

THE PHYSICS OF HETERODYNE DETECTION IN THE FAR-INFRARED: TRANSITION
FROM ELECTRIC-FIELD TO PHOTON-ABSORPTION DETECTION IN A SIMPLE SYSTEM*

Malvin Carl Teich
Columbia Radiation Laboratory
Department of Electrical Engineering
Columbia University
New York, New York 10027

SUMMARY

After briefly reviewing the history of heterodyne detection from the radiowave to the optical regions of the electromagnetic spectrum, we focus on the submillimeter/far-infrared by investigating the transition from electric-field to photon-absorption detection in a simple system. The response of an isolated two-level detector to a coherent source of incident radiation is calculated for both heterodyne and video detection. When the processes of photon absorption and photon emission cannot be distinguished, the relative detected power at double- and sum-frequencies is found to be multiplied by a coefficient, which is less than or equal to unity, and which depends on the incident photon energy and on the effective temperature of the system.

INTRODUCTION

Heterodyne detection has a long and august history in the annals of electrical engineering, reaching back to the earliest years of the century. The term has its roots in the Greek words "heteros" (other) and "dynamis" (force).

In 1902, Reginald Fessenden was awarded a patent (1) "relating to certain improvements ... in systems where the signal is transmitted by [radio]waves differing in period, and to the generation of beats by the waves and the employment of suitable receiving apparatus responsive only to the combined action of waves corresponding in period to those generated" The subsequent realization that one of these waves could be locally generated (the development of the local oscillator) provided a substantial improvement in system performance. Practical demonstrations of the usefulness of the technique were carried out between the "Fessenden stations" of the U.S. Navy at Arlington (Virginia) and the Scout Cruiser Salem, between the Salem and the Birmingham (1910), and at the National Electric Signaling Company. In

* Work supported by the Joint Services Electronics Program (U.S. Army, U.S. Navy, and U.S. Air Force) under Contract No. DAAG29-79-C-0079.

1913, John Hogan provided a thoroughly enjoyable account of the development, use, and performance of the Fessenden heterodyne signaling system in the first volume of the Proceedings of the IRE (2).

It was not long thereafter, in 1917, that Edwin H. Armstrong of the Department of Electrical Engineering at Columbia University carried out a thorough investigation of the heterodyne phenomenon occurring in the oscillating state of the "regenerative electron relay" (3). A major breakthrough in the field, the development of the superheterodyne receiver, was achieved by Armstrong in 1921 (4), and this famous invention is now used in systems as diverse as household AM radio receivers and microwave Doppler radars. The prefix "super" refers to "the super-audible frequency that could be readily amplified".

In the succeeding years, the application of heterodyne and superheterodyne principles followed the incessant march toward higher frequencies that culminated in the remarkable developments in microwave electronics about the time of World War II.

The marriage of heterodyning and the optical region took place in 1955. In a now classic experiment, Forrester, Gudmundsen, and Johnson (5) observed the mixing of two Zeeman components of a visible (incoherent) spectral line in a specially constructed photomultiplier tube. With the development of the laser, optical heterodyning became considerably easier to observe and was studied in 1962 by Javan, Ballik, and Bond (6) at $1.15 \mu\text{m}$ using a He-Ne laser, and by McMurtry and Siegman (7) at 6943 \AA using a ruby laser.

The development of new transmitting and receiving components in the middle infrared region of the electromagnetic spectrum led Teich, Keyes, and Kingston (8) in 1966 to perform a heterodyne experiment using a CO_2 laser at $10.6 \mu\text{m}$ in conjunction with a copper-doped germanium photoconductive detector operated at 4°K . Subsequent experiments with lead-tin selenide photovoltaic detectors confirmed the optimal nature of the detection process (9)-(11). The state of the art in applying these results to coherent infrared radar was reviewed by Kingston in 1977 (12). Elbaum and Teich (13) have recently shown the importance of properly choosing the performance measure for a heterodyne system. This can sometimes be critical to avoid faulty estimates of signal and noise levels leading to inconsistent or incorrect results, as pointed out by a number of authors (8), (14).

