

Catch-22 in Second Law, Part 2 Optimizing MIM Diode, Rectifying Johnson-Nyquist Noise

George Samuel Levy

Entropic Power, Irvine, CA, USA

Email: glevy@entropicpower.com

How to cite this paper: Levy, G.S. (2026) Catch-22 in Second Law, Part 2 Optimizing MIM Diode, Rectifying Johnson-Nyquist Noise. *Journal of Applied Mathematics and Physics*, 14, 4013-4048.
<https://doi.org/10.4236/jamp.2026.1410195>

Received: August 13, 2026

Accepted: October 17, 2026

Published: October 20, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution-NonCommercial International License (CC BY-NC 4.0).
<http://creativecommons.org/licenses/by-nc/4.0/>



Open Access

Abstract

Foremost arguments supporting the second law rely on a false dichotomy between the microscale (μ -Scale), where negentropy can be generated by T -Odd symmetry but is inaccessible due to the absence of thermodynamic interaction, and the macroscale (M -Scale), where massive coarse graining provides interactions but erases the negentropy. Operating at the mesoscale (m -Scale) bypasses this catch-22. Some negentropy is destroyed but the rest is harvested. A MIM diode used as a vehicle in this analysis is inherently spatially asymmetric but becomes T -Odd symmetric when a time-varying signal—such as the Johnson-Nyquist noise—modulates its asymmetry, causing carriers to move preferentially in one direction. An index of asymmetry Z_{asym} quantifying the operating position along m -Scale is defined as the ratio of the free energy generated by the T -Odd symmetry to the energy dissipation due to T -Even symmetry. Power output to the load is maximized at the geometric mean of m -Scale where $Z_{asym} = 1$. Thermal energy converted to electricity sent to the load cannot exceed $1/4$ of the free energy provided by the work function difference at the gap of the diode and cannot exceed $1.0829k_B T$ per carrier. The law is extended to $\Delta S \geq -\Phi/4T$ for m -Scale operation. This result is unopposed by theories and experiments all directed to μ -Scale and M -Scale, and has been previously reproduced twice—in optimizing Maxwell's Demon and the $\mathbf{E} \times \mathbf{B}$ Thermoelectric Effect. A MIM diode rectifying Johnson-Nyquist noise at 300 K generates up to 14.43 mW/mm^2 or nearly 42.8 times the power per unit area of a solar cell.

Keywords

Thermodynamic Threshold Optimization, Entropy, Time-Reversal Symmetry, Mesoscale, MIM Diode, MIIM Diode, Johnson-Nyquist, ExB Thermoelectric Effect, Second Law, Free Energy

You can't have your cake and eat it too... but you can eat half your cake and still have the other half.

1. Introduction

This is the second of a two-part paper. The first paper [1] presented an axiomatic analysis probing the foundational assumptions of the law and its applicability at the mesoscale (*m-Scale*), the borderline region between:

1. The microscale (*μ -Scale*) where fine entropy remains constant in time-reversal (*T-Even*) symmetric phenomena or may even decrease in time-reversal asymmetric (*T-Odd*) phenomena, and
2. The macroscale (*M-Scale*) where coarse entropy always increases due to overwhelming averaging or integrating that destroys microscopic information.

A previous publication [2] derived the inapplicability of the second law at the mesoscale by using Maxwell's demon as a vehicle. Part 1 of this paper [1] uses the $\mathbf{E} \times \mathbf{B}$ Thermoelectric Effect. The present paper uses a MIM diode. All three papers reach the same conclusions using different analytical methods. The following topics are covered:

- **False Dichotomy and Catch-22 in the Second Law**—All arguments supporting the law scrutinize systems at the microscale or the macroscale. No theory or experiment provides any evidence for the law at the mesoscale.
- ***T-Odd* Symmetry of Rectifiers**—All such devices are inherently configured as spatially asymmetric. However, they acquire their *T-Odd* symmetry when their spatial configuration is modulated by a time varying signal.
- **MIM-Class Diodes** are used as a vehicle for this analysis because of their outstanding bandwidth and sensitivity.
- **Thermodynamic cycle.** The operation of a diode is discussed as a thermodynamic cycle. The concept of configuration energy, entropic drive and free compression are introduced.
- **The Index of Asymmetry.** This index quantifies the concept of *m-Scale* and allows the precise determination of design parameters such as the work function difference in a MIM diode.
- **Optimizing Z_{asym} .** Three methods show that the optimal value for the index is $Z_{asym} = 1$ exactly at the geometrical mean of *m-Scale*. At that point, the thermal energy converted to the electrical energy sent to the load cannot exceed $\frac{1}{4}$ of the free energy provided by the diode in the form of a difference Φ in work function. **The change in entropy is limited to $\Delta S \geq -\Phi/4T$.** This result is congruent with the two earlier papers.
- **The Johnson-Nyquist Noise.** The peak spectral frequency of the noise is calculated and found to coincide with Stefan-Boltzmann's law and Planck's radiation law.
- **Dimension of MIM Diode.** A large area MIM diode spontaneously subdivides itself into small, correlated areas, each rectifying independently. No explicit lithographic division is needed.
- **Matching the Threshold to the Noise.** The conduction threshold of the diode

is defined to maximize the energy harvested from plasmons. The energy sent to the load cannot exceed 1/2 of the energy generated at the threshold.

- **Absolute Values of the Work Functions.** This section covers diode-specific parameters including optimizing the absolute value of work functions.
- **Deviation from Optimality Due to Multiple Quanta Threshold Interaction.** This section briefly discusses the impact on performance of multiple quanta combining at the diode's threshold.
- **Appendices A and B** discuss the isothermality of Maxwellian and Fermi gases in equilibrium in a potential field.
- **Appendix C** provides a bird's-eye view of the model required for rectifying very small amplitude and high frequency signals such as the Johnson-Nyquist noise.
- **Appendix D** describes the thermodynamic cycle of a MIM diode.
- **Appendix E** is a list of symbols used in this paper.

2. False Dichotomy and Catch-22 in the Second Law

This paper is a continuation of the “axiomatic analysis” study [1] aimed at probing the cornerstone logic supporting the foremost second law arguments. This analysis found that, without exception, all arguments supporting the second law consider two points of view:

1. The μ -Scale where collisions are absent. In systems with T -Even symmetry, fine entropy remains constant $dS_{fine}/dt = 0$; in those with T -Odd symmetry, fine entropy can decrease $dS_{fine}/dt < 0$.
2. The M -Scale where thermodynamic interactions such as collisions are present. From a macro-observer's point of view, massive averaging, integration and “coarse-graining” destroys microscale information and ensures that coarse-grained observed entropy always increases $dS_{coarse}/dt > 0$.

These two points of view present a catch-22:

1. At μ -Scale, systems with T -Odd symmetry can produce negentropy, but this negentropy is inaccessible because of a lack of thermodynamic interaction. With no input-output mechanism, no negentropy can be extracted.
2. Conversely, at M -Scale, massive thermodynamic interactions do provide an input/output capability, but they also destroy any μ -Scale negentropy, leaving none to extract.

This outlook presents a false dichotomy—a false choice between two extremes. Current arguments supporting the second law can be stated in colloquial language as “*You can't have your cake and eat it too.*” In the same colloquial language, this paper suggests “*You can eat half your cake and still have the other half.*” However, it is best eaten at the mesoscale (m -Scale), in half portions, between μ -Scale and M -Scale.¹

This framework is in full compliance with the conventional second law as it

¹This statement is an accurate reflection of the conclusion reached in this paper that the thermal energy converted to electricity by a “loaded” diode cannot exceed **half** the free energy defined by the voltage across the gap in the “loaded” state.

guarantees that negentropy is inaccessible at the endpoint (μ -Scale, M -Scale) of m -Scale **as proven by all H-Theorems and as confirmed by all experimental data** [1]. Conversely, no theoretical argument or experimental data forbids the existence and accessibility of negentropy at m -Scale. This paper applies these ideas to diodes as formerly presented in the optimization of Maxwell's Demon [2] and the $\mathbf{E} \times \mathbf{B}$ Thermoelectric Effect [1].

3. T -Odd Symmetry of Rectifiers

Diodes and fluidic valves are configured **at fabrication time as spatially asymmetric, not time asymmetric**. Therefore, at first glance, these devices should conform to all H-theorems and should not rectify anything. Being strictly spatially asymmetric, their time symmetry demands that, at equilibrium, *any transition in one direction must be exactly counterbalanced* by a transition in the opposite direction, making isothermal rectification of the Johnson-Nyquist noise an impossible task.

Upon closer examination, however, one realizes that *a diode becomes time asymmetric when it is exposed to a time varying signal that switches polarity with time*. The signal produces physical changes that alter its electronic configuration and conductivity, as the diode goes back and forth between forward mode and reverse mode.

1. In a PN diode, the thickness of the depletion layer changes with the direction of the bias: the layer is thin in forward mode and thick in reverse mode which leads to a change in resistance and to rectification.

2. Similarly, in a one-way fluidic check valve, a disk or piston moves, opening or closing the valve in response to fluid flow.

This dynamic change in spatial asymmetry is crucial in producing a T -Odd symmetry. Furthermore, it is associated with a change in a state variable, it is not instantaneous, and must use energy *borrowed from the signal*. This idea is obvious. **Yet it confirms that diodes and valves are not bound by the H-theorem.** Interestingly, very few experiments involving diodes have succeeded in producing *observable* rectification of thermal noise.

Nikulov [3] and Gurtovoi *et al.* [4] observed a DC voltage on an asymmetric microscopic superconducting ring when the ring or its segments are switched between superconducting and normal state by non-equilibrium noises (*i.e.*, an AC signal flowing through a spatially asymmetric device.) This voltage indicates that one of the ring segments operates as a DC power source.

Kondo *et al.* [5] developed a new organic thermoelectric device constructed as a MIM diode, capable of functioning at room temperature without any temperature gradient. Their device is driven by the Johnson-Nyquist noise. It produces an open-circuit voltage of 384 mV, a short-circuit current density of 1.1 $\mu\text{A}/\text{cm}^2$, and maximum output P_{max} of 94 nW/cm².²

²The power output (94 nW/cm²) of Kondo's device is far below what could be achieved. The device is not Z_{asym} optimized as recommended in this paper, its Φ is far from the optimum 111.98 meV, and its bandwidth is not matched to the Johnson-Nyquist noise.

Given the scarcity of experimental evidence, one could legitimately ask, “Why can’t PN diodes routinely produce an isothermal thermoelectric effect by rectifying thermally induced electrical Johnson-Nyquist noise?” In other words, why can’t diodes produce a DC output isothermally without requiring a temperature difference as in the Seebeck effect?

Even though these devices are time-reversal asymmetric while in operation, *they are enormously mismatched in energy and frequency to the noise*. The built-in potential of a silicon device is about 0.7V, about fifty times larger than the $\frac{1}{2}k_B T$ energy of the noise at 12.9meV, and the cutoff frequency of a diode is about 1MHz, seven orders of magnitude lower than the peak frequency of the noise at 18THz.

Accordingly, *two conditions must be met for a device to bypass the second law*:

1. Asymmetric Condition: The device must be *T-Odd* symmetric as required by the H-theorems. This condition is necessary to make bypassing the law *theoretically possible*. A diode meets this requirement in operation when a signal renders it *T-Odd* symmetric by modulating its spatial asymmetry.

2. Matching Condition: The device must be matched in energy and bandwidth to the noise being rectified. This condition is necessary to make their output *experimentally observable and economically useful*. A diode can meet this condition by operating at *m-Scale*, a *thermodynamic regime mostly unexplored both theoretically and experimentally*.

As shall be discussed in **Sections 6 and 7**, an index of asymmetry, Z_{asym} shall be defined that quantifies the position of a MIM diode at *m-Scale*. **Section 10** optimizes the diode to the Johnson-Nyquist noise.

4. MIM-Class Diode Model as Vehicle for Analysis

The first step in this analysis is to select a diode architecture with sensitivity and bandwidth matching the signal to be rectified.

4.1. MIM Frequency Response and Sensitivity

Figure 1 is a scatter diagram illustrating the capabilities of several diodes as a function of signal energy versus frequency [6].

