RC23493 (W0501-062) January 14, 2005
Electrical Engineering

IBM Research Report

Ni-NiO-Ni Tunnel Junctions for Terahertz and Infrared Detection

Philip C. D. Hobbs, Robert B. Laibowitz*, Frank R. Libsch
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, NY 10598

*Department of Electrical Engineering
Columbia University
New York, NY

Research Division
Almaden - Austin - Beijing - Haifa - India - T. J. Watson - Tokyo - Zurich

Ni-NiO-Ni Tunnel Junctions for Terahertz and Infrared Detection

Philip C. D. Hobbs, Robert B. Laibowitz*, and Frank R. Libsch

IBM T. J. Watson Research Center
PO Box 218,
Yorktown Heights NY 10598

* Department of Electrical Engineering,
Columbia University,
New York, NY

Abstract:

We present complete experimental determinations of the tunnel barrier parameters (two barrier heights, junction area, dielectric constant, and extrinsic series resistance) as a function of temperature for the Ni-NiO-Ni system, showing that when the temperature-invariant parameters are forced to be consistent, good-quality fits are obtained between *I-V* curves and the Simmons equation for this very low-barrier system ($\phi \approx 0.20$ eV). A splitting of ≈10 meV in the barrier heights due to the different processing histories of the upper and lower electrodes is clearly shown. We present a fabrication technique that produces high yields and consistent junction behaviour. We present preliminary results of inelastic tunnelling spectroscopy at 4 K that show a prominent peak at ~59 meV, shifted with respect to the expected transverse optic phonon excitation in bulk NiO but consistent with other surface-sensitive experiments. We discuss the implications of these results for the design of efficient detectors for terahertz and infrared radiation.

Keywords: tunnelling theory, metal-insulator-metal tunnelling, barrier heights, nickel oxide, optical detection, terahertz detection

Introduction:

Antenna-coupled tunnel junction (ACTJ) devices are of significant interest as fast detectors of terahertz and infrared radiation.[1,2] These consist of a metal-insulator-metal (MIM) tunnel junction coupled to a thin-film metal antenna. The Ni-NiO-Ni junction is especially useful due to its very low resistance-area products [as low as 1 $\Omega\cdot(\mu m)^2$]. The *RA* product is closely related to the RC time constant of the junction, and at 1 $\Omega\cdot(\mu m)^2$, the intrinsic time constant is near $3\cdot10^{-14}$ s. Fumeaux et al.[3] have shown that Ni-NiO-Ni junctions coupled to thin-film metal antennas can be used as infrared detectors and frequency mixers in the 10-μm band. More recently, Fumeaux, Boreman and co-workers have extended

similar results to the 3.39-μm region and the visible[4,5], though with very low sensitivity.

In this paper, we present fabrication methods for building repeatable devices with a high yield, as well as characterization data showing I-V curves and tunnelling spectroscopy over the range from 4.2 K to 300 K, with a preliminary extension to 400 K. In addition, we show extracted junction parameters (area, barrier thickness, barrier heights, dielectric constant, and parasitic series resistance) which enable these junctions to be specified with high accuracy. Fitting these parameters requires modifications of the standard Simmons-equation approach.[6]

Infrared and Terahertz Detection

In discussions of optical processes in metals and semiconductors, we typically use an approximate material model consisting of an ensemble of single-electron states. At low frequencies, we use a mean-field theory that takes into account the collective response of the electrons. In the terahertz and infrared regions, both the collective and single-particle pictures become problematical. Nonetheless, certain metals (particularly the free-electron metals Cu, Ag, and Au) still make good antennas at frequencies up to 200 THz, though since their dielectric constants are have large negative real parts and small imaginary parts (so that $k \gg n$, where $\tilde{n}=n+ik$ is the complex refractive index) they are not well described by a normal conductor model. For infrared and terahertz detection, the tunnel junction coupled to a metal antenna can also be regarded as a circuit element. Ni-NiO-Ni junctions have been shown to be fast enough to rectify light,[7] and their rectifying properties are thought to be essentially invariant from DC to visible wavelengths.[8,9] An antenna exhibits a radiation resistance R_A, which forms a voltage divider with the load resistance. If the load is a tunnel junction of differential resistance $r_J=R_A$ and negligible capacitance, the small-signal dc responsivity to an applied optical power P_{opt} is

$$\mathcal{R} \equiv \frac{I_{dc}}{P_{opt}} = \frac{r_J}{4}\frac{\mathrm{d}^2 I}{\mathrm{d}V^2} \qquad \textbf{(1)}$$

One factor of two comes from the voltage divider effect, and the other from the 2! in the Maclaurin series. This ratio is the responsivity, which is the appropriate figure of merit for an optical detector. Because numerical second derivatives are so noisy, it is difficult to measure this accurately, so the best procedure is to fit the *I-V* curve using the Simmons function, and compute the derivatives from the fit.

Junction Parameter Fitting

The fitting problem is to minimize the mean squared discrepancy between a measured *I-V* curve and a curve calculated from the junction parameters (area A, barrier heights ϕ_1 and ϕ_2, barrier thickness s, dielectric constant ε, and extrinsic series resistance R_s).

At $T = 0$, the WKB current density calculation is based on Eq. 3 of (Simmons

1963b)[10]:

$$J = J_s \exp\left[\left(\bar{\phi} + \frac{eV_{21}}{2}\right)\exp\left(-A\sqrt{\bar{\phi} + \frac{eV_{21}}{2}}\right) - \left(\bar{\phi} - \frac{eV_{21}}{2}\right)\exp\left(-A\sqrt{\bar{\phi} - \frac{eV_{21}}{2}}\right)\right] \qquad \textbf{(2)}$$

where $V_{21} \equiv V_2 - V_1$ is the junction bias voltage measured from side 2 to side 1, and J_s and A are given by

$$J_s = \frac{e}{h(\beta\,\Delta s)^2} \qquad \textbf{(3)}$$

and

$$A = \frac{4\pi\beta\Delta s}{h}\sqrt{2m} \qquad \textbf{(4)}$$

The quantity $\bar{\phi}$ is the mean value of the barrier height between the classical turning points of the motion, and the adjustment factor $\beta \approx 0.98$ is given by (11).

In the appendix, we discuss this further and give a method for evaluating $\bar{\phi}$ efficiently, which makes J computationally inexpensive, except for the evaluation of β. Fortunately, β is such a weak function of the junction parameters that for curve fitting purposes, a value of 0.98 can be used during most of the fit process, with only the last iterations needing the full numerical integration of (2) and (11).

The fit procedure adopted here is as follows. (A program written by one of the authors was used for the calculation.)

