A DOUBLE-FOURIER INTERFEROMETER TESTBED FOR SPECTRAL IMAGING IN THE FAR-INFRARED

JEREMY P. SCOTT
Master of Science, University of Lethbridge, 2017

A thesis submitted
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

EARTH, SPACE, AND PHYSICAL SCIENCE

Department of Physics and Astronomy
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

A DOUBLE-FOURIER INTERFEROMETER TESTBED FOR SPECTRAL IMAGING IN THE FAR-INFRARED

JEREMY P. SCOTT

Date of Defence: July 26, 2024

Dr. L. Spencer Thesis Supervisor	Associate Professor	Ph.D.
Dr. D. Naylor Thesis Examination Committee Member	Professor	Ph.D.
Dr. A. Tetarenko Thesis Examination Committee Member	Assistant Professor	Ph.D.
Dr. H. Matsuo External Examiner National Astronomical Observatory of Japan	Associate Professor	Ph.D.
Dr. C. Povey Chair, Thesis Examination Committee	Instructor III	Ph.D.

Dedication

This work is dedicated to my mother for her love and support.

Abstract

The far-infrared band of the electromagnetic spectrum is one of the least explored areas of astronomy despite its critical importance for understanding the formation and evolution of planets, stars, and galaxies. Part of this disparity is explained by the challenges associated with conducting observations with the sub-arcsecond spatial resolution required to resolve the structures of interest. Much of the utility of far-infrared astronomy is due to the wealth of spectral features residing within this band, which probe the full range of physical conditions within the interstellar medium. As such, the ideal instrument is one that provides broad spectral imaging capabilities with enhanced spatial resolution. A promising candidate is the double-Fourier interferometer that combines a spatial interferometer with a Fourier transform spectrometer. The two techniques have separately been studied extensively and used productively, however, their combination into a single instrument is relatively new and has yet to be fully validated.

The primary objective of this thesis, and the work presented within, is to validate the double-Fourier technique by demonstrating the key observing capabilities of the instrument. First, it must be shown that the data collected with the instrument can be used with traditional aperture synthesis techniques to produce images with enhanced spatial resolution. Second, it must be shown that the instrument is sensitive to spectral variation within the source, and thus, allow for the production of spectral images with enhanced spatial resolution. This is the main advantage a double-Fourier system has over transitional narrow band spatial interferometers. Third, it must be demonstrated that these techniques can be extended to wide field applications where multiple detectors, sensitive to different regions of the source plane, are used within a detector array.

The observing capabilities listed above were demonstrated independently using simple test cases. This thesis also includes a detailed report on the design, operation, and characterization of the particular double-Fourier interferometer used in this work. Analytical models and empirical measurements were leveraged to predict and better understand the results and limitations of the instrument observations. This analysis extends to the unconventional bolometer technology that was employed for the detector system of the interferometer. The full observation process is presented including data acquisition, calibration, and each step of image reconstruction. Both real data and simulated models are provided when discussing data reduction, and when presenting final results, which provide confidence in the accuracy and quality of the analysis. With the inclusion of a review of the underlying theoretical framework for spatial interferometry and Fourier transform spectroscopy, this thesis serves as an appreciably comprehensive end-to-end reference for double-Fourier interferometry.

Acknowledgments

This work would not have been possible without my thesis supervisor Dr. Locke Spencer. His academic, and financial contributions were critical factors in bringing the instrument and supporting infrastructure into existence, as well as facilitating the research presented in this document. His support early in my academic career was invaluable, and likely played a decisive role in my pursuit of an astronomy research-based career. I must also thank him for his patience and generosity throughout the years, without which this experience would have been significantly less enjoyable. Dr. David Naylor's contribution to the project, in association with Blue Sky Spectroscopy, largely consisted of the many pieces of equipment he allowed me to borrow for the purpose of alignment and characterizing the system. Perhaps more importantly, I would like to thank him for his engaging lectures on astronomy, which contributed to my interest in the subject, and for his leadership in the Astronomical Instrumentation Group, which is largely responsible for why the University of Lethbridge is a quality institution for this kind of research. I would like to thank Dr. David Leisawitz, and Dr. Lee Mundy for their leadership in the double-Fourier interferometry community. My interaction with these individuals as part of the Space Interferometer for Cosmic Evolution collaboration has resulted in conversations that have improved my understanding and confidence in the subject. I must also thank Dr. Berke Ricketti for sharing his experience and offering advice and encouragement on the subject of postdoctoral opportunities.

A number of fellow students have contributed to aspects of the interferometer system's development. Vincent Weiler developed the largely unmodified Python library that was used to interface with the Aerotech stages. He cut the aperture masks used as source targets, and he contributed significantly to the software used to monitor the laboratory environment.

Chris Benson performed the preliminary aliment of the Fourier transform spectrometer mirrors, and produced mylar beamsplitters that were used to facilitate further alignment. Dinula De Silva provided the initial Python framework to interface with the detector system which served as a foundation for the current implementation. Adam Sundberg assisted with the design and played a leading role in developing the model for the beamsplitter mounting assembly. Geoff Sitwell designed the single mirror mounts. I would like to thank them all for their contributions.

Finally, I would like to thank Rashmi Sudiwala and Carole Tucker from Cardiff University. Rashmi, in association with QMC Instruments helped significantly with the commissioning of the detector system and provided valuable insight into its operation. Carole, in association with Celtic Terahertz Technology provided the low-pass optical filter material used in certain experiments.

Contents

List of Tables

List of Figures

List of Abbreviations

AGN Active Galactic Nuclei.

ALMA Atacama Large Millimeter/submillimeter Array.

BETTII The Balloon Experimental Twin Telescope for Infrared Interferometry.

BSF Beam Spread Function.

CIB Cosmic Infrared Background.

DFI Double-Fourier Interferometer.

DFT Discrete Fourier Transform.

DTFT Discrete-Time Fourier Transform.

ESA European Space Agency.

FFT Fast Fourier Transform.

FIR Far-Infrared.

FIRI Far-InfraRed Interferometer.

FITE Far-Infrared Interferometer Telescope Experiment.

FOV Field of View.

FTS Fourier Transform Spectrometer.

FWHM Full Width at Half Maximum.

IDFT Inverse Discrete Fourier Transform.

ILS Instrument Line Shape.

IRAS Infrared Astronomical Satellite.

ISM Interstellar Medium.

ISO Infrared Space Observatory.

JWST James Webb Space Telescope.

KID Kinetic Inductance Detectors.

LISA Laser Interferometer Space Antenna.

MCF Mutual Coherence Function.

MIR Mid-Infrared.

NASA National Aeronautics and Space Administration.

NEP Noise Equivalent Power.

NIR Near-Infrared.

OAP Off-Axis Parabolic.

OPD Optical Path Difference.

PID Proportional-Integral–Derivative.

PSF Point Spread Function.

PSO Position Synchronized Output.

QCD Quantum Capacitance Detectors.

RHS Right-Hand Side.

RMS Root Mean Squared.

SMBH Super Massive Black Holes.

SNR Signal to Noise Ratio.

SPECS Submillimeter Probe of the Evolution of Cosmic Structure.

SPIRIT Space Infrared Interferometric Telescope.

SQUID Superconducting Quantum Interference Devices.

TES Transition Edge Sensor.

VDI Virginia Diodes.

YSO Young Stellar Objects.

ZPD Zero Path Difference.

Chapter 1

Introduction

Light, consisting of vibrations of the electromagnetic field, serves as the primary tool for investigating the universe. Indeed, the exotic cosmic environments manifest physical conditions on time scales that cannot be achieved in the lab, and the structures of interest are so incomprehensibly far away that there is little hope of probing them directly. Fortunately, a rich complement of information about the physical (e.g., temperature, density, etc.) and chemical characteristics of an emission source is encoded in light through its spectral content, intensity, polarization, and time variation. Information about the matter between the source of emission and observer can also be inferred from absorption and gravitational lensing. Although astronomy has made tremendous progress towards understanding the nature and evolution of the universe and the sub-structures within it, many questions are still outstanding. Some of the gaps in our understanding can be attributed to the lack of observations in the Far-Infrared (FIR) band of the electromagnetic spectrum. In particular, the kinds of observations that are most insightful require sub-arcsecond spatial resolution which can only be achieved with a spatial interferometer. A compelling implementation of such an instrument combines a spatial and spectral interferometer in order to provide spectral images with enhanced spatial resolution. Such an instrument has been associated with various names in the literature, but in this work, it will be consistently referred to as a Double-Fourier Interferometer (DFI).

Although the FIR and broader astronomy community have recognized the importance of such an instrument [1], resulting in National Aeronautics and Space Administration

(NASA) and European Space Agency (ESA) led mission concept studies [2, 3], no such mission has been funded to date. The cost of such a missions in comparison to competing proposals is certainly a relevant factor, but so is the demanding detector sensitivity requirements and the novel technology required for a space based interferometer. Significant advances are close to achieving the desired detector sensitivity on detector arrays with large pixel counts [4]. The planned Laser Interferometer Space Antenna (LISA) [5], will contribute significantly to overcoming engineering challenges associated with a space interferometer and provide confidence in its feasibility. Although the theory of a DFI rests on a firm foundation, there are conspicuously few experimental demonstrations of this observation technique. A comprehensive and compelling demonstration would be of significant value to the FIR community, both in terms of arguing the case of a DFI, but also for providing a reference that describes in detail the data reduction required to convert the instrument response into high resolution images. This thesis provides such a demonstration and probes independently the most important aspects of the DFI technique.

1.1 Infrared Universe

1.1.1 Overview

For the purposes of astronomy, the FIR is typically considered to consist of electromagnetic radiation within the $25-350\,\mu$m ($1.20-0.85\,$THz) band. An often repeated statement that captures the importance of FIR astronomy is that half of the energy emitted by stars and accreting objects comes to us in the form of FIR radiation [6]. This Cosmic Infrared Background (CIB) is due to the fact that stars are formed within, and Active Galactic Nuclei (AGN) are often obscured by, clouds of gas and dust that absorb or otherwise scatter light with energies greater than the FIR. For energy to escape these clouds, light must be converted, through various mechanisms, to a wavelength that has a low probability of interacting with the surrounding matter. The limiting factor is often interstellar dust: small ($0.1-1.0\,\mu$m) refractory solid particles composed of heavy elements like carbon, oxygen,

magnesium and iron [7]. In the Rayleigh limit, when the wavelength of light is larger than the diameter of dust grains, the scattering cross section depends strongly on wavelength $\propto \lambda^{-4}$ [8]. Beyond this limit is the Mie scattering regime, which is largely wavelength independent [8]. As such, light with a wavelength comparable to or smaller than dust grains will often be scattered while FIR radiation, with its longer wavelengths, pass largely unobstructed through dust (see Fig. 1.1). Since dust is a solid, it expresses thermal and optical properties similar to a black body. With each photon interaction is the possibility that the energy is absorbed and re-emitted as thermal radiation. As such, the most productive way to study the formation and evolution of galaxies, stars, and planets, is through FIR observations which allow the most direct measurements of the underlying structures.

Objects with temperatures in the range of $\sim 10 - 100\,\mathrm{K}$ that radiate as black bodies will express a spectrum with peak intensity in the FIR. These are typical temperature ranges for prestellar cores, protostars, and protoplanetary disks which represent some of the most important, yet poorly constrained, stages of star and planet formation. FIR observations provide a significant sensitivity advantage for such objects since they are brightest in this band. Dust is another important component that absorbs light emitted from protostars and re-emits this energy as thermal radiation that is brightest in the FIR. Although dust is considered a nuisance at shorter wavelength, it provides a powerful diagnostic for total luminosity in galaxies, star formation rates, dynamics in protoplanetary disks, and metalicity.

Another significant advantage of FIR astronomy is that this band contains the rest frame emission frequencies of many of the most interesting spectral lines and continuum features for the purpose of studying the chemistry and physical conditions of the Interstellar Medium (ISM) and Young Stellar Objects (YSO). In molecular clouds, interaction energies are sufficient to excite rotational energy states associated with CO, HD, OH, HF, H_2O, SO_2, NH_3, CH_3OH, CH_3OCH_3O, and more [9]. Of particular note is CO which is one of the most abundant molecules in the ISM and traces warm molecular gas, typically associated with feedback processes from other nearby stars [10]. The most abundant molecule in

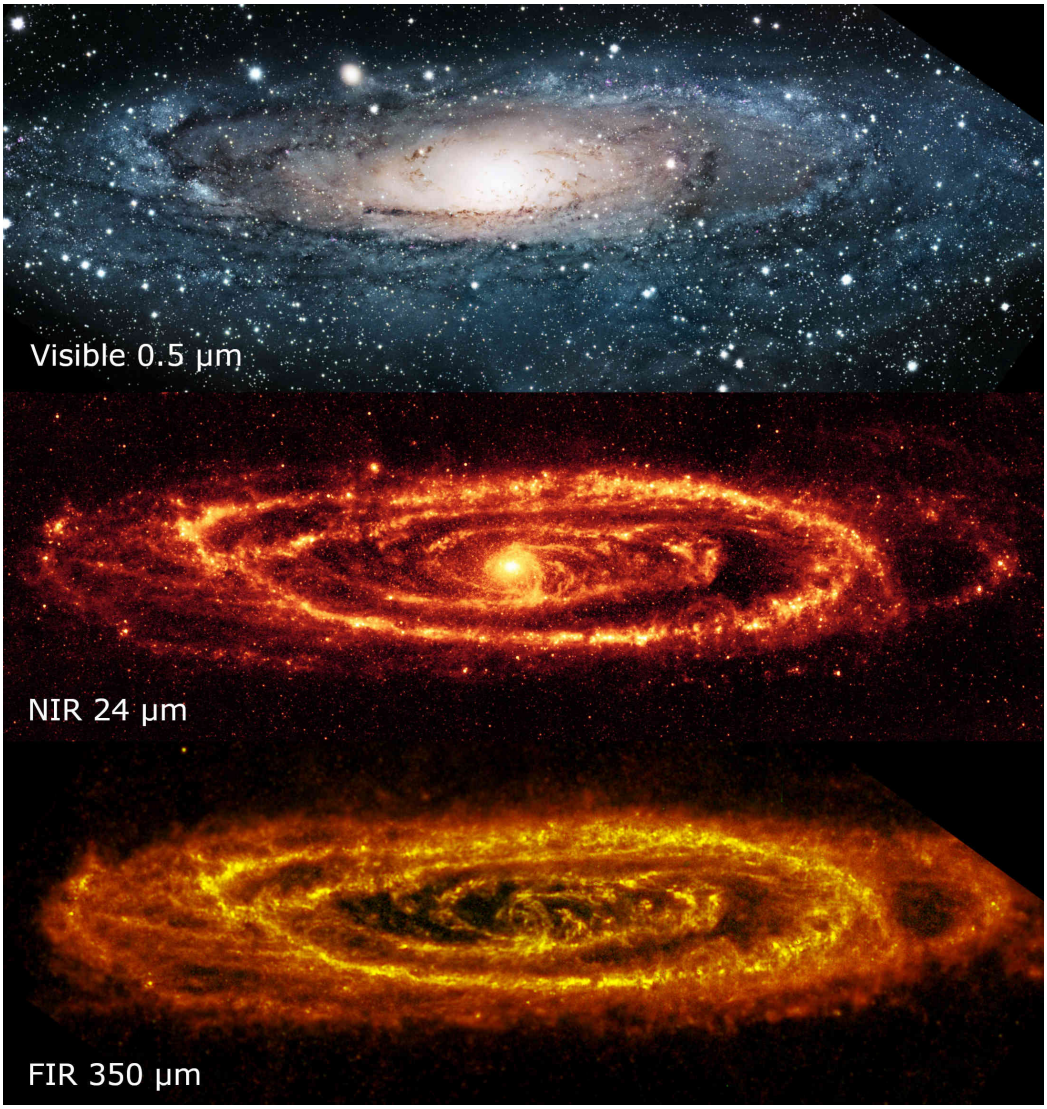


Figure 1.1: The Andromeda galaxy at different wavelengths. At visible wavelengths (top; credit: R. Gendler), much of the luminous matter in the galaxy is obscured by dust. Near-infrared images (NIR; middle; credit: NASA/JPL-Caltech/K. Gordon, University of Arizona), reveal the warm molecular clouds, while far-infrared images (bottom; credit: ESA/Herschel/PACS/SPIRE/J.Fritz, U.Gent), show cooler gas and dust. Much of the light emitted from stars, and particularly young stars embedded within molecular clouds, is absorbed and re-emitted at infrared wavelengths.

the universe is H_2 and is ubiquitous in molecular clouds. Observations of H_2 could provide unique access to the molecular mass budget of the ISM, especially in environments where other tracers are frozen onto dust grains. Unfortunately, this symmetric molecule lacks an electric dipole moment, possessing only a weak quadrupole, and does not express strong vibrational or rotational emission features. However, the asymmetric isotopologue HD has a slight dipole moment and its first two rotational states have corresponding features in the FIR. Perhaps the most interesting molecule in the ISM is H_2O. Water is a crucial component for life sustaining planets, and tracking the distribution of water in protoplanetary disks provides information about how and under what conditions water is deposited on these planets. The FIR has access to both rotational lines of water vapour and continuum emission features of water ice. As such, H_2O emission provides access to a wide range of physical conditions, from high energy feedback processes in molecular clouds to cold material beyond the snow line in protoplanetary disks [9].

The spectral features discussed so far originate from relatively cold environments with interaction energies low enough to sustain molecules. In warm photodissociation and hot ionized regions, the radiation field is dominated by higher energy photons. These photons are outside of the FIR and in many cases are obscured by gas and dust. However, under certain conditions neutral atoms and ions can have low energy excitations associated with the orientation of electron spin due to spin-orbit coupling [11]. Emission resulting from electron spin state transitions are called fine structure lines, and their energies are often low enough to escape the enshrouding gas and dust. For this reason, fine structure lines constitute the dominant cooling mechanism of the ISM and are particularly useful tracers for star formation and AGN accretion. Some of the most important species with rest frame emission in the FIR include [CII] (158 μm), [OI] (63, 146 μm), [OIII] (52, 88 μm), and [NII] (122, 205 μm). Other species that probe highly energetic regions like [ArII], [SIV], [NeV], and [OIV] have rest frame emission at lower wavelengths that are shifted into the FIR at moderate to high redshifts (see § 1.2.2). Such high energy regions are associated with AGN

and starburst galaxies and can be used to study how these structures formed and evolved in the early stages of reionization $z \approx 10$ [12] and during the era of peak star formation $z \approx 2$ [13]. A more extensive survey of modern FIR science cases can be found in the The PRobe for-Infrared Mission for Astrophysics (PRIMA) science book [14].

1.1.2 Limitations

Clearly the FIR has a unique advantage and significant discovery potential in astronomy, yet the FIR remains one of the most poorly studied bands. There are three main reasons for this which can be summarized as sensitivity requirements, the need for space based observations, and the spatial resolution constraints of single aperture telescopes.

Sensitivity

The development of FIR detectors was slow to start largely due to the lack of commercial applications to motivate funding such technology. Infrared detectors gained interest among astronomers in the 1960s and by the mid 1980s, state of the art could achieve detectors with a Noise Equivalent Power (NEP) of $\sim 10^{-17}\,\mathrm{W}/\sqrt{\mathrm{Hz}}$ [15]. However, the sensitivity requirement such that the dominant noise source is photon statistics from background emission is $\sim 10^{-20}\,\mathrm{W}/\sqrt{\mathrm{Hz}}$ [4]. Three promising FIR detector technologies with state of the art performance that can meet this requirement include Transition Edge Sensor (TES), Kinetic Inductance Detectors (KID), and Quantum Capacitance Detectors (QCD). These technologies are discussed in more detail below.

It wasn't until the refinement of TES technology in the mid 90s that significant progress was made towards the goal of background limited sensitivity. TES detectors exploit the rapid change in resistance as a function of temperature in the transition region of metals between superconducting and normal conduction states [16]. These detectors were initially developed in the 1940s but failed to gain wide spread use due to the difficulties associated with maintaining the sensors at the temperature of their transitions. Originally these detectors were current biased. As the sensor warmed under optical loading its resistance

increased, and since Joule heating is proportional to resistance in the current biased configuration, the detectors experienced thermal runaway resulting in poor stability. Another challenge was readout electronics. The available semiconductor amplifiers were poorly matched to the low impedance of TES circuits resulting in noisy readout. The first of these challenges was overcome using voltage biased TES elements that produced negative electrothermal feedback and self-regulated the thermal balance of the detector [17]. This self-regulation allowed for larger format detector arrays since negative feedback made the detectors less sensitive to their thermal environment. As such, detectors with slight manufacturing variations could obtain similar performance when sharing the same thermal bias. The second challenge was overcome with the introduction of very sensitive magnetic field sensors called Superconducting Quantum Interference Devices (SQUID), which were inductively coupled to the TES circuit to measure changes in electrical current [18]. Modern TES detectors can achieve an NEP of $\sim 10^{-19}\,\mathrm{W}/\sqrt{\mathrm{Hz}}$ and show potential for even more improvement [4]. In addition to sensitivity improvements, research is ongoing with respect to integrating several detectors into large format arrays which is primarily limited by the large number of wires required to do so. These wires introduce additional electrical and thermal effects which reduce performance.

Another promising detector technology is KID [19]. These detectors make use of charge carrier mass and the fact that currents resulting from the motion of these charge carriers has an associated kinetic energy. In order to change the current direction, the electric field must do work on the charge carriers proportional to the product of mass and velocity squared of the charge carriers. This description is equivalent to electrical inductance but is a result of the kinematics of the charge carrier. As such, the phenomena is called kinetic inductance. In superconductors, pairs of electrons are bound through electron-phonon interactions at low temperatures. The sum of electron spin for two electrons is an integer, and the electrons bound in this way form a composite boson called a Cooper pair. Cooper pairs are thus exempt from the Pauli exclusion principle and multiple Cooper pairs can share the same

quantum state. It is this property that allows them to move without resistance through a superconductor. Electrons, on the other hand, must wait for quantum states (i.e., adjacent positions in the conductor with an accessible energy level) to become available before they can move. As such, current is most efficiently facilitated by Cooper pairs, and due to their higher velocity, they are primarily responsible for the kinetic inductance within a superconducting wire. If the superconductor is exposed to photons with energy sufficient to break Cooper pairs, quasiparticles with the properties of electrons will be created. This effectively increases kinetic inductance, as Cooper pairs must travel faster to maintain the same current, and simultaneously increases the resistance of the conductor as a result of the quasiparticles [19].

To use these physical properties as a detector, the super conducting material, treated as a variable inductor, is integrated into an LC resonator circuit. The circuit is then capacitively coupled to a common read line that is shared by many other such resonators. The resonant frequency can be tuned by changing the capacitance of the LC circuit, with each detector on the read line having a unique resonant frequency. Detectors can be read by sending a broad spectrum electric signal down the read line with a bandwidth that encompass the resonant frequency of all detectors on the line. Each detector will parasitically extract energy from the input signal that is close to its resonant frequency with the amount of energy extracted inversely proportional to the quality factor of the resonator. At the output of the read line, the spectrum will be ideally flat with absorption features at the resonant frequencies of the detectors. When subjected to optical loading, inductance and resistance of the LC resonator will increases resulting in a lower resonant frequency and lower quality factor. These effects manifest in the output signal as a frequency shift and reduction in amplitude of the absorption feature. This technology, developed in the early 2000s, has achieved sensitivity that is competitive with TES detectors [4]. An attractive feature of KIDs is their inherent frequency division multiplexing capabilities. Multiple detectors can be read using a single read line, easily overcoming one of the main challenges of TES

detectors.

An even more recent detector technology with considerable potential is QCD [20]. In this case a superconducting island is coupled to a superconducting absorber through a tunnel junction. The island is further capacitively coupled to an LC resonator by a gate capacitor. The tunnel junction, island, and gate capacitor form a single Cooper pair box. The island is physically small and forms a quantum system that is sensitive to a small number of electrons such that when an electron is on the island, it significantly changes the capacitance of the gate capacitor. The operating principle of a QCD is that when light with sufficient energy to break Cooper pairs is incident on the absorber, quasiparticles, which are effectively electrons, will be created. An electron will then tunnel across the junction and occupy the absorber, significantly altering the capacitance of the gate capacitor. The electron will then tunnel back and forth between the absorber and island multiple times before relaxing back into a Cooper pair. The rate at which electrons tunnel across the junction, and their average occupancy time on the island, is proportional to electron density in the absorber, and thus the incident optical power. Detector readout is largely the same as with KIDs, but instead of a continuous frequency shift, the resonant frequency of each detector will be one of two discrete states corresponding to electron occupancy of the island. The average state is used as an estimate for the power incident on the absorber. QCDs possess the same frequency domain multiplexing advantage as KIDs, and since the island is sensitive to single electron occupancy, they can achieve single photon detection [21], making them the most sensitive FIR technology to date. However, QCDs were first demonstrated in 2009, and lag behind both TES and KID in terms of development. The are also quite easily saturated and presently have low dynamic range.

In addition to detector sensitivity, it is also important to consider the optical properties of the telescope. Every optical component of the optical system emits radiation that can typically be expressed as a black body with varying emissivity. Since the FIR corresponds to peak emission from a black body between $10-100\,\mathrm{K}$, components need to be cooled

below these temperatures to obtain and optimal thermal background. Although cooling small components to $< 4\,\mathrm{K}$ is routine, this requirement poses a significant challenge for the primary collecting mirror of telescopes. With appropriate shielding, space based telescopes can be passively cooled to 50 K [22], but even at this temperature, emission from the mirror is several orders of magnitude above background levels [4]. Clearly, actively cooled telescopes are required to unlock the full potential of FIR astronomy.

Atmospheric Opacity

Although the presence of water lines in the FIR band is a significant benefit to FIR astronomy, they are also responsible for one of the primary limitations. Earth's atmosphere contains a significant concentration of water vapour which absorbs and scatters almost all the light in the FIR band (see Fig. 1.2). Under ideal observing conditions at high elevation ground based observatories, atmospheric transmission is effectively zero with the exception of small windows at 35 and 350 μm, at the ends of the FIR band, where transmission may increases to $\sim$50%. Significant improvements can be obtained using scientific ballooning, however, observing campaigns are constrained to a duration of days, and there are significant risks associated with recovering and redeploying these facilities. The most effective way of overcoming atmospheric absorption is with space based observatories.

The compelling science case of FIR astronomy, and the need for space based observatories, led to the first space based FIR telescope called the Infrared Astronomical Satellite (IRAS) in 1983 [23]. Notably, it was the first cryogenic scientific satellite actively cooled to 10 K, and although detector sensitivity was relative low, it is still described as one of the most successful space astronomy missions ever conducted [24]. IRAS mapped 96% of the sky at 12, 25, 60, and 100 μm photometric bands and discovered 350 k infrared sources. It observed dust and gas in starburst galaxies resulting in a significant contribution to the first measurements of the CIB [25]. It also discovered Vega-type debris disks thought to be the result of colliding bodies and similar to the Kuiper belt in our solar system [26]. The

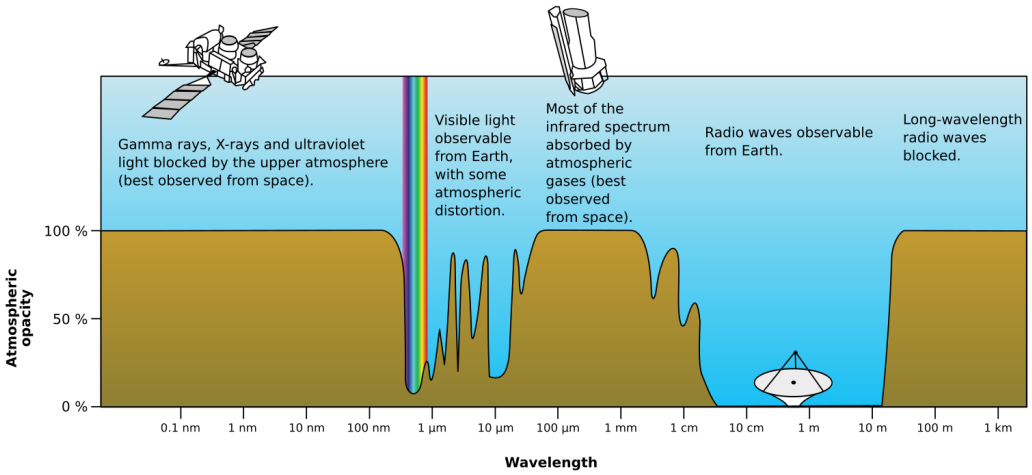


Figure 1.2: Atmospheric opacity at different wavelengths. Outside of a few narrow windows, the atmosphere is almost entirely opaque at FIR wavelengths, necessitating high altitude balloon observations or space based observatories. Figure obtained from NASA image gallery and is in the public domain.

success of IRAS motivated the development of ESAs Infrared Space Observatory (ISO) in 1995 [27], and later in 2003, the NASA led Spitzer space telescope [28]. Both of these telescopes were actively cooled to $\leq 6\,\mathrm{K}$ and benefited from improved sensitivity, detector arrays, and high resolution spectroscopy. The last space based FIR telescope was the ESA-led Herschel Space Telescope, launched in 2009 [29]. One of the most notable features of *Herschel* was its 3.5 m diameter primary mirror which significantly exceeded the previously largest 0.85 m mirror of Spitzer. This gave unprecedented spatial resolution, but due to its large size, was not actively cooled and observations were dominated by thermal emission from the $\sim 70\,\mathrm{K}$ mirror. In any case, these missions are all regarded as a tremendous success. It is impractical to list all their discoveries, but a few that are of particular interest include: detection of water lines in the ISM [30]; first observations of the galactic center [31]; detection of ultraluminous infrared galaxies and mapping the CIB [25]; the development of Mid-Infrared (MIR) diagnostics to distinguish AGN from starbursts [32]; first direct detection of light from extra solar planets [33], the detection of thousands of FIR sources including YSO at various stages of evolution [34]; and contributions to detection of some of the oldest galaxies in the universe [35]. Indeed, these observatories have exceeded

expectations and delivered unprecedented FIR science. However, their performance is drastically lacking in one key metric which will not be addressed by modern improvements in detector technology.

Spatial Resolution

One can recall the familiar experience where the apparent size of an object decreases as it moves further away. Equivalently, any two points on the object will appear to get closer together as the distance between the object and observer increases. As the distance continues to increase, eventually the two points will become indistinguishable. For an optical system, spatial resolution describes the minimum separation between two points that can still be distinguished. It is common to quantify this separation in terms of an angle, in which case it is called the angular resolution.

In the best case scenario, the angular resolution of an optical system will be limited by diffraction effects. If a point source is imaged by a diffraction limited system, radiant energy will always be spread over some finite region on the image plane. This intensity distribution is a consequence of the finite size of the optics which results in incomplete destructive interference in the neighbourhood of the nominal image point. This intensity distribution is called the Point Spread Function (PSF), and it depends significantly on the geometry of the optics. Using the Rayleigh criterion, two points are considered resolved when the central maximum of one PSF overlaps with the first minimum of the other [8]. With this definition, the angular resolution associated with a circular aperture is

$$\Delta\theta_{circ} \approx 1.22\frac{\lambda}{D}\,, \tag{1.1}$$

where λ is the wavelength of light being observed, and D is the diameter of the aperture. The PSF of a circular aperture is shown in Fig. 1.3. From Eq. 1.1, it is clear that resolution scales with wavelength. Since a small angular separation is desirable, somewhat confusingly referred to as high resolution, longer wavelengths are at a disadvantage compared

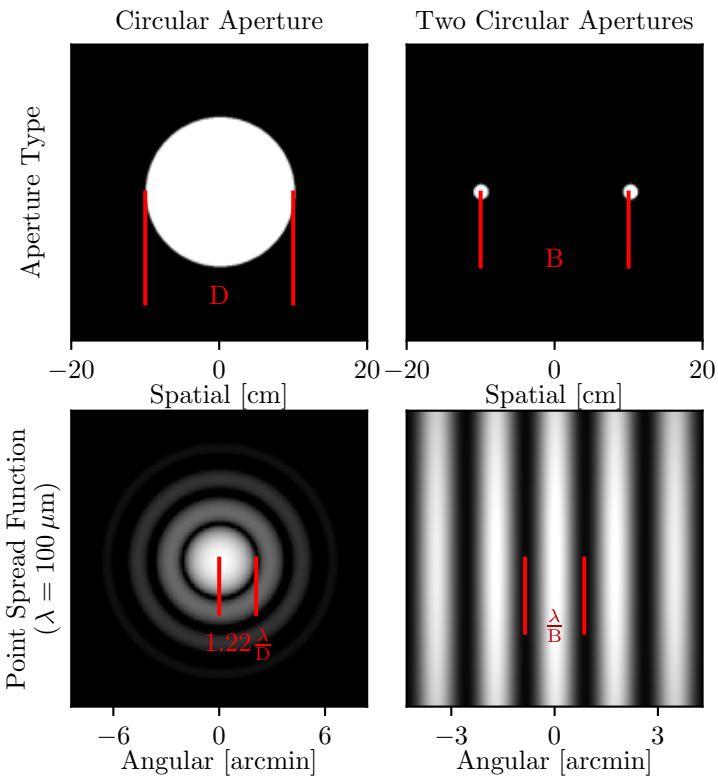


Figure 1.3: The points spread functions associated with different aperture geometries when viewing light with wavelength λ. The top row shows a circular aperture with diameter D (left), and two small circular apertures separated by a distance B (right). The bottom row shows the associated PSFs and includes labels for their diffraction limited angular resolutions.

to shorter wavelengths. For example, *Herschel*, with its 3.5 m primary mirror, the largest single mirror launched into space to date, achieves an angular resolution of 7.2 arcsec at 100 μm. This represents the same angular resolution that can be obtained with a 2.1 cm diameter telescope at visible wavelengths (e.g., 600 nm). Going in the other direction, *Hubble*, with its 2.4 m diameter primary mirror achieves a 0.06 arcsec at 600 nm. A 400 m diameter mirror would be required to obtain the same angular resolution at 100 μm, and a mirror over a kilometre in size would be required for the full FIR band, and it would need to be in space! Under these constraints, it is clear why angular resolution in the FIR lags behind other areas of astronomy resulting in the well-known "FIR gap" (see Fig. 1.4).

Low angular resolution leads to a situation called source confusions. Due to diffraction

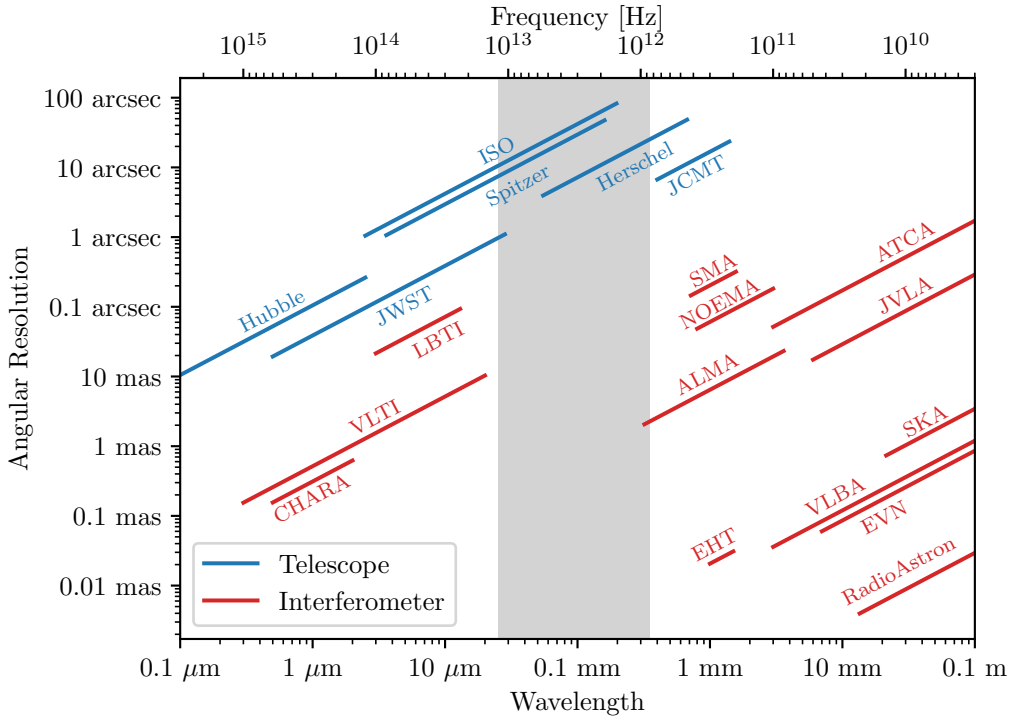


Figure 1.4: Angular resolution of various telescopes. With the exception of the FIR band (gray shaded region), all other relevant spectral bands have obtained sub-arcsecond angular resolution. Blue lines denote telescopes with a single aperture, while red lines denote spatial interferometer facilities. Image adapted from [36].

effects and the finite size of detectors, each detector in an optical system is sensitive to a region of the sky. The extent of this region is typically characterized by the Full Width at Half Maximum (FWHM) of the detector PSF and is called the "beam." Light originating from sources within the beam is integrated on the detector. Depending on the situation, light from sources that are not the target of interest is incident on the detector and can be interpreted as a kind of noise that pollutes measurements. It becomes difficult to determine if the signal you are measuring is a result of the target of interest or some other source in the beam. In addition to confusion, low spatial resolution simply prevents detailed mapping of the distribution of sources on the sky. Fig. 1.5 shows the effect of poor spatial resolution.

The most promising solution to the spatial resolution challenges of FIR astronomy is a spatial interferometer. Inspecting Fig. 1.4, it is clear that interferometers dominate in terms of spatial resolution and sub-arcsecond resolution has been achieved on both sides of

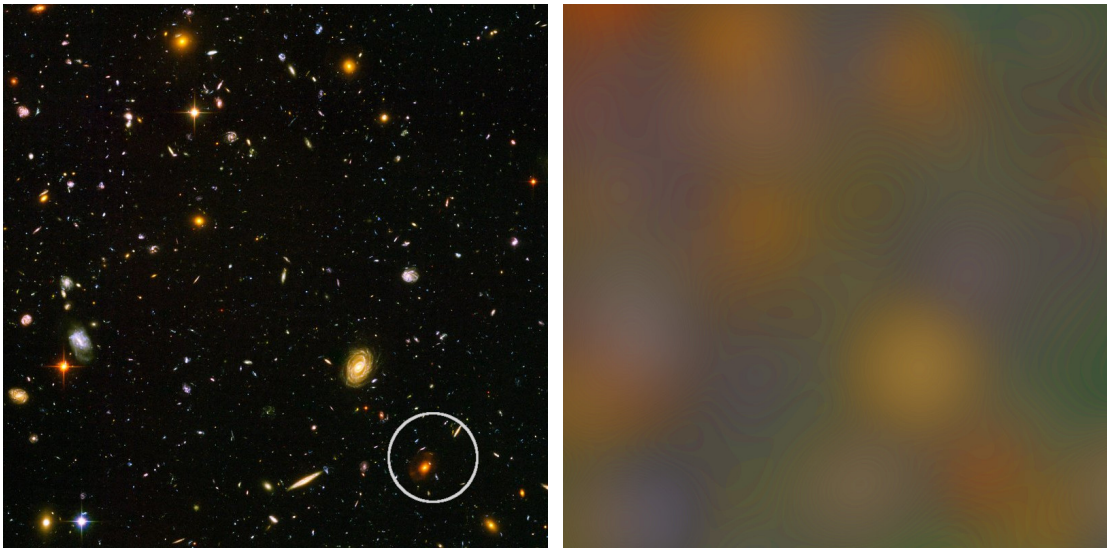


Figure 1.5: Effects of spatial resolution on image quality. The left panel shows a section of the *Hubble* ultra deep field (image adapted from NASA, ESA, and S. Beckwith (STScI) and the HUDF Team), with a circle representing the *Herschel* beam at 100 μm. Several individual galaxies reside within the *Herschel* beam and it becomes impossible to distinguish them when observed with the spatial resolution of *Herschel*. The left panel shows the same image but convolved with a synthetic *Herschel* beam and approximates what *Herschel* would see looking at the same patch of sky.

the FIR gap. The simplest form of a spatial interferometer consists of two input apertures separated by a displacement vector called the baseline. Each aperture samples the electric field at different regions of space which is relayed to a central combiner. Typically there is some mechanism that can vary the optical path difference between the two beams before reaching the combiner producing an interference pattern as a function of this optical path difference. This interference pattern gives information about the spatial distribution of the source intensity, and by measuring interference patterns at multiple baselines, an image of the source can be reconstructed using a technique called aperture synthesis. Spatial interferometry and aperture synthesis are discussed extensively in this thesis. In particular Ap. A includes a derivation and discussion of the van Cittert-Zernike theorem which provides the theoretical foundation of aperture synthesis. In Ap. B, the instrument response of the interferometer used in this work is derived, and demonstrates how the response is related to source characteristics.

Using a simple interferometer, the interference pattern obtained when viewing a point source, the PSF, is a sinusoid multiplied by the PSF of one of the assumed identical input apertures. Using the Rayleigh criterion, the diffraction limited angular resolution of a spatial interferometer is,

$$\Delta\theta_{intf} = \frac{\lambda}{2B}\,, \tag{1.2}$$

where B is the baseline length between the two input apertures, and λ is the wavelength of light being observed. In effect, the spatial resolution that can be obtained with a very large single mirror telescope with diameter D can also be obtained with a spatial interferometer consisting of two small apertures with diameters $d \ll D$, separated by a baseline $B \approx D$. This approach is a significantly more manageable engineering challenge than getting a 400 m telescope in space. Additionally, it is much easier to cool two small telescopes than a single large one. Once the sub-apertures are in space, the obtainable spatial resolution is only limited by the precision with which the baseline can be maintained, and the amount of observing time available. In principle, spatial resolution can be increased arbitrarily with a space based interferometer. For high quality interferometry, the baseline and any other distortions which affect optical path, must be maintained to within $\lambda/10$ of its nominal value. In this case the longer FIR wavelengths provide an advantage by somewhat relaxing instrument tolerances.

1.2 Science Cases for High Spatial Resolution

The apparent angular size of a distant object can be calculated using,

$$\theta \approx \frac{h}{L}\,, \tag{1.3}$$

where h is the linear size of the object, and L is the distance between object and observer. Due to the vast expanse of space, the distances to most astronomical sources of interest are large and the angular size of these sources are small. To unambiguously distinguish

these sources from their surrounding environments, an imaging system requires an angular resolution comparable to the angular size of the sources.

1.2.1 Stars and Planets

Stars form in cold dense molecular clouds permeated by dust. Density variations within these clouds serve as the nucleation site for gravitational collapse forming prestellar cores which radiate as black bodies with varying emissivities and peak emission in the FIR and sub-mm. As more matter accumulates on these cores they progress through various protostar classes. From class 0 to II, emission is dominated by the surrounding gas and dust in the FIR (see Fig. 1.6). It is not until the late class III phase when the star has consolidated or otherwise blown away its surrounding cloud that the star becomes unobstructed. While obscured, radiation with wavelengths comparable to or smaller than the size of dust grains $\lesssim 1\,\mu$m is absorbed by the grains and re-emitted thermally at low temperatures $\lesssim 25\,$K [37]. The study of prestellar cores and protostars is thus limited to continuum and line emission in the FIR. The characteristic scale of these structures is shown in Fig. 1.6 and decreases from $\sim 1,000\,$AU in the earliest stages to $\sim 10\,$AU in the later stages. In the context of planet formation, 1 AU is taken as the scale of interest.

The nearest star forming regions are Taurus [39] and the Rho Ophiuchi molecular complex [40], both about 140 pc away. Orion nebula is also particularly interesting for its high star formation activity and wide range of conditions making it one of the most studied star forming regions [41]. These star forming complexes are 400 pc away. At a distance of 100 pc, an angular resolution of 1 arcsec is required to resolve structures with a 100 AU scale, which is about an order of magnitude lower than what *Herschel* could achieve. As such, protostars remain unresolved in the FIR, even within the Milky Way.

