# SPECTROMETERS FOR HETERODYNE DETECTION

Andrew Harris
Department of Astronomy, University of Maryland, College Park MD 20742

## ABSTRACT

Heterodyne spectrometers split naturally into two sections: a front-end that shifts the input band to lower frequency and amplifies it, and a back-end which analyzes and detects the signal. This brief review discusses requirements for spectrometer back-ends at submillimeter and far-infrared wavelengths, describes basic spectrometer technology, and indicates some of the important areas for future development.

## INTRODUCTION

A heterodyne radiometer system consists of a sensitive, wideband front-end[1] that provides a strong signal at an appropriate frequency for analysis by a low sensitivity single- or multi-channel detector. Preamplification removes the overriding constraint of low signal loss for spectral processing, allowing flexible and powerful signal analysis. There are many practical advantages to this method. First, the signal may be divided to produce arbitrarily many copies for parallel processing without degrading sensitivity. This allows analysis of an entire spectral band with many detectors at once and is also of considerable value in interferometry. Second, the detectors themselves may be simple with low intrinsic sensitivity, including direct digital signal processing. Third, synthesizing filter bandpasses and filter shapes with sharp corners and steep skirts requires long pathlengths to allow signals with many wavelengths and phases to interfere. A general property of spectroscopy is that the analyzer length $L$ is related to the spectral resolution $R = \lambda/\Delta\lambda$ by $L \approx R\lambda/2$, where $\lambda$ is the signal's wavelength. Heterodyne's preamplification permits the delays to be warm transmission lines, digital shift registers, or any other convenient medium. This can greatly reduce the volume that must be cryogenically cooled in a spectrometer system. In addition, since amplification can overcome loss, filtering with multiple lossy structures opens a wide range of filter response synthesis techniques without degrading the system sensitivity. As a practical matter, heterodyne detection also provides high sensitivity with instruments at moderate cryogenic temperatures: SWAS and MAP are examples of spacecraft with passive cooling alone.

As a general rule, heterodyne observations from short-millimeter through far-infrared wavelengths are appropriate whenever detailed line shape information or precise knowledge of a filter bandpass is necessary. Since the technology for continuum detection is the same as for spectroscopy – a single instead of multiple frequency channels – much of this discussion applies to heterodyne systems in general, and not to spectroscopy alone.

Spectroscopy at wavelengths between the radio and infrared is key to understanding objects as distant as the earliest stars and galaxies to those as nearby as the Earth's atmosphere. Listing the possible applications is a paper in itself, but some of the highlights include the following topics. Submillimeter and far-IR atomic and molecular lines are major gas coolants and are bright from moderately excited gas. These lines are beacons of active star formation from our Galaxy to the earliest galaxies. Observations of bright submillimeter lines is possibly the best way to identify the redshift (age) of these galaxies while providing a valuable handle on interactions and metal production in the era of galaxy formation. Spectroscopy is the only probe of dynamics and kinematics of distant objects, providing measurements of momentum and mechanical energy densities and gravitational mass estimates. Spectral lines are the fingerprints that identify atoms and molecules in atomic and molecular clouds, and line surveys across wide spectral bands reveal the chemical and physical state of material that is forming stars. Detailed line shape information is a valuable tool for identifying and separating different structures that are spatially unresolved; protostellar collapse and accretion disks have unique spectral signatures that probe very small spatial scales. Detailed profile information from lines in planetary atmospheres probes temperature and pressure structure; abundance as a function of altitude measures the altitudes of different atmospheric layers.

*Contact information: Email: harris@astro.umd.edu, phone (301)405-7531*

Figure 1 illustrates some of the constraints on line spectroscopy for single spectral lines in the submillimeter and far-infrared. Spectral linewidths are generally set by dynamics within the source, with the exceptions of planetary atmospheres, whose lines are pressure broadened, and masers, whose lines are narrowed by amplification. The left and bottom scales in Figure 1 are typical velocity scales for single line observations of a range of different kinds of objects, shown by triangles in the body of the figure. To some extent the run of points shows that spectra of individual lines customarily have about a hundred resolution elements; an important exception is wideband line surveys with hundreds of lines and, correspondingly, many thousands of resolution elements. The top and right scales are the corresponding scales in the spectrometers' natural units of frequency, here connected to the velocity at a frequency of 1 THz. The values are representative and will vary by a factor of a few depending on wavelength and individual sources, but they show the resolution and bandwidth scales that the Doppler effect implies.

