Hot-Electron Effect in Superconductors and Its Applications for Radiation Sensors

Introduction

The term "hot electrons" was originally introduced to describe nonequilibrium electrons (or holes) in semiconductors (for a review see, e.g., Ref. 1). The term encompasses electron distributions that could be formally described by the Fermi function but with an effective elevated temperature. The concept is very fruitful for semiconductors, where the mobility of carriers can be shown to depend on their effective temperature. In metals, however, electrons do not exhibit any pronounced variation of the mobility with their energy. As a result, heating of electrons in a metal does not affect the resistance,[2] unless the change in the effective temperature is comparable with the Fermi temperature.

Schklovski[3] was the first to discuss the idea of combining the steady-state electron heating with the strong dependence of the resistance on the effective electron temperature in a metal film undergoing the superconducting transition. In the steady-state regime, however, electron heating is always masked by the conventional bolometric effect; therefore, experimental results on the heating of electrons by the dc current were not very convincing. The regime of dynamic electron heating by external radiation was studied in a series of experimental and theoretical works.[4–6] It was immediately realized that the very short relaxation time of electron excitations would make it feasible to design extremely fast radiation sensors with a sensitivity much better than that of conventional bolometers.

During the last decade, a new generation of hot-electron superconducting sensors has been developed. These include submillimeter and THz mixers, direct detectors, and photon counters for the broad spectral range from microwaves to optical radiation and x rays. Activity in the field of hot-electron superconducting sensors is growing rapidly. These sensors have already demonstrated performance that makes them devices-of-choice for many far-infrared (THz), infrared, and optical wavelength applications, such as plasma diagnostics, laser studies, ground-based and airborne heterodyne astronomy, and single-photon-detection and quantum communications. Parallel development of compact cryocoolers and THz radiation sources opens hot-electron sensors for satellite astronomy and communication applications. This article reviews the physical background of the hot-electron phenomenon in superconducting films and discusses various technical realizations of hot-electron radiation sensors.

Physics of Hot Electrons

Thermal dynamics in a superconducting film on a dielectric substrate can be thought of in terms of four co-existing subsystems: Cooper pairs, quasiparticles (electrons from broken Cooper pairs), phonons in the film, and phonons in the substrate. Thermal equilibrium exists when all of these can be described by equilibrium distribution functions with the same temperature. If any distribution does not satisfy these conditions, the situation is considered nonequilibrium. General treatment of a nonequilibrium state requires solution of the integral kinetic equations for space- and time-dependent distribution functions. To avoid the above complexity, various simplifying assumptions are used to reduce the general problem to analytically solvable rate equations.

1. Hot-Electron Cooling and Diffusion

The hot-electron model is most relevant for nonequilibrium superconductors maintained at temperature T near the superconducting transition T_c, where quasiparticles and phonons can be described by thermal, normal-state distribution functions, each with its own effective temperature. The electron and phonon effective temperatures (T_e and T_p) are assumed to be established instantly and uniformly throughout the whole specimen. This assumption implies that a rapid thermalization mechanism exists inside each subsystem.

The main steps of the hot-electron phenomenon that lead to the global equilibrium are depicted in Fig. 87.27. Introducing characteristic times of the energy exchange between subsystems reduces the problem of the global equilibrium recovery to a pair of coupled heat-balance equations for T_e and T_p. The intrinsic thermalization time τ_T should be short compared to energy exchange times. This two-temperature (2-T) approach was used for the first time by Kaganov *et al.*[2] to describe

steady-state electron heating in metals. Below T_c, the electron specific heat exhibits an exponential temperature dependence that makes equations nonlinear for even small deviations from equilibrium. The description can, however, be simplified in the vicinity of T_c. At this temperature the superconducting energy gap is strongly suppressed, concentration of Cooper pairs is very small, and unpaired electrons exhibit no significant superconducting peculiarities: they are regarded as normal electrons having the ordinary Fermi distribution function. In the normal state, the specific heat of electrons has a much weaker temperature dependence, which can be neglected for small deviations of T_e from the equilibrium. With these assumptions, the equations describing the hot-electron effect in superconductors become linear and can be written as

$$\frac{dT_e}{dt} = -\frac{T_e - T_p}{\tau_{ep}} + \frac{1}{C_e} W(t), \tag{1a}$$

$$\frac{dT_p}{dt} = \frac{C_e}{C_p}\frac{T_e - T_p}{\tau_{ep}} - \frac{T_p - T_0}{\tau_{es}}, \tag{1b}$$

where $W(t)$ represents the external perturbation (i.e., the power per unit volume absorbed by the electron subsystem); τ_{ep} and τ_{es} are the electron energy relaxation time via electron–phonon interaction and the time of phonon escape into the substrate; C_e and C_p are the electron and phonon specific heats, respectively;

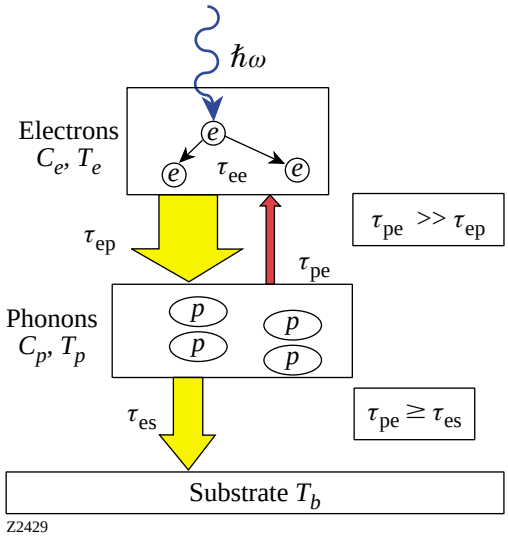

Figure 87.27
Thermalization scheme showing various channels of the energy transfer in a hot-electron device that relaxes toward global equilibrium.

and T_0 is the ambient (substrate) temperature. To derive the 2-T equations we used the condition of the energy-flow balance in equilibrium $\tau_{pe} = \tau_{ep}\left(C_p / C_e\right)$, where τ_{pe} is the phonon–electron energy relaxation time.

The first implementation of the electron-heating model to superconductors was made by Shklovski,[3] who used a more general, nonlinear form of the heat-balance equations to describe hysteresis of the critical current in a thin lead film. An analytical solution of Eq. (1) was first obtained for sinusoidal perturbations by Perrin and Vanneste[4] and for an optical pulse excitation by Semenov *et al.*[5] In the latter case, thermalization of electrons was interpreted as an increase of T_e. The increase was assumed to occur during a time interval that depended on both the duration of the optical pulse and the intrinsic thermalization time τ_T. The model was used to describe the response of superconducting NbN and $YBa_2Cu_3O_{7-\delta}$ (YBCO) films in the resistive state to near-infrared and visible radiation.[5,7] Figures 87.28 and 87.29 show a good agreement between experimental signals and the theoretical simulation.

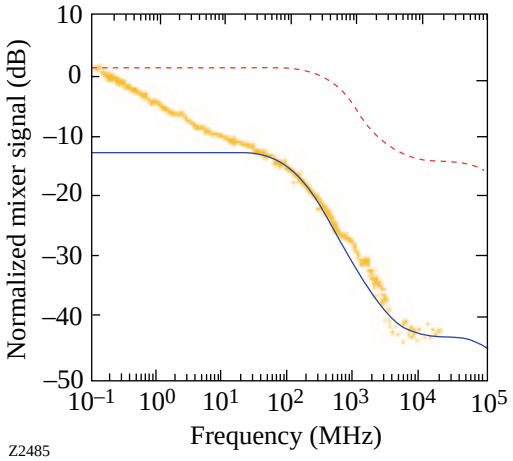

Figure 87.28
Response of a YBCO hot-electron photodetector (HEP) to optical radiation (dots) versus modulation frequency (Ref. 7). The solid line was calculated using Eqs. (1). The discrepancy at low frequencies is due to phonon diffusion in the substrate that was not accounted for in the model. The dashed line represents the thermal model.

Figure 87.30 presents the detailed thermalization diagrams for both YBCO [Fig. 87.30(a)] and NbN [Fig. 87.30(b)] thin films exited by 100-fs-wide optical pulses. The diagrams depict the process in the same manner as Fig. 87.27 but now include the actual values of the characteristic time constants for both materials. The values were obtained from the 2-T model via the fit of Eqs. (1) to the experimental photoresponse

data. The measurements were performed using the electro-optic sampling system, which allowed obtaining the intrinsic, time-resolved dynamics of the electron thermalization process in 3.5-nm-thick NbN[8] and 100-nm-thick YBCO films.[9] We note that, in general, the dynamic of the YBCO thermalization is roughly one order of magnitude faster than that of NbN. In both cases, the energy flow from electrons to phonons dominates the energy backflow due to reabsorption of nonequilibrium phonons by electrons; however, while the energy backflow in YBCO can be neglected because of the very large ratio $C_p/C_e = 38$, in NbN it constitutes a non-negligible 15% $\left(C_p/C_e = 6.5\right)$ of direct electron–phonon energy relaxation. Consequently, in YBCO film excited on the femtosecond time scale, the nonthermal (hot-electron) and thermal, bolometric (phonon) processes are practically decoupled, with the former totally dominating the early stages of electron relaxation. On the other hand, the response of NbN devices is determined by the "average" electron cooling time τ_e, which is given by $\tau_{ep} + \left(1 + C_e/C_p\right)\tau_{es}$[4,5] and corresponds to the time that elapses from the peak response until the magnitude of the response declines to $1/e$ of the maximum value. If the external perturbation is substantially longer than τ_{pe} (that is, >100 ps for YBCO films), the YBCO response is dominated by the bolometric process, as was shown by the bulk of the early photoresponse measurements.[10] The very large difference in the τ_{es} values for YBCO and NbN is mainly due to the drastic difference in thickness of the tested films. Additionally, ultrathin NbN films are a better acoustical match to the substrate. This significantly reduces τ_{es}.

Electron heating in the limiting case of a very short phonon escape time, $\tau_{es} << \tau_{ep}, \tau_{pe}$, was first studied by Gershenzon *et al.*[6] for Nb films. Although for this material[11] $C_p/C_e \approx 0.25$ and, consequently, $\tau_{ep} > \tau_{pe}$, the effective escape of phonons to the substrate prevents energy backflow to electrons. As a result, τ_{ep} alone controls the response of ultrathin (<10-nm) Nb films. Typical electron relaxation time in Nb is ≈1 ns at 4.2 K, which is over an order of magnitude larger than in NbN.

The 2-T model represented by Eqs. (1) is essentially the small-signal model. Deviations of the effective temperatures from equilibrium due to both the joule power dissipated by the bias current and absorbed radiation power are assumed small compared to their equilibrium values. The theory of operation of a hot-electron photodetector (HEP) was developed on the basis of this model by Gershenzon *et al.*,[12] and a novel hot-electron mixer (HEM) was proposed.[12,13]

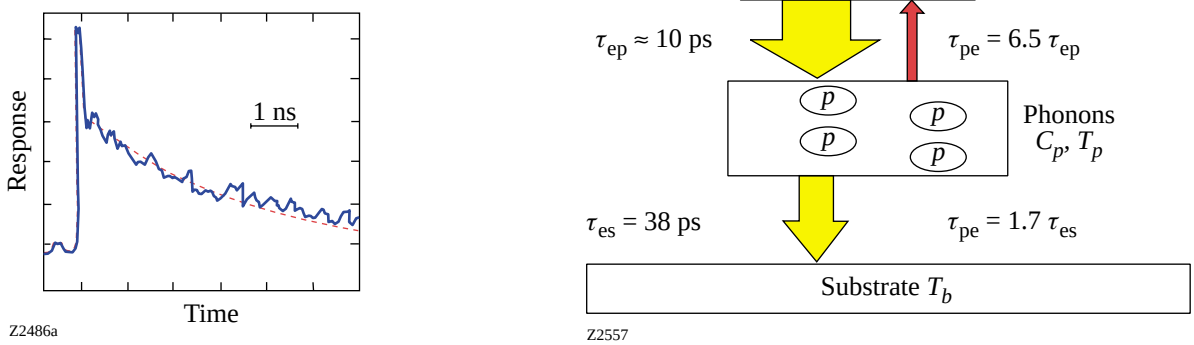

Figure 87.29
Response of a YBCO HEP to a femtosecond infrared pulse: experimental data (solid line) (Ref. 5) and simulations (dashed line) based on the 2-T model.

