

Extraction of emission area from Fowler–Nordheim plots

Richard G. Forbes,^{a)} Jonathan H. B. Deane, Nabil Hamid, and Hwa San Sim
*University of Surrey, School of Electronics and Physical Sciences, Guildford, Surrey GU2 7XH,
United Kingdom*

(Received 6 October 2003; accepted 9 February 2004; published 1 June 2004)

The process of deriving emission-area estimates from Fowler–Nordheim (FN) plots is investigated, using emission-area extraction functions and an iterative procedure suggested earlier [R. G. Forbes, *J. Vac. Sci. Technol. B* **17**, 526 (1999)]. Simulated FN-plot data have been prepared using free-electron theory and three different tunneling-barrier models (elementary triangular barrier, image-rounded “Schottky–Nordheim” barrier, and a quadratically enhanced barrier). These have been analyzed using two area-extraction spreadsheets with different models (elementary triangular barrier and Schottky–Nordheim barrier) embedded in them. It is confirmed that significant errors in the area estimate may occur if the emission model used to analyze the FN-plot data does not correspond well with the model/physics responsible for generating the data. It is also concluded that parameters used in the area-extraction process should correspond to emission variables in the midrange of the FN-plot data. © 2004 American Vacuum Society. [DOI: 10.1116/1.1691410]

I. INTRODUCTION

If the field electron emission (FE) theory assumed when extracting an emission area from a Fowler–Nordheim (FN) plot differs from the correct theory, then incorrect area values may be derived. Forbes and Jensen showed this when analyzing FE from a hemiellipsoidal emitter.¹ However, the need to integrate emission current density over the emitter surface obscured this conclusion. This note illustrates the effect more clearly, by using a flat emitter surface, free-electron theory and several models for the tunneling barrier.

We use an arbitrary value (10^{-14} m^2) for the “true emission area” A_{true} , and simulate FN-plot data using three models, namely an elementary triangular (ET) barrier, the standard image-rounded “Schottky–Nordheim” (SN) barrier, and a “quadratically enhanced” (QE) barrier discussed below.

The resulting FN plots are then analyzed, as if experimental data, using a linear regression program in the “XGRACE” software package. This yields the slope S and intercept $\ln\{R\}$ of the FN plot, together with their standard errors. Since a FN plot is, in principle, slightly curved (except for the ET barrier), these standard errors relate to the fitting of a straight line to well-defined data points lying on a slight curve. (With real experimental data, the standard errors would relate partly to experimental variability and measurement error.)

The values of R and S are then fed into spreadsheets developed from the one described in Ref. 2. These spreadsheets implement the iterative emission-area extraction procedure of Ref. 3, and yield an estimate A_{est} of emission area, together with uncertainty limits. For experimental data, these limits would relate to uncertainty about the emitter work-function and current-density values applicable to the experiments. We pretend that an equivalent lack of knowledge applies to our simulated data.

Each spreadsheet uses a specific barrier model. Two

spreadsheets have been developed, based on the ET barrier and the SN barrier. One can then investigate how A_{est} compares with A_{true} , for various combinations of the barrier models used in the FN-plot simulations and the spreadsheet analyses. The ratio $A_{\text{est}}/A_{\text{true}}$ is a measure of the consistency of a simulation and analysis process, and should be close to unity if the process is self-consistent. We show in Sec. IV that, if the models used in simulation and analysis are significantly different, then $A_{\text{est}}/A_{\text{true}}$ can differ significantly from unity.

An alternative demonstration of this uses the single-step area-extraction procedure, based on Eq. (11) below, that is described in outline in Ref. 3. We discuss this in Sec. III.