In the case of low mass stars, formation and evolution is reasonably well understood [42]. However, there is still significant uncertainty relating to the formation of high mass stars. These stars are very important to the evolution of galaxies and chemical enrichment

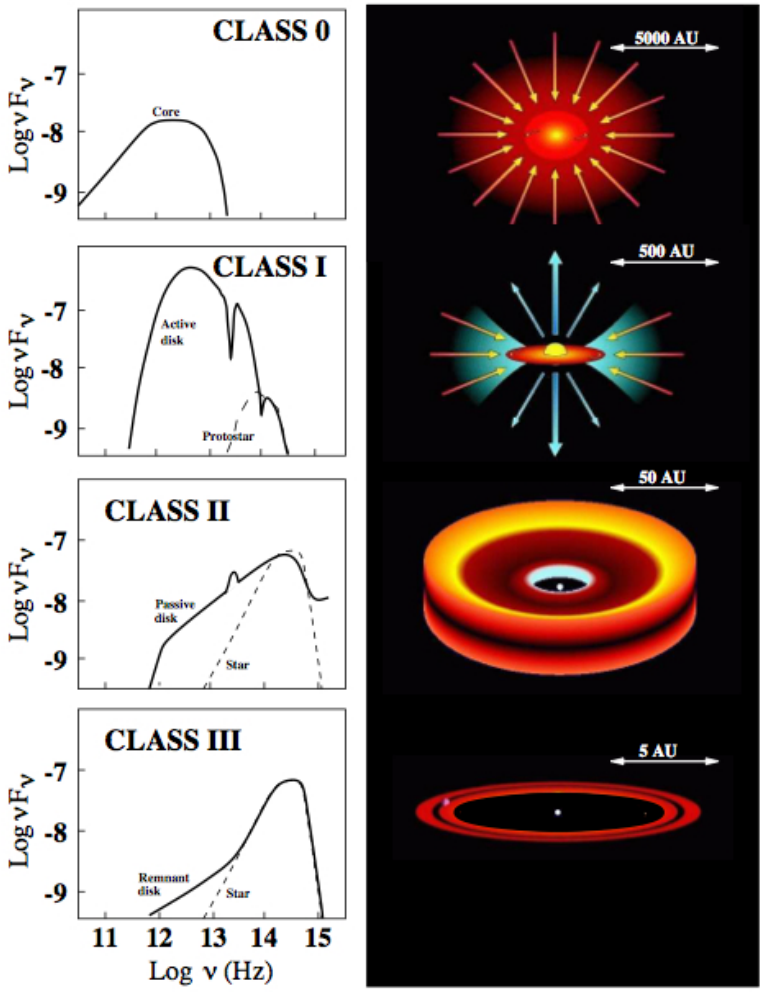


Figure 1.6: Formation and evolution of stars. Each row shows the different protostellar classes as stars progress though their development. The left column shows the SED associated with each class, and the right column shows morphology and scale. Young protostars are large and cold, emitting primarily in the FIR and sub-mm. As the stars mature, gravity pulls them into a smaller volume which heats up the surrounding gas and dust, moving the peak of the SED into the FIR. Conservation of angular momentum additionally results in the collapse of the cloud into a disk-like structure. It is not until the later stages, when stars are nearly main sequence, that most of the obscuring gas and dust is blown away by stellar winds. Even then, the remaining disk in which planets form emits primarily in the FIR. Figure obtained from [38], which was adapted from Andrea Isella's thesis (2006).

of the ISM. They produce the bulk of heavy elements which are critical to physical and chemical processes in the ISM, and their ionizing radiation, stellar winds, and supernovae have a strong feedback effect on the surrounding ISM, potentially triggering or quenching star formation. It is not known why massive stars are only born in clusters, and why mainly in the center of these clusters [43]. The two competing models are accretion of the surrounding gas vs single cloud collapse. If high mass stars form through accretion, why doesn't the intense radiation field and stellar wind quench accretion? What is the role of magnetic fields? If single cloud collapse, why don't these clouds fragment into smaller clouds? How important are stellar mergers in this process? To answer these questions, early stages of star formation must be observed on a scale of a few 100 pc. Since high mass star formation is rare, observations beyond the nearest star formation regions are needed to build up a relevant statistical sample. With an angular resolution of 0.1 arcsec, the relevant scales can be probed out to a distance of ~ 10 kpc. Furthermore, the best tracers for high mass star formation are the cooling lines [OIII] 52 μm associated with the highly ionized region around the star, [CII] 158 μm and [OI] 63 μm for the cooler photo dissociation region, and rotational emission from CO, H_2O, and OH to probed warm molecular clouds in the earliest stages of formation.

As a star's development progresses, a protoplanetary disk is formed. This disk is a dynamic and complicated structure with physical and chemical process that determine the conditions under which planets are formed. Fig. 1.7 shows a diagram that depicts the physical and chemical structure of a protoplanetary disk. The highest temperatures are obtained near the central star and decrease towards the periphery. The outer surface of the disk may be heated and ionized by uv radiation from other stars. Different phases of the gas are observed when progressing from the surface of the disk to the midplane, starting with low ionized gas, proceeding though neutral then molecular gas, and finally ices. Dust sublimates near the inner edge of the disk, but in the colder outer disk, dust grains can collide and grow to larger bodies. Dust grains can also serve as a substrate for chemical reac-

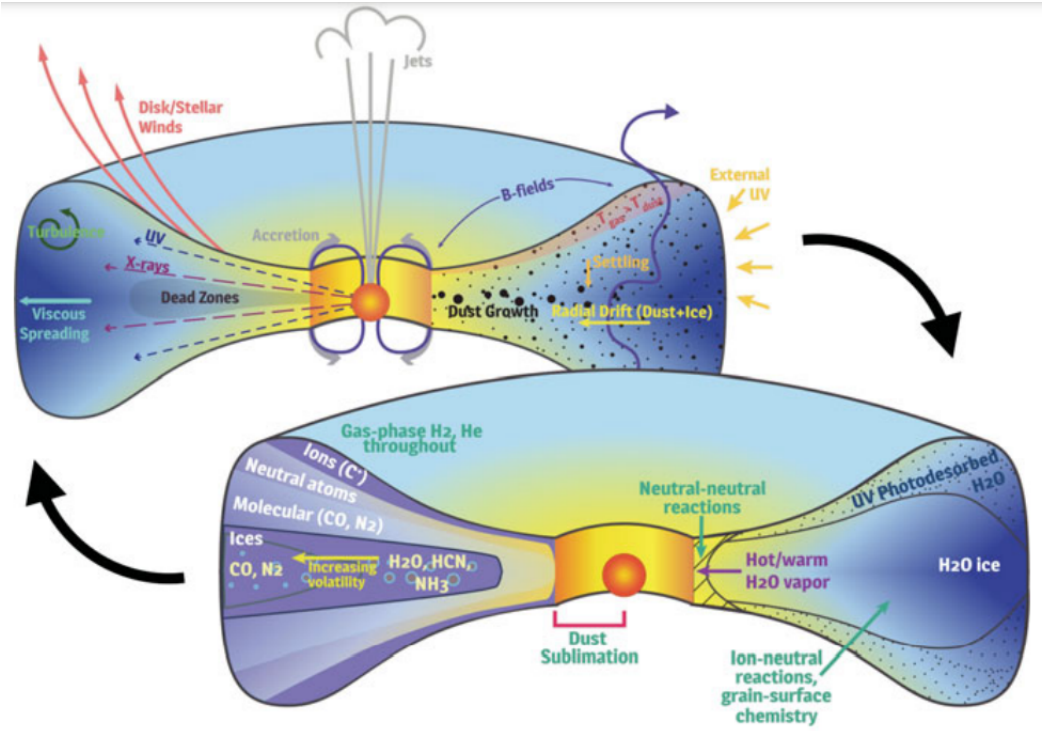


Figure 1.7: Physical and chemical structure of protoplanetary disks. The top figure shows the dynamic processes which regulate the physical environment of the protoplanetary disk. The physical conditions strongly influence the chemical structure of the disk, shown in the bottom figure. Chemical structure affects the cooling mechanism in the disk which then influences the temperature and physical structure as part of a feedback cycle. FIR observations can probe the full range of conditions within protoplanetary disks. Resolving these structures is critical to understanding how and where planets form, and what conditions determine the composition of those planets. Figure obtained from [44]; credit: Cleeves, International Astronomical Union.

tions facilitating the production of molecules, with more volatile molecules supported in the colder environments of the disk.

The FIR is unique in its ability to probe all the phases of protoplanetary disks. Ionized fine structure lines trace the hot inner region and outer surface, neutral fine structure lines give access to the neutral components, and molecular rotation lines probe cooler regions and of course provide direct measurements of the disk chemical structure. Of particular note is water ice continuum emission that is only observed in the FIR band. A model spectrum showing the important emission features of protoplanetary disks is shown in Fig. 1.8. The Near-Infrared (NIR) portion of the spectrum is observed by the James Webb Space

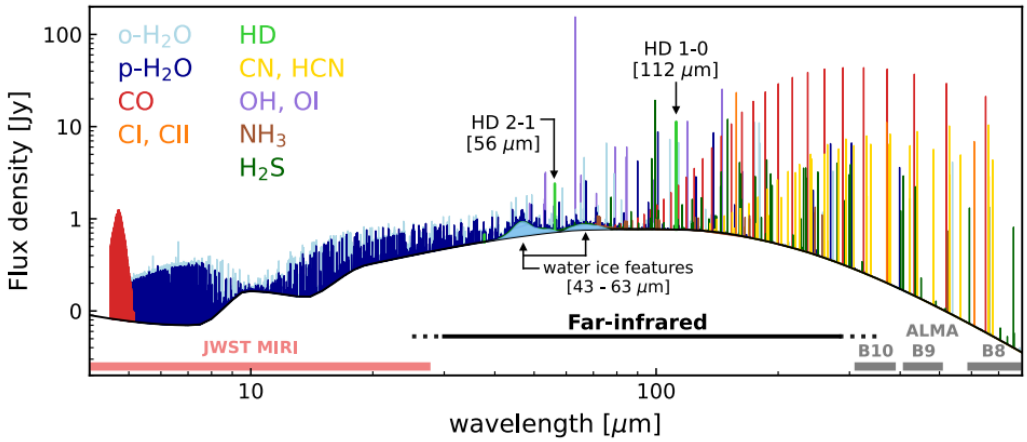


Figure 1.8: Model spectrum of a protoplanetary disk. FIR observations have access to a broad range of molecular and fine structure emission lines that probe the full range of conditions within the disk. Of particular note is the abundance of water vapour lines and water ice continuum features that allow mapping of the distribution and phase of water within disks. JWST observations probe the warm outer surface of the disk while ALMA probes the quiescent midplane. FIR can provide complementary observations and is the only band capable of observing the bulk of the disk. Figure obtained with permission from [46]. Credit: L. Trapman.

Telescope (JWST) [22], which maps the energetic outer surface of the disk. The Atacama Large Millimeter/submillimeter Array (ALMA) [45] is sensitive to the sub-mm part of the spectrum, which maps the cold quiescent parts of the midplane where molecules may be inactive and frozen onto dust grains. FIR observations bridge the gap between these two existing facilities and provide access to the majority of the disk while probing the full range of physical conditions.

Observations of water, both in it gaseous and solid forms are of particular importance. In the coldest environments, water frozen onto grains can be observed from their $43-63\,\mu$m continuum emission features. Rotational emission from water vapour can be used to trace thermal environments in the range $\sim 100-300\,$K where water is sublimated from grains. Higher energy transitions can also be used to trace shock regions with temperatures $>300\,$K [9]. As such, water traces the full range of conditions in the disk. Spatially resolving the different components is critical for understanding where and under what conditions planets form. Beyond the snow line, where water is frozen onto grains, the ice deposited

on the grain can help grains stick together to form larger bodies which are the precursors to planet formation. Resolving this region gives information about where and how efficiently this process occurs [47]. Resolving the distribution of water gives insight into what kinds of planets form and where they form, such as rocky planets like Earth close to the star, compared to gas giants and icy bodies in the outer disk. Such observations can also help answer the question of how water is deposited on rocky planets. Is water locked into porous grains, or is it delivered by comets? The ability for FIR observations to provide such rich insight about the phase and distribution of water make it one of the most powerful tools for answering the questions of where and how life sustaining planets form.

In the early stages of planet formation, the growing bodies are still embedded in the protoplanetary disk. The signatures of these planets can be inferred from the effects they have on the surrounding gas. They may created vortices or density variations that can be observed through continuum dust emission. Turbulence resulting from these vortices injects energy into the gas which can be observed using molecular lines from species such as CO and H_2O. These structures are expected to be observed with a characteristic scale of ~ 1 AU, which for the nearest star forming regions would require an angular resolution of 0.01 arcsec.

1.2.2 Galaxies and Cosmology

Cosmic distances are often quantified by "redshift" which relates the observed wavelength of light λ_{obs} to emitted wavelength λ_{emit} [48],

$$z = \frac{\lambda_{obs} - \lambda_{emit}}{\lambda_{emit}} \,. \tag{1.4}$$

The difference in wavelength is due to the observed expansion of the universe; the longer light stays in space, the more it gets stretched out. Since light travels at a finite speed that is constant in all inertial reference frames, the time in space is proportional the distance travelled. Note that since the expansion rate is not constant in time, redshift is not directly

proportional to distance.

Some important cosmological landmarks are the surface of last scattering $z \approx 1090$ when electrons combined with hydrogen nuclei [49, 50], and for the first time, light could travel though the universe without scattering off free charge carriers. At this time, baryonic matter in the universe had a nuclear abundance of 24% helium and 76% hydrogen [50]. Using these elements, the first stars are predicted to have formed at z = 20-30 with a mass $\gtrsim 100 M_\odot$ [51]. Fusion within these stars and their explosive deaths began to enrich the universe with heavier elements. Stars became more ubiquitous, and at $z \approx 10$ the energetic light of massive stars had re-ionized the rarefied intergalactic medium [49, 50]. The peak star formation rate is observed at $z \approx 2$ and has been decreasing ever since [13].

The formation and evolution of primordial stars is a subject in astronomy with considerable uncertainty. In the early universe devoid of heavier elements, fine structure lines associated with these elements, which provide the mechanism for efficient cooling of the ISM, were absent. As such, the mass of stars must be immense to overcome the thermo-kinetic pressure of the gas. Observations are required to constrain the mass and time scale associated with their formation. The main cooling mechanisms in such an environments are the H_2 quadrupole rotational emission lines at 12, 17, and 28 μm [51]. Although nominally emitted in the NIR, these lines are shifted into the FIR between $z \approx 2-20$. Resolving a molecular cloud with primordial star formation at $z \approx 20$ is challenging, both in terms of resolution and sensitivity. However, molecular clouds at $z \approx 10$ may have stars that form under similar conditions that can be resolved, and are expected to be bright enough to detect [3].

Primordial stars naturally leads to questions regarding the origin of dust. Significant quantities of dust have been measured in galaxies out to $z > 5$ [52], indicating that there must be a rapid mechanism for the production and distribution of dust. A plausible solution is the supernovae of massive stars, in particular type-II and pair-instability supernovae. These massive stars have short lifetimes and models predict that their energetic supernovae

could convert as much as 30% of the stellar mass into dust [53]. Traces of dust can be found with thermal emission from dust, and although some evidence has been found to support this model, the low resolution of current FIR measurements leads to doubts related to foreground confusion. To constrain these models, angular resolution needs to be comparable to resolutions of other telescopes at wavelengths which are known to originate from supernovae remnants. In this way, dust emission can be unambiguously correlated to the remnants.

The proliferation of dust and the chemical enrichment of the ISM provides efficient cooling pathways which could lead to rapid star formation. On its own, this does not explain why star formation peaked at $z \approx 2$ as significant quantities of dust were available before this time. It has been shown that the mass of Super Massive Black Holes (SMBH), thought to exist in the center of most galaxies, are directly related to the mass of the bulge in their host galaxies. This relation suggest that starbursts are correlated with accretion onto SMBH which produce AGN [54], and that this interplay strongly influences the evolution of galaxies. However, the extent of the galactic bulge is significantly larger than the Schwarzschild radius of the SMBH, suggesting co-evolution resulting from a common galactic gravitational potential. The question remains as to what triggers the active star formation, how is it regulated, and why does it stop? Galaxy mergers have been suggested as a trigger, and regulation is no doubt related to feedback processes between components of the AGN and host galaxy. It could even be the case that strong outflows from AGN push out much of the star formation material, thus quenching the rapid star formation phase [55].

A diagram showing the main components of the unified AGN model is presented in Fig. 1.9. The figure also shows the characteristic scale of some of the important components. Resolving these components is required for detailed analysis of the feedback processes in AGN. The first galaxies and subsequent mergers occur around $z \approx 5-10$. At this distance, an angular resolution of $0.02\,\text{arcsec}$ over an arcmin^2 Field of View (FOV) is required to distinguish between AGN excited narrow line region gas and circum-nuclear gas

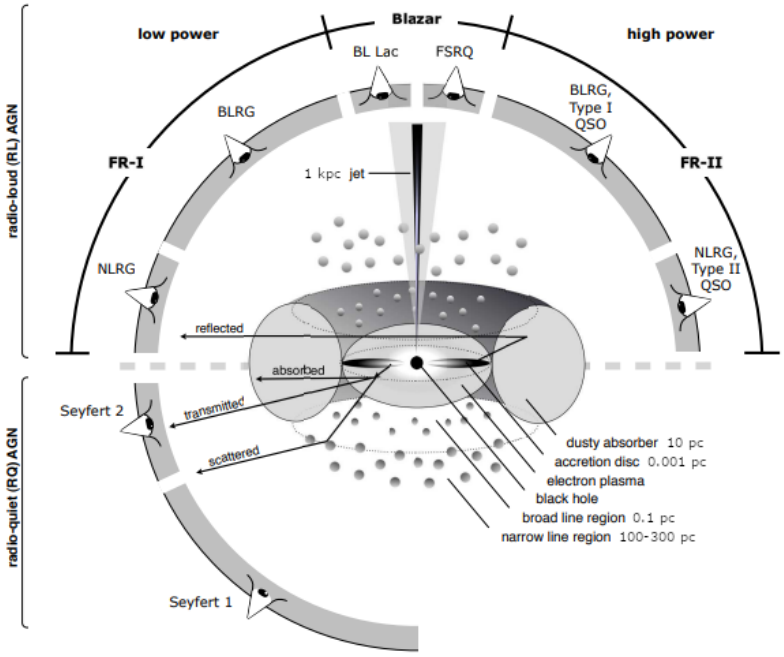


Figure 1.9: Structure and scale of AGN. AGN are often obscured by dust during their most active phase which correlates with starburst activity in the host galaxy, necessitating FIR observations. The various components can be observed in the FIR using the highly ionized fine structure lines. Resolving these structures would provide valuable information about feedback processes which might influence star formation in the host galaxy. Resolving the narrow line region in particular would be decisive in separating AGN excited gas from stellar radiation excited gas. Figure modified from [56].

excited by stars.

Evidence suggests that most galaxies progress through a dust obscured and FIR bright phase during their evolution. These FIR bright phases are signatures of starburst, are correlated to AGN, and contribute significantly to the CIB. Just like stars, the obscuring dust makes FIR the ideal band for studying the evolution of galaxies. AGN, being one of the brightest cosmic structures driven by accretion onto SMBH, express extremely energetic environments containing highly ionized gas. These ions produce fine structure lines within rest frame emission in the NIR and FIR that can be observed in the FIR at the redshifts of interest, $z \approx 5 - 10$. This statement applies equally well to fine structure lines resulting from lower energy ionization associated with massive star formation during the starburst phase (see Fig. 1.10). As such, FIR observations provide the means needed to distinguish

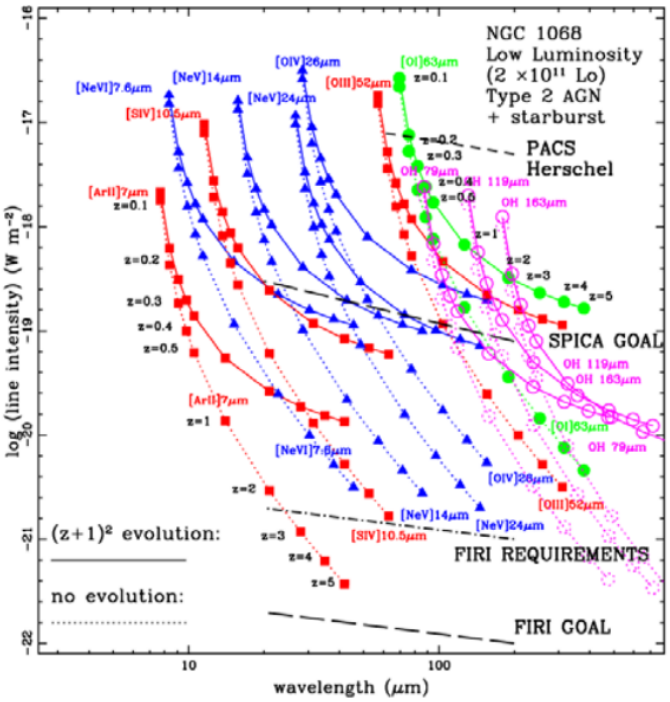


Figure 1.10: Important diagnostic lines for AGN and host galaxies. The predicted intensity and wavelengths of various spectral features are modelled for different redshifts assuming $(z+1)^2$ luminosity scaling (solid) or no scaling (dashed). Stellar radiation excited lines (red squares), AGN excited lines (blue triangles), photodissociation lines (green circles), and warm molecular gas lines (magenta circles) can all be detected at relevant redshfts withing the FIR. These lines are key tracers for separating AGN from starburst regions and studying their co-evolution. Figures obtained from [3] with permission.

between AGN and star formation components, and can track mergers and feedback at the appropriate epochs when these processes are most active.

With an angular resolution of 0.01 arcsec, the broad line region of AGN, associated with accretion and one of the smallest AGN structures, can be resolved out to ~ 16 Mpc. One of the largest structures is the narrow line regions, and emission from this region can be effectively used to trace AGN. With the proposed angular resolution, the narrow line regions are expected to be resolved at any redshift using the standard cosmological model [38]. FIR observations with sufficient angular resolution can thus trace the growth of SMBH through mergers obscured by dust, probing binaries separated by $0.1-100$ kpc out to $z \geq 10$. Studies of mergers can be used to infer the relative importance of mergers compared to accretion with respect SMBH growth. Additionally, distinguishing narrow line regions of AGN from

stellar radiation excited gas under the same angular resolution requirements can be used to study the co-evolution of AGN and starburst.

An orthogonal study of active star formation can be conducted by observing the local universe. In particular, the Magellanic clouds at 55 and 70 kpc have lower metalicity than solar abundance and demonstrate high levels of star formation activity. These clusters have conditions similar to those expected at $z \approx 2$ but are not subject to the effects of AGN discussed above. As such, the Magellanic clouds may provide insight as to why star formation peaked $z \approx 2$, and to what extent SMBH accretion is relevant. Due to their close proximity, star formation can be studied in greater detail with a higher resolution than in distant galaxies. For example, the $\sim 1,000\,\mathrm{AU}$ prestellar cores, representing the earliest stages of formation, can be resolved with sub-arcsecond angular resolution. Although the Magellanic clouds are subject to interaction with the gravitational potential of the Milky Way, they provide somewhat isolated test cases to study star formation in low metalicity environments.

1.3 Double Fourier Interferometry

In addition to meeting the spatial resolution requirements, spectroscopy is needed to observed the important spectral features. In most cases, the necessary spectral resolution is determined by the need to map gas dynamics. Gas velocity is revealed through the Doppler effect, where spectral lines are shifted in proportion to the radial velocity between source and observer. Typical requirements are $100\,\mathrm{km/s}$, which implies a resolving power of $\lambda/\Delta\lambda = 3,000$. Combining a spatial and spectral interferometer into a single instrument is thus a powerful tool for reaching FIR observation requirements. Although spatial and spectral interferometry have been studied separately to a high level of maturity, combining both techniques is still quite novel. The particular realization of a spatial-spectral interferometer used in this work will be called a DFI, as both spatial and spectral aspects of the instrument are directly related to Fourier analysis. In this section, the historical develop-

ment leading to this instrument is briefly reviewed.

Deciding when to start is difficult, but Thomas Young and Augustin-Jean Fresnel deserve some recognition due to their significant contribution to the wave theory of light, and they independently performed their own early interferometer experiments before 1820.

Albert A. Michelson arguably contributed most significantly to the DFI used in this work. In 1887, Michelson developed the "Michelson interferometer" that gained wide spread popularity as a result of its use in the famous Michelson–Morley experiment [57]. The instrument provided sensitive measurement of the optical path difference between two beams and is one of the simplest implementations of a Fourier Transform Spectrometer (FTS). The Mach–Zehnder interferometer was developed in 1891 [58], which used separated light paths and two input-output ports. Combining this design with an optical path delay mechanism similar to the Michelson interferometer results in a highly configurable instrument. The dual input and distinct optical paths make it amenable to spatial interferometry, while the dual outputs can be used to fully leverage the light input to the system.

Not long after developing his spectral interferometer, Michelson pioneered the field of stellar interferometry staring in 1890 when he used the wave nature of light to define, and further calculate, the visibility of interference fringes when viewing an extended source through two apertures [59]. He realized that by increasing the distance between the two apertures, the visibility of fringes would decrease to zero and that this could be used as a sensitive measure of the extent of a source or the separation between sources. The technique was first applied to Jupiter, but the work that was most iconic was when the technique was applied to measure the diameter of Betelgeuse in 1920 [60]. This was the first application of stellar interferometry and is associated with the famous "Michelson stellar interferometer," not to be confused with the Michelson interferometer. This experiment marked the beginning of spatial interferometry in astronomy.

The next major milestone was the discovery and further development of the van Cittert-Zernike theorem. In 1934, Pieter H. van Cittert introduced the "degree of consonance"

which described the correlation between the complex amplitudes of light at any two points in space [61]. By means of rather elaborate calculations, he succeeded in establishing a relationship between the source intensity and degree of consonance in the special case of a perfectly focused aberration-free optical system having a rectangular aperture. In 1938, Frits Zernike used a purely statistical approach to the topic and introduced the "degree of coherence" [62]. This result was a much more general and simple formulation which revealed the direct relationship between source intensity distribution and degree of coherence through a Fourier transformation. As a result of these efforts, what Michelson described as the visibility of fringes was generalized to a complex value.

The van Cittert-Zernike theorem was further developed by Harold H. Hopkins in 1951 who added an explicit description for homogeneous media [63]. This revised quantity was called the "phase-coherence factor" and was presented with an intuitive physical description in terms of the Young double slit experiment. Emil Wolf extended the formulation to include intensity distributions with arbitrary spectral width in 1955 by introducing a time delay between the two interferometer beams [64]. Wolf called the derived quantity the "mutual coherence function." With this addition, the foundational theory of DFI observation was complete. A derivation based on Wolf's work is presented in Ap. A. The normalized mutual coherence is called the "complex visibility," which is more commonly referenced in modern literature.

Spatial interferometry was first adopted by radio astronomers in the 1940s. This is unsurprising considering the resolution penalty imposed by diffraction limited optics in the radio band is orders of magnitude worse than visible or even FIR light. The longer wavelengths also relax the baseline precision required for interferometry, and the low atmospheric opacity allows for long baseline interferometers to be constructed on the ground. Another significant advantage, at least from a measurement perspective, is that radio frequencies are low enough that the electric field at each aperture can be recorded coherently and read out with a relatively low bandwidth. The correlation between fields at different

apertures can thus be directly obtained either digitally from data written to a storage device, or by transmitting and combining the signals through wires and analogue electronic devices. This provides a high level of flexibility compared to other wavelengths of light. Radiation with frequencies greater than tens of GHz must have their electric fields combined directly, and the correlation inferred from intensity measurements. However, photonic circuits, heterodyne systems, and fast readouts have helped to achieve more flexibility at higher frequencies. Perhaps the main drawback with radio interferometry is that coherent detection is typically limited to a narrow bandwidth and observations of distantly spaced frequencies require separate channels, which makes spectroscopy and the production of spectral images difficult.

The first demonstration of radio interferometry was in 1947 and detected sun spots using a sea-cliff interferometer [65]. This was a single aperture receiver that used reflection off sea water to produce a synthetic second aperture. The interferometer baseline was not configurable with such a system, which limited spatial resolution and what wavelengths of light could be used productively. The first double aperture radio interferometer was developed in 1948 [66], after which, multi-aperture systems became the norm. Almost all of the modern aperture synthesis algorithms and image deconvolution techniques grew out of radio astronomy due to its early monopoly on interferometry.

It wasn't until 1974 that interferometry at 10μm in the infrared was demonstrated [67]. The system employed heterodyning and coherent detection with limited bandwidth. In 1996, interferometry was demonstrated at visible wavelengths [68]. In this case, direct detection (as opposed to coherent detection) was used, which recorded visibilities and closure phase, a conserved quantity requiring three apertures that can be used to determine the phase of the complex visibility. The late development of visible light interferometry was largely due to the need for precise laser metrology and fast readout.

Up until the mid 1980s, interferometry and spectroscopy remained largely separate disciplines. One of the first demonstrations that combined the two was Itoh et al. in 1986 using

visible light [69]. The instrument, called a "volume interferometer," used a design similar to the Michelson interferometer with mirrors replaced by orthogonal wavefront folding prisms. At the output was a rectangular detectors array which recorded a two-dimensional interference pattern. The wavefront folding produced a situation where each point on the detector array effectively corresponded to a different baseline vector for a two-aperture system despite the instrument only having a single aperture. Like the Michelson interferometer, one of the wavefront folding prisms could be moved in order to affect and optical path difference between the two arms of the interferometer. In this way, the instrument probed both spatial and temporal coherence. However, spatial resolution was limited by the detector array and input aperture geometry, providing no spatial resolution enhancements over a single aperture system.

Soon after, Mariotti et al. independently developed their own NIR spatial-spectral interferometer in 1988 [70]. They employed separate apertures with beam combination in the pupil plane, like the Michelson stellar interferometer, which is amenable to high spatial resolution observations. The spectrometer side used a Mach-Zehnder-like design where the optical path length of one arm could be controlled independently of the other using a movable mirror. Modern designs closely resemble their design in most respects. Measurements where taken with the 4 m telescope at Kitt Peak National Observatory with additional measurements at Haute-Provence Observatory. They successfully extracted visibilities from interference fringes for resolved and unresolved stars.

In recent years, a number of groups have emerged that developed testbed systems to study spatial-spectral interferometry, specifically for the purpose of FIR astronomy. The earliest was the Leisawitz et al. group, who developed the Wide-field imaging interferometry testbed in 2003 [71, 72, 73]. This Mach-Zehnder-like system used visible light and could achieve relatively large effective baselines, and benefited from sensitive mature visible light detector technology at the cost of tighter wavefront error and alignment constraints. The camera was a multi-pixel array, which is what allows for wide-field imaging, and the source

was a highly configurable projector that could produce arbitrary source geometry with arbitrary spectral content. Source rotation, required for two dimensional reconstruction of the source, could thus be accomplished though software to high precision without manual intervention that might disrupt alignment. This testbed had the potential to simultaneously demonstrate all relevant aspects of a spatial-spectral interferometer, with the exception of measurements within the FIR band, but was unfortunately decommission. The most recent publication from the group shows interferograms and visibility measurements on a multipixel detector, but no complex visibilities or demonstration of aperture synthesis [74].

The Ohta et al. group used a Martin-Puplett type spectrometer in 2006 [75]. The instrument used direct detector bolometers sensitive to sub-mm and mm radiation, and could be configured to measure polarization. They successfully demonstrated complex visibility measurements and aperture synthesis using a single detector observing spectrally uniform sources with isosceles trapezoid shapes [76]. Their system was also used for solar measurements [77].

Finally, the Grainger et al. group developed a system sensitive to FIR radiation in 2012 [78]. They used a Mach-Zehnder-like design with a single direct detection bolometer. Interferograms were recorded using the step-and-integrated method where signal at a particular optical path difference was obtained by modulating the source and processing the detector signal using a lock-in amplifier. The group measured visibilities associated with sources consisting of narrow slits. They also introduced spectral variation into the source scene by masking one of two slits with a low-pass filter. This was the first demonstration of the DFI technique in the FIR, and the first to use sources with spectral variation. However, complex visibility measurements and aperture synthesis was not demonstrated. Their results were also confined to one-dimensional targets and a single detector.

Efforts have also been made to extend spatial-spectral interferometry beyond the lab using balloon-borne experiments. The two separate instances used instruments called Far-Infrared Interferometer Telescope Experiment (FITE) 2010 [79, 80] and The Balloon Ex-

perimental Twin Telescope for Infrared Interferometry (BETTII) 2014 [81]. The designs of both systems include an 8 m baseline. Although FITE appears to still be under active development, BETTII was irreparably damaged during its first flight and is no longer in development [82]. No science results were obtained, but the test did provide insight on the challenges and requirements of ballooning interferometry.

Considerable interest in a space based FIR interferometer has been growing since the 2000 decadal report "Astronomy and Astrophysics in the New Millennium" strongly recommended the development of space based infrared observing capabilities [1]. For the FIR, the Submillimeter Probe of the Evolution of Cosmic Structure (SPECS) was studied as a NASA Cosmic Vision mission. This was an ambitious project with a proposed 1 km baseline, 4 m sub-apertures cooled to 4 K, capable of producing spectral images with a resolving power of 10^4 over a 1 arcmin^2 field of view [83]. In 2003, the Community Plan For Far-IR/Submillimeter Space Astronomy recommended Space Infrared Interferometric Telescope (SPIRIT) as a pathfinder mission for SPECS [84, 2]. While in 2004, the ESA astronomy working group recommended their own interferometer study, which lead to the Far-InfraRed Interferometer (FIRI) mission concept study in 2009 [3], with contributing members from Europe, the USA, and Canada. Shortly after, FIRI was submitted as a proposal for the ESA Cosmic Vision 2015–2025 [85]. FIR interferometer recommendations have also been submitted to the 2010 NASA decadal survey [86], the ESA call for science themes for L2 and L3 missions (2013) [87], and the NASA long term plane (2021) [9].

1.4 Thesis Outline

The work presented in this thesis is expected to further enrich spatial-spectral interferometer development. A successful FIR space interferometer must be capable of generating wide-field spectral images with high spatial resolution. These requirements can be factored into independent tasks:

1. Demonstrate complex visibility extraction and aperture synthesis for a single detec-

tor. This item proves that integrating a spectrometer into the instrument does not compromise the imaging capabilities of a spatial interferometer.

   (a) If the relevant source variation is along one axis, this task can be performed with a single baseline orientation producing a one-dimensional image.

   (b) Generalize the procedure for a two-dimensional source requiring multiple baseline orientations.

2. Demonstrate that these images have spatial resolution better than can be obtained with the sub-apertures individually, and that image properties are consistent with theory. This item proves the spatial resolution enhancement capabilities of the DFI technique.

3. Demonstrate sensitivity to spectral variation in the source. This is the key advantage over traditional interferometers that allows the production of spectral images.

4. Demonstrate wide-field capabilities. A wide field of view significantly reduces the time required to observe extended sources and is of tremendous practical importance.

Fulfilling the items above is the main goal of the work presented in this thesis. Furthermore, this thesis is intended to provide a comprehensive resource for individuals interested in the DFI technique. To this end, Ch. 2 develops the underlying theory of Fourier analysis and its application to spectral and spatial interferometry. This chapter provides the tools needed to understand how a DFI works, and how to interpret and analyze the system response. Ch. 3 discusses the interferometer system design, focusing on the optical and mechanical characteristics. This chapter provides insight into the limitations and expected performance of the imaging system. Ch. 4 focuses specifically on the detector system, an unconventional Proportional-Integral–Derivative (PID) controlled TES bolometer array. The detector system is a critical component of the instrument that affects the spectral sensitivity and understanding its parameter space can help improve performance. Ch. 5 presents a variety of DFI observations and includes detailed discussions on data acquisition, calibration, complex visibility extraction, and aperture synthesis. These observations provide

insight into how source geometry affects complex visibility and directly addresses the items enumerated above. Concluding remarks and recommendations for future research are given in Ch. 7.

My work was conducted as part of a small research team. I joined the project early in its conception and I have played a primary role in virtually all aspects of the instrument development. With the exception of the specific items listed in the acknowledgments section, the work presented in this thesis is entirely my own. This thesis places little emphasis on the ancillary work that was needed to bring the project to fruition. In particular, the software libraries used for: data acquisition, data analysis, hardware control, simulation, and laboratory management are extensive and were critical to the success of the project, but are not presented. Additionally, this work does not focus on the commissioning phase which iterated over a number of proposed experimental designs, detector system data structures, and software versions. Although a significant amount of content is presented in this thesis, it does not capture the full scope of my contributions.

Aspects of this work have been presented at conferences including the Canadian Astronomical Society [88], Optica [89, 90], the International Symposium on Space Terahertz Technology [91], and the Society of Photo-Optical Instrumentation Engineers [92]. Some of the work relating to instrument design and characterization has been published in the Institute of Electrical and Electronics Engineers Transactions on Terahertz Science and Technology journal [93]. A future publication is expected for the experiments on complex visibility extraction and aperture synthesis discussed in Ch. 6.

Chapter 2

Fourier Analysis

Fourier analysis is the branch of mathematics concerned with representing general functions as linear combinations of orthonormal trigonometric functions. The intellectual heritage of such techniques date back to antiquity, however, the modern formalism is attributed to Jean-Baptiste Joseph Fourier who used, what are now called, Fourier series to model heat flow analytically [94]. Fourier series prove to be a powerful tool in mathematics do to the generality of functions that can be synthesized and the ease with which these representations can be manipulated. In physics, and optics in particular, representations of functions as a superposition of waves conforms precisely to the classical representation of polychromatic light. The Fourier series thus provides a model that expresses physical reality and not simply a convenient mathematical abstraction. However, the utility of a Fourier series reaches its full potential with the introduction of the Fourier transform, an operation which decomposes the general function into its frequency components. The implications for spectroscopy are evident, but more broadly, the associated theory provides the foundation for analogue and digital signal analysis. As such, adequate knowledge of Fourier analysis is essential for understanding and interpreting the output of the interferometer system used in this work.

This chapter is dedicated to providing an introduction to Fourier analysis and connecting it to the response generated by the Fourier Transform Spectrometer (FTS) and Double-Fourier Interferometer (DFI) used in this work. The Fourier series is developed in Sec. 2.1 followed by a derivation of the Fourier transform and its properties in Sec. 2.2. The dis-

cretization of signals and the resulting implications with respect to the Fourier transform are presented in Sec. 2.3. Connection to, and techniques associated with FTS are discussed in Sec. 2.4, with a similar discussion with respect to DFI in Sec. 2.5.

2.1 Fourier Series

Any function can be synthesized from the sum of an even and an odd function

$$f(x) = f_e(x) + f_o(x) \,, \tag{2.1}$$

where the even and odd functions can be found using,

$$f_e(x) = f_e(-x) = \frac{f(x) + f(-x)}{2} \,, \tag{2.2}$$

and

$$f_o(x) = -f_o(-x) = \frac{f(x) - f(-x)}{2} \,. \tag{2.3}$$

Functions that are periodic with period L can be represented as a linear combination of sinusoidal functions with frequencies consisting of the fundamental spatial frequency $k = 1/L$, and higher harmonics, $k_n = n/L$. Such a construction is formally a Fourier series [95],

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos(2\pi k_n x) + b_n \sin(2\pi k_n x) \,, \tag{2.4}$$

where a_0, a_n, and b_n are constants known as Fourier coefficients. The sum of weighted cosines and sine functions can be identified as the even and odd components of Eq. 2.1, respectively. Note the constant term originates from a zero-frequency cosine.

An equivalent representation of a Fourier series can be constructed using the amplitude and phase formalism,

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} d_n \cos(2\pi k_n x - \phi_n) \,, \tag{2.5}$$

where the amplitude is,

$$d_n = \sqrt{a_n^2 + b_n^2}\,, \tag{2.6}$$

and the phase is,

$$\phi_n = \arctan\left(\frac{b_n}{a_n}\right)\,. \tag{2.7}$$

This formalism is particularly useful when interpreting interferometric data, as discussed in § 2.4 and 2.5.

To construct a Fourier series, the function must satisfy the Dirichlet conditions [95]:

1. The integral of $|f(x)|$ over one period must converge.

2. The function must be single-valued and continuous, except at a finite number of finite discontinuities over one period.

3. The function must have a finite number of maxima and minima over one period.

If these condition are met, the Fourier series converges to $f(x)$ at all points where $f(x)$ is continuous. At points of discontinuity, the Fourier series converges to the average of the values the function approaches from either side of the discontinuity.

The coefficients of Eq. 2.4 are obtained using,

$$a_n = \frac{2}{L}\int_{-L/2}^{L/2} f(x)\cos(2\pi k_n x)\,dx\,, \tag{2.8}$$

$$b_n = \frac{2}{L}\int_{-L/2}^{L/2} f(x)\sin(2\pi k_n x)\,dx\,, \tag{2.9}$$

which can verified by inserting Eq. 2.4 into the above equations and using the orthogonality of trigonometric functions,

$$\int_{x_0}^{x_0+L} \cos\left(2\pi k_m x\right)\sin\left(2\pi k_n x\right)dx = 0 \qquad \text{for all } m \text{ and } n, \tag{2.10}$$

$$\int_{x_0}^{x_0+L} \cos(2\pi k_m x)\cos(2\pi k_n x)\,dx = \begin{cases} L & \text{for } m = n = 0, \\ \frac{1}{2}L & \text{for } m = n > 0, \\ 0 & \text{for } n \neq m, \end{cases} \tag{2.11}$$

$$\int_{x_0}^{x_0+L} \sin(2\pi k_m x)\sin(2\pi k_n x)\,dx = \begin{cases} 0 & \text{for } m = n = 0, \\ \frac{1}{2}L & \text{for } m = n > 0, \\ 0 & \text{for } n \neq m, \end{cases} \tag{2.12}$$

Since a Fourier series contains, in general, both sine and cosine functions, it can be expressed more conveniently using complex variables and Euler's formula,

$$e^{i\theta} = \cos(x) + i\sin(x)\ . \tag{2.13}$$

With this formalism, Eq. 2.4 can be expressed as,

$$f(x) = \sum_{n=-\infty}^{\infty} c_n e^{i2\pi k_n x} \tag{2.14}$$

where the Fourier coefficients are obtained using,

$$c_n = \frac{1}{L}\int_{-L/2}^{L/2} f(x)e^{-i2\pi k_n x}\,dx\ . \tag{2.15}$$

This equation can be verified by inserting Eq. 2.14 into Eq. 2.15 and using the orthogonality relation,

$$\int_{x_0}^{x_0+L} e^{-i2\pi k_m x}e^{i2\pi k_n x}dx = \begin{cases} L & \text{for } m = n, \\ 0 & \text{for } n \neq m, \end{cases} \tag{2.16}$$

which upon using Euler's formula, can be viewed as a generalized expression for Eqs. 2.10 to 2.12. Note that the Fourier coefficients in the complex representation now extend over

positive and negative integers and can be related to the real representation using,

$$\begin{aligned} c_n &= \frac{1}{2}(a_n - ib_n) \ , \\ c_{-n} &= \frac{1}{2}(a_n + ib_n) \ . \end{aligned} \tag{2.17}$$

It is clear from Eq. 2.14 that $f(x)$ has been generalized to a complex valued function, and the coefficients c_n represent the complex amplitude of the spectral components of the function.

2.2 Fourier Transform

Consider the spatial frequency harmonics, $k_n = \left\{\frac{1}{L}, \frac{2}{L}, \ldots, \frac{n}{L}\right\}$. The separation between these spatial frequencies is $\Delta k = 1/L$. With this consideration, inserting Eq. 2.15 into Eq. 2.14 gives,

$$f(x) = \sum_{n=-\infty}^{\infty} \Delta k \int_{-L/2}^{L/2} f(x') e^{-i2\pi k_n x'} \, dx' \, e^{i2\pi k_n x} \ . \tag{2.18}$$

If the period L is taken to infinity, the discrete summation in the above equation becomes an integral. This can be seen by noting the separation between spatial frequencies becomes a vanishingly small differential $\Delta k \to dk$, and the variable $k_n \to n \, dk = k$, becomes continuous. With these considerations, Eq. 2.18 becomes,

$$f(x) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x') e^{-i2\pi k x'} \, dx' \right] e^{i2\pi k x} dk \ , \tag{2.19}$$

which is formally Fourier's inversion theorem [95].

One can identify the expression in square parenthesis in Eq. 2.19 as the Fourier transform of $f(x)$,

$$\hat{f}(k) = \int_{-\infty}^{\infty} f(x) e^{-i2\pi k x} dx \equiv \mathcal{F}\left[f(x)\right] \ , \tag{2.20}$$

where the change of variable $x' = x$ is used. The inverse Fourier transform is then,

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(k) e^{i2\pi k x} dk \equiv \mathcal{F}^{-1}\left[\hat{f}(k)\right] \ . \tag{2.21}$$

In deriving these relations, the use of a positive or negative sign in the exponent is a matter of convention, and the relations hold if they are interchanged. It is also common for the frequency variable to be written as an angular frequency $\kappa = 2\pi k$, where the differential becomes $dk = d\kappa/2\pi$.