Figure 87.30
Hot-electron relaxation diagrams and characteristic times for (a) thin-film YBCO (Ref. 9) and (b) ultrathin NbN film (Ref. 8).

The 2-T approach neglects, however, diffusion of electrons and assumes that the effective temperatures remain uniform within the whole device. A different approach was proposed by Prober,[14] who considered diffusion of hot electrons out of the active area, rather than the energy transfer to phonons, as the main mechanism of the electron cooling. If the device length L is short compared to the thermal diffusion length $L_{th} = (D\tau_e)^{1/2}$, where τ_e is the electron cooling time and D is the electron diffusivity, relaxation of T_e is controlled by the electron out-diffusion time $\tau_d = \left(L^2/\pi^2 D\right)$. In the limiting case $L << L_{th}$, T_e remains almost uniform through the device length. The device can then be described by Eq. (1a), in which τ_{ep} and T_p should be substituted for τ_d and T_0, respectively. For longer devices, both the actual distribution of T_e along the device length and the phonon contribution to the electron relaxation should be taken into account.

2. Large-Signal Models

The common disadvantage of the small-signal model described above is that the optimal values of the bias current (for HEP's) and power of the local oscillator (for the HEM theory) are not derived in the framework of the model, but rather taken from the experiment or independently estimated. To include the bias current and the local oscillator (LO) power in a consistent manner, one should specify the structure of the resistive state and account for the dependence of the electron-cooling rate on the deviation from the equilibrium. For large deviations from equilibrium, heat-balance equations become nonlinear.

The large-signal mixer theory was developed by Nebosis *et al.*[15] for the uniform resistive state (which is, of course, a very crude approximation). The authors assumed a finite value of τ_{es} and introduced the superconducting critical current. Reasonable quantitative agreement (see Fig. 87.31) was found between the experimental data for NbN mixers and the theoretical results. Karasik *et al.*[16] implemented a similar approach for modeling a bolometric mixer fabricated from a high-temperature superconducting material. Floet *et al.*[17] considered the nonuniform resistive state of a hot-electron bolometer in the small-signal regime for $\tau_{es} = 0$, while Merkel *et al.*[18] developed the large-signal nonlinear model for a finite, nonzero value of τ_{es}. Both models described the resistive state of the mixer at optimal operation conditions in terms of a normal hot spot, maintained by self-heating. The hot spot occupies only a portion of the device length, thus assuring a mixer resistance between zero and the normal-state value. In this approach, the LO power is assumed to be uniformly absorbed in the mixer, whereas the joule power dissipation due to the bias current appears in the hot-spot region only. Since the diffusion of electrons is introduced in the basic equations, this model naturally covers all intermediate cases between the extreme diffusion cooling ($L << L_{th}$) and phonon-cooling ($L >> L_{th}$) regimes. Neglecting phonons ($\tau_{es} = 0$) and simulta-

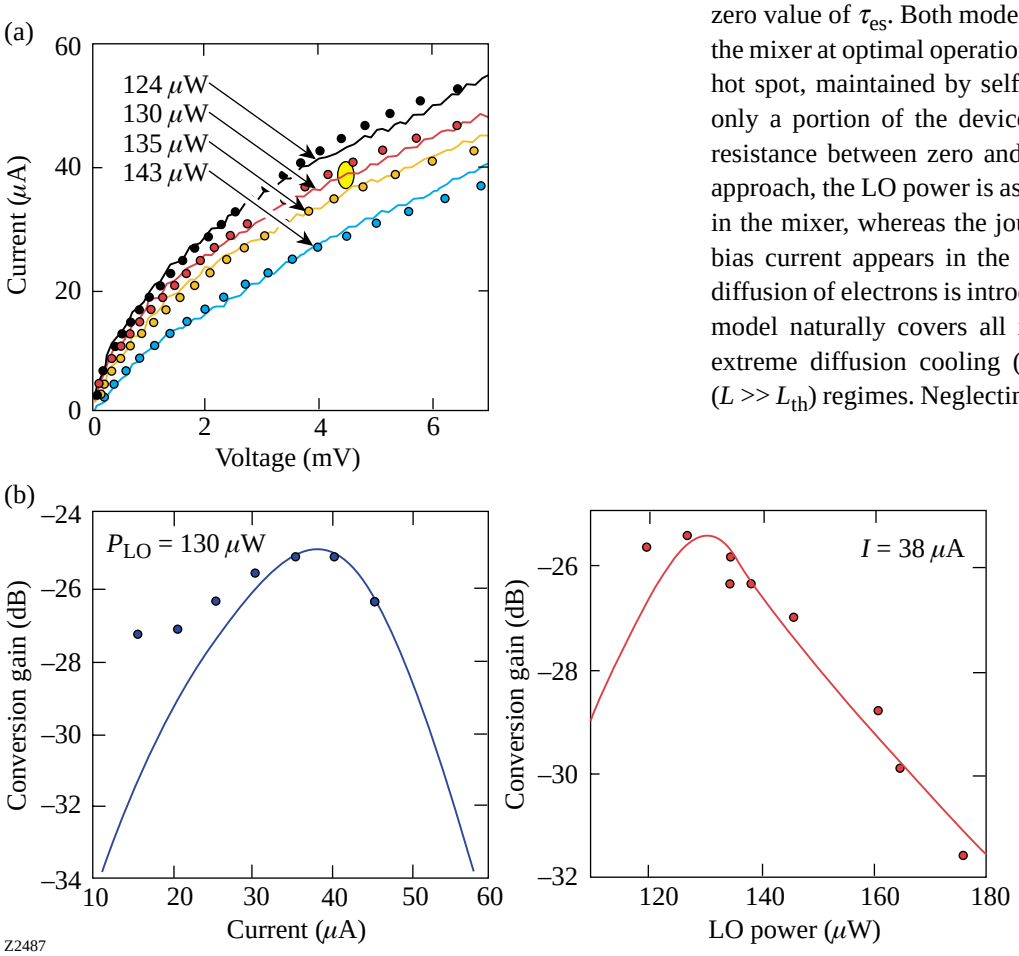

Figure 87.31
(a) Current–voltage characteristics for different LO power values. (b) Conversion gain curves for a NbN HEM compared with results of the uniform model (solid lines) (Ref. 15).

neously assuming $\tau_T = 0$, one can reduce the problem to the following system of equations[17] for T_e:

$$-K\frac{d^2T_e}{dx^2}+\frac{C_e}{\tau_e}\left(T_e-T_0\right)=j^2\rho_n+P_{\mathrm{RF}}\quad(\text{inside hot spot}),$$
$$-K\frac{d^2T_e}{dx^2}+\frac{C_e}{\tau_e}\left(T_e-T_0\right)=P_{\mathrm{RF}}\qquad(\text{outside hot spot}), \tag{2}$$

where K is the thermal conductivity, j is the bias current density, ρ_n is the resistivity of the mixer in the normal state, and P_{RF} is the LO power absorbed per unit volume. This description allows for an analytical solution, which returns the bias current as a function of the hot-spot length, and, thus, a voltage drop across the device. Results of simulations[17] are in good agreement with the experimental current–voltage (I–V) characteristics, especially for large P_{RF} values, which drive the mixer almost into the normal state. Surprisingly, results based on not only Eq. (2), but even on the more-accurate numerical model[18] shown in Fig. 87.32, do not differ much from simulations based on the uniform 2-T model (Fig. 87.31). With the appropriate set of fitting parameters, both approaches describe fairly well the I–V characteristics of the HEM and predict reasonable values of the conversion efficiency and noise temperature.

A nonthermal regime of the diffusion-cooled HEM was described by Semenov and Gol'tsman.[19] The authors considered a short device made from a clean material, in which τ_T is larger than τ_d. The device operated in the nonthermal regime and had the advantage of a short response time (or, equivalently, a large bandwidth) in the heterodyne mode. On the other hand, incomplete thermalization hampered the responsivity and increased the relative contribution of the Johnson noise to the total electric noise of the device. Compared to HEM's operated in the thermal regime, the nonthermal mixer required more power from LO. At low temperatures, however, the nonthermal regime of operation provided almost quantum-limited sensitivity.

The electric noise of a hot-electron sensor is comprised of the same components as the noise of any conventional bolometer: shot noise, Johnson noise, thermal noise, and flicker noise. To our knowledge, there is no consistent theory for flicker noise, so its contribution may be determined only experimentally. Unless the sensitivity of the bolometer reaches the quantum limit, the noise due to fluctuations in the background radiation can be neglected. The typical length of hot-electron devices studied so far was much larger than the diffusion length associated with the electron–electron scattering. In this limiting case, the superposition of Johnson noise and shot noise reduces to the Nyquist form, i.e., the spectral density of the voltage noise is $S_V = 4\,k_BT\,R$, where R is the resistance of the device. This noise has a "white" spectrum up to very high frequencies. The corresponding contribution to the system-noise temperature in the heterodyne regime increases rapidly when the conversion efficiency rolls off at intermediate frequencies (IF's) larger than $1/\tau_e$.

Thermal noise contributes to the total spectral density the amount $4k_BT^2I^2\,\tau_e\left(\partial R/\partial T_e\right)^2/\left(C_ev\right)$, where I is the bias current and v is the volume of the sensor. Since the conversion efficiency is proportional to $I^{-2}\tau_e^2P_{\mathrm{RF}}\left(\partial R/\partial T_e\right)^2/\left(C_ev\right)^2$ and

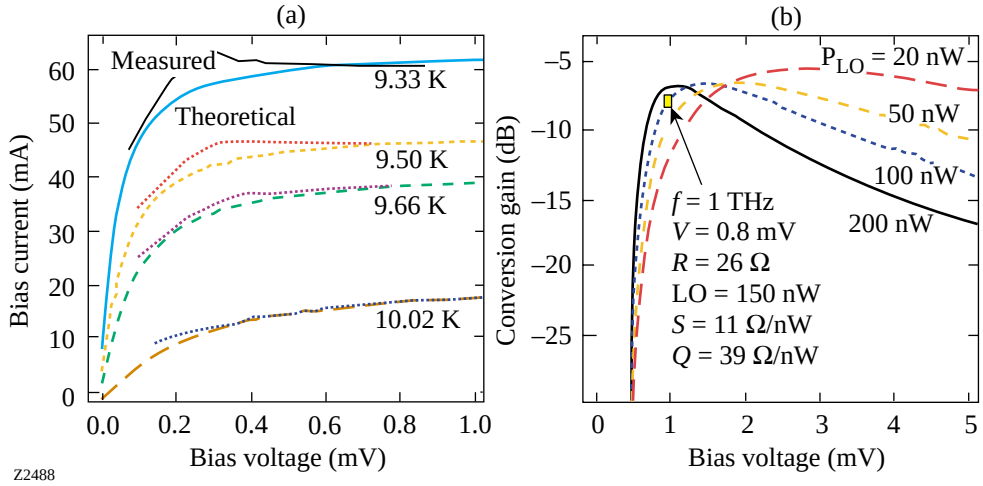

Figure 87.32
(a) Current–voltage characteristics and (b) conversion gain of a NbN HEM simulated in the framework of the hot-spot model (Ref. 18). Experimental characteristics are shown for comparison.

has the same roll-off frequency, the noise temperature of the mixer due to thermal fluctuations is given by $T_N \approx T_e^2 C_e v/\left(\alpha \tau_e P_{\mathrm{RF}}\right)$, where α is the optical coupling efficiency. The contribution to the noise temperature due to thermal fluctuations does not depend on the intermediate frequency; neither does the corresponding noise-equivalent power (NEP) in the direct-detection mode,

$$\mathrm{NEP} \approx \left(T_e/\alpha\right)\left(k_B C_e\, v/\tau_e\right)^{1/2}.$$

On the contrary, the contribution due to the Nyquist term increases rapidly at IF's larger than $1/\tau_e$ and usually limits the IF noise bandwidth of the mixer.