4.2. MIM Models and Reconfiguration Energy

Given the suitability of MIM-class diodes for high frequency, high sensitivity rectification, this kind of diode has been selected as a vehicle for the following analysis with the goal of optimally rectifying the Johnson-Nyquist noise.

The emergence of time-reversal asymmetry is due to a dynamic change in configuration associated with a *state variable* (thickness of depletion zone, disk position) modulated by the polarity and amplitude of the input signal. This state variable affects the resistance and bandwidth of the device. *A change in this state variable requires a reconfiguration energy provided by the signal and is limited by the bandwidth of the device.*

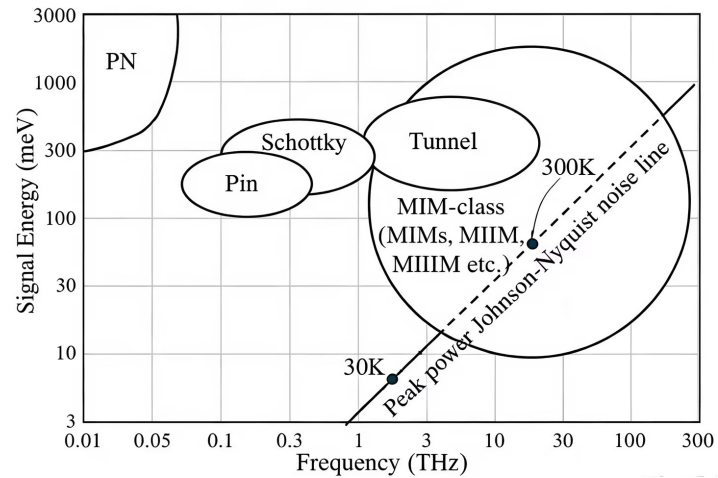


Figure 1. Performance map for several diode architectures. The diagonal line maps the peak spectral energy of the Johnson-Nyquist noise: $E_{Peak} = hf_{Peak} = n_{Peak}k_B T$, where the peak power frequencies and signal levels are located as a function of temperature. MIM-class diodes with one or more insulating layers, overlap this high frequency, low amplitude region. They include PN, Pin, Schottky, Tunnel, and MIM-class diodes (which include MIMs, MIIMs, MIIIMs and other multilayers variants). The line at the bottom right corner represents the frequency range and signal level of the Johnson-Nyquist noise as described by equation $E_{Peak} = hf_{Peak}$ where E_{Peak} is the energy of a plasmon at the peak spectrum frequency f_{Peak} (This topic is discussed in **Section 8**). As the diagram shows, the MIM-class diodes are the only ones capable of operating in that region. They stand out for their performance at terahertz and even optical frequencies.

MIM diode models by Tsu-Esaki [7], Simmons [8] [9], Grover [10] and others [11]-[13] do include a state variable in the form of the changing Fermi distribution across the gap, **but they do not account for the configuration energy and bandwidth** required to change this distribution. This omission renders them *instantaneous and inappropriate for modeling the rectification of Johnson-Nyquist fluctuations*. As far as this author knows, no diode model exists which is suitable for thermal noise rectification. It should include:

1. A *reconfiguration energy* φ_{Config} supplied by the signal, to modify the state variable.
2. A *time constant* linked to the change in the state variable.

For the sake of simplicity this paper discusses the simplest MIM diode which includes a single insulating layer. In such a diode, the state variable is associated with the movement of charges that affect the occupancy and vacancy distribution on either side of the gap and change its effective width (in analogy to the change in thickness of the depletion zone in a PN diode). **Figure 2** shows a MIM diode fabricated with:

1. On the left: a high density of occupied states and low density of vacant states, and
2. On the right: a low density of occupied states and high density of vacant states.

The effective width of the gap varies depending on the polarity of the applied

voltage V :

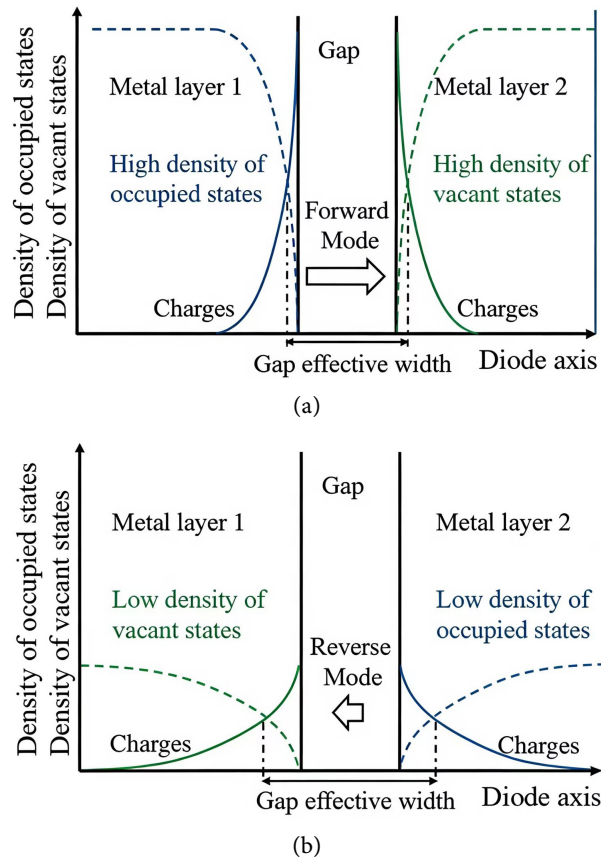


Figure 2. The sides of the gap can be designed to have high occupancy and low vacancy on the left, and low occupancy and high vacancy on the right. a) In forward mode, the charge layers on the surfaces are thin, causing the gap to have a thin effective width. b) In reverse mode they are thick, causing the gap to have a thick effective width.

1. A positive voltage $+V$ shifts the Fermi levels causing the diode to operate in forward mode with a high occupancy on the left and high vacancy on the right. Therefore, a carrier transiting across the gap, from left to right, originates near the surface of the metal on the left and ends up near the surface on the right. As a result, the gap has a narrow effective width, and its probability of transmission is high.

2. Conversely, a negative voltage $-V$ causes the diode to operate in reverse mode with a low occupancy on the right and low vacancy on the left. Therefore, a carrier transiting from right to left originates far from the surface on the right, and it ends up far from the surface on the left. The gap has a large effective width, and its probability of transmission is low.

Accordingly, tunneling requires traversing three regions: a left one of width d_L , the central gap of width d_G and a right one of width d_R . Therefore, the effective expected gap size is $d_{eff} = d_L + d_G + d_R$. Since $d_L + d_R$ is smaller in forward bias, d_{eff} is also smaller in forward bias. A sub-nanometer change in a one-nanometer gap

results in a large difference in its conductivity. Full integration of this concept into existing models is the subject of a future paper. It is briefly discussed in **Appendix C**.

5. Thermodynamic Cycle of Diode

5.1. Entropic Drive

The change in the configuration of the electronic states is ***T-Odd* symmetric because it is synchronized with the input signal**. It always favors a forward current over a reverse current. Unlike the Seebeck effect which requires a temperature difference, or Ohmic conduction which requires a voltage, this driving mechanism is **purely entropic**, caused by the synchronicity between the signal and the dynamic configuration of the states. At issue is not whether a diode can rectify a signal, it is how small this signal can be. This underlines the importance of the configuration energy φ_{Config} : it is commensurate with the signal to be rectified and must be included in the following analysis.

5.2. Free Compression across the Gap

It is assumed that carriers tunnel through the nanometer gap much faster than the average period of a Johnson-Nyquist fluctuation and that, **despite the fluctuations, they remain in local equilibrium**. When a fluctuation produces a forward voltage bias, carriers maintain their equilibrium by tunneling (reversibly) up the energy barrier across the gap. This process is both isothermal and adiabatic:

1. Carriers remain isothermal in compliance with the second law (a Fermi gas in equilibrium in a field is isothermal) and with the proof of the isothermality of a Fermi gas in **Appendix B**.

2. Carriers remain adiabatic because direct tunneling implies no collision that changes the carrier's energy.

Conventional thermodynamics admits one process which is both isothermal and adiabatic. It is called a free expansion. A gas freely expanding in a chamber retains its temperature and internal energy but increases its volumetric entropy.

In contrast to a free expansion, carriers crossing the gap **lower their positional entropy**, while remaining isothermal and adiabatic. Therefore, their change in position is the exact reverse of the free expansion. Let us call it "free compression." The process is not driven by a temperature difference as in the Seebeck effect, yet it does lower the temperature **equally on both sides** of the junction:

1. The emitting side of the gap drops in temperature because it loses the most energetic carriers which are more prone to tunneling.

2. The receiving side also drops in temperature because the incoming carriers lose kinetic energy as they tunnel adiabatically against the barrier potential.

Overall, both sides of the gap remain in isothermal equilibrium with respect to each other because tunneling affects them equally. However, the junction, as a whole, experiences a drop in temperature with respect to its surroundings. This temperature difference causes an inflow of heat that replenishes the thermal en-

ergy converted to electricity. This temperature difference and heat flow is a Fourier heat conduction effect and not a Carnot thermodynamic requirement. There is no need for a heat sink, but the device must still be colder than the surroundings.

In summary, the thermodynamic cycle includes a free (isothermal) compression of the carriers in the diode, followed by an isothermal expansion in the load. The work function difference Φ defines the *T-Odd* asymmetry of the diode and provides it with the free energy used in rectifying heat to electrical energy. The potential across the gap acts as a catalyst in the sense that it is not consumed in the heat to work conversion process.

6. The Index of Asymmetry as a Measure of *m-Scale*

The catch-22 dilemma posed by the second law can be summarized as such: At the μ -Scale, *T-Odd* systems can generate negentropy, but it is not accessible because of the dearth of thermodynamic interactions such as collisions. At *M-Scale*, these interactions are abundant, but they destroy any negentropy before it can be accessed. The theory advanced in [1] [2] and in this paper, is that these two processes compete with each other. The first generates negentropy but does not provide the means to access it, the second allows it to be accessed but destroys it. At *m-Scale*, interactions are moderate enough to leave some negentropy on the table, yet large enough to provide the input/output means for its harvesting.

The basic generator of free energy (and negentropy) in a MIM diode is the asymmetry produced by the difference Φ in the work functions of the metal plates across the gap. The generative and dissipative processes both compete for Φ . At μ -Scale, generation uses up all Φ and there is nothing left for the dissipative (through collisions) export of negentropy. At *M-Scale*, the reverse happens, negentropy can be accessed but there is none left to be exported. These two end points are the domain of the second law. The region in between, *m-Scale*, is off-limit to the law.

An index of asymmetry Z_{asym} is defined to quantify the position of a system along *m-Scale*. Z_{asym} is a dimensionless ratio that comprises a generative factor and a dissipative factor. It is infinite at μ -Scale and zero at *M-Scale*.

The Z_{asym} ratio includes:

1. The numerator, $\Phi - eV_{Load} - \phi_{Config} - \phi_{Resist}$, is the generative factor. It represents the *T-Odd* asymmetry of the device quantified by the **actual potential step** at the loaded gap. In an open-circuit unloaded state, the potential across the gap is Φ . However, when the diode is connected to a load, the rectified Johnson-Nyquist fluctuations produce a current through the diode and the load. The resulting load voltage, eV_{Load} , the configuration potential energy, ϕ_{Config} , and the internal resistance of the diode ϕ_{Resist} oppose this current and result in a reduction of the gap potential from Φ to $\Phi - eV_{Load} - \phi_{Config} - \phi_{Resist}$. Therefore, the actual free energy available at the gap is:

$$\text{Generated Free Energy} = \Phi - eV_{Load} - \phi_{Config} - \phi_{Resist} \quad (1)$$

2. The denominator, $eV_{Load} + \varphi_{Config} + \varphi_{Resist}$ is the dissipative factor. It represents the *T-Even* symmetry of the dissipative processes which includes: eV_{Load} dissipated in the load, φ_{Config} used to configure the forward/reverse electronic states of the diode, and φ_{Resist} dissipated by the internal resistance of the diode:

$$\text{Dissipated Free Energy} = eV_{Load} + \varphi_{Config} + \varphi_{Resist} \quad (2)$$

As originally defined in [1], Z_{asym} is the ratio of root-power quantities (such as v_x and v_y whose squares are directly proportional to power). Therefore, the **square** of Z_{asym} can be stated as the ratio of the energies associated with the generative and dissipative processes:

$$Z_{asym}^2 = \frac{\text{Generated Free Energy}}{\text{Dissipated Free Energy}} = \frac{\Phi - eV_{Load} - \varphi_{Config} - \varphi_{Resist}}{eV_{Load} + \varphi_{Config} + \varphi_{Resist}} \quad (3)$$

One should note that eV_{Load} , φ_{Config} and φ_{Resist} perform double duty:

1. In the numerator, they act as **potentials** that oppose the work function difference Φ .
2. In the denominator, they represent the **energy dissipated** in a full forward/reverse cycle driven by the AC Johnson-Nyquist noise.