II. THEORY

All three barrier models used generate variants of the generalized (phenomenological) FN equation introduced in Ref. 3. This may be written as

$$I = \lambda A a \phi^{-1} F^2 \exp[-\mu b \phi^{3/2}/F], \quad (1)$$

where ϕ and F are the local work function and the barrier field, a and b are the usual universal FN constants,³ A is emission area, and λ and μ are the generalized correction factors introduced in Ref. 3. In FN coordinates, where $\ln\{I/F^2\}$ is plotted against $1/F$, Eq. (1) becomes

$$\ln\{I/F^2\} = \ln\{\lambda R^{\text{el}}\} + \mu S^{\text{el}}/F, \quad (2a)$$

where

$$R^{\text{el}} \equiv A a \phi^{-1}, \quad S^{\text{el}} \equiv -b \phi^{3/2}. \quad (2b)$$

The parameter μ is usually a function of F , so usually Eq. (2) represents a slightly curved line. Further (except for plots associated with the elementary triangular barrier), it is known that, as $1/F$ reduces toward zero, a theoretical FN plot stops at the value $1/F_b$, where F_b is the barrier-field value at which $\mu(F)$ (and, hence, the barrier) becomes zero. The value of the “constant” $[\ln\{\lambda A a \phi^{-1}\}]$ in Eq. (2a) is thus given by the value of $\ln\{I/F^2\}$ at $1/F = 1/F_b$, not by its value

^{a)}Electronic mail: r.forbes@surrey.ac.uk

at $1/F=0$. However, an experimental plot will stop before $1/F$ reduces to $1/F_b$, because usually emitter melting or explosion would intervene, so the value of $\ln\{I/F_b^2\}$ is difficult or impossible to identify from the FN plot.

For all these reasons, an experimental FN plot (with its associated regression line) is best modeled by a *tangent* to Eq. (2). These tangents have the form

$$\ln\{I/F^2\} = \ln\{R\} + S/F = \ln\{rR^{\text{el}}\} + sS^{\text{el}}/F, \quad (3)$$

where $\ln\{R\}$ and S are the *predicted* intercept and slope of the tangent, and r ($\equiv R/R^{\text{el}}$) and s ($\equiv S/S^{\text{el}}$) are the corresponding intercept and slope correction factors, given in terms of λ and μ by³

$$s = \mu - Fd\mu/dF, \quad (4a)$$

$$r = \lambda \exp[S^{\text{el}}d\mu/dF]. \quad (4b)$$

From Eqs. (2b) and (3), it can be shown that

$$RS^2 = Aab^2rs^2\phi^2 \equiv AC_2\Gamma, \quad (5)$$

where the universal constant $C_2[\equiv ab^2 \cong 7.192489 \times 10^{13} \text{ A m}^{-2} \text{ eV}^2]$, and the emission-area extraction function $\Gamma[\equiv rs^2\phi^2]$ are defined by this equation. The parameters r and s vary with the point at which the tangent is taken: r , s , and Γ are functions of ϕ and F . Alternatively, and better, r , s , and Γ can be regarded as functions of ϕ and the emission current density J .

The values of r , s , and Γ also depend on the emission model used. In particular, for the elementary FN theory, $r = s = 1$; for the standard FN theory, s is about 1 but r is of order 100.³

For the elementary and standard FN theories, parameter values can be calculated from conventional “analytical” formulas^{2,4} including (for the standard FN theory) a formula^{2,5} for evaluating complete elliptic integrals that is accurate to two parts in 10^8 .

However, parameters for other models or barriers must be calculated numerically. A simple option for any metal free-electron theory is as follows.⁶ The probability D of electron escape through a barrier defined by the quasiclassical electron potential energy $V(x)$ can be written in terms of the “Gamow exponent” G by

$$D = \exp[-G]. \quad (6a)$$

In the Jeffreys–Wentzel–Kramers–Brillouin approximation,⁷ G is given by

$$G = \{(8m)^{1/2}/\hbar p\} \int V^{1/2} dx, \quad (6b)$$

where m is the free-space electron mass, $\hbar p$ is Planck’s constant divided by 2π rad, and $\{(8m)^{1/2}/\hbar p\}$ is a universal constant (for electrons) of value $10.24634 \text{ eV}^{-1/2} \text{ nm}^{-1}$.