Though the above simple treatment of the bolometer noise explains the main features, it does not provide an appropriate tool for computations. To obtain exact results, one should take into account the positive feedback via the load resistor and self-heating by the bias current. The former enhances the system output noise because the bolometer rectifies part of its own noise voltage drop across intrinsic resistance. The latter effect typically increases the IF bandwidth in the heterodyne regime and decreases the response time in the direct-detection mode. It is of little practical use, however, because operation in the vicinity of the thermal roll-off requires very precise stabilization of the ambient temperature. For a HEM with dc resistance R at the operation point and connected to the IF load with the impedance R_L, the dependence of the conversion efficiency $\eta(\omega)$ and single-sideband noise temperature $T_{\mathrm{SSB}}(\omega)$ on the IF was derived in the framework of the uniform model[15]

$$\eta(\omega) = \frac{2\alpha}{I^2}\frac{R_L}{\left(R_L + R_\infty\right)^2}\frac{C^2 P_{\mathrm{RF}}}{\left(C\dfrac{R - R_L}{R_L + R_\infty} + \xi\right)^2 + \varphi^2}, \quad (3)$$

$$T_{\mathrm{SSB}}(\omega) = \frac{2\,T_e\,R_\infty I^2}{\alpha C^2 P_{\mathrm{RF}}}\left(\xi^2 + \varphi^2\right) + \frac{2T_e^2\,\tau_e}{C_e V \alpha\, P_{\mathrm{RF}}}, \quad (4)$$

where

$$C = I^2 \tau_e \frac{\partial R/\partial T_e}{C_e V},$$

$$\xi(\omega) = \frac{1 + \omega^2\left(\tau_1\tau_3 + \tau_2\tau_3 - \tau_1\tau_2\right)}{1 + \left(\omega\,\tau_3\right)^2},$$

$$\varphi(\omega) = \frac{\omega\left(\tau_1 + \tau_2 - \tau_3\right) + \omega^3\tau_1\tau_2\tau_3}{1 + \left(\omega\,\tau_3\right)^2},$$

$$\tau_{1,2}^{-1} = \frac{1}{2}\left(\frac{1}{\tau_{\mathrm{ep}}} + \frac{C_e}{C_p\tau_{\mathrm{ep}}} + \frac{1}{\tau_{\mathrm{es}}}\right)$$

$$\times\left[1 \pm \sqrt{\frac{\dfrac{4}{\tau_{\mathrm{ep}}\tau_{\mathrm{es}}}}{\left(\dfrac{1}{\tau_{\mathrm{ep}}} + \dfrac{C_e}{C_p\tau_{\mathrm{ep}}} + \dfrac{1}{\tau_{\mathrm{es}}}\right)^2}}\right],$$

and

$$\tau_3 = \frac{\tau_{\mathrm{ep}}\tau_{\mathrm{es}}}{\tau_e}.$$

In the above equations, R_∞ is the impedance of the bolometer at very high IF, and $\partial R/\partial T_e$ is the slope of the superconducting transition at the operation point on the scale of the electron temperature. The slope of the transition cannot be derived from first principles in the framework of the uniform model. Its temperature dependence should be calculated in a phenomenological manner (see, e.g., Ref. 15), or the value at the specific operation regime should be concluded from the experiment. Ekström *et al.*[20] showed that the magnitude of the parameter C in Eqs. (3) and (4) can be determined from the experimental dc I–V characteristic as

$$C = \frac{\dfrac{dV}{dI} - R}{\dfrac{dV}{dR} + R}, \quad (5)$$

where dV/dI is the differential resistance of the HEM at the operation point. The advantage of the hot-spot model[18] is that it allows for numerical computation of the superconducting

transition slope for arbitrary values of the LO power, bias current, and ambient temperature.

3. Cooper-Pair, Kinetic–Inductive Photoresponse

Although the response of a superconductor that is kept well below T_c to external radiation cannot be adequately treated in the framework of the hot-electron approximation, we decided to include superconducting detectors operating at $T << T_c$ in our review. Rothwarf and Taylor[21] were the first to successfully develop the phenomenological description for nonequilibrium Cooper-pair recombination and breaking processes (so-called the RT model). At low temperatures, when energies of nonequilibrium quasiparticles after thermalization are spread over a narrow interval above the superconducting energy gap 2Δ, the appropriate parameters to characterize this nonequilibrium state are the number Δn_q of excess quasiparticles and the number Δn_p of excess, so-called, 2Δ phonons. The 2Δ phonons are emitted in the Cooper-pair recombination process and, since they have the energy of at least 2Δ, they are responsible for secondary breaking of Cooper pairs. For small perturbations, concentrations of Δn_q and Δn_p are given by the following linearized RT rate equations:

$$\frac{d}{dt}\Delta n_q = -\frac{\Delta n_q}{\tau_R} + \frac{2\Delta n_p}{\tau_B}, \tag{6a}$$

$$\frac{d}{dt}\Delta n_p = -\frac{\Delta n_p}{\tau_B} - \frac{\Delta n_p}{\tau_{es}} + \frac{2\Delta n_q}{\tau_R}, \tag{6b}$$

where τ_R and τ_B are the quasiparticle recombination time and the time of breaking Cooper pairs by 2Δ phonons, respectively. We note that Eqs. (6) are mathematically analogous to the 2-T model [Eqs. (1)]. Like the 2-T model, the RT approach assumes that there is a quick, intrinsic thermalization mechanism inside both the quasiparticle and phonon subsystems.

When photons with energy typically much larger than 2Δ are absorbed by a superconducting film maintained at $T << T_c$, they produce a time-dependent population $\Delta n_q(t)$ of nonequilibrium quasiparticles, leading to a temporary decrease in the superconducting fraction of electrons, $f_{sc} = \left(n_0 - n_q\right)\big/ n_0$, where $n_q = n_q(0) + \Delta n_q(t)$ is the instant concentration of the quasiparticles, $n_q(0)$ is their equilibrium concentration, and n_0 is the total concentration of electrons. Because the pairs are characterized by non-zero inertia, this process can be modeled as time-varying kinetic inductance:[22,23]

$$L_{kin}(t) = \frac{L_{kin}(0)}{f_{sc}}, \tag{7}$$

where $L_{kin}(0) = \mu_0\left(\lambda_L\right)^2 \big/ d$ is the equilibrium value per unit area of the film, λ_L is the magnetic penetration depth, and d is the film thickness. The change in time of L_{kin} in a current-biased superconducting film leads to a measurable voltage signal across the film edges.

For the limiting case of very fast thermalization, i.e., when τ_T is small compared to both τ_R and τ_B, the kinetic-inductive response was described by Semenov *et al.*[24] as the product of the analytical solution of Eqs. (6) and a fitting factor exponentially growing in time. The latter parameter corresponded to the multiplication cascade of quasiparticles during thermalization. The kinetic-inductive model describes well the experimental results obtained with pulsed and modulated cw excitations, for both the low-temperature superconductor (LTS) films (Fig. 87.33 and Ref. 24) and the high-temperature superconductor (HTS) films (Fig. 87.34 and Refs. 9 and 25).

4. Single-Photon-Detection Mechanisms

So far this discussion has been limited to integrating detectors in which the energy of a large number of absorbed photons is distributed among an even larger number of elementary thermal excitations in the detector. That is, individual photons cannot be distinguished, and only the average radiation power absorbed by the detector is measured. In the particular case of a thermal detector, e.g., a bolometer or a hot-electron detector near T_c, this average absorbed radiation power corresponds to enhanced effective temperatures of phonons and electrons, respectively. In a quantum (photon) detector, a single photon creates excitations that are collected and counted before they relax and before another photon is absorbed. Thus, the detector registers each absorbed photon, while the number of collected excitations measures the energy of absorbed photons.

The hot-electron quantum detector was first proposed by Kadin and Johnson.[26] In this model, a photon absorbed somewhere in the film initiates a growing hot spot. The resistance inside the hot spot is larger than in the surrounding area. Even if the size of the hot spot is much smaller than the size of the film, the voltage drop across the current-biased film "feels" the presence of the hot spot. The disadvantage of this approach for practical devices stems from the fact that the film is operated near its T_c and can withstand only a very small current density without being driven into the normal state. Since the

detector response is proportional to the bias current, the small operating current requires a complicated, SQUID-based read-out scheme.[27]

Semenov *et al.*[28] proposed a different quantum detection regime in a superconducting stripe that is operated well below T_c and carries a bias current only slightly smaller than the critical value at the operating temperature. Generation of a hot spot at the position where the photon has been absorbed creates a local region with suppressed superconductivity (normal region). The supercurrent is forced to flow around the normal (resistive) spot, through those parts of the film that remain superconducting. If the diameter of the resistive spot is such that the current density in the superconducting portion of the film reaches the critical value, a resistive barrier is formed across the entire width of the stripe, giving rise to a voltage pulse with the magnitude proportional to the bias current.

The physical difference of the quantum detection proposed in Ref. 28, as compared to Ref. 26, is that the resistive state and, thus, the response appear to be caused by the collaborative effect of the bias current and the hot-spot formation. In the hot

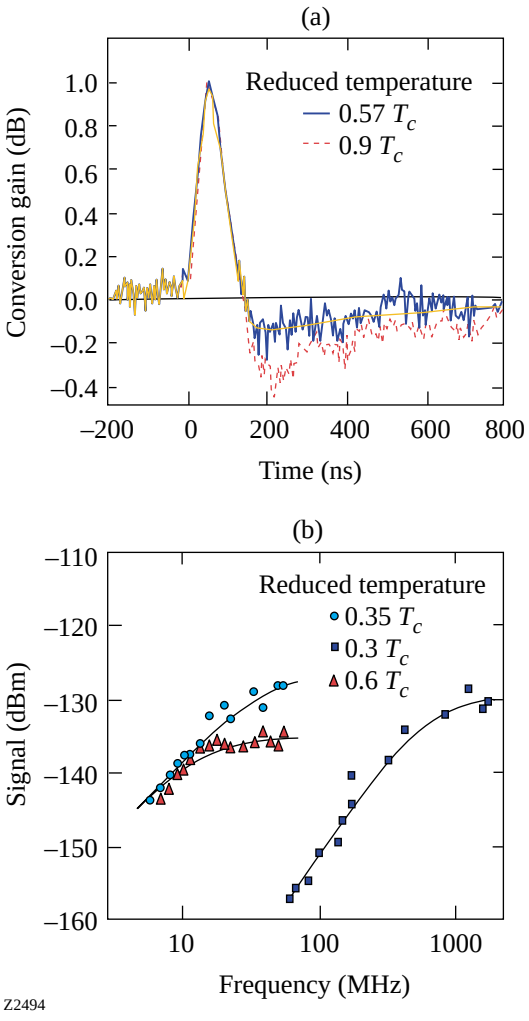

Figure 87.33
(a) Conversion gain and (b) signal response of a NbN HEP to pulsed and modulated cw optical radiation in comparison with model simulations based on Eqs. (6) (Ref. 24).