The foregoing discussion has dealt with more-or-less idealized systems; it must be kept in mind that there are a variety of effects that can alter heterodyne performance in important ways. Two examples were readily provided at this conference: Charles Townes discussed atmospheric perturbations of phase coherence in large diameter heterodyne receivers, and M. J. Mumma and T. Kostiuik discussed the effects of various sources of noise on system performance.

Finally, we point out that a number of system configurations employing different forms of nonlinear heterodyne detection have been proposed for various applications (15), (16). These make use of multiple frequencies,

nonlinear detectors, and/or correlation schemes. All are aimed at increasing the signal-to-noise ratio in situations for which usual operating conditions are relaxed in particular ways. One of these schemes (17), (15), using a two-frequency transmitter, is currently under consideration for possible application in over-water pilot rescue missions where the use of a light-chopping mechanism is interdicted because of its interruptive nature (private communication with W. Tanaka, Naval Weapons Center).

In the following section, we relax the assumption that heterodyne detection takes place by means of photon absorption, as it does in the middle infrared and optical regions (18), (19), or by means of electric-field detection, as it does in the radiowave and microwave regions. This will enable us to examine the heterodyne detection process in the submillimeter/far-infrared region of the spectrum in a simple way. The reader is cautioned that the treatment is heuristic in nature, and is only intended to provide a qualitative description of the underlying detection process.

TRANSITION FROM ELECTRIC-FIELD TO PHOTON-ABSORPTION DETECTION IN A SIMPLE SYSTEM

The increasing importance of heterodyning and electric-field detection in the submillimeter, infrared, and optical (20), (21) has encouraged us to examine the operation of such systems in relation to the more conventional photon-absorption detector (22)-(24).

For heterodyne mixing with coherent signals in a two-level detector, a simple argument indicates that double- and sum-frequency terms in the detected power are multiplied by the factor $\text{sech}(h\nu/2kT_e)$, which varies smoothly from unity in the electric-field detection regime to zero in the photon-absorption detection regime. This factor depends on both the incident photon energy $h\nu$ and on the effective excitation energy of the (two-level) detector, kT_e . It is observed that these terms only appear when it cannot be determined whether photon absorption or photon emission has taken place. Difference-frequency signals, on the other hand, arise from our inability to determine from which beam a photon is absorbed in a heterodyne experiment (18), (19).

We consider a simple hypothetical two-level system with an effective temperature T_e . It is assumed that the system responds linearly to the incident radiation intensity, and that its interaction with the field is sufficiently weak such that the state of the field is not perturbed by the presence of the detector.

We label the initial and final states of the system as $|\alpha\rangle$ and $|\Omega\rangle$ and of the radiation field as $|i\rangle$ and $|f\rangle$, respectively. For an electric-dipole transition, the transition probability W_{fi} (which is related to the detected power) is given approximately by

$$W_{fi} = |\langle f | \Omega | e q (E^+ + E^-) | \alpha i \rangle|^2, \quad (1)$$

where e is the electronic charge, q is the detector coordinate, and E^+ and E^- are the positive- and negative-portions of the electric-field operator, respectively. Since E^+ corresponds to photon absorption or annihilation, and E^- corresponds to photon emission or creation, the transition probability may be written as

$$W_{fi} = |\langle fu | eqa_{\ell} E^+ | \ell i \rangle + \langle fl | eqa_u E^- | u i \rangle|^2, \quad (2)$$

where $|\ell\rangle$ and $|u\rangle$ represent the lower and upper states of the two-level system, respectively. This equation assumes that the atomic system is, in general, in a superposition state. Using microscopic reversibility, the quantity $|\langle u | eq | \ell \rangle|^2$, which represents the quantum efficiency κ , may be factored out of Eq. (2). We then sum over the final states of the field (23), which are not observed, to obtain

$$\begin{aligned} W \propto & \langle i | a_{\ell}^* a_{\ell} E^- E^+ | i \rangle + \langle i | a_u^* a_u E^+ E^- | i \rangle \\ & + \langle i | a_{\ell}^* a_u E^- E^- + a_{\ell} a_u^* E^+ E^+ | i \rangle. \end{aligned} \quad (3)$$

The normally ordered first term corresponds to stimulated absorption, the antinormally ordered second term corresponds to photon emission (25), and the third is an interference term.