7. Optimizing Z_{asym}

Finding the optimal value of Z_{asym} that maximizes power transfer to the load is a three-level “nested” linear programming allocation problem:

1. The **inner allocation** requires selecting the best internal resistive dissipation φ_{Resist} that maximizes the available free energy $\Phi - eV_{Load} - \varphi_{Config} - \varphi_{Resist}$. Obviously, the best φ_{Resist} is the minimum one as determined by material limitations. The next two outer allocations assume this minimum fixed φ_{Resist} .

2. The **middle allocation** can be stated as a Maximin function that distributes a fixed amount of energy E_{dissip} earmarked for dissipation to eV_{Load} and to $\varphi_{Config} = E_{dissip} - eV_{Load} - \varphi_{Resist}$.

$$\begin{aligned} eV_{Load} &= \text{Maximin}(eV_{Load}, \varphi_{Config}) \\ &= \text{Maximin}(eV_{Load}, E_{dissip} - eV_{Load} - \varphi_{Resist}) \end{aligned} \quad (4)$$

The optimum result reassuringly concurs with the MPTT:

$$eV_{Load} = \varphi_{Config} = \frac{E_{dissip} - \varphi_{Resist}}{2} \quad (5)$$

Applying Equation (5) to Equation (3) simplifies Z_{asym} :

$$Z_{asym}^2 = \frac{\Phi - 2eV_{Load} - \varphi_{Resist}}{2eV_{Load} + \varphi_{Resist}} \quad (6)$$

3. The **outer allocation** distributes the unloaded gap potential Φ to the generative process represented in Equation (6) by the numerator, and to the dissipative process, by the denominator. Therefore, the maximum energy transferred to the dissipative process is:

$$2eV_{Load} + \varphi_{Resist} = \text{Maximin}(\Phi - 2eV_{Load} - \varphi_{Resist}, 2eV_{Load} + \varphi_{Resist}) \quad (7)$$

Since φ_{Resist} is fixed, this also yields the maximum energy transferred to the load. Expressing Z_{asym} as a tangent $Z_{asym} = \tan(\theta_{asym})$ of the asymmetry angle θ_{asym} and using Equation (6) yields:

$$\begin{aligned} 2eV_{Load} + \varphi_{Resist} &= \Phi \cos^2(\theta_{asym}); \text{ and} \\ \Phi - 2eV_{Load} - \varphi_{Resist} &= \Phi \sin^2(\theta_{asym}) \end{aligned} \quad (8)$$

Therefore, Equation (7) can be stated as

$$eV_{Load} = \frac{1}{2} \Phi \text{ Maximin}(\sin^2(\theta_{asym}), \cos^2(\theta_{asym})) \quad (9)$$

Hence the maximum energy transfer occurs at:

$$\theta_{asym} = \frac{\pi}{4}; \quad Z_{asym} = 1 \quad (10)$$

$$eV_{Load} = \varphi_{Config} = \frac{1}{4}(\Phi - 2\varphi_{Resist}) \quad (11)$$

As shown in **Figure 3**, the maximum energy transferred per carrier to the load occurs at $\theta_{asym} = \pi/4$, and $Z_{asym} = 1$. These results exactly match the ones in the optimization of Maxwell Demon in [2] and in the optimization of the $\mathbf{E} \times \mathbf{B}$ effect in [1] obtained using different phenomena as vehicles, and different techniques.

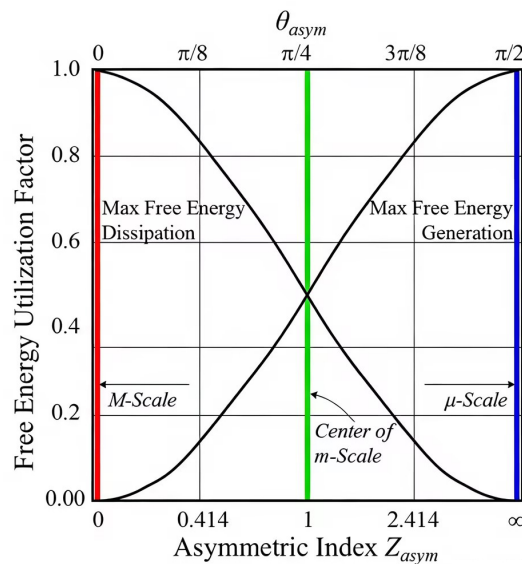


Figure 3. Free Energy Utilization Factor as a function of Z_{asym} . The Maximin maximum occurs at the intersection of the $\sin^2(\theta_{asym})$ and $\cos^2(\theta_{asym})$ function.

The role played by the configuration energy φ_{Config} is interesting. As mentioned in **Section 5.1 Entropic Drive**, it is associated with the polarization or chemical change required for electronic reconfiguration, and, therefore, is essential in providing the diode with *T-Odd* symmetry. It is drawn from the input signal as **heat** to SET the diode's **potential energy** into its forward bias configuration, thereby causing a lowering of entropy of the **surroundings**.

$$\Delta S_{\text{Surroundings}} (\text{Reverse} \rightarrow \text{Forward}) = -\frac{\varphi_{\text{Config}}}{T} \quad (12)$$

When the diode is RESET to its reverse bias configuration, **potential energy** φ_{Config} is dissipated as **heat** back to the **surroundings** causing a corresponding increase in entropy of the surroundings:

$$\Delta S_{\text{Surroundings}} (\text{Forward} \rightarrow \text{Reverse}) = +\frac{\varphi_{\text{Config}}}{T} \quad (13)$$

Over a whole cycle, the change in entropy **due to the change in configuration** is zero, but net heat Q has been converted to electrical work eV_{Load} . Therefore, **in the absence of any resistive loss** φ_{Resist} the overall change in entropy is:

$$\Delta S (\text{Cycle}) = -\frac{eV_{\text{Load}}}{T} = -\frac{\Phi}{4T} \quad (14)$$

These results represent a **limiting case** of a “perfect” MIM diode. It is assumed that a diode in forward bias uses φ_{Config} **to fully turn ON**, and that in reverse bias it dissipates φ_{Config} **to fully turn OFF**. The thermodynamic cycle of a “perfect” diode is discussed in **Appendix D**. Integrating these concepts into a realistic MIM diode model is briefly discussed in **Appendix C** and is the topic of future research.

8. The Johnson-Nyquist Noise

Equation (11) $eV_{\text{Load}} = \varphi_{\text{Config}} = 1/4(\Phi - 2\varphi_{\text{Resist}})$ relates design parameters eV_{Load} and Φ to each other but is silent about how they relate to the magnitude of the input Johnson-Nyquist fluctuations. This section describes this noise in terms of its amplitude and frequency for the purpose of providing a means for scaling these design parameters.

As described by Johnson, [14], the RMS voltage associated with this noise across a resistance R over a frequency band Δf is given by:

$$V_{\text{RMS}}^2 = 4k_B T R \Delta f \quad (15)$$

Maximum noise power ΔP is sent to a matching load, when the voltage across the load is $V_{\text{RMS}}/2$:

$$\Delta P = \frac{V_{\text{RMS}}^2}{4R} = k_B T \Delta f \quad (16)$$

This equation describes the power as boundlessly increasing with bandwidth Δf —the ultraviolet catastrophe. In fact, it does increase but eventually reaches a maximum and then drops because of a quantum cut-off not shown by the equation. Both the frequency at which this maximum occurs, and its magnitude are highly relevant in the rectification of this noise. Therefore, a more accurate equation is required.

A high frequency description of the Johnson-Nyquist noise can begin by describing it in terms of plasmons on a surface, each having energy hf . The power produced per plasmon is limited by the quantum cutoff [15]:

$$\Delta P = \frac{1}{e^{hf/k_B T} - 1} hf \Delta f \quad (17)$$

Assuming a plasmon wavelength λ , this power can be expressed as an intrinsic quantity per unit frequency and per unit area λ^2 :

$$\Delta P = \frac{1}{e^{hf/k_B T} - 1} \frac{hf}{\lambda^2} \Delta f \quad (18)$$

A normalized frequency n_{jn} (i.e., expressed as the number of units of " $k_B T$ " in one Johnson-Nyquist plasmon) can be defined as:

$$n_{jn} = \frac{hf}{k_B T} \quad (19)$$

and the wavelength λ can be written in terms of f and of the dielectric constant κ of the material:

$$\lambda = \frac{c}{f \kappa^{1/2}} = \frac{hc}{\kappa^{1/2} n_{jn} k_B T} \quad (20)$$

Combining equations (18), (19) and (20) yields the power of the noise per unit area and per unit normalized frequency n_{jn} :

$$\frac{dP}{dA} = \frac{\kappa (k_B T)^4}{h^3 c^2} \frac{n_{jn}^3}{e^{n_{jn}} - 1} \Delta n_{jn} \quad (21)$$

Note that Equation (21) shows a fourth-power dependency on temperature, and compares exactly with Planck's blackbody radiation spectrum $x^3/(e^x - 1)$ of [16]-[18] as expressed in [19] (eq. 12) with two adjustments:

1. Equation (21) assumes the dielectric constant $\kappa \neq 1$ to account for the reduction of the speed of light in the material.
2. Reference [19] includes a factor of 2 to account for the total radiated power in both polarizations, whereas Equation (21) describes the measured Johnson-Nyquist power which, even in the best case of perfect source-load matching, is half the available power.

The normalized power distribution in Equation (21) is shown in **Figure 4** as a function of n_{jn} .

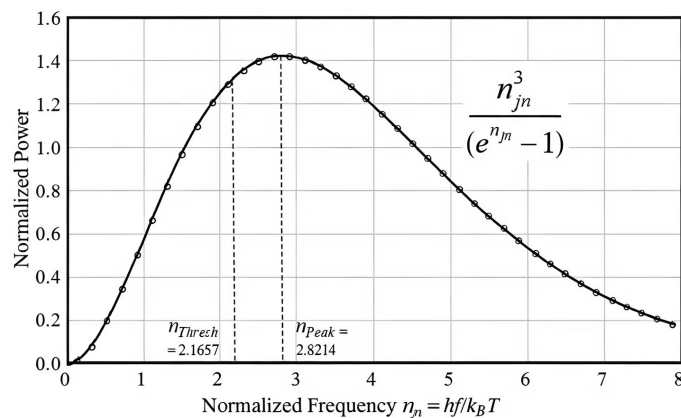


Figure 4. Normalized Johnson-Nyquist power spectrum. Peak power occurs at $n_{Peak} = 2.8214$.