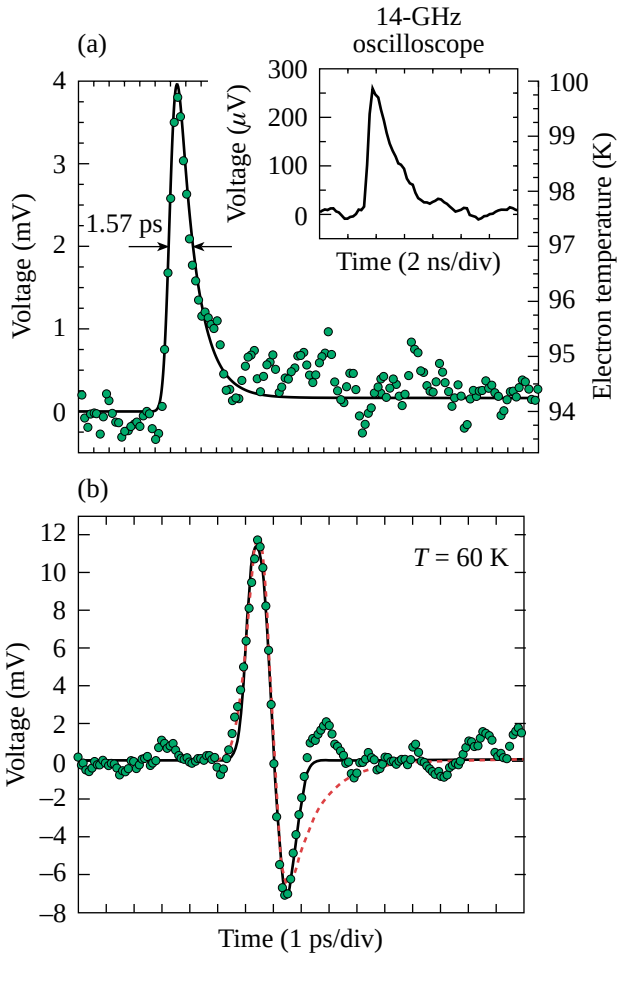

Figure 87.34
Experimental response (dots) of a YBCO HEP to 100-fs-wide optical pulses (Refs. 9 and 25). Simulated transients were obtained (a) with the uniform hot-electron model [Eqs. (1)] for the operation in the resistive state and (b) with the RT model [Eqs. (6)–solid line] and the 2-T model [Eqs. (1)–dashed line], for operation at low temperatures in the superconducting state. Inset in (a) shows a bolometric response.

spot, the nonequilibrium quasiparticle concentration increases due to hot-electron thermalization (multiplication) and decreases due to electron out-diffusion. The normal spot at the absorption site occurs when the concentration of nonequilibrium electrons exceeds the critical value corresponding to the local normal state. If the film thickness d is small compared to L_{th}, the concentration of nonequilibrium thermalized quasiparticles is given by

$$\frac{\partial}{\partial t}\Delta n_q = D\nabla^2 \Delta n_q + \frac{\Delta n_q}{\tau_e} + \frac{d}{dt}M(t), \tag{8}$$

where $M(t)$ is the multiplication factor and D is the normal-state electron diffusivity. The maximum value that $M(t)$ reaches during the avalanche multiplication process is called quantum yield or quantum gain; it is proportional to the energy of the absorbed quantum. Under assumptions that the $M(t)$ rate is much larger than the $1/\tau_e$ rate and that the photon is absorbed at $t = 0$ and $r = 0$, the solution for the time-dependent quasiparticle concentration profile takes the form

$$\Delta n_q(r,t) = \frac{M(t)}{4\pi Dd}\frac{1}{t}e^{-t/\tau_e}e^{-r^2/4Dt}. \tag{9}$$

The diameter of the normal spot is determined from the condition $n_q(0) + \Delta n_q(r,t) > n_0$. The maximum diameter of the normal spot increases with the quantum energy. The model[28] predicts an almost-Gaussian response pulse with a magnitude that, up to a certain extent, does not depend on the photon energy. On the other hand, the pulse duration is a function of the maximum spot size, providing the basis for spectral sensitivity of the device. Finally, the single-quantum detection regime should have a cutoff wavelength that depends on operating conditions (bias current and temperature) and the detector size. Since such a detector counts individual photons, it should have ultimate background-limited sensitivity through the whole range of operation conditions.

Gol'tsman *et al.*[29] experimentally demonstrated the supercurrent-assisted, hot-spot-detection mechanism for single optical (790-μm-wavelength) photons. Figure 87.35 shows a collection of "snapshots" recorded by a 1-GHz-bandwidth oscilloscope for different energies per laser pulse, incident on the NbN quantum HEP. Each snapshot presents an 80-ns-long record of the response to six successive 100-fs-wide pulses and was randomly selected out of a real-time detector output data stream. Trace A in Fig. 87.35 corresponds to an average of 100 photons per pulse hitting the detector. In this case, the HEP responded to each optical pulse in the laser train. The same 100%-efficient response was observed (trace B) when there were approximately 50 photons per pulse. As the incident laser intensity was further decreased (with other experimental conditions unchanged), the quantum nature of the detector response emerged. Instead of the linear decrease of the signal amplitude with incident light intensity, which is characteristic of a classical integrating detector, the response amplitude of the single-photon HEP remained nominally the same. In addition, some of the response pulses were missing because of the limited quantum efficiency of the device as well as fluctuations in the number of photons incident on the detector. The quantum voltage response of the HEP is most apparent in the bottom two pairs of traces: C and D (five photons/pulse) and E and F (one photon/pulse). Each pair corresponds to two different randomly selected records obtained under exactly the same experimental conditions. Note that in each case the detector response is very different. Averaging over a long observation time, however, showed that both the average number of captured pulses and their magnitude remained constant if the pulse energy was unchanged. This unambiguously demonstrated the single-photon operation of the device.

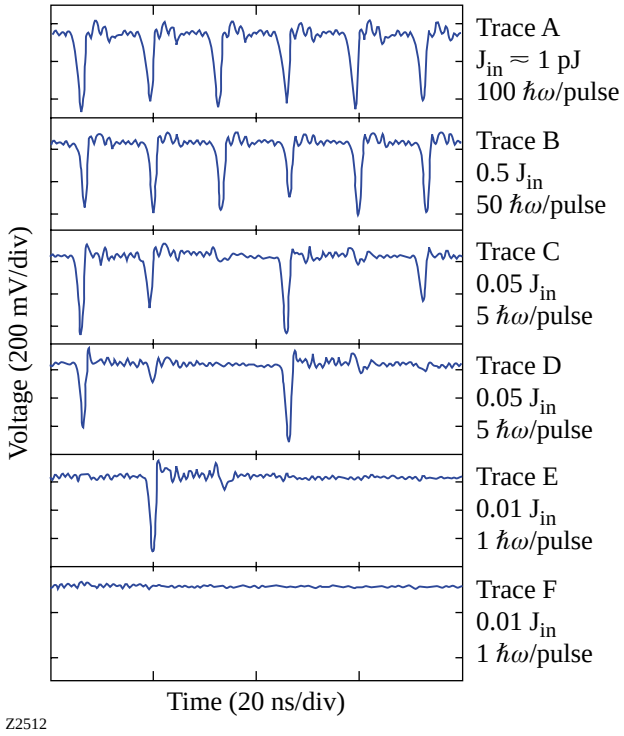

Figure 87.35
Response of a NbN quantum detector to trains of 100-fs optical pulses with a different number of photons per pulse (see text for details).

For a mean number of photons per pulse (m), the probability $P(n)$ of absorbing n photons from a given pulse is proportional to

$$P(n) \sim \frac{e^{-m}(m)^n}{n!}. \tag{10}$$

When the mean number of photons $m << 1$ (achieved, for example, by attenuating the radiation fluence to reduce the total number of photons incident on the detector to an average of much less than one photon per pulse),

$$P(n) \sim \frac{m^n}{n!}. \tag{11}$$

Consequently, for very weak photon fluxes, the probability of detecting one photon, two photons, three photons, etc., is

$$P(1) \sim m,\ P(2) \sim \frac{m^2}{2},\ P(3) \sim \frac{m^3}{6},\ \text{etc.} \tag{12}$$

Figure 87.36 plots the probability of the detector producing an output voltage pulse as a function of the number of photons per pulse, incident on the device area for two different values of the bias current. The left vertical axis indicates the mean number of detector counts per second. The right vertical axis corresponds to the probability of detecting an optical pulse. Open squares correspond to the bias current 0.92 I_c, where I_c is the critical current at the operation temperature. Saturation occurs at high incident photon fluxes. For smaller fluxes, as predicted by Eq. (11), the experimental data show the linear decrease of detection probability with the average number of incident photons over four orders of magnitude, clearly demonstrating the single-photon detection. At very low photon doses, experimental data points saturate at the level of 0.4-s^{-1} counts (probability 4×10^{-4}) since the experiment was performed in an optically unshielded environment. This level is regarded as the laboratory photon background. The solid squares in Fig. 87.36 correspond to the same device, operated under the same conditions as those for the solid-square data, but biased with 0.8 I_c. Experimental data points now follow a quadratic dependence of detection probability [see Eq. (12)], showing the two-photon detection. As expected for a two-photon process, the quantum efficiency is significantly lower than for the single-photon detection. At the same time, photon background is no longer observed since the probability of two uncorrelated, stray photons hitting the device within its response duration is negligibly small.

A nonequilibrium model of a single quantum x-ray detector with the readout via the superconducting tunneling junction was developed by Twerenbold.[30] Typically, a tunnel-junction detector consists of a relatively thick absorber film with an underlying thinner trapping layer, which forms one junction electrode. A photon captured in the absorber generates a high-energy photoelectron that relaxes via hot-electron multiplication into the energy gap of the absorber. Nonequilibrium quasiparticles excited during the cascade diffuse to the adjacent trapping layer, which has a smaller energy gap. There, quasiparticles scatter inelastically, reaching an energy level corresponding to the trapping-layer energy gap. The latter process is called "trapping" because it confines the charge to the region close to the tunnel barrier. The tunnel junction is externally biased in such a way that trapped quasiparticles can tunnel directly to the electrode characterized by the lower-energy gap. The same potential barrier prevents them from returning. They can, however, break Cooper pairs in the low-gap electrode and then form new pairs with unpaired electrons in their own electrode. Thus, the process returns unpaired electrons to the initial electrode, increasing the number of tunneling events per quasiparticle and providing intrinsic charge amplification. The time integral of the current transient

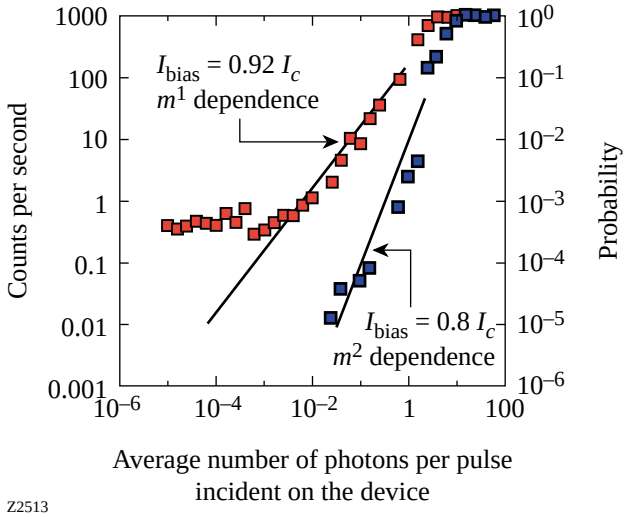

Figure 87.36
Count rates and the corresponding counting probability for a NbN quantum detector as a function of the radiation intensity. Depending on bias current, the detector can count single-photon (red squares) or two-photon (blue squares) events (Ref. 29).

gives, with no free parameters, the charge that has been transferred through the tunnel junction. This latter value is proportional to the number of quasiparticles created in the cascade and, consequently, to the x-ray quantum energy.

