experiment but the better agreement at the lower onset fields in a traditional $\ln J$ versus $1/F$ plot. This was achieved with few adjustable parameters (other than the work function φ, electron affinity χ and a factor σ to account for empirical electron loss) for which the theory does not seem to be very sensitive to, within reasonably acceptable values from experiments. This is to be contrasted with the need for assumptions about large field enhancement factors $\Gamma \sim 50$–350 within current FN–JWKB-based theories. However, we do not dispute that field enhancement factors have an important role in nano-structured systems. Indeed, for some (metallic based) systems such as MoS_2 nano-flowers, we have found quite good agreement between a classical field-enhanced FN–JWKB theory and experiments [20], but this does not appear to be the case for epoxy based carbon field emitters [12, 13, 21]. This suggests that the role of field enhancement in such systems is more subtle and less straightforward, perhaps embedded via dynamical polarization effects in a time-dependent tunnelling mechanism. In section 6 we speculate on this new physics which our results seem to indicate. Here we will suggest an interpretation of a massive (field-induced) dielectric effect at electromagnetic frequencies, of the order of the inverse transit time across the surface, for electrons near the Fermi level. Some open questions remain and suggestions for both further extensions of the theory and, in particular, experiments to verify our observations will also be made. In the conclusion we summarize our main results.

2. Preliminaries

As in the FN–JWKB theory, the role of the Fermi level, Fermi–Dirac statistics and, in particular, the unique wave–particle duality of quantum mechanics are key concepts in our model. The essentially equilibrium distribution of electrons is modelled by a one-dimensional (x-dependent

only) Hamiltonian:

$$H = \frac{p_x^2}{2m} + V(x), \tag{2}$$

whose eigenvalue W is the x part of the energy measured with reference to the Fermi level [9]. At this stage some preliminary remarks about the potential $V(x)$ (see equation (1)) are in order. Our definition of the surface potential follows that of Good and Müller [9] and differs from that of Guth and Mullin [22], who followed Nordheim [2] by connecting the outside classical image potential to the bulk at some value x_0 such that

$$\chi = \mu + \varphi = \frac{e^2}{16\pi\varepsilon_0 x_0}. \tag{3}$$

We do not propose to do this, on the grounds that x_0 should be limited to no less than a Bohr radius $a_B \sim 0.5$ Å. This would then artificially restrict the theory to systems with $\chi \leqslant 6.8$ eV, an inconsistency dating back to the original paper of Nordheim, who subsequently quoted values of χ (which is C in his notation) of between 10 and 20 eV [2]. For example, for tungsten, as noted by Good and Müller [9], $\chi \approx 10.3$ eV and they must have found Nordheim's proposal troublesome. Nordheim's, and later Guth and Mullin's, inconsistency signals to us that on the ångström scale the potential given by the classical expression equation (1) is an over-simplification, which modern techniques such as LDA/GGA have attempted to address [23]. However, we do propose later on that there is an $x_0 \sim 1$ Å as the distance of closest approach to the surface for the matching of the wavefunctions. There is no inconsistency here, though, since length scales less than $x_0 \sim 1$ Å are outside the regime of physics considered in our model. We propose that the behaviour of the potential near x_0 is highly significant, which is in contrast to Good and Müller [9] whose 'potential is not expected to have a significance at the point $x = 0$'. This was earlier also noted by Guth and Mullin [22], but they did not pursue the matter further.

Next the current density J of the emitted electrons is usually given by an integral over the energy W:

$$J = e \int_{-\infty}^{\infty} P(W)\, dW, \tag{4}$$

where $P(W)\, dW$ gives the number of electrons within dW that emerge from the surface per second per unit area. However, it is a fundamental assumption in our model that we shall, in fact, replace equation (4) by

$$J = e \int_{-\vartheta}^{0} P(W)\, dW, \tag{5}$$

as the dominant term, since all other contributions are many orders of magnitude less. As we shall see later, this includes the finite-temperature $T \neq 0$ K corrections to our theory, which are down by at least five orders of magnitude at room temperature, for the field strengths we consider. The quantity $P(W)$ is usually expressed as the product of a quantum mechanical tunnelling transmission coefficient $\mathcal{T}(W)$ and a Fermi-type supply function $N(W)$ so that

$$N(W) = \frac{4\pi m k_B T}{h^3} \ln\left\{1 + \exp\left[-\left(\frac{W}{k_B T}\right)\right]\right\}, \tag{6}$$

in which k_B is the Boltzmann constant. We have no doubt that a substantial modification of the JWKB tunnelling probability:

$$\mathcal{T}_{\text{JWKB}}(W) = \exp\left[-\int_{x_1}^{x_2} \left(\frac{8m}{\hbar^2}(V(x) - W)\right)^{1/2} dx\right], \tag{7}$$

in which x_1 and x_2 are the classical turning points of the potential $V(x)$, is a prerequisite for any viable theory for the carbon emitters [21]. There are at least two reasons for this.

The first reason is that a reduction of the tunnelling barrier height is inevitably necessary for any significant emission current. In the classical theory this is the Schottky image reduction given by

$$\Delta\phi = e\sqrt{\frac{eF}{4\pi\varepsilon_0}} = 3.7947 \times 10^{-2}\sqrt{eF} \text{ eV}^{1/2} \ \mu\text{m}^{1/2}, \tag{8}$$

which for a work function φ would require fields of the order of

$$F_b = 694\varphi^2 \text{ eV}^{-2} \text{ V } \mu\text{m}^{-1}. \tag{9}$$

In our theory we shall mimic this barrier reduction in a different way, see section 6 later. For now, let us examine the classic FN–JWKB formula [9] for the emission current density:

$$
\begin{aligned}
J_{\text{FN}} &= \frac{\sigma a F^2}{t(y_F)^2\varphi}\exp[-b\upsilon(y_F)\varphi^{3/2}/F], \\
a &= 1.5414 \times 10^2 \text{ A cm}^{-2} \text{ eV } \mu\text{m}^2 \text{ V}^{-2}, \\
b &= 6.8309 \times 10^3 \text{ eV}^{-3/2} \text{ V } \mu\text{m}^{-1}, \\
y_F &= 3.7947 \times 10^{-2} F^{1/2}/\varphi \text{ eV } \mu\text{m}^{1/2} \text{ V}^{-1/2};
\end{aligned}
\tag{10}
$$

in which $t(y_F)$ and $\upsilon(y_F)$ are given in terms of elliptic integrals [8, 9] and all fields F are measured in V μm^{-1} in this paper. We have included a factor σ to account for empirical loss factors, such as those due to cathode–anode geometrical factors, back-contact resistance and imperfect supply, since for composite and granular systems we expect a larger fraction of electron loss than for a clean metal surface in general. Equation (10) shows that for a typical work function $\varphi \sim 2$ eV, and applied fields in the range 1–16 V μm^{-1}, a field enhancement factor $F \to \Gamma F$ with Γ of up to 50 is necessary to produce currents of the order found for carbon emitters [21] and up to $\Gamma \sim 350$ for carbon nanotubes [13], see later. More importantly, in view of the weak dependence of the elliptic-function-based $\upsilon(y_F)$, the behaviour of $\ln J$ remains essentially that of $1/F$ which is not in accord with data for carbon field emitters [13, 21]. However, we must emphasize that there are systems such as MoS_2 nano-flowers [20] that are in accord with the classical field-enhanced FN–JWKB description. Our concern in this paper is with systems such as epoxy graphite and certain graphitized nanotube carbon field emitters, which appear to be in disagreement with FN–JWKB theory (see later).