We now assume that before the interaction, the probability amplitudes of the two possible states, a_{ℓ} and a_u , were related by the Boltzmann factor, with lower and upper level energies represented by E_{ℓ} and E_u , respectively, and with excitation energy kT_e defining the effective temperature of the detector. Thus,

$$|a_u|^2 / |a_{\ell}|^2 = \exp[-(E_u - E_{\ell}) / kT_e], \quad (4)$$

yielding

$$a_{\ell} = (1 + e^{-x})^{-1/2} e^{i\phi} \quad (5)$$

and

$$a_u = e^{-x/2} (1 + e^{-x})^{-1/2} e^{i\theta}, \quad (6)$$

with $x \equiv h\nu/kT_e$ and $e^{i\theta}$, $e^{i\phi}$ representing phase factors. Then, generalizing to an arbitrary radiation field represented by the density operator ρ , we cavalierly obtain the expression

$$W \propto \left\{ [e^{x/2} \operatorname{sech}(x/2)] \operatorname{tr}\{\rho E^- E^+\} + [e^{-x/2} \operatorname{sech}(x/2)] \operatorname{tr}\{\rho E^+ E^-\} \right. \\ \left. + [\operatorname{sech}(x/2)] \operatorname{tr}\{\rho (e^{-i\gamma} E^- E^- + e^{i\gamma} E^+ E^+) \} \right\}, \quad (7)$$

with $\gamma \equiv \phi - \theta$.

Considering ideal heterodyne detection, i.e., two parallel, monochromatic, and coherent waves of frequencies ν_1 and ν_2 impinging normally on the detector, and neglecting spontaneous emission so that the antinormally ordered and the normally ordered terms are equal in magnitude (25), the first two terms above generate dc and difference-frequency signals, while the third term contributes double- and sum-frequency signals. This may be clearly seen by explicitly rewriting Eq. (7) as

$$W \propto |\epsilon_1^0|^2 + |\epsilon_2^0|^2 + 2|\epsilon_1^0||\epsilon_2^0| \cos[2\pi(\nu_1 - \nu_2)t + (\beta - \alpha)] \\ + [\operatorname{sech}(h\nu/2kT_e)] \left\{ |\epsilon_1^0|^2 \cos(4\pi\nu_1 t - 2\alpha - \gamma) + |\epsilon_2^0|^2 \cos(4\pi\nu_2 t - 2\beta - \gamma) \right. \\ \left. + 2|\epsilon_1^0||\epsilon_2^0| \cos[2\pi(\nu_1 + \nu_2)t - (\alpha + \beta) - \gamma] \right\} \quad (8)$$

where $\epsilon_1^0 = |\epsilon_1^0| e^{i\alpha}$ represents the complex electric-field amplitude of the constituent field with frequency ν_1 and phase α . Double- and sum-frequency terms in the heterodyne signal are therefore multiplied by the factor $\operatorname{sech}(h\nu/2kT_e)$.

For $h\nu/kT_e \rightarrow 0$, this factor approaches 1 and the classical electric-field heterodyne signal obtains

$$W_{\text{elec}} \propto \left[|\epsilon_1^0| \cos(2\pi\nu_1 t - \alpha - \frac{\gamma}{2}) + |\epsilon_2^0| \cos(2\pi\nu_2 t - \beta - \frac{\gamma}{2}) \right]^2. \quad (9)$$

For $h\nu/kT_e \rightarrow \infty$, $\operatorname{sech}(h\nu/2kT_e) \rightarrow 0$ and the photon-absorption (optical) heterodyne signal obtains (18), (19),

$$W_{\text{abs}} \propto \left\{ |\epsilon_1^0|^2 + |\epsilon_2^0|^2 + 2|\epsilon_1^0||\epsilon_2^0| \cos[2\pi(\nu_1 - \nu_2)t + (\beta - \alpha)] \right\}. \quad (10)$$

A graphical presentation of the function $\operatorname{sech}(h\nu/2kT_e)$ vs. $(h\nu/kT_e)$ is provided in Fig. 1.

The result for the heterodyne case is easily reduced to the direct detection (video) case for a coherent signal by setting $|\epsilon_2^0| = 0$, which yields

$$W_{\text{dir}} \propto |\epsilon_1^0|^2 \left\{ 1 + [\text{sech}(h\nu/2kT_e)] [\cos(4\pi\nu_1 t - 2\alpha - \gamma)] \right\}. \quad (11)$$

Thus, double-frequency intensity fluctuations are discerned for electric-field direct detectors, while they are suppressed for photon-absorption direct detectors which respond simply as $|\epsilon_1^0|^2$.

