A Thermal Infrared Heterodyne Receiver with Applications to Astronomy

David D. Snyder Hale
Space Sciences Laboratory, University of California at Berkeley, Berkeley, CA 94720

ABSTRACT

A method for detection of thermal infrared radiation from astronomical sources is discussed from an instrumentation point of view. This method utilizes heterodyne detection, whereby light collected by a telescope is spatially mixed with single-mode radiation from an infrared laser local oscillator onto a HgCdTe photodiode. The operating wavelength range of the receiver is adjustable from approximately 9 to 12 μm, with an intermediate frequency response of the photodiode and associated electronics spanning from 200 to 2800 MHz, or 2.6 GHz in each sideband. The instrumentation which will be discussed is currently in use by the UC Berkeley Infrared Spatial Interferometer (ISI) at the Mount Wilson Observatory. Each of the three ISI elements consists of a 1.65m telescope with complete heterodyne receiver, and is capable of operating as the front end to a variety of different instruments. The sensitivity of the receiver to measuring weak stellar sources will be discussed and contrasted with homodyne receiver performance. Actual spectroscopic and interferometric measurements of astronomical sources will be mentioned.

INTRODUCTION

In principle, heterodyne detection is a process of mixing two signals. For astronomical detection, a stellar signal is combined, or mixed with a stable, fixed-frequency signal (the "local oscillator") onto a nonlinear detector; the frequency of the local oscillator (LO) is chosen to lie at the center of the band to be detected in the source. The multiplication of these two signals produces a detector response which contains frequencies equal to the sum and difference of the two input frequencies, while preserving the amplitude and phase information contained in the original stellar signal. Typically the difference frequency between the stellar and local oscillator signal frequencies (termed the intermediate frequency) is that which is desired because of the often greater flexibility of handling a relatively lower frequency signal.

The heterodyne frequency conversion technique is particularly well-suited for detection of thermal infrared radiation, i.e. frequencies on the order of tens of THz. A mid-infrared stellar signal collected by a telescope may then be spatially mixed with radiation from a laser (which serves as the local oscillator), onto a photodiode which serves as the nonlinear (square law) detector. Intermediate frequencies on the order of GHz are typically obtained, and standard radio techniques are utilized for amplification, transmission, filtering, recording, etc.

Heterodyne receivers are typically characterized by their effective bandwidth and detection efficiency. The detectors themselves exhibit an intermediate frequency resistive-capacitive (RC) time response which is characteristic of the detector construction. Because the intermediate frequency (IF) signal is then amplified, the frequency response of electronic circuits must also be factored into the overall response of the system. The combined frequency response of detector and electronics determines the total effective bandwidth. Effective detection efficiency accounts for not only the quantum efficiency of the detector, but also other factors such as net transmission efficiency of the atmosphere and telescope optics. At present, 11 μm heterodyne receivers have (double sideband) bandwidths of nearly 6 GHz (0.2 cm^{-1}), and net efficiencies of about 30%, which accounts for quantum efficiencies of approximately 40-50% and transmission efficiencies of ~ 60% from star to detector. Larger bandwidths (~ 1 cm^{-1}) and higher quantum efficiencies (~ 80%) may possibly be realized with quantum-well infrared photodetector (QWIP) devices, but these have yet to be proven in the application of astronomical heterodyne detection.