The definite integral of $x^3/(e^x - 1)$ from $x = 0$ to ∞ in Equation (21) is equal to $\pi^4/15 = 6.4939$. Therefore, an upper bound for the noise power per area for all frequencies is:

$$\frac{dP}{dA} = \frac{\pi^4}{15} \frac{\kappa (k_B T)^4}{h^3 c^2} \quad (22)$$

The noise power in Equation (21) peaks at the normalized frequency n_{Peak} which is obtained by differentiating and setting to zero $x^3/(e^x - 1)$ in Equation (21), yielding:

$$n_{Peak} = 2.8214 \quad (23)$$

Substituting this value into equations (19) yields the peak-power frequency:

$$f_{Peak} = n_{Peak} \frac{k_B T}{h} = 2.8214 \frac{k_B T}{h} \quad (24)$$

As an example, setting $T = 300\text{K}$ into Equation (24) yields the peak power frequency of the Johnson-Nyquist noise:

$$f_{Peak} = 17.64 \text{ THz} \quad (25)$$

and substituting this value in Equation (20) yields the peak power wavelength in the material:

$$\lambda_{Peak} = \frac{c}{f_{Peak} \kappa^{1/2}} = \frac{1}{n_{Peak}} \frac{hc}{\kappa^{1/2} k_B T} = 0.3544 \frac{hc}{\kappa^{1/2} k_B T} \quad (26)$$

The noise energy for a plasmon at f_{Peak} is therefore:

$$E_{Peak} = h f_{Peak} = n_{Peak} k_B T \quad (27)$$

and at 300K it is:

$$E_{Peak} = n_{Peak} k_B T = 72.94 \text{ meV} = 11.68 \times 10^{-21} \text{ Joules} \quad (28)$$

This is the peak energy of a Johnson-Nyquist plasmon at a given frequency (temperature) per unit area λ^2 .

9. Dimensions and Power Output of the Diode

Such diode stacks could be integrated into semiconductor chips to provide the chips with power in situ. How small should such diodes be? This section covers the following topics:

- Area of diode.
- Thickness of diode
- Power Output

9.1. Maximum Area Diode

To determine the best area of a diode, one may begin with the observation that the Johnson Nyquist noise is uncorrelated at different points on a surface if these points are *spaced further apart than the coherence length of the noise or the wavelength λ_c of a plasmon*. To calculate λ_c in the material, one can first observe that in free space, at $T = 300\text{K}$, the wavelength at peak power is,

$$\lambda_{peak} = \frac{c}{f_{peak}\kappa^{1/2}} = 16.99 \times 10^{-6} \text{ m} \quad (29)$$

where the dielectric constant κ is assumed to be unity. However, in a MIM structure with a nanometer-scale insulator (1 - 5 nm), the surface plasmon polaritons (SPPs) are strongly compressed by the effective mode index which is about $n_{eff} \approx 30$ for copper at that frequency [20]-[26]. Therefore, the plasmon wavelength for the copper example along the surface is about:

$$\lambda_c = \frac{c}{f_{peak} n_{eff}} \approx 0.566 \times 10^{-6} \text{ m} \quad (30)$$

The area of a coherent pixel is therefore:

$$\lambda_c^2 = \frac{c^2}{f_{peak}^2 n_{eff}^2} \approx 0.320 \times 10^{-12} \text{ m}^2 \quad (31)$$

One could imagine a large flat MIM diode subdivided into little pixels of area $A = \lambda_c^2$, and thickness λ_p , with each pixel connected in parallel between two commonly shared conductive plates, each operating as a small diode, each operating uncorrelated and independently of each other, each rectifying its own local correlated noise, and each feeding its DC outputs in parallel to the plates. **Figure 5** shows a stack of MIM diodes.

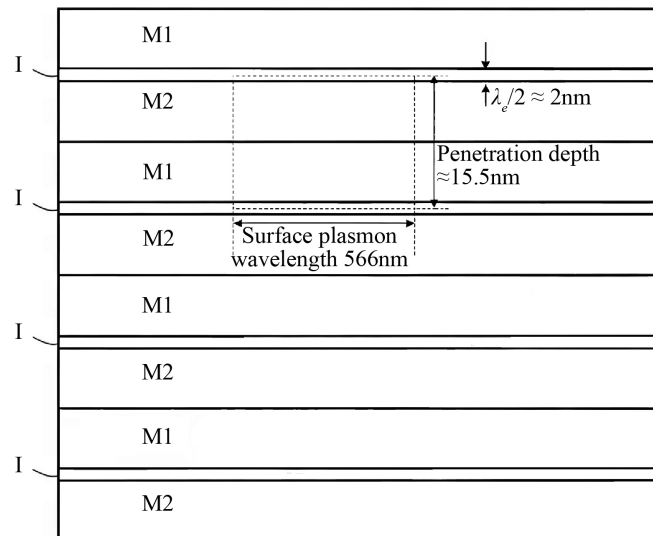


Figure 5. Large area MIM-class diodes can be stacked in layers forming a multiplicity of independently operating diodes, the volume of each diode defined by the correlation area of a surface plasmon and the penetration depth of the peak power frequency.

This visualization of a diode as an assembly of distinct pixels is useful as an explanatory tool but in practice there is no need for *explicitly* dividing the diode because *separate correlation areas are formed spontaneously by the signal itself*. The lithographic complexity of fabricating individual diodes is thereby eliminated. In principle, there is no limit to the size of the area of the diode.

9.2. Thickness of Gap. A Fabry-Perot Resonator for Electrons

Conductance through the gap can be promoted by favoring the transmission of electrons. Since the dielectric constant of the gap is lower than that of the metal, this can be achieved by setting the thickness of the gap to one half the de Broglie wavelength of an electron to cancel destructive interference due to reflections as in a Fabry-Perot resonator. Setting the width of the gap d_G to half the de Broglie wavelength, yields the thickness for maximum transmission:

$$d_G = \frac{1}{2} \lambda_{Peak} = \frac{h}{2\sqrt{2m_{eff} E_{Peak}}} = \frac{h}{2\sqrt{2m_{eff} n_{Peak} k_B T}} \quad (32)$$

Using $m_{eff} = m_0$ at $T = 300\text{K}$, provides an upper bound for the optimal thickness of the gap³:

$$d_G = 2.27 \text{ nm} \quad (33)$$

This thickness is an upper bound because:

1. The effective mass of the electrons is assumed to be m_0 and
2. The change in effective gap thickness due to variations in occupancy and vacancy states as discussed in **Section 4.2** and shown in **Figure 2** is not included in d_G .

Fabricating a device with this dimension is challenging but can be achievable with tailored self- assembling monolayers (SAM), atomic layer deposition (ALD), and other semiconductor technology.

9.3. Thickness of Diode

The diode's metal layers need to be as thick as the skin penetration depth λ_p of the Johnson-Nyquist noise at the peak power frequency f_{Peak} :

$$\lambda_p = \sqrt{\frac{\rho}{\mu_{mag} \pi f_{Peak}}} \quad (34)$$

where μ_{mag} is the magnetic permeability, and ρ is resistivity. For a non-magnetic metal, one can assume $\mu = 4\pi \times 10^{-7} \text{ H/m}$, reducing the equation to:

$$\lambda_p = \frac{3162}{2\pi} \sqrt{\frac{\rho}{f_{Peak}}} \quad (35)$$

For example, using copper with a resistivity of about $\rho = 0.167 \times 10^{-7} \Omega \cdot \text{m}$, the penetration depth at 17.64 THz is:

$$\lambda_{pCopper} = 0.0155 \times 10^{-6} \text{ m} \quad (36)$$

The classical skin-depth formula in Equation (34) is marginal at 17.64 THz, where copper's Drude relaxation time makes conductivity complex. The true penetration depth there is set partly by the plasma frequency. Accordingly, the result of Equation (36) shall be treated as an approximation. Diode layers arranged in a

³Carriers near the Fermi level have eV-scale kinetic energies and sub-nm wavelengths inside the gap region significantly smaller than the one provide by the Fabry-Perot condition.

stack, much thinner than this roughly estimated penetration depth would have to share thermal plasmons, and, therefore, become less efficient.

10. Matching the Threshold to the Noise

10.1. Variability of the Gap Potential

Unlike a battery with a **fixed chemical potential**, the free energy available from a diode is not fixed: it is a function of the load. An **unloaded** diode has an open-circuit potential difference Φ across the gap due to the different work functions of its metal plates. However, if a load is connected to the diode, the rectification of the Johnson-Nyquist noise produces a current. This current generates potentials eV_{Load} across the load and φ_{Config} across the diode and φ_{Resist} due to internal diode resistance. These three potentials oppose and reduce the potential across the gap from Φ to $\Phi - eV_{Load} - \varphi_{Config} - \varphi_{Resist}$. Hence, the gap potential is variable:

$$\text{Loaded Gap Potential} = \Phi - eV_{Load} - \varphi_{Config} - \varphi_{Resist} \quad (37)$$

Therefore, a diode optimized such that $Z_{asym} = 1$ and according to equations (3) and (11) has an actual gap potential:

$$\text{Loaded Optimized Gap Potential} = \frac{1}{2}\Phi = 2eV_{Load} + \varphi_{Resist} \quad (38)$$

10.2. Total Power of Johnson-Nyquist Noise

The power spectrum of the Johnson-Nyquist noise per unit normalized frequency Δn_{jn} is given by Equation (21):

$$\frac{dP}{dA} = \frac{\kappa(k_B T)^4}{h^3 c^2} \frac{n_{jn}^3}{e^{n_{jn}} - 1} \Delta n_{jn} \quad (39)$$

The total power over all frequencies is obtained by integrating from 0 to ∞ as shown in Equation (22).

$$\frac{dP}{dA} = \frac{\pi^4}{15} \frac{\kappa(k_B T)^4}{h^3 c^2} = 6.4939 \frac{\kappa(k_B T)^4}{h^3 c^2} \quad (40)$$

Using copper plates with $\kappa = n_{eff}^2 = 900$, [20] [26] and the total power at 300K is:

$$\frac{dP}{dA} = 6.4939 \times 10.130 = 65.780 \text{ mW/mm}^2 \quad (41)$$

10.3. Matching Gap Potential to Noise for Maximum Power

The diode is a threshold device, conductive only when the energy of a fluctuation $n_{jn}k_B T$ exceeds the threshold **potential** $n_{Thresh}k_B T$ determined by the **energy barrier** across the gap as per Equation (38):

$$n_{jn}k_B T \geq n_{Thresh}k_B T = \frac{1}{2}\Phi = 2eV_{Load} + \varphi_{Resist} \quad (42)$$

Furthermore, each successful rectification event delivers a fixed energy $n_{Thresh}k_B T = eV_{Load} + \varphi_{Config} + \varphi_{Resist}$ to be shared by the load, the configuration process,

and the internal resistive loss, **regardless of how much the plasmon energy exceeds the threshold**; the excess is thermalized. The work function difference Φ acts as a catalyst and remains unchanged.

Accordingly, the power rectified at the threshold is the product of the **rate of above-threshold events** and the **fixed energy delivered** per event.

The rate of all **candidate** events (regardless of the threshold) within a normalized bandwidth Δn_{jn} is obtained by dividing the plasmon energy spectrum of Equation (22) by the energy per plasmon $n_{jn}k_B T$:

$$\frac{d^2 N}{dA dt} = \frac{\kappa (k_B T)^3}{h^3 c^2} \frac{n_{jn}^2}{e^{n_{jn}} - 1} \Delta n_{jn} \quad (43)$$

where $d^2 N/dA dt$ is the number of plasmon crossing a unit area. A single plasmon with energy $n_{jn}k_B T > n_{Thresh}k_B T$ traverses the diode and generates an energy fixed by the threshold potential $n_{Thresh}k_B T$. Taking the product of $n_{Thresh}k_B T$ and the plasmon distribution in Equation (43) and integrating from n_{Thresh} to infinity yields the total power carried by rectified plasmons through the diode:

$$\text{Power through diode} = \frac{\kappa (k_B T)^4}{h^3 c^2} \int_{n_{Thresh}}^{\infty} \frac{n_{Thresh} n_{jn}^2}{e^{n_{jn}} - 1} dn_{jn} \quad (44)$$

To maximize power, we maximize the definite integral with respect to n_{Thresh} . A numerical analysis yields the optimum normalized threshold frequency:

$$n_{Thresh} = 2.1657 \quad (45)$$

which corresponds to a maximum integral of 2.8494. Note that $n_{Thresh} = 2.1657$ is not equal to $n_{Peak} = 2.8214$ as per Equation (23) and shown in **Figure 4**. Hence

$$\text{Maximum Power through Diode} = 2.8494 \frac{\kappa (k_B T)^4}{h^3 c^2} \quad (46)$$

Comparing Equation (46) with Equation (40) yields the maximum power efficiency through the threshold:

$$\text{Fraction of Total Power through Diode} = \frac{2.8494}{6.4939} = 0.43878 \quad (47)$$

For the sake of simplicity, we shall assume in the rest of this paper that the diode resistive loss $\phi_{Resist} = 0$. Therefore, as established in Equation (42), the actual power that reaches the load is only one half of the power through the diode: ($eV_{Load} = \frac{1}{2} n_{Thresh}k_B T$). Dividing Equation (46) by two yields:

$$\text{Power to Load} = 1.4247 \frac{\kappa (k_B T)^4}{h^3 c^2} \quad (48)$$

At $T = 300\text{K}$, with copper layers and $\kappa = n_{eff}^2 \approx 900$, the power to the load is:

$$\text{Power to Load} = 1.4247 \frac{\kappa (k_B T)^4}{h^3 c^2} = 1.4247 \times 10.130 = 14.432 \text{ mW/mm}^2 \quad (49)$$

This is approximately 42.8 times the power per unit area defined by the Shockley-Queisser efficiency limit [27] [28] of 0.337 mW/mm² of solar cells.