The theoretical energy resolution of the tunnel junction detector is given by $2.4\left[h\nu\,\Delta\left(F+1+1/n\right)\right]^{1/2}$, where $h\nu$ is the quantum energy, n is the number of tunneling events per one quasiparticle, and F is the Fano factor that describes the statistical fluctuations of the charge-generation process. The Twerenbold model incorporates the two-dimensional diffusion equation for Δn_q and the general nonlinear form of the RT equations.

A more general approach, including time evolution of nonequilibrium distribution functions of quasiparticles and phonons, was developed by Nussbaumer *et al.*[31] The authors solved the Chang-Scalapino equations numerically for the quasiparticle and phonon distribution functions in a spatially homogeneous situation and supplemented the solution by one-dimensional diffusion. The full theory includes the parameters that are important for the real detector, such as back tunneling and losses of quasiparticles at the edges of the device, resulting in good agreement between the calculated transient response signals and the experimentally measured pulse shapes.

Hot-Electron Detectors

A minor, but physically very important, difference exists between a superconducting HEP and a conventional superconducting bolometer when they are operated in the transition-edge regime. In the bolometer, thermal equilibrium between electrons and phonons is established instantly, whereas in the hot-electron detector these two systems are not in equilibrium. In this review, we restrict ourselves to publications where the nonequilibrium state between the electron and phonon subsystems was clearly observed. Basically, there are two ways to decouple electrons from phonons: nonequilibrium phonons should leave the detector at a time scale that is short compared to τ_{pe}, or the intensity of external radiation should vary faster than $1/\tau_{pe}$. Depending on the superconductor and experimental arrangement, a real hot-electron detector falls somewhere between these two extremes.

1. Transition-Edge Superconducting Detectors

Historically, the first HEP's were developed and studied in the early 1980s by Gershenzon *et al.*,[32] using ultrathin Nb films as the detector body. Niobium is characterized by relatively long τ_{pe}, typically a few hundred nanoseconds at liquid helium temperature, so that $\tau_{es} < \tau_{ep}$ for films thinner than 10 nm.[11] Therefore, detectors based on thin Nb films belong to the first limiting case in that their response time is approximately equal to τ_{ep}. The best performance that the Nb HEP's can achieve[33] is NEP = 3×10^{-13} W/Hz$^{1/2}$, detectivity $D^* = 4 \times 10^{11}$ cm s$^{1/2}$ J^{-1}, and a response time of 4.5 ns. Thus, these devices are less sensitive, although much faster, than semiconductor bolometers. When the detector area was adjusted properly, Nb HEP's demonstrated a constant value of sensitivity in the range from microwaves (150 GHz) to ultraviolet (10^{15} Hz). This is actually their greatest advantage when compared to semiconductor counterparts. A Nb-based HEP was implemented to study the emission of a cyclotron *p*-germanium laser.[34] The combination of large sensitivity and short response time made it possible to identify the Landau levels responsible for lasing.

In the late 1990s, the Gershenzon group developed a HEP based on NbN superconducting films.[35] NbN has much shorter τ_{ep} and τ_{pe} than Nb; thus, even for 3-nm-thick films, NbN HEP's operate in the mixed regime (i.e., τ_{ep} and τ_{es} jointly determine the response time of the detector). Detectors made from ultrathin NbN films are much faster than Nb-based devices. The intrinsic $\tau_{ep} \approx 10$ ps, while the overall response time is about 30 ps near T_c.[8] The best-demonstrated NEP $\approx 10^{-12}$ W/Hz$^{1/2}$ (Ref. 36). In spite of a rather-complicated electronic band structure,[37] the quantum yield in NbN reaches above 300 for near-infrared photons,[38] which corresponds to one-third of the upper theoretical limit. Detectors fabricated from NbN were used to study the emission of optically pumped

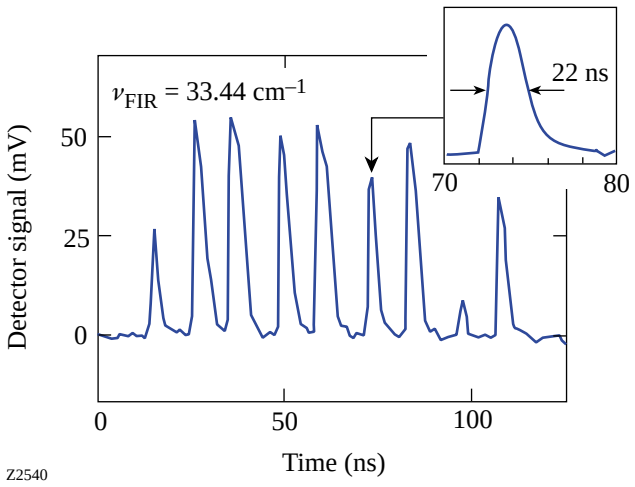

Figure 87.37
Pulses from a single-shot, optically pumped, far-infrared gas laser recorded with a NbN HEP (Ref. 39). The inset shows one of the pulses on an expended time scale.

infrared gas lasers, in particular, pulsed lasers.[39] Figure 87.37 shows far-infrared laser pulses recorded with a NbN hot-electron detector. The unique combination of response time and sensitivity made it possible to detect and identify very weak emission lines.

Miller *et al.*[40] have demonstrated a photon counter based on the transition-edge, hot-electron, direct detector. The device was a 20 × 20-μm^2 square of 40-nm-thick tungsten film (Fig. 87.38) having $T_c = 80$ mK with a transition width of 1 mK. The device was operated at a bath temperature of 40 mK in a voltage-bias regime that maintained the sensor within the transition region via negative electrothermal feedback.[41] This mode of operation was shown to increase the transition-edge sensor sensitivity and to decrease its time constant to $\tau_0/(1+\alpha/n)$. Here τ_0 is the intrinsic time constant of the sensor, n is the power of the temperature dependence of the thermal conductance between the film and substrate, and α is the dimensionless sharpness parameter of the superconducting transition.[41] A photon absorbed in the sensor heats the electron system above its equilibrium temperature, leading to an increase of the sensor's resistance and, consequently, to the decrease of the bias current and dissipated joule power. The integral of the drop in current (read out by an array of dc SQUID's) gives the energy absorbed by the sensor with no free parameters. The detector described in Ref. 40 exhibited a time constant of about 60 μs and was able to register 0.3-eV (4-μm-wavelength) single photons with an energy resolution of 0.15 eV. To test the detector, the authors observed the planetary nebula NGC 6572, using the 8-in. telescope. The energy resolution was somewhat lower than in the laboratory, although it was high enough to detect the strong emission lines.

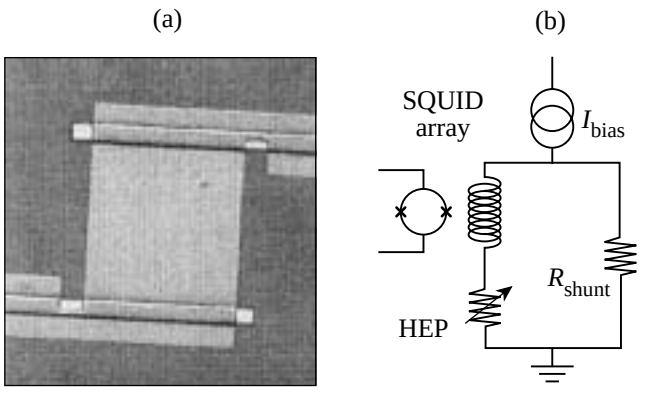

Figure 87.38
(a) Microphotograph of a transition-edge, hot-electron quantum detector and (b) the corresponding equivalent circuit (Ref. 40).

A hot-electron microcalorimeter was developed by Nahum and Martinis.[42] In this type of device, photon absorption gives rise to T_e in a metal absorber and is measured using the *I–V* characteristics of a normal-insulator-superconductor tunnel junction, in which part of the absorber forms the normal electrode. Figure 87.39 shows a schematic of the tested device. The current through the junction was measured with a low-noise dc SQUID. The absorber had an area of 100 × 100 μm^2 and was deposited on a silicon nitride membrane. In this configuration, the phonons that escaped from the absorber were reflected back from the membrane and were further available for the energy exchange. Thus, the Si_3N_4 membrane prevented energy loss from the electron subsystem in the absorber. The microcalorimeter operated at 80 mK with a time constant of 15 μs and demonstrated an energy resolution of 22 eV for 6-keV photons.

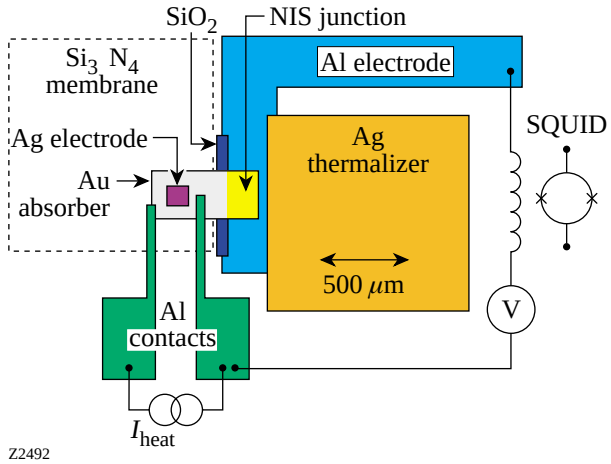

Figure 87.39
Detailed schematic of the hot-electron microcalorimeter developed by Nahum and Martinis (Ref. 42) (see text for explanation).

In another version, Nahum and Martinis[43] proposed a microbolometer that consisted of a normal metal stripe connected to superconducting electrodes (Fig. 87.40). The device relied on Andreev reflections of low-energy, thermal quasiparticles at the edges of the stripe and on weak electron–phonon coupling at low temperatures. Both effects confined the energy delivered by the photons, providing a large rise of T_e. This was subsequently read out by the superconductor-insulator-normal metal junction, for which the metal strip formed the normal electrode. Projected responsivity and NEP of the device with the Cu absorber operated at 100 mK were about 10^9 V/W and 3×10^{-18} W/Hz$^{1/2}$, respectively, which is at least a factor of 10 better than the performance of any currently available detectors. The time constant of the microcalorimeter is determined

by the rate of energy transfer from electrons to phonons that corresponds to τ_{ep} at the Fermi level. For the device under consideration in Ref. 43, the computed response time τ = 20 μs. Since, for a bolometer, NEP scales as $\tau^{-1/2}$, the device performance can be further improved by increasing the response time up to a value only slightly less than that required by a specific application.

