Thermal-to-electrical energy conversion by diodes under negative illumination

Parthiban Santhanam and Shanhui Fan
Edward L. Ginzton Laboratory and Department of Electrical Engineering, Stanford University, Stanford, California 94305, USA
(Received 4 May 2015; revised manuscript received 23 October 2015; published 25 April 2016)

We consider an infrared photodiode under negative illumination, wherein the photodiode is maintained at a temperature \( T \) and radiatively exposed to an emissive body colder than itself. We experimentally demonstrate that a diode under such conditions can generate electrical power. We show theoretically that the efficiency of energy conversion can approach the Carnot limit. This work is applicable to waste heat recovery as well as emerging efforts to utilize the cold dark universe as a thermodynamic resource for renewable energy.

DOI: 10.1103/PhysRevB.93.161410

The semiconductor photodiode is widely used as a device for electrical power generation. In a typical configuration, a photodiode at ambient temperature \( T \) is radiatively exposed to a hot surface with a temperature \( T_{\text{surf}} > T \), which illuminates the diode with thermal radiation [Fig. 1(a), top]. The diode absorbs the incident photons to generate electrical power. This configuration is used in both solar and thermophotovoltaics.

In this Rapid Communication we investigate and experimentally demonstrate an alternative configuration through which a photodiode can be used to generate electrical power. In this configuration, depicted in the bottom of Fig. 1(a), the diode is instead radiatively exposed to a cold surface with a temperature \( T_{\text{surf}} < T \). We refer to a diode in such a condition as being under “negative illumination.”

As we will show, an ideal photodiode absorbing only within a narrow band can in fact, when placed under negative illumination from a cold surface \( T_{\text{surf}} < T \), extract work from thermal energy at the Carnot limit \( \eta_{\text{Carnot}} \equiv (T - T_{\text{surf}})/T \).

We apply Eq. (1) to both the positive illumination regime, where \( T_{\text{surf}} > T \), and the negative illumination regime, where \( T_{\text{surf}} < T \); these regimes correspond to the illustrations at the top and bottom of Fig. 1(a), respectively. Here we assume that \( \epsilon(\omega, \theta) = \Theta(\omega - E_{\text{gap}})/\hbar \) with \( \Theta \) being the unit step function and \( E_{\text{gap}} \) being the band gap of the semiconductor.

In Fig. 1(b), the black curve passing through the origin represents the current-voltage relation when \( T_{\text{surf}} = 295 \text{ K} = T \). Equation (1) in this case gives rise to the well-known diode law. The red curve represents the relation when \( T_{\text{surf}} = 315 \text{ K} > T \) (i.e., positive illumination). Here the greater flux of incident photons causes the generation rate of electrons and holes to exceed their recombination rate; thus a reverse current is seen when the diode’s terminals are shorted. At open circuit, an increased concentration of electrons and holes is required to raise the recombination rate to match the elevated generation rate. Thus a positive open-circuit voltage is obtained.

The blue curve in Fig. 1(b) represents the characteristic current-voltage relation of the diode when \( T_{\text{surf}} = 255 \text{ K} < T \) (i.e., negative illumination). With negative illumination, at short circuit the generation rates of electrons and holes fall short of their recombination rate, so that a forward current is seen. At open circuit a reverse bias is now required to decrease the steady-state carrier concentrations to in turn decrease the recombination rate to match generation.

We begin our analysis of photodiodes under negative illumination by applying the principle of detailed balance [1]. For simplicity we first consider an ideal photodiode with zero nonradiative recombination. We assume the diode is maintained at temperature \( T \) and is radiatively exposed to a perfectly absorptive surface at temperature \( T_{\text{surf}} \) over the entire hemisphere which it faces. We may express the current density \( J \) through such a diode in terms of the difference between the outgoing and incoming photon fluxes \( J_{\text{N, out}} \) and \( J_{\text{N, in}} \).

\[
J = q(J_{\text{N, out}} - J_{\text{N, in}}) = \frac{q}{2\pi} \int_0^\infty d\omega \int_0^{\pi/2} d\theta \cos(\theta) \sin(\theta) \epsilon(\omega, \theta) \left[ e^{(\hbar\omega - qV)/(k_BT\text{surf})} - 1 \right] \left[ e^{(\hbar\omega)/(k_BT)} - 1 \right] \]

where the emissivity \( \epsilon \) is in general a function of angular frequency \( \omega \) and polar angle \( \theta \), and we have invoked Kirchhoff’s law to set the diode’s absorptivity equal to its emissivity.

\[
 \epsilon(\omega, \theta) = \Theta(\omega - E_{\text{gap}})/\hbar \]

We apply Eq. (1) to both the positive illumination regime, where \( T_{\text{surf}} > T \), and the negative illumination regime, where \( T_{\text{surf}} < T \); these regimes correspond to the illustrations at the top and bottom of Fig. 1(a), respectively. Here we assume that \( \epsilon(\omega, \theta) = \Theta(\omega - E_{\text{gap}})/\hbar \) with \( \Theta \) being the unit step function and \( E_{\text{gap}} \) being the band gap of the semiconductor.