Since the thickness of MIM diodes is typically in the submicron range, stacking

such devices as shown in **Figure 5** could achieve substantial power output per unit volume. The power per unit volume is obtained by dividing the power per unit area in Equation (48) by the optimal thickness of the metal layers λ_p . At 300K, and for $\lambda_{pCopper} + d_G = 15.5 \text{ nm} + 2.27 \text{ nm}$ as per Equation (36) and (33):

$$\left(\frac{dP}{dV} \right)_{Load} \approx 812 \text{ W/mm}^3 \quad (50)$$

Of course, the rapid conversion of thermal energy to DC power quickly lowers the temperature of the device. Heat exchangers interleaved within the stack then become the real limitation to the power generated per unit volume.

10.4. Optimal Design

When a plasmon's energy is equal to $n_{Thresh}k_B T$, it is exactly and fully allocated: some of it, φ_{Resist} is dissipated in the diode—The rest is equally shared by the load, and by the configuration, and nothing is left over at threshold.

$$n_{jn}k_B T \geq n_{Thresh}k_B T = \frac{1}{2}\Phi = 2eV_{Load} + \varphi_{Resist} = 2\varphi_{Config} + \varphi_{Resist} \quad (51)$$

This equation is illustrated by **Figure 6** for an ideal diode with $\varphi_{Resist} = 0$ and operating under optimal design conditions:

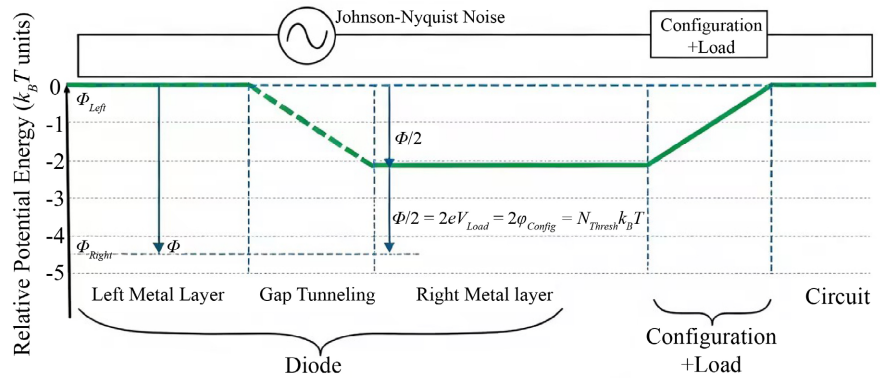


Figure 6. Potential energy profile around circuit for an ideal diode with $\varphi_{Resist} = 0$. The load potential eV_{Load} and configuration potential φ_{Config} decrease the potential across the gap from the work function difference, Φ , to $\Phi - eV_{Load} - \varphi_{Config}$. The left work function Φ_L is arbitrarily selected at zero relative potential.

From Equation (45), $n_{Thresh} = 2.1657$, and with Equation (51) assuming an ideal optimized diode with $\varphi_{Resist} = 0$:

$$eV_{Load} = \varphi_{Config} = 1.0829k_B T \text{ and } \Phi = 4.3314k_B T \quad (52)$$

At 300K,

$$\begin{aligned} eV_{Load} = \varphi_{Config} &= 27.995 \text{ meV}; \\ n_{Thresh}k_B T &= 55.990 \text{ meV}; \text{ and} \\ \Phi &= 111.98 \text{ meV} \end{aligned} \quad (53)$$

Thus, a stack of about 36 diodes, each less than a micron thick, connected in

series could produce up to 1V.

11. Absolute Values of the Work Functions

Equation (51) relates the work function difference $\Phi = \Phi_R - \Phi_L$ between the right and left plates to other design parameters but is silent on the absolute values of Φ_R and Φ_L . The plates and the gap present a tunneling barrier to carriers. The height of this barrier is the energy separation between the plate's Fermi level and the conduction band edge of the insulator filling the gap. In the ideal (Schottky-Mott) limit, this barrier height equals the plate's work function minus the electron affinity, χ_{ins} , of the insulator; real interfaces carry dipoles and defect states that can shift the actual barrier from this nominal value. Maximizing the transmitted current while preserving rectification requires operating near the boundary between quantum tunneling and classical over-barrier transport, referred to here as the marginal tunneling mode. Too large a mean barrier is detrimental because it exponentially suppresses the tunneling current. Too small a mean barrier is also detrimental because it moves the diode into classical over-barrier transmission and eliminates its rectifying function. Optimally, the mean barrier should therefore be approximately equal to the threshold energy of the plasmons:

$$\frac{(\Phi_R + \Phi_L)}{2} - \chi_{ins} = n_{Thresh} k_B T \quad (54)$$

Combining this equation with the optimal $\Phi = \Phi_R - \Phi_L = 2n_{Thresh} k_B T$ from Equation (51) yields absolute work functions in terms of $n_{Thresh} k_B T$:

$$\Phi_R - \chi_{ins} = 2n_{Thresh} k_B T; \quad \Phi_L - \chi_{ins} = 0 \quad (55)$$

Equations (51) and (55) links many design parameters together:

$$n_{Thresh} k_B T = \frac{1}{2} \Phi = \frac{1}{2} (\Phi_R - \chi_{ins}) = 2eV_{Load} + \varphi_{Resist} = 2\varphi_{Config} + \varphi_{Resist}; \quad (56)$$

$$\Phi_L = \chi_{ins}$$

The equality $\Phi_L = \chi_{ins}$ can be approximated using pure metals and insulators as shown in **Table 1** and **Table 2**. If desired, a more precise match can be achieved by alloying metals and using compound or mixed-oxide insulators or by altering fabrication methods.

Table 1. Insulators.

Insulator (Oxides)	Electron Affinity χ_{ins}
TiO ₂	~3.9 - 4.3 eV
Nb ₂ O ₅	~3.9 - 4.3 eV
SnO ₂	~4.5 eV
ZnO	~4.2 - 4.35 eV
In ₂ O ₃	~3.5 - 4.50 eV (typically ~4.45 eV)

Table 2. Metals.

Metal	Work Function Φ_L
Hafnium (Hf):	~3.9 eV
Niobium (Nb):	~3.95 - 4.3 eV
Aluminum (Al):	~4.06 - 4.28 eV
Cadmium (Cd):	~4.07 - 4.08 eV
Indium (In):	~4.09 eV
Tantalum (Ta):	~4.00 - 4.25 eV
Lead (Pb):	~4.14 - 4.25 eV
Titanium (Ti):	~4.33 eV
Silver (Ag):	~4.26 - 4.74 eV
Copper (Cu):	~4.5 - 5.10 eV

The above values are approximate depending on preparation, fabrication and orientation of crystal faces and on the surface treatment of the plates comprising the diode. Their variability can be of benefit in fine tuning the work functions:

1. $\Phi = \Phi_R - \Phi_L$ to twice the threshold energy $2n_{Thresh}k_B T$, and
2. $\Phi_L = \chi_{ins}$.

A future detailed analysis will have to include the influence of the effective mass, image-force corrections, and an over-barrier transport model.

12. Deviation from Optimality Due to Multiple Quanta Threshold Interaction

The model presented herein assumes a single quantum interacting with the gap threshold potential. In practice multiple plasmons may combine to overcome the threshold. Such addition and subtraction of boson quanta is well established. For example, multi-quantum barrier crossing by incoherent noise such as photon-assisted tunneling with n-photon sidebands, is discussed in Tien-Gordon [29], Tucker-Feldman [30], Grover-Joshi-Moddel [31] and Asada, M. and Suzuki [32].

The analysis of section 10.3 can be extended to cover multiple quanta interactions. In the single plasmon interaction, the energy generated across the threshold is the product of the threshold energy $n_{Thresh}k_B T$ and the plasmon distribution of Equation (43) integrated from $n_{Thresh}k_B T$ to ∞ as in Equation (46):

$$P_1 = \frac{\kappa(k_B T)^4}{h^3 c^2} \int_{n_{Thresh}}^{\infty} \frac{n_{Thresh} n_{jn}^2}{e^{n_{jn}} - 1} dn_{jn} = 2.8494 \frac{\kappa(k_B T)^4}{h^3 c^2} \quad (57)$$

The proposed extension uses the product of $n_{Thresh}k_B T$ and the **joint distribution** of multiple plasmons whose individual energy is smaller than $n_{Thresh}k_B T$ but whose total energy is equal or greater than $n_{Thresh}k_B T$. This product is doubly integrated:

$$P_2 = \frac{1}{2} \frac{\kappa(k_B T)^4}{h^3 c^2} \int_0^{n_{Thresh}} \frac{n_{Thresh} n_{jn1}^2}{e^{n_{jn1}} - 1} \left[\int_{n_{Thresh} - n_{jn1}}^{n_{Thresh}} \frac{n_{Thresh} n_{jn2}^2}{e^{n_{jn2}} - 1} dn_{jn2} \right] dn_{jn1} \quad (58)$$

The total power produced by combining a single and a double plasmon channel is the sum of the two:

$$P_{1+2} = P_1 + P_2 \quad (59)$$

The power generated by such multiquanta threshold activation has an upper bound: it cannot exceed the total power of the Johnson-Nyquist spectrum for all channels combined as per Equation (40):

$$\frac{dP_{total}}{dA} = \frac{\pi^4}{15} \frac{\kappa(k_B T)^4}{h^3 c^2} = 6.4939 \frac{\kappa(k_B T)^4}{h^3 c^2} \quad (60)$$

which, from Equation (46), is $6.4939/2.8494 \approx 2.28$ times greater than the single plasmon interaction model of Equation (57). Refining this analysis is left for future study.

13. Conclusions

As highlighted by the axiomatic analysis in [1], all theories and experiments supporting the second law, without exception, investigate two limiting conditions: μ -Scale and M -Scale.

1. At μ -Scale negentropy can be produced by T -Odd symmetry but remains inaccessible due to the lack of thermodynamic interactions.

2. At M -Scale abundant interactions provide the means for accessing this negentropy but destroy it before it can be accessed.

This catch-22 constrains conventional thermodynamics. However, it is a false dichotomy between the μ -Scale and M -Scale extremes, which gives no consideration to the m -Scale middle ground. No theory or experiment supporting the law probes m -scale where some negentropy is destroyed but the rest can be harvested.

A diode is intrinsically spatially asymmetric due to an imbalance in the configuration of occupied and vacant states across the gap. The diode becomes T -Odd symmetric when a time-varying signal—for example, the Johnson-Nyquist noise—modulates this asymmetry, thereby allowing carriers to move preferentially in one direction. A state variable with an associated energy is defined to describe changes in this configuration.

In the absence of any externally forced signal—and only driven by the Johnson-Nyquist noise—carriers remain in thermal equilibrium as they tunnel back and forth across the gap. They do so isothermally (there is no temperature difference across the gap), and adiabatically (they do not experience any collision in direct tunneling). However, the emerging T -Odd symmetry caused by an input signal favors tunneling in the forward direction over the reverse direction. This unidirectional tunneling lowers the positional entropy of the carriers. Being isothermal, adiabatic and negentropic, this process is a free compression, the exact reverse of the well-known free expansion. A full thermodynamic cycle comprises a free (isothermal) compression in the diode and an isothermal expansion in the load.