1. Initialization
    a. Assign a range to each variable (the N variables to be optimized are given non-zero ranges)
    b. Choose random starting values within each range: N+1 sets for the Nelder-Mead method[11,12], one set for Powell's method

2. Preliminary Iteration
    a. Find the intrinsic junction bias voltage V_{21} from the measured bias by subtracting IR_s, where I is the measured tunnel current. Assuming the residuals to be small, this ensures a self-consistent solution without requiring a sub-iteration to find the R_s correction at each point on each iteration.
    b. Use (2) at each value of V_{21} to find I from the current junction parameters.
    c. Compute the mean square residual.
    d. Generate the next set of parameters according to the optimization method chosen, or exit if converged.

3. Final Iteration:

Using the preliminary values of the parameters, generate new guesses in their neighbourhood and iterate again, but this time using the numerical solution of (11) for β. This makes each iteration much slower, but often only a few iterations are needed, and this polishing is necessary only once—exploration of the parameter space can be done with β=0.98. (In the present work, the main effect of setting β=0.98 is an underestimation of R_s.)

In general, the mean square residual function will exhibit multiple minima, and this tendency becomes worse as the temperature increases, the measurements become noisier, or the accuracy of the current density calculation deteriorates. Unfortunately, there is no guarantee that the global minimum at a given temperature corresponds to the real device parameters—but forcing the geometric parameters to be constant with temperature does lead to reliable results, as will be shown.

Experimental Procedure

Device Fabrication

The devices used here were fabricated by a germanium shadow mask technique[13] similar to that of Jackel et al.[14], which uses a 3-layer structure consisting of a polymethyl methacrylate (PMMA) bottom layer, germanium middle layer, and PMMA top layer. The desired pattern is written on the top layer using electron-beam lithography. Reactive-ion etching is used to transfer the pattern to the Ge layer, and an oxygen ash step then removes the top layer and undercuts the bottom layer, leaving a suspended Ge shadow mask in the form of slots with bridges. Nickel is then put down in two steps, using angled evaporation near room temperature. After the first metal level, the sample is removed from the evaporator and NiO barrier is formed by oxidizing the nickel briefly in an O_2 plasma similar to that used to ash the PMMA. The second Ni layer is then deposited on top, evaporating from a different angle, and the Ge lifted off in hot acetone. Ambient or thermal oxidation of the Ni produces unstable junctions with poor repeatability and low yield, but the O_2 plasma produces repeatability of ± 10% or better in junction resistance, with yields > 80%. Furthermore, it affords excellent control of the junction areas, since the two metal layers are defined by a single layer of lithography and the overlap depends only on the PMMA thickness and the angle of deposition; with a 600-nm PMMA layer, a 5° rotation about a 40° nominal deposition angle shifts the metal edge by less than 100 nm.

Figure 1 shows an electron micrograph of one of the junctions used in this work, and Figure 2 shows a similar junction with the Ge bridge still in place. The sharpness of the shadows is apparent. The geometric area of this junction is 0.40 ± 0.02 $(\mu m)^2$. All junctions were fabricated on the same chip, with a single level of electron-beam lithography, so all of their overlap areas are closely similar. The lumps on the metal were due to nucleation occurring at sites of minor surface damage to the thin SiO_2 underlayer which seems to have occurred during photoresist ashing.

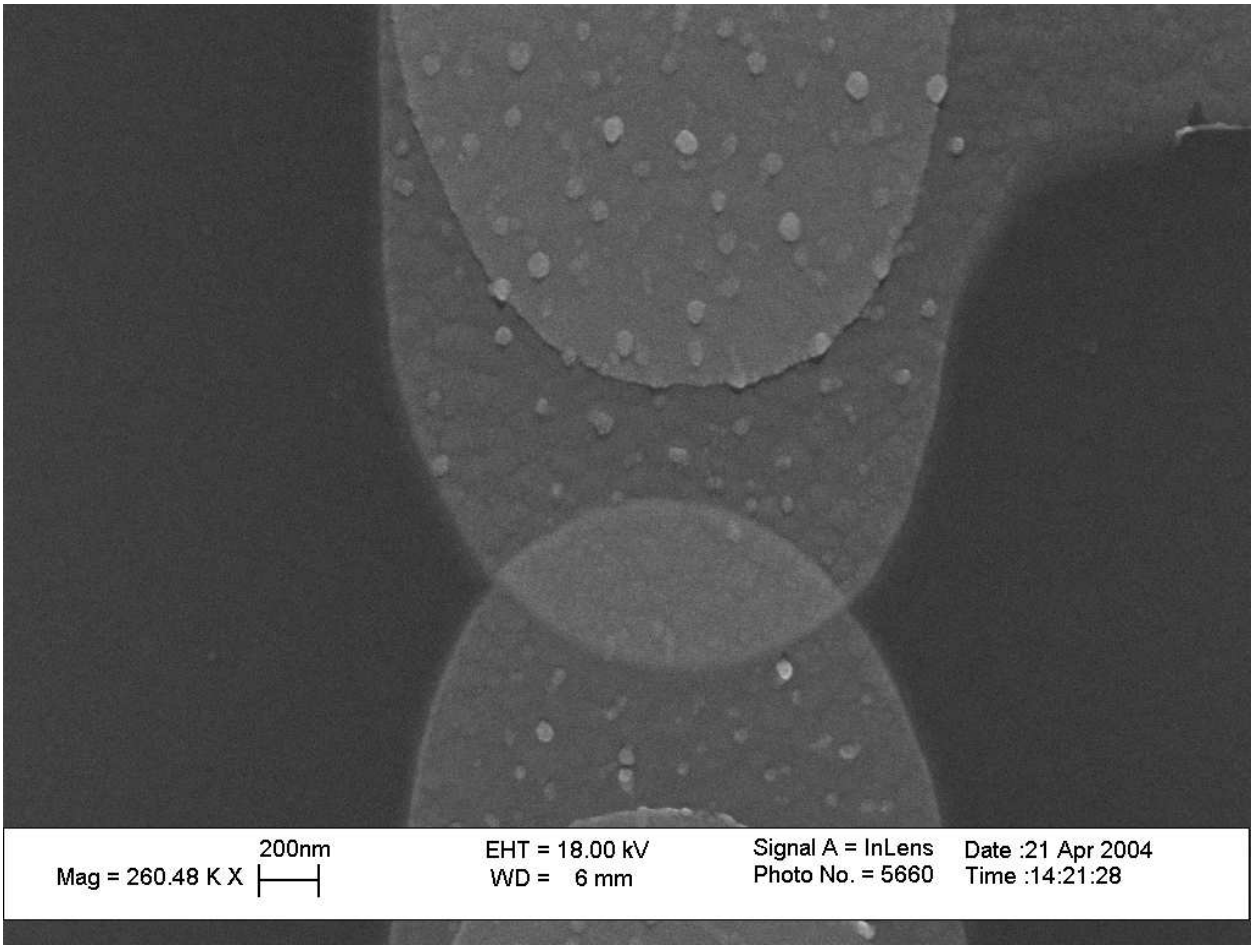

Figure 1: Scanning electron micrograph of a Ni-NiO-Ni tunnel junction fabricated using the Ge shadow mask technique. The junction area is 0.40 ± 0.02 $(\mu m)^2$