CONCLUSION

The photon-absorption detector is, by definition, initially in its ground state and functions by the annihilation of a single (in general nonmonochromatic) photon (18). The presence of the difference-frequency signal is understood to arise from our inability to determine from which of the two constituent beams the single photon is absorbed (19). The two-level electric-field detector, on the other hand, has equal probability of being in the lower and in the upper state, so that the pure processes of photon annihilation and photon absorption occur with equal likelihood, and we must add the effects of both. When we are unable to determine which of these processes is occurring, we must add amplitudes rather than squares of amplitudes, thereby allowing interference to occur. Any attempt, in this case, to determine whether photon emission or photon absorption takes place would randomize the phase γ , and thereby wash out the sum- and double-frequency components. In general, then, a photon incident on a video detector induces upward and downward transitions with different probabilities. This, in turn, creates a quantum-mechanical probability density (and charge distribution) that varies in time, producing a current at the incident radiation frequency. The power absorbed then contains a double-frequency signal. For heterodyne detection, in the general case, sum-frequency signals are observed as well.

The foregoing heuristic model yields a simple result for an idealized two-level system. Replacing the operator eq by the non-relativistic Hamiltonian $(\hat{p} - e\hat{A})^2/2m$, where \hat{p} and \hat{A} represent the momentum and vector-potential operators, respectively, would allow transitions more general than electric-dipole, and absorptions of more than one quantum, to occur. A rigorous treatment should consider a collection of such systems (as a model for a bulk photodetector or metal antenna), in the presence of a surrounding reservoir, and should be carried out using the density matrix formalism. It is expected that the effects described here are important in a broad range of systems, including the Josephson detector (26), (27).

Tucker (28) has recently carried out a rigorous analysis of quantum-limited detection in tunnel junction mixers. He demonstrates that nonlinear tunneling devices are predicted to undergo a transition from energy detectors to photon counters at frequencies where the photon energy becomes comparable to the voltage scale of the dc nonlinearity. It is apparent that the character of this result is closely related to the discussion presented here.

REFERENCES

- (1) Fessenden, R. A.: Wireless Signaling. U. S. Patent No. 706,740, August 12, 1902.
- (2) Hogan, J.L.: The Heterodyne Receiving System, and Notes on the Recent Arlington-Salem Tests. Proc. IRE, vol. 1, pp. 75-102, July 1913.
- (3) Armstrong, E. H. : A Study of Heterodyne Amplification by the Electron Relay. Proc. IRE, vol. 5, pp. 145-168, April 1917.
- (4) Armstrong, E. H. : A New System of Short Wave Amplification. Proc. IRE, vol. 9, pp. 3-27, February 1921.
- (5) Forrester, A. T.; Gudmundsen, R. A.; and Johnson, P.O.: Photoelectric Mixing of Incoherent Light. *Phys. Rev.*, vol. 99, pp. 1691-1700, September 1955.
- (6) Javan, A; Ballik, E. A.; and Bond, W. L.: Frequency Characteristics of a Continuous-Wave He-Ne Optical Maser. *J. Opt. Soc. Amer.*, vol. 52, pp. 96-98, January 1962.
- (7) McMurtry, B. J.; and Siegman, A. E. : Photomixing Experiments with Ruby Optical Maser and Traveling Wave Microwave Phototubes. *Appl. Opt.*, vol. 1, pp. 51-53, January 1962.
- (8) Teich, M. C.; Keyes, R. J.; and Kingston, R. H.: Optimum Heterodyne Detection at 10.6 μm in Photoconductive Ge:Cu. *Appl. Phys. Letters*, vol. 9, pp. 357-360, 15 November 1966.
- (9) Teich, M. C. : Infrared Heterodyne Detection. *Proc. IEEE*, vol. 56, pp. 37-46, January 1968 [reprinted in Infrared Detectors, Hudson, R. D., Jr., and Hudson, J. W. Eds. (Dowden, Hutchinson, and Ross, Stroudsburg, 1975)].
- (10) Teich, M. C. : Homodyne Detection of Infrared Radiation from a Moving Diffuse Target. *Proc. IEEE*, vol. 57, pp. 786-792, May 1969.
- (11) Teich, M. C. : Coherent Detection in the Infrared. In Semiconductors and Semimetals, Willardson, R. K., and Beer, A. C. Eds. (Academic, New York, 1970), vol. 5, Infrared Detectors, ch. 9, pp. 361-407.
- (12) Kingston, R. H.: Coherent Optical Radar. *Optics News*, vol. 3, no. 3, pp. 27-31, Summer 1977.
- (13) Elbaum, M; and Teich, M. C.: Heterodyne Detection of Random Gaussian Signals in the Optical and Infrared: Optimization of Pulse Duration. *Opt. Commun.*, vol. 27, pp. 257-261, November 1978.