The second reason is the WKB criteria. It is a well known result that the validity of the WKB approximation involves a criterion, that can be derived from the requirement that the variation of the deBroglie wavelength of the electron must be less than the dimension of the region of tunnelling [24]. This restriction on the validity of the FN–JWKB theory is well known to workers in the field, for example in the treatment of thermionic emission by Guth and Mullin [22] who attempted to improve the FN–JWKB theory by an exact analytic treatment of a parabolic approximation to the top of the barrier. Our concern is, however, with regions close to the surface of the order of ångströms to which tunnelling presumably occurs. Then the WKB validity criteria can be written as [24]

$$m\hbar\left|\frac{\partial V}{\partial x}\right| \ll p_F^3 = \hbar^3 k_F^3, \tag{11}$$

which, for fields $1 < F < 20$ V μm^{-1}, is dominated close to the surface by the image term, so that the above can be rewritten as

$$\frac{e^2}{16\pi\varepsilon_0 x_0} \ll \frac{x_0}{\lambda_F}\frac{1890.19}{\lambda_F^2} \text{ eV Å}^2. \tag{12}$$

For $x_0 \sim 1$ Å and $\lambda_F \sim 7$ Å the ratio ζ (of the LHS to the RHS) for the inequality in equation (12) is about 0.6, as also noted previously by Guth and Mullin [22] for the case of tungsten but using a different analysis. The situation is worse in our case when we shall later propose a strongly enhanced surface potential by at least two orders of magnitude. Unfortunately we are not aware of any detailed analysis for the degradation in the JWKB approximation as a function of the ratio ζ, nor of any systematic means for improving it when $\zeta \gg 1$, although some recent studies are beginning to emerge in the mathematical literature, due to advances in resurgent analysis [19].

The reader should note that our previous remarks are not meant to be a criticism of the JWKB approximation nor the FN–JWKB theory in general (see also remarks in the third paragraph of the last section). We merely need to stress that in the region close to the surface, and especially when the surface potential is strong, then there are reasons not to invoke a simplistic JWKB approximation. In fact, as pointed out by Guth and Mullin [22], a more sophisticated treatment of the FN–JWKB is possible. In their model they made use of connection formulae derived via an exact analytic solution of the parabolically approximated barrier top potential (which unfortunately involves rather complicated parabolic cylinder functions). Using the exact ($F = 0$) Whittaker function solution of Nordheim [2] and later MacColl [18] as a test case, they claim to have found transmission coefficients that are in agreement numerically. We do not know to what extent this agreement still holds for *large* fields and in particular what other modifications are necessary for strong image potentials ($\zeta \gg 1$). There is also a puzzling point in that the ($F = 0$) case (being of $1/r$ only) does not really have a parabolic barrier top and the agreement could therefore perhaps be fortuitous. This subject is no doubt worth further studies in view of our exact analytic solution in the form given in appendix B.

3. Imageless triangular barrier

We shall review the image-free tunnelling problem given by the following Schrödinger equation which is important for orientation:

$$\psi''(x) + \frac{2m}{\hbar^2}[W - V_0(x)]\psi(x) = 0, \tag{13}$$

where again W is the energy eigenvalue of the Hamiltonian H but now with the potential $V_0(x)$, which does not contain the image term:

$$V_0(x) = \begin{cases} -\mu & \text{for } x < 0\text{—region 1}; \\ \varphi - eFx & \text{for } x > 0\text{—region 2}. \end{cases} \tag{14}$$

This is a well known problem to theoretical workers in the field dating back to Fowler and Nordheim (FN) [1] and also recent review articles [10, 11], including more recently Jensen *et al*, who used these exact results as a basis for analytical approximations for the image rounded barrier problem, which they claim to have been validated by numerical studies [25].

For later convenience we shall recast equation (13) into the more compact form

$$\psi''(x) + (\epsilon + \alpha x)\psi(x) = 0, \tag{15}$$

with

$$\alpha = \frac{2meF}{\hbar^2} = c_\alpha eF,$$

$$\epsilon = \frac{2m(W - \varphi)}{\hbar^2} = c_\alpha(W - \varphi); \tag{16}$$

$$c_\alpha = 2.6246 \times 10^7 \text{ eV}^{-1} \text{ }\mu\text{m}^{-2}$$

in which F is measured in $V \, \mu m^{-1}$ as usual. To the best of the authors' knowledge this graduate-level quantum exercise does not seem to have found its way into many textbooks, although variants of it have been scattered around [27–30]: Modinos [26] simply cites Gadzuk and Plummer [10]. Following Landau and Lifshitz [31], we shall transform to the new variable [26] $\xi = (x + \epsilon/\alpha)\alpha^{1/3}$. Then equation (15) in region 2 becomes the typical Airy equation[3]:

$$\psi''(\xi) + \xi \psi(\xi) = 0. \tag{17}$$

However, some care needs to be exercised in selecting the appropriate type of Airy function, since the standard Ai function has asymptotic form [31, 32]:

$$Ai(-\xi) \approx \xi^{-1/4} \sin(\tfrac{2}{3}\xi^{3/2} + \pi/4). \tag{18}$$

As this also contains a reflected wave, it is not appropriate to our boundary conditions. The correct approach is to revert to the original 1/3 fractional-order Bessel function definition for the Airy functions and then construct appropriate linear combinations to yield the forward travelling wave solution only. This approach, of course, leads to an appropriate modified Hankel function which, as we shall see, has the correct asymptotic form required. Other combinations of Bessel functions are possible: for these the reader may refer to the literature and texts [10, 11, 27, 29, 30]. With this proviso, we can write down the solution for the wavefunctions of equation (17) as

$$\psi(x) = \begin{cases} e^{ikx} + re^{-ikx} & \text{in region 1;} \\ t[u_1(\xi) - e^{-i\pi/3}u_2(\xi)] & \text{in region 2,} \end{cases} \tag{19}$$

where $k = (2m(W + \mu)/\hbar^2)^{1/2}$ and the u_1 and u_2 are given by Bessel functions of order 1/3, here defined as

$$u_1(\xi) = \tfrac{1}{3}(\pi\xi)^{1/2} J_{-1/3}(\tfrac{2}{3}\xi^{3/2}), \tag{20}$$

and

$$u_2(\xi) = \tfrac{1}{3}(\pi\xi)^{1/2} J_{1/3}(\tfrac{2}{3}\xi^{3/2}), \tag{21}$$

respectively. Note, however, that analytical continuation to negative ξ requires some care: see appendix A. It is then straightforward to obtain the reflection and transmission amplitudes by matching the wavefunctions and their derivatives at the interface $x = 0$. We provide the results here in terms of $H_{1/3} = u_1 - e^{-i\pi/3}u_2$, which is a slightly modified Hankel function:

$$t = \frac{2}{H_{1/3}(\xi_0) - i\lambda_0^{1/3} H'_{1/3}(\xi_0)}$$

$$r = \frac{H_{1/3}(\xi_0) + i\lambda_0^{1/3} H'_{1/3}(\xi_0)}{H_{1/3}(\xi_0) - i\lambda_0^{1/3} H'_{1/3}(\xi_0)}, \tag{22}$$

where $\xi_0 = \xi(x = 0) = \bar{\epsilon} = \epsilon/\alpha^{2/3}$ here, while $\lambda_0^{1/3} = \alpha^{1/3}/k$. An important check must be made to these formulae which requires that the transmission coefficient:

$$\mathcal{T} = \frac{\lambda_0^{1/3}}{4}|t|^2, \tag{23}$$

and reflection coefficient, $\mathcal{R} = |r|^2$, satisfy the unitarity property $\mathcal{T} + \mathcal{R} = 1$. This is easily checked to be the case when appropriate use is made of the Wronskian identity:

$$u_1 u'_2 - u_2 u'_1 = \frac{1}{2\sqrt{3}}. \tag{24}$$

[3] The advantage of using the Airy function formulation is that it naturally takes care of the long range boundary condition posed by the electric field.