I-V Curve Tracing

The device was bonded out and placed in a helium cryostat. Current and voltage were measured using a four-wire apparatus. Figure 3 shows four I-V curves from a single test device over the range of 4 K to 295 K. The nonlinearity is about four times larger at low temperature. Figure 4 shows a typical fitted I-V curve, together with the fit residuals plotted as $10(I_{exp}-I_{fit})$. Apart from very small hooks at the edges (possibly due to Joule heating) the fit residual is limited by measurement noise. In all of these curves, a positive V2-V1 means that the top metal *M*2 is more positive than the bottom metal *M*1. Although the device tested here did not survive long enough to be tested above 300 K, another similar device produced fitted barriers near 0.18 eV at 423 K.

Parameter Extraction

Because most of these data come from a single junction, measured over a wide temperature range, it is appropriate to constrain the fits to make the geometric properties (e.g. the junction area and barrier thickness) constant over *T*. Besides internal logical consistency, this allows an important check: when the geometric parameters are fixed, does the scatter in the other fitted parameters improve or deteriorate?

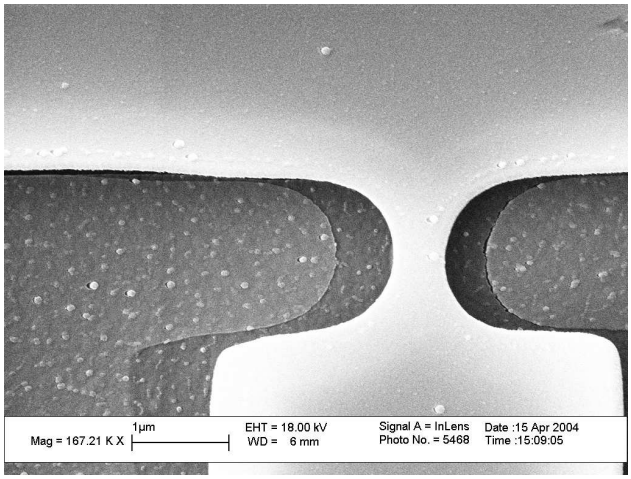

Figure 2: Suspended germanium shadow mask of the type used in this work.

Figure 5 shows a typical fitted experimental *I-V* curve measured at 156.5 K. In (a) the experimental and fit curves are plotted together; (b) is the differential resistance $r_J = dV/dI$; (c) is the second derivative d^2I/dV^2 of the fit function; and (d) is the responsivity calculated from the fit function using (1). Here V_{21} is the voltage measured from M2 to M1 as before. The data shown are from two runs, which lie on top of each other to within the noise level. The responsivity is quite respectable for an infrared detector, even though the curves appear nearly straight.

Figure 6 shows a summary of fitted data from several temperatures, using Nelder-Mead downhill simplex optimization on all the junction parameters (junction area *A*, oxide dielectric constant *K*, oxide thickness *s*, barrier heights ϕ_1, ϕ_2, and extrinsic series resistance R_s). Most of the points lie on well defined trajectories, but with significant scatter and a few outliers. The barrier heights show no consistent trend except that ϕ_2 is always less than ϕ_1. Geometric parameters are almost temperature-independent, and the dielectric constant of insulating metal oxides is typically a weak function of temperature as well. None of these shows any clear trend. Fixing the values A=0.3 (µm)^2, s=2.5 nm, and K=8.4 produced the improved fits of Figure 7. The outliers have disappeared, the scatter around the trend curves has decreased significantly, and the fit residuals have not increased, all of which supports the hypothesis that these parameters represent the true physical parameters of the junction. Interestingly, the fitted area in Figure 6(e) is about 75% of the geometric area, suggesting that the tunnelling is occurring rather uniformly over the overlap area. Aluminum oxide junctions typically behave as though their active areas were a small fraction of the geometric area, suggesting that the tunnel current is dominated by thin spots in the barrier.[15] This is physically reasonable, because the wave function dies off very much more steeply in the junction region with typical Al_2O_3 barrier heights of a few eV than with the 0.2 eV Ni-NiO-Ni barrier. This effect may also be partly responsible for the improved consistency of these results compared with high-barrier fits.

The metal resistance curve Figure 7(b) is also reasonable. Comparing it with the scaled resistance of a nearby metal line (also shown), R_s rises more steeply with *T*, looking

much more like the tabulated curve for bulk Ni[16] plus an offset, as shown. More work is needed to try to determine whether the steeply rising part of R_s is a real metal resistance or a fit artifact, due e.g. to the effects of finite temperature on the tunnelling probability, which are not considered here.

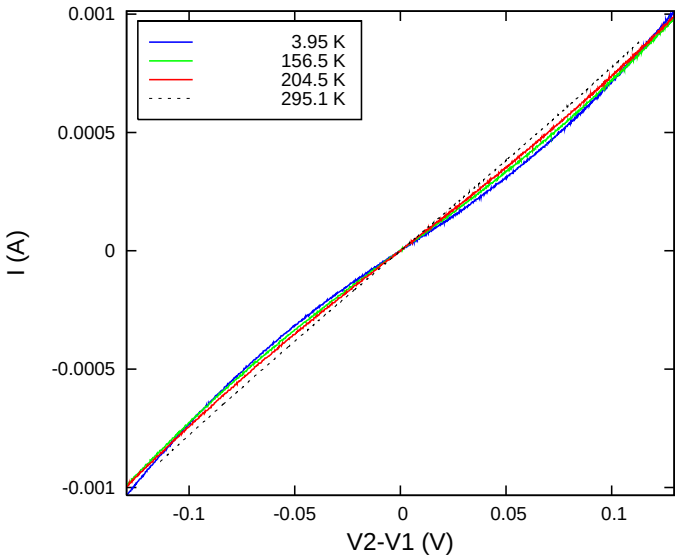

Figure 3: Some of the I-V curves for a junction similar to that in Figure 1, at various temperatures. As the temperature rises, the curves get straighter.