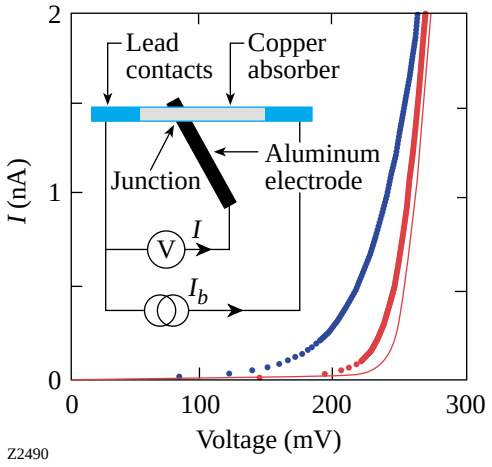

Figure 87.40
A hot-electron microbolometer using Andreev reflections of quasiparticles from superconducting contacts and the corresponding *I–V* characteristics (Ref. 43).

Finally, Karasik *et al.*[44] proposed the use of the dependence of the electron–phonon scattering time on the electron mean free path to control the intrinsic response time of a transition-edge detector. Increase of the intrinsic response time results in the decrease of the minimum detectable power, while at the same time, the device response time can be decreased to a reasonable value by exploiting the negative electrothermal feedback. According to estimates in Ref. 44, using this approach, a detector could be fabricated with NEP = 10^{-20} W/Hz$^{1/2}$ and the millisecond τ at 100-mK bath temperature.

2. Superconducting Kinetic-Inductive Detectors

The detectors described in the preceding section produce a response when the device, or at least part of it, is in the resistive state. Kinetic-inductive integrating detectors represent their superconducting counterpart. The L_{kin} [see Eq. (7)] of a superconducting film makes it possible to monitor the concentration of Cooper pairs. In a constant current-biased superconducting film, after the destruction of a certain number of Cooper pairs, the remaining pairs accelerate to carry the same bias current. Because of non-zero inertia of pairs, or L_{kin}, acceleration requires an electric field. This intrinsically generated electric field is seen from the exterior as a voltage pulse developing across the film. Mathematically, this voltage transient is given by

$$V_{kin} = I\frac{dL_{kin}}{dt}. \tag{13}$$

Figure 87.34(b) presented earlier the V_{kin} transient, recorded for a YBCO microbridge excited by 100-fs optical pulses. The numerical fit was based on Eq. (13) and either Eqs. (1) or (6).

The main advantage of superconducting kinetic-inductive detectors is their low noise power. To realize this advantage, a SQUID readout should be used. Grossman *et al.*[45] described the design of a kinetic-inductive detector/mixer with an estimated NEP = 2.5 × 10^{-17} W/Hz$^{1/2}$ and a bandwidth of 5.5 MHz at 100 mK. Unfortunately, a laboratory prototype showed only NEP = 4.4 × 10^{-11} W/Hz$^{1/2}$ (Ref. 46). Sergeev and Reizer[47] performed thorough calculations for both *s*-wave and *d*-wave superconductors, including the appropriate quasiparticle distribution function and scattering times. They found NEP and D^* close to those reported in Ref. 45. Bluzer[23] proposed a balanced-bias scheme for a kinetic-inductive photodetector with directly coupled SQUID readout, intended to eliminate the losses inherent in inductively coupled readouts and increase the responsivity of the detector. Performance of the detector was simulated for a 0.1-μm-thick YBCO film at 9 K, resulting in NEP = 2.5 × 10^{-15} W/Hz$^{1/2}$ and 10-μs response time. It is believed that the use of a LTS material should result in a two- to three-orders-of-magnitude decrease in NEP.

3. Superconducting Quantum Detectors

A number of novel approaches proposed during the last decade have been aimed at the realization of detectors with ultimate quantum sensitivity. Kadin and Johnson[26] introduced the quantum detection regime in ultrathin resistive films. In the proposed mechanism, an absorbed photon induces a resistive hot spot, centered at the point where the photon hits the film. If the photon flux is sufficiently low, hot spots do not overlap until they disappear. Using material parameters of NbN, the authors estimated that a 0.1-μm^2 size sample should respond to 1-eV photons with 1-mV-amplitude pulses and 10-GHz bandwidth. For technological reasons, practical detectors would require significantly larger areas and, consequently, much smaller responsivities, forcing the implementation of a sophisticated readout scheme such as an array of SQUID's.[27]

A photon counter using the quantum detection regime in a current-carrying superconducting film[28] was recently demonstrated by Gol'tsman *et al.*[29] The counting element consisted of a 1.3-μm-long, 0.2-μm-wide microbridge, formed from a 6-nm-thick NbN film deposited on a sapphire substrate. The detector was operated at 4.2 K, with a bias current of approximately 90% of I_c. Voltage pulses generated by the bridge in response to absorbed photons were further amplified by a cooled, low-noise amplifier (see Fig. 87.35). The output pulses were time limited by electronics and had a duration of approximately 100 ps. The intrinsic dark count rate for the detector was measured to be below 0.001 s^{-1} (probability 10^{-6}), which corresponds to zero detected responses over 1000 s when the input was completely blocked. Table 87.I presents the basic parameters of the device operated at the 790-nm wavelength. Single-photon counting was observed in the photon-wavelength range from 0.4 μm to 2.4 μm.[48] We note that the device represents a unique combination of the picosecond response time and very high responsivity. These characteristics of NbN HEP's should lead to their practical implementation in areas ranging from free-space satellite communication,[49] through quantum communication and quantum cryptography,[50] to ultraweak luminescence observations and semiconductor integrated circuit testing.[51] Another exciting application for this type of detector can be background-limited direct detector arrays[52] for submillimeter astronomy.

Table 87.I: Experimental performance of a NbN photodetector at 790 nm.

Response time–intrinsic/measured	10 ps/100 ps
Quantum gain factor	340
A/W responsivity	220 A/W
V/W responsivity	4×10^4 V/W
Device quantum efficiency	~20%
Operating temperature	~4 K
Dark counts per second	<0.0001
Device noise temperature	~15 K

The most-advanced superconducting quantum detectors are tunnel-junction detectors, which are being developed for a wide range of applications from materials science and microanalysis to particle physics and astrophysics. Only a few recent publications are mentioned here because a full review of the activities in this field is beyond the scope of this article. Nb-based tunnel-junction detectors with Al trapping layers have reached, for photon energies of about 70 eV, an energy resolution of 1.9 eV. This performance is limited by the statistics of quasiparticle multiplication.[53] A typical device had an area of 50×50 μm^2. The smallest-detectable, 0.3-eV (4.1-μm-wavelength) photon energy was achieved with Ta-based devices[54] since this material has an energy gap smaller than that of Nb. An energy resolution of 0.19 eV was demonstrated for 2.5-eV (0.5-μm-wavelength) photons, using Ta-based devices with an area of 20×20 μm^2 and 12-μs response time.

Hot-Electron Mixers

Historically, HEM's have been divided into two large categories: lattice- or phonon-cooled[13] and diffusion-cooled[14] devices. As presented earlier, the physics for these two types of HEM's is essentially the same. Both types can be described by Eqs. (2) using temperature-dependent parameters and proper boundary conditions. The analysis becomes easier, however, when the device is designed to be close to one of two extremes, namely, the lattice- or the diffusion-cooling regime. Typically, lattice-cooled mixers are made from thin films of NbN, whereas diffusion-cooled devices use Nb or Al.

1. Lattice-Cooled Mixers

Current state-of-the-art NbN technology is capable of routinely delivering 3.0-nm-thick devices that are 500×500 nm^2 in size with T_c above 9 K. Near T_c, τ_{pe} is close to τ_{es}, which is about 40 ps for 3-nm-thick film [see Fig. 87.30(b)]. The τ_{ep} at 8 K is below 20 ps, which results, with the diffusivity of 0.5 $cm^2 s^{-1}$, in a thermal healing length of about 30 nm. Since the device length is typically much larger, the mixer operates in the phonon-cooled regime. The mixer's intrinsic IF bandwidth is determined by the combination of τ_{ep} and τ_{es} time constants. In real devices, however, the measured bandwidth depends strongly on the bias regime. This makes it difficult to compare published data and reach meaningful conclusions. For HEM's on Si substrates, the best reported gain and noise bandwidths are 3.5 GHz[55] and 8 GHz,[56] respectively. Further increases in the bandwidth for lattice-cooled HEM's can be achieved by using a substrate material that is better thermally coupled to the superconducting film. One promising candidate is MgO. Recent measurements have shown[57] that MgO provides, for a 3.5-nm-thick bolometer, a 4.8-GHz gain bandwidth and 5.6-GHz noise bandwidth, respectively. Further progress in increasing the bandwidth may be achieved by decreasing the bolometer thickness. Recently a 9-GHz gain bandwidth was reported[58] for a 2.5-nm-thick device on MgO. Unfortunately, this direction is limited because NbN films thinner than 2.5 nm become inhomogeneous and lose their superconductivity.[59]

A waveguide version of the receiver with the lattice-cooled NbN HEM has been installed and operated successfully in the frequency range of 0.6 to 0.8 THz[60] and 1.04 THz[61] at the 10-m Sub-mm Telescope Facility on Mount Graham in Arizona. At this telescope, the measured noise temperature of the receiver was 560 K at 0.84 THz and 1600 K at 1.035 THz over a 1-GHz IF bandwidth centered at 1.8 GHz. The receiver was used to detect the CO molecular line emission in the Orion nebula (Fig. 87.41). It is worth noting that this was the first ground-based observation at a frequency above 1 THz. A quasi-optical version of the HEM receiver for the THz range is currently under preparation for test flights on a stratospheric airplane observatory.[62] The mixer will be incorporated into a planar logarithmic spiral antenna (Fig. 87.42), which is integrated with an extended hyperhemispherical silicon lens.

Practical advantages of the lattice-cooled devices are their stability and the weak sensitivity of their noise temperature to operation parameters. Figure 87.43 shows that, indeed, the noise temperature of a NbN hot-electron mixer does not vary noticeably over a broad range of LO power and bias voltage.[63]

2. Diffusion-Cooled Mixers

The bulk of diffusion-cooled mixers has been realized based on Nb films. At a 4.2-K bath temperature, the 10-nm-thick Nb film typically has τ_{ep} of about 1 ns and a diffusivity of 2 $cm^2 s^{-1}$,[11] which results in $L_{th} \approx 0.15\ \mu m$. Therefore, Nb devices having a length of 0.1 μm or less operate in the diffusion-cooled regime. It has been shown experimentally[64] that the transition to diffusion cooling of electrons occurs at a

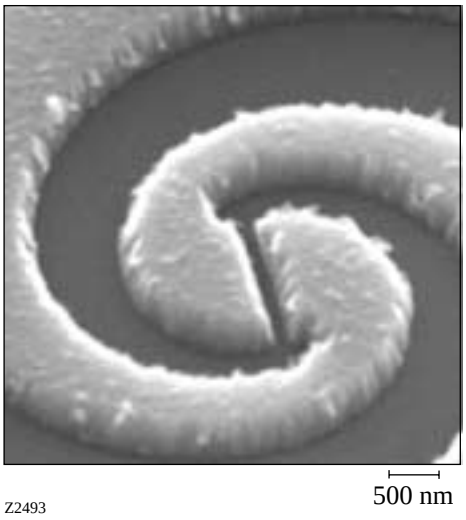

Figure 87.42
Central part of a planar logarithmic spiral antenna with the NbN hot-electron microbridge.