Contact information for D.D.S. Hale: e-mail: david@isi.mtwilson.edu

INSTRUMENTATION

A heterodyne receiver designed for detection of stellar mid-infrared signals has been constructed as follows. Collection of light from an astronomical source is performed with a 1.65m aperture telescope of effective focal length *f*/89. A few centimeters beyond the focal point, the telescope beam is split into two parts by a dichroic beamsplitter which sends mid-infrared (N band) radiation on to the heterodyne detection system, and near-infrared (K band) radiation on to a camera for guiding and tip/tilt wavefront correction. Because the atmospheric index of refraction does not change significantly between the K and N bands, wavefront correction in the near-infrared is adequate to stabilize the mid-infrared signal. The mid-infrared stellar beam reflected off of the dichroic beamsplitter is then sent to a beamsplitter where it is combined with light from a CO_2 laser (which has been properly shaped by a series of lenses in order to ensure the laser and telescope beams are of the same f-number). The combined beams propagate to a mirror which reflects the beam into a signal detection dewar through an antireflection (AR) coated *f*/1.5 aspheric ZnSe lens and an AR coated ZnSe window tilted at 5° from normal incidence to the optical axis in order to prevent back-reflections. Figure 1 shows a schematic[1] of these various beams. The radiation is focused by the lens on a liquid nitrogen cooled HgCdTe photodiode, which produces the intermediate frequency (IF) signal.

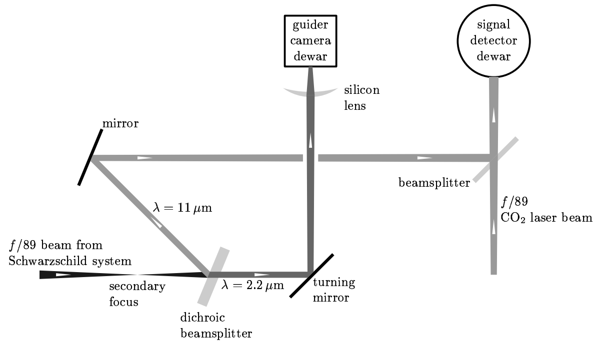

Figure 1: *Schematic diagram of heterodyne mixing optics. Additional optics for chopping are not shown.*

The CO_2 laser local oscillator (LO) is similar to the Freed[2] design, and operates with a meter-long semi-confocal cavity having a concave mirror at one end with a radius of curvature of twice the cavity length, and a grating at the other end in a Littrow configuration acting as a plane mirror where the beam waist is located. Oscillation on undesirable modes is prevented or reduced by apertures inside the cavity in front of the grating and the concave mirror. A variety of CO_2 isotopes can be used for operation on any one of a series of laser frequencies between approximately 9 and 12 µm. Different laser lines are easily selected by setting the tilt of the 80 groove/mm original ruled diffraction grating. Passive frequency stabilization is integral to the laser design. Four invar rods define the cavity length and the gas discharge operates in a sealed tube surrounded by a jacket of circulating and thermally stabilized water. Laser output power is stabilized via a computer-controlled servo loop which is able to make small adjustments to the cavity length by means of a piezoelectric element.

The IF signal from the photodiode, produced by the product of the stellar and LO signals, encompasses a wide range of frequencies within the time responsiveness of the detector. Those frequencies within the detector bandwidth above the CO_2 laser LO frequency form the upper sideband, while frequencies below the LO frequency form the lower sideband. A low-noise FET pre-amp mounted on the same cold finger in the signal dewar as the HgCdTe photodiode amplifies the IF signal before reaching a room temperature amplifier. The combined frequency response of the amplifier electronics and photodiode extends from approximately 200 to 2800 MHz, and is equalized to within a few dB over the entire 2.6 GHz passband. Since both upper and lower sidebands contribute to the IF signal equally, an instrument will have an effective infrared bandwidth of twice the IF bandpass, or 5.2 GHz (~ 0.17 cm^{-1}).

The heterodyne detection system just described has primarily been in use as the receiver for an interferometer[3], but has also served as the front-end for two different spectrometers[4,5]. These three instruments will not be described in detail here, but a few of their measurements will be mentioned in a later section on results.

PERFORMANCE

The fundamental noise power for a heterodyne detector (which detects only one polarization, i.e. that of the LO) is equivalent to an average of one quantum per second per unit bandwidth in the same polarization as the LO. Since heterodyne detection has the ability to measure the phase of a wave, and phase is complimentary to energy or number of quanta, this noise is an inescapable result of quantum mechanics and the uncertainty principle[6,7]. The average signal-to-noise for an ideal photodiode using heterodyne detection is thus[8]

$$\left(\frac{S}{N}\right)_{het} = \frac{P_\nu}{h\nu}\sqrt{\Delta\nu\, t} \tag{1}$$

where P_ν is the signal power in a single polarization per unit bandwidth per unit time, $h\nu$ is the quantum energy, $\Delta\nu$ is the full double sideband (DSB) detector bandwidth, and t is the integration time.