For sufficiently small fields, we can obtain an asymptotic expansion (assuming $\xi_0 \to \infty$ *but* $\lambda_0 \to 0$) that yields

$$\mathcal{T}_{\mathrm{asymp}}(W) \approx \frac{4\lambda_0^{1/3}|\xi_0|^{1/2}}{(1 + \lambda_0^{2/3}|\xi_0|)}\mathrm{e}^{-\frac{4}{3}|\xi_0|^{3/2}}$$

$$= \frac{4[|(W - \varphi)|(W + \mu)]^{1/2}}{\mu + \varphi}\mathrm{e}^{-\frac{4}{3}|\xi_0|^{3/2}}, \tag{25}$$

the pre-factor being a result not obtainable within ordinary WKB. The reader may like to note that this result is identical with equation (18) of FN [1], who also gave a more accurate asymptotic formula in their earlier equation (17) FN [1], and more recently a higher-order asymptotic expansion of this pre-factor has been given by Jensen [11]. For the sake of illustration, let us set $\mu = \varphi$ for the moment. Then we see that the above pre-factor for $W = 0$, i.e. $\mathcal{T}_{\mathrm{asym}}(0)$, differs from the WKB result by a mere factor of two, i.e.

$$\mathcal{T}_{\mathrm{asymp}}(0) \approx 2\mathrm{e}^{-\frac{4}{3}\xi_0^{3/2}}. \tag{26}$$

That this asymptotic result *could* be problematic as we approach the top of the barrier can be seen from the fact that if we naively allow the fields to become arbitrarily large and adopt equation (26), then we would arrive at an unphysical result in which $\mathcal{T} > 1$, which is absurd. No doubt this can only occur for very large fields $F \sim 10^4 \varphi^{3/2}$ V $\mu\mathrm{m}^{-1}$ in general for the imageless case, but if a significant barrier reduction mechanism operates (which effectively reduces φ even to millielectronvolts), then it would be necessary to use the original formula equations (22) and (23). These observations are not new and are well known to workers in semiconductor devices [27, 33] but are worth reiterating in any case.

Once again we emphasize that our purpose is not to criticize the WKB approach in general, but to merely note that its prediction of large transmission coefficients at high applied fields, without knowledge of the pre-factor, can be problematic. Unfortunately this caution does not seem to have been emphasized by earlier workers such as Good and Müller [9] nor by Murphy and Good Jr [34]. The latter have derived widely used criteria [35] for the validity of FN–JWKB theory on the basis of the WKB integrand properties equation (7) alone, which may be inadequate. Indeed any sophisticated beyond-WKB procedure, whose purpose it is to recover the *correct* pre-factor, as in equation (26), may not be an altogether worthwhile exercise in light of the above remarks. The reader should note in particular the well known result from elementary quantum mechanics that the transmission coefficient is, in general, not unity at the top of the square barrier [36].

4. The full image rounded triangular barrier

We now turn to the quantum tunnelling problem including the effect of the image term as in the original Nordheim [2, 18] theory. As far as the authors are aware, there have not been many serious attempts in the past at seeking an exact analytic solution of the *full* one-dimensional Schrödinger equation. Our purpose here is twofold. Firstly, we shall document our attempts and comment on the successes and failures. This work has led us to interesting territory and also educated us in some important advances in modern mathematical analysis, not widely known to standard mathematical physics [19, 37] and differential equations [38, 39]. Secondly, and more importantly, our analysis here has motivated us to innovate on a choice of surface potential that will be applied to carbon emitters in the next section.

The full image + field Schrödinger equation (15) now becomes

$$\psi''(x) + \left[\epsilon + \alpha x + \frac{\beta}{x}\right]\psi(x) = 0, \tag{27}$$

where the image potential is characterized by the parameter $\beta = me^2/(8\pi\varepsilon_0\hbar^2)$, but in view of our lack of knowledge of surface potentials, we might perhaps consider β as a free parameter, see later. Once again, transforming to the parameter ξ as before, we now have the equation

$$\psi''(\xi) + \left[\xi + \frac{\bar{\beta}}{\xi - \bar{\epsilon}}\right]\psi(\xi) = 0, \tag{28}$$

where $\bar{\beta} = \beta\alpha^{-1/3}$ and $\bar{\epsilon} = \epsilon\alpha^{-2/3}$. Clearly the image-free case $\bar{\beta} = 0$ reduces to the usual Airy differential equation (17). Equation (28) can be rewritten as

$$(\xi - \bar{\epsilon})\psi''(\xi) + [\xi(\xi - \bar{\epsilon}) + \bar{\beta}]\psi(\xi) = 0. \tag{29}$$

We shall in fact further rewrite this as

$$\psi''(\xi) + \left[\frac{(\xi - \lambda_1)(\xi - \lambda_2)}{(\xi - \bar{\epsilon})}\right]\psi(\xi) = 0, \tag{30}$$

in which the two roots of the quadratic are given by

$$\lambda_1 = \tfrac{1}{2}\left[\bar{\epsilon} + \sqrt{\bar{\epsilon}^2 - 4\bar{\beta}}\right] \qquad \text{and} \qquad \lambda_2 = \tfrac{1}{2}\left[\bar{\epsilon} - \sqrt{\bar{\epsilon}^2 - 4\bar{\beta}}\right]. \tag{31}$$

Let us first discuss the WKB approximation [4–7, 24], which is obtained by approximating the wavefunction as

$$\psi(\xi) = \frac{1}{\sqrt{p(\xi)}}\exp\left[i\int^{\xi} p(\xi')\,d\xi' + \frac{i\pi}{4}\right], \tag{32}$$

where

$$p(\xi) = \left[\frac{(\xi - \lambda_1)(\xi - \lambda_2)}{(\xi - \bar{\epsilon})}\right]^{1/2}; \tag{33}$$

and the lower limit in the integral in equation (32) is the smaller of λ_1 or λ_2, which depends on the sign of $\bar{\epsilon}$, i.e. above or below the barrier, respectively. The evaluation of the integral in equation (32) is then an exercise in elliptic integral reduction which we need not repeat here [2, 9, 34], which fortunately nowadays is greatly aided by tools such as Mathematica v4.1. Equation (30) itself does not appear to have a closed form solution in terms of known functions, although several series methods of solution are available, depending on the parameters; see appendix B. We may as well write down the exact closed form solution for the case ($F = 0$) in terms of the appropriate Whittaker function, due to Nordheim [2], as a start. This can be written in terms of the scaled parameters $z = 2\bar{\epsilon}^{1/2}y$, where $y = \xi - \bar{\epsilon}$ (which will be important to us later):

$$\psi(z) = \mathcal{W}_{i\gamma, \frac{1}{2}}(iz), \qquad \gamma = \frac{\bar{\beta}}{2\bar{\epsilon}^{1/2}}; \tag{34}$$

where the Whittaker functions $\mathcal{W}_{\lambda,\mu}(z)$ are known to satisfy the differential equation [40]

$$\frac{d^2\mathcal{W}}{d^2z} + \left(-\frac{1}{4} + \frac{\gamma}{z} + \frac{\frac{1}{4} - \mu^2}{z^2}\right)\mathcal{W} = 0. \tag{35}$$

In fact, in terms of the variable z the differential equation we seek (equation (27) or (28)) is less complicated looking than equation (35), being

$$\psi''(z) + \left(\frac{1}{4} + \frac{\gamma}{z} + \bar{\lambda}z\right)\psi(z) = 0, \qquad \bar{\lambda} = \frac{1}{8\bar{\epsilon}^{3/2}}; \tag{36}$$

and it is a puzzle to us that it has received so little attention. However, our attempts at a series solution based on the Whittaker functions have not met with success, in view of the incompatibility of the recursion and orthogonality properties of these functions with the
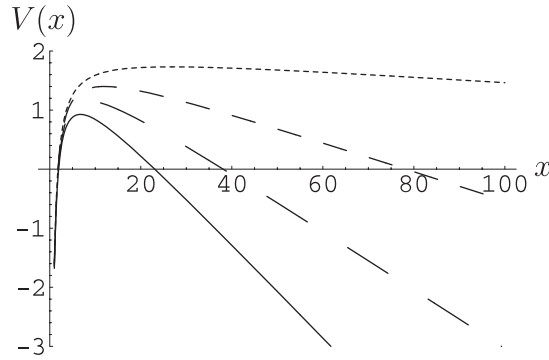
Figure 1. Plot of the classical surface image potential for a field enhancement factor of $\Gamma = 50$. Curves are drawn in decreasing order for $F = 16$ (full), 10 (long dash), 5 (short dash) and 1 (dotted), respectively, in units of V μm^{-1}. $V(x)$ is in electronvolts, with the work function $\varphi = 2.0$ eV and the distance x from the surface is in ångströms. Note the Schottky barrier is reduced by about 53.6% for the largest field.