- (14) Brockman, P.; Hess, R. V.; Staton, L. D.; and Bair, C. H.: DIAL With Heterodyne Detection Including Speckle Noise: Aircraft/Shuttle Measurements of O₃, H₂O, NH₃ With Pulsed Tunable CO₂ Lasers. Heterodyne Systems and Technology, NASA CP-2138, 1980. (Paper 39 of this compilation.)
- (15) Teich, M. C.: Nonlinear Heterodyne Detection. In Topics in Applied Physics, Keyes, R. J. Ed. (Springer-Verlag, Berlin/Heidelberg, 1977), vol. 19, Optical and Infrared Detectors, ch. 7, pp. 229-300.
- (16) Teich, M. C.: Heterodyne Correlation Radiometry. Opt. Engineering, vol. 17, pp. 170-175, March 1978.
- (17) Teich, M. C. : Three-Frequency Heterodyne System for Acquisition and Tracking of Radar and Communications Signals. Appl. Phys. Letters, vol. 15, pp. 420-423, December 1969.
- (18) Teich, M. C. : Field-Theoretical Treatment of Photomixing. Appl. Phys. Letters, vol. 14, pp. 201-203, March 1969.
- (19) Teich, M. C. : Quantum Theory of Heterodyne Detection. In Proc. Third Photoconductivity Conf., Pell, E. M. Ed. (Pergamon, New York, 1971), pp. 1-5.
- (20) Hocker, L. O.; Sokoloff, D. R.; Daneu, V.; Szöke, A.; and Javan, A.: Frequency Mixing in the Infrared and Far-Infrared Using a Metal-to-Metal Point Contact Diode. Appl. Phys. Letters, vol. 12, pp. 401-402, June 1968.
- (21) Abrams, R. L.; and Gandrud, W. B.: Heterodyne Detection of 10.6- μ Radiation by Metal-to-Metal Point Contact Diodes. Appl. Phys. Letters, vol. 17, pp. 150-152, August 1970.
- (22) Mandel, L.: Fluctuations of Photon Beams: The Distribution of the Photo-Electrons. Proc. Phys. Soc. (London), vol. 74, pp. 233-243, 1959.
- (23) Glauber, R. J.: The Quantum Theory of Optical Coherence. Phys. Rev., vol. 130, pp. 2529-2539, June 1963; Coherent and Incoherent States of the Radiation Field. Phys. Rev., vol. 131, pp. 2766-2788, September 1963.
- (24) Kelley, P. L.; and Kleiner, W. H.: Theory of Electromagnetic Field Measurement and Photoelectron Counting. Phys. Rev., vol. 136, pp. A316-A334, October 1964.
- (25) Mandel, L: Antinormally Ordered Correlations and Quantum Counters. Phys. Rev., vol. 152, pp. 438-451, December 1966.
- (26) McDonald, D. G.; Risley, A. S.; Cupp, J. D.; Evenson, K. M.; and Ashley, J. R.: Four-Hundredth-Order Harmonic Mixing of Microwave and Infrared Laser Radiation Using a Josephson Junction and a Maser. Appl. Phys. Letters, vol. 20, pp. 296-299, April 1972.

- (27) Zimmerman, J. E.: Heterodyne Detection with Superconducting Point Contacts and Enhanced Heterodyne Signals from Tightly Coupled Contacts. J. Appl. Phys., vol. 41, pp. 1589-1593, March 1970.
- (28) Tucker, J. R.: Quantum Limited Detection in Tunnel Junction Mixers. IEEE J. Quantum Electron., vol. QE-15, pp. 1234-1258, November 1979.