In Fig. 1(b), the black curve passing through the origin represents the current-voltage relation when \( T_{\text{surf}} = 295 \text{ K} = T \). Equation (1) in this case gives rise to the well-known diode law. The red curve represents the relation when \( T_{\text{surf}} = 315 \text{ K} > T \) (i.e., positive illumination). Here the greater flux of incident photons causes the generation rate of electrons and holes to exceed their recombination rate; thus a reverse current is seen when the diode’s terminals are shorted. At open circuit, an increased concentration of electrons and holes is required to raise the recombination rate to match the elevated generation rate. Thus a positive open-circuit voltage is obtained.

The blue curve in Fig. 1(b) represents the characteristic current-voltage relation of the diode when \( T_{\text{surf}} = 255 \text{ K} < T \) (i.e., negative illumination). With negative illumination, at short circuit the generation rates of electrons and holes fall short of their recombination rate, so that a forward current is seen. At open circuit a reverse bias is now required to decrease the steady-state carrier concentrations to in turn decrease the recombination rate to match generation.

In both the cases of positive and negative illumination, at operating points between open circuit and short circuit the product \( JV \) is negative and corresponds to the generation of electrical power. The device is extracting work from the flow of heat across the temperature difference between one body at temperature \( T \) and another at \( T_{\text{surf}} \). The existence of the operating regime explored here, in which the photodiode acts as a heat engine located on the hot side of a temperature difference, has been alluded to multiple times in the existing literature but does not appear to have been analyzed in detail [2,3]. In particular, the efficiency with which a diode could extract electrical work from thermal energy has not been analyzed and no experiments have been reported to verify the basic existence of the phenomenon. These items are therefore the primary aims of this Rapid Communication.

We may use Eq. (1) to calculate the energy conversion efficiency of an ideal diode under negative illumination. For each electronic charge that flows through the combined diode-load circuit, the amount of work extracted is \(-qV\) and the amount of heat rejected is \(h\omega\), where \(q\) is the magnitude of the electron’s charge and \(h\omega\) is the energy of an average photon
RAPID COMMUNICATIONS

PARTHIBAN SANTHANAM AND SHANHUI FAN
PHYSICAL REVIEW B 93, 161410(R) (2016)

FIG. 1. (a) The upper diagram depicts the net inflow of photons from a hot surface giving rise to a reverse current in a conventional p-i-n photodiode. The lower diagram indicates that the net outflow of photons to a cold surface can result in a positive current. (b) Current-voltage characteristics for a photodiode exposed to an emissive surface with temperature below (blue), at (black), and above (red) that of the diode. The green shaded boxes represent the potential for electrical power generation in the two nonequilibrium scenarios.

emitted by the diode. Since the input energy is simply the sum of these two quantities, we may express the thermal-to-electrical energy conversion efficiency as

$$\eta = \frac{-qV}{-qV + \hbar\omega}.$$  

(2)

We compute the efficiency \(\eta\) when the emissivity of the diode is in the narrow-band limit [i.e., \(\epsilon(\omega, \theta)\) is only nonzero within a small window of frequencies around some \(\omega_0\)] [3]. In this limit, setting \(J = 0\) in Eq. (1) yields the open-circuit voltage

$$V_{oc} = \frac{\hbar\omega_0}{q} \left(1 - \frac{T}{T_{surf}}\right).$$  

(3)

Near this open-circuit operating point then, \(\eta\) approaches the Carnot efficiency.

$$\eta_{\max} = \frac{-qV_{oc}}{-qV_{oc} + \hbar\omega_0} = \frac{T - T_{surf}}{T} = \eta_{\text{Carnot}}.$$  

(4)

In principle, the emissivity of a body can be tailored with a variety of photonic structures [4–9] which could be utilized to approximate this narrow-band limit. However, because the maximum power point can be far from the open-circuit condition, a wide emissivity spectrum may prove more ideal in practice if the objective is to maximize the power generated.

We now present an experimental demonstration of the physics of a photodiode operating under negative illumination. In particular, we seek to demonstrate that when a photodiode’s field of view is covered by a highly emissive surface, the sign of the temperature difference \((T_{surf} - T)\) determines the sign of the short-circuit current density \(J(V = 0)\) as portrayed in Fig. 1. Consequently when \((T_{surf} - T) < 0\) (i.e., when the photodiode is placed under negative illumination) in the presence of a finite resistive load connected to the diode’s terminals, the diode will operate in the second quadrant and therefore generate electrical power to drive the load.