The paradigm shift from Fourier series to Fourier transforms is effectively a generalization from periodic functions to arbitrary aperiodic functions. As a consequence of extending the period of the series to infinity, only the first of the Dirichlet conditions must be satisfied by the function $f(x)$ for the transform to exist.

For the derivations of general relationships between Fourier domains, conjugate Fourier pairs of position x, and spatial frequency k, are used. However, these results apply equally to functions described in time t, with temporal frequencies f. Most practically from the perspective of spectral interferometry, the Fourier pairs of Optical Path Difference (OPD) z, and wavenumber $k = 1/\lambda$, where λ is the optical wavelength, are used. For spatial interferometry, the Fourier pairs of angle θ in radians, and wavelength normalized baseline $u = B/\lambda$, where B is the interferometer baseline, are used.

The Fourier transform can be extended to two-dimensions by taking the Fourier transform over the orthogonal components within a given space independently,

$$\begin{aligned} \hat{f}(k_x, k_y) &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f(x,y) e^{-i2\pi k_x x}\, dx \right] e^{-i2\pi k_y y}\, dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) e^{-i2\pi(k_x x + k_y y)}\, dx\, dy\,. \end{aligned} \tag{2.22}$$

This approach can be generalized to arbitrarily many dimensions n, using,

$$\hat{f}(\vec{k}) = \int_n f(\vec{x}) e^{-i2\pi \vec{k}\cdot\vec{x}}\, d\vec{x}\,, \tag{2.23}$$

where $\vec{x}$ and $\vec{k}$ are n-dimensional vectors with components describing orthogonal spatial coordinates and spatial frequencies, respectively.

2.2.1 Symmetry

If a function is *real*, as is the case with any physically realizable signal, its Fourier transform is,

$$\begin{aligned}\hat{f}(k) &= \int_{-\infty}^{\infty} [f_e(x) + f_o(x)]\, e^{-i2\pi kx}\, dx \\ &= \int_{-\infty}^{\infty} f_e(x)\, [\cos(2\pi kx) - i\sin(2\pi kx)]\, dx \\ &+ \int_{-\infty}^{\infty} f_o(x)\, [\cos(2\pi kx) - i\sin(2\pi kx)]\, dx\,,\end{aligned} \tag{2.24}$$

where Eq. 2.1 has been used to represent the general function $f(x)$ in terms of its even and odd components.

Recall that the definite integral of an odd function vanishes when integration is performed with symmetric limits about the origin. Additionally, the product of two even functions, or two odd functions, produce an even function, while the product of an even and odd function produces an odd function. With these considerations, Eq. 2.24 reduces to,

$$\hat{f}(k) = \int_{-\infty}^{\infty} [f_e(x)\cos(2\pi kx) - i f_o(x)\sin(2\pi kx)]\, dx\,. \tag{2.25}$$

It is observed that $\cos(2\pi kx)$ is an even function of k, while $\sin(2\pi kx)$ is an odd function of k. The main point of interest is that the Fourier transform of a real function results in a spectrum with real and imaginary components. The real component is symmetric and encodes the even part of the function. Conversely, the imaginary component is anti-symmetric and encodes the odd part of the function. When the function is real, its transform possesses Hermitian symmetry,

$$\hat{f}(-k) = \hat{f}^*(k)\,, \tag{2.26}$$

as is evident from Eq. 2.25, where * denotes the complex conjugate.

Using a similar analysis, if a function is *imaginary*, the Fourier transform gives,

$$\hat{f}(k) = \int_{-\infty}^{\infty} [i f_e(x) \cos(2\pi kx) + f_o(x) \sin(2\pi kx)] \, dx \, . \tag{2.27}$$

Once again, there are real and imaginary components. In this case, however, the real component is anti-symmetric and encodes the odd part of the function, while the imaginary component is symmetric and encodes the even part of the function. When the function is imaginary, its transform possesses the following symmetry,

$$\hat{f}(-k) = -\hat{f}^*(k) \, . \tag{2.28}$$

2.2.2 Properties and Transform Pairs

Fourier transforms possess a number of useful properties and transform pairs that make analysis more intuitive. A selection of these properties are presented in Tab. 2.1, and a selection of transform pairs are presented in Tab. 2.2. These tables are presented here for convenience, and the equations within are derived in Ap. C. The application of these equations in the context of interferometry are presented in Sec. 2.4 and 2.5.

Parseval's Theorem

Parseval's theorem relates the integrals of the magnitude of a given function squared in both Fourier domains and is given by the equation,

$$\int_{-\infty}^{\infty} |f(x)|^2 \, dx = \int_{-\infty}^{\infty} \left|\hat{f}(k)\right|^2 \, dk \, . \tag{2.29}$$

Table 2.1: A selection of Fourier transform properties [96].

Property	Function	Fourier Transform
Linearity	$af(x)+bg(x)$	$a\hat{f}(k)+b\hat{g}(k)$
Scaling	$f(ax)$	$\frac{1}{\lvert a\rvert}\hat{f}(\frac{k}{a})$
Translation	$f(x-a)$	$e^{-i2\pi ak}\hat{f}(k)$
Phase Shift	$e^{i2\pi ax}f(x)$	$\hat{f}(k-a)$
Multiplication	$f(x)g(x)$	$[\hat{f}*\hat{g}](k)$ †
Convolution	$[f*g](x)$	$\hat{f}(k)\hat{g}(k)$
Correlation	$[f\star g](x)$ ‡	$\hat{f}^*(k)\hat{g}(k)$
Duality	$f(x)$ $\hat{f}(x)$	$\hat{f}(k)$ $f(-k)$

† The symbol $*$, represents the convolution operator.
‡ The symbol $\star$, represents the cross-correlation operator.

Table 2.2: A selection of Fourier transform pairs [96].

Name	Function	Fourier Transform
Cosine	$\cos(2\pi ax)$	$\frac{1}{2}[\delta(k-a)+\delta(k+a)]$
Sine	$\sin(2\pi ax)$	$\frac{1}{2i}[\delta(k-a)-\delta(k+a)]$
Rectangle	$\Pi\left(\frac{x}{a}\right)$	$a\,\mathrm{sinc}(ak)$ †
Sinc	$\mathrm{sinc}\left(\frac{x}{a}\right)$	$a\Pi(ak)$
Exponential Decay	$e^{-a\lvert x\rvert}$	$\frac{2a}{a^2+(2\pi k)^2}$
Step Function Decay	$u(x)e^{-ax}$ ‡	$\frac{1}{a+i2\pi k}$
Gaussian	e^{-ax^2}	$\sqrt{\frac{\pi}{a}}e^{-\frac{(\pi k)^2}{a}}$
Dirac Delta	$\delta(x)$	1
Unit	1	$\delta(k)$
Dirac Comb	$\mathrm{Ш}_L(x)$	$\frac{1}{L}\sum_{n=-\infty}^{\infty}\delta\left(k-\frac{n}{L}\right)$

† The normalized sinc function is defined as $\mathrm{sinc}(x)=\frac{\sin(\pi x)}{\pi x}$
‡ $u(x)$ is the Heaviside step function.

This equality can be verified by inserting the definition of the Fourier transform Eq. 2.20, into the Right-Hand Side (RHS),

$$\begin{aligned}\int_{-\infty}^{\infty} \left|\hat{f}(k)\right|^2 dk &= \int_{-\infty}^{\infty} \left\{ \left[\int_{-\infty}^{\infty} f(x) e^{-i2\pi kx}\, dx \right] \left[\int_{-\infty}^{\infty} f(x') e^{-i2\pi kx'}\, dx' \right]^* \right\} dk \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x) f^*(x') \left[\int_{-\infty}^{\infty} e^{-i2\pi k(x-x')}\, dk \right] dx'\, dx \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x) f^*(x') \delta(x-x')\, dx'\, dx \\ &= \int_{-\infty}^{\infty} f(x) f^*(x)\, dx\,. \end{aligned} \tag{2.30}$$

The third line can be obtained as a combination of the phase shift property and the Fourier transform of the unit function (see Tab, 2.1 and 2.2).

Parseval's theorem is often interpreted as an expression of conservation of energy. In a general sense, the signal squared is referred to as the signal power, regardless of the units of the signal. However, in the context of electronics, signals are usually recorded as, or can be equivalently converted to, a voltage drop across a resistor as a function of time. As such, the units of both sides of Eq. 2.29 have units of $\mathrm{V}^2 \cdot \mathrm{s}$, and if this voltage is measured across a 1Ω resistor, the result is equivalent to energy with units of joules (J). The energy in both reciprocal domains of the Fourier transform are thus equivalent.

2.3 Discretization

Many signals of interest manifest as continuous variables which have defined values at every instant of time, within some time range. Although analog systems exist which can be used to perform various kinds of analysis on continuous signals, digital computers provide greater flexibility and convenience. Information stored in the form of discretized values is also more resistant to degradation. Any amount of noise on a continuous variable degrades the signal. However, a digital signal can be perfectly reconstructed as long as the noise is below the threshold required to resolve different discrete values. Although discretization provides many benefits, there are disadvantages. For example, rounding errors result from

taking a range of continuous values and putting them all in the same bin. However, it is often the case that rounding errors are small compared to other sources of noise present in the system. A more practical concern is that signals must be sampled at finite intervals and the signal is not known during the time between these intervals. In general, this results in loss of information, however, under certain conditions the signal can be sampled without loss of information.

2.3.1 Sampling Theorem

Mathematically, the process of discretizing a continuous signal is accomplished by multiplying the signal with a Dirac comb,

$$\begin{aligned} f_n &= f(x)\text{Ш}_{\Delta x}(x) \\ &= f(x)\sum_{n=-\infty}^{\infty}\delta(x-n\Delta x)\,, \end{aligned} \tag{2.31}$$

where Δx is the sampling interval, which can be either the time or distance between samples depending on the context. Taking the Fourier transform of this equation gives,

$$\begin{aligned} \mathcal{F}\left[f_n\right] &= S(k) * k_s \sum_{n=-\infty}^{\infty}\delta\left(k-nk_s\right) \\ &= k_s \sum_{n=-\infty}^{\infty} S(k-nk_s)\,, \end{aligned} \tag{2.32}$$

where the convolution theorem has been used, and $S(k)$ is the spectrum of the continuous function obtained through its Fourier transform. The Fourier transform of the Dirac comb Eq. C.27, produces another Dirac comb with delta functions separated by intervals of the sampling frequency $k_s = 1/\Delta x$. Eq. 2.32 effectively describes an infinite number of $S(k)$ copies separated by intervals of the sampling frequency, as shown in Fig. 2.1.

If the signal is band-limited such that the minimum frequency is zero and the maximum frequency is k_m, the repeated copies of the signal spectrum will have no overlapping

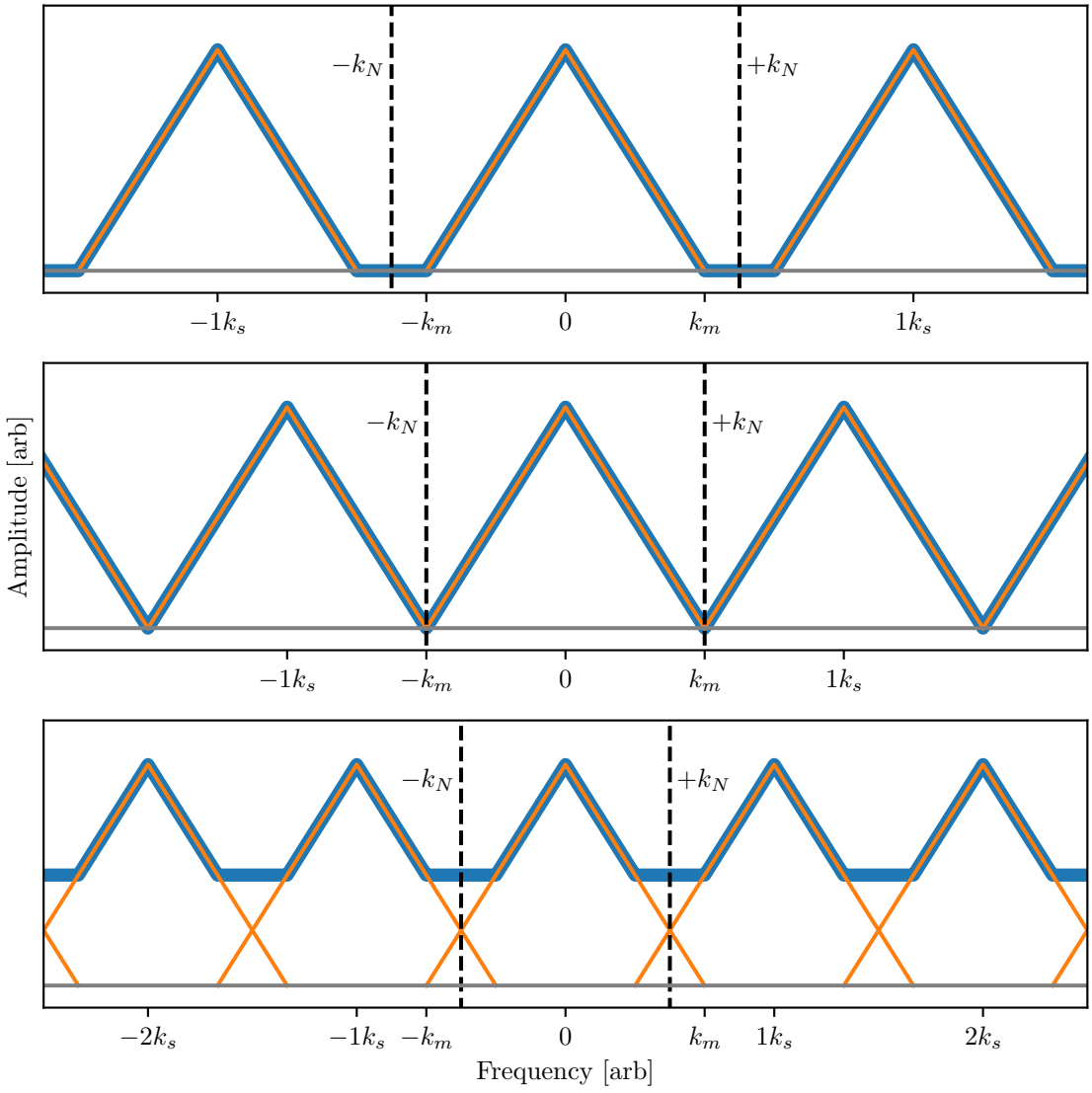


Figure 2.1: Spectra resulting from the Fourier transform of a band-limited signal with various sampling rates. The spectral content of the signal is band-limited to a maximum frequency k_m, as shown by the orange triangles. The signal is sampled at a frequency k_s, which produces the blue curve after transformation, and is consistent with convolving a Dirac comb with the signal spectrum Eq. 2.32. It is observed that when $k_s < 2k_m$, the spectra overlap and higher frequency content is folded over the Nyquist frequency $k_N = k_s/2$, indicated by the dashed line.

segments if

$$k_s \geq 2k_m \, . \tag{2.33}$$

In other words, to reconstruct a signal without distortion, the signal must be sampled at a frequency twice that of the highest frequency present in the signal. This condition is formally the Nyquist–Shannon sampling theorem [97], which motivates the concept of the Nyquist frequency,

$$k_N = \frac{k_s}{2} \, , \tag{2.34}$$

and describes the highest frequency resolvable using a given sampling rate.

When Eq. 2.33 is violated, the signal reconstructed from the sampled points is distorted. The resulting distortion is known as aliasing, an example of which is shown if Fig. 2.2. Effectively, frequency components above Nyquist present in the continuous signal manifest in the sampled signal as frequency components that are less than the Nyquist frequency. This can be understood in terms of a folding effect demonstrated in Fig. 2.1. To avoid aliasing, either the sampling rate must be increased, or high frequency components must be attenuated using a low-pass filter such that the sampling theorem is respected.

2.3.2 Discrete Fourier Transform

Eq. 2.32 is formally the Discrete-Time Fourier Transform (DTFT). It maps an infinite sequence of discrete signal samples, uniformly spaced in time, to a continuous periodic summation of the signal's Fourier transform (i.e., its spectrum) in frequency space. The DTFT is analogous to the Fourier series which maps an infinite number of uniformly spaced discrete samples in frequency space to a continuous and periodic function in time.

Neither paradigm is suitable for real applications since signals can only be measured over a finite amount of time. In practice, signals are discretely sampled over a finite time,

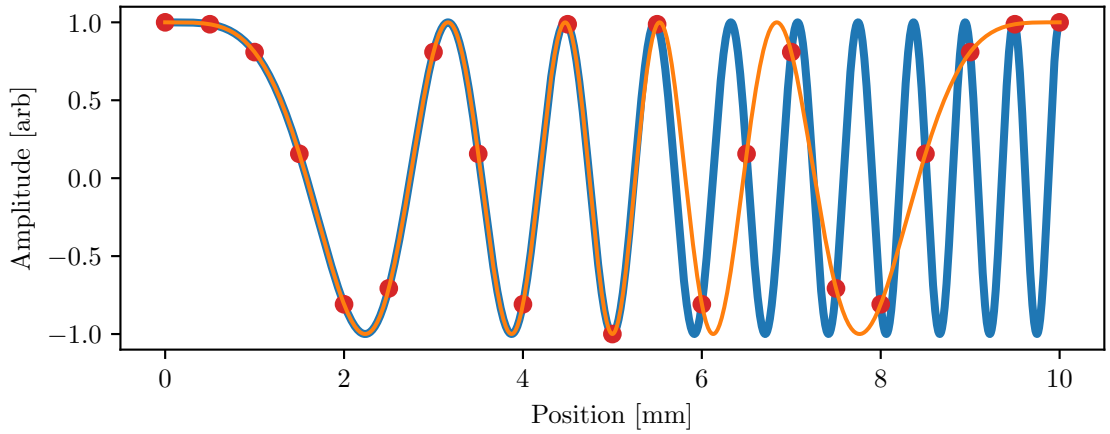


Figure 2.2: Aliasing resulting from undersampling a high frequency signal. The blue curve shows a chirp signal where the frequency increases linearly from $0\,\mathrm{mm}^{-1}$ to $2\,\mathrm{mm}^{-1}$. The signal is sampled at a frequency of $2\,\mathrm{samples mm}^{-1}$, as shown by the red circles, which results in a Nyquist frequency $k_N = 1\,\mathrm{mm}^{-1}$. The orange curve results from plotting the same chirp signal as the blue curve, but the frequency linearly decreases after reaching the Nyquist frequency. Effectively, folding the high frequency components of the blue curve over the Nyquist frequency produces the orange curve. It can be observed that the orange curve is an accurate representation of the sampled points and matches the blue curve perfectly until 5 mm when the frequency of the chirp equals the Nyqyist frequency. After this point, the blue curve is aliased by the samples, meaning the samples can be equally attributed to the lower frequency orange signal.

and frequency components are obtained using the Discrete Fourier Transform (DFT),

$$\hat{f}_r = \sum_{n=0}^{N-1} f_n e^{-i2\pi rn/N}\,, \tag{2.35}$$

with the Inverse Discrete Fourier Transform (IDFT) given by,

$$f_n = \frac{1}{N}\sum_{r=0}^{N-1} \hat{f}_r e^{i2\pi rn/N}\,. \tag{2.36}$$

In this case, f_n are assumed to be N uniformly spaced samples of one cycle of a continuous periodic signal. Similarly, $\hat{f}_r$ are assumed to be N uniformly spaced samples of one cycle of a continuous periodic spectrum.

The validity of these equations can be verified in the following way. Consider the

Fourier transform of a discretized signal,

$$\begin{aligned}\hat{f}(k) &= \int_{-\infty}^{\infty} \left[f(x) \sum_{n=-\infty}^{\infty} \delta(x - n\Delta x) \right] e^{-i2\pi kx}\, dx \\ &= \sum_{n=-\infty}^{\infty} f(n\Delta x) e^{-i2\pi kn\Delta x} \,,\end{aligned} \tag{2.37}$$

where n is the sample index and Δx is the distance between samples. Note that Eq. 2.37 is just an alternative form of Eq. 2.32 and produces a continuous periodically summed spectrum as shown in Fig 2.1. The next step is to discretely sample the spectrum at points $k = r\Delta k$, where r is the sample index and Δk is the distance between frequency samples. Eq. 2.37 is then,

$$\hat{f}(r\Delta k) = \sum_{n=-\infty}^{\infty} f(n\Delta x) e^{-i2\pi rn\Delta k\Delta x} \tag{2.38}$$

Now impose the condition that the function $f(x)$ is periodic and that it has been sampled such that one cycle is measured with N uniformly spaced samples; that is

$$L = N\Delta x \,, \tag{2.39}$$

where L is the period of the signal. Note that a finite signal that is not periodic can be made periodic by appending copies of itself at its ends. Similarly, assume one cycle of the already periodic spectrum $\hat{f}(k)$ has been measured with N uniformly spaced samples; that is

$$k_s = N\Delta k = 1/\Delta x \,, \tag{2.40}$$

where k_s is both the "period" for the spectrum and the sampling rate of the signal as dis-

cussed in Sec. 2.3.1. With these considerations Eq. 2.38 is,

$$\begin{aligned}\hat{f}_r &= \sum_{n=0}^{N-1} f_n e^{-i2\pi rn\Delta k\Delta x} \\ &= \sum_{n=0}^{N-1} f_n e^{-i2\pi rn\Delta k/k_s} \\ &= \sum_{n=0}^{N-1} f_n e^{-i2\pi rn/N} ,\end{aligned} \tag{2.41}$$

which is equivalent to the DFT, Eq. 2.35. A similar process can be performed from the perspective of a Fourier series to obtain the IDFT. The additional factor of $1/N$ is a normalization constant required to make the IDFT the inverse of the DFT.

The discretization process, as described above, and the required sampling of the signal and spectrum for correct use of the DFT and IDFT is shown in Fig. 2.3. In the context of FTS, the relevant signal is an interferogram that is effectively sampled according to the sequence $\{-z_{max}, -z_{max}+\Delta z, ..., 0, ..., z_{max}-\Delta z\}$, where z_{max} is the maximum OPD and Δz is the spacing between samples. This sequence is similar to the orange curve in the top panel of Fig. 2.3. However, the required sampling for correct use in the DFT is as shown by the red circles, thus, the samples must be swapped such that the sequence is $\{0, ..., z_{max}-\Delta z, -z_{max}, -z_{max}+\Delta z, ...\}$. The output spectrum samples are oriented in the same way.

2.3.3 Fast Fourier Transform

For each frequency component of the DFT, N multiplications and additions are required. Since there are N frequency components, the DFT has a computational complexity that scales as $O(N^2)$. Such complexity becomes prohibitively expensive for large input sequences. Fortunately, through clever factorization of the DFT calculations, the complexity scaling can be reduced. Any algorithm that offers efficient calculation of the DFT is referred to as a Fast Fourier Transform (FFT).

The most popular strategy involves recursively splitting the full N-point DFT into two

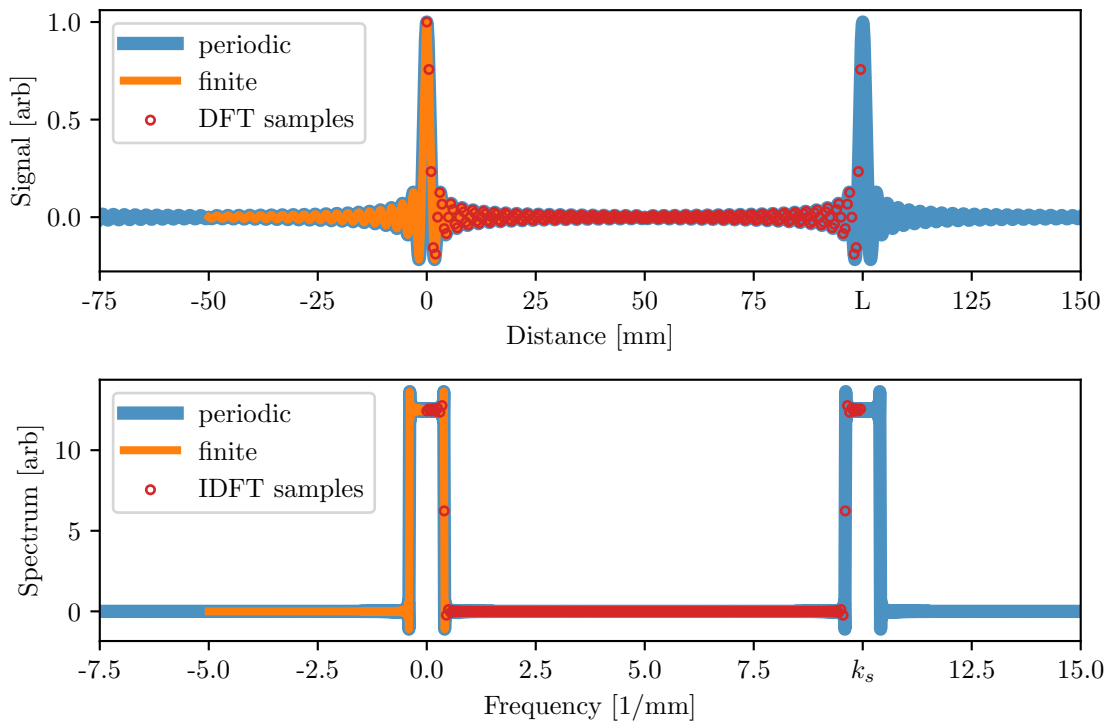


Figure 2.3: Sampling required for use of DFT and IDFT. The top panel shows a finite signal (orange) that has been artificially made infinite and periodic (blue). Red circles show one cycle worth of samples drawn from the infinite periodic signal that are used as inputs to the DFT. Transformation of these values produces a sequence of frequency samples shown by red circles in the lower panel. In applying the IDFT, one proceeds in the reverse direction, from the spectrum to the signal, where the color of the curves in the lower panel have the same interpretation as the top panel.

$N/2$-point DFTs, then combining the results [98]. This is accomplished by separating the DFT into even and odd parts,

$$\begin{aligned}
\hat{f}_r &= \sum_{n=0}^{N-1} f_n e^{-i2\pi rn/N} \\
&= \sum_{m=0}^{N/2-1} f_{2m} e^{-i2\pi r(2m)/N} + \sum_{m=0}^{N/2-1} f_{2m+1} e^{-i2\pi r(2m+1)/N} \\
&= \sum_{m=0}^{N/2-1} f_{2m} e^{\frac{-i2\pi rm}{N/2}} + e^{\frac{-i2\pi r}{N}} \sum_{m=0}^{N/2-1} f_{2m+1} e^{\frac{-i2\pi rm}{N/2}} \\
&= \hat{f}_r^e + e^{\frac{-i2\pi r}{N}} \hat{f}_r^o \,,
\end{aligned} \tag{2.42}$$

where $\hat{f}_r^e$ is the DFT of the even-index parts of f_n, and $\hat{f}_r^o$ is the DFT of the odd-index parts

of f_n. From the periodic nature of the DFI,

$$\begin{aligned} \hat{f}_r &= \hat{f}_r^e + e^{\frac{-i2\pi r}{N}} \hat{f}_r^o \\ \hat{f}_{r+\frac{N}{2}} &= \hat{f}_r^e - e^{\frac{-i2\pi r}{N}} \hat{f}_r^o \ , \end{aligned} \tag{2.43}$$

so that the even and odd index transforms can be reused to calculate the two halves of the full transform. Note that Eq. 2.43 represents N two-point DFTs. The FFT algorithm described is represented visually in Fig. 2.4.

To summarize, the N-point DFT is split into two $N/2$-point DFTs resulting in a computational complexity $O(N^2/2)$. The results are then combined with N two-point DFTs with computational complexity $O(2N)$. The full complexity is $O(2N + N^2/2)$ which for $N = 1024$ results in $\sim 5 \times 10^5$ multiplications and additions which should be compared to the $\sim 10^6$ such operations required for the DFT. When this process is repeated recursively, the computational complexity approaches $O(N \log_2 N)$. As such, when N is large, the FFT can compute the transform $N/\log_2 N$ times faster than the DFT.

2.3.4 Comparison to Fourier Transform

The DFT preserves many of the same characteristics of the Fourier transform, however, a few subtle changes should be noted.

Units

The Fourier transform multiplies the units of the input sequence with the units of the independent variable that is being integrated over. For example, a voltage signal measured at regular intervals of time will have a Fourier transform with units of $V \cdot s$. The DFT is a weighted summation of the input sequence, and as such, the output units are the same as the input units.

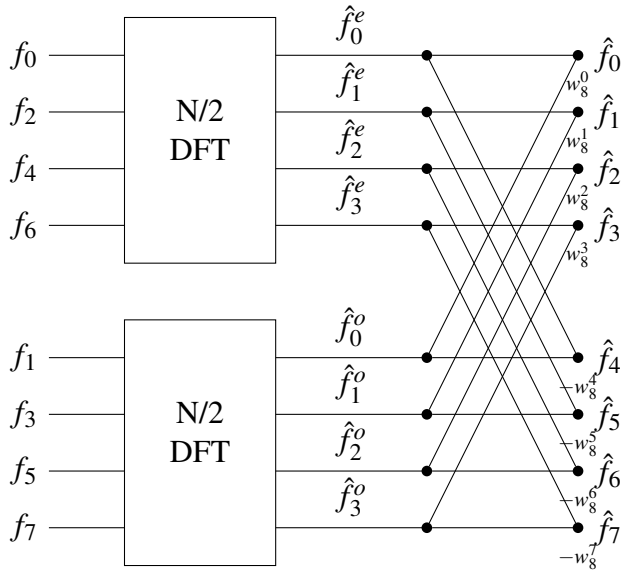


Figure 2.4: Visual representation of one iteration of the FFT. The full N-point DFT is split into two $N/2$-point DFTs consisting of the even and odd index points of the input sequence. The outputs are then combined using what is effectively N two-point DFTs which computes the full transform. Computational efficiency is obtained by re-using the calculated $\hat{f}_r^e$ and $\hat{f}_r^o$ values through linear combinations using the coefficients $w_N^r = e^{-i2\pi r/N}$ (Eq. 2.43).

Normalization

The Fourier transform pairs presented in Tab. 2.2 are mostly valid when using the DFT. However, the expressions for $\hat{f}(k)$ must be normalized by the number of point in the input sequence for strict equality, that is,

$$\hat{f}_r = \hat{f}(k)/N \,. \tag{2.44}$$

Parseval's Theorem

As a consequence of the items above, Parseval's theorem Eq. 2.29 for the DFT is modified to,

$$\sum_{n=0}^{N-1} |f_n|^2 = \frac{1}{N} \sum_{r=0}^{N-1} \left|\hat{f}_r\right|^2 \,. \tag{2.45}$$

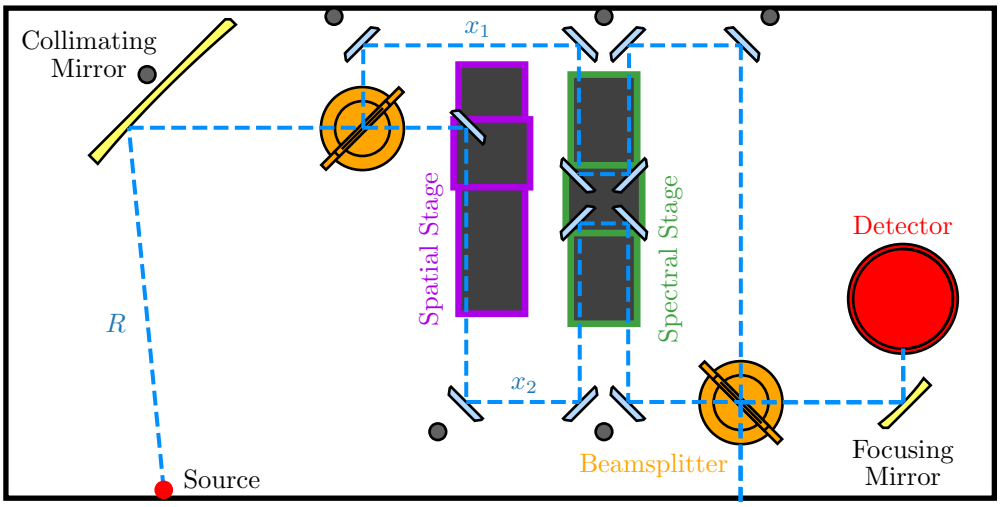


Figure 2.5: FTS configuration used for this work. Input light is split by a 50:50 intensity beamsplitter then propagated along two paths with optical path lengths x_1 and x_2. The beam are combined using a second beamsplitter and imaged onto the detector plane. OPD is controlled by the spectral stage and the resulting signal is an interference pattern recorded as a function of OPD.

2.4 Fourier Transform Spectroscopy

An FTS is an instrument that probes the spectral content of light. Input radiation is split into two beams and an OPD is induced between the two beam before they are recombined and imaged onto a detector (see Fig. 2.5). The resulting signal is an interference pattern as a function of OPD, often called an interferogram. OPD can be interpreted as a time delay $\tau = z/c$, between the two beams, and the recorded intensity can be interpreted as the correlation between electric fields in the two beams with that particular time delay. As such, an FTS measures the temporal coherence of input radiation which is encoded in an interferogram. These interferograms consist of a superposition of sinusoids that are discretely sampled over finite intervals, and the analysis of these data is, to a large extent, applied Fourier analysis. In this section, a discussion of interferograms and their analysis is presented in the context of Fourier analysis.

2.4.1 Monochromatic Source

The ideal response of an FTS to a monochromatic source of frequency k_0, measured as a function of OPD $z = x_2 - x_1$, is

$$I_{FTS}(z) = 2I_0 \cos(2\pi k_0 z) \ , \tag{2.46}$$

where I_0 is the intensity of light incident on the detector from each arm of the interferometer. The intensity in each arm is assumed to be equal and proportional to the intensity of light emitted by the source. The constant term has been omitted for convenience, and a full derivation of the response for the system shown in Fig. 2.5 is presented in Ap. B. Evidently, the monochromatic response is a cosine function with amplitude $2I_0$. The sinusoidal nature of the response is a consequence of the two beams passing through regions of constructive and destructive interference as OPD changes. The Fourier transform of a cosine function is (see C.28),

$$\mathcal{F}\left[2I_0 \cos(2\pi k_0 z)\right] = I_0 \left[\delta(k-k_0) + \delta(k+k_0)\right] \ . \tag{2.47}$$

The Fourier transform of the monochromatic response is two delta functions centred on $\pm k_0$, the optical frequency of the incident light, and with amplitude equal to I_0, the intensity in each arm of the interferometer.

If some arbitrary phase offset ϕ is applied to the cosine function, the result is,

$$\begin{aligned} \mathcal{F}\left[2I_0 \cos(2\pi k_0 x - \phi)\right] &= I_0 e^{-i\phi k/k_0} \left[\delta(k-k_0) + \delta(k+k_o)\right] \\ &= I_0 \left[e^{-i\phi}\delta(k-k_0) + e^{i\phi}\delta(k+k_o)\right] \ , \end{aligned} \tag{2.48}$$

where the translation property is used in the first line and the filtering property of the delta functions is used in the second line. Applying an offset to the cosine effectively rotates the delta functions within the complex plane with the delta function at $\pm k_0$ rotated by $\mp\phi$.

Consider the complex value of the delta function at $k = k_0$. Expanding the first term on

the RHS of Eq. 2.48 gives,

$$I_0 \left[\cos(\phi) - i\sin(\phi)\right] = S_{Re} + iS_{Im} \,, \tag{2.49}$$

where S_{Re} and S_{Im} are the real and imaginary components of the complex value, respectively. By noting that $\tan(\phi) = \sin(\phi)/\cos(\phi)$, the phase can be recovered using,

$$\phi = \arctan\left(\frac{S_{Im}}{S_{Re}}\right) \,. \tag{2.50}$$

As a consequence of the symmetry relation discussed in Sec. 2.2.1, if a general asymmetric function is fully real, the phase spectrum will be anti-symmetric due to the anti-symmetry of the imaginary component. This statement is also verified by repeating the previous steps for the delta function centered at $-k_0$, which results in a different sign for the imaginary component in Eq. 2.49. Furthermore, by noting $\sin^2(\phi) + \cos^2(\phi) = 1$, the magnitude of the complex value can be obtained using,

$$I_0 = \sqrt{S_{Re}^2 + S_{Im}^2} \,. \tag{2.51}$$

2.4.2 Polychromatic Source

The ideal response of an FTS when observing a polychromatic source is a superposition of sinusoids,

$$I_{FTS}(z) = 2\int_0^\infty I(k)\cos\left(2\pi kz\right) dk \,. \tag{2.52}$$

According to the linearity property, the Fourier transform of this response produces a sum of complex value delta functions each individually possessing the characteristic previously outline in this Sec. 2.4.1, as shown in Fig. 2.6 for two frequencies. This means $S_{Re}(k)$ and $S_{Im}(k)$ from Eq. 2.49 can be interpreted as real and imaginary components of a complex spectrum, respectively. These components can be used to calculate a phase spectrum $\phi(k)$ and amplitude spectrum $I(k)$, using Eq. 2.50 and 2.51, respectively. In particular, the

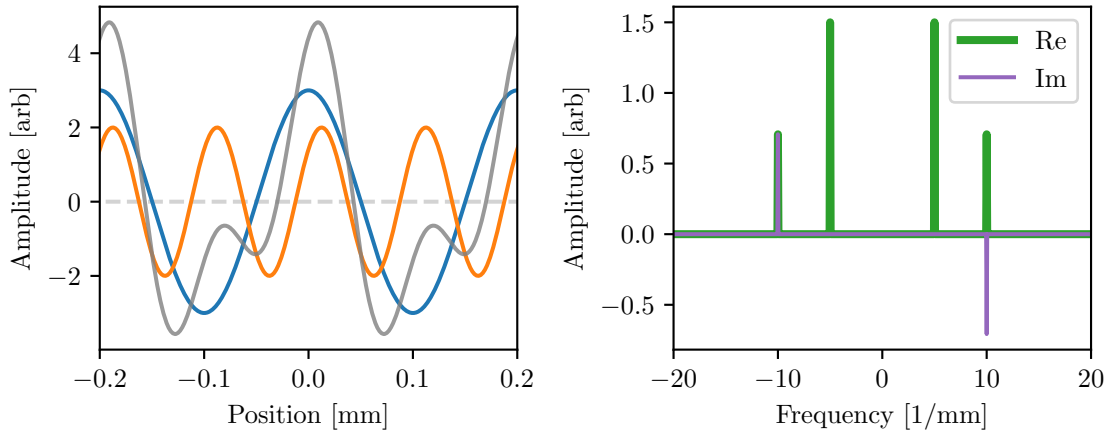


Figure 2.6: Cosines and their Fourier transforms. The left panel shows two cosine functions with different amplitude, frequency, and phase. Their sum is shown in gray. The right panel shows the Fourier transform of the gray curve where contributions from the constituent cosines are evident. Peaks at $k = \pm 5\,\mathrm{mm}^{-1}$ with amplitude 1.5 are fully real and correspond to the blue cosine with the same frequency and twice the amplitude. Peaks at $k = \pm 10\,\mathrm{mm}^{-1}$ are complex with real and imaginary components resulting from the phase shift applied to the orange cosine.

Fourier transform provides a powerful tool for simultaneously extracting spectral components from the response of an FTS which correspond to the optical frequencies present in the source.

Assuming the amplitudes of the sinusoids are constant over a finite bandwidth $2I(k) = A$, and there are no phase errors, the polychromatic response is,

$$\begin{aligned} I(z) &= \int_0^{k_{max}} A\cos(2\pi kz)\,dk \\ &= A\left[\frac{\sin(2\pi kz)}{2\pi z}\right]_0^{k_{max}} \\ &= Ak_{max}\,\mathrm{sinc}(2k_{max}z)\ , \end{aligned} \tag{2.53}$$

as shown in Fig. 2.7. Although a flat source spectrum is not realistic, this result demonstrates many important characteristics of interferograms obtained when observing sources with broad spectra. In particular, the interferogram is symmetric and expresses a central lobe, often called the white-light fringe, with oscillating side-lobes that decay asymptoti-

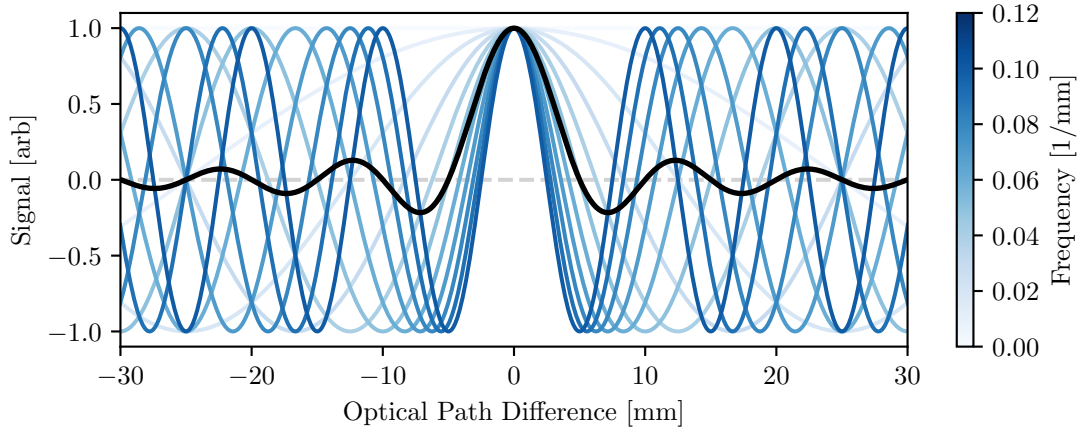


Figure 2.7: Interferogram generated from the superposition of cosine functions. The cosine functions have uniform amplitude and frequency indicated by their shade of blue. The average of these cosines gives the black sinc function.

cally to zero. The width of the main lobe is proportional to the coherence length of the input radiation, and is inversely proportionate to the bandwidth. In general, as the bandwidth increases, the interferogram becomes taller and narrower as expected from the scaling property of Fourier transforms (see Fig. 2.8).

2.4.3 Spectral Resolution

Spectral resolution describes the minimum frequency separation between statistically independent samples in a spectrum. A higher spectral resolution (lower frequency separation) makes it easier to distinguish between closely spaced spectral features and provides greater accuracy on feature positions. The spectral resolution that can be achieved with a particular interferogram is determined by the nature of the DFT and the symmetric property of the interferogram. It can be inferred from Eqs. 2.39 and 2.40 that the spectrum is sampled at intervals of,

$$\Delta k = \frac{1}{L} \rightarrow \frac{1}{2z_{max}} \,. \tag{2.54}$$

Here, as in Sec. 2.3.2, L represents the full length of the sampled signal (interferogram in this case), and z_{max} is the maximum optical path difference. In practice, it is often the

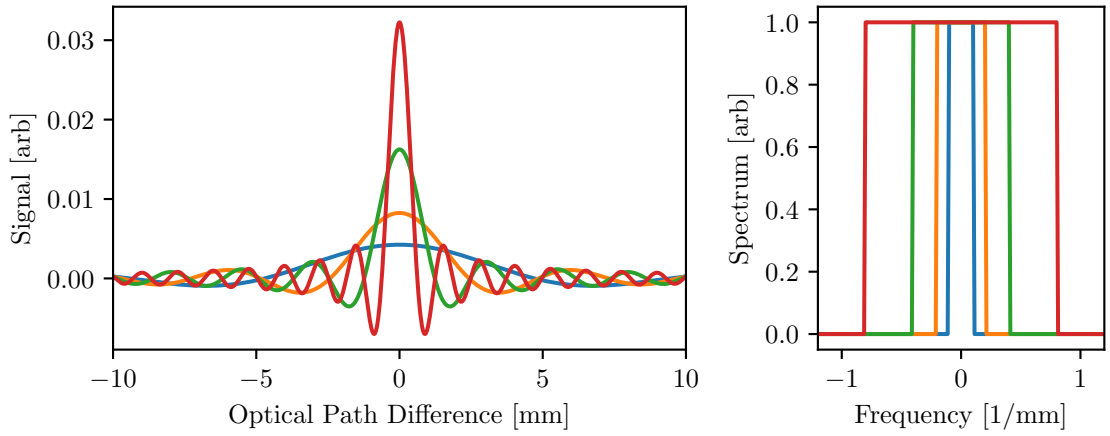


Figure 2.8: Interferogram (left) generated from spectra with different bandwidths (right). As the bandwidth increases, the resulting interferogram increases in amplitude and becomes more narrow.

case that $2z_{max} < L$, implying a one-sided interferogram where more samples are taken on one side of the interferogram compared to the other side. However, the symmetry of an ideal FTS interferogram, $I_{FTS}(-z) = I_{FTS}(z)$, allows for the missing measurements to be populated by the complementary value on the other side of the interferogram. In principle an FTS interferogram can always be constructed such that $2z_{max} = L$.