Barrier Asymmetry

Since the top and bottom metal layers are nominally identical, one might expect the junctions to be highly symmetric electrically, but Figure 7(a) shows that this is not so; the barrier heights differ by as much as 20 meV. We conjecture that this is due to a composition gradient in the lower Ni-NiO interface owing to interdiffusion in the oxidation step. This interdiffusion is not present in the upper interface because there is no oxygen plasma bombardment of the upper layer. Apart from the small asymmetry in the *I-V* curve that results, the barrier height inequality is not significant for detector operation, but this mechanism may provide a way to control the barrier heights to optimize responsivity. It is interesting that the lower barrier is on the side with the more abrupt junction. This is consistent with the image potential being stronger there, due to the more nearly ideal metal surface.

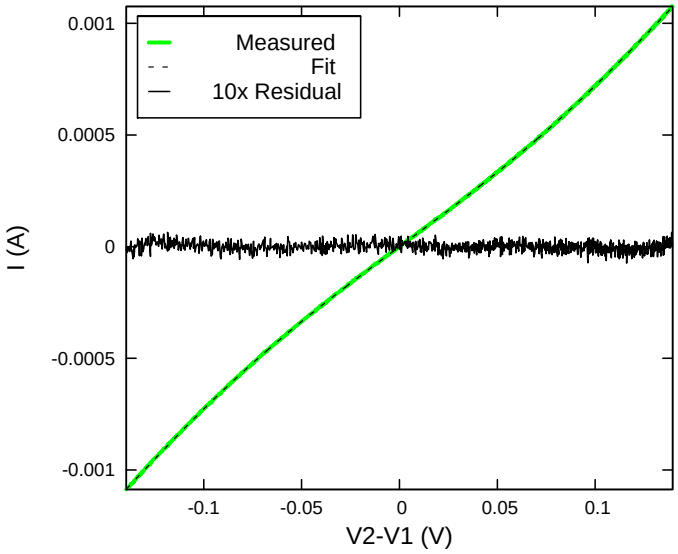

Figure 4: Typical fitted I-V curve with residual $10(I_{exp}-I_{fit})$. This was the 156.5 K curve of Figure 3.

Calculated Detector Responsivity

These devices are potentially good detectors for infrared and terahertz signals. Figure 8 is a plot of responsivity vs bias voltage for the same temperatures as Figure 3, using the extracted junction parameters from the constrained fits. If the barrier heights can be kept near 0.21 eV at room temperature, and the metal resistance reduced, room-temperature responsivity values $R > 1$ A/W may result.

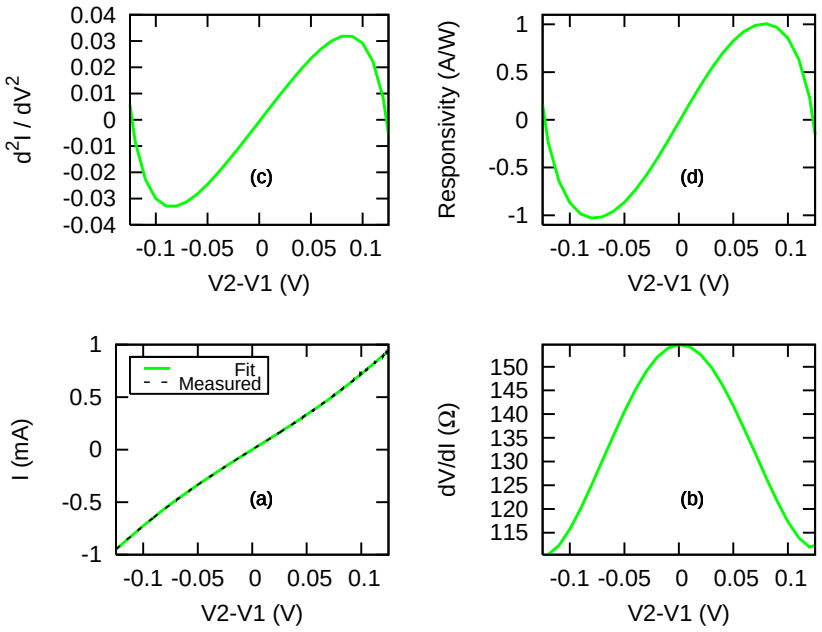

Figure 5: Parameter extraction from *I-V* data: *T*=156.5 K.

Junction Time Constant

Besides the intrinsic speed of the tunnelling itself, the *RC* time constant of the junction may limit its response speed, as is common in semiconductor detectors. Since the fits give a good value for the *K* seen by the tunnelling electrons, we can readily estimate that the capacitance per square metre is

$$C/A \approx \epsilon_0 \frac{K}{s} = 0.03\, F/m^2 \qquad \textbf{(5)}$$

so that at an *RA* product of 1 Ω(μm)^2, the junction time constant is *RC* ≈ 30 fs. This is only a rough estimate, since on these length scales it is not certain that the "plate separation" is the same as *s* in the tunnelling equation. Still, since a 3 dB loss at 200 THz corresponds to a time constant of 0.8 fs, it is clear that efficient ACTJ detectors in the near infrared will have to use reactive or travelling wave techniques to reduce the effects of capacitance.

Inelastic Tunnelling Spectroscopy

In extrapolating from dc measurements to junction properties at 10^{14} Hz, it is important to be certain that tunnelling is the dominant conduction mechanism. The results of earlier sections have shown that this is the case, owing to the very close fits to the Simmons equation, but other lines of evidence are reassuring. Optical detection is one of these, and another is inelastic tunnelling spectroscopy. Nickel oxide has a complicated phonon spectrum, since its unit cell has a small rhombohedral distortion and significantly noncubic behaviour due to magnetic coupling effects, as reported by Chung *et al.*[17] Figure 9 shows a preliminary