linear field term in equation (36). This state of affairs is puzzling which is in contrast with the analogous problem of the hydrogen atom in an electric field, already noted by the early pioneers (see references in FN [1]). The latter can be treated exactly by a rapidly convergent infinite series expansion in terms of Laguerre/confluent hypergeometric functions as in the Stark effect/van der Waals molecule problem, see, for example, [41]. We do not know if this merely reflects our ignorance of the properties of Whittaker functions or otherwise, for in the appendix B we shall show that other series methods of solution exist. We will therefore not discuss these results further here except to say that they are, in general, series expansions in the parameter $\bar{\beta}/\bar{\epsilon}$. Unfortunately for large fields or enhanced $\bar{\beta}$, which are the main interest to us here, these series are not useful without some sophisticated re-summation procedure. Fortunately for us this is not necessary for the purpose we have at hand, which is to apply the enhanced potential theory to a model for carbon field emitters, for which an alternative (perturbation) approach is more suitable.

5. Model for carbon field emitters

In recent years there has been widespread excitement in carbon based field emitters [12, 13, 21] which have the potential, (already realized) for cost-effective large area display systems. The motivation for our work is based on the observed high current density and, as we shall see, the non-exponential behaviour of these emitters. Our first observation is the relative magnitudes of the various coefficients, especially in the form equation (36). In particular we shall focus on the barrier top (see figure 1) which occurs when $\bar{\beta} = \bar{\epsilon}^2/4$ and hence also when $\gamma = \bar{\epsilon}^{3/2}/8$ becomes large. We note that this can be achieved through a large field enhancement effect of the order of F_b/F in FN–JWKB theory, see equation (9) as well as from $y_F = 1$ in equation (10). However, the current will be too large in general (since JWKB predicts $\mathcal{T} = 1$) and the field dependence of $\ln J$ will remain that of $1/F$, which falls too steeply as we shall see later. This is a consequence of the fact that, in the classic FN–JWKB theory, transport is mainly via quantum tunnelling *through* the classically forbidden region, even if the potential has been significantly lowered (see figure 1).

We will not discuss the mechanism by which the barrier lowering actually occurs in our model. Instead we shall adopt it as a *working hypothesis*, but we do not expect that *all*
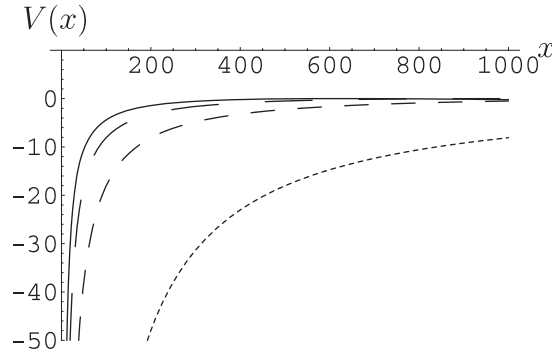
$V(x)$

Figure 2. Plot of our effective surface potential as in figure 1. Curves are drawn in decreasing order for $F = 16$ (full), 10 (long dash), 5 (short dash) and 1 (dotted) V μm^{-1}, respectively. $V(x)$ is in electronvolts, with the work function $\varphi = 2.0$ eV and the distance x from the surface is in ångströms. Note that there is no classical Schottky barrier here and the magnitude of $V(x)$ suggests that this potential can only be of very short duration, of the order of femtoseconds (see the text).

electrons in the band shall enjoy this opportunity, except for a narrow strip from $W = 0$ to ϑ *very* close to the Fermi level. Indeed we do not even expect that this (enhanced potential) barrier lowering phenomenon takes place in a general static sense. It is probably dynamic and needs to occur only over a timescale of the order of the transit time, which is about a few femtoseconds (10^{-15} s). Previous workers, such as Guth and Mullin [22], have noted the dominance of the image term for small distances of the order of ångströms from the surface. Our hypothesis here is to allow for a large enhancement in β, and hence $\bar{\beta}$, that effectively wipes out the barrier during the transit across the surface layer for a special group of near-Fermi-level electrons, which we shall call the ϑ electrons. We shall discuss the physical interpretation of this assumption later. For now we shall merely state our key hypothesis that, *for the purpose of calculating tunnelling probabilities* for the ϑ electrons, the parameter $\bar{\beta}$ shall be replaced by $\bar{\beta} = \bar{\epsilon}^2/4$ or equivalently $\gamma = \bar{\epsilon}^{3/2}/8$ in equation (36). In view of the magnitude of the parameter $\bar{\epsilon}$ it then suffices to obtain a solution in the limit in which this parameter is large and then treat the rest of the terms in the Schrödinger equation, such as in equation (36), as perturbations. *This is the sole mathematical basis of our theory, as well as its main underlying physical assumption.* For $\varphi = 2$ eV and $1 < F < 16$ V μm^{-1}, then $\bar{\epsilon}$ ranges from about 594 to 94, which should indeed justify treating the neglected terms as perturbations. In fact, a plot of the now enhanced potential (see figure 2) shows that the neglected terms are insignificant for the largest field $F = 16$ V μm^{-1} considered, up to some 1000 Å. This figure further shows that such a potential cannot be interpreted in the usual static sense, but we shall defer these discussions to section 6.

We expect that, over the distance of a few nanometres, tunnelling would have occurred from the surface if it is possible at all. Thus the long range departures of our approximate effective potential from the full potential will only contribute in a perturbative way. In particular, for the matching conditions we are most interested in the solution near to the origin $x = 0$, which corresponds to $\xi(0) = \xi(x = 0) = \bar{\epsilon}$, so that we can go back to the variable $y = \xi - \bar{\epsilon}$, which is small for our purpose $y \ll \bar{\epsilon}$ and then equation (28) can be well approximated by

$$y\psi''(y) + (y + \bar{\epsilon}/2)^2 \psi(y) \approx y\psi''(y) + (\bar{\epsilon}/2)^2 \psi(y) = 0. \tag{37}$$

The appropriate solution for the above is now given by the Hankel function of order 1 which,

upon transforming back to the original variables ξ, can be written as

$$\psi(\xi) = e^{i\pi/4}\sqrt{\frac{\pi}{2}}\bar{\epsilon}(\xi - \bar{\epsilon})^{1/2}H_1^{(1)}(\bar{\epsilon}(\xi - \bar{\epsilon})^{1/2}). \tag{38}$$

We can now obtain the transmission coefficient as was done in the previous section by matching the wavefunction and its derivative, but now at $\xi_0 = \xi(x = x_0)$, whose value we shall determine later. The result is then given by

$$\mathcal{T} = \frac{2\bar{\epsilon}^2\lambda_0^{1/3}}{|\psi(\xi_0)|^2 + \bar{\epsilon}^2\lambda_0^{1/3} + \lambda_0^{2/3}|\psi'(\xi_0)|^2}, \tag{39}$$

where the derivative of ψ is easily shown to be given in terms of the Hankel function of order 0:

$$\psi'(\xi) = e^{i\pi/4}\sqrt{\frac{\pi}{8}}\bar{\epsilon}^2 H_0^{(1)}(\bar{\epsilon}(\xi - \bar{\epsilon})^{1/2}). \tag{40}$$