2.4.4 Bandwidth

Bandwidth describes the range on frequencies present in a signal. Through application of the phase shift property of the Fourier transform, the minimum frequency can be translated to zero, and after this translation, the maximum frequency k_{max}, is equal to the bandwidth. The following statements are given in the context of a bandwidth that ranges from $[0, k_{max}]$.

For an interferogram, the bandwidth properties are determined by the sampling rate as described in Sec. 2.3.1. In particular, the Nyquist frequency or maximum resolvable frequency is calculated using Eqs. 2.34 and 2.40,

$$k_N = \frac{1}{2\Delta z}\,, \tag{2.55}$$

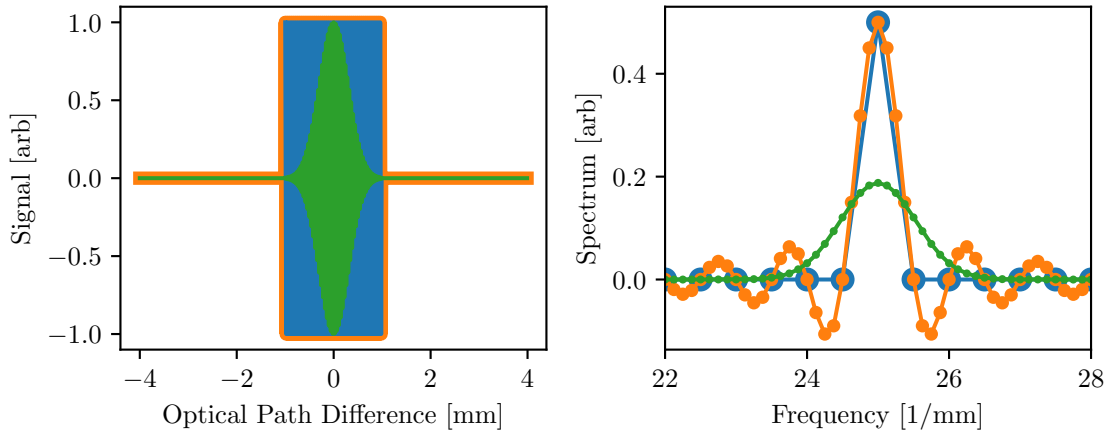


Figure 2.9: Effects of zero padding and apodization. The left panel shows a sinusoidal interferogram (blue) that is zero padded (orange) and then apodized with a Gaussian function (green). The right panel shows the resulting spectra. The spectrum resulting from the zero padded signal is a sinc interpolation of the spectrum resulting from the unpadded signal. The spectrum resulting from the apodized signal shows how side-lobes are attenuated using apodization.

where Δz is the separation between OPD samples in the interferogram. To avoid aliasing, Δz should be selected such that $k_N \geq k_{max}$. If the signal is known to be band-limited over a bandwidth that does not include zero frequency, then aliasing can be used to reduce the sampling conditions required for full signal recovery [99].

2.4.5 Zero Padding

Since spectral resolution is determined by the maximum OPD, it is possible to synthetically increase the spectral resolution of a spectrum by appending zeros to the ends of the interferogram. This process is called zero padding and is demonstrated in Fig. 2.9. Although zero padding does not add information to the interferogram and does not increase the true spectral resolution, it can still be useful for mapping interferograms of different lengths, thus of different spectral resolutions, onto the same frequency grid. Additionally zero padding can be used to increase the number of samples to a multiple of two, facilitating the use of the FFT (Sec. 2.3.3). The increased sampling density in the spectral domain also provides advantages when using regression techniques to analyze spectral features.

The left panel of Fig. 2.9 shows interferograms, and the blue curve represents a monochromatic interferogram consisting of a cosine function with $k = 25\,\mathrm{mm}^{-1}$. The right panel shows the corresponding spectra obtained using the Fourier transform, which for the blue curve, is a delta function centred at $k = 25\,\mathrm{mm}^{-1}$, as expected. The blue interferogram is zero padded by symmetrically appending zeros to both sides, increasing the maximum OPD by a factor of four, which produces the orange interferogram. The resulting spectrum is a sinc function with a sample density four times the unpadded interferogram. Notably, the zero crossings of the sinc functions are at the sampled points of the original spectrum. In particular, the width of the sinc function, as measured by the distance between the first zero crossings, is identical to the width of the delta function. As such, zero padding has not increased the true spectral resolution and the resulting effect is a sinc interpolation of the original spectrum.

An alternative interpretation of the above result is to suppose the blue interferogram initially extended the full range of the padded interferogram. The orange interferogram is then obtained by multiplying the blue interferogram with a truncating rectangle function. As a consequence of the multiplication property of Fourier transforms, the spectrum of the orange signal is obtained as the convolution of the transform of the blue signal (a delta function) and the transform of the rectangle function (a sinc function). In the context of FTS, the transform of the truncating function gives the Instrument Line Shape (ILS), and the spectrum is obtained as the convolution of the ILS with delta functions corresponding to different spectral components of the source radiation.

2.4.6 Apodization

In a general sense, any operation that modifies a signal is a form of apodization. In FTS, apodization is typically accomplished by multiplying the interferogram by some kind of window function. For example, multiplying the interferogram by a rectangle function, as discussed in Sec. 2.4.5, is a kind of apodization. A consequence of this apodization was

to convert the delta function to a sinc function. In many cases, this result is not desirable as less prominent spectral features may be obscured by the side-lobes of more prominent spectral features.

The presence of side-lobes, consisting of decaying oscillations, is a phenomena known as ringing and results from discontinuities in the interferogram. Such discontinuities typically result from measuring interferograms that do not decay to zero at the maximum measured OPD. Ringing is suppressed for interferograms from coherent sources, which do not decay, if the frequency components present in the interferogram are sampled exactly in the spectral domain, as is the case with the blue curve in Fig. 2.9. For a broadband source, ringing can be suppressed using an appropriate apodization function.

The main characteristic of an apodization function that attenuates ringing is that, when multiplied with the interferogram, the ends of the interferogram are smoothly tapered to zero. One such function is a Gaussian, the use of which is demonstrated ind Fig. 2.9. The green curve is constructed by multiplying the orange interferogram by a Gaussian function such that the interferogram is tapered sufficiently to zero before the discontinuity resulting from zero padding. Ringing has been significantly attenuated in the spectral domain by this process at the cost of a wider spectral feature with lower amplitude. All such apodization functions express this trade-off, and optimal functions provide maximal side-lobe suppression at the cost of minimal line width distortion [100].

In practice, the best way to minimize ringing is by sampling the interferogram out to an OPD where the signal is below the noise. This has the added benefit of maximizing Signal to Noise Ratio (SNR) with less processing complexity. Even when extending the OPD is not possible, the use of an apodization function is contentious as information is inevitably lost and may complicate spectral feature analysis [100].

2.5 Spatial Interferometry

While FTS probes the temporal coherence of light in order to determine the source spectrum, spatial interferometry probes the spatial coherence of light in order to determine the source intensity distribution. In this section, the principles of spatial interferometry are discussed in the context of Fourier analysis.

2.5.1 Fundamentals

Spatial coherence $\Gamma_0(\vec{B})$, is quantified by correlating the electric fields sampled at two regions of space. The vector separation between the two sampled regions is called the base line $\vec{B}$. By measuring spatial coherence at multiple baselines, the monochromatic source intensity distribution $I(\vec{\theta}, \nu_0)$, can be reconstructed using a process called aperture synthesis [101]. The theoretical foundation of this technique is provided by the van Cittert-Zernike theorem,

$$\Gamma_0(\vec{B}) = \int_\Omega I(\vec{\theta}, \nu_0) e^{-i2\pi\vec{B}\cdot\vec{\theta}/\lambda_0} \, d\Omega \,, \tag{2.56}$$

which is derived from physical principles in Ap. A. In effect, the spatial coherence of the electric field in the Fraunhofer regime is given by the two-dimensional Fourier transform of the source intensity distribution. The reciprocal coordinates for this transform are angles $\vec{\theta}$, in radians, which describes the source plane, and the *uv*-plane defined in terms of spatial frequencies, or equivalently, wavelength normalized baseline $\vec{B}/\lambda_0$. The decomposition of an image into its spatial frequencies is shown in Fig. 2.10.

Since the van Cittert-Zernike theorem is a Fourier transform, the properties developed in this chapter apply to the analysis of the spatial coherence. In particular, much of the intuition about DFT carries over with two main conceptual changes. First, the analysis is generally performed in two spatial dimensions instead of one temporal dimension. Second, the observation technique measures spatial frequency components in order to reconstruct a signal rather than measuring a signal to recover spectral components.

An arbitrary one-dimensional signal can be constructed as a Fourier series consisting of

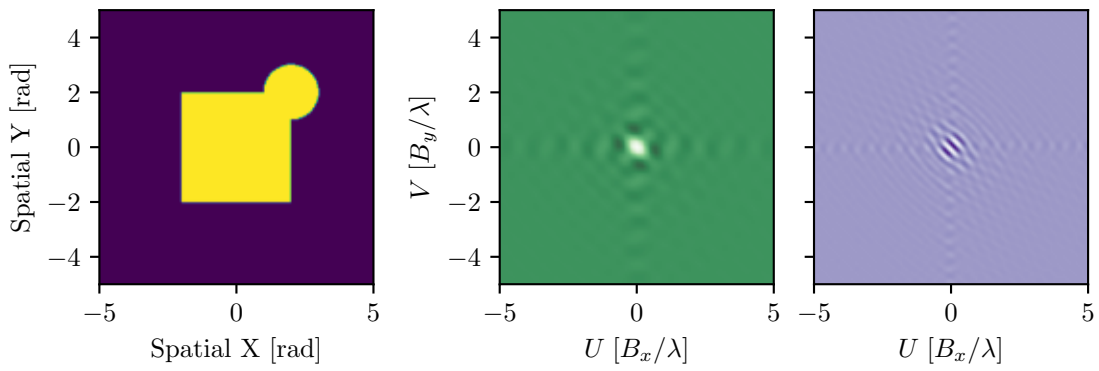


Figure 2.10: Two-dimensional Fourier transform of an image. A simple image (left) defined using angular coordinates is decomposed into its real (center) and imaginary (right) spatial components. The spatial components are defined on the *uv*-plane using wavelength normalized baseline units. Real components are symmetric about the origin while imaginary components are anti-symmetric showing Hermitian symmetry.

a superposition of sinusoids possessing various amplitudes, frequencies, and phases. Similarly, a two-dimensional signal can be constructed in the same way, except in this case, the sinusoids are expressed in two-dimensions, as shown in Fig. 2.11. These two-dimensional sinusoids are referred to as spatial frequency components and are mapped onto the *uv*-plane based on their orientation and the number of cycles per radian. Components closer to the origin of the *uv*-plane have a smaller spatial frequency, and as these components are rotated, the orientation of the waves in the spatial domain is rotated in the same way. The amplitude and phase of the spatial components are given by the magnitude and phase of its corresponding value in the *uv*-plane, and calculated using Eqs. 2.51 and 2.50, respectively.

Baseline Sampling

The source intensity distribution is a continuous function and so is its Fourier transform, represented as a complex spatial frequency distribution in the *uv*-plane. The *uv*-plane must be sampled discretely, and assuming the source intensity distribution is band-limited, sampling must be performed according to the Nyquist sampling theorem 2.33 in order to reconstruct the source without aliasing. Practical limitations prevent Nyquist sampling of the *uv*-plane, and sampling is often sparse and non-uniform.

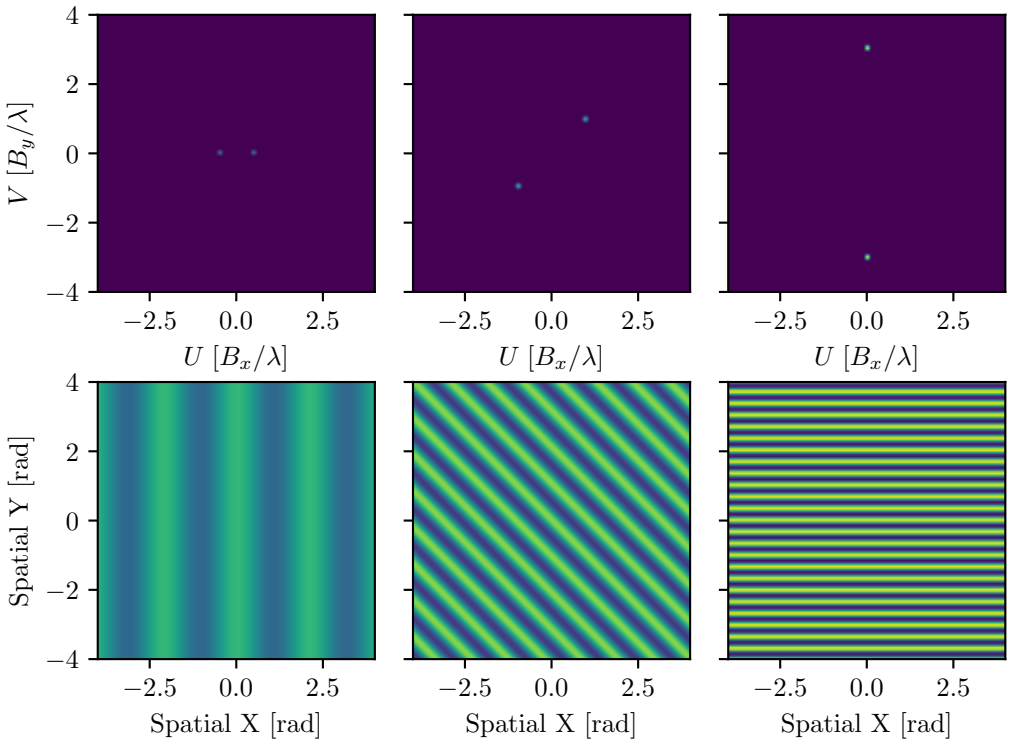


Figure 2.11: Relationship between *uv*-samples and spatial frequency components. The top row shows the *uv*-plane which have been populated by a single spatial frequency component of varying magnitude and orientation. The bottom row shows the resulting spatial distribution, obtained as the inverse double-Fourier transform of the above *uv*-plane. From left to right is shown increasing magnitude, rotation, and distance from the origin in the *uv*-plane resulting in increasing amplitude, rotation, and spatial frequency in the spatial domain.

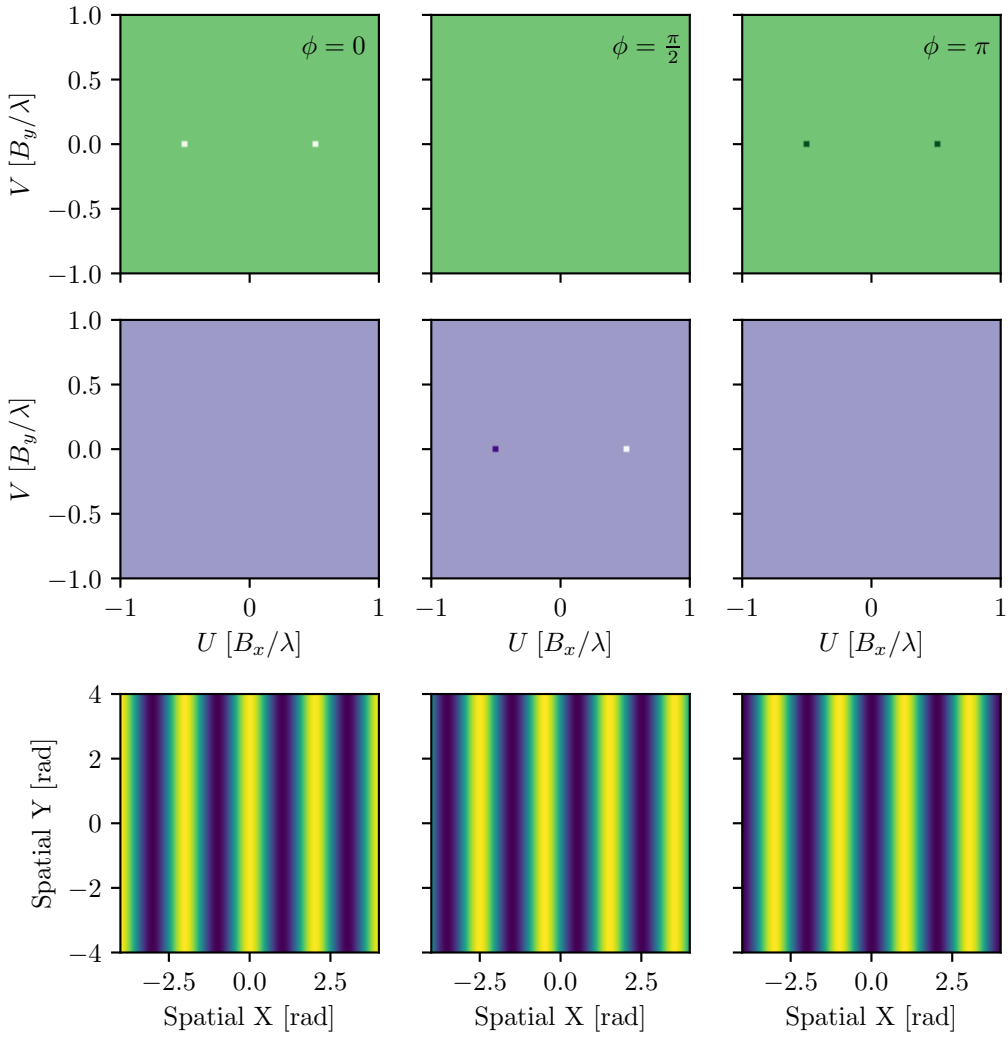


Figure 2.12: The effect of phase in the *uv*-plane on spatial signal distribution in the spatial domain. The real (top) and imaginary (middle) components of the *uv*-plane with various induced phases are shown along the the resulting spatial distributions (bottom). Non-zero phase results in a rotation of the positive baseline *uv* samples in the complex plane. Negative baseline samples are inferred from Hermitian symmetry. The resulting spatial distributions obtained as the inverse two-dimensional Fourier transform of the *uv*-plane show a two-dimensional sinusoid offset by an amount equivalent to the induced phase.

Sparse sampling can be modelled as applying a mask to the *uv*-plane. By applying a mask to Eq. 2.56 and taking the inverse Fourier transform, the resulting source reconstruction using the multiplication property is,

$$\begin{aligned}\mathcal{F}^{-1}\left\{M_{u,v}\Gamma_0(\vec{B})\right\} &= \mathcal{F}^{-1}\left\{M_{u,v}\right\} * \mathcal{F}^{-1}\left\{\Gamma_0(\vec{B})\right\} \\ &= \tilde{M}_{\vec{\theta}} * I(\vec{\theta},\nu_0)\ ,\end{aligned} \tag{2.57}$$

where $M_{u,v}$ is the masking function in the *uv*-plane, and $\tilde{M}_{\vec{\theta}}$ is its inverse Fourier-transform. In the context of aperture synthesis, $\tilde{M}_{\vec{\theta}}$ is called the dirty beam. Sparse sampling in the *uv*-plane thus produces a dirty image consisting of the source intensity distribution convolved with the dirty beam. An equivalent interpretation is that sparse sampling in the *uv*-plane applies a spatial frequency filter to the source intensity distribution that attenuates spatial frequencies not included in the masking function. The effects of such sampling is demonstrated in Fig. 2.13.

Since the source intensity distribution is always real for physical sources, its Fourier transform possesses Hermitian symmetry $\Gamma_0(-\vec{B}) = \Gamma_0^*(\vec{B})$. Thus, when measuring spatial coherence, only half the *uv*-plane must be sampled since the other half is obtained for free. Intuitively, this makes sense as the baseline vector is arbitrarily drawn from one sampled region in space to the other, and inverting the direction would not change the useful information obtained from a measurement. Hermitian symmetry can be observed in Figs. 2.10, 2.11, and 2.12.

Sparse sampling typically truncates the distribution in the *uv*-plane resulting in significant ringing in the spatial domain. Removing this ringing is equivalent to deconvolving the dirty beam from the dirty image to produce a clean image. Many cleaning algorithms have been developed with most being variations of the clean [102] and maximum entropy method [103].

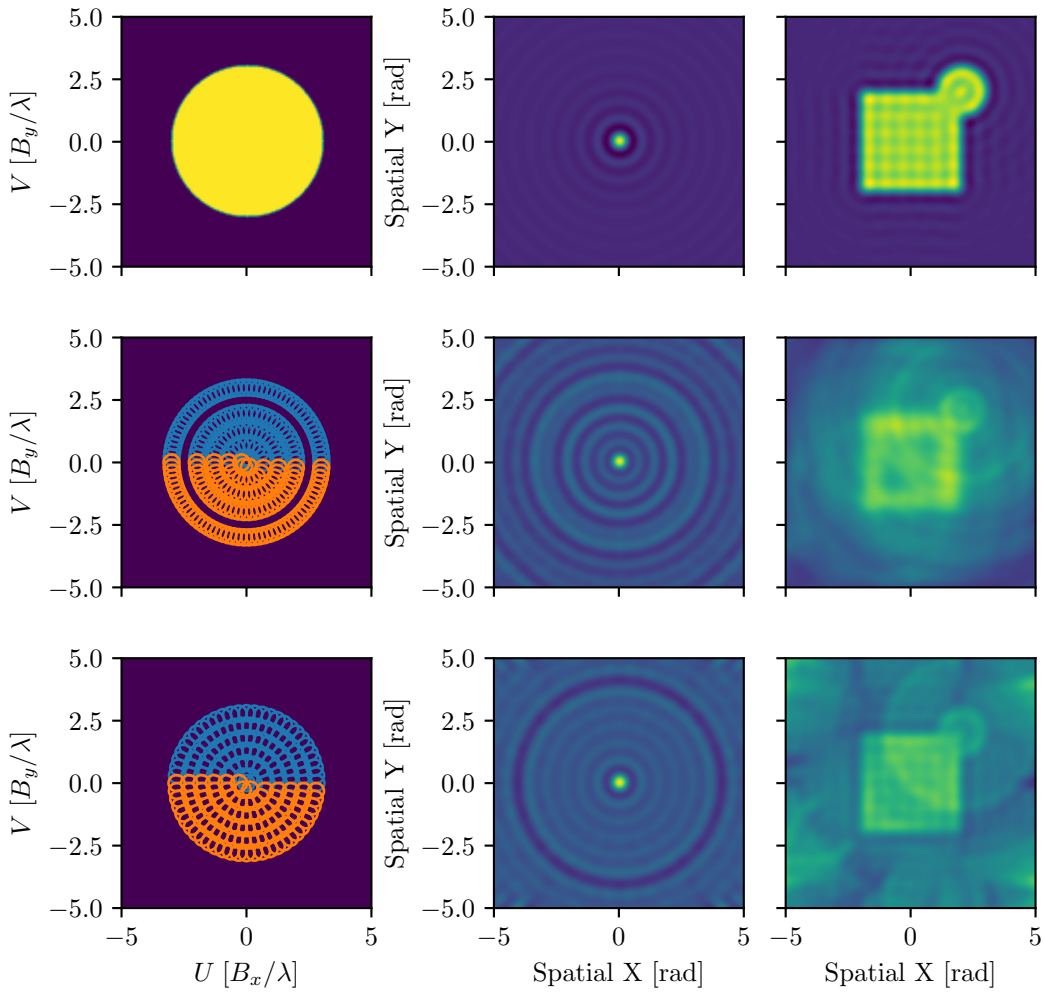


Figure 2.13: The effects of sparse *uv*-plane sampling. Sampling of the *uv*-plane (left), resulting dirty beam (middle), and reconstructed dirty image (right) are shown for various sampling strategies. The top row shows the sampling characteristic of a single aperture telescope with dirty beam given by an Airy pattern. The second and third rows show sampling strategies that are practical for a spatial interferometer with blue circles for measured values and orange circles for symmetric values. The second row uses a Gaussian weighted sampling density while the third row uses a uniform sampling density. The undistorted image used for these simulations is shown in Fig. 2.10.

Spatial Resolution

Spatial resolution describes the minimum separation between two point sources such that the sources can still be distinguished individually. At separations less than the spatial resolution, the sources are blended and appear as a single source. The finite nature of spatial resolution results from the wave characteristics of light and and diffraction effects of finite apertures. When a point source is imaged, its light is spread over some regions of the image plane where the intensity distribution is given by the diffraction pattern resulting from the electric field diffracting through the aperture. For this reason, the intensity distribution obtain in this way is called the Point Spread Function (PSF). In the Fraunhofer regime, and assuming the aperture is uniformly illuminated, the PSF is obtained as the square of the Fourier transform of the aperture geometry. A familiar example is the Airy pattern, which is the PSF associated with a circular apertures. More complicated PSFs result when using a spatial interferometer. They vary from sinusoids for a single baseline configuration, similar to that shown in Fig. 2.11, to the square of the dirty beams shown in Fig. 2.13.

A common metric used for quantifying the spatial resolution of an optical system is the Rayleigh criterion [8]. It states that two sources are considered resolved if the maximum of one PSF coincides with the first minimum of the other. For a spatial interferometer with baseline magnitude B, observing light with wavelength λ, this criterion gives,

$$\theta_{min} = \frac{\lambda}{2B}\,, \tag{2.58}$$

This result is consistent with Fourier analysis and Nyquist sampling where the spatial frequency $k = B/\lambda$, requires a spatial sampling $\Delta\theta = 1/2k$. More complicated uv sampling strategies produce more complicated PSFs. However, due to the design of a practical spatial interferometer, it is most convenient to vary the baseline length and orientation. This naturally produces sampling in the form of concentric circles or spirals similar to Fig. 2.13. In the dense sampling limit, uv coverage approaches a circle where the outer perimeter cor-

responds to samples taken with a maximum baseline B_{max}. The PSF associated with such coverage is an Airy pattern and is the same as would be obtained with a single aperture telescope with a primary mirror of diameter $D = B_{max}$. As such, in the dense sampling limit, the spatial resolution or a spatial interferometer is,

$$\theta_{min} = 1.22\frac{\lambda}{B_{max}} \, . \tag{2.59}$$

Of course, dense sampling is rarely obtained due to time constraints, and the spatial resolution associated with an observation tends to reside between the limiting cases of Eq. 2.58 and 2.59.

2.5.2 Interferometer Response

The spatial interferometer used for this work is shown in Fig. 2.14. In fact, the design is a DFI that combines a spatial interferometer with the OPD varying infrastructure of an FTS. Electric fields sampled at two regions of space are propagated though the optical system then combined at the beamsplitter and imaged onto the detector array. The spectral stage controls the OPD $z = x_2 - x_1$, between the two beams, and the spatial stage controls the baseline separation B, between the two input apertures. A full derivation of the instrument response is given in Ap. B, and the main result is,

$$I_{DFI}(\vec{B},z) = -e^{-i\phi_{bs}}\left[\int_{\Omega}\int_{-\infty}^{\infty} I_{12}(\vec{\theta},k)e^{-i2\pi k\vec{B}\cdot\vec{\theta}}e^{i2\pi kz}\,dk\,d\Omega\right] \, , \tag{2.60}$$

where $I_{12}(\vec{\theta},k)$ is the geometric mean of the power spectrum of the light incident on the detector resulting from each beam, and is proportional to the source power spectrum. Symmetry has also been imposed onto the source spectrum to extend the frequency integration limits to $\pm\infty$. The quantity in square brackets is identified as the Mutual Coherence Function (MCF) Eq. A.12, and is a generalized version of the van Cittert-Zernike theorem Eq. 2.56, suitable for sources of arbitrary spectral width. The detector can only measure

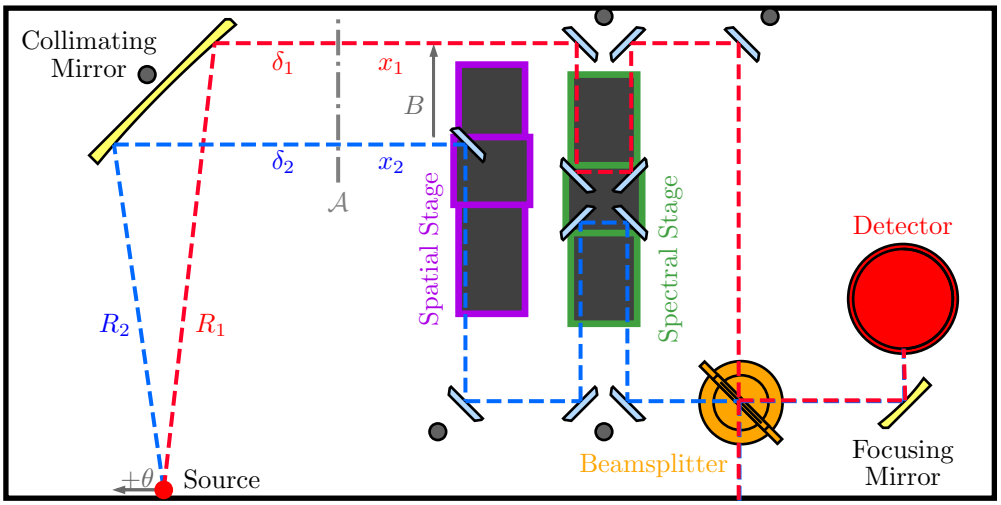


Figure 2.14: DFI configuration used for this work. Input light is sampled by two apertures then propagated along two paths with optical path lengths x_1 and x_2. The beam are combined using a second beamsplitter and imaged onto the detector plane. The baseline is controlled by the spectral stage, and OPD is controlled by the spectral stage. The resulting signal is an interference pattern recorded as a function of OPD at a particular baseline.

real intensity resulting in a detected signal,

$$\begin{aligned} I_{DFI}(\vec{B},z) &= -\int_{\Omega}\int_{-\infty}^{\infty} I_{12}(\vec{\theta},k)\cos\left(2\pi k(z-\vec{B}\cdot\vec{\theta})-\phi_{bs}(k)\right)\,dk\,d\Omega \\ &\rightarrow \int_{\Omega}\int_{-\infty}^{\infty} I_{12}(\vec{\theta},k)\sin\left(2\pi k(z-\vec{B}\cdot\vec{\theta})\right)\,dk\,d\Omega\,, \end{aligned} \tag{2.61}$$

where the second line is obtained from the approximation $\phi_{bs}(k)\approx\pi/2$ and the fact that the detector signal is inverted such that an increase in incident power produces a lower signal. Eq. 2.61 describes the response at a particular baseline $\vec{B}$, as a superposition of monochromatic signals with frequency $k=1/\lambda$, resulting from point sources at angular positions $\vec{\theta}$. A conceptual simplification can be made by taking the two-dimensions source intensity distribution and projecting it onto the axis parallel to the baseline vector. This reduced the situation to a one-dimensional problem where vector quantities can be replaced by scalar quantities.

To develop intuition on the character of the response function, simulated outputs when observing various on-axis sources with a 130 mm interferometer baseline are presented in

Fig. 2.15. A monochromatic point source results in a single sine function with a constant amplitude and frequency k. A polychromatic point source with a rectangular spectrum produces multiple monochromatic sine waves of linearly increasing frequency. The average of the individual sine waves results in the final response, an antisymmetric function with an envelope that decreases asymptotically to zero as OPD magnitude increases.

Now consider an extended source represented as a line of identical point sources uniformly distributed along the baseline direction and centered on the optical axis. When the source is monochromatic, each point source will produce a sine function with constant frequency and amplitude. If a point source is displaced from the optical axis by an angle θ, the sine wave will be offset on the OPD axis by $\Delta z = B\theta$. The average of these sine waves is itself a sine wave with the same frequency as its constituents but with decreased amplitude. When calculated this way, the decreased amplitude is indicative of the decreased visibility or contrast of fringes. The variation of fringe visibility is one of the main observables of a spatial interferometer, and gives information about the magnitude of the complex visibility. As shown by this example, the visibility is related to source structure. If the source is polychromatic, the result is the same as the polychromatic point source but reduced in amplitude due to the same visibility effects just described.

When the source is displaced from the optical axis by an angle θ, the resulting interference pattern will also be shifted along the OPD axis by $\Delta z = B\theta$, as previously discussed. This time, note the dependence on baseline which amplifies the OPD offset. As the baseline increases, the separation between the responses of individual point sources which constitute an extended source will widen, typically resulting in lower visibility fringes. If the center of an extended source is displaced from the optical axis, the central feature of the interference pattern is also shifted by an amount proportional to the baseline. This displacement of fringes along the OPD axis is the other main observable of a spatial interferometer and gives information about the phase of the complex visibility. These effects are shown in Fig. 2.16 for an off-axis polychromatic extended source.

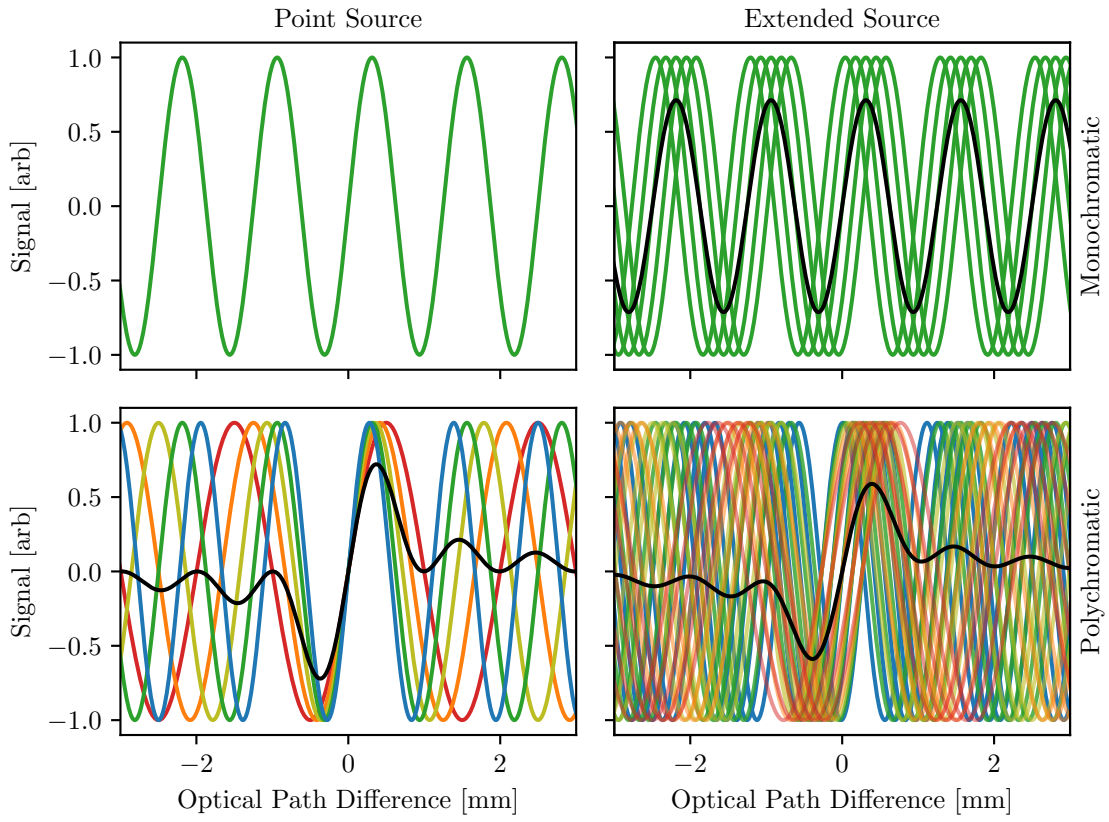


Figure 2.15: DFI response to various on-axis sources. The left column is associated with point source observations and the right column is associated with extended source observations. The general consequence of sources extent is to reduce the visibility of fringes owing to the spread of interference patterns resulting from the individual, uncorrelated, point sources that make up an extended source. The top row shows monochromatic observations and the bottom row shows polychromatic observations. The superposition of monochromatic interference patterns with different frequencies results in a prominent central feature that decays with increasing OPD magnitude. The frequency of monochromatic interference patterns are indicated by colors with the net response indicated by black curves.

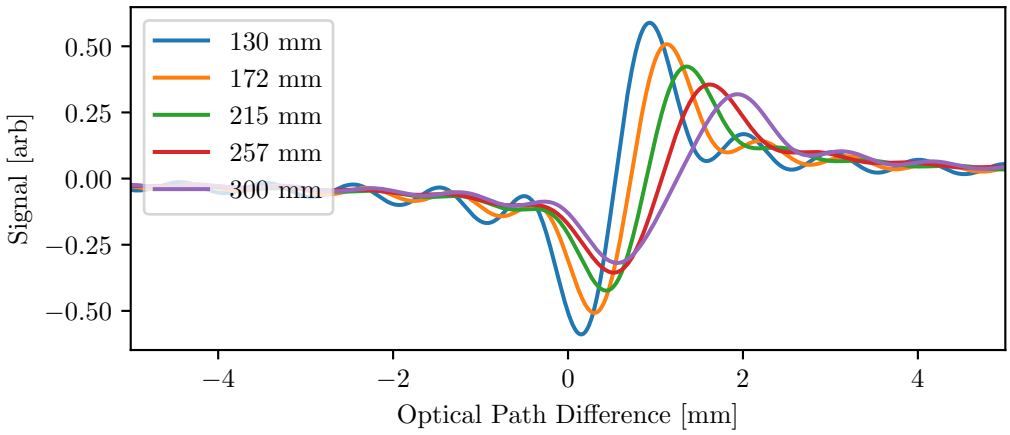


Figure 2.16: Effects of baseline length on DFI response. In general, as the baseline length increases when observing an extended source, the visibility of fringes decrease. If the center of the source is not aligned with the optical axis, the interferogram will be offset along the OPD axis for non-zero baselines. This offset will increase as the baseline length increases.

Integrating the instrument response over the source plane, Eq. 2.61 can be represented as,

$$I_{DFI}(\vec{B},z) = \int_{-\infty}^{\infty} I_{12}(k)|\gamma_0(\vec{B},k)| \sin\left(2\pi kz + \phi_0(\vec{B},k)\right) dk \,, \tag{2.62}$$

where $|\gamma_0(\vec{B},k)|$ is the magnitude of the complex visibility manifested as fringe visibility, and $\phi_0(\vec{B},k)$ is the phase of the complex visibility manifested as a shift of the fringes along the OPD axis. The full response is a superposition of such modulated and shifted functions corresponding to each frequency present in the source intensity distribution. As discussed in Sec. 2.4.1, the Fourier transform of a sinusoid produces a delta function with real and imaginary components. The amplitude and phase can be recovered from these components using Eqs. 2.51 and 2.50, respectively (keeping in mind the $\pi/2$ phase difference between sine and cosine functions). The utility of this techniques comes from the linear property of Fourier transforms in that real and imaginary components are computed simultaneously for all frequencies present in the source. Thus the Fourier transform of the instrument response Eq. 2.62, quantifies the complex visibility, at a particular baseline, for all frequencies present in the source.

To reconstruct the source intensity distribution, the complex visibilities $\gamma_0(\vec{B},k)$, ob-

tained as described above, are projected into the *uv*-plane at coordinates (kB_x, kB_y). In general, when there are frequency depended variations in the source intensity distribution, a *uv*-plane is constructed for each frequency. Repeating this process for multiple baseline lengths and orientations, the *uv*-plane can be suitably sampled to meet observation requirements. The inverse Fourier transform of the *uv*-plane produces a dirty image similar to that depicted in Fig. 2.13, specific to each frequency.

If the source is spectrally uniform, in the sense that the source preserves the same shape when viewed at any frequency, complex visibilities from different frequencies can be used to populate a single *uv*-plane. Recall that *uv* coordinates correspond to different spatial frequency components. At a given baseline, higher optical frequencies probe higher spatial frequencies and a single baseline observation can be used to populate multiple points in the *uv*-plane resulting in higher quality images. In order to combine *uv* samples from all frequencies, the spectral variations in source intensity $I_{12}(k)$, must be normalized using a zero baseline measurement $\vec{B}$=0, which removes the effects of source extent and visibility.

2.6 Summary

Classical theory models light as vibrations of the electromagnetic field. As such, these disturbances and their superpositions can be described as a sum of harmonic oscillators with various amplitudes, frequencies, and phases. The Fourier series constructs arbitrary functions using the same formalism. Associated with the Fourier series is the Fourier transform, which decomposes the arbitrary function into its frequency components and provides information about the amplitude and phase associated with each frequency component. This relationship between the signal and spectral domain makes Fourier analysis the ideal framework with which to study interferometry.

In this chapter, a review of Fourier analysis was presented. The Fourier series was introduced, and from this starting point, the Fourier transform was derived. A number of convenient symmetries, theorems, and Fourier transform pairs where discussed which help

make the relationship between Fourier domains more intuitive. Detailed derivations for these theorems and transforms are presented in Ap. C. In practice, continuous signals must be sampled discretely at periodic intervals. The way in which the signal is sampled affects the resulting bandwidth and spectral resolution of the spectrum obtained by evaluating the DFT. The effects of discretization were discussed in detail. Perhaps the most important result from this study of discretization is that, to avoid aliasing in the relevant spectral or spatial domain, the signal must be sampled at greater than or equal to twice the maximum spectral or spatial frequency present within the signal. The DFT was derived, and the way in which the signal must be sampled for correct use of the DFT was emphasized. An overview of the FFT, an efficient algorithm for fast computation of the DFT, was also presented.

After the fundamentals of Fourier analysis were presented, these ideas were further developed in the context of FTS and spatial interferometry. In particular, it was shown how the spatial and spectral structure of the source being observed is encoded in the instrument response and recorded interferogram. It was also shown how parameters associated with the source structure can be extracted from the interferogram with the application of the Fourier transform and the analysis of the resulting spectrum. Concepts including zero padding and apodization were discussed in the context of FTS. Furthermore, concepts of *uv*-plane sampling and how this relates to synthesized image quality was discussed in the context of spatial interferometry.

Chapter 3

Interferometer System

In this chapter, the interferometer system used in this work is presented along with descriptions and characterizations of all the relevant components, with the exception of the detector system which is the subject of Ch. 4. These characterizations are used to determine many of the observing capabilities of the instrument, and are used to verify that the individual components perform with a precision suitable for interferometry. The optical design is also discussed and modelled to first order with the paraxial approximation in order to derive equations that describe the mapping between conjugate points on the source plane to the detector plane. This model is further refined with ray tracing experiments to study aberrations, and beam modelling to study diffraction effects. Field of View (FOV) measurements are then used to assess the validity of the optical model.

3.1 Overview

The interferometer system can be configured to operate in two different observing modes; Double-Fourier Interferometer (DFI) mode, and Fourier Transform Spectrometer (FTS) mode, as shown in Fig. 3.1. In DFI mode, a source is placed at the focal point of an off-axis parabola, and emits light as a spherical wave. On incidence with the collimating mirror, the spherical wave is converted to a plane wave and propagates towards the input apertures of the interferometer system. The apertures are circular mirrors at 45° incidence, and through a series of reflections off identical shaped mirrors, the light is relayed to a beamsplitter which combines the light. Half the power of the light is directed out of the in-

strument, while the other half is directed towards a focusing mirror which images the light onto a 5×5 feed horn coupled array of detectors. The superposition of electric fields from both arms of the interferometer result in an intensity measurement on the detectors.

A double-rooftop mirror configuration is mounted on the spectral stage, (i.e., the one closest to the detector system (see. Fig. 3.1)). The position of this stage controls the optical path difference between the two beams before they are combined at the beamsplitter. This design results in a folding factor of four, meaning that every one unit of mechanical stage travel results in four units of Optical Path Difference (OPD) change. A single flat mirror is mounted on the spatial stage (the one closest to the collimating mirror (see. Fig. 3.1)), and acts as an input aperture to the interferometer and is called the dynamic aperture. The position of the spatial stage controls the separation between the two input apertures, effectively setting the spatial interferometer baseline. The second input aperture is called the static aperture and is the first flat mirror along the optical path in upper arm of the instrument. When the spectral stage is scanned at a constant speed, the detector system records a signal that encodes both the temporal and spatial coherence of the electric fields sampled by the input apertures.

In FTS mode, an additional beamsplitter is placed in front of the dynamic aperture and an additional flat mirror is placed in front of the static mirror. This configuration is functionally equivalent to a DFI system with a baseline length equal to zero. In this configuration, the detector system encodes exclusively the temporal coherence of the electric field entering the interferometer.