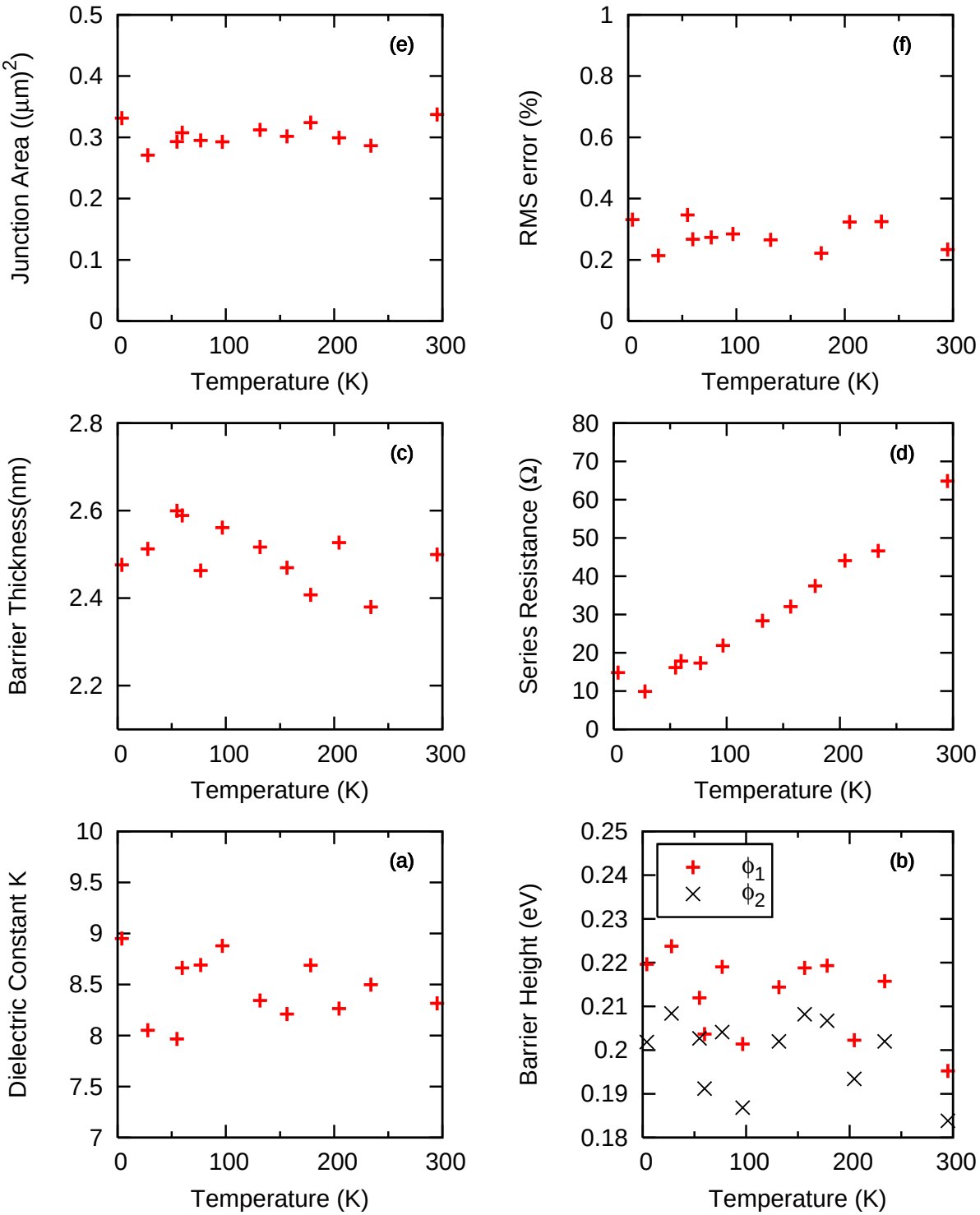

Figure 6: Tunnel junction parameters extracted from the tested junction with all parameters free. Note the general consistency of the barrier thickness (c), area (e), and dielectric constant (a): these show scatter but no clear trend.

inelastic tunnelling spectrum of a junction on the same chip as that used for the parameter extraction. The spectrum shows a pronounced peak near 59 mV with a shoulder on the side nearest 0 V. The peak position differs by some 9 mV from transverse optic (TO) phonon dispersion curves of Chung et al, but interestingly agrees much better with a peak found in room-temperature scanning tunnelling microscope data by Olejniczak and Bieniecki,[18] who find a large peak near 59 mV with the NiO surface negatively biased with respect to the tungsten STM tip.

Discussion

Fit Accuracy

There is a great improvement in the scatter of the fitted barrier heights and series resistances when the consistency conditions on *A*, *s*, and *K* are applied. In general, faulty constraints will increase the fit residuals and the scatter of the fitted values. The low scatter suggests that the values used are close to correct. These values were taken as an approximate average of the free fits of Figure 6, with outliers disregarded. A more systematic procedure would be to use two nested iterations, with the outer loop adjusting the constraints to minimize the total rms residual over all *T* values, and the inner loop optimizing the remaining free parameters for each *T*. Although such a program is not difficult to write, its run time would be prohibitive unless further significant improvements in efficiency were available. In order to achieve full confidence in the fitted results, it would be necessary to perform these fits over a larger set of junctions, measure R_s independently, and seek independent experimental support, e.g. from spin polarization and optical detection. Further work in optical detection using these devices is under way.

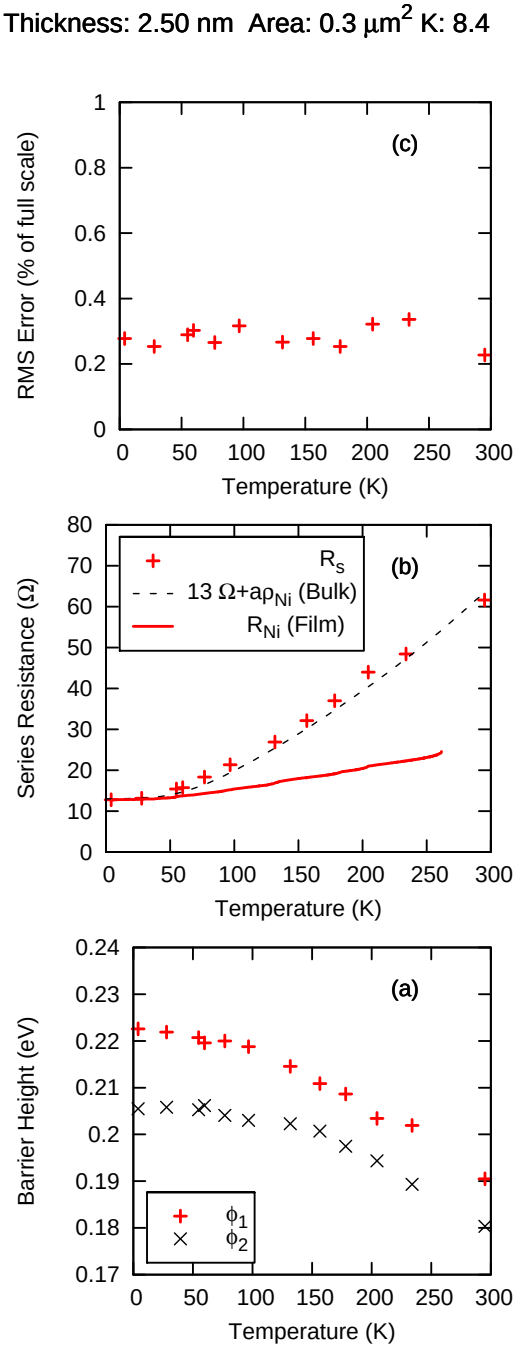

Figure 7: Junction parameters extracted from the same data set as before, but with *A, s, K* fixed. Note the reduced scatter and the good fit of R_s to the resistivity of bulk Ni.