Operation at the m -Scale is optimized by defining a dimensionless index of

asymmetry Z_{asym} that quantifies the position of a system along this scale. This index is a ratio of two factors, 1) a free energy generative factor which dominates at μ -Scale where $Z_{asym} = \infty$, and 2) a dissipative factor which dominates at M -Scale where $Z_{asym} = 0$. It is shown that power output to the load is maximized at the geometric mean of m -Scale where $Z_{asym} = 1$.

The free energy provided by the potential Φ across the gap of a MIM diode, can be used to convert heat to electrical work W sent to a load. Φ operates as a catalyst and is not consumed. The linear programming maximin function is applied several times in a nested fashion to maximize power output: 1) to the allocation of free energy to the dissipation due to internal electrical resistance, 2) to the allocation of Φ between generation and dissipation of **free energy** and 3) to the allocation of **energy** between the diode as a source, and to the load.

Accordingly, the optimal configuration for maximum power transfer requires an ideal diode with zero internal resistance, and a work function difference Φ defined by $1/4 \Phi = eV_{Load} = \varphi_{Config}$ where eV_{Load} is the load potential and φ_{Config} is the configuration energy. Hence the maximum work transferred to the load is limited to $1/4 \Phi$.

Furthermore, the maximum amount of energy that can be rectified from the Johnson-Nyquist noise by a MIM diode and sent to the load is shown to be $1.0829k_B T$. Accordingly, the optimal configuration for maximum power transfer per plasmon is $1/4 \Phi = eV_{Load} = \varphi_{Config} = 1.0829k_B T$.

A MIM diode rectifying the above-threshold portion of the Johnson-Nyquist noise at 300K can produce up to 14.432 mW/mm² or nearly 42.8 times the power per unit area of a solar cell (with $\kappa = 900$ for copper plates).

Considering Φ as a Helmholtz energy A_{Helm} , the second law, extended for m -Scale operation, becomes:

$$\Delta S \geq -\frac{\Phi}{4T} = -\frac{A_{Helm}}{4T} \quad (61)$$

When the diode is “loaded,” its available free energy decreases from Φ to $\Phi_{Loaded} = \Phi - eV_{Load} - \varphi_{Config} - \varphi_{Resists}$, which under fully matched optimal conditions is $\Phi_{Loaded} = 1/2 \Phi$. The second law extension of Equation (61) with a loaded diode becomes:

$$\Delta S \geq -\frac{\Phi_{Loaded}}{2T} = -\frac{A_{Helm-Loaded}}{2T} \quad (62)$$

This extension of the second law does not contradict any evidence—theoretical or experimental—supporting the law, simply because such evidence does not exist at m -Scale. This result has been reproduced with different methodologies applied to two different vehicles: The optimization of Maxwell’s Demon [2] and the optimization of the $\mathbf{E} \times \mathbf{B}$ Thermoelectric Effect [1].

Data Availability Statement

Please contact the author for additional information.

Acknowledgements

I thank my wife Penny for her unwavering support, and my children, grandchildren and all future generations of the world who motivate me to pursue this research for tikkun olam.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- [1] Levy, G.S. (2026) Catch-22 in Second Law Revealed by Axiomatic Analysis, Part 1, Optimizing ExB Effect, Extending Law. *Journal of Applied Mathematics and Physics*, **14**, 3968-4012. <https://doi.org/10.4236/jamp.2026.1410194>
- [2] Levy, G.S. (2026) The Second Law Extended to Time-Reversal Asymmetric Systems—Thermodynamic Threshold Optimization. *Journal of Applied Mathematics and Physics*, **14**, 961-994.
- [3] Nikulov, A.V. (2021) Dynamic Processes in Superconductors and the Laws of Thermodynamics. *Physica C: Superconductivity and Its Applications*, **589**, Article ID: 1353934. <https://doi.org/10.1016/j.physc.2021.1353934>
- [4] Gurtovoi, V.L., Antonov, V.N., Exarchos, M., Il'in, A.I. and Nikulov, A.V. (2019) The DC Power Observed on the Half of Asymmetric Superconducting Ring in Which Current Flows against Electric Field. *Physica C: Superconductivity and Its Applications*, **559**, 14-20.
- [5] Kondo, S., Kameyama, M., Imaoka, K., Shimoi, Y., Mathevet, F., Fujihara, T., *et al.* (2024) Organic Thermoelectric Device Utilizing Charge Transfer Interface as the Charge Generation by Harvesting Thermal Energy. *Nature Communications*, **15**, Article No. 8115. <https://doi.org/10.1038/s41467-024-52047-5>
- [6] Ourednik, P., Picco, G., Tuan Nguyen, D., Spudat, C. and Feiginov, M. (2023) Large-Signal Dynamics of Resonant-Tunneling Diodes. *Journal of Applied Physics*, **133**, Article 014501. <https://doi.org/10.1063/5.0134223>
- [7] Tsu, R. and Esaki, L. (1973) Tunneling in a Finite Superlattice. *Applied Physics Letters*, **22**, 562-564. <https://doi.org/10.1063/1.1654509>
- [8] Simmons, J.G. (1963) Generalized Formula for the Electric Tunnel Effect between Similar Electrodes Separated by a Thin Insulating Film. *Journal of Applied Physics*, **34**, 1793-1803. <https://doi.org/10.1063/1.1702682>
- [9] Simmons, J.G. (1963) Electric Tunnel Effect between Dissimilar Electrodes Separated by a Thin Insulating Film. *Journal of Applied Physics*, **34**, 2581-2590. <https://doi.org/10.1063/1.1729774>
- [10] Grover, S. and Model, G. (2013) Metal Single-Insulator and Multi-Insulator Diodes for Rectenna Solar Cells. Springer Science + Business Media. <https://www.colorado.edu/faculty/model/sites/default/files/attached-files/grover13c.pdf>
- [11] Brinkman, W.F., Dynes, R.C. and Rowell, J.M. (1969) Tunneling Conductance of Asymmetrical Barriers. *Journal of Applied Physics*, **41**, 1915-1921. https://research.physics.unc.edu/ftsui/shared_data/papers/tunneling%20refs/BDR%20model%20JApplPhys_41_1915.pdf
- [12] Datta, S. (1995) Electronic Transport in Mesoscopic Systems. Cambridge University Press. <https://doi.org/10.1017/cbo9780511805776>

- [13] Ashcroft, N.W. and David Mermin, N. (2021) Beyond the Independent Electron Approximation. In: *Solid State Physics*, Cengage.
- [14] Johnson, J. (1928) Thermal Agitation of Electricity in Conductors. *Physical Review*, **32**, 97-109.
- [15] Watson Dan. M. Lectures.
https://www.pas.rochester.edu/~dmw/ast203/Lectures/Lect_20.pdf
- [16] Young, H.D. and Freedman, R.A. (2016) University Physics. 14th Edition, Pearson, 1256-1257.
- [17] Planck, M. (1914) The Theory of Heat Radiation. Translated by Masius, M. 2nd Edition, P. Blakiston's Son & Co.
- [18] Kramm, G. and Molders, N. (2009) Planck's Blackbody Radiation Law: Presentation in Different Domains and Determination of the Related Dimensional Constants. *Journal of the Calcutta Mathematical Society*, 1-25.
<https://arxiv.org/pdf/0901.1863>
- [19] Marr, J.M. and Wilkin, F.P. (2012) A Better Presentation of Planck's Radiation Law. Union College. <https://arxiv.org/abs/1109.3822v3>
- [20] Davies, J.H. (1997) The Physics of Low-Dimensional Semiconductors: An Introduction. Cambridge University Press. <https://doi.org/10.1017/cbo9780511819070>
- [21] Jordan, E.C. (1968) Electromagnetic Waves and Radiating Systems. Prentice Hall.
- [22] Han, Z. (2010) Ultracompact Plasmonic Racetrack Resonators in Metal-Insulator-Metal Waveguides. *Photonics and Nanostructures—Fundamentals and Applications*, **8**, 172-176. <https://www.sciencedirect.com/science/article/pii/S1569441010000258>
- [23] Maier, S.A. (2007) Plasmonics: Fundamentals and Applications. Springer.
- [24] Economou, E.N. (1969) Surface Plasmons in Thin Films. *Physical Review*, **182**, 539-554. <https://doi.org/10.1103/physrev.182.539>
- [25] Berini, P. (2009) Long-Range Surface Plasmon Polaritons. *Advances in Optics and Photonics*, **1**, 484-588. <https://doi.org/10.1364/aop.1.000484>
- [26] Oulton, R.F., Sorger, V.J., Zentgraf, T., Ma, R., Gladden, C., Dai, L., *et al.* (2009) Plasmon Lasers at Deep Subwavelength Scale. *Nature*, **461**, 629-632.
<https://doi.org/10.1038/nature08364>
- [27] NASA (2026) Solar Irradiance Science.
<https://earth.gsfc.nasa.gov/climate/projects/solar-irradiance/science>
- [28] Shockley, W. and Queisser, H.J. (1961) Detailed Balance Limit of Efficiency of p-N Junction Solar Cells. *Journal of Applied Physics*, **32**, 510-519.
<https://doi.org/10.1063/1.1736034>
- [29] Tien, P.K. and Gordon, J.P. (1963) Multiphoton Process Observed in the Interaction of Microwave Fields with the Tunneling between Superconductor Films. *Physical Review*, **129**, 647-651. <https://doi.org/10.1103/physrev.129.647>
- [30] Tucker, J.R. and Feldman, M.J. (1985) Quantum Detection at Millimeter Wavelengths. *Reviews of Modern Physics*, **57**, 1055-1113.
<https://doi.org/10.1103/revmodphys.57.1055>
- [31] Grover, S., Joshi, S. and Moddel, G. (2013) Quantum Theory of Operation for Rectenna Solar Cells. *Journal of Physics D: Applied Physics*, **46**, Article 135106.
<https://doi.org/10.1088/0022-3727/46/13/135106>
- [32] Asada, M. and Suzuki, S. (2023) Simple Model for Frequency Response of a Resonant Tunneling Diode Caused by Potential Change of Quantum Well Due to Electron Charge. arXiv: 2305.17396.

- [33] Brillouin, L. (1950) Can the Rectifier Become a Thermodynamical Demon? *Physical Review*, **78**, 627.
- [34] van Kampen, N.G. (1960) Non-Linear Thermal Fluctuations in a Diode. *Physica*, **26**, 585.

Appendix

Appendix A: Maxwellian Gas Column Is Isothermal

Two proofs that a Maxwellian gas column is isothermal are presented herein.

A1. Isothermality of Maxwellian Gas Due to Exponential Nature of Maxwellian Distribution.

At ground level Maxwellian gas molecules have a given Maxwellian distribution:

$$f_{MB}(E) = Ae^{-E/k_B T} \quad (63)$$

where A is the normalization factor, and E is the average kinetic energy of the particles. At a given elevation h , potential energy mgh must be added to E , yielding:

$$f_{MB}(E, h) = Ae^{-(E+mgh)/k_B T} \quad (64)$$

Factoring out mgh produces:

$$f_{MB}(E, h) = Ae^{-mgh/k_B T} e^{-E/k_B T} \quad (65)$$

and renormalizing yields:

$$f_{MB}(E, h) = A'e^{-E/k_B T} = A'f_{MB}(E, 0) \quad (66)$$

where:

$$A' = Ae^{-mgh/k_B T} \quad (67)$$

Comparing equations (63) and (66) indicates that the *exponential Boltzmann factor is left unchanged*, but the normalization constant is altered. Hence, the gas column is isothermal, but its density changes with elevation.