Most relevant, in terms of design considerations that affect the instrument's ability to meet observation goals, is DFI mode. With respect to the development of the instrument, the available sources, linear translation stages, beamsplitters, and detector system precede the optical design phase, and are taken as constraints rather than free parameters. The design space is then defined in terms of the sizes, shapes, and positions of the collimating, focusing, and flat mirrors. These parameters are then constrained by the following obser-

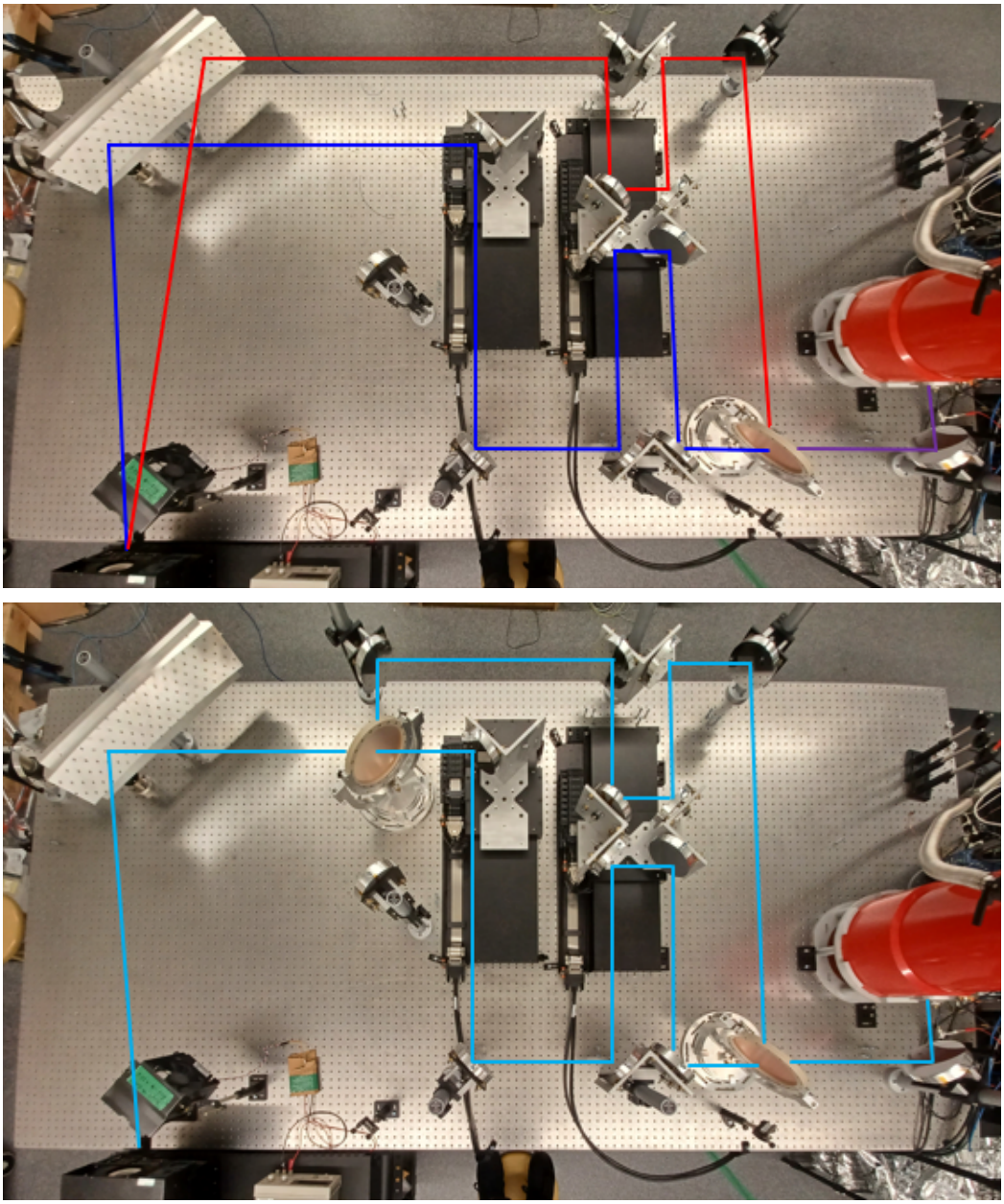


Figure 3.1: Pictures of the interferometer system in DFI mode (top), and FTS mode (bottom). These pictures have complementary diagrams presented in the Ch. 2, as Fig. 2.14 and Fig. 2.5, respectively. In these pictures are shown the main components of the interferometer system including: Pegasus thermal source, collimating mirror, flat mirrors, Aerotech translations stages, beamplitters, focusing mirror, and cryostat housing the detector array.

vation objectives:

- Resolve a 1 mm source at 1 THz in DFI mode.
- A DFI spatial resolution that is at least 5 times better than the primary beam of the detectors (see § 3.4.2).

The first item is somewhat arbitrary, but was chosen to be as small as is practical for the manufacturing of aperture masks, and is of dimensions large enough that a $1,000\,^{\circ}$C source of the same size is bright enough to be visible with the detector system. The second item is a necessary condition for convincingly demonstrating the DFI technique, which attempts to resolve spatial features smaller than the primary beam. As further constraints, all components should fit on the 3×1.5 m optics table and be justified with the table grid to facilitate alignment. The aperture mirrors should be as small as possible in order to sample the smallest baseline possible, but large enough that they do not limit throughput. Throughput should be optimized for the central detector for best case Signal to Noise Ratio (SNR) and should be limited by the detector feed horn acceptance angle. Finally, as always, manufacturing of the components must fit within the available budget. This point is most relevant to the collimating mirror as it is a large monolithic structure that must be machined to a custom shape with high precision. Although it is the most important mirror in the design, it must be minimized to smallest acceptable dimensions that allow for the target spatial resolution and sufficient throughput.

Instrument parameters and performance metrics are summarized in Tab. 3.1. For such quantities that can be estimated based on the system optical design, calculations and measurements are presented in this chapter. Parameters related to interferometer performance are developed in this chapter, with validation details left until Ch. 5. Detector specific quantities are presented in Ch. 4.

Table 3.1: Interferometer system parameters.

Parameter	Value	Units
collimating mirror focal length	$1,200$	mm
focusing mirror focal length	240	mm
flat mirror diameter	135	mm
field of view (pixel/array)[†]	$14/75$	mm
field of view (pixel/array)	$40/215$	arcmin
magnification	-0.2	-
detector feed horn $f_{\#}$	4.5	-
optical NEP	7.1	$\mathrm{pW}/\sqrt{\mathrm{Hz}}$
detector time constant	1.0	ms
sampling rate	13	kHz
optical bandwidth	$0.3-2.2$	THz
baseline range	$135-310$	mm
spectral resolution[‡]	0.19	GHz
spatial resolution ($\nu = 1.2\,\mathrm{THz}$)[†]	1.0	mm
spatial resolution ($\nu = 1.2\,\mathrm{THz}$)	3.0	arcmin

[†] Defined on the source plane Eq. 3.18.
[‡] Maximum resolution for an on-axis point-source.

3.2 Instrument Components

In this section, descriptions and characterizations of the individual interferometer components are presented. Technical drawing for appropriate components are given in Ap. D.

3.2.1 Detector

The detector system is the most complicated part of the interferometer and is critical to its overall performance. As such, Ch. 4 is dedicated to a detailed description and characterization of the detector operation and performance. A summary of the most important aspects are presented here.

The detector system was manufactured by QMC Instruments and consists of a 5×5 array of Transition Edge Sensor (TES) bolometers housed within a $\sim 4\,\mathrm{K}$ cryostat. Bolometers are coupled to the incoming electric field by a rectangular feed horn array. Each horn has an f-number $f_{\#} = 4.5$, and a square entrance aperture with dimensions $3.63\,\mathrm{mm}$. Along

the horizontal or vertical axis, the horn centers are separated by 4.0 mm, thus the full array accepts light incident on the feed horn plane defined by a square with dimensions 19.63 mm. The f-number of the horns define a cone with an acceptance angle of $\sim 12.7°$. All other optical components in the system were oversized such that this acceptance angle would be the limiting factor in throughput for the central detector, and its projection onto the focusing mirror is taken as the aperture stop. Due to the proprietary nature of the detector system, some important information is missing which affects optical design and system characterization. One such piece of information is the feed horn beam shape. In the absence of this information, it is assumed the beam is approximately a radially symmetric Gaussian with a Full Width at Half Maximum (FWHM) defined by the acceptance angle. Discussion of the beam profile incorporating square feed horns is presented in § 3.4.2.

After the square entrance aperture, the feed horns smoothly taper to a square waveguide with dimensions 0.59 mm. The size and shape of the waveguide determines which modes of light can be propagated through the waveguide and eventually reach the detector. An analysis of which modes can be propagated is presented in § 4.3.4. Following the waveguide is an integrating cavity where the TES resides. The TES itself consists of a pure Niobium filament as the superconducting material, coupled to a heating wire, all on a gold absorber. Mode rejection at lower frequencies results in a cutoff at ~ 250 GHz, while reduced coupling efficiency begins to attenuate sensitivity at ~ 1 THz with no signal observable after ~ 2.2 THz. Additionally, there is an aggressive frequency attenuation at higher frequencies that is not explained by the beamsplitter efficiency. Although the beamsplitter transmission strongly influences the spectral characteristics of the overall interferometer system, the limiting factors on the instrument bandwidth seem to be a result of the detector system.

Unlike most TES bolometers which operate based on electrothermal feedback, the detectors used in this work are regulated by a Proportional-Integral–Derivative (PID) controller feedback system. Performance of the detector system strongly depends on the PID gains and other configuration parameters. The parameters used when making science mea-

surements in this work result in a sensitivity characterized by the Noise Equivalent Power (NEP) of, $7.1\,\mathrm{pW}/\sqrt{\mathrm{Hz}}$, and a detector response time of 1.0 ms representing a low-pass cutoff frequency of 174 Hz.

The 25 detectors are time-division multiplexed over 4 channels such that 7 readout cycles are required to measure the full detector array. Nominally, each readout cycle takes $11\,\mu$s and a full frame requires $77\,\mu$s. This measurement time gives a detector sampling rate of $12,914$ Hz. In addition to incident optical power, the detector system also records a high/low signal on an ancillary digital channel. This channel is useful for synchronization with external hardware and is sampled each readout cycle.

3.2.2 Linear Translation Stages

High precision Aerotech ALS20045 linear translation stages were used for the spatial and spectral stages of the interferometer. As stated by the manufacturer, the stages possess a 450 mm linear travel range, with $\pm 1.0\,\mu$m accuracy and $\pm 0.5\,\mu$m repeatability. Using an analogue sine wave encoder, the internal metrology obtains a resolution of 2 nm. The stage controller provides a digital signal that can be configured to produce a variety of digital waveforms that relate to stage motion. Typically this signal is configured to send out a pulse at regular intervals of stage travel, and is fed into the detector system's digital channel to register detector data with stage position. As such, this signal is called the Position Synchronized Output (PSO).

Stage calibration and complicated motion functions can be constructed using a suite of software provided by the manufacturer. Stage motion is controlled though a PID feedback system, and the controller parameters must be calibrated for optimal performance when nominal speed or loads on the stages are changed. Calibration was always performed using the manufacturer software, however, stage motion was facilitated using custom software developed in Python. The Python software manages a network connection that sends and receives ASCII strings conforming to the Aerotech Ensemble syntax. In addition to many

other features, stage motion can be controlled and the stage position can be queried in this way, providing a more integrated platform for interacting with the interferometer system.

As discussed in Sec. 2.4, processing interferometer data involves the use of the Fast Fourier Transform (FFT) algorithm, which requires signals to be sampled at regular OPD intervals. The detector system used in this work samples the signal at regular time intervals, and to ensure equally spaced OPD samples, the stage must travel at constant speed. In most cases, the sources being observed result in low SNR. This is due to the small area of the source being observed, reduced visibility of fringes resulting from the nature of a spatial interferometer measurement, sensitivity limitations of the detector, and imperfections in the optical system. To improve SNR, many interferograms must be averaged, usually more than a thousand. For averaging to be effective, the stage motion must have a high degree of repeatability. These are the two qualities required from the stage: constant speed and repeatability, both to high accuracy. More quantitatively, the OPD errors resulting from stage motion should be less than $\lambda/10$ which, accounting for the folding factor of four of the double-rooftop mirror configuration on the spectral stage, implies positional errors of the stage must be less than $3.4\,\mu$m at $2.2\,$THz.

To develop some confidence in the internal stage metrology, interferograms obtained when observing a thermal source are considered. The thermal source acts as an external reference of stage position in that the temporal coherence of the radiation is low and results in an interferogram with a localized feature. This feature, called the white light fringe, and has a center that corresponds to Zero Path Difference (ZPD) of the FTS. Using this feature, and assuming good SNR, interferograms can be aligned without reference to the stage metrology. Alternatively, interferograms can be aligned using the stage metrology then averaged. If the difference between the white light fringe aligned average and the metrology aligned average is negligible, the stage metrology can be considered sufficiently reliable.

An interferogram obtained in FTS mode when viewing a $1,200\,^{\circ}$C is shown in Fig. 3.2.

The signal is then passed though a band-pass filter that removes frequency components that are outside of the optical band. The peak of the filtered signal is taken as the estimate for ZPD, which is subject to a $\pm 1\,\mu$m rounding error. Subsequent scans are subjected to the same processing and then the ZPD aligned interferograms are averaged producing the white light fringe result in the lower panel of Fig. 3.2. For each scan, the stage was configured to produce a PSO pulse every time the stage mechanical position moves $15\,\mu$m, corresponding to a Nyquist frequency of 2.5 THz, which is fed into the detector system's digital channel. Interferograms are then aligned to the 15th PSO pulse, excluding the acceleration phase of the stage motion. After averaging all scans aligned in this way, the result is passed though the same band-pass filter described above. This PSO aligned interferogram average is also shown in the lower panel of Fig. 3.2. Both methods represent the average of 500 forward scans and produce nearly identical results.

For a quantitative analysis of the stage motion repeatability, the start and end positions of the spectral stage for multiple scans are considered. These data are obtained during a DFI mode observation and consist of $1,200$ forward and reverse scans after 8 hours of continuous operation. Position measurements are queried from the stage controller and depend on the internal stage metrology. Mean subtracted distributions of the start and end positions for forward and reverse scans are shown in Fig. 3.3. Also shown is the mean subtracted distribution of the total stage travel for the forward and reverse directions. The distributions related to start and end positions are all similar with approximately Gaussian shapes and standard deviations of 2.3 nm. The stage travel distributions are also approximately Gaussian with standard deviations of 3.2 nm. These results show that the stage does not experience any systematic drift in scan range during long observations and that the errors associated with random variations in stage travel is about three orders of magnitude lower than required.

High precision in stage motion repeatability is particularly important for DFI observations. Individual interferograms are of low SNR and averaging hundreds of interferograms

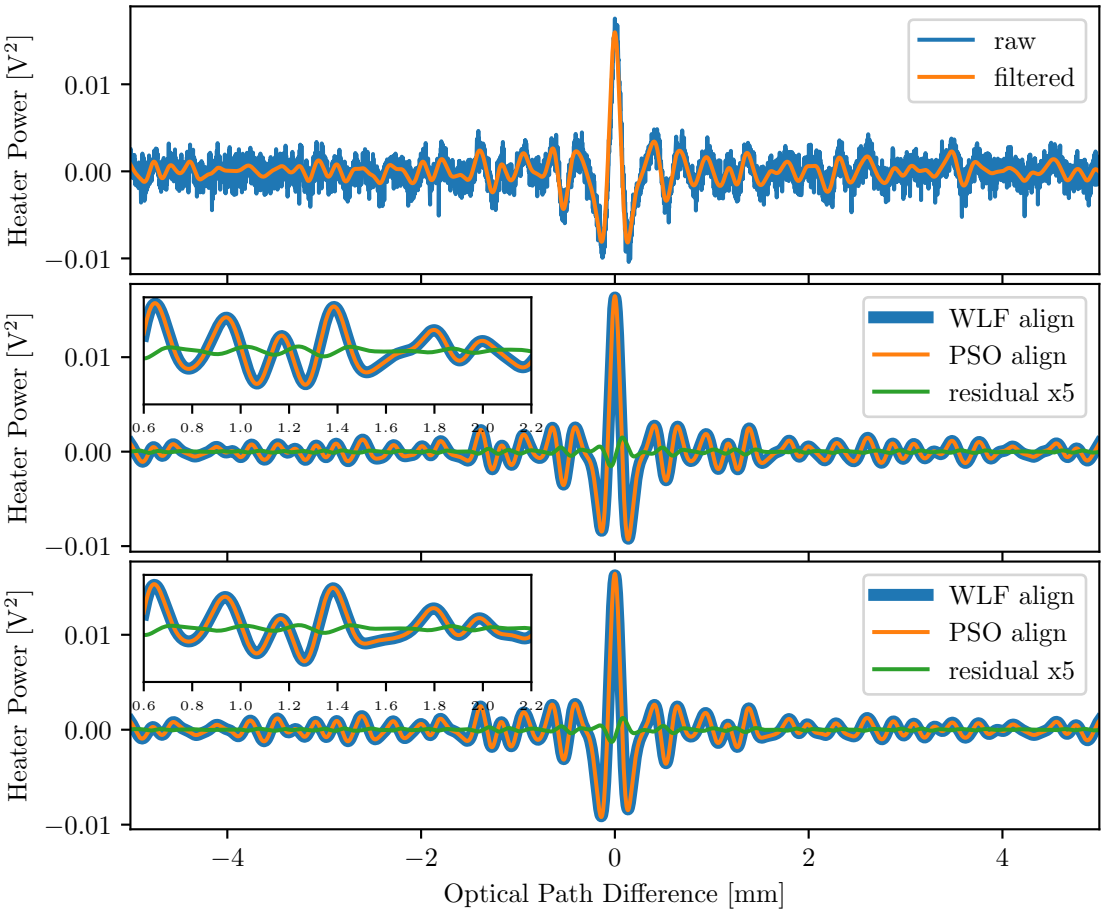


Figure 3.2: Comparison of averaging using external and internal metrologies. The top panel shows unaltered (blue) interferometer data that has been smoothed (orange) using a band-pass filter to remove frequency components outside of the optical band. The peak of the smoothed signal is used to align 500 interferograms resulting in an average interferogram shown in the lower panels (blue). Averaging the same interferograms, but aligned using PSO pulses based on the internal stage metrology, is also shown (orange) along with the residual between the two methods (green). The center and bottom panels show results for forward and reverse scans, respectively.

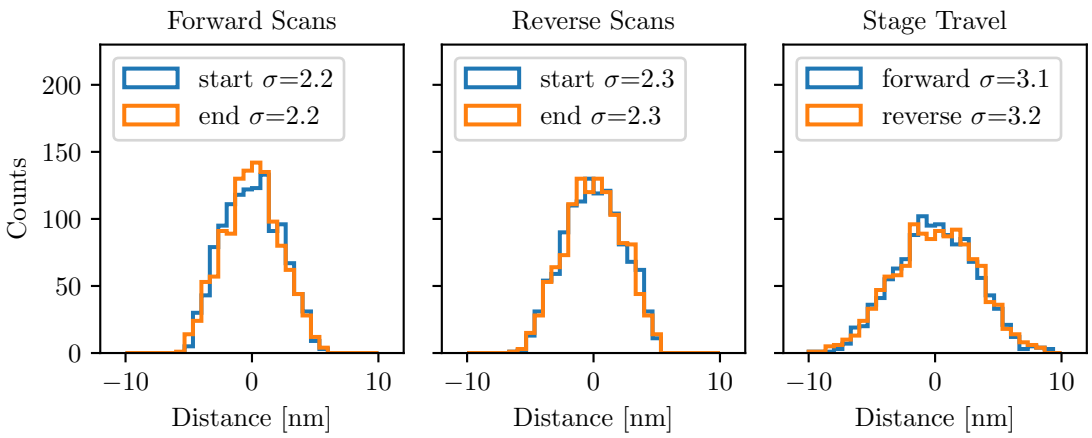


Figure 3.3: Distribution of stage positions and travel distance. The left and center panels show the mean subtracted distribution of start (blue) and end (orange) positions of forward and reverse scans, respectively. The right panel shows the mean subtracted distribution of the stage travel distance the forward (blue) and reverse (orange) directions. Values given are the standard deviations of associated distributions in nanometers.

are required before any signal becomes evident. Even if SNR was high, interferograms generally express complicated structure that changes with baseline. As such, it is generally impossible to align interferograms based on regular features present in the signal. Thus, high quality metrology is required in order to accurately align interferograms and average down noise effectively.

Using the PSO signal recorded on the detector digital channel, a speed profile of the stage motion can be constructed. The stage was commanded to travel at the typical speed of 6.7 mm/s and the controller was configured to send out a PSO pulse ever 15 μm of stage travel. By identifying the time difference between PSO pulses, the stage speed can be reconstructed. Reconstructed speed profiles for a forward and reverse scan are shown in Fig. 3.4. Deviations from the commanded speed at the start and end of each scan result from the acceleration of the stage. The dashed lines indicate the position of the 15th PSO pulse after the stage starts moving in either direction. After the 15th pulse, the stage speed reaches a constant speed and detector measurements are take at approximately equal OPD intervals. It is for this reason that the first 15 pulses are skipped when using the stage metrology based averaging strategy described above.

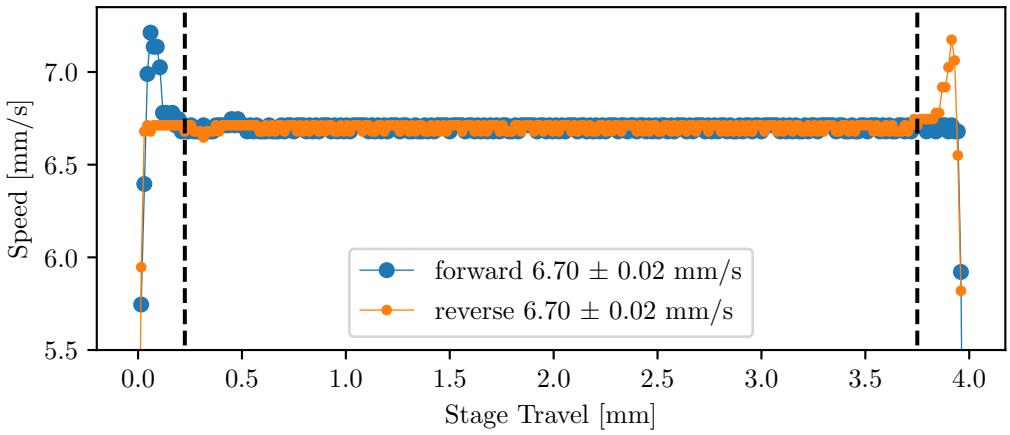


Figure 3.4: Stage speed profiles when commanded to travel at 6.7 mm/s. Forward (blue) and reverse (orange) directions profiles show constant speed after initial acceleration regions. The dashed lines indicate the 15th PSO pulse after the stage starts moving and is taken as the point at which the stage reaches constant speed. Average and standard deviations for data between the dashed lines are shown.

The mean and standard deviation of the data residing within the dashed lines are calculated giving 6.700 ± 0.017 mm/s. This uncertainty represents a relative error of 0.25%. The average OPD sampling interval under the conditions for this experiment is 2.075 μm, and using the same relative error, the OPD intervals are expected to have an associated random error of 5.2 nm. Variations resulting from random sampling error are well bellow the 3.4 μm requirement.

3.2.3 Beamsplitters

The beamsplitters used in this work are composed of metal mesh layers separated by dielectric spacers [104]. By combining inductive low-pass, and capacitive high-pass meshes, the beamsplitters achieve a 50:50 split between reflected and transmitted light intensity. The cut-off frequency for each kind of mesh is dependent on the mesh geometry. In this way, the beamsplitters can be tuned to operate over a broad bandwidth. The beamsplitters used in the interferometer are circular with a 175 mm diameter clear aperture.

The most important characteristics of a beamsplitter are its transmission and reflection spectra, and the associated phase properties of the transmitted and reflected light. The pho-

tomixer system, described in § 3.2.4, is used to measure the transmission spectrum. In the absence of a reflection measurement, it is assumed that the beamsplitters are approximately lossless,

$$R \approx 1 - T \ , \tag{3.1}$$

where R is the reflectance, and T is the transmittance of the beamsplitter. These quantities describe the fraction of radiant power reflected or transmitted, respectively. Beamsplitter efficiency is then expressed as,

$$\varepsilon_{bs} = 4RT \ , \tag{3.2}$$

so that in the ideal case $R = T = 0.5$ and $\varepsilon_{bs} = 1$. The calculated efficiency from the measured transmission spectrum is shown in Fig. 3.5. Beamsplitter efficiency is ~ 1.0 between $0.4 - 2.1\,\mathrm{THz}$ with a notable drop in efficiency at $\sim 1.7\,\mathrm{GHz}$. At frequencies $\geq 2.1\,\mathrm{THz}$, the efficiency drops nearly to zero before returning to appreciably high yet variable values. Although the beamsplitter seems functional at frequencies greater than $2.4\,\mathrm{THz}$, no interferometer signal is observable after $2.2\,\mathrm{THz}$. This lack of signal may be due to the invalidity of the lossless assumption at higher frequencies. However, significant attenuation of signal is observed after $\sim 1\,\mathrm{THz}$ and suggests the high-frequency cutoff of the interferometer system is due to reduced detector coupling efficiency at higher frequency. At the lower frequency edge, the beamsplitter efficiency reaches 0.5 at $\sim 0.25\,\mathrm{THz}$ which corresponds approximately to the frequency at which the first propagation mode of the detector waveguides becomes active.

The general phase properties of the beamsplitters are discussed in Ap. B.3 in detail. To summarize, transmission results in a constant phase shift. Depending on whether reflected light interacts with the inductive or capacitive side of the beamsplitter, the resulting phase shift is $\phi_L \approx +\pi/2$ or $\phi_C \approx -\pi/2$, respectively. The main utility of these quantities is to establish the symmetry properties of measured interferograms in FTS or DFI observation modes. In FTS mode, interferograms are approximately symmetric, while in DFI mode, in-

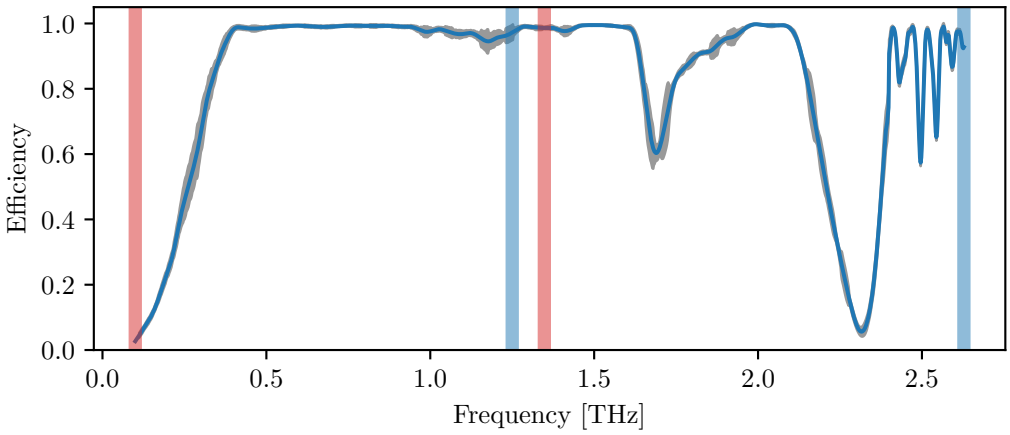


Figure 3.5: Measured beamsplitter efficiency. The red and blue bands indicate the edges of the frequency ranges obtained using different pairings of input lasers to the photomixer. Efficiently is near unity between $0.4 - 2.1\,\mathrm{THz}$ with a significant loss at $1.7\,\mathrm{THz}$. Variable but appreciably high efficiency is observed at frequencies greater than $2.4\,\mathrm{THz}$.

terferograms are approximately anti-symmetric when observing symmetric sources. These symmetries are generally preserved with some variation due to dispersion resulting from the beamsplitters and detector electronics. For FTS measurements, the phase is well compensated as a result of the instrument design using two beamsplitters, and due to the relatively high SNR measurements, the absolute value of the spectrum can be used and the phase need not be considered. The phase spectrum in DFI mode is not compensated and is highly significant for data analysis. A detailed discussion of the phase spectrum measurement and why it matters is presented in § 5.2.1, while the resulting phase spectrum is shown in Fig. 3.6.

3.2.4 Sources

A variety of sources are used to probe different aspects of the interferometer performance including coherent and incoherent thermal sources. These sources are summarized in Tab. 3.2 and are discussed in more detail in this section.

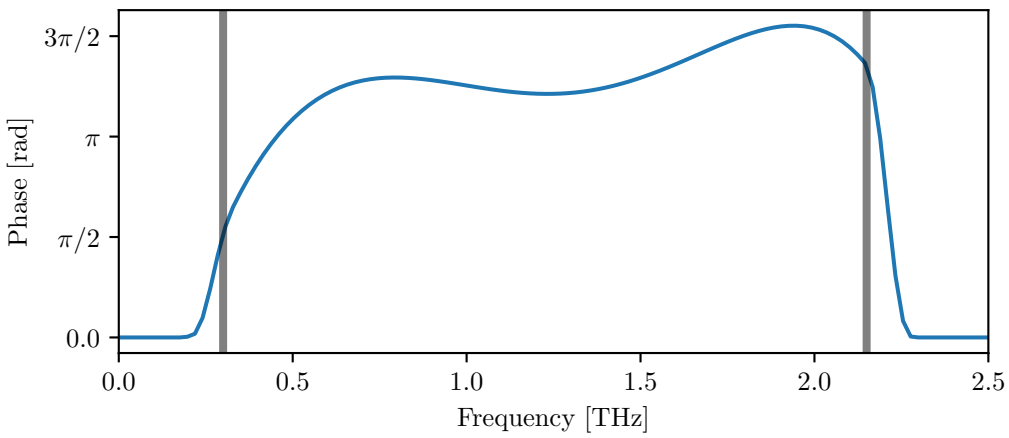


Figure 3.6: DFI mode phase spectrum. This figure shows the polynomial fit to the measured phase spectrum when operating in DFI mode. Horizontal lines show the bandwidth of the system.

Table 3.2: Instrument sources.

Name	Description
Pegasus	Thermal source with a 35 mm aperture and maximum temperature of $1,200\,^{\circ}$C.
Wide Field	Thermal source with a 90 mm aperture and maximum temperature of $100\,^{\circ}$C.
Photomixer	Tunable coherent source that emits radiation in the $0.04 - 3.30$ THz range.
Virginia Diodes	Tunable coherent source that emits high power ~ 16 mW, radiation in the $320 - 330$ GHz range.

Pegasus Thermal Source

The Pegasus model R source from Isotech produced thermal radiation with a broad spectrum approximated by Planck's law. The temperature of the source is stabilized by a PID controller and operates within the temperature range $150 - 1,200\,^{\circ}$C. Structurally, the source consists of a ceramic tube wound with nichrome wire and further surrounded by insulating material. Inside the tube resides a thermometer coupled to an iron thermal load 70 mm in length. When current is passed through the wire, the temperature of the wire increases and heat is conducted to the iron. Current through the wire is moderated by the PID controller using the thermometer reading as the process variable. Behind the iron is an insulating ceramic cylinder, and in front of the iron is 70 mm of length where a thick

ceramic tube can be inserted to improve thermalization. With the insert, the exit aperture is 20 mm in diameter, and without the insert, the aperture is 35 mm.

When making DFI observations, a mask is placed in front of the source aperture and at the focal point of the collimating mirror. The masks have various shapes cut out to allow light through which illuminates the collimating mirror. For the DFI technique to work, the source must be accurately modelled as a collection of incoherent point sources, and the experimental setup simulates this physical situation for the cutout shapes. However, the simulation is only valid when the surface of the backlighting thermal source, as viewed from the perspective of the collimating mirror is at a constant temperature. Although the iron load can be reasonably assumed to be a uniform temperature, the temperature of walls of the ceramic tube just in front of the iron are at a different temperature that decreases towards the exit aperture of the tube. This is to say, if some region of the collimating mirror sees light originating from the ceramic walls while another region of the mirror sees light originating from the iron load when looking through the same point on the mask, the DFI requirement begins to lose validity. To minimize this effect, the ceramic insert was removed. In this configuration, an on-axis mask point can be regarded as emitting a cone of uniform light with half angles $\theta_{1/2} = \pm 14^{\circ}$. As the mask point moves by an amount x in millimeters orthogonal to the instrument optical axis, the half angles change asymmetrically and are given by,

$$\theta_{1/2} = \arctan\left(\frac{\frac{35}{2} \pm x}{70}\right), \tag{3.3}$$

which is approximately linear for small x with a $-0.74\,^{\circ}\mathrm{mm}^{-1}$ slope. Given the angle that the collimating mirror subtends from its focal point, performance relating to uniform illumination begins to degrade when the mask point is shifted by $+3.2$ mm or -4.5 mm where the negative sign indicates the direction away from the detector (see Fig. 2.14). This range effectively defined the maximum extent of the mask features for optimal observing conditions. However, if the mask feature is not extended, the source can be re-positioned to accommodate for off-axis points further away from the optical axis than this range.

The Pegasus black body was the source of choice when attempting to perform imaging with the DFI system due to its broad spectrum, high intensity (at least compared to other available broadband sources), and the control over source intensity distribution using masks. Using the highest temperature possible would be preferable for measurements to maximize the SNR, however, the prolonged use of the source at its highest temperature lead to failure of the nichrome wire. As such, the source was most commonly used at $1,100°$C for long observations, but found use at other temperatures for measurements relating to detector system characterization.

Wide Field Thermal Source

A second thermal source was used in this work for wide field applications. It is a custom black body produced to study beamsplitter emission for the Herschel/SPIRE instrument [105], and consists of heating elements coated with a mixture of thermal epoxy, graphite, and carborundum grains. This coating results in high emissivity while reducing specular reflection. Coupled to the black body are two temperature sensors which feed into a PID controller that regulates the heating elements. The black body obtains a high level of temperature uniformity over a circular emission surface 90 mm in diameters, and can operate at temperatures between $30-100\,°$C. The relatively low maximum temperature of the black body makes it unsuitable for DFI observations. However, the large profile and thermal uniformity of the emission surface makes it ideal for simultaneous characterization of all detectors within the detector array.

TeraScan 1550 Photomixer

The TeraScan 1550 system is a commercial product developed by TOPTICA Photonics and is capable of producing temporally coherent radiation at moderate power over a large tunable range [106]. A diagram of the TeraScan system is shown in Fig. 3.7. Distributed feedback lasers are used to produce radiation at closely spaced wavelengths $\sim 1.5\,\mu$m, which are coupled to single mode polarization maintaining optical fibers. The fibers are then

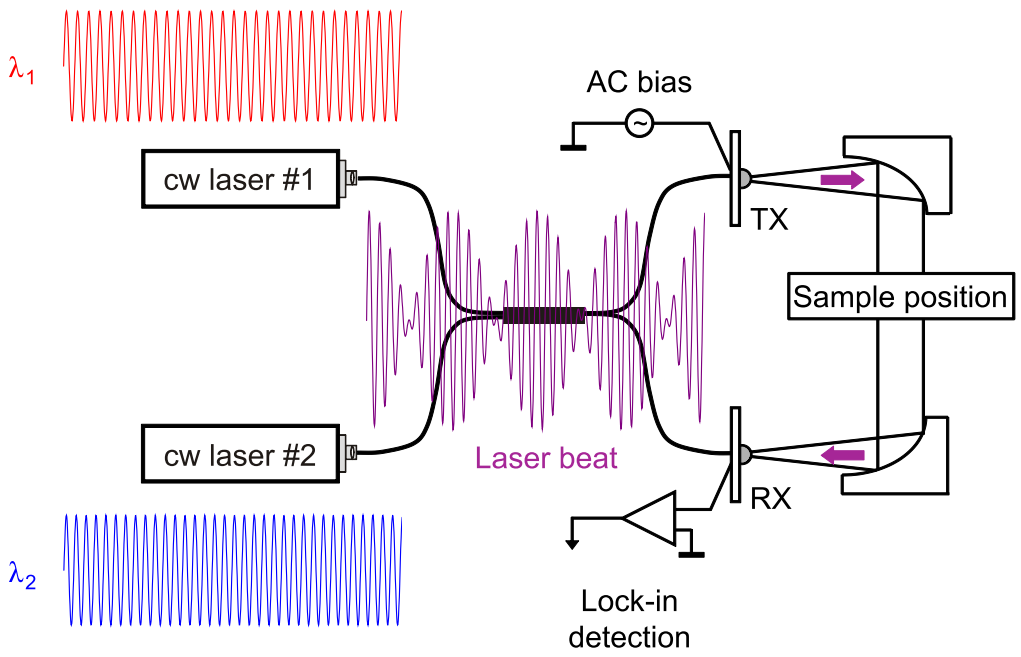


Figure 3.7: TeraScan 1550 photomixer diagram [106]. The system consists of two distributed feedback laser sources emitting radiation with closely spaced wavelengths coupled to single mode polarization maintaining optical fibers. Light is then combined then separated using a fiber optic splitter. The resulting beams are guided to the transmitter and receiver elements. The photomixers proper reside within the transmitter and is used to emit THz radiation at the beat frequency of the combined lasers. Image used with permission from TOPTICA Photonics.

brought to a fiber splitter, combining and then separating the electric fields, producing a high frequency carrier wave modulated at a beat frequency given by the sum and difference of the input radiation frequencies, respectively. The resulting electric fields are propagated to the transmitter and receiver elements.

Inside the transmitter is the photomixer containing an InGaAs photodiode emitter coupled to a log-spiral antenna. The photodiode consists of an intrinsic layer embedded between p-doped and n-doped semiconductors with a band gap of 1.68 μm. When light from the lasers is incident on the intrinsic layer, charge carriers are produced in a time varying fashion dependent on the modulating wave. By applying a bias voltage across the junction, a photocurrent results which oscillates at the beat frequency of the combined lasers. The moving charges are coupled to the antenna producing a THz wave.

The THz frequency can be adjusted by tuning the output frequencies of the lasers. Distributed feedback lasers use a semiconductor architecture where periodic variations in the semiconductor structure produce periodic variations in the index of refraction [107]. This periodic structure supports resonance at wavelengths that satisfy $\lambda = 2n_{eff}\Lambda$, where n_{eff} is the effective refractive index of the semiconductor and Λ is the pitch or wavelength of the periodic structure. As such, only one longitudinal mode is supported within a much longer active cavity providing a high level of mode stability. The pitch, and thus frequency of emitted radiation, can be adjusted through thermal expansion with precise control over the semiconductor temperature.

The TeraScan system comes with three lasers, and through the different combinations, the transmitter can emit light over $0.04 - 3.30\,\text{THz}$, with a line width of $\sim 1\,\text{MHz}$. Power output decreases with frequency and depends on the bias voltage. With default settings, THz power values of $100\,\mu\text{W}$ at $0.1\,\text{THz}$ and $10\,\mu\text{W}$ at $0.5\,\text{THz}$ are typical. In front of the antenna is a hemispherical lens 10 mm in diameter, with a 22 mm back focal distance, which produces a cone of light with a FWHM divergence angle of 15°.

On its own, the transmitter functions as a coherent THz source that can be used as the input to the interferometer system. The bias voltage can also be varied in some periodic way in order to electronically modulate the output power. When combined with the receiver element, the TeraScan system becomes an interferometer that is particularly useful for measuring transmission spectra. The receiver element is similar in structure to the transmitter element, and charge carriers are created within the InGaAs semiconductor intrinsic layer with a time variation dependent on the combined lasers beat frequency. However, instead of an external bias voltage, photocurrent is induced by a time varying voltage resulting from the THz wave interacting with the receiver antenna. Recall that the THz wave frequency results from the beat frequency in the transmitter which has a common origin with the beat frequency in the receiver. Photocurrent in the receiver is then a measure of the phase difference $\Delta\phi$ between the THz wave and the modulating wave in the receiver with

frequency ν. This situation can be modelled as,

$$I_{RX}(\nu) \propto E_{THz}\cos(\Delta\phi) = E_{THz}\cos(2\pi z\nu/c)\ , \tag{3.4}$$

where z is the OPD between the transmitter arm, including the THz path, and the receiver arm, both measured initially from the fiber splitter. Note that photocurrent is proportional to the amplitude of the THz wave electric field E_{THz}, not power. Eq. 3.4 is similar to the response on an FTS, however, in this case, the frequency is varied by tuning the lasers while the optical OPD remains constant. The receiver signal can be amplified by modulating the bias voltage of transmitter at some frequency, then processing the receiver signal using a lock-in amplifier.

True measures of the THz field amplitude, unmodified by phase, are given when $\Delta\phi = \pm n\pi$, which correspond to the extrema of the sinusoidal photocurrent response that are separated by $\Delta\nu = c/z$. As such, the quantity of interest is the envelope of the absolute value of photocurrent response. This function is typically obtained by averaging several datasets, followed by smoothing and peak identification. The result is usually a relatively smooth curve which has been sampled densely enough so that it can be interpolated onto whatever frequency axis that is convenient. To measure the transmission spectrum of a particular material, a photcurrent spectrum is measured with the sample placed in the path of the THz wave $I_{samp}(\nu)$, in addition to a background measurement with no sample $I_{bkg}(\nu)$. After interpolating both measurements onto the same frequency grid, which may be necessary due to phase shifts on transmission, the transmission spectrum is calculated as,

$$T(\nu) = \left(\frac{I_{samp}(\nu)}{I_{bkg}(\nu)}\right)^2\ . \tag{3.5}$$

Note this is a ratio of photocurrents proportional to electric field, and squaring the ratio gives transmitted power. By dividing by the background measurement, the calculation effectively removes absorption due to atmosphere or systematic errors.

Virginia Diodes

Commercial millimeter-wave transmitters produced by Virginia Diodes (VDI) are capable of achieving high output power radiation with narrow linewidths. These transmitters exploit a technique where microwave frequency voltage signals are frequency multiplied up to THz frequencies. In particular, the VDI Tx-148 accomplishes frequency multiplication using planar GaAs Schottky diode technology [108] in order to obtain $\sim 16\,\mathrm{mW}$ emission over $320 - 330\,\mathrm{GHz}$ with an advertised line width on the order of tens of Hz. The transmitter includes a diagonal horn antenna which produces an approximately Gaussian beam with a 10° FWHM. The VDI sources used in this work were provided by Dr. David Naylor in association with Blue Sky Spectroscopy.

3.2.5 Mirrors

The components previously discussed are commercial products that were selected to be of the highest quality available while staying within budget. Interferometer mirrors, however, were custom made and designed to meet observation objectives outlined earlier in this chapter. In particular, the mirror designs affect interferometer baseline coverage, magnification, beam shape, and spectral resolution. These characteristics also influence field of view and throughput. Inevitably, the optical design will result in some level of degraded performance for some region of the source plane due to the finite size and surface quality of the optics. Mirror dimensions, focal lengths, and surface quality were chosen such that these distortions would be minimized for the central detector.

All the mirrors used for the interferometer are constructed from Aluminum 6061 which is an alloy known to have high reflectivity at the wavelengths of interest [109]. Technical drawings for all the mirrors are presented in Ap. D.

Collimating Mirror

The collimating mirror is derived from a paraboloid surface as illustrated in Fig. 3.8. The three-dimensional paraboloid is a surface of revolution where all points satisfy the

equation,

$$z = \frac{r^2}{4f} \,, \tag{3.6}$$

where $r = \sqrt{x^2 + y^2}$ is the distance from the axis of symmetry, and f is the focus point of the paraboloid. To facilitate alignment, a 90° Off-Axis Parabolic (OAP) geometry was chosen such that the central ray would be justified to the grid of the optics table. As such, the nominal center of the mirror is where the slope of the paraboloid is one. Solving for this condition, the mirror center is at $(r,z) = (2f, f)$. A light ray originating from the focus point of the paraboloid must then travel

$$f_{eff} = 2f \,, \tag{3.7}$$

before intersecting the center of the mirror. To obtain the mirror surface, a plane tangent to the mirror center is constructed such that the center of the plane coincides with the center of the mirror. Projecting the plane onto the paraboloid surface, along the direction of the plane normal vector, produces the mirror surface. For the collimating mirror, the normal plane has dimensions 590×135 mm. From the perspective of an on-axis source, the collimating mirror subtends the angles $\sim 11.5°$ to $-10.5°$, referenced from the mirror center with positive counter-clockwise rotation. Projected onto the aperture plane, the collimating mirror is ~ 420 mm in in extent.