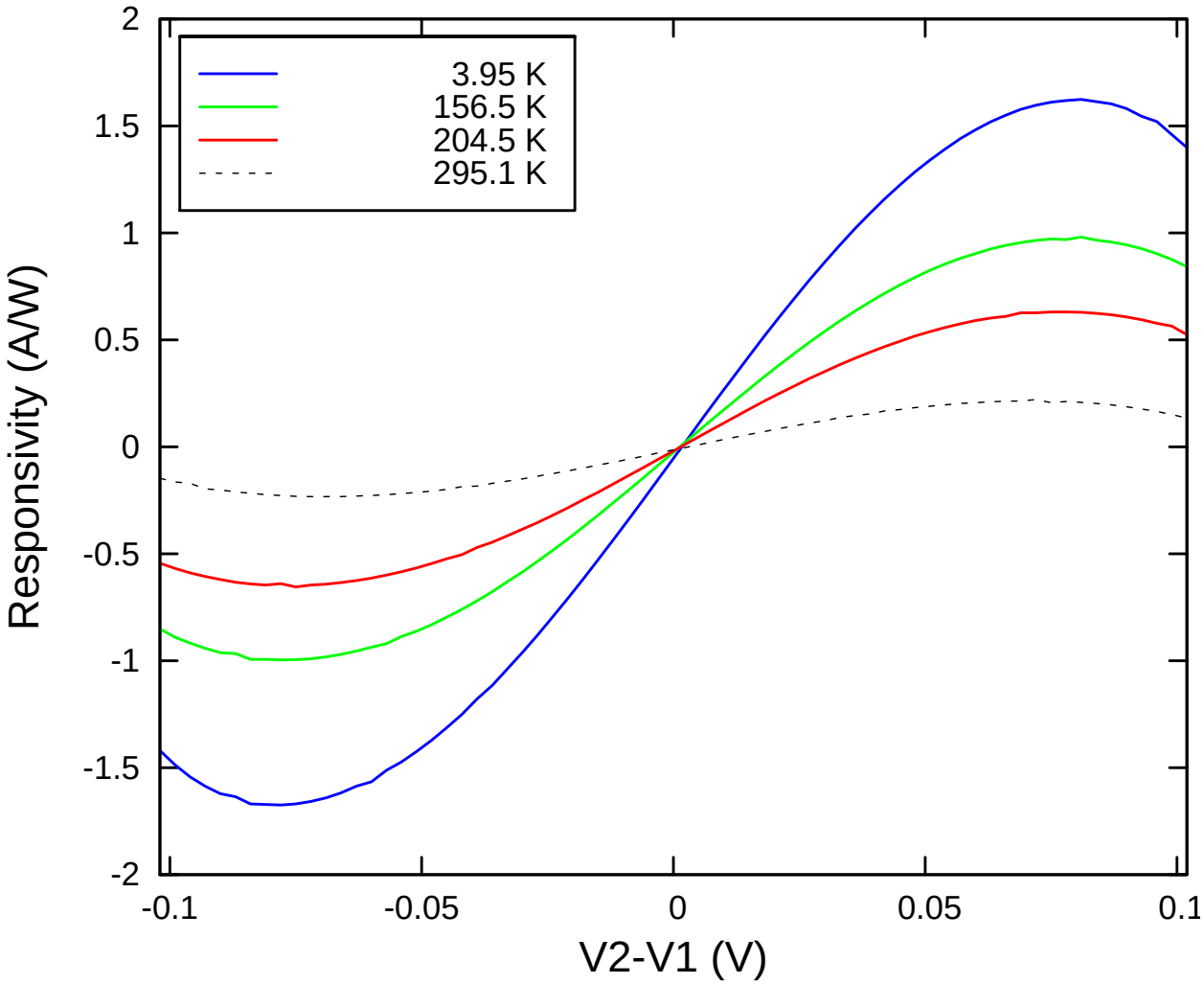

Figure 8: Calculated responsivity of the test junction as a detector vs T. The low responsivity at 298 K (≈0.2 A/W) is due to increased metal resistance and reduced barrier heights, which can both be addressed in device design.

Systematic Errors

The major sources of systematic error in this work are believed to be Joule heating and electromigration. A structure of characteristic size *a* cooled by 3-D conduction into a substrate of thermal conductivity α will have a junction-to-ambient thermal resistance of about

$$\Theta_{JA} \sim \frac{1}{a\alpha} \qquad \textbf{(6)}$$

which for a 1-μm junction on doped silicon is 10^4 K/W at room temperature [α≈100 W/(m·K)[19]] and much less at low temperature, reaching 200 K/W at 20 K [α≈5

kW/(m·K)]. Thus the room-temperature ΔT for a 100-Ω junction with 100 mV of bias is

$$\Delta T \sim \frac{(0.1\text{V})^2}{100\Omega} \cdot (10^4 \text{K/W}) = 1\text{K} \quad (7)$$

corresponding to a resistance change of a few tenths of a percent in the heated region, which is itself a small part of the total extrinsic resistance. The effect is also at least an order of magnitude less at low temperature (T < 100 K) due to the rapid increase of α.

Electromigration is a somewhat more serious problem, potentially leading to the formation of voids and shorts. Since the junctions used here are planar overlaps, the current density is highest in the region outside the junction area. Thus the expected major effect of electromigration would be a systematic change in R_s as the run proceeds. This is not believed to be a significant short-term effect for junction biases below 120 mV, because the *I-V* curves from several runs at a given temperature are identical within the measurement noise. Many of these junctions underwent irreversible breakdown when the bias voltage was above 0.2 V, and their characteristics changed slowly with time at lower voltages. Each of the experimental data sets consists of two *I-V* runs at each temperature, and the room temperature *I-V* curves were closely similar before and after the low temperature runs.