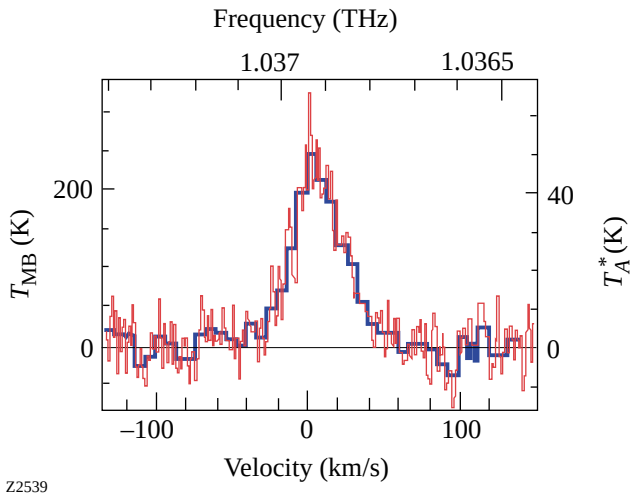

Figure 87.41
Terahertz CO line in the Orion IRc2 nebula recorded with a NbN hot-electron mixer at a ground-based telescope in Arizona (Ref. 61). The thick solid line shows a smoothed spectrum at a resolution of 25 MHz. The temperature scale of the spectrum is calibrated by taking into account the receiver noise temperature, the estimated atmospheric opacity, and the estimated efficiency of the telescope.

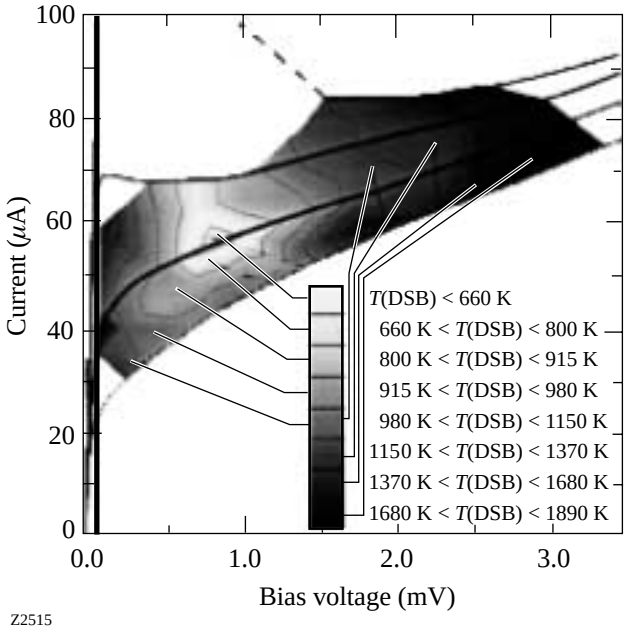

Figure 87.43
Double-sideband (DSB) noise temperature of a laboratory heterodyne receiver with NbN HEM at various bias regimes (Ref. 63).

device length ≈ 0.2 μm. Expected gain bandwidth for a 0.1-μm-long device is about 7 GHz, if one assumes uniform electron heating through the length of the device. Laboratory tests at sub-THz frequencies confirmed theoretical expectations, and a 9-GHz gain bandwidth was measured for a 0.1-μm-long HEM.[65] No noise bandwidth data have been reported so far for diffusion-cooled mixers. Traditionally, quasi-optical, diffusion-cooled HEM's use a twin-slot or double-dipole planar antenna and a hemispherical lens to couple the LO and signal radiations to the mixer. The best reported noise temperatures for Nb diffusion-cooled mixers are presently almost twice as large as those of lattice-cooled devices.

Another apparent difference between the two types of HEM's is the optimal bias regime, i.e., the regime resulting in the lowest noise temperature. For a lattice-cooled HEM, the optimal bias point is within the linear portion of the nonhysteretic *I–V* characteristics,[63] whereas optimal operation of diffusion-cooled devices corresponds to the nonlinear portion of a hysteretic *I–V* curve.[65] The difference stems from boundary conditions imposed on the normal domain. Movement of the domain walls caused by signal radiation is not influenced by the contacts[66] if they are located far enough from the domain borders. One can envision such a domain as a freestanding domain in a stable equilibrium state. In the opposite case, when domain walls are confined near the contacts, the temperature profile at the walls slopes more steeply and the wall movement is restricted by the contacts. This hampers the responsivity of the HEM. As a result, the length of a diffusion-cooled mixer is smaller than the thermal diffusion length L_{th} and corresponds to the length of the smallest freestanding domain. Therefore, in a diffusion-cooled HEM, the conversion loss and, consequently, the noise temperature are smaller when the domain is "overcooled" and is slightly shorter than the smallest freestanding domain. The actual domain length, as seen from the resistance in the normal state at the optimal operation point,[65] is about 0.6 of the mixer physical length, whereas for phonon-cooled HEM's,[63] the domain length is 0.2 of the device length. Since the total noise power at the HEM output is partly due to Nyquist noise, smaller responsivity should result in a somewhat larger noise temperature. Another disadvantage of the diffusion-cooled HEM is that its hysteretic regime may cause additional instability[67] when accessed by a practical receiver.

For both mixer types, it is common that optimal operation, aimed at the minimal noise temperature, does not provide the largest-possible IF bandwidth. Both the bandwidth and the noise temperature increase with the bias current. Thus, varying the bias regime allows a compromise between the desired bandwidth and the noise temperature acceptable for a particular application.

A diffusion-cooled Al mixer has been recently proposed[68] as an alternative to Nb devices. Measurements at 30 GHz[69] showed that a diffusion-cooled Al mixer exhibits reasonably good performance, but these data are not conclusive for the desired THz operation since the quantum energy of 30-GHz photons remains smaller than the Al energy gap. Moreover, there are concerns[19] that Al HEM's at THz frequencies would require a large LO power.

Table 87.II and Fig. 87.44 summarize the current state-of-the-art noise temperatures for both the lattice-cooled and diffusion-cooled HEM's. The rapid increase in noise temperature with frequency is inconsistent with the hot-electron model. The model suggests that the noise temperature, when corrected for optical losses, should not depend on frequency unless it approaches the quantum-limited value $h\nu/k_B$. A proper account of losses in coupling optics does not eliminate the above discrepancy; the noise temperature of the mixer alone increases with frequency, following closely the 10 $h\nu/k_B$ law in the frequency range from 0.6 THz to 5.2 THz. It has been shown recently[64] that the nonuniform distribution of the high-frequency current across the device may account for this effect.

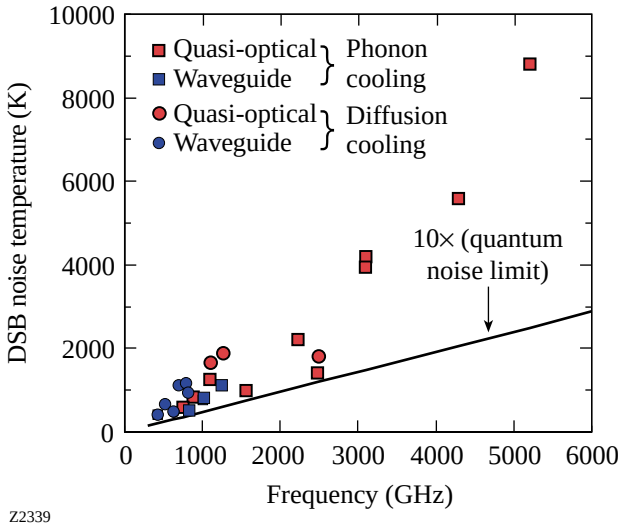

Figure 87.44
Best double-sideband (DSB) noise temperatures for various types of superconducting hot-electron mixers as a function of signal frequency. The solid line is the hot-electron model prediction.

In Fig. 87.45, simulated frequency dependence of the conversion efficiency is compared with the noise temperature corrected for optical losses. Good agreement between the experimental and theoretical results up to 4 THz suggests that the increase in the noise temperature should be less pronounced for narrower HEM's.

Conclusions

Superconductor hot-electron radiation sensors, operated as either THz-frequency mixers or optical single-photon detectors, promise a revolutionary approach for diagnostics, radio astronomy, and quantum cryptography and communications. The unique performance of these devices in heterodyne as well as in the direct-detection regime results from a combination of the hot-electron phenomenon with the high sensitivity of a superconductor to nonequilibrium electronic states. To take full advantage of this combination, devices are routinely fabricated from ultrathin superconducting films and feature submicron lateral dimensions. They are also operated in the very-low-noise cryogenic environment.

HEM's proved their reliability and advantageous features during a two-year test on a ground-based telescope. In the frequency range from 1 THz to 5.2 THz, HEM's outperformed Schottky diodes, making them the device-of-choice for THz astronomy and communications.

HEP's demonstrated excellent performance in the spectral range from far-infrared wavelengths to x rays when operated in either integrating or quantum regimes. Their future applications are expected in areas ranging from background-limited detector arrays for submillimeter astronomy and x-ray spectroscopy, through practical, high-speed quantum cryptography, to digital integrated-circuit diagnostics.

ACKNOWLEDGMENT

The authors thank S. Cherednichenko for helpful discussions and S. Svechnikov for systemizing published data. This research was made possible in part by Award No. RE-2227 of the U.S. Civilian Research and Development Foundation for the Independent States of the Former Soviet Union. G. N. Gol'tsman acknowledges support of the INTAS (#97-1453, #99-569). R. Sobolewski acknowledges support of the U.S. Office of Naval Research grant N00014-00-1-0237.

Table 87.II: Best double-sideband (DSB) noise temperatures reported in the literature for lattice-cooled and diffusion-cooled mixers.

Lattice-cooled mixers					
Quasi-optical layout			Waveguide layout		
Frequency (GHz)	DSB noise temperature (K)	Reference	Frequency (GHz)	DSB noise temperature (K)	Reference
620	500	70	430	410	73
750	600	65	636	483	73
910	850	65	840	490	61
1100	1250	65	1017	750	61
1560	1000	71	1030	800	61
1620	700	58	1260	1100	61
2240	2200	71			
2500	1100	58			
3100	4000	72			
4300	5600	72			
5200	8800	72			
Diffusion-cooled mixers					
Quasi-optical layout			Waveguide layout		
Frequency (GHz)	DSB noise temperature (K)	Reference	Frequency (GHz)	DSB noise temperature (K)	Reference
630	470	64	530	650	76
1100	1670	74	700	1100	17
1267	1880	75			
2500	1800	64			

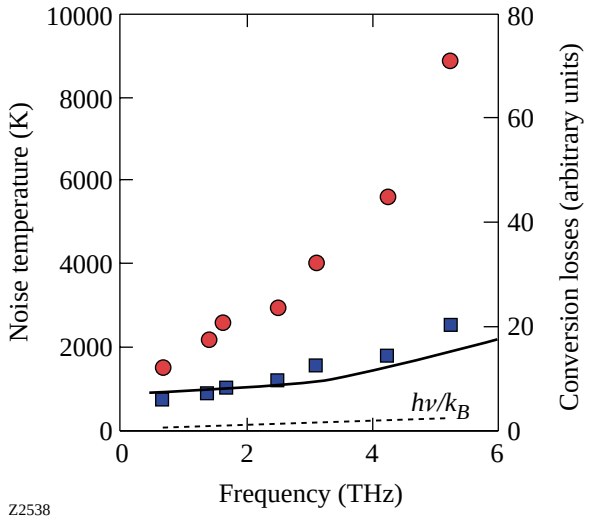

Figure 87.45
Frequency dependence of the noise temperature (circles) and conversion losses (squares) of a NbN HEM (Ref. 68). The solid line shows the calculated conversion losses that account for the skin effect in the device. The dashed line represents quantum-limited noise temperature $h\nu/k_B$. The scale of the right axis was adjusted to match calculated conversion losses and corrected noise temperature.