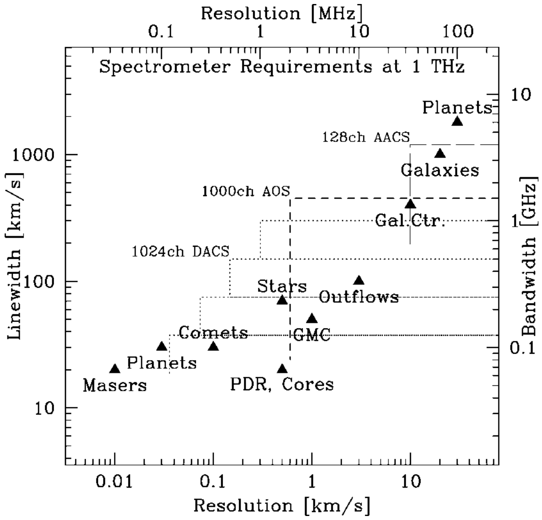


***Figure 1:*** *Connection between velocity and frequency scales for a set of typical sources. Frequencies scale linearly from the 1 THz (300 μm) example shown here.*

The Doppler relation shows that the spectrometer bandwidth must increase linearly with line frequency (a useful relationship for evaluating spectrometers is that the velocity range contained within a 1 GHz bandwidth is numerically equal to the center wavelengths in micrometers). Technology developed for longer wavelength radio astronomy is useful for moderate bandwidth applications. Spectrometer development specifically for submillimeter and far-infrared observations has tended to concentrate on increasing spectrometer bandwidth. Dashed and dotted lines indicate the resolution and bandwidth covered by three broadband spectrometer technologies: digital autocorrelator spectrometers (DACS), acousto-optical spectrometers (AOS), and analog autocorrelation spectrometers (AACS). I discuss these and other technologies below.

## SPECTROMETER TECHNOLOGIES

### Filter banks

Filter bank spectrometers[2] are the most straightforward generalization of a single channel spectrometer: a power divider at the input splits the signal into several or many channels, each of which has a separate bandpass filter and power detector (Figure 2). Filter banks contain many individual parts, some of which require tight mechanical tolerances for manufacture. This tends to result in a complex assembly which is large, massive, and expensive. Because of this, large filter banks have been supplanted by other technologies in many applications. Atmospheric and solar observations are important exceptions, since filter banks can sparsely sample a wide band with different bandwidths, and filter bank systems fly in space on microwave limb sounder instruments.

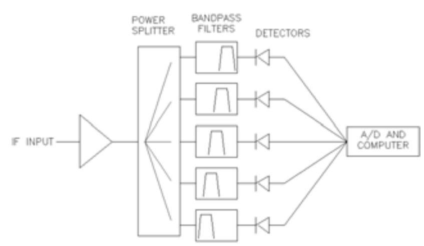


***Figure 2:*** *Schematic view of a filter bank spectrometer*

### Acousto-optical spectrometers

Acousto-optical spectrometers (AOS) have attractive attributes of compact size, low power consumption, and maximum bandwidths of 1 to 2 GHz[3]. In addition to widespread use at ground and airborne millimeter and submillimeter observatories, AOS are in space on SWAS and ODIN, are planned for the Herschel satellite, and have been parts of proposals for other space sciences and planetary missions.

Figure 3 is a schematic drawing of an AOS. The basic principle of operation is the deflection of an optical or infrared laser beam by high-frequency sound waves propagating in a crystal. A small piezoelectric transducer converts the microwave signal to be analyzed into a pressure wave in an optically crystal. Monochromatic laser light scatters from this sound wave to produce the spectrum. One way to understand the scattering is to recognize that the pressure wave consists of alternating dense and rarified regions in the crystal that form a transmission grating. Different spatial wavelengths, corresponding to different frequencies in the input signal, have different grating constants and diffract the laser beam by different angles (this is also known as Bragg diffraction, and the acousto-optical cells are often called Bragg cells). A signal with many frequency components produces pressure waves with a corresponding mix of wavelengths, diffracting laser power over a range of angles. A lens focuses the diffracted beam onto a CCD or other linear array of photodetectors, recording the intensity of diffracted light as a function of angle: the brightness as a function of angle is the spectrum itself.