All the mirrors used in this work were manufactured by B-Con Engineering. The collimating mirror shows visible striations resulting from the machining process, which are most prominent about half way towards the edges as measured from the center of the mirror. This local surface roughness is expected to degrade performance and is baseline dependent. To estimate surface roughness, a 4D Technology PhaseCam 4030 was used, which employs a Twyman–Green interferometer to measures OPD variations with respect to a flat reference surface [8]. Measurements suggest the ridges on the mirror result in deviations ranging from 20 nm to 100 nm. A sample of these measurements is shown in Fig. 3.9. Typically

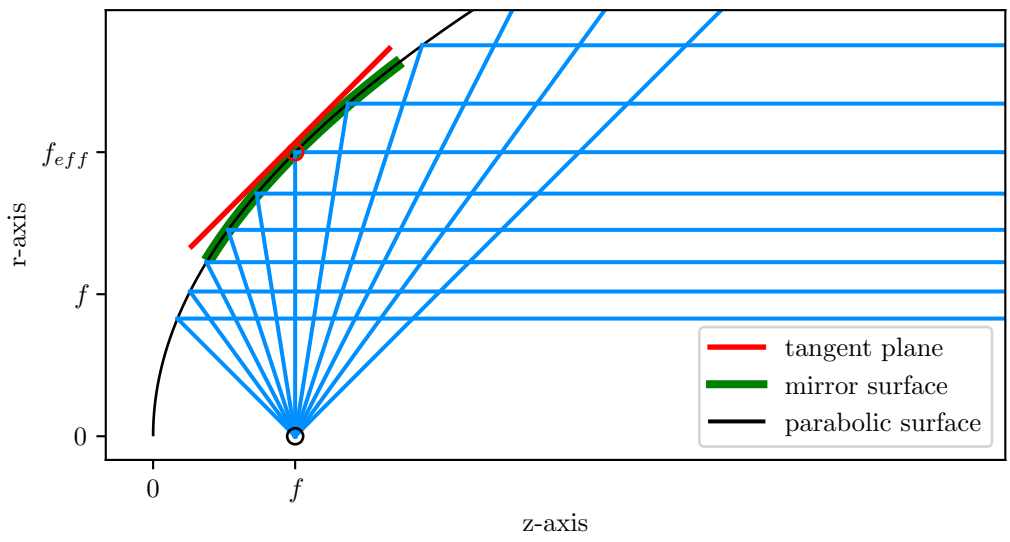


Figure 3.8: Parabolic mirror geometry. The black parabolic curve represents a cross section of a three dimensional paraboloid with focal length f. Light rays emitted from a point source at the focal point, represented by blue lines, are made parallel after reflection off the parabolic surface. A tangent plane, shown as a red line, is constructed and centered on the point of the parabola where the slope is one. The mirror surface in green results from projecting the plane, along the direction of the plane's normal vector, onto the paraboloid surface. The distance between the paraboloid focus point and the mirror center is the effective focal length of the mirror $f_{eff} = 2f$.

when assessing surface roughness and modeling the effects on instrument performance, the surface variations are assumed to be random. In this case, the surface variations express regular structure and such analysis is not valid. A more appropriate analysis would require detailed numerical simulations, which are not performed in this work. In the absence of such analysis, it is still encouraging that average surface deviations are an order of magnitude lower than required.

In addition to local surface deviations, the accuracy of the broader mirror curvature was measured. To accomplish this, a laser was placed on one of the translation stages with the laser beam incident on the collimated side of the mirror. The system was aligned such that the laser beam would propagate parallel to collimated light, and the reflected beam would be focused at the focal point of the mirror. A target was placed at the focal point and the position of the focused light on this target was recorded as the laser was

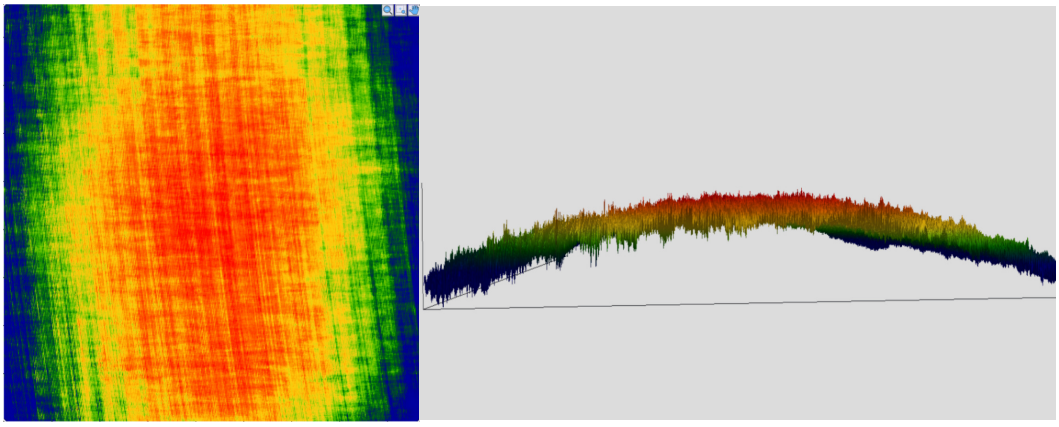


Figure 3.9: Collimating mirror surface deviations. These images were obtain using a phase camera which reconstructs the surface under observation using interferometric measurements combined with proprietary analysis software. This measurement was taken $\sim$40 mm away from the center of the mirror and shows the striations visible on the surface. The left panel shows an image of a 7×7 mm region of the mirror, while the right panel shows a two dimensional reconstruction of the same region. The color scale indicates deviations from a flat surface with a maximum value of $\sim$160 nm. Analysis of the Root Mean Squared (RMS) surface roughness and measurements of the peak to peak deviations suggest ridges are approximately 40 nm in height. The overall curvature is attributed to slight misalignment.

scanned across the mirror surface. Results for this experiment are shown in Fig. 3.10 and indicate that collimated light is focused to a roughly circular region with a 1 mm radius. These measurements also provide an estimate of the collimating mirror's effective focal length $f_{col} = 1198 \pm 2$ mm. Variation in the focus point as a function of aperture position will degrade instrument performance by introducing baseline dependent OPD offset errors. Measured intensity variations resulting from source plane coupling to the interferometer beams are also expected. Deviation from the ideal collimating mirror curvature results in the sensitivity of a detector to different points on the source plane varying as a function of aperture position.

Focusing Mirror

The focusing mirror takes collimated light at the end of the interferometer system and focuses it onto the detector feed horn plane. Geometric considerations for this mirror are similar to that of the focusing mirror Fig. 3.8. In this case the tangent plane is a circle with

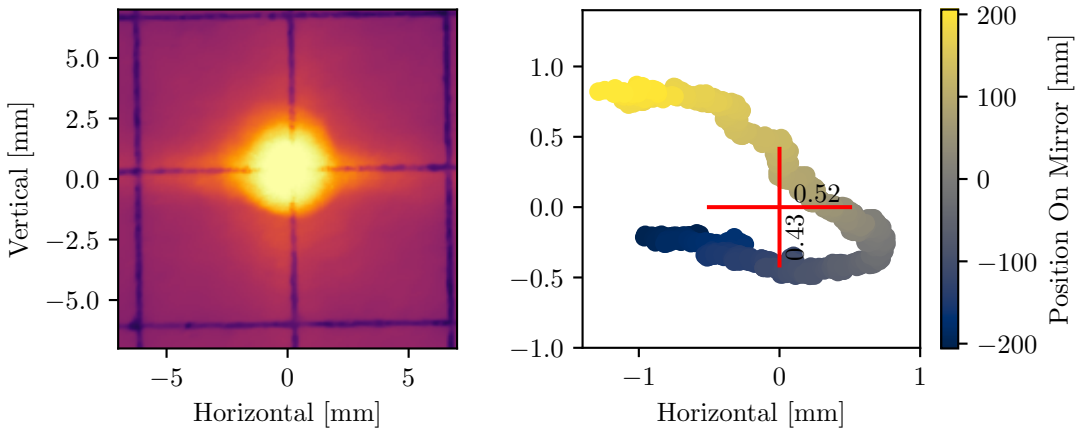


Figure 3.10: Collimating mirror curvature. A laser beam parallel to the optical axis was scanned over the collimated side of the mirror and its reflected position on the focal plane was recorded. The left panel shows the average of all recorded positions on the focal plane target and the right panel shows the position on the target as a function of mirror position. The red cross parameterizes the mean and standard deviation of the distribution of points.

a 135 mm diameter, and the paraboloid focal length is $f = 120$ mm.

Striations resulting from the machining process are evident on the mirror, however, in the active region of the mirror, the ridges are < 30 nm in height. More prominent ridges are located > 30 mm away from the center of the mirror, but the mirror is oversized and for the central detector only light incident on the inner 27 mm radius is efficiently coupled to the feed horns. The mirror curvature was also measured with results shown in Fig. 3.11. In this case, light is focused to a circular region ~ 0.5 mm in radius. However, much of this variation results from light incident with the mirror outside the active region. Light incident on the inner 30 mm radius is focused to a circular region 0.2 mm in radius. This spread is expected to give good performance since the target is the 3.63 mm square entrance apertures of the detector feed horns. With this experiment, the effective focal length of the focusing mirror was measured to be $f_{foc} = 240 \pm 2$ mm.

Flat Mirrors

Eleven flat mirrors are required for DFI observations with an additional mirror required for FTS observations. These mirrors are all identical circular flats 135 mm in diameter.

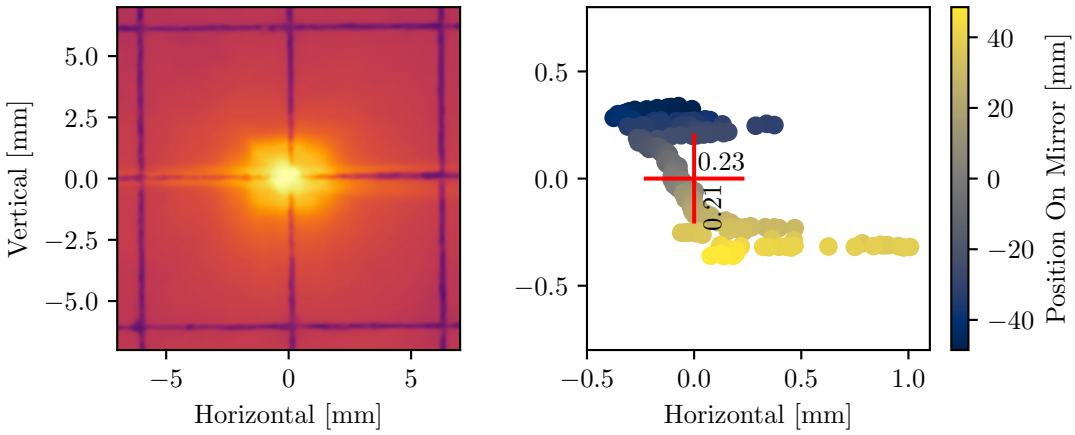


Figure 3.11: Focusing mirror curvature. A laser beam parallel to the optical axis was scanned over the collimated side of the mirror and its reflected position on the focal plane was recorded. The left panel shows the average of all recorded positions on the focal plane target and the right panel shows the position on the target as a function of mirror position. The red cross parameterizes the mean and standard deviation of the distribution of points.

Surface measurements were made for one test mirror and it was found to have a surface roughness RMS of 28 nm with no visible striations.

Although the flat mirrors are all circular, they are rotated by 45° around an axis parallel to the optics table normal vector. When considered as apertures that the interferometer beams pass through, they take the shape of ellipses with a 135 mm major axis in the vertical direction, and a $135/\sqrt{2} \approx 95$ mm minor axis in the horizontal direction.

Like the focusing mirror, but unlike the collimating mirror, the flat mirrors use a mounting strategy where their backs have three contact points oriented as a right triangle. These contacts rest against micrometers, which given the triangular geometry, provide precise control over the tip-tilt degrees of freedom, and allow for translation along the direction of the mirror plane normal vector. All mirrors are designed to be justified with the optics table grid, and the mounting apparatus are designed such that all optical components are aligned to sub-millimetre precision when in their default positions on the optics table.

3.3 Optical Design

In the section, the individual component parameters discussed in the previous section are used to estimate figures of merit for the interferometer system. In some cases, calculations are facilitated by a number of simplifying assumptions that will be highlighted when they are used.

3.3.1 Focusing Optics

The optical component that limits the light incident on the detector is the $f_{4.5}$ feed horn. To determine an estimate of the aperture stop, the acceptance cone of the feed horn is projected onto the focusing mirror. The aperture stop is then represented as a circle with diameter,

$$d_{AS} = \frac{f_{foc}}{f_{4.5}} \approx 53 \quad [\text{mm}] \tag{3.8}$$

centered on the optical axis and intersects the interferometer beam after the beamsplitter and just before the focusing mirror. The most important approximation used here is the assumption that the feed horn acceptance cone has a hard cutoff. Although the beam shape of the feed horns have yet to be measured, they are most likely approximately Gaussian with a gradual cutoff.

The solid angle associated with a cone with half angle $\theta_{1/2}$ is [110],

$$\Omega = 2\pi \left[1 - \cos(\theta_{1/2})\right] \ , \tag{3.9}$$

where the half angle can be expressed in terms of the f-number,

$$\theta_{1/2} = \arctan\left(\frac{1}{2f_{\#}}\right) \ . \tag{3.10}$$

Using the feed horn f-number, the solid angle for a detector is $\Omega_{det} \approx 38 \times 10^{-3}$ sr. Throughput describes the amount of light that propagates through an optical system. Using the paraxial approximation and assuming the radiating and receiving surfaces are parallel, the

system throughput can be calculated as [110],

$$S = A\Omega \,, \tag{3.11}$$

where A is the area of the radiating surface and Ω is the solid angle or the receiving surface as viewed by the radiating surface. Throughput is a conserved quantity in a well designed optical system, and the roles of the emitting and receiving surfaces can be interchanged. In this way, the throughput of a detector in the interferometer system is calculated as $A_{det} \times \Omega_{det} \approx 5.06 \times 10^{-7}\,\mathrm{m^2 sr}$.

3.3.2 Collimating Optics

Since the light travelling between the collimating and focusing mirrors is ideally parallel, the aperture stop can be projected to just in front of the collimating mirror unmodified. The collimating mirror in conjunction with the aperture stop results in an f-number $f_{\#} = f_{col}/d_{AS} = 22.5$. Using Eq. 3.9, the solid angle associated with the input aperture is $\Omega_{ap} \approx 1.55 \times 10^{-3}\,\mathrm{sr}$. Using conservation of throughput, this implies a square on the source plane with side lengths 18 mm, couples to the central detector.

Each arm of the interferometer can be modelled as double converging lens optical system where the first lens is synthesizes from the collimating mirror and flat aperture mirror, and the second lens is synthesized from the focusing mirror and aperture stop (see Fig. 3.12). This model is valid in the paraxial regime where sources are close to the optical axis and the angles at which light intersect the lenses are small.

To estimate magnification and FOV, the thin lens equation is used [8],

$$\frac{1}{s_o} + \frac{1}{s_i} = \frac{1}{f} \,, \tag{3.12}$$

where s_o and s_i are the distances to the source and image objects, respectively, measured from the center of the lens, and f is the focal length of the lens. When a source is located

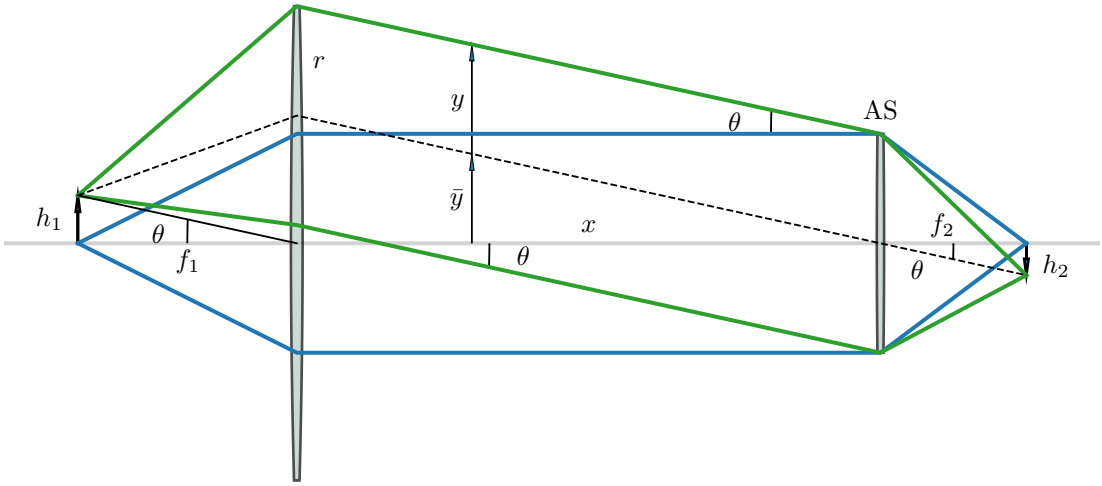


Figure 3.12: Thin lens model used to estimate optical performance. The collimating and entrance aperture mirrors are combined and represented as a converging lens. This lens has a focal length f_1 equal to the effective focal length of the collimating mirror, and has the same extent as the entrance aperture mirror which is length r taller than the aperture stop (AS). The AS is another converging lens that results from the acceptance cone of a detector feed horn projected onto the focusing mirror. This lens has a focal length f_2 equal to the effective focal length of the focusing mirror. An unvignetted source of height h_1, located on the focal plane of the first lens, is imaged onto the focal plane of the second lens with height h_2. Two bundles or rays are shown, one bundle for an on-axis point which trace marginal rays (blue), and one bundle for an off axis point (green). The chief ray is also shown with a dashed line where $\bar{y}$ is the height of the chief ray and y is the height of the off-axis ray bundle at any point along the optical axis. Angles θ are labeled to show that they are equivalent.

on the foal plane of a lens, $s_o = f$, the image point is at infinity, implying parallel rays. Additionally, the transverse magnification of a lens is given by [8],

$$m_T \equiv \frac{h_i}{h_o} = \frac{-s_i}{s_o} \,, \tag{3.13}$$

where h_i and h_o are the heights of the image and source objects, respectively. For a multi-lens system, the image point for the first lens is used as the source point for the next lens $s_{2o} = x - s_{1i}$, where x is the distance between lenses. The total transverse magnification for a multi-lens system is also obtained as the product of the magnifications of the individual lenses. With these considerations, the total transverse magnification of the double lens system that represents the interferometer is,

$$\begin{aligned} M_T &= \lim_{s_{1i}\to\infty} \frac{s_{1i}}{s_{1o}} \times \frac{s_{2i}}{s_{2o}} \\ &= \lim_{s_{1i}\to\infty} \frac{s_{1i}}{s_{1o}} \times \frac{s_{2i}}{x - s_{1i}} \\ &= \frac{-f_2}{f_1} \,, \end{aligned} \tag{3.14}$$

which when $f_{1/2}$ take the values of the effective focal length of the collimating and focusing mirrors, respectively, gives $M_T = -0.2$. With the magnification, the extent of objects on the image plane and the source plane are related by,

$$h_{img} = -h_{src} \frac{f_{foc}}{f_{col}} \tag{3.15}$$

and can be used to calculate the FOV of a single detector and the entire detector array. Using the detector extent parameters discussed in Sec. 3.2.1, the detector and detector array FOVs are $\mathrm{FOV}_{det} \approx 18\,\mathrm{mm}$ and $\mathrm{FOV}_{ary} \approx 98\,\mathrm{mm}$, respectively, and represent the dimensions of square regions on the source plane centered on the optical axis.

Perhaps more important than the total FOV is the unvignetted FOV. This represents the region of the source plane where light emitted from the points within this region are not

obstructed by any apertures in the optical system, with the exception of the aperture stop. In other words, any point source within the unvignetted FOV with the same intensity as an on-axis point source will appear just as bright to the detector as the on-axis source. Sources outside the unvignetted FOV will be gradually attenuated as they move further from the optical axis. Fig. 3.12 shows the geometry for an unvignetted source using the thin lens model for an arm of the interferometer. The blue lines represent the marginal rays which describe a bundle of rays emitted from an on-axis source point that pass through the aperture stop. The green lines represent a bundle of rays that originate from an off-axis point and also pass through the aperture stop. Also drawn as a black dashed line is the chief ray, a ray that by definition originates from an off-axis point and passes through the center of the aperture stop. An off-axis source is unvignetted if at every point along the optical axis the aperture radius is,

$$r_{ap} \geq |\bar{y}| + |y| \,, \tag{3.16}$$

where $\bar{y}$ and y are the heights of the chief ray and off-axis ray bundle, respectively. The aperture at any point along the optical axis can be regarded as some abstract region of space that light passes through, but more practically, it is the bounding region of some optical component that light passes through. The relevant apertures in this case are the first and seconds lenses, and from Fig. 3.12, it is clear that Eq. 3.16 is satisfied. In this model, an unvignetted source point a distance h_1 away from the optical axis requires that the first aperture be a height r greater than the apertures stop height. Using similar triangles, the relationship between these quantities is,

$$h_1 = \frac{f_1 r}{x} \,, \tag{3.17}$$

where f_1 is the focal length of the first lens, and x is the separation between lenses. Given the effective focal length of the collimating mirror, the difference in size between the aperture stop and flat mirror aperture, and the average distance between collimating and focusing

mirrors $\sim$4 m, the unvignetted FOV is estimated to be an ellipse with major and minor axis in the vertical and horizontal directions of lengths 24.0 mm and 12.4 mm, respectively. The average of these axes is 18.2 mm, which is the same as the nominal full FOV of a single detector.

The source plane must be defined in terms of angles in order to make the intensity distribution amenable to Fourier transformation. From Fig. 3.12, it is clear that the mapping between souce or detector plane spatial extent h_i, and angular extent θ, is,

$$\theta \approx \tan(\theta) = \frac{h_i}{f_i} \,. \tag{3.18}$$

In the context of true mirror based optical system, f_i represents the effective focal length of the OAP mirror and applies to both the collimating a focusing optics.

3.3.3 Resolution

As discussed in § 3.2.5, the projected extents of the collimating and flat aperture mirrors onto the aperture plane are 420 mm and 95 mm, respectively. With the constraint that the aperture mirrors must reside entirely within the projected collimating mirror region, the maximum separation between the centers aperture mirrors is $420 - 95 = 325$ mm. However, some space is left on either side of the apertures as a buffer and the maximum separation used in observations was 310 mm. Theoretically, the minimum separation is 95 mm, however, due to the depth of the mirrors and their associated mounting infrastructure, the minimum separation is 135 mm. These limits constrain the allowable baseline range for DFI measurements to $135 \leq B \leq 310$ mm. Minimum baseline increments are set by the positional accuracy of the translation stage, but in practice, the desired increments are determined by the observing objective and affect the extent of the reconstructed image in the same way that OPD steps affect bandwidth (see § 2.3.2).

One way to quantify the spatial resolution of an instrument is by using the Rayleigh criterion to describe the minimum separation between point sources where the sources can

still be distinguished (see § 2.5.1). Another metric that applies to spatial interferometry is the minimum source extent that that can be resolved. In this case, resolved means the visibility for fringes drops to zero at the maximum baseline of the interferometer. From the van Cittert-Zernike theorem (see Ap. A), the visibility is given by the absolute value of the normalized Fourier transform of the source intensity distribution. If the source is a square of angular width θ, then its associated visibility curve is the Fourier transform of the rectangle function with $a = \theta$ Eq. C.31,

$$\mathcal{V} = \mathrm{sinc}(\theta B/\lambda) \,, \tag{3.19}$$

where B is the interferometer baseline, and λ is the wavelength being observed. The spatial extent of θ on the source plane can be obtained using Eq. 3.18. When using the normalized sinc function, the first zero occurs when the argument is equal to one. With these considerations, the minimum source extent that can be resolved is,

$$h_{min} = \frac{\lambda f_{col}}{B_{max}} \,. \tag{3.20}$$

With a target resolution $h_{min} = 1.0\,\mathrm{mm}$, focal length $f_{col} = 1,200\,\mathrm{mm}$, and maximum baseline $B_{max} = 310\,\mathrm{mm}$, the source is resolved for $\lambda \leq 258\,\mu\mathrm{m} \rightarrow 1.16\,\mathrm{THz}$.

Spectral resolution is determined by the maximum OPD between the two beams of the interferometer Eq. 2.54. With a DFI system, the concept of ZPD is somewhat ambiguous since, at a particular baseline, a source that is off the optical axis by an angle θ will have its ZPD shifted by (see Fig. A.2),

$$\delta_z = B\sin(\theta) \,, \tag{3.21}$$

The term "phase center" is used to avoid this ambiguity and refers to ZPD for an on-axis point source. Since phase is such an important quantity with spatial interferometers, double-

sided interferograms are recorded such that the set of OPD samples are,

$$z \in \{-z_{max}, -z_{max} + \Delta z, ..., 0, ..., +z_{max} - \Delta z\} \ , \tag{3.22}$$

where Δz is the separation between OPD samples, and OPD is measured with respect to the phase center.

Increasing the baseline results in a decrease in the path length of the beam associated with the dynamic aperture (see Fig. 2.14). This means the position of the phase center on the spectral stage is baseline dependent. To compensate for the phase center shift, for every four units the baseline is increased, the spectral stage is moved up by one unit. This factor of four results from the folding factor associated with the double-rooftop mirror design on the spectral stage. Accounting for these effects, the maximum symmetric OPD that can be achieved consistently at all baselines is,

$$z_{max} = 4L_{max} = 4 \times \frac{L_{stage} - \frac{B_{max} - B_{min}}{4}}{2} \approx 800 \quad [\text{mm}] \ , \tag{3.23}$$

where L_{max} is the maximum symmetric mechanical stage travel, and $L_{stage} = 450\,\text{mm}$ is the total travel distance of the translation stage. With this maximum OPD, with respect to the phase center, the best achievable spectral resolution is $\Delta\nu \approx 0.18\,\text{GHz}$. In practice, however, averaging thousands of interferograms is required to obtain good SNR. To confine observations to a reasonable amount of time, scan ranges were significantly shortened for DFI observations resulting in a more typical spectral resolution of 25 GHz.

Maximum OPD also imposes restrictions on the instrument FOV through Eq. 3.21. If a source is too far off axis, its ZPD will be shifted outside of the sampled interferogram, and this shift is proportional to baseline. If a point source is defined to be within the FOV when its white light fringe is sampled by the interferogram, then OPD limits the FOV to,

$$\text{FOV}_{OPD} = \frac{z_{max} f_{col}}{B_{max}} \ , \tag{3.24}$$

and this quantity is defined with spatial units on the source plane.

3.4 Optical Characterization

In the previous section, estimates for the interferometer performance were derived using the paraxial approximation. This formalism is convenient for its simple analytical form, and produces good estimates for conjugate points close to the optical axis. However, the instrument is configured to operate over a wide FOV where the paraxial approximation loses validity. This section is dedicated to characterization of the interferometer in the regime where the paraxial approximation fails. Note that the horizontal axis of the source plane is inverted compared to the more conventional orientation. This is done to keep results consistent with the sign convention used when deriving the DFI response equation (see Ap. B, and Fig. B.1 in particular).

3.4.1 Ray Tracing

One way to assess the performance of an optical system is to use ray tracing. Ray tracing describes a method for calculating the path light takes as it originates from some point and propagates through the optical system. In these calculations, light is treated as a ray which propagates in some direction characterized by a unit vector, and travels in a straight line when propagating in a medium with constant index of refraction. Rays may also be defined with properties like wavelength, amplitude, and polarization, which are appropriate for quantifying electromagnetic waves. As the ray interacts with optical components in the system, their properties are modified according to physical laws. In particular, the law of reflection, Snell's law, and Fresnel equations [111], constitute the core of any physical ray tracing algorithm. Since this analysis is based on physical interactions with optical surfaces, it is suitable for light originating from any point in space, overcoming the limitations of paraxial theory.

To facilitate this analysis, analytic ray tracing software was written in the Python pro-

gramming language [112], with support from the third-party packages: NumPy [113], SciPy [114], Astropy [115], and Matplotlib [116]. The software leverages the equations of geometric optics to calculate interaction between rays, described as vectors, and surfaces modelled as planes and conic surfaces. Since the rays and surfaces are analytic, the paths rays take as they propagate through an optical system can be modelled with a high degree of numerical precision. The software accommodates reflective and refractive interactions while tracking quantities like wavelength, amplitude, path length, optical path length, and phase associated with rays. The software was not designed to accommodate stochastic deviations of ray interactions resulting from surface roughness statistics.

A ray trace of the interferometer system in a typical state is shown in Fig. 3.13. The only dispersive optic in the system is the beamsplitter. Since its boundaries with air are parallel and the beamsplitter membrane is thin, the wavelength dependent properties of the beamplitter have little effect on the trace. As such, the trace is wavelength independent. In this figure, rays are emitted from points on the source plane that are conjugate to the center of detectors on the detector plane, as calculated using the paraxial approximation. The rays are propagated through the optical system and terminated on the detector plane. Only rays that successfully propagate though the optical system and terminate on the detector plane are shown. As can be seen from the figure, source points further from the optical axis have emission cones that are constrained to smaller angles than the on-axis point. This result is an expression of the vignetting associated with off-axis source points.

On the detector plane, it is observed that the on-axis source point focuses perfectly on the central detector in accordance with the paraxial approximation. Additionally, the image is inverted, as expected. However, for more distant source points, the rays do not achieve a proper focus, and the nominal focus points are offset from the expected detector centers. This offset appear to be asymmetric with a dilation on the left side and a compression on the right side.

A more complete picture of the rays in the detector plane is shown in Fig. 3.14. In

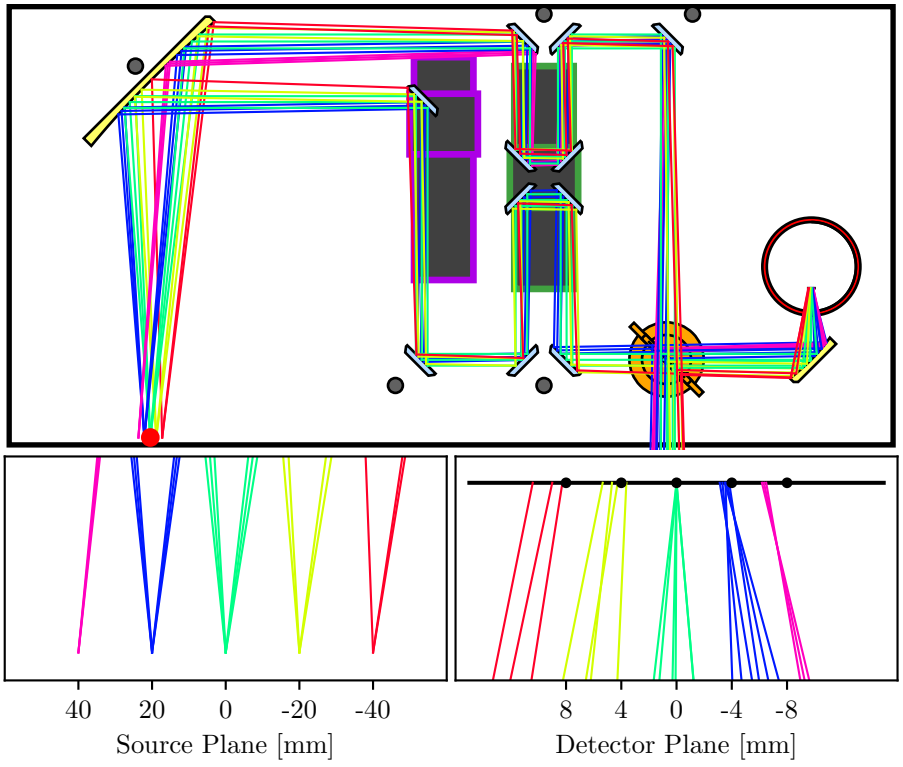


Figure 3.13: Interferometer ray trace. Bundles of rays originating from positions on the source plane, that coincide with the projected positions of detector centers, are propagated through the interferometer system to the detector plane. The top panel shows the entire system while the bottom panels show close up views of the source plane (left) and detector plane (right). The off-axis source points are not perfectly imaged onto the detector plane, nor do their effective focus points coincide with the detector centers, indicated by black circles in the bottom right panel. Vignetting is also present and no ray vector could be found for the far left source point that would propagate through the lower arm of the detector system. These distortions become more significant with increasing distance from the optical axis.

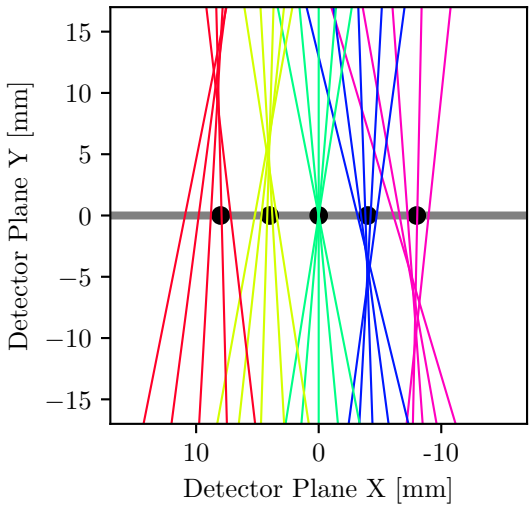


Figure 3.14: Aberrations resulting from the parabolic mirrors. This trace is similar to Fig. 3.13, but was contrived to be free of any vignetting, and demonstrates exclusively the distortions associated with the collimating and focusing mirrors. The presence of coma, and approximately parabolic field curvature, are evident. Black circles correspond to centers of detectors.

this case the system is contrived to prevent vignetting by the flat mirrors. This figure clearly demonstrates the deviations from the paraxial approximation resulting from the use of parabolic collimating and focusing optics. Such deviations are more generally referred to as aberrations [8]. The traces show the characteristic signature of coma, and the heights of the nominal focus points across the detector plane indicates the presence of field curvature which is approximately parabolic. This two-dimensional trace does not give information along the vertical axis of the optics system. However, it is expected that similar effects occur along the vertical axis, in that degradation increases symmetrically away from the on-axis point. As such, astigmatism is expected for source points that are unequally distant from the horizontal and vertical axes.

3.4.2 Interferometer Beam

In assessing the optical properties of the instrument, diffraction effects associated with the aperture geometry were not considered. However, diffraction effects are important for determining what a detector sees on the source plane. The input apertures are structurally

circular mirrors rotated by a 45° angle. The resulting elliptical shape does not give an accurate representation of how the electric field couples to the detector system since coupling is strongly influenced by feed horn characteristics. In order to determine the effective aperture geometry, it is imagined that the detector is replaced by a transmitter and its emission is reverse propagated through the optical system to aperture plane.

Consider the situation where one beam of the DFI is blocked. After replacing the detector with a transmitter, the electric field passes through a rectangular waveguide, which is then coupled to an approximately pyramidal feed horn. The feed horn projects the electric field out of the detector system as a spherical wave with an approximately Gaussian amplitude distribution. After incidence with the focusing mirror, the field is truncated at the edges of the mirror and converted to a plane wave that is propagated to the aperture plane.

The f-number of the feed horn $f_{\#}$, is assumed to define the FWHM of the Gaussian amplitude distribution. The spherical wave emitted from the feed horn travels the distance of the focusing mirror's effective focal length f_{foc}, before being converted to a plane wave and propagating to the aperture plane. As a plane wave, the Gaussian amplitude distribution has FWHM $= f_{foc}/f_{\#}$, and can be written as,

$$A(\vec{r}) = T(\vec{r})e^{-\alpha|\vec{r}|^2}\,, \tag{3.25}$$

where $\alpha = (2\lambda f_{\#}/f_{foc})^2 \ln(2)$, and $\vec{r}$ are Cartesian coordinates defined in the aperture plane. The elliptical truncation function is,

$$T(\vec{r}) = \begin{cases} 1 & \text{if } \left(\frac{r_x}{a}\right)^2 + \left(\frac{r_y}{b}\right)^2 \leq 1 \\ 0 & \text{if } \left(\frac{r_x}{a}\right)^2 + \left(\frac{r_y}{b}\right)^2 > 1 \end{cases} \tag{3.26}$$

where the semi-axes are $b = D/2$ and $a = b/\sqrt{2}$, with D being the circular mirror diameter. The aperture plane electric field distribution is shown in Fig. 3.15.

Converting the aperture plane field distribution $A(\vec{r})$, to source plane distribution is ac-

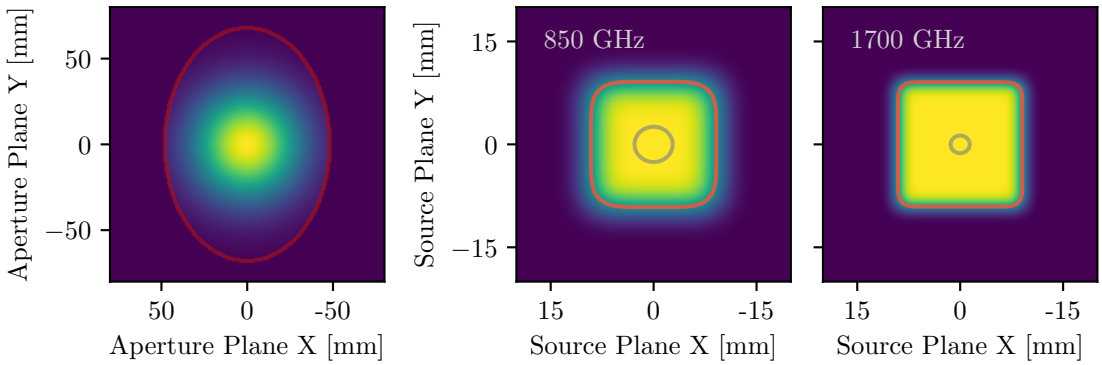


Figure 3.15: Aperture illumination and beam models. The left panels shows the frequency independent electric field amplitude distribution on the aperture plane resulting from the Gaussian beam of the feed horn. The red curve truncates the distribution and corresponds to the elliptical boundary of the rotated flat mirror. The center and right panels show the detector beam models at 850 GHz and 1.7 THz, on the source plane, respectively. The gray curve shows the 50% contour resulting from squaring the Fourier transform of the field distribution in the left panel, representing the Fraunhofer diffraction pattern associated with the aperture illumination. This pattern is then convolved with a square on the source plane which is the projection of feed horn aperture. The red curve shows the 50% contour of the synthesized beam.

complished using a Fourier transform [117],

$$\hat{A}(\vec{\theta}) = \int_{\mathcal{A}} A(\vec{r}) e^{-i\kappa\vec{\theta}\cdot\vec{r}}\, d\vec{r} \approx e^{-(\pi\vec{\theta})^2/\alpha}\,, \tag{3.27}$$

where $\vec{\theta}$ are source plane angular coordinates (see Fig. A.2), and the integral is taken over the aperture plane $\mathcal{A}$. The right-hand side is obtained as the Fourier transform of the Gaussian component of Eq. 3.25 (see Ap. C), and the approximate equality results from neglecting the truncation function. Since truncation occurs where the field amplitude is low, its effect on the Fourier transform is negligible (see § 2.4.6).

Now consider the situation where both beams of the DFI are allowed to propagate to the source plane. The resulting intensity pattern is $I(\vec{\theta}) = \hat{A}_1(\vec{\theta})\hat{A}_2^*(\vec{\theta})$, where subscripts denote the electric field pattern produced by the individual beams. Due to the symmetry in both beams of the DFI, $\hat{A}_1(\vec{\theta}) \approx \hat{A}_2(\vec{\theta})$, when the system is well aligned. The normalized source

plane intensity pattern is then,

$$\left|\hat{A}(\vec{\theta})\right|^2 \approx \exp\left[-\left(\frac{\pi\vec{\theta}f_{foc}}{\lambda f_{\#}}\right)^2 \frac{1}{2\ln(2)}\right] . \tag{3.28}$$

This equation pertains to a point source located at the centre of the feed horn plane. However, the feed horn aperture is rectangular with finite extent. To account for this, imagine a collection of point sources evenly distributed over the feed horn plane. On the source plane, one would observe the sum of Eq. 3.28 over these points, with angular offsets $\delta\vec{\theta} = -\vec{\theta}_f$, where $\vec{\theta}_f$ is the angular position of an individual point source on the feed horn plane. This process is functionally equivalent to convolving Eq. 3.28 with the projection of the feed horn aperture onto the source plane. The source plane sensitivity, or beam, of a particular detector is then,

$$\xi(\vec{\theta}) = \left|\hat{A}(\vec{\theta})\right|^2 * rect\left(\frac{\vec{\theta}-\vec{\theta}_{det}}{\theta_h}\right) , \tag{3.29}$$

where $rect[(\vec{\theta}-\vec{\theta}_{det})/\theta_h]$ is a two-dimensional rectangle function of width $\theta_h = 52\,\text{arcmin}$, centered on the projected detector position $\vec{\theta}_{det}$, and represents the projection of the rectangular feed horn aperture onto the source plane. A visual representation of the result of this process is shown in Fig. 3.15.

Reverse propagating the square of the electric field, from the detector to the source plane, produces a model of the interferometer beam that is equivalent to the source plane sensitivity of the detector. For the DFI, this pattern is approximately independent of the optical path lengths of the two arms of the interferometer but is affected by vignetting. There is a modest frequency dependence on the beam shape over the bandwidth of the instrument. At lower frequencies, the corners of the 50% contour describing the beam shape are more rounded than at higher frequencies, however, the width and height of the rectangular beam shape are essentially constant at 18 mm. Note that this modeling assumes the Gaussian amplitude distribution resulting from the feed horn is frequency independent.

3.4.3 Field of View Measurements

To empirically validate the preceding analysis, the FOV of the detector array is mapped using the photomixer as a source. Data is acquired by setting the photomixer to a particular optical frequency and electronically modulating the output power by varying the bias voltage sinusoidally with a frequency of 10 Hz. The photomixer is then raster scanned over the detector plane using 5 mm steps in the vertical and horizontal directions. At a particular optical frequency and photomixer position, the power incident on each detector is quantified by extracting the amplitude of the sinusoidal signal resulting from the modulated source. This amplitude is extracted by taking the Fourier transform of the signal and fitting the resulting feature in the frequency domain with a sinc function. The amplitude of the sinc function gives the amplitude of the sinusoid. A demonstration of this process is presented in § 4.3.3. This approach was chosen as it is less affected by noise in low SNR signals than peak-to-peak identification. By combining the results from different source positions, a sensitivity map over the entire source plane is generated for each detector. Since this observation is carried out at multiple optical frequencies, a frequency dependent sensitivity map is also obtained. However, maps were limited to a maximum frequency of 1050 GHz, since the power output of the photomixer decreases rapidly with frequency, and the detector sensitivity begins to decrease after ~ 1 THz. The results for all detectors can be co-added producing a sensitivity map for the entire detector array. Sensitivity maps obtained in this way are shown in Fig. 3.16 for the static aperture at multiple optical frequencies, and at multiple aperture positions with a constant 850 GHz optical frequency. Contours are added to the figure representing the 50% level of the sensitivity.