An interesting comparison can be made between heterodyne and homodyne (i.e. direct) detection. For homodyne detection the fundamental noise is dominated by detector dark current and read noise only for very narrow (~ 0.1 cm^{-1}) bandwidths, and is more frequently dominated by thermal background radiation striking the detector when larger bandwidths are used. The average signal-to-noise for an ideal direct detector is

$$\left(\frac{S}{N}\right)_{\text{hom}} = \frac{P_\nu}{h\nu}\sqrt{2\Delta\nu\, t\,\frac{\left(e^{h\nu/kT}-1\right)}{1-\varepsilon}} \tag{2}$$

where T is the temperature of the optics and atmosphere through which the signal is received, and ε is the fractional transmission of radiation reaching the telescope.

From these equations it appears that direct detection may have an advantage by a factor of ~ 45 over heterodyne detection, when comparable narrow (0.2 cm^{-1}) bandwidths are used. However, this advantage is not always realized to this extent. In practice, direct detectors may utilize bandwidths as large as 100 cm^{-1} where they are predominantly background noise limited. Present mid-infrared cameras can have 1σ detection limits of ~ $6x10^{-14}$ W in 1 second of averaging time, about a factor of 10 below the theoretical limit given above, negating much of the expected gain from the broad bandwidth. For bandwidths as narrow as about 1 cm^{-1}, sensitivities of about ~ $2x10^{-15}$ W can be achieved in 1 second by direct detection, roughly 1.5 times the theoretical noise of an ideal heterodyne detector with a DSB bandwidth of ~ 5 GHz.

Heterodyne receivers are of course also not perfectly efficient, but with their much narrower bandwidths they are limited primarily by fundamental quantum fluctuations rather than those caused by sources such as amplifier noise; hence, they are able to achieve sensitivities much closer to their theoretical limit. The constructed heterodyne receiver is routinely able to detect about ~ $2x10^{-15}$ W in a 1 second average, assuming a net efficiency of about 30% and 5.2 GHz DSB bandwidth, which is within a factor of ~ 1.5 of that predicted by theoretical equation (1).

RESULTS

While the heterodyne receiver described herein was designed with the versatility in mind to accept a variety of instruments, it was principally built for use with an interferometer, namely the UC Berkeley Infrared Spatial Interferometer (ISI), a more thorough description of which can be found by Hale[3] *et al.* Interference fringes are obtained by multiplying the IF signal from two (or more) telescopes in an RF correlator. Since interference is carried out at the relatively narrow-band IF frequencies, delay requirements for compensating geometrical pathlength differences between the source and the receiving elements are greatly relaxed, allowing the ISI to integrate for many hours by calculating delays without fringe tracking (as a direct-detection interferometer must do). The ISI has been used primarily to study the spatial distribution of dust around late-type stars by measuring visibility curves and applying spherically symmetric radiative transfer models in order to estimate dust shell inner radii, temperature, and optical depth. The inner radii of many dust shells have been clearly resolved by the ISI, and in some cases measurements of the diameter of the central star have been obtained to 1% precision. Continual monitoring has identified changes in visibilities due to variations in stellar luminosity and also movements and changes in the surrounding dust.