As usual we can verify that this solution satisfies unitarity $\mathcal{T} + \mathcal{R} = 1$ through the use of the appropriate Wronskian identities for the Hankel functions. That this is satisfied exactly in equation (37) for the modified potential is important and future perturbative studies should respect this unitarity condition. In fact, we are not aware if other approximate procedures [22, 25] are so respectful of the unitarity condition. We believe this is necessary for any physically self-consistent theory for large transmission. We should remind the reader that the WKB approximation violates unitarity in an exponentially small way, *provided* the transmission coefficient is exponentially small, see section 50 [24]. Equation (39) has a weak logarithmic behaviour (due to the $1/x$ potential near the origin) in ψ'. However, the physics of our model does not go below the order of 1 Å. We shall therefore invoke this by letting $x_0 = 1$ Å as the matching point on the surface. Thus we may write to the leading order in large quantities

$$\mathcal{T} \approx \frac{4\pi}{|\bar{\epsilon}|^{3/2}[\ln(\delta^{1/2}|\bar{\epsilon}|^{3/2})]^2}\sqrt{\frac{W + \mu}{|W - \varphi|}}, \tag{41}$$

where $\delta(x_0)$, which we emphasize is a function of the matching point x_0, will be determined below. Physically it is reasonable that we may choose to connect the wavefunctions on the left to the right at the point x_0, since our theory cannot describe sub ångström processes, which must involve deBroglie wavelengths $\lambda \ll \lambda_F$. The equality of $\delta(x_0)$ so chosen with the energy of the ϑ electrons, i.e. $\vartheta \approx \delta(x_0)\varphi$, is an additional hypothesis, which seems an obvious choice. However, its justification lies outside the realm of the present theory, for which we can only speculate at this stage, see section 6. Equation (41) is the key result of our theory, which differs from the JWKB approximation ($\mathcal{T} = 1$) at the top of the barrier. As we have taken x_0 as about two Bohr radii, i.e. $x_0 \approx 2a_B = 1$ Å,[4] then by definition we can now determine $\xi_0 = \xi(x = x_0)$ from $\xi_0 - \bar{\epsilon} = \delta\bar{\epsilon} = \alpha^{1/3}x_0$ and thus we have

$$\delta = \delta(x_0) = \frac{\vartheta}{\varphi} = \frac{eFx_0}{\varphi} = \frac{10^{-4}eF}{\varphi}\ \mu\text{m}. \tag{42}$$

Using equation (41) we can now estimate the current density near the barrier top. We shall do this at finite temperatures using equations (5) and (6). The former can now be written as

$$J = \frac{me\varphi^2}{2\pi^2\hbar^3}\int_{-\delta}^{0}\mathcal{T}(\varphi\omega)\frac{k_BT}{\varphi}\ln\left(1 + \exp\left[-\frac{\varphi\omega}{k_BT}\right]\right)d\omega, \tag{43}$$

---

[4] By way of comparison, the typical tunnelling length in FN–JWKB theory (without field enhancement) is of the order of $\ell = \frac{\varphi}{eF} \sim 10^4$ Å to 10^3 Å for $1 < F < 16$ V μm^{-1}, with $\varphi = 3$ eV.

where we have introduced a dimensionless integration variable: $\omega = W/\varphi$. Note that due to the assumptions (1) and (2) in section 1, with all electrons belonging to the same band, the parameters δ, and hence also ϑ, do not enter the derivation of the supply function $N(W)$ in equation (6).

Now in view of the smallness of δ, and thus the exponent of the exponential throughout the region of integration, the above can be rewritten as

$$J \approx \frac{me\varphi^2}{2\pi^2\hbar^3} \int_{-\delta}^{0} \mathcal{T}(\varphi\omega)\left(\frac{k_B T}{\varphi} \ln 2 - \frac{1}{2}\omega\right) d\omega. \qquad (44)$$

An estimate of each term can now be made from equation (42) which shows that the finite temperature term would go as

$$\frac{\delta k_B T}{\varphi} \sim \frac{k_B T(eF) \times 10^{-4}}{\varphi^2}, \qquad (45)$$

which is at least five orders of magnitude down from the leading term for the temperature and fields we consider. We have to remark that the temperature dependence of δ, which we should emphasize as $\delta(x_0, T)$, is outside the realm of the present theory. However, we believe it is unlikely that the already small value of δ can be sustained to very high temperatures of the order of φ/k_B. Hence any serious discussions of the high temperature dependence are probably premature at this stage.

In view of the weak dependence of the logarithmic term in equation (41) the integration can proceed by extracting the leading field dependence in its argument. The final result we obtain is given by

$$J_{CH} \approx \frac{\sigma c_A \varphi^{-3/2}}{[\ln(c_B \varphi/(eF)^{1/2})]^2} \sqrt{\frac{\mu}{\varphi}} (eF)^3 \text{ A cm}^{-2}, \qquad (46)$$

where the constants are $c_A = 9.924 \times 10^{-2} \text{ eV}^{-3/2} \text{ } \mu\text{m}^3$ and $c_B = 51.231 \text{ eV}^{-1/2} \text{ } \mu\text{m}^{-1/2}$, respectively, and σ is the same empirical electron loss factor as introduced in the FN–JWKB formula equation (10) earlier.

Note that the field dependence of the current is in this case F^3 which differs from the FN–JWKB prefactor which goes as F^2 and there is *no* exponential dependence. We can now compare equation (46) with the standard FN–JWKB formula equation (10) and with experimental data [21]. To do this we need to take into account some empirical loss factors. This must include the cathode–anode probe geometry since ours is a one-dimensional model. Even if the cathode emitting surface is kept very small, comparable to the probe tip ($\sim$1 mm) in most cases, we do not expect all electrons to be collected. The imperfect supply is a further problem for composite granular systems, since the Fermi-supply function is an ideal case: some electrons could be trapped or recombined with holes at various sites, especially near contacts. These studies require more detailed knowledge of the experimental conditions. However, we expect that they do not vary significantly in order-of-magnitude terms with the field over the range considered here. Hence for simplicity we have adopted a consistent approach for both the FN–JWKB and our model by including a factor σ in the two formulae. For the FN–JWKB formula we further include a field enhancement factor Γ, i.e. $F \rightarrow \Gamma F$, see figure 3. We found that we need to have $\Gamma \sim 50$ to get the order of magnitudes shown for the FN current density. An interesting observation is that the larger departures at the higher fields in our case may be due to the neglected terms in equation (36), that is the 1/4 factor and the $\bar{\lambda}$, which a future perturbative calculation may be able to improve. We emphasize that we have only two adjustable parameters, namely φ and μ, to which the results are not too sensitive, within the reasonable values expected. The parameter σ is set by the FN–JWKB current so that we are comparing like with like. For FN–JWKB, there is an additional parameter Γ, which is the
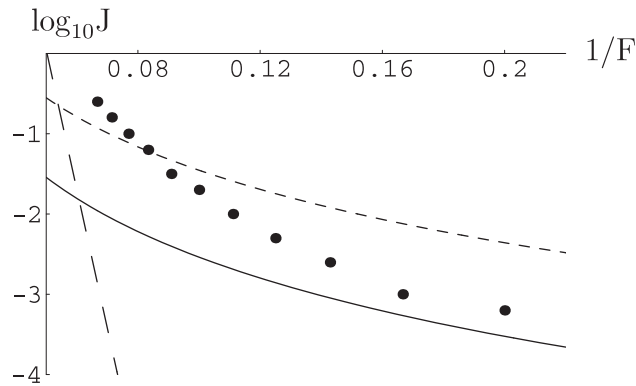
Figure 3. Logarithmic $\log_{10} J$ versus $1/F$ plot of the emission current in our model for epoxy-graphite PFE material. The black dots are experimental points from [21]. The full curve is our result which is given in equation (46). The dotted curve is the same but does not include the logarithmic field factor in the denominator. The broken straight line is the FN–JWKB result J_{FN}, as given in equation (10) with a field enhancement factor of $\Gamma = 50$. J is in units of A cm^{-2} and F is in units of V μm^{-1}. The work function $\varphi = 2.0$ eV, with $\mu = 2\varphi$ and $\sigma = 10^{-3}$ in both cases.
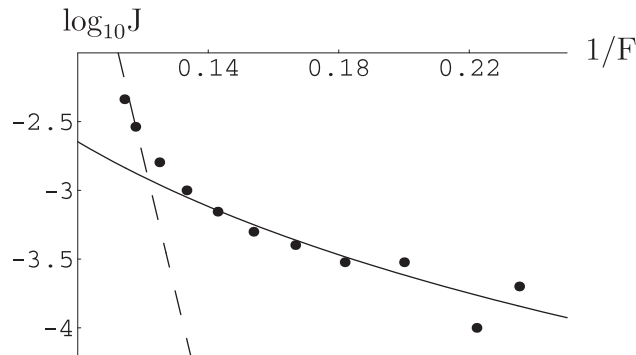