The experimental setup is depicted in Fig. 2. An uncooled HgCdZnTe photodiode (Vigo’s PVI-3TE-6), whose room-temperature band gap \(E_{gap}\) is 218 meV, is set facing a temperature-controlled infrared-emissive surface (Acktar’s Metal Velvet) [10–12]. A hyper-hemispherical GaAs solid immersion lens is attached to the photodiode’s detecting surface so that the field of view is constrained to a cone with an 18° half-angle. The area of the diode’s optical cross section is a 1 × 1 mm² square. The emissive surface is placed close enough to the photodiode that it covers the diode’s entire field of view. The temperature of the emissive surface is

FIG. 2. Diagram of the experimental setup. An infrared photodiode was exposed to a high-emissivity surface which covered its entire field of view. A chopper was used to allow a lock-in measurement of the photocurrent.
varied from 285 to 305 K while the temperature of the diode remains at 295 K, and the diode’s short-circuit photocurrent is measured.

In order to characterize the photocurrent as a function of emissive surface temperature, a lock-in technique is employed. An optical chopper (415 Hz) is introduced between the diode and surface so that only during the open phase of the chopper is the diode optically exposed to the emissive surface. The diode is connected directly to a transimpedance amplifier (TIA) whose input impedance is negligible due to a virtual null. The photocurrent is amplified with a total transimpedance gain of $10^6$ V/A; it also passes through broadband low-pass and high-pass filters to keep the amplifier from overloading. This amplified signal, two examples of which are shown in the main plot of Fig. 3, is then read in by an analog-to-digital converter at $10^5$ samples/s for a duration of 1 s; this digital signal is subsequently post-processed using a fast Fourier transform (FFT). Note that both of the raw data signals in Fig. 3 resemble zero-mean square waves because their dc components are filtered out to avoid saturating the TIA.

Before we examine the quantitative data produced from the FFT, it is useful to examine the raw data in Fig. 3 qualitatively. When the emissive surface is hot, as in the case of the red curve, the photocurrent seen during the open phase of the measurement is approximately 150 nA below its value during the closed phase. During the open phase, the photodiode experiences a net influx of photons from the hot surface; during the closed phase, it sees only the chopper blade with which it is in equilibrium. Thus we can infer that the effect of the hot surface is to produce a reverse current of 150 nA. By contrast when the emissive surface is cold, as in the case of the blue curve, the photocurrent seen during the open phase of the measurement is about 130 nA above its value during the closed phase. Since the diode sees the cold surface during the open phase but only the chopper blade during the closed phase, we may infer that the effect of the cold surface is to produce a forward current of 130 nA. The distinctive $180^\circ$ phase shift seen in Fig. 3 as $T_{surf}$ changes from above $T$ to below $T$ is the key feature of the measurement.

The short-circuit current $I_{sc}$ was computed as the averaged peak-to-peak amplitude of the photocurrent square wave in Fig. 3 using an FFT algorithm. The zero-bias resistance of the diode was measured at equilibrium to be $R_{ZB} = 220 \Omega$, and the maximum extractable power is $P = \frac{1}{4} I_{sc}^2 R_{ZB}$. $I_{sc}$ and $P$ are plotted versus emissive surface temperature in Fig. 4. When the emissive surface was colder than the diode, a positive photocurrent was seen, but when the emissive surface was hotter than the diode, a negative photocurrent was observed. In both cases the extractable electrical power is a positive quantity, indicating that the diode can generate electricity from a temperature difference of either sign.

FIG. 4. (a) Short-circuit current and (b) extractable power versus temperature of the emissive surface facing the diode. The discrete points are the results of a lock-in measurement experiment described in the text. The solid line represents a model including nonidealities.
where $A_{\text{optical}}$ is the effective optical area of the diode. $R_{\text{shunt}}$ and $R_{\text{series}}$ offer a convenient way to capture the effects of nonradiative recombination, finite carrier mobility, and finite contact resistance. $R_{\text{shunt}}$ is equal to the product of the zero-bias resistance of an ideal diode of the same band gap and the ratio of the radiative recombination lifetime to the nonradiative lifetime. $R_{\text{series}}$ is equal to the difference between the zero-bias resistance of the full device and the parallel combination of $R_{\text{shunt}}$ with the zero-bias resistance of an ideal diode. A more complete model of electron transport in a photodiode can be reduced to these two quantities insofar as the total extractable power is concerned. This is possible largely because unlike diodes under positive illumination, the maximum power point under negative illumination is at sufficiently low voltage that the current-voltage characteristic can be linearized.