This heterodyne receiver is also well-suited for spectral line research. A spectrometer constructed by Betz[4] *et al*. was used with one of the ISI receivers to measure infrared emission lines of stratospheric ammonia (NH_3) produced by the collision of Comet Shoemaker-Levy 9 with Jupiter in July 1994. Lineshapes of three different NH_3 emission lines at 10.7 μm were measured with a resolving power of ~ 10^7; a detailed analysis of the temporal behavior and NH_3 abundance distributions has recently been made by Fast[9] *et al.*

The ISI heterodyne receivers have also been used with a spectral filterbank, described by Monnier[5] *et al.* Using off-the-shelf 60 MHz RF filters, spectral resolutions of $\lambda/\Delta\lambda$ = 27THz/60MHz = 4.5×10^5 are readily obtained, enough to resolve features arising from Doppler shifts as small as ~ 0.7 km/s. A bank of 32 such filters was used to measure the line profiles of CO_2 in absorption in the Martian atmosphere, and NH_3 in the carbon star IRC+10216. Additionally, this bank of filters may be used in conjunction with the ISI's RF correlator, allowing for interferometry on spectral lines to be carried out. Using this filterbank with the interferometer, Monnier[10] *et al.* were able to locate the molecular formation regions of silane (SiH_4) and ammonia (NH_3) around the carbon stars IRC+10216 and VY CMa.

SUMMARY

A heterodyne receiver of ~ 0.2 cm^{-1} bandwidth for thermal infrared (~ 9 – 12 μm) wavelengths has been constructed, capable of 1σ detection of ~ 2×10^{-15} W from an astronomical source in 1 second of integration. To date, three different instruments have been used with the receiver for spectroscopic and interferometric observations of planetary atmospheres, late-type stars, and the dusty environment surrounding stars. While direct detection will always remain the most sensitive method of detection of mid-infrared astronomical sources for wide bandwidths, heterodyne detection can be competitive with, and for narrow bandwidths such as required for spectral line research, can even surpass direct detection.

REFERENCES

1. E.A. Lipman, *Studies of Evolved Stars with a Mid-Infrared Interferometer*, Ph.D. Dissertation, University of California at Berkeley (1998).
2. C. Freed: *Design and Short-Term Stability of Single-Frequency CO_2 Lasers*, IEEE Journal of Quantum Electronics, vol. QE-4, no.6, 404 (1968).
3. D.D.S. Hale, M. Bester, W.C. Danchi, W. Fitelson, S. Hoss, E.A. Lipman, J.D. Monnier, P.G. Tuthill, and C.H. Townes: *The Berkeley Infrared Spatial Interferometer: A Heterodyne Stellar Interferometer for the Mid-Infrared*, Astrophysical Journal, vol. 537, 998-1012 (2000).
4. A. Betz, E.C. Sutton, and R.A. McLaren: *Infrared Heterodyne Spectroscopy in Astronomy*. In *Laser Spectroscopy III*, Springer Series in Optical Sciences vol.7, eds. J.L. Hall and J.L. Carlsten, 31 (1977).
5. J.D. Monnier, W. Fitelson, W.C. Danchi, and C.H. Townes: *Mid-Infrared Interferometry on Spectral Lines. I. Instrumentation*, Astrophysical Journal Supplement Series, vol. 129, 421-429 (2000).
6. R. Serber & C.H. Townes: *Quantum Electronics*, ed. C.H. Townes, Columbia Univ. Press (1960).
7. H.J. Kimble & D.F. Wells: *Squeezed States of the Electromagnetic Field*, Journal of the Optical Society of America B, vol.4, issue 10, 1450 (1987).
8. R.H. Kingston: *Detection of Optical and Infrared Radiation*, Springer Series in Optical Sciences vol.10, ed. D.L. MacAdam, 25 (1979).
9. K. Fast, T. Kostiuk, P. Romani, F. Espenak, T. Hewagama, A. Betz, R. Boreiko, and T. Livengood: *Temporal Behavior of Stratospheric Ammonia Abundance and Temperature Following the SL9 Impacts*, Icarus, vol. 156, 485-497 (2002).
10. J.D. Monnier, W.C. Danchi, D.S. Hale, P.G. Tuthill, and C.H. Townes: *Mid-Infrared Interferometry on Spectral Lines. III. Ammonia and Silane around IRC+10216 and VY Canis Majoris*, Astrophysical Journal, vol. 543, 868-879 (2000).