One application of these maps is to clarify the positions and orientations of each detector within the detector array. These results suggest the alignment of the mirrors can be improved in that the array sensitivity seem to be centered below and to the left of the A0 detector, and with ideal alignment, it should be centered on the A0 detector. It also seems as though the point of maximum sensitivity varies as a function of baseline, but is largely con-

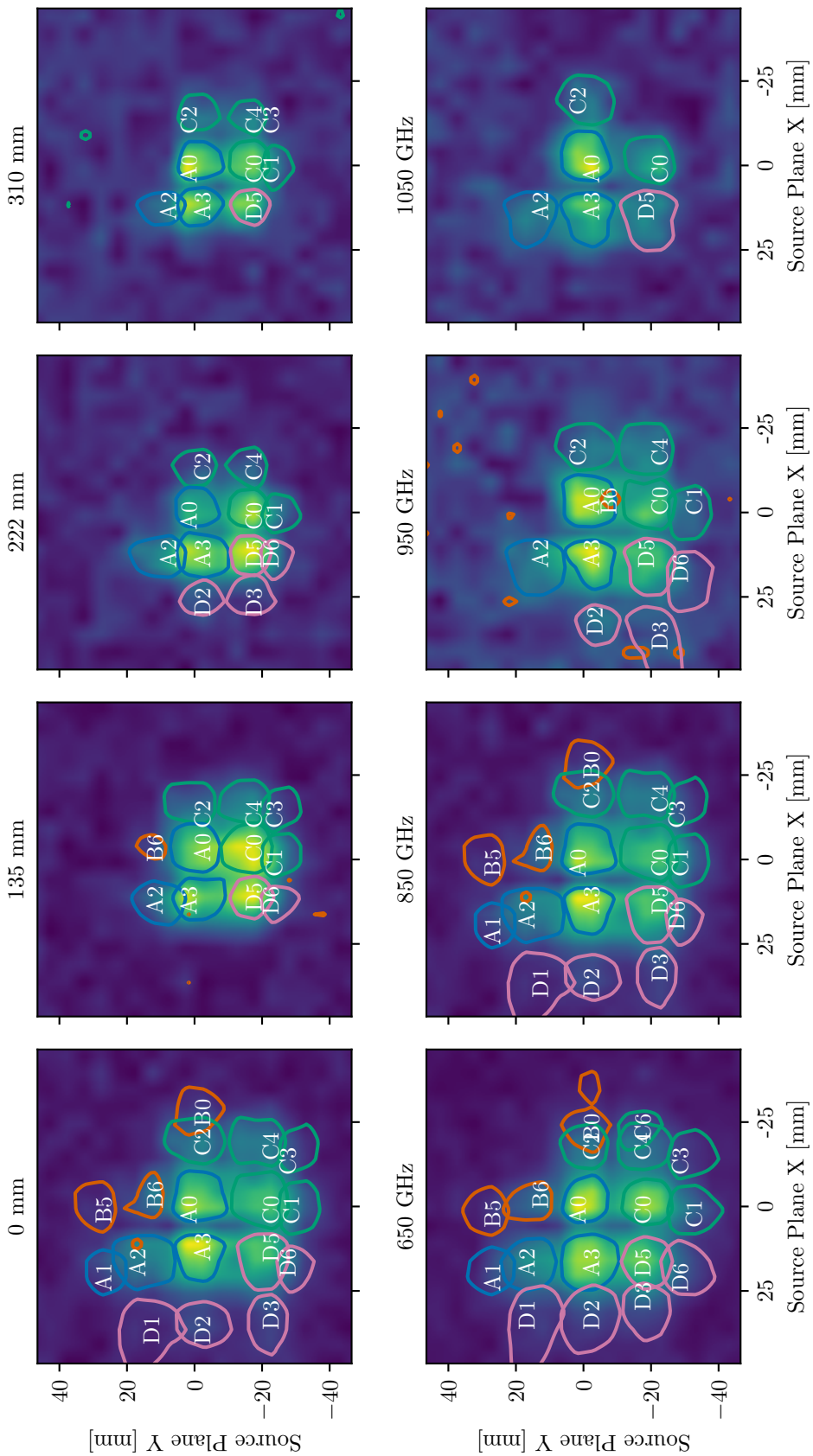


Figure 3.16: Interferometer field of view. This figure shows sensitivity maps obtained by raster scanning the photomixer across the source plane in 5 mm increments, then interpolated onto a finer grid. The sensitivity measurements for the detectors were co-added to show the overall sensitivity of the detector array. 50% contours are added to show the individual detector beams and detector labels are positioned to be centered on their respective point of maximum sensitivity. Only detectors whose maximum sensitivity values was above a particular threshold are included. The top row shows maps for different aperture positions at a constant 850 GHz optical frequency. The bottom row shows maps resulting from the static aperture using different optical frequencies.

stant as a function of optical frequency. This effect can be explained by the curvature of the collimating and focusing mirror as shown in Figs. 3.10 and 3.11. Furthermore, only a few detectors in the full array are suitable for interferometer measurements. In particular, the detectors with beams indicated in the lower right panel of Fig. 3.16 contain detectable signal over all the sampled aperture positions and optical frequencies. With improved alignment, it is expected that the central nine detectors should be usable.

It is observed that the detector beam positions on the left side of the source plane are spread out, while beam positions on the right are compressed. This behaviour is expected from the ray trace in Fig. 3.13. The ray trace shows uniformly distributed source points where source points on the left get compressed and source points on the right get dilated, when imaged on the detector plane. It is then expected that going in the reverse direction, from uniformly distributed points on the detector plane to the source plane, would produce the opposite effect, as observed in the sensitivity maps. Additionally, there appears to be a general compression of detector beams further away from the optical axis along the vertical direction.

Although the unvignetted region of the source plane is difficult to quantify from these results, an estimate from visual inspection gives $18 \pm 5\,\text{mm}$. More easily quantified is the beam size of each detector beam, represented by the 50% contour shown in the figure, and the separation between beams. For A0, the nominally on-axis beam, the width is $14.5 \pm 1.5\,\text{mm}$. The position of the A3, A0, and C0 beams, are respectively, $-13.9 \pm 2\text{mm}$, $1.5 \pm 1.5\text{mm}$, and $18.9 \pm 1.3\text{mm}$, which gives an average separation of $16.4\,\text{mm}$. The full FOV of the detector array is approximately square with a width of $75 \pm 5\,\text{mm}$. All the beam properties, including widths, separations, and full FOV, deviate from expected values by roughly -20%.

All the A0 contours from Fig. 3.16 are re-plotted in Fig. 3.17 for closer inspection. The left panel shows beams at various optical frequencies with respect to the static aperture. As expected, the beams are roughly rectangular and largely independent of frequency with

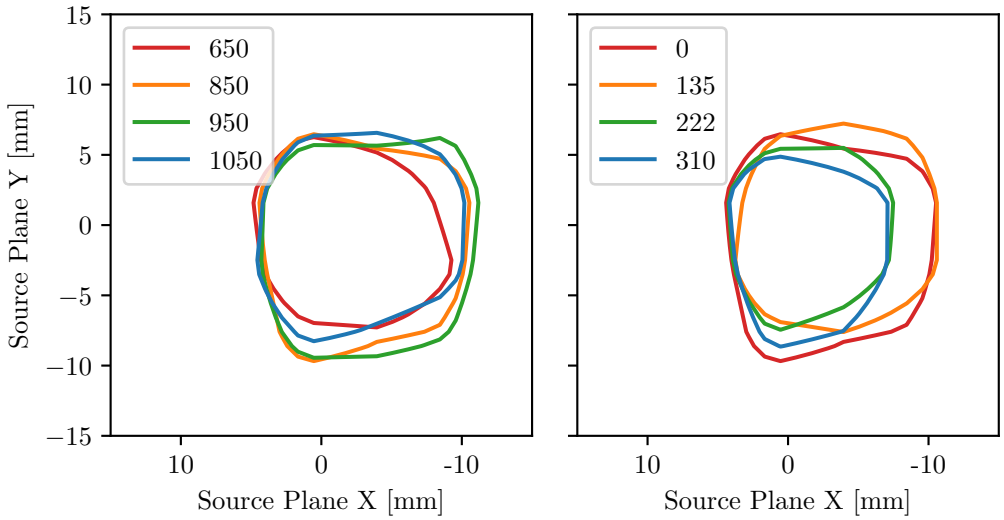


Figure 3.17: Field of view of the A0 detector. These curves are the 50% contours of the A0 detector taken from Fig. 3.16. The left panel shows the A0 beam at multiple optical frequencies using the static aperture, and are all similar with the 650 GHz beam being somewhat more circular. The right panel shows beams resulting from different aperture positions using a constant 850 GHz optical frequency. These beams get somewhat smaller with increasing aperture position, likely resulting from a bias in the photomixer alignment. Apart from a 20% reduction in size, the beam shapes seem to be in good agreement with beam models Fig. 3.15.

the lowest frequency beam expression having somewhat more rounded corners. The left panel shows beams resulting from different aperture positions at a constant 850 GHz. These results suggest the beam gets smaller with greater aperture position and this is attributed to vignetting or misalignment.

Other than the overall 20% reduction in beam extent and separation, the sensitivity mapping experiment is in good agreement with the optical modelling discussed in previous sections. A significant uncertainty in the experiment is physical characteristics of the photomixer. The back focal length of the transmitter lens is not well defined, and although care was taken to obtain best alignment, the focal point of the lens may not have coincided with the focal plane of the collimating mirror. The direction of the photomixer beam is also uncertain, and it is possible the beam direction favoured the far side of the collimating mirror, explaining why the sensitivity map using the static aperture produced the best result while aperture positions further away from the static aperture gradually degraded performance.

To address these concerns, a followup experiment was conduced where a 1 mm slit, back lit by the Pegasus thermal source, was moved across the source plane in 2 mm increments. Light passing through the slit was modulated by a mechanical chopper, and the analysis proceeded in the same way as the previous experiment. This version of the experiment benefits from having a well defined source structure that is known to be point-like with respect to the detector beam along the horizontal axis. Additionally, there is a higher degree of confidence that the slit coincided with the focal plane of the collimating mirror. However, this experiment required manual intervention, taking more time, and had a limited range in the slit positions. As such, these measurements were confined to central detector and fewer baselines. Results for this experiment, conducted at two different dynamic aperture positions, are shown in Fig. 3.18. The curves show similar sensitivity profiles at the two aperture positions, and are in good agreement with beam width estimates using the photomixer experiment. In this case, the FWHM is 13.0 ± 0.5mm. Both experiments suggest a 20% reduction from expected values in beam widths, separations, and full FOV. One factor which explains this discrepancy is that the feed horn array provided by the manufacturer is 20% smaller than indicated by the technical drawings in Ap. D. This hypothesis is plausible as there are often differences between design and fabrication dimensions for custom components.

3.5 Summary

In this chapter, the design and characterization of the interferometer instrument was discussed in detail. The individual components of the system were described and their figures of merit were calculated and measured when possible, as summarized in Tab. 3.1. Each component was found to be within acceptable performance limits with the collimating mirror contributing most significantly to deviations from ideal performance. Given the component parameters, figures of merit were calculated including the spatial and spectral resolution of the interferometer, and optical properties including magnification, vignetting,

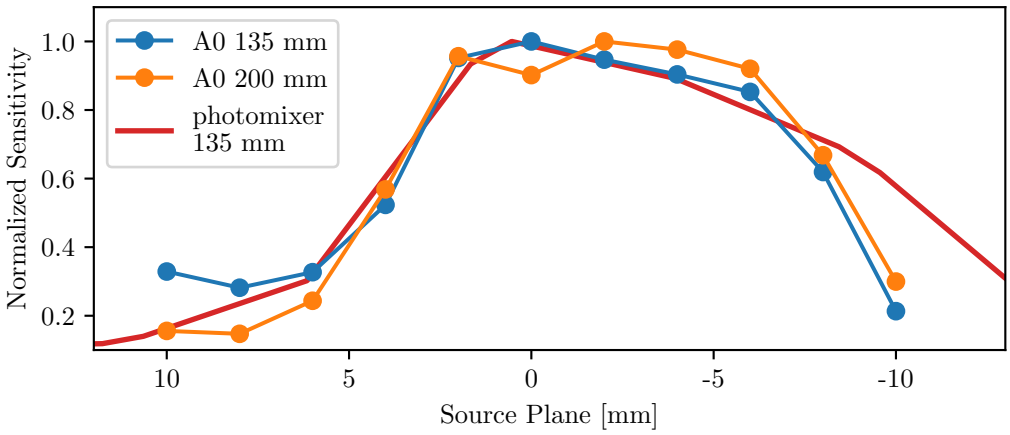


Figure 3.18: Field of view measurements obtained using a 1 mm slit point-like source. The slit was moved across the source plane in 2 mm increments to produced these sensitivity curves for the A0 detector at two aperture positions. Both curves are similar in shape and largely agree with the photomixer measurements.

and FOV. The properties were estimated using the paraxial approximation, which were further refined using ray tracing simulations to study aberrations, and beam modelling analysis to study diffraction effects. These optical properties were then verified through detector sensitivity measurement that largely confirmed the validity of the models, with the exception of an overall 20% reduction in expected beam dimensions.

The interferometer system meets its design goals of being able to resolve a 1 mm source at ~ 1.2 THz, and the primary beam of the detectors are $> 10\times$ larger than this resolution. The optical design has been well optimized for the central detector in the detector array, and interferometric measurements are expected to be possible for the central and a few of the surrounding detectors.

Chapter 4

Detector System

In this chapter, the operating principles and characterization of the detector system used in this work are presented. The detectors used in any optical system are of critical importance when discussing the performance capabilities of the system. The detectors used in this work belong to the category of Transition Edge Sensor (TES) bolometers. When radiant energy is incident on the detectors, that energy is converted to heat which increases the temperature of the TES element. The TES element is a pure niobium filament, the resistance of which changes rapidly as a function of temperature in the region where the filament transitions from a normal conducting state to a super conducting state. As such, when operated within this transition region, the TES is a sensitive thermometer capable of measuring small changes in optical power.

Although TES bolometers have been studied extensively [16, 118], the operational principle of the detectors used in this work is significantly different from conventional TES detectors. Conventional TES are voltage biased and maintain a particular operating temperature primarily through negative electrothermal feedback. As such, the sensor is kept near thermal equilibrium with typically small temperature deviations when subjected to variable optical power. The TES used in this work are current biased, resulting in positive electrothermal feedback, and are at risk of thermal runaway and detector instability. Instead of relying on electrothermal feedback, the thermal state is moderated primarily by a heater wire coupled to the TES and regulated by a Proportional-Integral–Derivative (PID) controller operating at $\sim 13\,\mathrm{kHz}$. Throughout the duration of a detector readout cycle, the TES

is never in thermal equilibrium and temperature variations are relatively large and rapid. These conditions explicitly violate the assumptions made when deriving standard bolometer theory, so the detector system used in this work merits empirical characterization given the lack of literature on the subject. Although operation of the detectors used in this work is more complicated with lower sensitivity than conventional TES bolometers, these detector have some advantages when used in a laboratory setting. The detector parameters are highly configurable and can be optimized for different observing conditions and the PID-based control loop prevents the detectors from saturating when exposed to the laboratory environment.

Direct application of standard bolometer theory to the detectors used in this work is potentially misleading. However, the concepts of balancing power entering and exiting the system, how this energy budget affects the temperature of the TES, the basic components of a bolometer, and understanding the primary sources of noise for such detectors all remain relevant topics. As such, Sec. 4.1 is dedicated to a review of standard bolometer theory. In Sec. 4.2, an overview of the detector system is presented along with a discussion pertaining to the operating principles and readout cycle of the detectors. Experimental characterization of the central detector in the detector array, the one primarily used in this work, is presented in Sec. 4.3.

4.1 Bolometers

4.1.1 Response

A bolometer consists of an absorbing element that is coupled to a thermometer, and these thermally active components are weakly linked to a heat sink with temperature T_s. In some cases, the absorber and thermometer are mounted on a common substrate. The combination of the substrate, absorber, and thermometer has a bulk heat capacity C and are mutually connected with high thermal conductivity such that they can always be assumed to be in thermal equilibrium at temperature T. To investigate the properties of a bolometer

under optical loading, suppose that radiant energy with intensity

$$P_{optical} = P_0 + \delta P e^{i2\pi ft} \quad [\mathrm{W}] \ , \tag{4.1}$$

is incident on the absorber, which includes a constant term P_0, and a harmonic term with amplitude δP, that varies in time t, with frequency f. This power influences the temperature of the bolometer according to

$$T = T_0 + \delta T e^{i2\pi ft} \quad [\mathrm{K}] \ . \tag{4.2}$$

In many cases, the thermometer component of the bolometer is realized by a resistor which varies with temperature $R(T)$. In such a case, the resistor is integrated into an electrical bias circuit that can be simplified into a Thevenin equivalent circuit consisting of a voltage source V_b, and load resistor R_L, in series with the thermometer resistor. A voltage biased resistor is more common, but in this work the resistor is current biased by setting $R_L \gg R(T)$ so that variations in resistance do to a changes in temperature have little effect on the bias current $I_b \approx V_b/R_L$. The load resistor is decoupled from the thermal circuit so that the relevant electrical heating includes only the thermometer resistor. Using the current biased configuration, and expanding the temperature dependence of the resistor to first order, the electrical power dissipated is

$$P_{electrical} = I_b^2 R(T) = I_b^2 \left[R(T_0) + \frac{dR}{dT} \delta T e^{i2\pi ft} \right] \ . \tag{4.3}$$

Heat leaves the thermally active part of the bolometer to the thermal sink through a weak thermal link. The power removed is given by

$$P_{conductance} = \int_{T_s}^{T} G(T)\, dT = \overline{G}(T_0 - T_s) + G\delta T e^{i2\pi ft} \tag{4.4}$$

where $G = dP/dT = \kappa A/l$ is the thermal conductance, which can be related to the physical properties of the link described by its cross-sectional area A, length l, and thermal conductivity κ. The average thermal conductance between the steady state temperature T_0 and the temperature of the thermal sink T_s is denoted by $\overline{G}$. Energy can also be exchanged with the thermally active parts of the bolometer through the bulk heat capacity at a rate given by

$$P_{capacitance} = \frac{dT}{dt} C = i2\pi f \delta T e^{i2\pi ft} C \,, \tag{4.5}$$

where the time derivative is calculated from Eq. 4.2.

The thermal balance of the bolomter is given by $P_{opt} + P_{el} = P_{cond} + P_{cap}$, which is expanded to

$$P_0 + \delta P e^{i2\pi ft} + I_b^2 R(T_0) + I_b^2 \frac{dR}{dT} \delta T e^{i2\pi ft} = \overline{G}(T_0 - T_s) + G\delta T e^{i2\pi ft} + i2\pi f \delta T e^{i2\pi ft} C \,. \tag{4.6}$$

Equating the steady state terms give

$$P_0 + I_b^2 R(T_0) = \overline{G}(T_0 - T_s) \,, \tag{4.7}$$

and equating the time dependent terms resulting from variations in optical loading gives

$$\frac{\delta P}{\delta T} = G + i2\pi f C - I_b^2 \frac{dR}{dT} \quad \left[\frac{\mathrm{W}}{\mathrm{K}}\right] \,. \tag{4.8}$$

With a current biased bolometer, temperature is inferred from the measured voltage drop across the thermometer resistor. The voltage responsivity describes the change in voltage for a change in incident optical power, $S_V = \delta V/\delta P = I_b(dR/dT)(\delta T/\delta P)$, and can be obtained from Eq. 4.8 as

$$S_V = \frac{I_b \frac{dR}{dT}}{G + i2\pi f C - I_b^2 \frac{dR}{dT}} \quad \left[\frac{\mathrm{V}}{\mathrm{W}}\right] \tag{4.9}$$

It is convenient to introduce an effective thermal conductance $G_{eff} = G - I_b^2 \frac{dR}{dT}$. The second term is included to account for how a change in thermometer resistance changes the power dissipated in the bolometer through Joule heating, subsequently influencing temperature. The sensitivity of the bolometer to temperature variations is parameterized using $\alpha = (dR/dT)/R$ which is conventionally evaluated at the steady state temperature, but is generally temperature dependent. The effective thermal conductance is then modified to $G_{eff} = G - I_b^2 R\alpha$. Furthermore, an effective thermal time constant $\tau_{eff} = C/G_{eff}$ can be defined which describes how quickly the bolometer responds to variations in optical loading. With these quantities, Eq. 4.9 can be written as

$$S_V = \frac{I_b R\alpha}{G_{eff}} \frac{1}{1 + i2\pi f\tau_{eff}} , \tag{4.10}$$

which has the form of a first order low-pass filter transfer function.

In the current biased bolometer configuration, if $\alpha > 0$, it is possible for $G_{eff} = 0$, in which case Eq. 4.10 diverges. This situation corresponds to thermal runaway resulting from positive electrothermal feedback. Since Joule heating is given by $I_b^2 R$ and resistance increases with temperature, $\alpha > 0$, increasing temperature results in even more Joule heating. This positive feedback cycle results in thermal runaway if the thermometer resistor generates more heat than can be conducted to the heat sink.

4.1.2 Noise

Photon Noise

Photon noise describes the signal variations resulting from the statistical properties of photon arrival rates from the source being observed. Thermal sources are common experimental and astronomical targets, so photon noise is derived with respect to thermal radiation. The spectral density associated with a thermal source is described by Planck's law,

$$B(\nu, T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/k_B T} - 1} \quad \left[\frac{\mathrm{W}}{\mathrm{m}^2 \cdot \mathrm{Hz} \cdot \mathrm{sr}}\right] , \tag{4.11}$$

where h is the Planck constant, k_B is the Boltzmann constant, c is the speed of light in the medium, ν is the frequency of light, and T is the temperature of the source.

From the antenna theorem, single moded throughput of a diffraction limited optical system is given by [117]

$$(A\Omega)_0 = \lambda^2 = \frac{c^2}{\nu^2} \quad \left[\mathrm{m}^2 \cdot \mathrm{sr}\right] , \tag{4.12}$$

where λ is the wavelength of light. Associated with each photon is energy

$$E_{phot} = h\nu \quad [\mathrm{J}] . \tag{4.13}$$

Multiplying Eq. 4.11 by Eq. 4.12, dividing by Eq. 4.13, and dividing by a factor of two (accounting for orthogonal polarization states), gives the number of photons in each vibration mode (i.e., frequency ν),

$$n = \frac{1}{e^{h\nu/k_BT} - 1} \quad \left[\frac{1}{\mathrm{s} \cdot \mathrm{Hz}}\right] , \tag{4.14}$$

which can be identified as a special case of the familiar Bose-Einstein distribution for photons called the Planck distribution.

Associated with the Planck distribution is a mean variance [119]

$$\left\langle (\delta n)^2 \right\rangle = n + n^2 . \tag{4.15}$$

With respect to thermal radiation, this represents the variance in the number of photons arriving each second within a particular mode. When $h\nu/k_BT \gg 1$, the variance in the arrival rate of photons is $\left\langle (\delta n)^2 \right\rangle \approx n$ and is well described by Poisson statistics. In this regime, the statistical properties of radiation are similar to particles and is associated with shot noise. In the opposing limit, when $n \gg 1$, the square term dominates and photons arrive in clusters emphasizing the wave-like nature of light.

The mean squared fluctuation in energy arriving each second is then

$$P_N^2 = \int_{\nu_{min}}^{\nu_{max}} (h\nu)^2 2N(n+n^2)\, d\nu \,, \tag{4.16}$$

where $2N = 2A\Omega\nu^2/c^2$ is the number of modes accounting for two polarization states (see Eq. 4.12). From the Nyquist-Shannon sampling theorem (see § 2.3.1), integration over one second is the sampling rate required to unambiguously determine signal in a $1/2\,\mathrm{Hz}$ bandwidth Δf. Normalizing the mean squared noise power Eq. 4.16, by bandwidth gives,

$$\frac{P_N^2}{\Delta f} = 2\int_{\nu_{min}}^{\nu_{max}} h\nu P(\nu,T)\, d\nu + \frac{1}{A\Omega}\int_{\nu_{min}}^{\nu_{max}} \frac{c^2}{q\nu^2} P^2(\nu,T)\, d\nu \,, \tag{4.17}$$

where $P(\nu,T) = 2Nnh\nu$ is the spectral power absorbed by the bolometer. The factor $q = A\Omega\Delta\nu\Delta t\nu^2/c^2$ represent the number of modes in one polarization state with frequency ν and detected in the bandwidth $\Delta\nu$ during a time Δt. Although q does not follow from the derivation, it is often added to satisfy the central limit theorem where the fluctuations from many modes should approach a Gaussian distribution [120].

The bandwidth normalized mean squared noise power resulting from an optical system viewing black body emission is presented in Eq. 4.17. As such, an equivalent bandwidth normalized mean squared signal power is required to produce an Signal to Noise Ratio (SNR) = 1. From the definition of Noise Equivalent Power (NEP) (see § 4.3.2), it can be concluded that Eq. 4.17 gives the square of the photon noise NEP. It is necessary to introduce quantities which describe the optical system including the source emissivity ε, the transmissivity of the optical system from source to detector $\mathcal{T}$, and the absorption efficiency of the detector η. This equation is often written using the unitless quantity $x = h\nu/k_BT$ as

$$\mathrm{NEP}_{photon}^2 = \frac{4(k_BT)^5}{c^2h^3}\frac{A\Omega\varepsilon\mathcal{T}}{\eta}\left(\int_{x_{min}}^{x_{max}} \frac{x^4\,dx}{e^x-1} + \frac{\varepsilon\mathcal{T}\eta}{q}\int_{x_{min}}^{x_{max}} \frac{x^4\,dx}{(e^x-1)^2}\right) \quad \left[\frac{\mathrm{W}^2}{\mathrm{Hz}}\right] \,, \tag{4.18}$$

and describes NEP referenced to the input of the instrument.

Johnson–Nyquist Noise

Johnson–Nyquist noise describes fluctuations of electric potential resulting from thermal variations within electrical components. In particular, fluctuations within the thermometer resistor produces noise in the bolometer voltage signal. If motion of electrons within the resistor are modelled as harmonic oscillations, then noise can be attributed to random phases associated with these oscillations. By the equipartition theorem, each quadratic degree of freedom receives energy $k_B T/2$ when in thermal equilibrium [119]. For a single vibrational mode, there are two quadratic degrees of freedom, resulting in $k_B T$ energy per mode, with the total power over all modes given by $k_B T \Delta f$.

To understand how this noise power is dissipated in an external load, it is conventional to represent the noisy resistor as a noise free resistor R in series with a voltage noise source $V_N = \sqrt{4 k_B T R \Delta f}$. The external load has resistance $R_L = R$ and is connected in series with the noise circuit with a lossless transmission line for maximum power transfer. The power dissipated in the load by the noise voltage is then $I_N^2 R_L = V_N^2 R_L/(R+R_L)^2 = k_B T \Delta f$. This is the noise power that would be measured by an impedance matched instrument connected to the noisy resistor, and is equivalent to the noise power obtained in the previous paragraph.

Voltage noise is converted to power noise through the voltage responsivity V_N/S_V (see Eq. 4.10). As such, the bandwidth normalized mean square power noise is

$$\mathrm{NEP}^2_{J/N} = \frac{V_N^2}{\Delta f \, |S_V|^2} = \frac{4 k_B T R}{|S_V|^2} \,, \tag{4.19}$$

which is the square of the Johnson–Nyquist noise NEP.

Thermal Fluctuation Noise

When the bolometer is in thermal equilibrium, energy will be transferred to or away from the thermally active part through the random interactions of electrons, photons, or phonons (quantized lattice vibrations). Although thermal fluctuations are facilitated by various mechanism, the noise they produce is typically called phonon noise. The mean

variance of thermal fluctuations is [121]

$$\left\langle (\delta T)^2 \right\rangle = \frac{k_B T^2}{C} = \int_0^\infty B_T(f)\, df \ , \tag{4.20}$$

where $B_T(f)$ is the temperature spectral density. Ignoring electrothermal feedback, this quantity can be related to the power spectral density B_P, as $B_T(f) = B_P/[G^2 + (2\pi f C)^2]$ using Eq. 4.8. In this case the power fluctuations are treated as white noise and are independent of frequency. The power spectral density can be understood as the bandwidth normalized mean variance of power fluctuations and can be identified with a squared NEP. Using the substitution in Eq. 4.20 and solving for B_P gives

$$\mathrm{NEP}^2_{phonon} = 4 G k_B T^2 \ . \tag{4.21}$$

When multiple thermal elements at different temperatures are connected, Eq. 4.21 is evaluated as a sum over these elements.

Excess Noise

The noise contributions discussed above are well understood, however, there are poorly understood noise sources as well. Notoriously, there is low frequency noise with spectral dependence $1/f^a$ present in most electronic systems that has not been rigorously explained. For a bolometer, some of this noise may be explained by low frequency fluctuations in the bias current or temperature variations of the heat sink. A common noise component also results from amplifier electronics, however, this is usually quite low for a well designed system. In TES bolometers specifically, some excess noise can be explained by internal thermal fluctuations within the TES filament itself instead of between the thermally active part of the bolometer and the heat sink. Noise with the same spectral dependence as Johnson-Nyquist noise has been observed which tends to increase with lower normal resistance and when biased lower on the TES transition curve. This noise is also strongly

correlation with the bolometer sensitivity parameter α. Attempts have been made to explain this noise in terms of magnetic field interactions [122], and the fluctuations of normal and superconducting regions within the TES filament [16].

4.2 Operating Principle

The detector system used in this work was manufactured by QMC Instruments, and due to the proprietary nature of the unconventional design, some technical details about the system remain unknown. Perhaps the most important missing information for the purpose of detailed characterization is knowledge of the electrical circuit. An effort was made to map the circuit design from the exposed electronics, but this exercise was insufficient to fully render the circuit. Although this limits the ability to infer some physical characteristics of the detector, there is still much to discuss in terms of detector operation. Additionally, performance with respect to most important figures of merit can still be measured without reference to the circuit.

4.2.1 Overview

The detector system consists of a 5×5 array of feed horn coupled TES bolometers (see Fig. 4.1). Detectors are read out using a time division multiplexing scheme where as many as seven detectors sequentially share time on a single channel with a total of four channels. Each detector is designated using a letter and a number (e.g., A0), where the letter indicates the channel and the number indicates the order of detectors on the channel. Only four detectors share the A channel, and it should be noted that the D2 and D3 detectors are unstable. Various metadata are output from the detector system including entries related to detector readout timing, the state of the detector system's external digital channel, and a running frame count. A "frame" describes the information acquired during a full iteration of the multiplexing cycle.

Mapping detector sensitivity on the source plane of the Double-Fourier Interferometer

(DFI) system was discussed in § 3.4.3, and results were shown in Fig. 3.16. With reference to these results, the orientation of detectors in situ can be inferred and is presented in the top left panel of Fig. 4.1. The top right panel shows the physical detector array and the bottom left panel shows the feed horn array. Technical drawing for these two components are shown in Ap. D. Each horn has an f-number 4.5, and a square entrance aperture with dimensions 3.63 mm. The f-number of the horns defines a cone with an acceptance angle of $\sim 12.7^{\circ}$. The bottom right panel shows a close up view of one of the TES bolometers. The absorber of the detector is made of gold. One of the roles of the feed horn coupling device is to convert the impedance of free space at the entrance of the horn to the impedance of the absorber at the exit of the horn. When the impedance is matched, absorption efficiency is maximized [123]. Two concentric circles are observed on the detector which are joined to wires that leave the detector. One of these circles corresponds to the TES element, a pure niobium filament that transitions between normal and superconducting states and serves as a sensitive thermometer. The other circle is a heater wire that is used to regulate the temperature of the TES device.

The detector and feed horn assembly is housed within a cryostat that is cooled using a PT405 pulse tube cryocooler manufacture by Cryomech (see Fig. 4.2). The cryostat uses a two-stage design where one region is cooled to ~ 65 K and a second region nested within the first is cooled to ~ 4 K. The detector assembly is mounted onto a detector block within and is thermally biased by the $T_s = 4$ K environment. Since the detectors are thermally active, the detector block is typically a few Kelvin above the bias at ~ 6 K. Before cooling, the cryostat must be evacuated to a pressure $\leq 10^{-2}$ mbar to prevent ices from forming within the cryostat. The cool-down and warm-up periods can take several hours, and are monitored using temperature sensitive diodes mounted to the 65 K plate, 4 K plate, and detector block. A typical thermal cycle is shown in the right panels of Fig. 4.2. These diodes are current biased at $10\,\mu$A. Temperature is inferred by measuring the voltage drop across these diodes and referencing these measurements to a temperature calibration curve provided by QMC

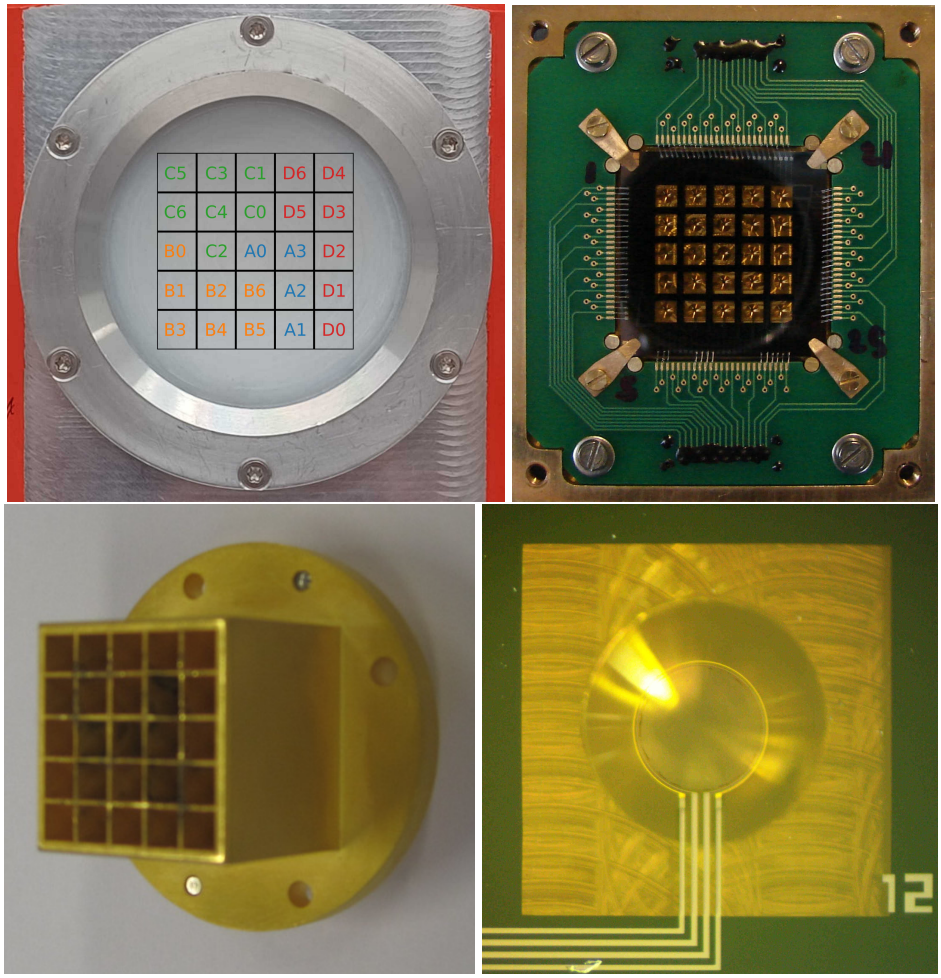


Figure 4.1: Pictures of the detector system. The top left picture shows the physical orientation of the detectors, with their designations, on the cryostat window. Top right shows the detector array which is coupled to the incident electric field using the feed horn array shown in the bottom left. A close up view of an individual detector is shown in the bottom right which notably has two circuit elements: a niobium TES filament and a heater resistor.

Instruments. Data acquisition is facilitated by an external teensy microcontroller which is also connected to sensors which monitor the temperature, pressure, and humidity of the laboratory environment. The software that interfaces with the teensy controller also periodically queries the state to of the cryocooler compressor and logs various metadata that can be used to assess whether the equipment is functioning properly.

A high-level diagram that shows the electrical and thermal circuit of a single TES bolometer is shown in Fig. 4.3. The absorber is represented as a red rectangle on which is mounted a variable resistor R_{TES}, which represents the niobium filament, and a heater

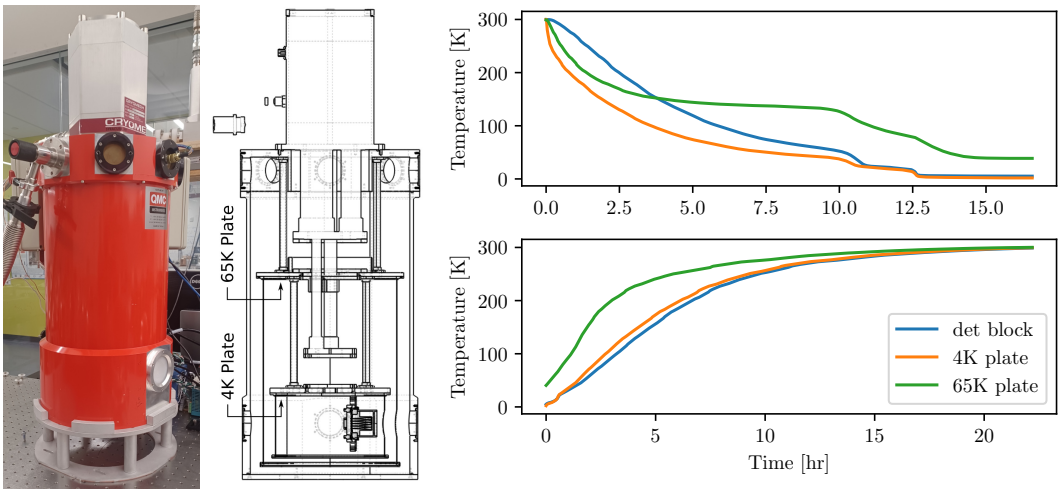


Figure 4.2: The cryostat system. The left panel show a picture of the cyrostat in which the detectors are located. The center panel shows a cross sectional view of the cryostat interior which notably consists of a 65 K and a nested 4 K environment. The detector array is mounted on the detector block within the 4 K volume. The top right and bottom right panels show a cooldown and warmup thermal cycle of the cryostat, respectively.

resistor R_{htr}. It is assumed that the TES resistor and absorber are always in thermal equilibrium at temperature T. The heater is generally not at this temperature, but can efficiently conduct heat to the absorber. The absorber is connected to a thermal ground at temperature T_s, through a weak link with thermal conductance G. The absorber assembly has a bulk heat capacity C. The TES resistor is current biased. Studying the exposed electronics, it appears as though this bias is accomplished using two voltage buffers that maintain a constant potential across the TES in series with a load resistor $R_L \gg R_{\mathrm{TES}}$, such that the current through the circuit is largely unaffected by changes in R_{TES}. Voltage across the TES resistor is proportional to its resistance and is measured in a way that includes additional resistance from the voltmeter leads R_{leads}. The measured voltage V_{TES}, is used as an input to a PID controller, the output of which drives the voltage applied across the heater resistor. The PID controller output is used as the detector signal for this system and is adjusted by the controller such that a predefined V_{TES} is maintained.

When a voltage is applied across the heater resistor, power $P_{htr} = V_{htr}^2/R_{htr}$ is contributed to the absorber. At the same time, the current bias of the TES resistor contributes

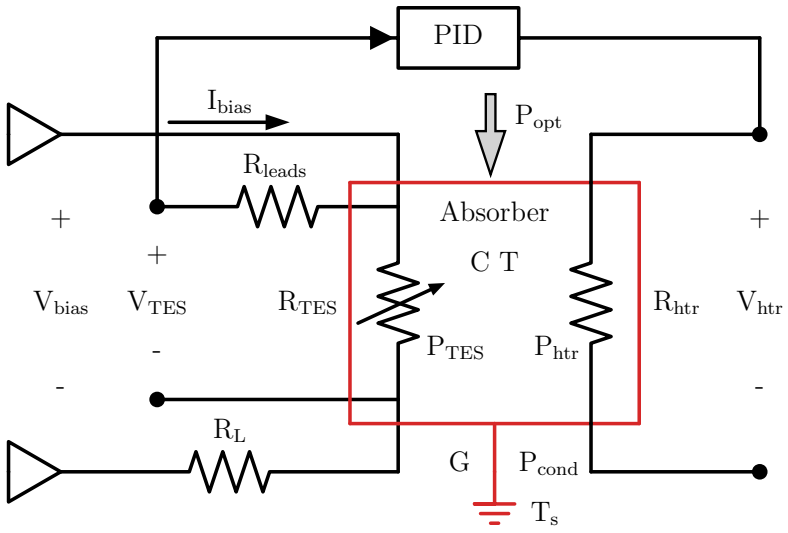


Figure 4.3: High-level thermal and electrical detector circuit diagram. The TES resistor R_{TES} and heater resistor R_{htr} are mounted on the red bolometer absorber with bulk heat capacity C and temperature T. The temperature of the absorber/TES configuration can increase due to power loading from incident optical radiation P_{opt}, and Joule heating from the TES or heater resistors P_{TES} and P_{htr}, respectively. Temperature may decrease when power is conducted P_{cond} to the heat sink at temperature T_s through a thermal conductance G. The TES resistor is current biased I_{bias} using two voltage buffers which maintain a voltage bias V_{bias} across R_{TES} and a load resistor R_L in series. $R_L \gg R_{\mathrm{TES}}$, such that changes in TES resistance do not significantly affect the current. The voltage drop across the TES resistor V_{TES} is measured, which includes a voltage drop due to resistance in the voltmeter leads R_{leads}, and is input to a PID controller. The PID controller outputs a voltage V_{htr} used to drive the heater resistor which is used as the detector signal. The PID controller attempts to maintain a particular TES voltage setpoint which is proportional to the TES resistance and absorber temperature. When the absorber temperature rises due to optical loading, the heater voltage is decreased and vice versa. As such, the detector signal is inverted with respect to the incident optical power.

power $P_{\mathrm{TES}} = I_{bias}^2 R_{\mathrm{TES}}$. If under optical loading, an additional P_{opt} is included which describes the optical power absorbed by the absorber. It is assumed that energy leaves the absorber through a link with thermal conductance G at a rate of $P_{cond} = G(T - T_s)$. Combining these effects, the energy change of the system over a small time increment dt is,

$$\begin{aligned} dE &= \left(P_{opt} + P_{\mathrm{TES}} + P_{htr}\, dt - P_{cond}\right) dt \\ &= \left(P_{opt} + I_{bias}^2 R_{\mathrm{TES}} + \frac{V_{htr}^2}{R_{htr}} - G(T - T_s)\right) dt \ . \end{aligned} \tag{4.22}$$

This equation is analogous to Eq. 4.6, but assumes constant optical loading. Using this change in energy, the change in temperature of the absorber is

$$dT = \frac{dE}{C} \ , \tag{4.23}$$

and this change in temperature changes the resistance of the TES resistor,

$$R_{\mathrm{TES}}(T) = R(T_0) + \left.\frac{dR}{dT}\right|_{T_0} dT \ , \tag{4.24}$$

where $T = T_0 + dT$. Using these equations, the response of the TES to thermal loading can be simulated. One such simulation is shown in Fig. 4.4, in the absence of optical loading. The simulation uses typical parameters that are not necessarily equivalent to the unknown parameters for the detector system used in this work, however, the results are qualitatively representative. When the heater is activated, temperature rises sharply in a largely linear fashion as the change in energy is dominated by the heater term. When the heater is deactivated, temperature decreases through conduction, which depends on the difference in temperature between the absorber and the thermal sink. The rate of change of temperature thus depends on the temperature of the absorber resulting in the characteristic exponential decay curve. The resistance curve is similar in shape to the temperature curve when resistance values are within the approximately linear transition region. When the temperature is

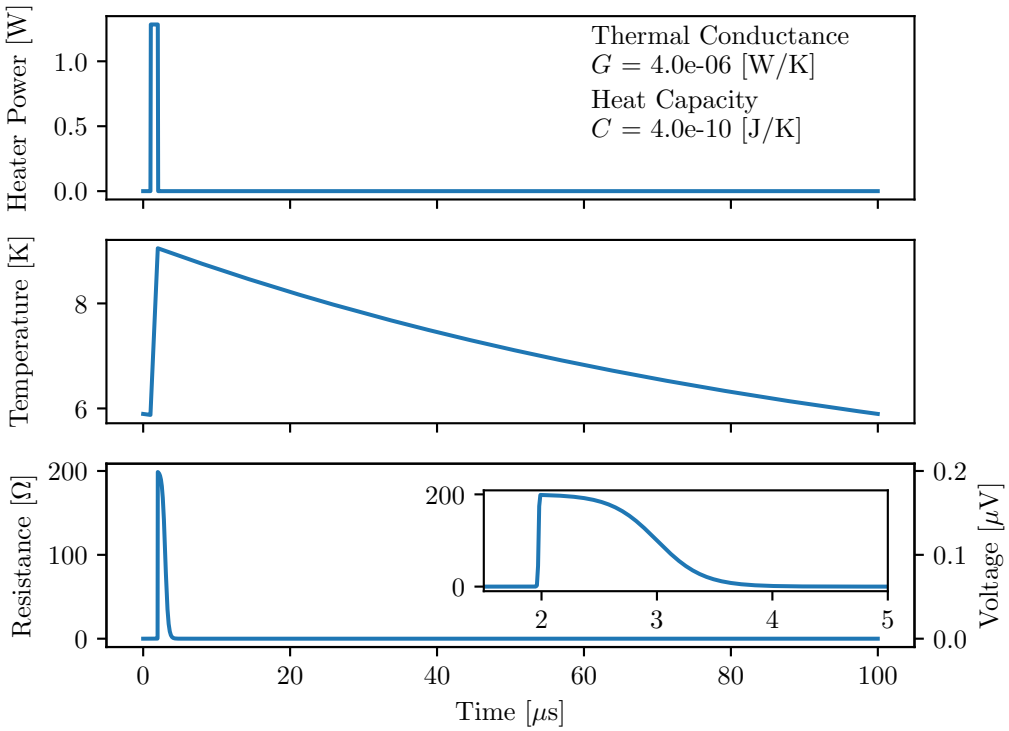


Figure 4.4: Detector response simulation. When the heater is activated (top), power is conducted to the absorber and TES filament which rapidly increases their temperature (center). The increase in temperature results in a rapid increase of the TES resistance (botom) as it passes through its transition region and saturates after the filament enters its normal conducting state. When the heater is deactivated, temperature deceases exponentially with time constant $\tau \approx C/G$. TES resistance decreases slowly while temperature is high enough to maintain the normal conducting state producing a plateau-like region. After additional cooling, the TES passes through the transition region and returns to its superconducting state.

bellow the superconducting temperature, resistance is zero and insensitive to temperature change. Similarly, when the resistor temperature is high enough and the resistor expresses normal resistance, the resistance curve is effectively saturated and further increases in temperature have little effect on resistance. This saturation results in the apparent plateau-like region in the resistance curve.