Figure 4. Logarithmic $\log_{10} J$ versus $1/F$ plot of the emission current in our model for graphitized carbon nanotube samples on an AAO template. The black dots are experimental points from [13]. The full curve is our result which is given in equation (46). The broken straight line is the FN–JWKB result J_{FN} as given in equation (10) with a field enhancement factor of $\Gamma = 350$. J is in units of A cm^{-2} and F is in units of V μm^{-1}. The work function $\varphi = 5.0$ eV, with $\mu = 2\varphi$ and $\sigma = 4.91 \times 10^{-3}$ in both cases.

enhancement factor, and the FN–JWKB current density is very sensitive, in order-of-magnitude terms, to its value. Especially interesting is a comparison with data from graphitized carbon nanotubes, see figure 4. These intriguing results inspire us to seek an interpretation for the strong potential enhancement, which we shall now discuss.

6. Discussions

In this section we will make no pretense that we are mainly speculating. The two key ingredients in the theory require some physical interpretation. These are (i) the strong surface potential enhancement and (ii) the connection between the matching point x_0 and the ϑ electrons. We shall first address the potential enhancement factor by noting that our barrier lowering assumption $\bar{\beta} = \bar{\epsilon}^2/4$ in the last section can be rewritten in terms of a suppression of the

effective dielectric constant κ_s for the surface potential, see equation (27). This can be deduced from

$$\phi(r) = -\frac{\beta}{\kappa_s r}, \tag{47}$$

whose *classical* image theory value for κ_s is given in terms of the bulk material *static* dielectric constant κ_b as

$$\kappa_s = \frac{\kappa_b + 1}{\kappa_b - 1}. \tag{48}$$

The well known value for κ_b is 11.9 for the case of silicon, making $\kappa_s \approx 1.18$. The factor of dielectric suppression in our model potential is therefore of the order of (from $\bar{\beta} = \bar{\epsilon}^2/4$)

$$\kappa_s = \frac{e^3 F}{(W - \varphi)^2} = \frac{F}{F_b} \frac{\varphi^2}{(W - \varphi)^2} \approx \frac{F}{F_b}, \tag{49}$$

for the ϑ electrons. This massive dielectric suppression is therefore of the same order as the field enhancement factors required in classical field-enhanced FN–JWKB theory to obtain the same magnitude of current density. Such a suppression cannot occur in a static theory, since static dielectric constants must be greater than unity on thermodynamic grounds [42]. Thus we now need to consider a dynamical picture, whose key parameters may be the Fermi wavelength λ_F and a characteristic angular frequency ω_F, which will be typically of the order of the inverse transit time $\tau_F = \lambda_F/v_F \sim 10^{-15}$ s, as mentioned before. It is well known that for the case of silicon (see, for example, [33] section 5.3), the electron transit time $\tau_F \sim 10^{-14}$ s is slow by comparison with the dielectric relaxation time, which is controlled by the plasmon frequency $\tau_p \sim 10^{-16}$ s. Hence the static polarizability prevails and the classical Schottky mechanism holds for silicon. Since graphite is known to have a low density of states near the Fermi level, we would expect a significant modification of this picture. We note in passing that the transit time is still a subject of some contention [43]. Indeed a proper time-dependent theory of quantum tunnelling that goes beyond WKB is still not available at present, to the best of our knowledge. Our work here reveals that this development is a matter of some urgency. Therefore we expect our estimate could be out by orders of magnitude. Nevertheless if we accept this femtosecond estimate, then the electromagnetic wavelengths concerned would be of the order of 3000 Å which is in the near-ultraviolet range. This effect is likely to be confined only to the order of λ_F from the surface, since if not, there would be consequences more easily detectable than in field emission. An examination of equation (49) further shows that the effect is non-linear. A polarizability that is linear in F implies a quadratic order dipole polarization from a time-dependent perturbation theory perspective [44]. Hence optical measurements that probe the surface, perhaps using evanescent waves for example, might be extremely revealing. We do not know to what extent this property is related to the graphitization of the carbon particle surfaces, as details have not yet been systematically investigated to the best of our knowledge.

Next we wish to state clearly that we do not yet know the underlying physics, at this stage, for relating the energy of the ϑ electrons through the closest approach matching point x_0 as defined by $\delta(x_0)$ in equation (42). We adopted this ansatz as the most obvious choice. It is intriguing, however, that the latter is merely the ratio between an atomic polarization energy $\vartheta = eFx_0$ versus the work function φ, see equation (42). Heuristically this implies that one can expect the dielectric suppression, which must only involve a fraction δ of surface atoms, would also involve the same fraction of electrons near the Fermi level in any dynamical dielectric processes. However, one has to note that this choice may well be peculiar only to the $1/x$ potential, which could also be modified at distances of the order of the atomic

size. Indeed a systematic study using a more sophisticated enhanced potential, which includes dipole, quadrupole and higher-order multipole terms, would be very interesting, since this would lead to angular predictions of the emission current density that might be detectable experimentally. The relevant parameters and their connection with the ϑ electrons will remain a challenge for a more sophisticated dynamical theory. The order of the onset field for the proposed dielectric phenomenon to occur, which is $F \sim 1$ V $\mu\mathrm{m}^{-1}$, is particularly important to further elucidate and must not be forgotten.

7. Conclusion

In conclusion, we have presented a relatively high current field emission model which is based on a strongly enhanced surface potential. The enhancement leads to a significant barrier lowering that mimics the classical field-enhanced Schottky barrier lowering mechanism. In view of this strong enhancement, which grossly violates the WKB validity criteria near the surface, we are compelled to pursue an analytic solution of the one-dimensional Schrödinger equation. Fortunately we were able to do this, with less effort than the classical image (zero-field) Nordheim/MacColl solution [2, 18]. By treating the smaller terms as perturbations and the assumption that only a thin layer of electrons near the Fermi level (which we call the ϑ electrons) are involved in the process, we found an intriguing agreement with experimental data. This includes epoxy graphite composite PFE [21] and graphitized carbon nanotubes on an AAO template [13] field emitters, both in terms of the order of magnitude as well as the shape of the curve in a ln J versus $1/F$ plot, particularly at lower onset fields. We further interpret this result as a massive dielectric suppression involving only the near-Fermi-level ϑ electrons and the surface atoms. Further work is necessary, which should include more accurate calculations, perhaps employing the results from the exact infinite series solutions given in the appendix. Understanding the microscopic physics of this phenomenon might involve dynamical effective medium concepts [45] and also shape enhancement factors [46]. We may perhaps further speculate that, in a microscopic theory, our effective potential, see figure 2, is only meaningful as a relevant Fourier component of a time-dependent fluctuating potential. Hence a proper time-dependent tunnelling theory that is coupled to an underlying mechanism for the dielectric suppression involving the ϑ electrons will be a highly worthwhile challenge.

Postscript: after this work was completed, we have found that our solution equation (38) has also been previously obtained by MacColl [18] in the context of zero glancing angle reflection from metals with no applied fields. His reflection coefficient $\mathcal{R}$ is given by $\mathcal{R} = 1 - \mathcal{T}$ in terms of our $\mathcal{T}$ in equation (39).