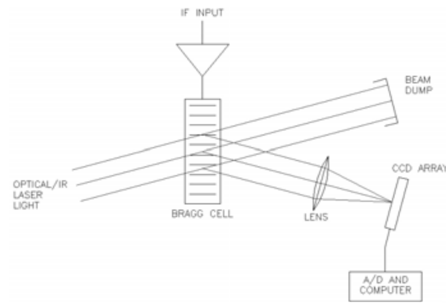


***Figure 3:*** *Schematic view of an AOS*

Groups throughout the world have built and operated AOS. Bragg cells that twist the polarization of the diffracted light, allowing the diffracted light to be separated from other scattered light, enabled the stability and noise performance of AOS to became truly competitive with other types of spectrometers[3]. Multiple spectrometers in a single compact unit are now possible with multiple transducers on a single Bragg cell[3]. A quad-AOS from the University of Cologne that packages four AOS using a single set of optical components as shown in Figure 4; Figure 5 shows the development model of one of these quad-AOS for the FIRST/Herschel observatory, a module 400×170×130 mm in size with 7650 channels covering 4 GHz[3].

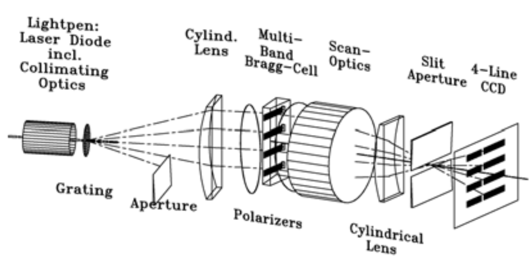


**Figure 4:** Quad-AOS

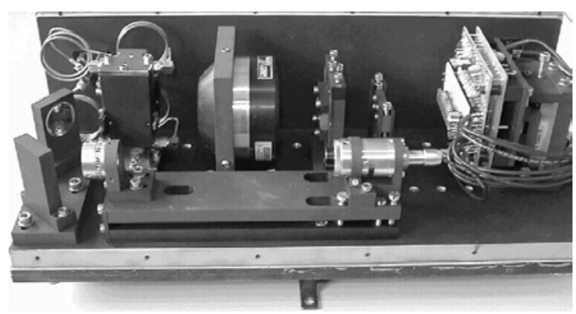

**Figure 5:** Development model of the Herschel AOS

Acoustic wave attenuation limits bandwidths to 1 to 2 GHz with resolution of about 2000 in standard Bragg cell materials. (As for any grating, the resolution depends on the length of the active region, and this is set by attenuation in the crystal.) The University of Cologne and BAe Systems have begun investigating new materials for AOS with bandwidths of about 3 to 4 GHz and approximately 3 MHz resolution. Optical processing may be useful for even wider bandwidths, and one possibility is by spectroscopy of a frequency-modulated laser beam. Scattering of photons off phonons in a modulator's optically active crystals causes a frequency shift by the Doppler effect, adding frequency sidebands to the laser's spectrum. It may be possible to recover the modulator's input signal by high-resolution optical spectroscopy of the laser light. Experiments with a Fabry-Perot interferometer are underway at the University of Cologne, with the goal of a few MHz resolution over 20 GHz, and should soon identify some of the noise and optical resolution limits to this method.