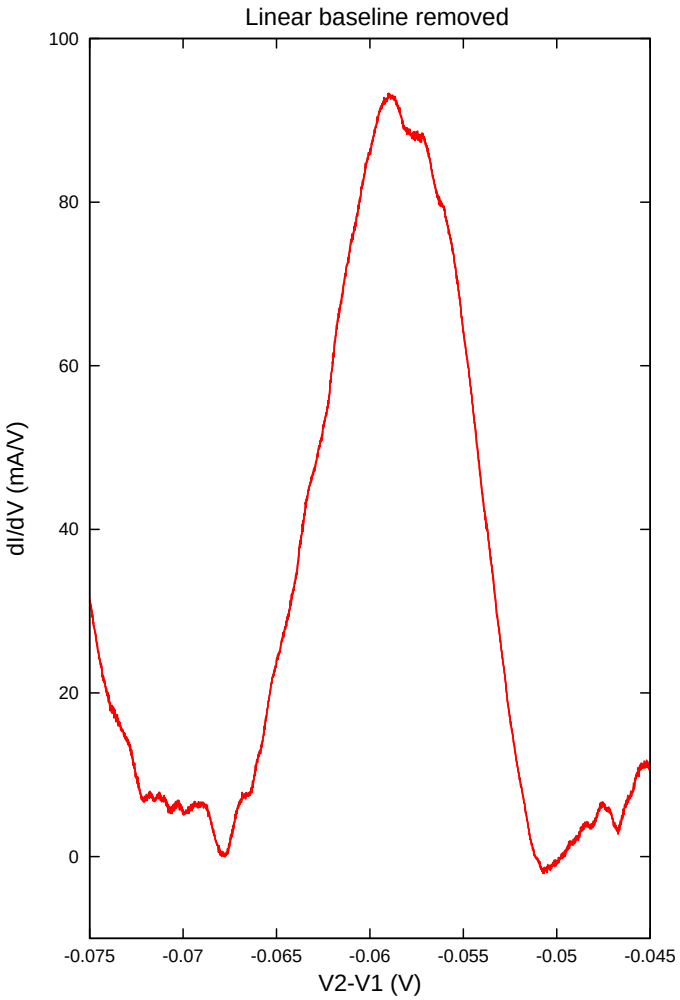

Figure 9 4-K inelastic tunnelling spectra of a sample similar to the one used in the parameter extractions, showing what appear to be the TO phonon peaks of NiO.

Finally, the current density at 1 mA in 0.3 $(\mu m)^2$ is $3 \cdot 10^5$ A/cm^2, and much higher with smaller junctions, so current crowding effects and associated nonlinear resistances cannot be ruled out.

Conclusion: Implications for Optical Detection

Having a repeatable process for making small Ni-NiO-Ni junctions is a prerequisite for making these devices into technologically useful optical detectors. Work is under way to make efficient antenna-coupled tunnel junction detectors in the 1.55-μm optical

communication bands. The major potential advantages of ACTJ detectors at those wavelengths are their very high speed and that they replace III-V devices, requiring only metal, oxide, and dielectric. The low-temperature responsivity of these junctions is much better than that at room temperature, due to metal resistance and to the decline in the effective barrier heights from ~0.22 eV at 4 K to ~0.19 eV at 295 K. It may be possible to tune the composition of the metal and oxide so as to raise the barrier height slightly, and it is certainly possible to reduce the metal resistance by using thicker films, higher-conductivity metals, and mask layouts that reduce the lead length. With these improvements, responsivity values > 1 A/W may well be attainable at room temperature. At 1.55 μm, unit quantum efficiency corresponds to R=1.25 A/W, so it may be possible to make efficient NIR detectors provided the poorly understood details of the rectification process do not preclude it. In order to realize this goal, however, the junction capacitance must be compensated for, either by a travelling-wave design or a reactive matching approach. Further theoretical investigation is needed to elucidate the physics of terahertz and infrared rectification in MIM junctions.

Acknowledgements

The authors wish to thank George A. Keefe, John R. Kirtley, and Roger H. Koch for samples and helpful discussions, and Niranjana Ruiz, S. J. Chey, and Dennis Martinez for device fabrication.

Appendix A: Simmons equations for the very low barrier case

Most tunnel barrier fitting attempts in the literature are based on the two classic 1963 papers of Simmons, treating symmetric[20] and asymmetric[21] barriers. There are two key errors in the second paper, and both papers rely on crude analytical approximations necessary for hand calculation. These approximations are suitable for low-accuracy work when the barrier heights are large, but at the Ni-NiO barrier height of 0.2 eV they lead to significant inaccuracies and even to unphysical results such as large negative currents flowing at positive bias. In addition, since we need to extract several parameters from the fitted curves, it is vital that the fits be as close as possible: since low-resistance junctions have I-V curves that are nearly straight, the fits are very sensitive to poor-quality approximations. Good fits require lots of data points and accurate fit functions, which means a lot of computation. Hence we seek a computationally more efficient algorithm than brute force numerical integration of (2).

The WKB expression for the tunnelling probability of an electron of energy E through a one-dimensional barrier is[22]

$$D(E_x) = \exp\left[-\frac{4\pi}{h}(2m)^{1/2}\int_{s_1}^{s_2}\left(\eta + \phi(x) - E_x\right)^{1/2}dx\right] \qquad \textbf{(8)}$$

where η is the Fermi energy. The square root in the integral is inconvenient to handle. A useful approximation is

Figure 10: 1-dimensional image potential geometry

$$\int_{s1}^{s2} f^{1/2}(x)\,dx \approx \sqrt{\bar{f}}\,\Delta s \tag{9}$$

which leads to an alternative expression for the tunnelling probability

$$D(E_x) = \exp\left(-\frac{4\pi\beta\Delta s}{h}\sqrt{2m\left(\eta + \bar{\phi} - E_x\right)}\right) \tag{10}$$

where $\bar{f}$ is the mean value of the total potential over the interval Δs separating the two classical turning points s_1 and s_2, at which the potential function crosses the Fermi levels of the electrodes. The adjustment factor β is given by

$$\beta = 1 - \frac{1}{8\bar{f}^2\Delta s}\int_{s_1}^{s_2}\left[f(x) - \bar{f}\right]^2 dx \tag{11}$$

where $f(x)$ is the expression inside the square root in the integrand of (8). It is nearly constant and within a few percent of unity. If β is computed exactly, (10) is closely equivalent to (8), but β can also be set constant for more approximate work.

These expressions apply at T=0; this is a good approximation for high barriers, but a much poorer one where $\phi \approx 8$ kT. Significant improvements in accuracy at very low RA products will require computing the tunnel current at finite temperature.