Appendix A. Analytical continuation of Bessel functions

Confusion arises with Bessel functions of order 1/3 due to the special care needed to treat the phase factors: some popular mathematical software uses incorrect continuation formulae. Firstly our convention is the same as Watson's so that the appropriate continuation formulae

are given by

$$I_\nu(z) = \begin{cases} e^{-i\nu\pi/2} J_\nu(ze^{i\pi/2}) & \text{if } -\pi < \arg(z) \leqslant \pi/2 \\ e^{3i\nu\pi/2} J_\nu(ze^{-3i\pi/2}) & \text{if } \pi/2 < \arg(z) \leqslant \pi. \end{cases} \tag{A.1}$$

An approach to perform the continuation for negative z is

$$u_1 = \frac{\sqrt{\pi}\sqrt{x}J_{-1/3}(\frac{2}{3}x^{3/2})}{3} = \frac{i}{3}\sqrt{\pi}\sqrt{|x|}J_{-1/3}\left(\frac{-2i}{3}|x|^{3/2}\right)$$
$$= \frac{i}{3}e^{i\pi/6}\sqrt{\pi}\sqrt{|x|}I_{-1/3}\left(\frac{2|x|^{3/2}}{3}\right), \tag{A.2}$$

which is incorrect. The negative argument prevents the use of the first form of equation (A.1). Instead the branch cut properties require the continuation formula [32]

$$J_\nu(e^{im\pi}z) = e^{im\nu\pi}J_\nu(z) \tag{A.3}$$

which holds similarly for the I_ν functions. Then

$$u_1 = \frac{\sqrt{\pi}\sqrt{x}J_{-1/3}(\frac{2}{3}x^{3/2})}{3} = \frac{i}{3}\sqrt{\pi}\sqrt{|x|}J_{-1/3}\left(\frac{-2i}{3}|x|^{3/2}\right)$$
$$= \frac{1}{3}\sqrt{\pi}\sqrt{|x|}I_{-1/3}\left(\frac{2|x|^{3/2}}{3}\right), \tag{A.4}$$

without any complex phase factors. Similarly u_2 picks up a minus sign:

$$u_2 = \frac{\sqrt{\pi}\sqrt{x}J_{1/3}(\frac{2}{3}x^{3/2})}{3} = \frac{i}{3}\sqrt{\pi}\sqrt{|x|}J_{1/3}\left(\frac{-2i}{3}|x|^{3/2}\right) = -\frac{1}{3}\sqrt{\pi}\sqrt{|x|}I_{1/3}\left(\frac{2|x|^{3/2}}{3}\right). \tag{A.5}$$

Hence one has to manually continue the J_ν to the I_ν functions for negative arguments when, if using the incorrect continuation, equation (A.2) results. We have found that Mathematica v4.1 uses equation (A.2).

Appendix B. Infinite series solutions

The generic form of the full image finite field problem is given by the second-order differential equation:

$$(\xi - \bar{\epsilon})\psi'' + \xi(\xi - \bar{\epsilon})\psi + \bar{\beta}\psi = 0. \tag{B.1}$$

Unfortunately this does not belong to the standard hypergeometric class and cannot therefore be treated by the usual mathematical physics functions. However, using the standard Frobenius method one can obtain the following power series solutions. Let

$$u_m(\xi) = \sum_{n=0}^{\infty} a_n \xi^{m+n}, \tag{B.2}$$

then equation (B.2) leads to the following five-term recursion relation for the determination of the coefficients a_n:

$$a_{n-1}(m+n-1)(m+n-2) - a_n\bar{\epsilon}(m+n)(m+n-1) + a_{n-4} - \bar{\epsilon}a_{n-3} + \bar{\alpha}a_{n-2} = 0. \tag{B.3}$$

All the coefficients are well determined in terms of the coefficient a_0 which is a normalization factor that can be set to 1. The first term satisfies

$$\bar{\epsilon}m(m-1)a_0 = 0, \tag{B.4}$$

giving $m = 0$ or 1, which shows, as expected, that we shall have two linearly independent solutions. Carrying on we have $a_1 = 0$ and

$$a_2 = \frac{\bar{\beta}}{\bar{\epsilon}} \frac{a_0}{(m+2)(m+1)}. \tag{B.5}$$

Notice that $a_1 = 0$ regardless but a_2 is only finite for $\bar{\beta} \neq 0$, i.e. it originates from the image potential. The following coefficients can be sequentially determined:

$$a_3 = \frac{a_2(m+1)}{\bar{\epsilon}(m+3)} - \frac{a_0}{(m+3)(m+2)}, \tag{B.6}$$

$$a_4 = \frac{a_3(m+2)}{\bar{\epsilon}(m+4)} + \frac{a_0}{\bar{\epsilon}(m+4)(m+3)} + \frac{\bar{\beta}}{\bar{\epsilon}} \frac{a_2}{(m+4)(m+3)}, \tag{B.7}$$

and so on. The general five-term recursion relation can be re-written as

$$a_n = \frac{1}{(m+n)(m+n-1)} \left[\frac{1}{\bar{\epsilon}} a_{n-1}(m+n-1)(m+n-2) + \frac{\bar{\beta}}{\bar{\epsilon}} a_{n-2} - a_{n-3} + \frac{1}{\bar{\epsilon}} a_{n-4} \right]. \tag{B.8}$$

Unfortunately there is no way to simplify this any further. When $\bar{\beta} = 0$ (image free case) the above in fact reduces to a two-term relation given by

$$a_n = -\frac{a_{n-3}}{(m+n)(m+n-1)}. \tag{B.9}$$

This immediately leads us to the two independent solutions, for $m = 0$:

$$u_0(\xi) = 1 + \sum_{n=1}^{\infty} \frac{(-1)^n 1 \cdot 4 \cdot 7 \cdots (3n-2)}{(3n)!} \xi^{3n}$$

$$= \frac{\Gamma(2/3)}{3^{1/3}} \xi^{1/2} J_{-1/3}\left(\frac{2}{3}\xi^{3/2}\right), \tag{B.10}$$

and for $m = 1$:

$$u_1(\xi) = \xi + \sum_{n=1}^{\infty} \frac{(-1)^n 2 \cdot 5 \cdot 8 \cdots (3n-1)}{(3n+1)!} \xi^{3n+1}$$

$$= 3^{1/3} \Gamma(4/3) \xi^{1/2} J_{1/3}\left(\tfrac{2}{3}\xi^{3/2}\right), \tag{B.11}$$

which are the series expansions for the Bessel functions of order $1/3$. The oscillatory nature of these functions means that the convergence is, in general, poor and some more sophisticated re-summation method must be used to obtain, in particular, the asymptotic properties of these functions. Also an independent solution does not seem to exist for the zero-field finite image case using the above power series equation (B.2), which is perhaps not surprising, since the Whittaker functions [2, 18, 40] have a branch point at the origin. However, the re-summation procedure described below does yield an independent solution, which is unfortunately a divergent series.