## Correlation spectrometers

It is possible to analyze signals in the time domain as well as in the spectral domain, with the Fourier transform linking the two. Correlation spectrometers that vary the time lag between signals, of which the Michelson interferometer may be the best known, obtain the power spectrum of input signals. The power spectrum $S_{XX}(f)$ of a signal is related to the its autocorrelation function by a Fourier transform:

$$S_{XX}(f) = \int_{-\infty}^{\infty} R_{XX}(\tau)\cos(2\pi f \tau)\, d\tau \quad , \tag{1}$$

where the autocorrelation function $R_{XX}(\tau)$ for signal $X$ as a function of time delay $\tau$ is

$$R_{XX}(\tau) = \langle V_X(t) \cdot V_X(t+\tau) \rangle \ . \tag{2}$$

Here the angle brackets indicate a time average. A correlator system accordingly has four distinct sections: a multiplier to form the product of two voltages, a range of delays, an averaging, and a Fourier transform. Since the signal power may be divided for parallel processing, a radio "Michelson" measures the signals at all lags simultaneously instead of serially. Cross-correlation spectroscopy, for example with spatial interferometry, is possible by simply substituting a signal with voltage $V_Y(t)$ in equation (2) to produce the cross-correlation $R_{XY}(\tau)$ and the cross-spectrum $S_{XY}(f)$.

## Digital correlator

Figure 6 is a schematic diagram of a simple two-level digital autocorrelator. A fast digitizer at the input of the system converts the input signal to a small digital word; all subsequent processing is done in digital logic. This digital signal feeds to a series of multipliers both directly (no time lag) and after a time delay produced in a shift register. Logic multiplies these "prompt" and delayed signals at each clock cycle, with an accumulator summing the products. After a given integration time a computer reads the digital words in the accumulator and Fourier transforms the correlation function to produce the spectrum.

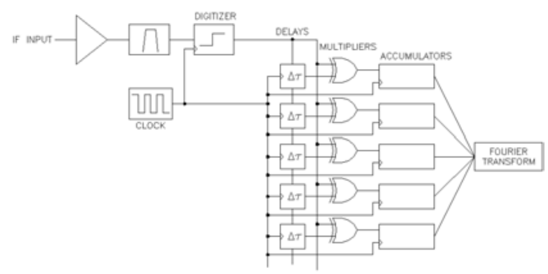


***Figure 6:*** *Schematic layout of a digital autocorrelator*

The advantage of digital correlation is that almost all of the processing is purely digital. Timing is simple to control in a digital system, allowing synthesis of different channel bandwidths without changing spectrometer hardware; this reconfiguration by changing a single clock frequency is a major advantage for digital systems. Other advantages include simple fabrication and hardware replication and freedom from drifts once the signal is in digital form. The disadvantage of digital correlation, ironically, is the same as its strength: much of the processing is purely digital. Proper sampling of the signal requires digitization at a frequency at least twice the signal's input bandwidth. The fastest digitizers currently available have sample rates of approximately 2 gigasamples per second, limiting individual correlators to approximately 1 GHz bandwidths. Power consumption for digital systems scales as $P = VCf$, where $V$ is the bias voltage, $C$ is the device capacitance, and $f$ is the operating frequency; processing with high clock frequencies is an inherently high power technology. Digital logic generally runs at lower speed than the digitizer, so the signal must be farmed out to a number of sub-correlators whose outputs combine to produce the total correlation. The number of digitization levels is a compromise between limiting the number of data products and signal to noise degradation. A common compromise, three level digitization, gives 81% of the signal-to-noise compared with an undigitized signal[4], or a loss of nearly one third of the observing time.

Speed requirements have led to a number of custom integrated circuit designs that incorporate many multipliers and the accumulators with the shift register in a single package. Recent performance gains in field programmable gate arrays (FPGA), digital hardware packages whose internal configurations can be downloaded from software, allow flexible correlators made with high performance general purpose logic packages. Although interconnections circumscribe the applications, this approach essentially couples a fast digitizer with a very flexible and fast processor that can be configured for many different kinds of processing.

The freedom from many drifts, ease of fabrication, and excellent control of spectral channel frequencies and shapes make digital correlators the technology of choice for observations with thousands of narrowband channels over moderate bandwidths. Relatively narrowband correlators are part of the instrument complement on ODIN and Herschel[5].

## Analog correlator

Analog correlators exchange the configuration flexibility of digital correlators for high efficiency and moderate spectral resolution across very wide bandwidths. Single-lag correlators of this type in a filter bank array are at the heart of the CBI and DASI cosmic background imaging spectrometers[6]. Figure 7 is a schematic diagram of an analog lag correlator for spectroscopy: multiplication is provided by fast transistor circuits and time delays are provided by short sections of transmission line. Fully analog processing of all the high frequency signals leaves digital processing to low-frequency signals alone. In addition to high efficiency, these spectrometers have low power dissipation.