Indefinite Integral of the Image Potential

In the 1-D case, the total potential is the sum of the trapezoidal potential due to the applied bias voltage and the tunnelling barriers, plus the image potential, which is the attraction felt by a charge carrier for a conductive surface. Because of the hall-of-mirrors geometry (see Figure 10), there is an infinite sum over images of images. In the perfect-

conductor limit, the potential V_{img} is given by

$$V_{\text{img}}(x) = \frac{-e^2}{4\pi\epsilon_0 K}\left(\frac{1}{2x} + \sum_{n=1}^{\infty}\frac{ns}{(ns)^2 - x^2} - \frac{1}{ns}\right) \tag{12}$$

where the metal surfaces are at x=0 and x=s, and K is the relative dielectric constant (Simmons 1963a, eq. 32—Simmons 1963b eq. 21 is erroneous). This sum has a rather distressingly slow convergence: to achieve double-precision accuracy (a 48-bit significand), almost 60,000 terms are required. In the expression for the tunnel current, we require the integral of the total potential between the classical turning points of the motion (where the potential crosses the Fermi level). Since these points move as the junction bias is changed, we require the indefinite integral of the image potential,

$$F_{img}(x) = \int^{x} V_{img}(x')\,dx' \tag{13}$$

The integrand is symmetric with respect to s/2, so

$$F_{img}(x) = 2A_{img}\left(\frac{s}{2}\right) - A\left(\frac{s}{2} - x\right) \qquad \left\{x > \frac{s}{2}\right\} \tag{14}$$

For $x \in [0,s/2]$ the sum is uniformly convergent and its terms are regular, so we can decompose it in partial fractions,

$$F_{img}(x) = -\frac{\alpha}{2}\ln x - \frac{\alpha}{2}\left(r\left(\frac{x}{s}\right) - r\left(-\frac{x}{s}\right)\right) \tag{15}$$

where $\alpha = e^2/(4\pi\varepsilon_0 k)$ and r is given by

$$r(y) = \frac{1}{2}\int^{y} dy \sum_{n=1}^{\infty}\left(\frac{1}{n-y} - \frac{1}{n}\right) \tag{16}$$

Taking the binomial expansion of the first term and collecting terms, this becomes

$$r(y) = \frac{1}{2}\int^{y} dy' \sum_{j=1}^{\infty}(-y')^j \sum_{n=1}^{\infty}\left(\frac{1}{n}\right)^{j+1} \tag{17}$$

Expressing the last sum in terms of the Riemann ζ-function, we can rewrite the image potential integral as

$$F_{img}(x) = \left(-\frac{e^2}{4\pi\epsilon_0 k}\right)\left(\frac{1}{2}\ln\left(\frac{x(1+x/s)}{1-x/s}\right) - \frac{x}{s} + \sum_{m=0}^{\infty}\left(\frac{x}{s}\right)^{2m+1}\frac{(\zeta(2m+1)-1)}{2m+1}\right) \quad \left\{x < \frac{s}{2}\right\} \tag{18}$$

The advantage of this more complicated formula over (13) is the very rapid

convergence of the infinite sum. Because $\zeta(m)-1 \sim 2^{-m}$ for large m, the m^{th} term of the sum makes a maximum relative contribution of about $0.5^{4m+1}/m$, so that double precision accuracy needs at most 11 terms.

1. A. Sanchez, C. F. Davis, Jr., K. C. Liu, and A. Javan, "The MOM tunneling diode: Theoretical estimate of its performance at microwave and infrared frequencies" *Journal of Applied Physics* **49**, 5270-5277 (1978)

2. B. Michael Kale "Electron tunneling devices in optics", *Optical Engineering* **24**(2), 267-274 (1985)

3. C. Fumeaux, W. Herrmann, H. Rothuizen, P. De Natale, F.K. Kneubuehl, "Mixing of 30 THz laser radiation with nanometer thin-film Ni-NiO-Ni diodes and integrated bow-tie antennas" *Applied Physics* **B 63,** 135-140 (1996)

4.C. Fumeaux, M. Gritz, I. Codreanu, W. Schaich, F. Gonzalez, G. Boreman, "Measurement of the resonant lengths of infrared dipole antennas,"
Infrared Physics and Technology 41, 271-281 (2000).

5. C. Fumeaux, J. Alda, G. Boreman, "Lithographic antennas at visible frequencies,"
Optics Letters 24, 1629-1631 (1999).

6. Dorneles, L. S., D. M. Schaefer, M. Carara, and L. F. Schelp, "The use of Simmons' equation to quantify the insulating barrier parameters in Al/AlO_x/Al tunnel junctions", Appl. Physics Lett., **82**, 17, pp. 2832-2834 (28 April 2003)

7. J. G. Small, G.M. Elchinger, A. Javan, A. Sanchez, F.J. Bachner, D.L. Smythe
"Ac electron tunneling at infrared frequencies: Thin-film M-O-M diode structure with broad-band characteristics" *Applied Physics Letters* **24**(6), 275-279 (1974)

8. M. Heiblum, S. Wang, J.R. Whinnery, T.K. Gustafson "Characteristics of integrated MOM junctions at dc and at optical frequencies" *IEEE Journal of Quantum Electronics* QE-**14**(3), 159-169 (1978)

9. S. K. Masalmeh, H. K. E. Stadermann, and J. Korving, "Mixing and rectification properties of MIM diodes", *Physica B* **218**, pp. 56-59 (1996)

10.Simmons, John G., "Electric Tunnel Effect between Dissimilar Electrodes Separated by a Thin Insulating Film", J. Appl. Physics, **34**, 9 pp 2581-2590 (September 1963). (Simmons 1963b)

11. Nelder, J. A., and Mead, R., 1965 Computer Journal **7**, 308.

12. Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T., *Numerical Recipes in C.* (Cambridge, 1988), section 10.4

13. George A. Keefe, private communication

14. L. D. Jackel, R. E. Howard, E. L. Hu, D. M. Tennant, and P. Grabbe, "50-nm silicon structures fabricated with trilevel electron beam resist and reactive-ion etching", *Appl Physics Lett* **39**(3) pp. 268-270. August 1, 1981

15. J. D. R. Buchanan, T. P. A. Hase, B. K. Tanner, N. D. Hughes, and R. J. Hicken, *Appl Physics Lett* **81**, 751 (2002)

16. David Lide, ed., *CRC Handbook of Chemistry and Physics, 81st Ed.*, CRC, 2000, p. 12-46

17. E. M. L. Chung, D. McK. Paul, G. Balakrishnan, M. R. Lees, A. Ivanov, and M. Yethiraj, "Role of electronic correlations on the phonon modes of MnO and NiO", *Phys Rev B* **68**, 140406 (2003)

18. W. Olejniczak and M. Bieniecki, "Fine structure in differential conductance of oxidized nickel observed in a room temperature stm experiment", *Solid State Communications*, **101**, 12 pp. 877-882 (1997)

19. Thermal conductivity values are from the CRC Handbook of Chemistry and Physics, 2001 ed.

20. Simmons 1963a

21.Simmons, John G., "Electric Tunnel Effect between Dissimilar Electrodes Separated by a Thin Insulating Film", J. Appl. Physics, **34**, 9 pp 2581-2590 (September 1963). (Simmons 1963b)

22. Simmons, John G., "Generalized Formula for the Electric Tunnel Effect between Similar Electrodes Separated by a Thin Insulating Film", J. Appl. Physics, **34**, 6 pp 1793-1803 (June 1963), eq. 1.