B.1. A method of re-summation

An alternative solution for the full image problem which is an effective re-summation of the previous power series solution that will allow us to perform numerical calculations, will now be shown. The approach is to first use a Laplace transform on equation (B.1). Let

$$\psi(\xi) = \int_C e^{\xi t} f(t) \, dt, \tag{B.12}$$

over some contour C, then the resulting differential equation in ψ becomes converted into a second-order differential equation in f:

$$f'' - (t^2 - \bar{\epsilon}) f' + (\bar{\beta} - \bar{\epsilon}t^2 - 2t) f = 0. \tag{B.13}$$

It is easy to show that the solution for $\bar{\beta} = 0$ reduces to the Airy integral where $f(t) = e^{t^3/3}$, so we shall factorize f in the form: $f = ug$, where $u = e^{t^3/3}$. The equation for g now simplifies tremendously:

$$g'' + (t^2 + \bar{\epsilon})g' + \bar{\beta}g = 0. \tag{B.14}$$

Unfortunately there is still no closed form solution for this equation although modern methods of analysis have now yielded some interesting new results. For it can be proved that solutions for g must be an entire function, by which we can develop an expansion as an infinite series of the form [38, 39]

$$g = \sum_{n=0}^{\infty} a_n t^{m+n}. \tag{B.15}$$

The boundary condition requires that when $\bar{\beta} \to 0$ then $g \to 1$. With this we can show that $m = 0$ is the only allowed index. Then we have the following four-term recursion relation:

$$(n-1)(na_n + \bar{\epsilon}a_{n-1}) + (n-3)a_{n-3} + \bar{\beta}a_{n-2} = 0. \tag{B.16}$$

The first few solutions are easily obtained: $a_0 = 1$, $a_2 = \frac{\bar{\beta}}{2}$, $a_3 = \frac{\bar{\epsilon}\bar{\beta}}{6}$ and $a_4 = \frac{\bar{\beta}}{24}(\bar{\beta} - \bar{\epsilon}^2)$ and so on. As required only the first term survives in the limit $\bar{\beta} = 0$. Interestingly the four-term relation equation (B.16) can actually be reduced to a three-term one. Let $\gamma_n = \frac{a_n}{a_{n-1}}$ then

$$(n-1)(n\gamma_n + \bar{\epsilon})\gamma_{n-1} + \frac{(n-3)}{\gamma_{n-2}} + \bar{\beta} = 0, \tag{B.17}$$

which may be more convenient for computation. Finally we have now effectively re-summed the series to obtain

$$\psi(\xi) = \sum_{n=0}^{\infty} a_n \int_C t^n e^{\xi t + \frac{1}{3}t^3} \, dt. \tag{B.18}$$

We now see that

$$\begin{aligned}
\int_C t^n e^{\xi t + \frac{1}{3}t^3} \, dt &= \frac{d^n}{d\xi^n} \int_C e^{\xi t + \frac{1}{3}t^3} \, dt \\
&= \frac{d^n}{d\xi^n}\left[\xi^{1/2} Z_{\pm 1/3}\left(\frac{2}{3}\xi^{3/2} \right) \right],
\end{aligned} \tag{B.19}$$

the last term following from the relation between the Airy functions and the Bessel functions of order $1/3$. In view of the recursion properties for the derivatives of the Bessel functions, we have now converted the solution effectively into an infinite series of Bessel functions. Only the first term survives in the image-free case, which is now trivial, and all the oscillatory bits are now absorbed into the Bessel functions. A further re-summation appears to be necessary to recover the Whittaker functions in the zero-field finite image $\bar{\beta} \neq 0$ case.

References

[1] Fowler R H and Nordheim L W 1928 *Proc. R. Soc.* A **119** 173–81
[2] Nordheim L W 1928 *Proc. R. Soc.* A **121** 626–39
[3] Millikan R A and Lauritsen C C 1928 *Proc. Natl Acad. Sci.* **14** 45
[4] Jeffreys H 1924 *Proc. Lond. Math. Soc.* **23** 428–36

Priority for the WKB method appears to go back to

Liouville J 1837 *J. Math.* **2** 16

Liouville J 1837 *J. Math.* **2** 418

and later

Lord Rayleigh 1912 *Proc. R. Soc.* **86** 207

 [5] Wentzel G 1926 *Z. Phys.* **38** 518
 [6] Kramers H A 1926 *Z. Phys.* **39** 828
 [7] Brillouin L 1926 *C. R. Acad. Sci.* **183** 24
 [8] Burgess R F, Kroemer H and Houston J M 1953 *Phys. Rev.* **90** 515
 [9] Good R H and Müller E W 1956 *Handbuch der Physik* vol 21, ed S Flugge (Berlin: Springer) pp 176–91
[10] Gadzuk J W and Plummer E W 1973 *Rev. Mod. Phys.* **45** 487
[11] Jensen K L 2001 Theory of field emission *Vacuum Microelectronics* ed W Zhu (New York: Wiley)
[12] Burden A P 2001 Materials for field emission displays *Int. Mater. Rev.* **46** 213–31
[13] Gao H *et al* 2003 *J. Appl. Phys.* **93** 5602–5
[14] Hastas N A *et al* 2002 *Semicond. Sci. Technol.* **17** 662–7
[15] Mott N F and Gurney R W 1948 *Electronic Processes in Ionic Crystals* (Oxford: Oxford University Press)
[16] Herring C H 1950 *Rev. Mod. Phys.* **21** 187
[17] Dearnaley G, Stoneham A M and Morgan D V 1970 *Rep. Prog. Phys.* **33** 1129
[18] MacColl L A 1939 *Phys. Rev.* **56** 699–702
[19] Sternin B Y and Shatalov V E 1996 *Borel–Laplace Transform and Asymptotic Theory—Introduction to Resurgent Analysis* (Boca Raton, FL: CRC Press)
[20] Li Y B, Bando Y and Goldberg D 2003 *Appl. Phys. Lett.* **82** 1962–4
[21] Tuck R A, Taylor W and Latham R V 1997 *Proc. 4th Int. Display Workshop* (Nagoya: Society for Information Display) pp 723–6
[22] Guth E and Mullin C J 1940 *Phys. Rev.* **59** 575–84
[23] See, for example, Zangwill A 1992 *Physics at Surfaces* (Cambridge: Cambridge University Press) pp 57–63
[24] Landau L D and Lifshitz E M 1991 *Quantum Mechanics* (Oxford: Pergamon) p 165 (section 49)
[25] Jensen K L *et al* 2002 *Appl. Phys. Lett.* **81** 3867–9
[26] Modinos A 1984 *Field, Thermionic and Secondary Emission Spectroscopy* (New York: Plenum)
[27] Roy D K 1977 *Tunnelling and Negative Resistance Phenomena in Semiconductors* (*Int. Series on the Science of the Solid State* vol 11) (Oxford: Pergamon) p 53
[28] Delbourgo R 1977 *Am. J. Phys.* **45** 1110–2
[29] Duke C B 1969 *Tunnelling in Solids* (*Solid State Physics Suppl.* 10) ed F Seitz, D Turnbull and H Ehrenreich (New York: Academic)
[30] Moll J L 1964 *Physics of Semiconductors* (New York: McGraw-Hill)
[31] Landau L D and Lifshitz E M 1991 *Quantum Mechanics* (Oxford: Pergamon) pp 74–9 (section 24)
[32] Watson G H 1944 *Theory of Bessel Functions* (Cambridge: Cambridge University Press) p 77
[33] Sze S M 1981 *Physics of Semiconductors* 2nd edn (New York: Wiley) pp 254–65 and references quoted therein
[34] Murphy E L and Good R H Jr 1956 *Phys. Rev.* **102** 1464–73
[35] Stratton R 1962 *Phys. Rev.* **125** 67–82
[36] Schiff L I 1965 *Quantum Mechanics* (New York: McGraw-Hill) p 103
[37] Shawyer B and Watson B 1994 *Borel's Methods of Summability* (Oxford: Oxford University Press)
[38] Gundersen G G 1988 *Trans. Am. Math. Soc.* **305** 415–29
[39] Belaïdi B and Hamouda S 2001 *Electron. J. Diff. Eqns* **2001** 1–5
[40] Gradshteyn I S and Ryzhik I M 1980 *Table of Integrals, Series and Products* (New York: Academic) pp 1059–64 (sections 9.22–23)
[41] Choy T C 2000 *Phys. Rev.* A **62** 12506–16 and references quoted therein
[42] Landau L D, Lifshitz E M and Pitaevskii L P 1984 *Electrodynamics of Continuous Media* (Oxford: Pergamon) p 59 (section 15)
[43] Landauer R and Martin Th 1994 *Rev. Mod. Phys.* **66** 217–28
[44] Langhoff P W, Epstein S T and Karplus M 1972 *Rev. Mod. Phys.* **44** 603–42
[45] Choy T C 1999 *Effective Medium Theory—Principles and Applications* (Oxford: Oxford University Press) pp 47–71 and references therein
[46] Jaeger D L, Hren J J and Zhirnov V V 2003 *J. Appl. Phys.* **93** 691–7