The best-developed example of this technology for spectroscopy is the WASP family of lag spectrometers[7], 128 lag spectrometers with approximately 4 GHz of bandwidth in a compact unit that dissipates about 50 W. Microwave transistors in a Gilbert cell configuration are the multipliers, and microstrip transmission lines form the delays. Correlators are intrinsically phase sensitive, and phase switching removes most internal drifts to give excellent stability.

Spectrometer bandwidth is limited mainly by the multiplier circuit. Figure 8 is shows the frequency response of a Gilbert-cell multiplier custom MMIC designed by S. Maas for wideband correlators. Results from the prototype show that the response and device noise are suitable for correlators covering at least 2-16 GHz. Analog lag correlators using other multiplier devices are also being investigated for wideband millimeter spectroscopy[8] and cosmic background imaging[9].

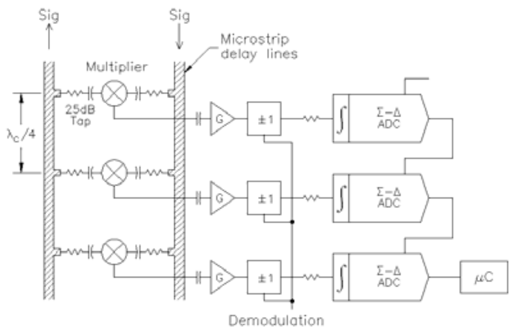


***Figure 7:*** *Schematic view of a WASP analog lag correlator segment.*

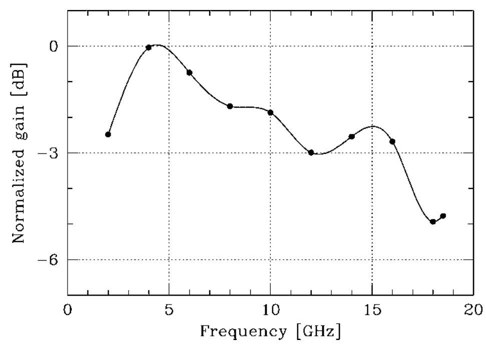


***Figure 8:*** *Response of a new Gilbert-cell multiplier custom MMIC for ultra-wideband analog correlators.*

# FOCUS FOR FUTURE DEVELOPMENT

The goal of heterodyne spectroscopy is clear: measure all the photons in the receiver band at all times. Technologies suitable for submillimeter and far-infrared spectrometers must continue to develop to keep pace with the possibilities presented by new generations of wideband front-ends. The current state of affairs, in order of bandwidth but noting that there are no sharp dividing lines, are analog correlation spectrometers with bandwidths of a few GHz; acousto-optical spectrometers with bandwidths to 1.5 GHz; and digital correlators with bandwidths to 1 GHz. Stacking multiple spectrometers in frequency with analog preprocessing covers wider bandwidths. Small nonlinearities can generate steps between these sub-bands, however, a problem that is especially acute for digital systems.

In some order of bandwidth the following lists areas ripe for future development.

- Spectrometers for extragalactic line-resolved spectroscopy and line searches from objects in the early universe will need hundreds to thousands of channels over bandwidths from about five to fifty gigahertz. Stability and freedom from low-level systematic spectral features are critical for observations of these weak sources. Eliminating structure from sub-band stacking suggests that spectrometers with inherently broad bandwidths are needed, and cross-correlation may be important. The likely technology for this application are analog correlators or optical schemes such as laser sideband generation. Spectrometers with these properties will also be useful for observations of lines in planetary tropospheres.
- Spectrometers with many thousands of channels over a few gigahertz are necessary for efficient line surveys that characterize the chemical and physical state of a molecular cloud. Surveys in bands crowded with many lines require very high dynamic range and stable and well-defined channel shapes

to accurately measure and separate components from different regions. The high resolution but relative insensitivity to broad spectral baseline errors suggests that stacking digital correlators or acousto-optical spectrometers is a good solution. Increasing the bandwidth of the individual spectrometers is important to reduce system size and the complexity of the frequency stacking hardware.