For the ET barrier, one writes $V = h - eFx$, where h is the barrier height and e is the elementary positive charge, and evaluates Eq. (6b) analytically, obtaining

$$G = G^{\text{el}} = bh^{3/2}/F. \quad (7)$$

For a more general barrier, one evaluates the JWKB integral in Eq. (6b) numerically and writes the result as $G = \nu G^{\text{el}}$, where ν is a correction factor defined by this equation. The corresponding emission current is then evaluated from the “general-barrier FN equation”:⁶

$$I = \tau_F^{-2} Aa\phi^{-1} F^2 \exp[-\nu_F b\phi^{3/2}/F], \quad (8a)$$

where τ is a correction factor defined by

$$\tau = \nu + hd\nu/dh, \quad (8b)$$

and ν_F and τ_F are the values of ν and τ evaluated for $h = \phi$. Obviously, τ_F^{-2} and ν_F are specific instances of the parameters λ and μ in Eq. (1).

The elementary FN theory can be derived from Eq. (8) by putting $\tau = \nu = 1$, and the standard FN theory by substituting for ν and τ the well-known special FE elliptic functions ν and t .^{2,4} Our third model, involving a QE barrier was defined by writing

$$V_n = h - eFx - e^2/16\pi\epsilon_0 x + (eF/a_t)x^2, \quad (9)$$

where ϵ_0 is the electric constant, and a_t is a parameter (the “tip radius”) that was made equal to 10 nm. Values of emission current I , as a function of field, were then evaluated for us by Koh and Yap as part of their BEng projects. They used Eqs. (6) to (9) above and a program initially developed by one of us (J. H. B. D.). These values provided the simulated FN plot for the QE barrier.

We note in passing that the precision of the integration program, estimated as about one part in 10^{15} , is much better than the precision of the usual formulas^{2,5} for the approximate evaluation of complete elliptic integrals. So, it is more precise to use the integration program for the SN barrier, too. It thus seems likely that use of these elliptic-integral formulas, and the associated tiresome mathematics,^{2,4} may become obsolete in the context of FE.

III. ILLUSTRATION USING THE SINGLE-STEP FORMULA

For a simulated FN plot based on a chosen emission model and a chosen value A_{true} of an emission area, and for given work-function and field values (hence, known values of J and I), we know the true values of emission area and of r , s , Γ , R , and S , and can rearrange Eq. (5) as

$$A_{\text{true}} = R_c S_c^2 / C_2 \Gamma_{\text{true}}, \quad (10)$$

where Γ_{true} , R_c , and S_c are known correct values for the emission process responsible for the FN-plot data and for the ϕ , J , and I values involved. [One would normally consider a field (and corresponding J and I values) near the middle point of the data range, because a linear regression fit is best modeled by a tangent to Eq. (2) taken near this middle point.]

With experimental data, or with simulated data where we pretend we have no information other than the data points, we are presented only with “empirical” values (R_e and S_e) of R and S obtained by linear regression or manual line fitting. There are then several difficulties that may prevent us

from reliably extracting the emission area. (1) We may not know the true emission model. (2) There may be uncertainty about the relevant value of local work function ϕ . (3) Usually, we do not know to which current-density (J) value (“in the middle of the data range”) the values R_e and S_e correspond. Also, if there is no reliable knowledge concerning emission area, then we cannot reliably convert measured current into true local current density J .

Therefore, we seek to restrict the error (in estimating the emission area) that results from uncertainty in ϕ or J , by making an estimate A_{est} of the emission area from the formula

$$A_{\text{est}} = R_e S_e^2 / C_2 \Gamma_{\text{model}}, \quad (11)$$

and placing uncertainty limits on Γ_{model} . These limits are derived from calculations using the assumed emission model and assumptions about the possible ranges of ϕ and J .

For example, consider analyzing FN plots from metals using the standard FN theory. If we assume that $3.5 \leq (\phi/\text{eV}) \leq 6.5$ and $10^5 \leq (J/\text{A m}^{-2}) \leq 10^{12}$, then Ref. 3 shows that we need to take $\Gamma_{\text{model}} (= \Gamma_{\text{st}}) = 2000 \text{ eV}^2 \pm 50\%$.

Experimental data are usually measured/reported as a function of voltage or macroscopic field, because the barrier field is not easily measurable. The use of the product $R_e S_e^2$ in Eq. (11) ensures that conversion factors or enhancement factors drop out of the theoretical argument,³ certainly in regard to first-order effects. This means that normal experimental data can be analyzed using the methods discussed here. However, for the present investigations, it is easier to use the barrier field as the independent variable.