A2. Isothermality of Maxwellian Gas Due to Relative Nature of Potential Energy

Consider gas particles observed at elevation h with a distribution defined by the Boltzmann factor:

$$f_{MB}(E, h) = Ae^{-(E+mgh)/k_B T} \quad (68)$$

Since potential energy mgh is relative to the observer, a more accurate version of Equation (68) should include the observer's reference "ground" energy mgh_{Ref} :

$$f_{MB}(E, h) = Ae^{-[E+mg(h-h_{Ref})]/k_B T} \quad (69)$$

A change in potential energy $mg\Delta h$ of gas particles is equivalent to a change in elevation of the frame of reference of the observer. To remain invariant across multiple observers this change must be applied equally to the potential mgh and to the reference potential mgh_{Ref} for any observer. Hence:

$$\begin{aligned} mgh &\rightarrow mg(h + \Delta h) \\ mgh_{ref} &\rightarrow mg(h_{ref} + \Delta h) \end{aligned} \quad (70)$$

Applying these changes to Equation (69) yields:

$$f_{MB}(E, h) = Ae^{-[E+mg((h+\Delta h)-(h_{ref}+\Delta h))]/k_B T} \quad (71)$$

leaving the distribution invariant in any frame:

$$f_{MB}(E, h) = Ae^{-[E + mg(h - h_{Ref})]/k_B T} \quad (72)$$

Hence a Maxwellian gas column at equilibrium is isothermal. The key idea in this proof is that potential energy is relative to the frame of the observer, leaving the Boltzmann factor invariant.

Appendix B: Fermi Gas Column Is Isothermal

This appendix presents a proof that fermions remain isothermal in a potential energy gradient. To avoid circular reasoning, this proof does not depend on the second law. It relies on the *relative nature of potential energy* which requires *observers in frames with different elevations to perceive changes in potential energy equally*.

For the purpose of providing an educational introduction to this topic, **Appendix A** provides two proofs that a column of Maxwellian gas remains isothermal in the absence of any forcing agent.

B1. A Fermi Gas Column is Isothermal

Consider a change in potential energy $\Delta\mu$ in a Fermi gas characterized by the distribution:

$$f_{FD}(E, \mu) = \frac{1}{1 + e^{(E - \mu - \Delta\mu)/k_B T}} \quad (73)$$

Does the change $\Delta\mu$ cause a change in the temperature of the gas? Since the distribution is not strictly exponential, one cannot prove isothermality by factoring a change in potential energy term $\Delta\mu$ out of the distribution as done in **Appendix A** for a Maxwellian gas.

The proof presented below relies on the *relative nature of potential energy with respect to an observer's frame*. This property leads to the invariance of the Boltzmann factor for any observer. The dependency of the Fermi-Dirac distribution on the Boltzmann factor renders the distribution isothermal with elevation. A detailed explanation follows.

In a Fermi gas, quantum state can only be in one of two conditions: occupied by an electron or empty (occupied by a hole), and these two distributions must add up to unity:

$$f_{FD_e}(E, \mu) + f_{FD_h}(E, \mu) = 1 \quad (74)$$

Furthermore, *the ratio of occupied to empty states follows the Boltzmann factor*. Physically, this means the probability of finding an occupied state relative to the probability of finding an unoccupied state is governed by the energy difference represented by chemical potential μ :

$$\frac{f_{FD_e}(E, \mu)}{f_{FD_h}(E, \mu)} = \frac{f_{FD}(E, \mu)}{1 - f_{FD}(E, \mu)} = e^{-(E - \mu)/k_B T} \quad (75)$$

Since potential energy μ is *relative to the observer*, a more accurate version of

Equation (75) should include the *observer's reference "ground" energy* μ_{Ref} :

$$\frac{f_{FD}(E, \mu)}{1 - f_{FD}(E, \mu)} = e^{-[E - (\mu - \mu_{Ref})]/k_B T} \quad (76)$$

The Fermi-Dirac distribution can now be derived using Equation (76) which includes the ground energy. Rearranging terms yields:

$$f_{FD}(E, \mu) = \frac{1}{1 + e^{[E - (\mu - \mu_{Ref})]/k_B T}} \quad (77)$$

A change in potential energy $\Delta\mu$ of the fermions is equivalent to a change in elevation of an observer. To remain invariant across multiple observers, this change must be applied equally to the potential μ and to the reference potential μ_{Ref} for any observer. Hence a change $\Delta\mu$ produces a change in μ and in μ_{Ref} : $\mu \rightarrow \mu + \Delta\mu$ and $\mu_{Ref} \rightarrow \mu_{Ref} + \Delta\mu$. Applying these changes to Equation (77) yields:

$$f_{FD}(E, \mu) = \frac{1}{1 + e^{[E - ((\mu + \Delta\mu) - (\mu_{Ref} + \Delta\mu))]/k_B T}} \quad (78)$$

The $\Delta\mu$'s cancel each other, thereby leaving the distribution invariant:

$$f_{FD}(E, \mu) = \frac{1}{1 + e^{[E - (\mu - \mu_{Ref})]/k_B T}} \quad (79)$$

Therefore, a Fermi gas column in equilibrium has been proven to be isothermal, independently of the second law. The key concept in this proof is that potential energy is relative to the observer, allowing the Boltzmann factor to be invariant across observers, and independent of the potential energy measured in any given frame of reference. The proof also applies to any gas such as bosons, whose statistics can be derived from the Boltzmann factor as shown in Equation (75).

Appendix C: A Dynamic MIM Model for Thermal-Noise Rectification

This appendix provides a bird's eye view of the model required for rectifying very small amplitude and high frequency signals such as the Johnson-Nyquist noise. The formulation of an actual model is left for a future paper.

C1. Basic Requirements for Johnson-Nyquist Noise Rectification

Requirements for the model include:

1. It must be *T-Odd* symmetric: the configuration of the diode must change in response to the polarity of the signal. In a PN diode, the thickness of the depletion zone changes depending on forward or reverse bias. In a fluidic valve, the disk or piston moves in response to the direction of fluid flow.
2. The configuration should be described by a state variable which changes between forward and reverse bias and affects the transmittivity of the diode. This change requires energy φ_{Config} provided by the signal and it occurs over a time period τ_{Config} .

Without such a state variable, current models cannot properly assess how a diode behaves when the signal—the Johnson-Nyquist noise—has an energy $E_{Thermal}$

and characteristic frequency $\omega_{Thermal}$ respectively matching the diode's φ_{Config} and $1/\tau_{Config}$. For example:

1. If φ_{Config} is much greater than $E_{Thermal}$, the noise cannot pay the reconfiguration cost; the configuration remains frozen, the device is merely spatially asymmetric, and it does not rectify.
2. If φ_{Config} is much smaller than $E_{Thermal}$, the noise flips the configuration freely in both directions; with no energetic cost distinguishing SET from RESET, the modulation is itself an equilibrium fluctuation, and the device does not rectify.
3. If $1/\tau_{Config}$ is much smaller than $\omega_{Thermal}$, the diode cannot follow the signal; it remains static and does not rectify.
4. If $1/\tau_{Config}$ is much greater than $\omega_{Thermal}$, the diode follows the signal quasi-statically; it reduces to the instantaneous static model of **Section C.2** and cannot rectify.

These four failure modes are the two endpoints of *m-Scale* approached along two different axes. Cases 1 and 3, in which the configuration is frozen, and no reconfiguration energy is dissipated, correspond to μ -Scale where $Z_{asym} \rightarrow \infty$. Cases 2 and 4, in which reconfiguration is free or instantaneous, and all free energy is dissipated, correspond to *M-Scale* where $Z_{asym} \rightarrow 0$. Rectification is only possible between these endpoints. As shown in this paper, the optimum between the configuration energy and the thermal noise is $\varphi_{Config} = \frac{1}{2}E_{Thermal}$ and the optimal point occurs at *m-Scale* where $Z_{asym} = 1$.

C2. Why Existing Models Are Inapplicable

No diode model currently satisfies these requirements. The models of Tsu-Esaki [7], Simmons [8] [9], Grover [10] and Brinkman-Dynes-Rowell [11] compute the current from the instantaneous voltage across a barrier of fixed geometry. Writing the Tsu-Esaki current explicitly in terms of the transmission function T_{Tunnel} and the gap width d_G :

$$J(V) = \frac{4\pi m_e e}{h^3} \int_0^\infty T_{Tunnel}(E_x; d_G, \bar{\varphi}_B) S_{Supply}(E_x, V) dE_x \quad (80)$$

where E_x is the carrier energy normal to the gap and the supply function S_{Supply} counts the net carriers available for tunneling at that energy:

$$S_{Supply}(E_x, V) = k_B T \ln \left(\frac{1 + e^{\frac{E_F - E_x}{k_B T}}}{1 + e^{\frac{E_F - E_x - eV}{k_B T}}} \right) \quad (81)$$

The transmission function is computed in the WKB approximation through a barrier profile $\varphi_B(x)$ of *fixed* width d_G :

$$\begin{aligned} T(E_x; d_G, \bar{\varphi}_B) &= \exp \left[- \left(\frac{2}{\hbar} \right) \int_0^{d_G} d_G \sqrt{(2m_e (\varphi_B(x) + E_F - E_x))} dx \right] \\ &\approx \exp \left[- 2d_G \frac{\sqrt{(2m_e \bar{\varphi}_B)}}{\hbar} \right] \end{aligned} \quad (82)$$

where $\bar{\varphi}_B$ is the mean barrier height and the last form holds for carriers near the

Fermi level. Three observations show why this formulation cannot describe thermal-noise rectification:

1. Equation (80) is “*Voltage Asymmetric*” and not *T-Odd*, because all parameters except the voltage are fixed. The width d_G appears only as a fixed parameter in the exponent of Equation (82): the barrier geometry cannot respond to the signal. When the device is in thermal equilibrium and in the absence of an input voltage V , the equation is perfectly time symmetric (*T-Even*). Therefore, by the argument of **Section 3**, it cannot rectify the equilibrium Johnson-Nyquist noise.

2. The response is instantaneous. The equations do not include any device time constant τ_{Config} or noise frequency $\omega_{Thermal}$.

3. No energy is assigned to the signal, and no cost in energy φ_{Config} is extracted from the signal to achieve a change in configuration. Furthermore, the density of states enters only through the free-electron supply function of Equation (81) and is completely decoupled from the transmission function—yet it is precisely the asymmetry in the DOS that generates the space-charge layers whose modulation renders the device *T-Odd*.

These are precisely the omissions identified in **Section 4.2**. It is worth noting how sensitively the transmission depends on the gap width: for $\bar{\varphi}_B \approx 1\text{eV}$ the decay constant in Equation (82) is $2\sqrt{2m_e\bar{\varphi}_B}/\hbar \approx 10.2\text{nm}^{-1}$, so a modulation of d_{eff} by only 0.2nm changes the transmission by a factor of about 8. A sub-nanometer configurational change is sufficient to produce a strong rectification ratio, as anticipated in **Section 4.2**.

C3. The Brillouin Diode Paradox

Furthermore, Equation (80) shows its inadequacy by giving the appearance of a second law violation. If a voltage-asymmetric static I-V curve is naively driven by a zero-mean Johnson-Nyquist voltage, time-averaging produces a spurious DC output—the well-known Brillouin diode paradox [33]. The conventional resolution is that a self-consistent kinetic treatment—one that includes the diode’s own thermal fluctuations and its back-action on the signal by van Kampen—restores detailed balance and yields zero output [34].

But this argument negates the very essence of a diode: its configuration is not static; it is modulated by the input signal. The present framework agrees with the conventional resolution of the paradox at the endpoints of *m-Scale* (μ -Scale and *M-Scale*) but shows that a dynamic device—one that borrows reconfiguration energy from the signal—retains a net output within *m-Scale*, bounded by equations (42) and (46).

C4. Proposed Model Outline

It is proposed that the configuration that changes dynamically, as required by *T-Odd* symmetry, is the thickness of the space charges on either side of the gap. These thicknesses depend on the spatial asymmetry in the density of states (DOS) between the two metal plates across the gap. In combination they become *T-Odd*—thin in forward mode and thick in reverse mode, as illustrated in **Figure 2**—because of the input signal. This change in configuration changes the effective

width d_{eff} of the gap, causing preferential tunneling in the forward mode. Hence, for a proper description, d_{eff} must be expressed in terms of the DOS, and the change in d_{eff} must require energy φ_{Config} and occur over a time period τ_{Config} .