Although it is possible to directly model a complete device and thereby generate values for $R_{\text{shunt}}$ and $R_{\text{series}}$ for use in Eq. (5), doing so falls outside the scope of this work. Instead, $R_{\text{shunt}}$ and $R_{\text{series}}$ for the experiment presented previously were determined by the responsibility implied by the short-circuit current measurements in Fig. 4(a) and an independent measurement of the zero-bias resistance of the diode. For the diode used in the experiment then $A_{\text{optical}} = 0.1$ mm$^2$, $R_{\text{shunt}} = 23.4$ $\Omega$, and $R_{\text{series}} = 197$ $\Omega$. Note that the use of a solid immersion lens results in an $A_{\text{optical}}$ exceeding the diode’s true physical area. As seen in Fig. 4, the model agrees well with our measurements, including the asymmetry of extractable power between positive and negative temperature differences.

The model in Eq. (5) can be used to evaluate and compare new designs in terms of their values for $R_{\text{shunt}}$ and $R_{\text{series}}$. For an ideal device, $R_{\text{series}} = 0$ and $R_{\text{shunt}} = \infty$.

In the case of the nonideal diode in our experiment, operating at 295 K while being optically exposed to a 285 K emissive surface led to just over 1 $\text{pW}$ of extractable electrical power. In principle, one could redesign the structure so its electrical contacts are much less than a diffusion length from all relevant radiative recombination. Such a change would lead to a nearly ideal short-circuit current and could be captured by our model via a significantly reduced $R_{\text{series}}$. If $R_{\text{series}}$ could be effectively eliminated while $R_{\text{shunt}}$ remained unchanged, our model predicts that a square decimeter of a 295 K surface exposed to the coldness of space at 3 K could generate more than a microwatt of electrical power.

The shunt parasitic $R_{\text{shunt}}$ in our model primarily reflects the presence of Auger recombination and its time-reversed process electron impact ionization (sometimes called Auger generation). The presence of these nonradiative processes is a fundamental aspect of any diode in any semiconductor with a narrow direct band gap. In $p$-type $\text{Hg}_1-x\text{Cd}_x\text{Te}$, the Auger lifetime at low excitation is shorter than the radiative lifetime by a factor of $O(100)$ and accounts for the low $R_{\text{shunt}}$ measured in our experiment [13].

Moderate improvements of the model parameter $R_{\text{shunt}}$ may be possible using techniques which have not been fully exploited here. One well-established method to reduce the relative strength of Auger processes is to employ quantum wells [14]. Potential also exists to exploit the Purcell effect using optical concentration in space and frequency; by enhancing the strength of radiative interactions in this way the relative importance of nonradiative processes can be diminished [15]. The use of a solid immersion lens in our experiment represents a step in this direction [16]. Furthermore, certain material systems under development are promising candidates for more efficient devices. For example, III-V semiconductor alloys with dilute concentrations of nitrogen are known to have less Auger generation and recombination due to the increased effective mass of conduction-band electrons [17].

One potential implication of the effect observed here lies in the extraction of power from the cold dark universe [2,18–30]. Byrnes et al. [2] considered a power extraction scheme in which a heat engine extracts work from between the $T = 300$ K ambient and a radiatively cooled blackbody facing a 3 K cold reservoir. In the ideal limit, in which the heat engine is a Carnot engine, this system has a maximum power density of 48.4 W/m$^2$. We emphasize that this result does not represent the fundamental limit of power extraction. This may be understood by noting the constraint in Ref. [2] that the radiatively cooled surface is a classical blackbody. Instead, if the body radiating into the 3 K cold reservoir is composed of several gray bodies with narrow emissivity spectra and the temperature of each body is then optimized independently, the maximum power density is 55.0 W/m$^2$, which defines a multispectral limit [31]. If the negative illumination effect observed in this Rapid Communication were exploited to extract power from the same two thermal reservoirs, Eq. (1) predicts a maximum power density of 54.8 W/m$^2$ in the ideal limit. This figure significantly exceeds that of the equilibrium isothermal configuration from Byrnes et al. and in fact is fairly close to the ideal multispectral limit.

In conclusion, we have analyzed a method for extracting electrical work from a hot body radiating into a cold optical far field, such as the Earth radiating into outer space, using a long-wavelength semiconductor photodiode. In the ideal case we find that the thermal-to-electrical energy conversion efficiency approaches the Carnot limit and the maximum extractable power density from a 300 K object is found to be $\approx 55$ W/m$^2$. An experimental test of the basic principle yielded good agreement with theory. The conceptual principle provides an opportunity to tap into a major unused source of renewable energy. Improved devices with large emitting areas could have implications for off-grid power generation at night and waste-heat recovery.

This work was supported by the DOE “Light-Material Interactions in Energy Conversion” Energy Frontier Research Center under Grant No. DE-SC0001293.