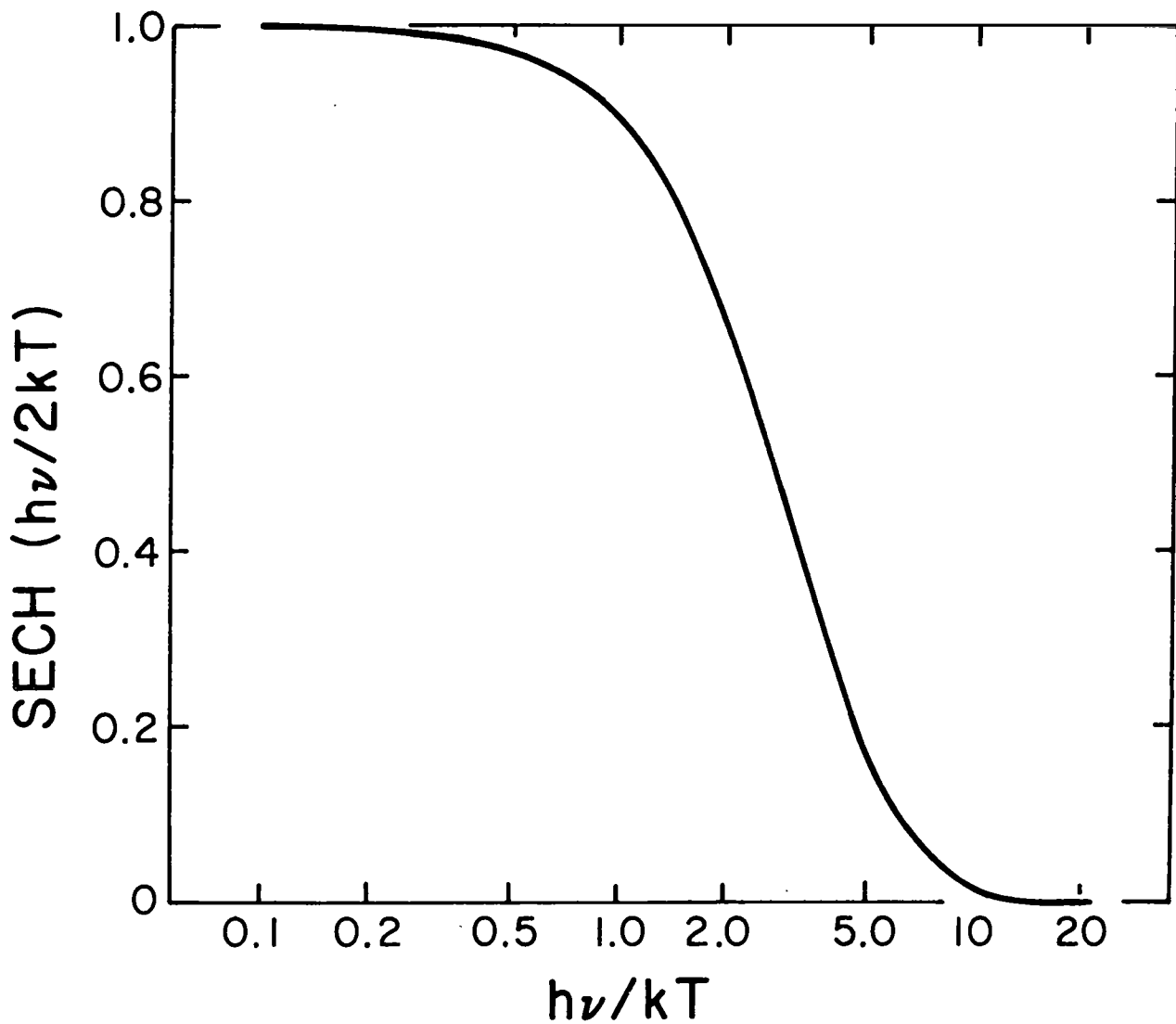


Figure 1.- Factor $\text{sech}(h\nu/2kT)$ vs. $h\nu/kT$. This quantity appears as a coefficient in the absorption/emission interference term.