Perhaps, the most suitable starting point is the transfer-matrix formulation used by Grover and Modell [10] for multi-insulator diodes, typically solved by numerical methods. The potential profile across the junction is decomposed into N thin slices, each with constant potential U_j and thickness δj . Within slice j the wave function is a superposition of forward and backward waves:

$$\psi_j(x) = A_j e^{(ik_j x)} + B_j e^{(-ik_j x)} \text{ where } k_j = \frac{\sqrt{(2m_e(E_x - U_j))}}{\hbar} \quad (83)$$

where k_j is imaginary inside classically forbidden slices. Continuity of ψ and $d\psi/dx$ at each interface relates adjacent amplitude pairs through a 2×2 matrix M_j , and the product over all slices,

$$M = M_N M_{N-1} \cdots M_1 \text{ and } T(E_x) = \frac{k_N}{k_1} \left| \frac{1}{M_{11}} \right|^2 \quad (84)$$

yields the transmission through the entire structure.

The strength of this formulation for the present purpose is that the potential is already discretized: the space-charge layers on either side of the gap are naturally represented as additional slices whose potentials U_j depend on the local density of states and occupancy, and whose extents d_L and d_R depend on the configuration state. It is therefore suggested to modify this model so that:

1. The slice potentials adjacent to the gap are functions of a configuration state variable determined by the DOS asymmetry of **Figure 2**, giving the effective width $d_{eff} = d_L + d_G + d_R$ its bias dependence.
2. Each change of configuration draws energy φ_{Config} from the signal.
3. The configuration relaxes over a time τ_{Config} set by the charge-relaxation dynamics of the space-charge layers—a time constant distinct from, and much longer than, the femtosecond transit time of a tunneling carrier. The brevity of the transit time is often invoked to justify applying static models at terahertz frequencies; that argument addresses the wrong time constant, since it is the reconfiguration of the space-charge layers, not the transit of a single carrier, which governs the dynamic response.

C5. Relation to Photon-Assisted Tunneling

Dynamic tunneling models do exist, and the best known is photon-assisted tunneling (PAT), introduced by Tien and Gordon [29], generalized by Tucker and Feldman [30], and applied to rectenna solar cells by Grover, Joshi, and Modell [31]. However, PAT is dynamic in a fundamentally different sense. In PAT, the time-varying potential modulates the *energy* of the carriers, producing sidebands at $E \pm n\hbar\omega$ weighted by Bessel functions, while the barrier *geometry* remains fixed and no energy cost is charged for any reconfiguration. PAT correctly describes the rectification of coherent, non-equilibrium illumination. Driven by equilibrium

noise, it respects detailed balance and yields zero output. In the language of this paper, PAT is an endpoint theory: it governs the μ -Scale and M -Scale limits where the conventional second law holds, while the interior of m -Scale is opened only by a configurational state variable carrying an energy cost.

A structural analog of such a state variable appears in the theory of resonant-tunneling diodes, where the frequency response is modeled through the potential change of the quantum well caused by its own electron charge, with an associated relaxation dynamic [32]. That approach, however, is directed at oscillator bandwidth and includes no thermodynamic accounting of the reconfiguration energy. To the author's knowledge, no published model combines a configurational state variable tied to the DOS asymmetry, an explicit reconfiguration energy φ_{Config} drawn from the signal, and an explicit relaxation time τ_{Config} applied to the rectification of equilibrium noise. The formulation of such a model is the subject of a future paper.

Appendix D: Thermodynamics Cycle of a MIM Diode

The thermodynamic cycle of the diode shall assume an idea diode with zero internal resistance dissipation, $\varphi_{Resist} = 0$.

1. Optimal index of asymmetry $Z_{asym} = 1$ in the center of m -Scale, and source load matching as in Equation (3):

$$Z_{asym}^2 = \frac{\Phi - eV_{Load} - \varphi_{Config}}{eV_{Load} + \varphi_{Config}} = 1 \quad (85)$$

2. Source load matching:

$$eV_{Load} = \varphi_{Config} \quad (86)$$

3. Johnson-Nyquist noise matching threshold level at gap:

$$n_{Thresh} k_B T = \Phi - eV_{Load} - \varphi_{Config} \quad (87)$$

which implies:

$$eV_{Load} = \varphi_{Config} = \frac{1}{4}\Phi = \frac{1}{2}n_{Thresh} k_B T \quad (88)$$

Furthermore, since electrons cross the gap of a MIM diode much quicker than the period of the Johnson-Nyquist noise, electrons shall be assumed to be in quasi thermal equilibrium and isothermal across the gap as proven in **Appendix B**.

As illustrated in **Figure 7**, the cycle comprises two stages, BA, and AB.

Initial Conditions.

To monitor the changes in entropy of the electrons, and the environment around the cycle, we begin by setting them to zero:

$$\Delta S_{Electrons} = 0; \quad \Delta S_{Surroundings} = 0 \quad (89)$$

D1. Stage BA

This stage begins when a fluctuation with energy $n_{Thresh} k_B T = 1/2 \Phi$ has barely enough energy to clear the gap potential step $1/2 \Phi = \Phi - eV_{Load} - \varphi_{Config}$. This stage comprises 1) a configuration process and 2) a compression of the electron gas.

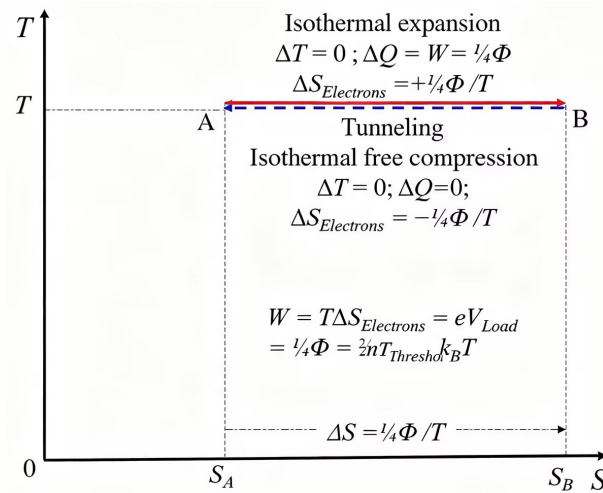


Figure 7. Thermodynamic cycle of Temperature versus Entropy. Electrons cycling through the diode and load follow two stages: 1) A free compression in blue due to the tunneling through the gap where they go from low potential and high entropy, to high potential and low entropy, and 2) an isothermal expansion in green through the rest of the circuit from high potential and low entropy, to low potential and high entropy.

D.1.1 Configuration BA

Half of the fluctuation energy $\varphi_{Config} = 1/4 \Phi = 1/2 n T_{Thresh} k_B T$ is used to configure the electronic distribution. It SETs the diode into a FORWARD mode that favors forward tunneling of electrons through the gap. Since the configuration energy is potential, the only change in entropy is due to the transfer of heat from the surroundings:

$$\Delta S_{Electrons} = 0; \quad \Delta S_{Surroundings} = -\frac{1}{4} \frac{\Phi}{T} \quad (90)$$

D.1.2. Compression BA

This upward tunneling step is the exact reverse of the well-known “free expansion” which is isothermal, adiabatic, and entropic. This upward step can be called a “**free compression**” because it is isothermal, adiabatic and negentropic:

1. *Isothermal* because electrons undergo unforced equilibrium tunneling as explained in **Appendix B**. (Isothermality of a gas column **in equilibrium** is also mandated by the second law.)

$$\Delta T = 0 \quad (91)$$

2. *Adiabatic*. Electrons are not in thermal contact with their surroundings, as they undergo direct tunneling through the gap, therefore:

$$\Delta Q = 0 \quad (92)$$

3. *Negentropic*. Electrons lower their positional entropy by $-1/4 \Phi / T$ as they narrow down their position, and step into a higher potential across the gap. Only the $eV_{Load} = 1/4 \Phi$ share is stored in the electron gas; the φ_{Config} share is stored in the diode’s configuration and returned to the surroundings on RESET —The cumulative changes in entropies are:

$$\Delta S_{\text{Electrons}} = -\frac{1}{4} \frac{\Phi}{T}; \quad \Delta S_{\text{Surroundings}} = -\frac{1}{4} \frac{\Phi}{T} \quad (93)$$

D2. Stage AB

This stage comprises a reconfiguration of the diode and an expansion of the electron gas:

D.2.1 Reconfiguration AB

This stage is initiated when the fluctuation changes sign and RESETS the diode to REVERSE mode. The potential configuration energy $\varphi_{\text{Config}} = 1/4 \Phi = 1/2 n_{\text{Thresh}} k_B T$ is dissipated in the environment. Therefore, the cumulative changes in entropies are:

$$\Delta S_{\text{Electrons}} = -\frac{1}{4} \frac{\Phi}{T}; \quad \Delta S_{\text{Surroundings}} = -\frac{1}{4} \frac{\Phi}{T} + \frac{1}{4} \frac{\Phi}{T} = 0 \quad (94)$$

D.2.2 Expansion AB

As electrons expand isothermally against the load voltage, they receive $Q = e V_{\text{Load}} = 1/4 \Phi$ from the surroundings and perform work $W = e V_{\text{Load}} = 1/4 \Phi$. Therefore, the cumulative change in entropy is:

$$\Delta S_{\text{Electrons}} = -\frac{1}{4} \frac{\Phi}{T} + \frac{1}{4} \frac{\Phi}{T} = 0; \quad \Delta S_{\text{Surroundings}} = -\frac{1}{4} \frac{\Phi}{T} \quad (95)$$

The diode is back to its starting point, but work $W = 1/4 \Phi$ has been performed, heat $Q = 1/4 \Phi$ has been extracted from the surroundings, and the entropy of the surrounding has decreased by $1/4 \Phi / T$.

Appendix E: List of Variables

μ : Potential energy

μ_{mag} : Magnetic permeability

A: Constant

A_{Helm} : Helmholtz free energy

$\mathbf{B}$: Magnetic field vector

c : speed of light

d_{eff} : Effective thickness of gap

d_G : Nominal thickness of gap

d_L : Thickness of left addition to gap

d_R : Thickness of right addition to gap

e : Charge of electron

$\mathbf{E}$: Electric field vector

E_{dissip} : Energy earmarked for dissipation

E_{Peak} : Energy of Johnson-Nyquist plasmon at the spectrum peak

E_{Thermal} : Energy of a thermal plasmon

$e V_{\text{Load}}$: Potential energy across the load

f : Frequency

f_{FD} : Fermi-Dirac distribution

f_{Peak} : frequency at the spectrum peak

g : acceleration of gravity

h : Planck's constant
 J : Current through MIM diode
 k_B : Boltzmann's constant
 m : mass
 m_{eff} : Effective mass of electron
 M -Scale: Macroscale
 m -Scale: Mesoscale
 n_{eff} : square root of dielectric constant κ
 n_{jn} : normalized frequency expressed as the number of " $k_B T$ " units in one Johnson-Nyquist plasmon
 n_{Thresh} : normalized frequency for threshold potential in $k_B T$ units
 n_{Peak} : value of n_{jn} at the peak power in the spectrum
 N : number of plasmon quanta
 P : Power
 P_{Load} : Power to the load
 R : Resistance
 S : Entropy
 T : Temperature
 T -Even: time-reversal symmetric
 T -Odd: time-reversal asymmetric
 V_{RMS} : RMS voltage of the Johnson-Nyquist noise
 Z_{asym} : Index of asymmetry
 Δf : bandwidth
 θ_{asym} : asymmetry angle = $\text{atan}(Z_{asym})$
 κ : dielectric constant
 λ : Wavelength
 λ_c : Plasmon's wavelength for Johnson-Nyquist noise
 λ_p : Penetration depth
 λ_{Peak} : Wavelength at the spectrum peak
 μ -Scale: Microscale
 ρ : resistivity
 τ_{Config} : time period of the configuration
 Φ : Work function difference between metal plates
 $\bar{\phi}_B$: Mean barrier height
 φ_{Config} : Configuration energy of the diode
 φ_{Resist} : Potential energy drop due to internal diode resistance
 χ_{ins} : electron affinity of the insulator
 $\omega_{Thermal}$: Characteristic frequency of thermal noise