The above approach restricts the uncertainty about ϕ and J values, but does not deal with the uncertainty about the emission model. Consider, for simplicity, a hypothetical situation in which the correct emission model is a free-electron model based on the ET barrier, but the empirical data are analyzed using the standard FN theory. As before, take the uncertainty in the work function as $3.5 \leq (\phi/\text{eV}) \leq 6.5$. For the elementary FN theory, $r = s = 1$, so from the definition of Γ via Eq. (5), this range of ϕ leads to the range $12 \leq (\Gamma_{\text{el}}/\text{eV}^2) \leq 42$, and the “correct” substitution into Eq. (11) should be $\Gamma_{\text{model}} (= \Gamma_{\text{el}}) = 27 \text{ eV}^2 \pm 55\%$. Clearly, if instead the analysis uses the “incorrect” standard FN theory, which takes $\Gamma_{\text{model}} (= \Gamma_{\text{st}}) = 2000 \text{ eV}^2 \pm 50\%$, then this would produce an emission-area estimate that is too small by a factor of about 75.

IV. ILLUSTRATION USING ITERATIVE PROCEDURES

A. Methods

The uncertainties about ϕ and J values, but particularly those about J values, can be reduced by using an iterative procedure as outlined in Ref. 3. This is based on knowing a “typical midrange current” I_{mid} for the experimental or simulated FN-plot data. An assumption in the discussion in Sec. III is that there is initial uncertainty by a factor of up to $10^{3.5}$ in our knowledge of J . This, and the uncertainty in ϕ ,

TABLE I. Mean values of the area ratio $A_{\text{est}}/A_{\text{true}}$, and associated uncertainties, for simulated FN-plot data analyzed using the standard spreadsheet. The upper error limits are associated with uncertainty relating to ϕ and J ; the lower error limits (if present) relate to the regression fit.

FN-plot data source	$\phi =$		
	4 eV	4.5 eV	5 eV
Elementary data	0.0089 $\pm 5.5\%$	0.011 $\pm 5.5\%$	0.013 $\pm 5\%$
Standard data	0.42 $\pm 2.5\%$ $\pm 3.8\%$	0.49 $\pm 2.1\%$ $\pm 4.6\%$	0.56 $\pm 2.0\%$ $\pm 4.7\%$
QE data	1.12 $\pm 1.5\%$ $\pm 12\%$	1.69 $\pm 0.8\%$ $\pm 14\%$	2.94 $\pm 0.8\%$ $\pm 18\%$

lead (for analyses using the standard FN theory) to an uncertainty of $\pm 50\%$ in the estimated emission area. If we combine this result with the known current value I_{mid} , using the formula $J = I_{\text{mid}}/A_{\text{est}}$, then (much-tighter) limits on J are obtained, and these in turn yield a smaller uncertainty in A_{est} . The process may be iterated until the result converges, usually after about five steps: The effect is to reduce the uncertainty in the area-estimate A_{est} .

For the standard FN theory, the iteration can be carried out using Table II in Ref. 3, but, in general, the iteration is better implemented via a spreadsheet similar to that discussed in Ref. 2. This enables precise values of Γ to be calculated for the four limiting pairs of (ϕ, J) values, using the emission model embedded in the spreadsheet.

As already indicated, sets of simulated FN-plot data were prepared for three free-electron emission models, using ET, SN, and QE barrier models, taking in each case a “true” emission-area A_{true} of 10^{-14} m^2 . These sets were prepared for each of three work-function values, namely 4 eV, 4.5 eV, and 5 eV.

These data sets were then analyzed using two spreadsheets, one with the elementary FN theory embedded in it, one with the standard FN theory. Each dataset/spreadsheet-analysis combination leads to a mean emission-area estimate A_{est} , and a consequent value of the ratio $A_{\text{est}}/A_{\text{true}}$. Values of these ratios are shown in Table I, for analyses using the “standard” spreadsheet, and in Table II, for analyses using

TABLE II. Mean values of the area ratio $A_{\text{est}}/A_{\text{true}}$, for simulated FN-plot data analyzed using the elementary spreadsheet.