REFERENCES

1. E. M. Conwell, *High Field Transport in Semiconductors*, Solid State Physics Supplement 9 (Academic Press, New York, 1967).
2. M. L. Kaganov, I. M. Lifshitz, and L. V. Tanatarov, Sov. Phys.-JETP **4**, 173 (1957).
3. V. A. Schklovski, [Sov. Phys.] FTT **17**, 3076 (1975).
4. N. Perrin and C. Vanneste, Phys. Rev B **28**, 5150 (1983).
5. A. D. Semenov *et al.*, Phys. Rev. B **52**, 581 (1995).
6. E. M. Gershenzon *et al.*, Sov. Phys.-JETP **59**, 442 (1984).
7. M. Lindgren *et al.*, Appl. Phys. Lett. **65**, 3398 (1994).
8. K. S. Il'in, M. Lindgren, M. Currie, A. D. Semenov, G. N. Gol'tsman, R. Sobolewski, S. I. Cherednichenko, and E. M. Gershenzon, Appl. Phys. Lett. **76**, 2752 (2000).
9. M. Lindgren, M. Currie, C. Williams, T. Y. Hsiang, P. M. Fauchet, R. Sobolewski, S. H. Moffat, R. A. Hughes, J. S. Preston, and F. A. Hegmann, Appl. Phys. Lett. **74**, 853 (1999).
10. R. Sobolewski, in *Superconducting and Related Oxides: Physics and Nanoengineering III*, edited by D. Pavuna and I. Bozovic (SPIE, Bellingham, WA, 1998), Vol. 3481, pp. 480–491.
11. E. M. Gershenzon *et al.*, Sov. Phys.-JETP **70**, 505 (1990).
12. E. M. Gershenzon *et al.*, Sov. Phys.-Tech. Phys. **34**, 195 (1989).
13. E. M. Gershenzon *et al.*, Supercond., Phys. Chem. Technol. **3**, 1582 (1990); B. S. Karasik and A. I. Elantiev, Appl. Phys. Lett. **68**, 853 (1996).
14. D. E. Prober, Appl. Phys. Lett. **62**, 2119 (1993).
15. R. S. Nebosis *et al.*, in *Proceedings of the Seventh International Symposium on Space Terahertz Technology* (University of Virginia, Charlottesville, VA, 1996), pp. 601–613.
16. B. S. Karasik, W. R. McGrath, and M. C. Gaidis, J. Appl. Phys. **81**, 1581 (1997).
17. W. D. Floet *et al.*, in *Proceedings of the Tenth International Symposium on Space Terahertz Technology* (University of Virginia, Charlottesville, VA, 1999), pp. 228–236.
18. H. Merkel *et al.*, IEEE Trans. Appl. Supercond. **9**, 4201 (1999).
19. A. D. Semenov and G. N. Gol'tsman, J. Appl. Phys. **87**, 502 (2000).
20. H. Ekström *et al.*, IEEE Trans. Microw. Theory Tech. **43**, 938 (1995).
21. A. Rothwarf and B. N. Taylor, Phys. Rev. Lett. **19**, 27 (1967).
22. E. N. Grossman, D. G. McDonald, and J. E. Sauvageau, IEEE Trans. Magn. **27**, 2677 (1991).
23. N. Bluzer, J. Appl. Phys. **78**, 7340 (1995).
24. A. D. Semenov *et al.*, IEEE Trans. Appl. Supercond. **7**, 3083 (1997).
25. M. Lindgren, M. Currie, C. A. Williams, T. Y. Hsiang, P. M. Fauchet, R. Sobolewski, S. H. Moffat, R. A. Hughes, J. S. Preston, and F. A. Hegmann, IEEE J. Sel. Top. Quantum Electron. **2**, 668 (1996).
26. A. M. Kadin and M. W. Johnson, Appl. Phys. Lett. **69**, 3938 (1996).
27. D. Gupta and A. M. Kadin, IEEE Trans. Appl. Supercond. **9**, 4487 (1999).
28. A. D. Semenov, G. N. Gol'tsman, and A. A. Korneev, Physica C **351**, 349 (2001).
29. G. N. Gol'tsman, O. Okunev, G. Chulkova, A. Lipatov, A. Semenov, K. Smirnov, B. Voronov, A. Dzardanov, C. Williams, and R. Sobolewski, Appl. Phys. Lett. **79**, 705 (2001).
30. D. Twerenbold, Phys. Rev. B **34**, 7748 (1986).
31. Th. Nussbaumer *et al.*, Phys. Rev. B **61**, 9719 (2000).
32. E. M. Gershenzon *et al.*, JETP Lett. **34**, 268 (1981).
33. E. M. Gershenzon *et al.*, Sov. Tech. Phys. Lett. **15**, 118 (1989).
34. A. V. Bespalov *et al.*, Solid State Commun. **80**, 503 (1991).
35. B. M. Voronov *et al.*, Supercond., Phys. Chem. Technol. **5**, 960 (1992).
36. Yu. P. Gousev *et al.*, J. Appl. Phys. **75**, 3695 (1994).

37. S. V. Vonsovskii, IU. A. Iziumov, and E. Z. Kurmaev, *Superconductivity of Transition Metals: Their Alloys and Compounds*, Springer Series in Solid-State Sciences (Springer-Verlag, Berlin, 1982).

38. K. S. Il'in, I. I. Milostnaya, A. A. Verevkin, G. N. Gol'tsman, E. M. Gershenzon, and R. Sobolewski, Appl. Phys. Lett. **73**, 3938 (1998).

39. P. T. Lang *et al.*, Appl. Phys. B **B53**, 207 (1991); P. T. Lang *et al.*, Opt. Lett. **17**, 502 (1992).

40. A. J. Miller *et al.*, IEEE Trans. Appl. Supercond. **9**, 4205 (1999).

41. K. D. Irwin, Appl. Phys. Lett. **66**, 1998 (1995).

42. M. Nahum and J. M. Martinis, Appl. Phys. Lett. **66**, 3203 (1995); K. D. Irwin *et al.*, Nucl. Instrum. Methods Phys. Res. A **444**, 184 (2000).

43. M. Nahum and J. M. Martinis, Appl. Phys. Lett. **63**, 3075 (1993).

44. B. S. Karasik *et al.*, Supercond. Sci. Technol. **12**, 745 (1999).

45. E. N. Grossman, D. G. McDonald, and J. E. Sauvageau, IEEE Trans. Magn. **27**, 2677 (1991).

46. J. E. Sauvageau, D. G. McDonald, and E. N. Grossman, IEEE Trans. Magn. **27**, 2757 (1991).

47. A. V. Sergeev and M. Yu. Reizer, Int. J. Mod. Phys. B **10**, 635 (1996).

48. G. N. Gol'tsman, O. Okunev, G. Chulkova, A. Lipatov, A. Dzardanov, K. Smirnov, A. Semenov, B. Voronov, C. Williams, and R. Sobolewski, IEEE Trans. Appl. Supercond. **11**, 574 (2001).

49. G. G. Ortiz, J. V. Sandusky, and A. Biswas, in *Free-Space Laser Communication Technologies XII*, edited by G. S. Mercherle (SPIE, Bellingham, WA, 2000), Vol. 3932, pp. 127–138.

50. G. Gilbert and M. Hamrick, "Practical Quantum Cryptography: A Comprehensive Analysis (Part One)," to appear in Phys. Rep.; see also http://xxx.lanl.gov/abs/quant-ph/0009027.

51. J. C. Tsang and J. A. Kash, Appl. Phys. Lett. **70**, 889 (1997).

52. G. Gol'tsman *et al.*, "Background Limited Quantum Superconducting Detector for Submillimeter Wavelengths," to be published in the *Proceedings of the Twelfth International Symposium on Space Terahertz Technology*.

53. S. Friedrich *et al.*, IEEE Trans. Appl. Supercond. **9**, 3330 (1999).

54. P. Verhoeve *et al.*, IEEE Trans. Appl. Supercond. **7**, 3359 (1997).

55. S. Cherednichenko *et al.*, in *Proceedings of the Eighth International Symposium on Space Terahertz Technology* (Harvard University, Cambridge, MA, 1997), pp. 245–252.

56. H. Ekström *et al.*, Appl. Phys. Lett. **70**, 3296 (1997).

57. S. Cherednichenko *et al.*, in *Proceedings of the Eleventh International Symposium on Space Terahertz Technology* (Ann Arbor, MI, 2000), pp. 219–227.

58. M. Kroug *et al.*, IEEE Trans. Appl. Supercond. **11**, 962–965 (2001).

59. B. M. Voronov, Moscow State Pedagogical University, private communication (2000).

60. J. H. Kawamura *et al.*, IEEE Trans. Appl. Supercond. **9**, 3753 (1999); J. H. Kawamura *et al.*, Publ. Astron. Soc. Pac. **111**, 1088 (1999).

61. C.-Y. E. Tong *et al.*, in *Proceedings of the Eleventh International Symposium on Space Terahertz Technology*, (Ann Arbor, MI, 2000), pp. 49–59.

62. H.-W. Huebers *et al.*, in *Airborne Telescope Systems*, edited by R. K. Melugin and H.-P. Roeser (SPIE, Bellingham, WA, 2000), Vol. 4014, pp. 195–202.

63. P. Yagoubov *et al.*, Appl. Phys. Lett. **73**, 2814 (1998).

64. P. J. Burke *et al.*, Appl. Phys. Lett. **68**, 3344 (1996).

65. R. A. Wyss *et al.*, in *Proceedings of the Tenth International Symposium on Space Terahertz Technology* (University of Virginia, Charlottesville, VA, 1999), pp. 215–228.

66. A. D. Semenov and H.-W. Hubers, IEEE Trans. Appl. Supercond. **11**, 196 (2001).

67. A. D. Semenov *et al.*, J. Appl. Phys. **88**, 6758 (2000).

68. B. S. Karasik and W. R. McGrath, in *Proceedings of the Ninth International Symposium on Space Terahertz Technology* (1998), p. 73.

69. I. Siddiqi *et al.*, in *Proceedings of the Eleventh International Symposium on Space Terahertz Technology*, (Ann Arbor, MI, 2000), pp. 82–94.

70. P. Yagoubov *et al.*, Supercond. Sci. Technol. **12**, 989 (1999).

71. E. Gerecht *et al.*, in *Proceedings of the Tenth International Symposium on Space Terahertz Technology* (University of Virginia, Charlottesville, VA, 1999), pp. 200–207.

72. J. Schubert *et al.*, Supercond. Sci. Technol. **12**, 748 (1999).

73. J. Kawamura *et al.*, in *Proceedings of the Eighth International Symposium on Space Terahertz Technology* (Harvard University, Cambridge, MA, 1997), pp. 23–26.

74. A. Skalare *et al.*, in *Proceedings of the Ninth International Symposium on Space Terahertz Technology* (1998), pp. 115–120.

75. A. Skalare *et al.*, IEEE Trans. Appl. Supercond. **7**, 3568 (1997).

76. A. Skalare *et al.*, Appl. Phys. Lett. **68**, 1558 (1996).