- Spectrometers for heterodyne focal plane arrays need a few hundred megahertz to cover sources within the Galaxy. A subsequent generation of spectrometers with bandwidth sufficient for external galaxies will be interesting, but the number of pixels can be much smaller than a galactic mapper: the size of extragalactic sources becomes small in the submillimeter and far-infrared in most lines. Cost, power consumption, and volume will be very important for fielding simple individual spectrometers that will be combined to make a large system.
- Detection with cross-correlation will become more important with time. One obvious application is spatial interferometry, which requires either single- or multi-channel correlator back-ends. Sensitive interferometry is only possible with large numbers of photons, so heterodyne detection's quantum noise will not be a disadvantage for spectral line observations. (Lossless signal division between baselines may be a substantial advantage.) An additional advantage of cross-correlation is high stability, a fact familiar from single-dish continuum instruments and spatial interferometers. Cosmic background imagers already use cross-correlation partly for this reason. Some spectral resolution is needed to retain a large field of view: smearing by fringes with different wavelengths limits an interferometer's synthesized field of view for fractional bandwidths smaller than the antenna diameter divided by the baseline. This can become a significant limit for compact interferometers that image the CMB. Single dish instruments based on correlation or continuous comparison architectures will have the high efficiency and stability needed for long integrations on weak distant sources.
- Wideband microwave device development is an important support technology. Devices might include fast digitizers; wideband multipliers, amplifiers, and passive components; and optical deflectors and modulators. Technologies with low power consumption, small volume, and low parts count (for high reliability) will be a priority for airborne and space instruments.

As receiver technology pushes to higher frequencies and wider instantaneous bandwidths the bulk of development effort for suitable back-ends for submillimeter and far-infrared heterodyne spectroscopy will continue to be in producing larger bandwidths and adaptation of technologies from millimeter and centimeter wavelength developments.


## ACKNOWLEDGEMENTS

I thank R. Schieder for providing information on acousto-optical spectrometers developments and D. Hawkins for discussions of FPGA-based digital correlators. Work on wideband analog correlators at the University of Maryland is sponsored by NASA grant NAG5-9075, NSF grant AST-9819747, USRA funds for SOFIA heterodyne spectroscopy, and by the University of Maryland.


## REFERENCES


1. Zmuidzinas, J., this Proceedings; Gaier, T., this Proceedings
2. H. Penfield, *Multichannel-Filter Spectrometers*, in Meth. Expt. Phys. vol. 12B, ed. M.L. Meeks, Academic Press (1976)
3. R. Schieder, V. Tolls, and G. Winnewisser, *The Cologne Acousto-Optical Spectrometers*, Exp. Astronomy 1, 101 (1989); R. Schieder et al., *Design of Large Bandwidth Acousto-Optical Spectrometers*, Proc. SPIE 3357, 359 (1998); R. Schieder, *State of the Art of AOS*, Proceedings of the COST-712 Workshop (1999)
4. B.F. Cooper, *Autocorrelation Spectrometers*, in Meth. Expt. Phys. vol. 12B, ed. M.L. Meeks, Academic Press (1976); A.R. Thompson, J. Moran, and G.W. Swenson, *Interferometry and Synthesis in Radio Astronomy*, Wiley, ch. 8, 1986
5. L. Ravera et al., *Wide Band Digital Auto Correlator for FIRST*, Proc. SPIE 3357, 368 (1998)
6. S. Padin, J.K. Cartwright, M.C. Shepherd, J.K Yamasaki, and W.L. Holzapfel, *A wideband analog Correlator for Microwave Background Observations*, IEEE Trans. Instrum. Meas. 50, 1234 (2001)
7. A.I. Harris and J. Zmuidzinas, *A Wideband Lag Correlator for Heterodyne Spectroscopy of Broad Astronomical and Atmospheric Spectral Lines,* Rev. Sci. Inst. 72, 1531 (2001)
8. N. Erickson, priv. comm..
9. AMiBA correlator at http://www.asiaa.sinica.edu.tw/amiba/