FN-plot data source	$\phi =$		
	4 eV	4.5 eV	5 eV
Elementary data	0.84 $\pm 55\%$	1.06 $\pm 55\%$	1.31 $\pm 55\%$
Standard data	49 $\pm 55\%$ $\pm 3.8\%$	59 $\pm 55\%$ $\pm 4.6\%$	68 $\pm 55\%$ $\pm 4.7\%$
QE data	140 $\pm 55\%$ $\pm 12\%$	215 $\pm 55\%$ $\pm 14\%$	380 $\pm 55\%$ $\pm 18\%$

the “elementary” spreadsheet. [Use of the elementary spreadsheet is, in practice, equivalent to the use of Eq. (11).]

In Tables I and II, two percentage error limits are associated with most ratio values. The first is the residual uncertainty associated with the iteration, after convergence has occurred, and relates to uncertainties in ϕ and J . The second, where given, relates to regression fitting of a straight line to “curved” data, and is the standard percentage error in (RS^2) , evaluated using conventional techniques but ignoring the correlation between the values of R and S .

In the iterations that led to Table I, no attempt was made to optimize the “typical midrange emission current” I_{mid} for the data set being used. I_{mid} is a changeable parameter in the spreadsheets, but was left unchanged at (for historical reasons) a value of $0.192 \mu\text{A}$. In comparison with other effects, results were not expected to be sensitive to the value of I_{mid} (but, see below).

In Table II, the spreadsheet uncertainty derives solely from the arbitrary assumption made about the uncertainty in a work function [$3.5 \leq (\phi/\text{eV}) \leq 6$], which is the same in all cases.

B. Results and discussion

The central result that stands out from Tables I and II is that, if there is a large difference between the emission model used in the simulation and the emission model used in the emission area extraction procedure, then the ratio $A_{\text{est}}/A_{\text{true}}$ can be significantly different from unity and, consequently, there may be a serious error in the estimate of emission area. In this context, the SN and QE barrier models are relatively close to each other, and both are significantly different from the ET barrier model. It is seen that the standard spreadsheet yields area ratios near unity for the SN and QE simulated data, but ratios close to 0.01 for the ET simulated data. By contrast, the elementary spreadsheet yields ratios close to unity for the ET simulated data, and ratios of between 50 and 400 for the SN and QE simulated data.

Although the SN and QE barrier models are relatively close, the standard spreadsheet handles the SN data in a slightly more consistent manner than it does the QE data.

Two features of the stated error limits in Table I deserve comment. First, the “regression fitting” error component is noticeably higher for the QE simulated data than for the SN simulated data. This is because the FN plot of the QE simulated data is noticeably more curved than that of the SN simulated data.⁸ This suggests the tentative possibility that, if errors of experimental measurement could be kept very low (which is a big “if”), and if there were no other distorting factors, then something about the nature of the barrier (and/or the size of the emitting object) might be learned from the curvature of an experimental FN plot.

Second, when the SN simulated data is analyzed using the standard spreadsheet, then one might hope that the resulting $A_{\text{est}}/A_{\text{true}}$ should be very close to 1. In fact, derived ratios are around 0.5. The combined error limit (due to both sources of error) is, however, typically about $\pm 7\%$. So, in statistical terms, the derived ratios are significantly different from 1.0

TABLE III. To show how the mean estimated emission area A_{est} , and the area ratio $A_{\text{est}}/A_{\text{true}}$, vary with I_{mid} , for data generated using the standard FN theory and analyzed using the standard spreadsheet.

I_{mid} (A)	A_{est} (m^2)	$A_{\text{est}}/A_{\text{true}}$
1.92×10^{-7}	4.9×10^{-15}	0.49
1.92×10^{-6}	5.6×10^{-15}	0.56
1.92×10^{-5}	6.4×10^{-15}	0.64
1.92×10^{-4}	7.5×10^{-15}	0.75
1.92×10^{-3}	9.0×10^{-15}	0.90
1.92×10^{-2}	1.14×10^{-14}	1.14
1.92×10^{-1}	1.53×10^{-14}	1.53

(if the method of error estimation is adequate). This implies that the procedures described here are not completely self-consistent, and suggests that some relatively small discrepancy remains to be determined.

Finally, we note that, in all cases, the ratio $A_{\text{est}}/A_{\text{true}}$ increases as the assumed work-function value increases. The reason for this is not currently understood, but might be associated with the effect discussed next.

V. INFLUENCE OF THE VALUE OF I_{mid}

When analyzing SN simulated data using the standard spreadsheet, it was noticed by two of us (H. S. S. and N. H.) that the derived value of A_{est} was slightly dependent on the value of I_{mid} . More systematic exploration yielded results such as those in Table III, which relate to the $\phi=4.5 \text{ eV}$ data set. Table III shows that, as I_{mid} increases from $1.92 \times 10^{-7} \text{ A}$ to 0.192 A , A_{est} increases from $4.9 \times 10^{-15} \text{ m}^2$ to $1.5 \times 10^{-14} \text{ m}^2$. Obviously, this range contains the “true” emission area 10^{-14} m^2 .

The midrange current for the data set is about $5 \times 10^{-4} \text{ A}$; from Table III, the corresponding emission area is about $8 \times 10^{-15} \text{ m}^2$. The corresponding ratio value $A_{\text{est}}/A_{\text{true}}$ now becomes 0.8, rather than the 0.49 shown in Table I. This suggests that we may have discovered the main cause of the residual inconsistency remarked earlier in Table I. It suggests (not surprisingly) that, for self-consistency in the procedures discussed here, it is necessary to use in the spreadsheet an I_{mid} value corresponding to the actual midrange current of the FN plot.

VI. GENERAL CONCLUSIONS

The main conclusion to be drawn from this work is a strong confirmation that, in the extraction of emission area from experimental FN plots, it is highly important to use an emission model that corresponds well to the actual emission physics of the situation. This is not a new conclusion, but the simulations here do seem to bring out the effect much more clearly than some previous work.

Admittedly, the difference between the “real” barrier encountered in a FE process and the standard SN barrier may be less extreme than the difference in the model (between the standard SN barrier and the ET barrier) that is responsible for the numerical effects discussed here. But, es-

pecially for small emitting features, we do not clearly know how well or poorly the SN barrier models the actual tunneling barrier.

The investigations described here concentrated on “getting the barrier right”, but it is equally important to “get the pre-exponential right”, particularly if a free-electron model was not a good representation of the emitter’s electronic structure—which might well be the case if localized surface states or very small emitting features strongly influence the emission.

Subsidiary conclusions are: (1) When using Eq. (11), the best result is obtained if the Γ value used corresponds to a current density in the midrange of the experimental data, and (2) when using an iterative procedure that relies on converting preliminary estimates of emission area to revised estimates of current density, it is best to use a “typical current” in the midrange of the experimental data. There is scope for further work to determine how best to choose these midrange values.

There also seems a (remote) theoretical possibility that interesting information might be derived from study of the

degree of curvature of FN plots, if well-defined experimental data were available.

ACKNOWLEDGMENTS

This work is based on undergraduate BEng projects carried out by two of the authors (N. H. and H. S. S.).^{8,9} The authors wish to thank Yee Fan Yap and Kenneth Koh for providing the simulated FN-plot data for a QE barrier, as part of their own undergraduate BEng projects.

¹R. G. Forbes and K. L. Jensen, *Ultramicroscopy* **89**, 17 (2001).

²R. G. Forbes, *J. Vac. Sci. Technol. B* **17**, 534 (1999).

³R. G. Forbes, *J. Vac. Sci. Technol. B* **17**, 526 (1999).

⁴E. L. Murphy and R. H. Good, *Phys. Rev.* **102**, 1464 (1956).

⁵M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1965).

⁶R. G. Forbes, *Surf. Interface Anal.* (to be published).

⁷H. J. Jeffreys, *Proc. London Math. Soc.* **23**, 428 (1924).

⁸H. S. Sim, BEng Project Report, University of Surrey, Department of Electronic Engineering, 2003.

⁹N. Hamid, BEng Project Report, University of Surrey, Department of Electronic Engineering, 2003.