It should be easy to understand that changing the amount of power supplied by the heater will change the resulting TES resistance curve: if more power is supplied, it will take more time for the TES resistor to relax back into its super conducting state, and vice versa. With the detectors used in this work, the heater voltage is controlled by a PID controller that

adjusts the applied voltage so that the TES resistance is at a predefined value at a specific time. Critically, if power is supplied by optical loading, the heater voltage must be lowered in order to meet this condition. Since the heater voltage is used as the detector signal, the detector signal is inversely proportional to the incident optical power.

The PID controller is a feedback system which accepts a process variable $r(t)$ and returns a control variable $u(t)$ which is input to some system that affects the process variable. The control function is written as

$$u(t) = K_p\, e(t) + K_i \int_0^t e(t')\, dt' + K_d \frac{de(t)}{dt} \,, \tag{4.25}$$

where K_p, K_i, and K_d are positive coefficients corresponding to the proportional, integral, and derivative terms, respectively. The error value is calculated as

$$e(t) = r(t) - y(t) \,, \tag{4.26}$$

where $y(t)$ is a particular setpoint. The PID controller seeks to minimize this error through repeated application of the control function. Although these equations are represented as continuous variables, they are typically sampled and evaluated at periodic intervals. The proportional term has more influence in the control function when the error value is large and acts much like the restoring force of a spring trying to return to equilibrium. If the proportional gain is too large, the process variable may overshoot the setpoint resulting in oscillations which may grow over time and produce instability. The derivative term is effectively a damping term that is most relevant when the error function is changing rapidly and acts in a way to reduce this change. On their own, these two terms effectively model a damped harmonic oscillator. The integral term is most significant when error with the same sign persists over a relatively long period of time. This term can correct for residual error that is not corrected by the other terms. Such a situation may result when there is some bias force on the process variable that equals the restoring force at a particular error value.

The integral term is proportional to the total accumulated error. If the accumulated error is positive and the error value itself becomes negative, the integral term will continue to force the process variable away from the setpoint until the accumulated error becomes zero. In this way, an integral gain that is too high may lead to oscillations and instability. One of the key benefits of PID controllers is that they require no information about how the system responds to the control variable and still manage to achieve good, although generally not optimal, performance in a variety of applications.

4.2.2 Readout

For the detector system used in this work, the process variable is the TES state which is physically a 32-bit number that is a digitized representation of V_{TES} (see Fig. 4.3) and is related to R_{TES} through a linear transformation. The control variable is the heater voltage applied to the heater circuit. In addition to the PID gains, there are a number of detector parameters that define the operation of the system. The active phase for an individual detector within the full time domain multiplex detection cycle is called the readout cycle. This readout cycle is segmented into discrete time steps called loops, with the duration of each loop corresponding to the time it takes for the hardware to read the TES state. The loop time can be artificially increased with software, but the minimum time is constrained by the time it takes the hardware to execute the various software instruction within a given loop. The total time of a readout cycle is the product of the loop time and the number of loops, and is referred to as the pulse width. The number of loops is a free parameter. At the beginning of the the readout cycle, the heater is switched on and remains on for a specified amount of time. As previously discussed, the voltage supplied to the heater is determined by the PID controller. At some point in the readout cycle, measured TES states start being used as inputs for the PID controller, and at some later time, measured TES states stop being used as inputs to the PID controller. The points are called TES read start and TES read end, respectively. Both are free parameters but restricted such that the end point is

greater than the start point and neither can be greater than the pulse width. If averaging is enabled, then the average of these samples will be used as input to the PID controller. If averaging is disabled, only the sample obtained at TES read start will be used.

Additional parameters allow for different operating modes. The multiplexing readout parameter controls whether multiplexing is used. If multiplexing is disabled, then one detector on each channel is read repeatedly. Under nominal operation, only the control variable (heater voltage) of the PID controller is output from the detector system. When the full-pulse readout parameter is enabled, the system will additionally output the sampled process variable (TES state) for each loop count. The various detector parameters previously discussed are summarized in Tab. 4.1. The PID specific parameters affect the control function of the PID controller, while the other parameters affect the readout timing.

An example readout cycle is shown in Fig. 4.5. The top panel shows the sampled process variable at each loop count, obtained when using the full-pulse readout mode. The data is also presented in terms of TES voltage after subtracting the voltage drop associated with the lead resistance. The initial $>$10k value at the start of the readout cycle is attributed to switching noise which has contributions related to heater activation and the multiplexer connecting the TES filament to the bias circuit. The heater remains on until the heater off point, indicated by a star. During this time, the temperature of the filament increases. However, for most of this time, the process variable is constant since the temperature of the filament remains below the superconducting temperature. When the heater is turned off, the heater resistor still has an elevated temperature and takes some time before the excess heat is transferred to the absorber, and subsequently, the filament. As a result of this heating, the filament moves through its entire transition region and well into the normal conducting region. As the fillament cools, it moves back through its transition region into the superconducting region. During some range of the readout cycle, the process variable is read and used as input to the PID controller, indicated by the triangles in Fig. 4.5. In this case, three points are read, and the average is used as the input. The PID controller varies

Table 4.1: Detector system parameters. The first category includes parameters relevant to the PID controller. The second category includes parameters relevant to the detector readout cycle. The thrid category includes parameters which alter the operating mode of the detector system.

Parameter	Description	Typical Range
K_p	PID proportional gain.	$100-1,000$
K_i	PID integral gain.	$100-1,000$
K_d	PID derivative gain.	$100-1,000$
Setpoint (SP)	TES state that the PID loop tries to maintain.	$10,000-18,000$
Loop Time (t_l)	The time it takes for the system to read the TES state.	$\sim 1\,\mu$s
Pulse Width (PW)	The length of an individual detector element readout cycle.	$[5,50]\,t_l$
Heater Time (t_{htr})	Time an individual TES heater is active within the readout cycle.	$[3,25]\,t_l$
TES Read Start ($\mathrm{R_s}$)	Time within the readout cycle when TES states start being input to the PID loop	$[3t_l, \mathrm{PW}]$
TES Read End ($\mathrm{R_e}$)	Time within the readout cycle when TES states stop being input to the PID loop	$[\mathrm{R_s}+t_l, \mathrm{PW}]$
TES Averaging	When enabled, TES state measurements from R_s to R_e are averaged within a single readout cycle and input to the PID loop. When disabled, only the sample at $\mathrm{R_s}$ is used.	Enabled/Disabled
Multiplexed Readout	When enabled, the readout cycles through all detectors within a given channel. When disabled, the readout cycle only includes a single detector per channel.	Enabled/Disabled
Full-pulse Readout	When enabled, the detector system provides the TES state at each loop count within a readout cycle in addition to the PID output value.	Enabled/Disabled

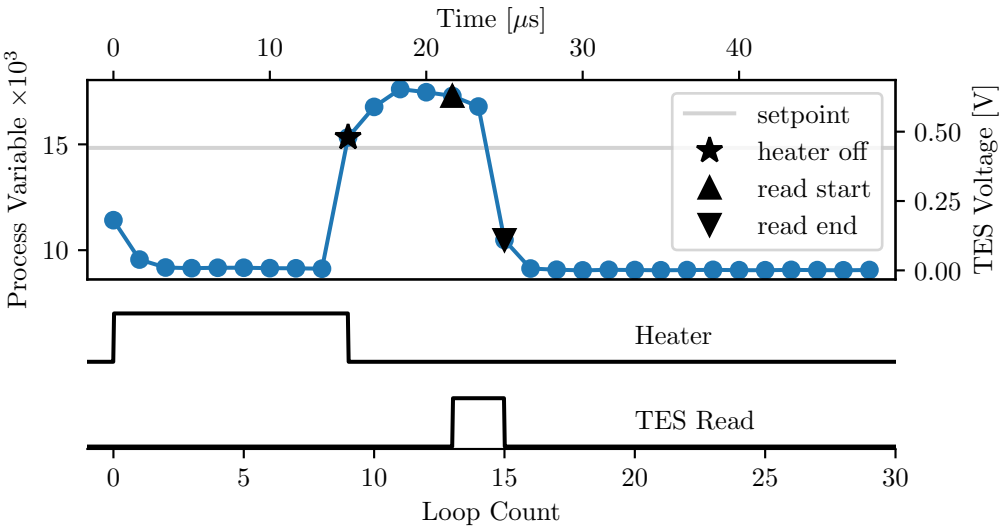


Figure 4.5: Detector readout cycle. The top panel shows the process variable (i.e., TES voltage with leads contribution subtracted) during a readout cycle. The bottom panel shows timing diagrams associated with the heater state and TES read state. At the start of the cycle, the bias circuit and heater are activated producing switching noise. TES temperature begins to increase which results in an increase of the process variable only once the temperature is above the superconducting temperature. When the heater is deactivated, temperature continues to increase as the still warm heater conducts residual heat to the absorber. When the read state is active, sampled process values are averaged and input to the PID controller, which attempts to vary the heater voltage in the next cycle so that this average is equal to the setpoint.

the heater voltage such that the average of these points is equal to the setpoint indicated in the figure by the horizontal line. The top panel of Fig. 4.5 should be compared to the bottom panel of Fig. 4.4, noting the similar shape due to a similar operating principle.

The readout cycles for each detector in the array over a full multiplexing cycle is shown in Fig. 4.6. In this case, the readout cycle is configured using nominal parameters. The A4 to A6 detectors show no signal as these detectors do not exit. Detectors B2 and B3 correspond to unstable detectors and do not provide useful signal under nominal operation.

With the preceding discussion, the difference between conventional TES bolometers and the detectors used in this work can be fully appreciated. With conventional bolometers, the response function and noise equations are derived under the assumption of thermal equilibrium and small power perturbations. These assumption are clearly violated with

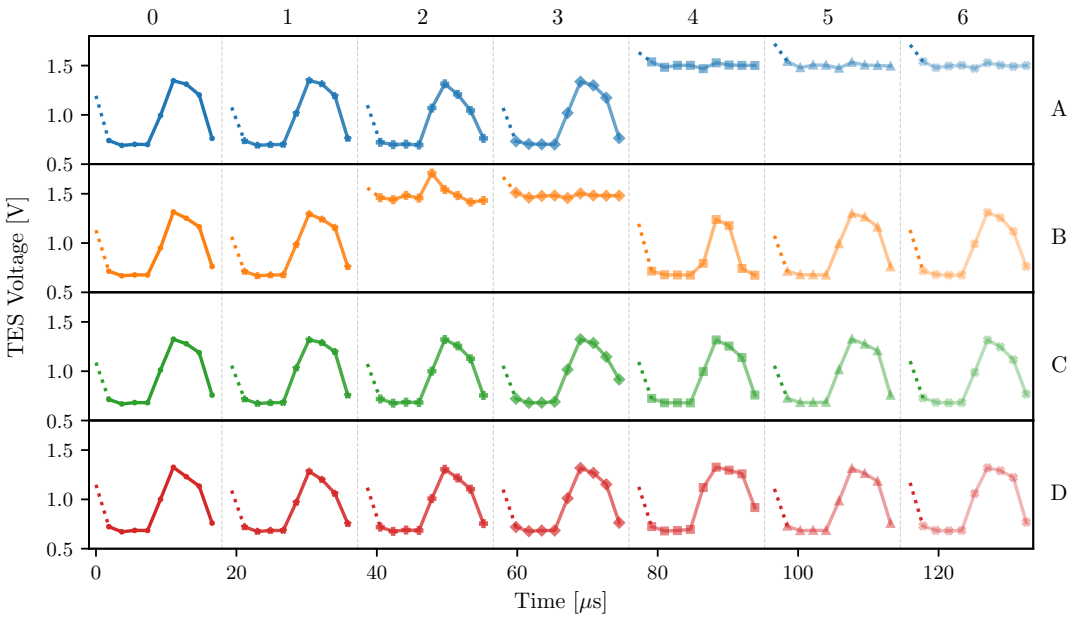


Figure 4.6: Time division multiplexed readout using nominal readout parameters. Up to seven detectors are read out sequentially on a single channel. Note that the A4 to A6 detectors do not exist, and the B2 and B3 detectors are unstable.

this system where the temperature varies rapidly and substantially over a readout cycle. In conventional systems, electrothermal feedback plays the main role in regulating the temperature of the TES. Although electrothermal feedback is relevant with these detectors, their temperatures are primarily regulated by PID control function. Detector performance in terms of response amplitude, response time, and noise properties have been observed to change significantly based on variations of the detector parameters shown in Tab. 4.1. This is a layer of complexity not included in the equations discussed in Sec. 4.1. Although the detector response has been reasonably well modeled using numerical simulations based on the equations presented in this section, convenient analytical expressions for the response and noise are not provided. Instead, figures of merit and other performance metrics are measured experimentally and presented in the next section.

For a conventional current biased TES, the response variable is the voltage drop across the TES filament, which is proportional to the incident optical power. Furthermore, the voltage response of the TES is proportional to the temperature derivative of the TES fila-

ment transition region $S_V = dV/dP_{opt} \propto dR/dT$. That is, a small change in optical power results in a large change in the response variable. In this way, the TES bolometer is effectively a signal amplifier with a gain dR/dT. For the detectors used in this work, the signal is the voltage applied across a heater resistor V_{htr}, but the response variable is better understood as V_{htr}^2 as this is the quantity that is proportional to incident optical power. It is reasonable to assume that the proportionality constant between heater power and optical power is largely independent of TES resistance and is instead determined by some effective coupling efficiency between the TES filament, heater, and absorber. The way in which a sharp transition region benefits sensitivity is that a small change in heater power results in a large change in the TES state. As such, precise adjustment of the response variable is required to move the TES state to its setpoint.

4.3 Characterization

The science observations used in the work were conducted with the detector in multiplexing mode with full-pulse readout disabled (see Tab. 4.1). This is expected to be the nominal configuration for future work as well since multiplexing is required to access the full detector array, and the loop time is typically shorter when full-pulse readout is disabled resulting in a higher frame rate. With these constraints, the detector system still has effectively eight parameters that can be adjusted to vary performance. A broad survey of this parameter space was conducted and detector performance was assessed for the entire array in terms of noise characteristics, SNR, and frequency response. Although collecting, organizing, and analyzing this data represents a significant amount of work, it is out of scope of the present thesis. In this section, detector performance is presented for the central A0 detector using the largely unmodified default parameters provided by QMC Instruments that were used for science observations. The PID configuration was $(K_p, K_i, K_d, \text{SP}) = (1000, 200, 0, 10585)$, and the readout configuration was $(\text{PW}, t_{htr}, R_s, \text{Averaging}) = (10t_l, 4t_l, 9t_l, \text{Disabled})$ with multiplexing enabled.

4.3.1 Transition Curve

The physical property responsible for the high sensitivity of TES bolometers is the rate at which the resistance of the TES filament changes with respect to temperature in the region where the filament transitions between normal and superconducting states. To measure the transition curve, the detector system was turned off and a mirror was placed in front of the cryostat window. This allowed the detectors to reach thermal equilibrium with the cryostat environment and resulted in a largely isolated system since the mirror had a low emissivity and blocks external radiation from entering the system. The detector system was then turned on, however, the readout cycle was configured such that the heater was never active. With these changes, the most significant sources of detector heating were removed. The cryostat was then turned off, and the temperature within the cryostat began to rise slowly at a rate of $\sim 1\,\mathrm{K/min}$. Temperature within the cryostat was monitored using the detector block mounted diode (see Fig. 4.2). Due to the close physical proximity to the detectors, and the slow rate at which temperature increased, it is assumed the detectors closely approached thermal equilibrium with diode thermometer as the cryostat warmed. In this configuration, the main source of deviation from equilibrium was the Joule heating from the TES element as a result of the bias current. The detector system was configured using the full-pulse readout parameter enabled. As such, data output from the system included the process variable (i.e., the TES voltage). A teensy microcontroller was programed to sample the detector block temperature on the rising and lowering edge of an external trigger. This external trigger was also sent to the digital channel of the detector system. In this way, temperature samples from the teensy and process variable samples from the detector data could be synchronized in time. By co-aligning the two datasets, the process variable as a function of temperature was estimated and is presented in Fig. 4.7.

Although the methodology for this measurement was reasonable, it is likely that the temperature of the TES is under estimated. When the resistance of the TES filament is non-zero, current biasing the filament results in Joule heating which raises the TES temperature

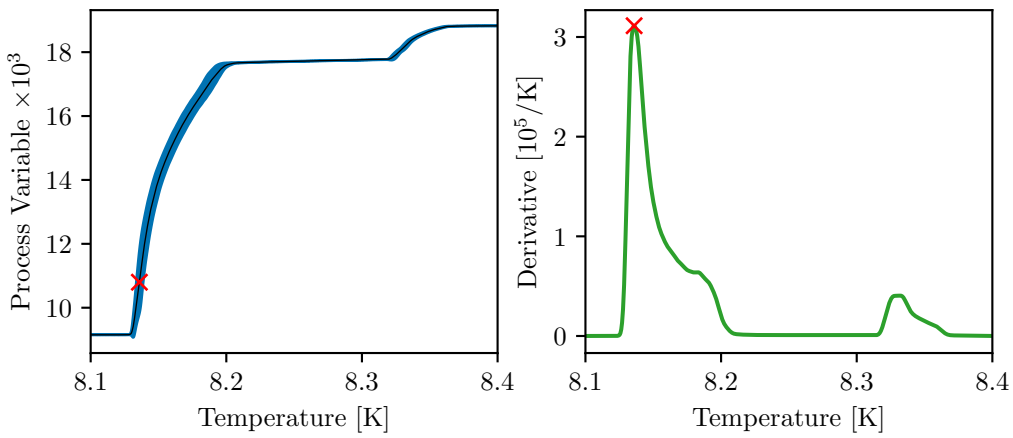


Figure 4.7: Detector transition region. The left panel shows the process variable as a function of temperature where the blue curve indicates the distribution of measurements among four trials. The right panel shows the temperature derivative of the process variable with a peak value at 8.14 K indicated by the red cross. The corresponding 10,585 process variable is used as the setpoint within the PID controller.

above the cryostat environment. This further raises the TES resistance, and subsequently, the sampled process variable. A such, the curve presented in the left panel of Fig. 4.7 is expected be be somewhat more steep than the real transition. However, since the readout was conducted with time domain multiplexing enabled, the detector was not subject to the heating effects of the bias current for $6/7$-ths of the time, which allows for the TES filament to, at least partially, cool back down to the cryostat temperature. An obvious suggestion for improving the methodology would be to decrease the bias current and minimized self-heating. Modifying the bias circuitry was deemed too great a risk and was not adjusted as part of this work.

The temperature derivative of the transition curve is shown in the right panel of Fig. 4.7. This result was used to identify the most sensitive part of the transition, indicated by the red cross at 8.14 K. Heuristically in association with Eq. 4.9, it was assumed that using this value as the setpoint in the PID loop would provide the best performance. Two transition regions are evident in Fig. 4.7. The smaller transition at $\sim$8.35 K is not well understood but is speculated as resulting from impurities and/or variations in geometry of the TES filament where it is joined to the electrical circuit.

4.3.2 Noise Equivalent Power

Noise within the detector signal is an important factor in determining detector sensitivity. To estimate noise properties, a mirror was placed in front of the cryostat window, and signal was recorded for 30 s. A mean subtracted segment of this data is shown in the top panel of Fig. 4.8. The right side of this panel shows the mean subtracted distribution of the full measurement. It is observed that the noise is well approximated as a Gaussian distribution with standard deviation $\sigma_{\mathrm{RMS}} = 8.1 \times 10^{-4}\,\mathrm{V}^2$. Repeated trials of the same measurement produces similar results in terms of the average value and standard deviation, with variations of $10^{-3}\,\mathrm{V}^2$ and $5 \times 10^{-6}\,\mathrm{V}^2$, respectively. Variations in the mean between trials is comparable to the standard deviation of the signal within each trial, indicating noticeable low frequency drift on the order of minutes.

The bottom panel of Fig. 4.8 shows the amplitude spectral density of the noise represented in term of both optical frequency (left/bottom axes) and electrical frequency (top/right axes). The electrical frequency describes the rate at which the signal oscillates physically, while the optical frequency represents where these oscillations would show up if these oscillations were attributed to the source optical spectrum. Noise amplitude is predictably greater at low frequencies due to the typical $1/f$ noise. The rest of the spectrum is relatively flat with a forest of unidentified noise features.

The detector NEP is a quantity often used to estimate sensitivity limits. NEP is functionally defined as the amount of optical power incident on the detector required to achieve an SNR=1 in a 1 Hz bandwidth. Due to the Nyquist sampling theorem (see § 2.3.1), this represents half a second of integration time. A common way to compute this value is by using the ratio of noise amplitude spectral density and power responsivity. In principle, this is a frequency dependent quantity, but is typically presented as a single number by integrating over the instrument bandwidth. In this sense, NEP can be calculated as the ratio of the bandwidth normalized Root Mean Squared (RMS) noise amplitude and the integrated

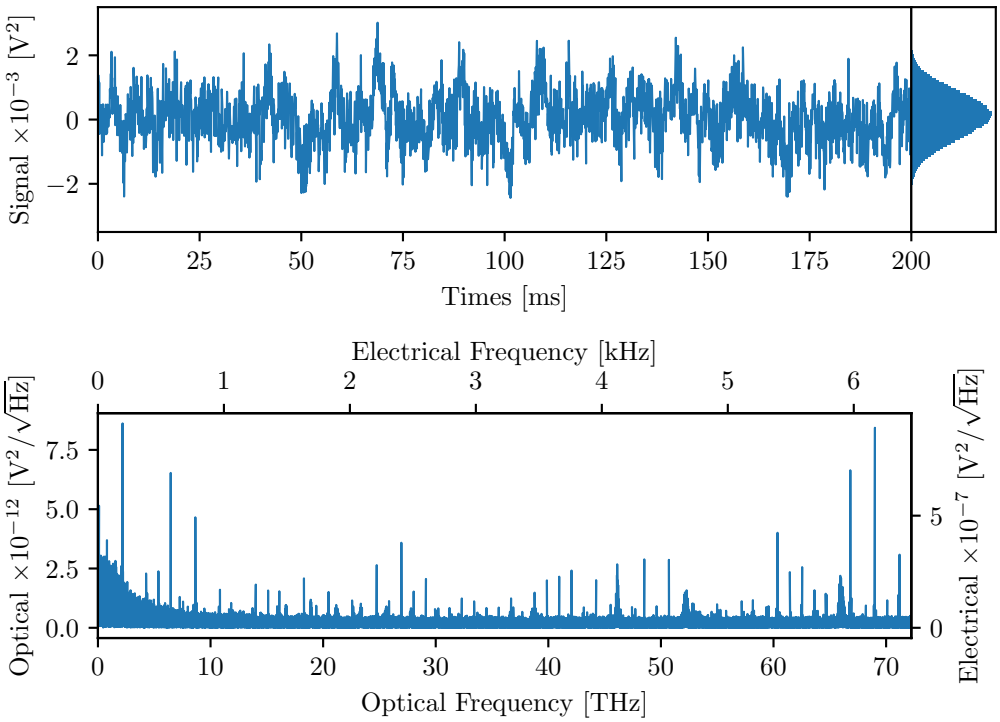


Figure 4.8: Detector noise properties. The top panel shows a short, mean subtracted, time stream of detector signal when the cryostat window was covered by a mirror. A histogram of a longer version of this noise signal is shown on the right which is approximately Gaussian with a $8.1 \times 10^{-4}\,\text{V}^2$ standard deviation. The bottom panel shows the amplitude spectral density of this noise signal and is presented in terms of optical and electrical frequencies. The noise spectrum shows a forest of unidentified noise features and a characteristic $1/f$ low frequency noise profile.

power response,

$$\text{NEP} = \frac{\sigma_f}{S_P}\,. \tag{4.27}$$

Here $\sigma_f = \sigma_{\text{RMS}}/\sqrt{\Delta f}$ represents the bandwidth normalized RMS noise amplitude. The denominator uses the square root of the electrical bandwidth $\Delta f \approx 6.5\,\text{kHz}$ for normalization, which may seem strange. The square root of the bandwidth is used because σ_f is conventionally computed as the square root of the bandwidth normalized RMS noise power. Note that the electrical bandwidth in this case is given by half the detector frame rate. The integrated power response is given by $S_P = \Delta S/\Delta P$, which is obtained by measuring the change

in signal amplitude ΔS for a given change in optical power ΔP . Using the substitution $\mathrm{SNR} = \Delta S/\sigma_{\mathrm{RMS}}$,

$$\mathrm{NEP} = \frac{\Delta P}{\mathrm{SNR}\sqrt{\Delta f}}\,, \tag{4.28}$$

and is the expression used to estimate NEP in this work.

To measure ΔS, the wide field thermal source discussed in § 3.2.4 was placed $\sim 10\,\mathrm{mm}$ in front of the cryostat window and set to a particular temperature. After a few minutes, when the source reached equilibrium, data was recorded for a duration of $\sim 3\,\mathrm{min}$. A second trial was conducted where the source was modulated by a mechanical chopper at $10\,\mathrm{Hz}$. This procedure was conducted for temperatures between $50 - 110\,^{\circ}\mathrm{C}$ in increments of $10\,^{\circ}\mathrm{C}$. In the static case, the signal difference is computed as the difference between the average value of two trials at different temperatures. It was anticipated that this approach may be subject to low frequency errors resulting from signal drift on the time scale of a few minutes. This is why the modulated trial was conducted. For the modulated trials, the periodic waveforms were averaged with reference to the chopper generated trigger signal input to the detector system's digital channel. The peak to peak amplitude of the resulting average is used as an estimate for the modulated signal. Taking the difference between the modulated signals at two temperatures can be modelled as $(S_{T2} - S_{bkg}) - (S_{T1} - S_{bkg}) = S_{T2} - S_{T1}$, and is expected to be similar to the static measurement in the absence of low frequency noise $< 10\,\mathrm{Hz}$.

To estimate ΔP, the power incident on the detector resulting from the thermal source at temperature T is modelled as

$$P(T) = \int_{0.3\,\mathrm{THz}}^{2.2\,\mathrm{THz}} \mathcal{T}_{win} A\Omega\, B(\nu, T)\, d\nu \tag{4.29}$$

where $B(\nu, T)$ is the Planck function (see Eq. 4.11), $\mathcal{T}_{win}$ is the cryostat window transmission (see § 4.3.4), $A\Omega$ is the optical throughput, and the integral is taken over the instrument optical bandwidth. This model does not account for spurious reflection losses and resonant

cavities which may reduce the amount of absorbed radiation and modify the spectral response. Throughput is taken to be independent of frequency and is defined by the product of the detector feed horn aperture area A, and the solid angle Ω, defined by a cone with the same f-number as the feed horn $f_{\#} = 4.5$. Since the wide field thermal source fills the detector aperture, this throughput parameterization is expected to be reasonably accurate. Power difference with different source temperatures is then $\Delta P = P(T_2) - P(T_1)$.

Following the steps outlined in the previous discussion, NEP was calculated for every combination of source temperature. Results are shown in Fig. 4.9. The right panel shows a heat map where the top left triangle corresponds to static measurements with $\mathrm{NEP}_{static} = 2.0 \pm 0.3\,\mathrm{pW}/\sqrt{\mathrm{Hz}}$, and the bottom right triangle corresponds to the modulated measurements with $\mathrm{NEP}_{mod} = 7.1 \pm 0.2\,\mathrm{pW}/\sqrt{\mathrm{Hz}}$. The right panel show a histogram of the results in which an outlier is evident for the static measurements. This outlier is attributed to poor thermalization of the source and is excluded from statistical calculations. The static measurements have a slightly higher spread than the modulated values, as expected from the low frequency noise present in the static measurements. NEP measurements using both methods are of the same order of magnitude, and are reasonably consistent with the $2.8\,\mathrm{pW}/\sqrt{\mathrm{Hz}}$ reported by QMC Instruments. Since the detector is not operated under static loading during science observations, the modulated estimates is considered a better representation of the instrument NEP. How the instrument response changes as a function of modulation frequency is discussed in the next subsection.

4.3.3 Frequency Response

Frequency response describes the sensitivity of the detector to different modulation frequencies and is related to the speed with which the detector can responds to variations in the optical loading. One way to understand this relationship is to consider the impulse response of a hypothetical system. The impulse response can be modelled as the product of the Heaviside step function, and an exponential decay. The step function represents an

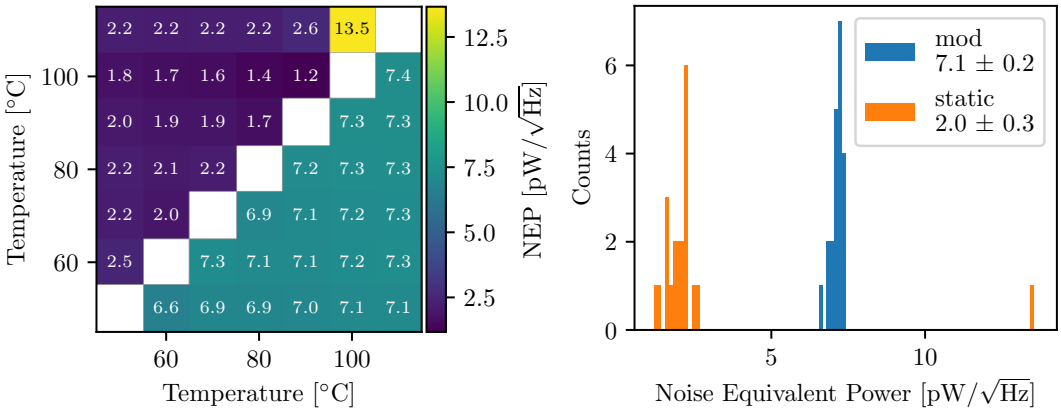


Figure 4.9: Noise equivalent power measurements. The left panel shows a grid of measured NEPs using different temperature differentials. The top left triangle represents measurements using static optical loading while the bottom right triangle represents measurements using modulated optical loading. The right panel shows histograms for these two datasets. Statistics for the static measurements exclude the outlier.

impulse or rapid increase in temperature of the TES, and the exponential decay expresses the characteristic dependence of a variable who's time derivative is proportional to the variable. Such is the case for a bolometer who's cooling rate is proportional to the temperature difference between the TES filament and the thermal sink. The impulse response can be written as

$$h(t) = u(t)e^{-at} \,, \tag{4.30}$$

where $u(t)$ is the Heaviside step function, and a parameterizes the exponential decay. This parameter is often associated with a time constant $\tau = 1/a$ which characterizes the time it takes for the system to relax back to its equilibrium state. The impulse response of an arbitrary system modelled in this way is shown in the left panel of Fig. 4.10.

The Fourier transform of Eq. 4.30 is derived in Sec. C.2 and is given by

$$\hat{h}(f) = \frac{1}{a + i2\pi f} \,, \tag{4.31}$$

which has the same form as the transfer function of a first-order lower pass frequency filter. Note the similarly of this equation with the voltage response of a conventional TES

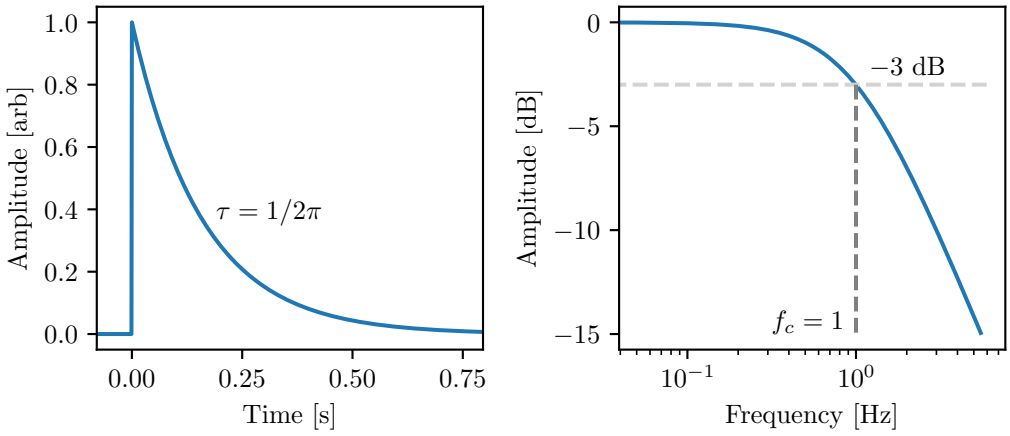


Figure 4.10: Fourier transform of a truncated exponential decay function. The left panel shows a signal representative of a system that is rapidly excited and allowed to decay with a characteristic time constant τ. The right panel shows the corresponding Fourier transform plotted using a decibel scale (see Eq. 4.34). Notice that the frequency response reaches the $-3\,\mathrm{dB}$ level at the cutoff frequency f_c, as expected for a low-pass filter.

bolometer Eq. 4.10. Taking the absolute value give

$$|\hat{h}(f)| = \frac{a}{\sqrt{1+\left(\frac{2\pi f}{a}\right)^2}}\,, \tag{4.32}$$

which is the frequency response curve of a first-order low-pass filter [124]. The cutoff frequency f_c, is defined to be the frequency at which the power spectrum drops to half its initial value, $|\hat{h}(f_c)|^2 = |\hat{h}(0)|^2/2$. This condition is met when $a = 2\pi f_c$, and the resulting relationship between the exponential decay time constant and the power spectrum cutoff frequency is,

$$\tau = \frac{1}{2\pi f_c}\,. \tag{4.33}$$

Often the frequency response curve is presented on a decibel scale computed as

$$L(f) = 10\log\left(\frac{|\hat{h}(f)|^2}{|\hat{h}(0)|^2}\right) = 20\log\left(\frac{|\hat{h}(f)|}{|\hat{h}(0)|}\right) \quad [\mathrm{dB}]\,, \tag{4.34}$$

which has a value of $\sim -3\,\mathrm{dB}$ at the cutoff frequency. The frequency response associated

with an arbitrary system is shown in the right panel of Fig. 4.10.

An Fourier Transform Spectrometer (FTS) or DFI instrument measures interferograms which typically express a superposition of harmonic functions. The detector will have low sensitivity to variations within the interferogram if the frequencies of these oscillations are above the detector cutoff frequency. As such, knowledge of the detector frequency response and the harmonic content of interferograms must be understood to obtain good instrument performance. The FTS system used in this work possesses a spectral stage with a folding factor of four, meaning that every one unit of stage travel results in four units of Optical Path Difference (OPD) variation. When viewing a monochromatic source with wavelength λ, the detector records a sinusoidal interferogram as a function of OPD also with wavelength λ. As such, if the stage is moving at a constant speed v_{mech}, the frequency of the signal recorded on the detector is

$$f_{det} = \frac{v_{opt}}{\lambda} = \frac{4v_{mech}}{\lambda} \,. \tag{4.35}$$

To directly measure the frequency response of the detector system, the photomixer described in § 3.2.4 was placed at the input of the FTS with its output radiation set to 850 GHz. The speed of the spectral stage was then configured to produce a particular frequency. An interferogram was recorded and analyzed with the intent of estimating the amplitude and frequency of the resulting sinusoidal function. Analysis was performed in the frequency domain by zero padding the interferogram and taking the absolute value of the Fourier transform. The resulting feature is well approximated as the absolute value of a sinc function. Data was fit using such a sinc model to extract the amplitude and frequency of the sinusoidal interferogram. Performing analysis in the frequency domain provides a simpler and more accurate way of determining the average amplitude and frequency of the sinusoid than in the signal domain where the amplitude of each peak is difficult to determine because of noise. Zero padding the interferogram before transformation results in sinc interpolation in the frequency domain (see § 2.4.5), which provides more points to constrain the fitting algorithm and generally results in better performance. In this case, the signal was padded

by a factor of three. Sample analysis for an interferogram and fit result are shown in the top panels of Fig. 4.11.

The procedure described above was repeated for a number of frequencies to build up a sample set sufficient to measure the detector frequency response. These data were then fit to Eq. 4.32 to determine the detector cutoff frequency. Results for forward and reverse scans are shown in the bottom panels of Fig.4.11. The cutoff frequency was determined to be $f_c \approx 174\,\mathrm{Hz}$ which corresponds to a time constant $\tau \approx 1\,\mathrm{ms}$. By using $f_{det} = f_c$, and by knowing the maximum optical frequency ν_{max} or equivalently minimum wavelength $\lambda = \lambda_{min}$, the optimal stage speed can be calculated using Eq. 4.35. Initially this calculation was performed for $\nu_{max} = 2.5\,\mathrm{THz}$, resulting in an estimate of $v_{mech} = 6.7\,\mathrm{mm/s}$. However, this calculation initially included an erroneous estimate of the cutoff frequency resulting from an error in the frequency response fitting model. Science measurements were consistently performed with the stage speed given above. With this stage speed and the correct detector cutoff frequency, the DFI system achieves an optical cutoff frequency of $\sim 2\,\mathrm{THz}$, which is relatively close to the $2.2\,\mathrm{THz}$ upper edge of the instrument bandwidth. Although this error attenuates signal at the higher optical frequencies, the complex visibility measurement, the main quantity of interest for this work, is obtained by dividing one spectrum by another. As long as both spectra have the same attenuation error, the results are still valid even though SNR is moderately compromised.

4.3.4 Spectral Response

The detector signal is the voltage applied across the heater resistor in the detector circuit (see Fig. 4.3). However, since the power output by the heater is used to compensate for changes in the incident optical power, heater voltage squared is the more useful response variable. When computing the Discrete Fourier Transform (DFT) of a recorded interferogram, the resulting spectrum also has units of volts squared. The function which converts the instrument units to physical units that describe the incident optical power is called the

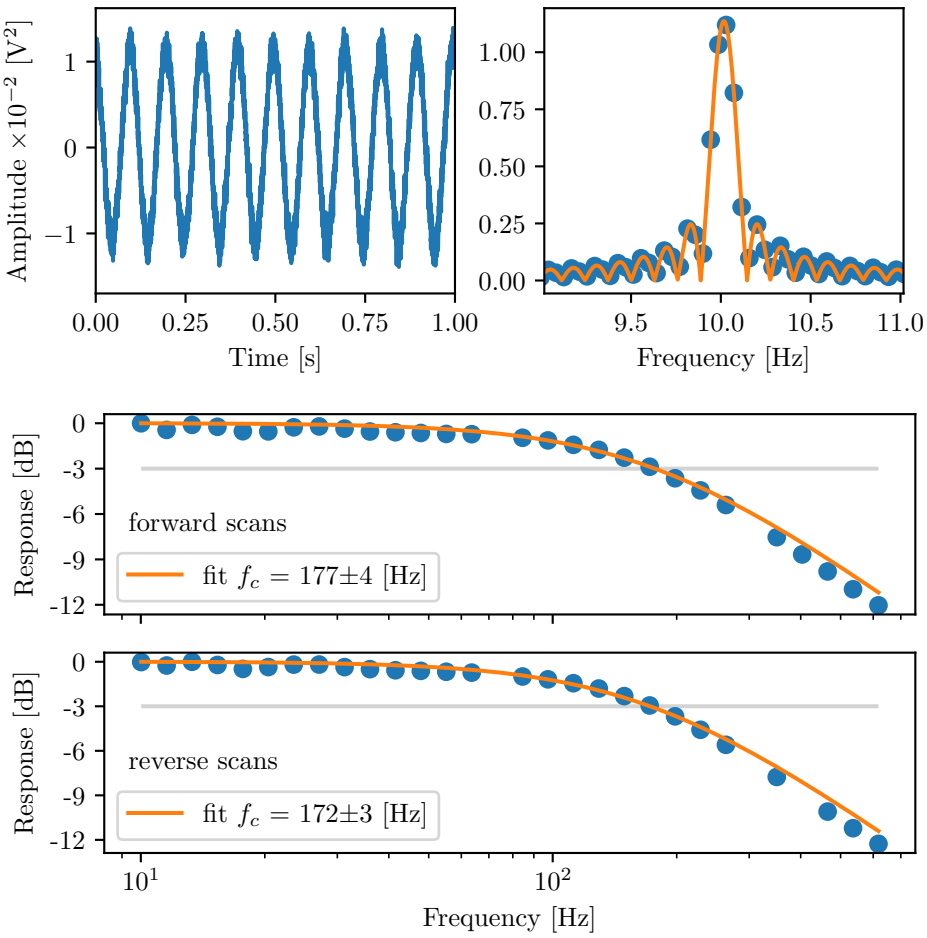


Figure 4.11: Detector frequency repose. Sinusoidal optical loading was achieved using the interferometer response associated with a monochromatic source, the signal of which is shown in the top left panel. The absolute vale of the Fourier transform of this zero padded signal is shown in the top right panel. This spectral feature was fit using the absolute value of the sinc function to extract the amplitude and frequency of the sinusoidal signal. Results for multiple trials analyzed in this was are plotted in the bottom panels using both forward and reverse scans of the interferometer spectral stage. This data is fit using a first order low-pass filter equation (see Eq. 4.32) to estimate the cutoff frequency of the detector $f_c = 174 \pm 4$ Hz. Missing data points are due to corrupted data.

spectral response.

The spectral response function can be obtained by computing

$$\mathcal{R}(\nu) = \frac{S(\nu,T)}{B_{eff}(\nu,T)} \qquad \left[\frac{\mathrm{V}^2}{\mathrm{W}}\right] , \tag{4.36}$$

where $S(\nu,T)$, with units $[\mathrm{V}^2/\mathrm{Hz}]$, is the spectral density obtained experimentally when observing a thermal source at temperature T, and $B_{eff}(\nu,T)$, with units $[\mathrm{W}/\mathrm{Hz}]$, is the effective spectral density of the thermal source. The effective spectral density is modelled as

$$B_{eff}(\nu,T) = \mathcal{T}_{eff}(\nu)\, n\lambda^2\, B(\nu,T) , \tag{4.37}$$

where $\mathcal{T}_{eff}(\nu)$ is the effective transmission spectrum which accounts for all the ways in which the spectral intensity may be reduced when radiation propagates from the source to the detector through the interferometer system. The Planck function is denoted by $B(\nu,T)$ (see Eq. 4.11), and $n\lambda^2$ describes the n-moded throughput of the optical system at wavelength λ. Eq. 4.37 is intended to model the spectral density of light that is incident on the detector system. Any additional inefficiencies not captured by this equation that manifest in the experimental spectrum are attributed to detector effects and are incorporated into the spectral response.

For a diffraction limited single-mode optical system, throughput is given by λ^2 [117]. When the system can support the propagation of multiple modes, throughput is increased to $n\lambda^2$, where n is the number of modes. The number of modes a system can support for a particular wavelength of light λ is largely determined by the system waveguide geometry which allows for the efficient transmission of electromagnetic waves. In this work, the detector feed horns accept electromagnetic waves through a square input aperture with side dimensions 3.63 mm. The horns gradually taper to a square output aperture with side dimensions 0.59 mm and this smaller shape is approximately maintained over a length of $\sim$7 mm (see Fig. D.6). Within this region, the electromagnetic waves are regarded as prop-

agating though an approximately rectangular waveguide.

The number of modes supported within a rectangular waveguide is obtained through solutions of Maxwell's equations with appropriate boundary conditions for the rectangular conductor [123]. To summarize, when an electromagnetic wave enters the waveguide, the wave number vector will have components along the horizontal and vertical directions of the waveguide $\vec{k} = \langle k_x, k_y \rangle$, where $k^2 = k_x^2 + k_y^2$. For a rectangular waveguide with horizontal and vertical dimensions a and b, respectively, the wave is supported when $a = m/2k_x$ and $b = n/2k_y$, where m and n are positive integers. Effectively, an integer multiple of half the projected wavelength along the horizontal or vertical direction must fit within the horizontal or vertical extent of the waveguide, respectively. This is the familiar resonance condition for a wave bound by two nodes within a cavity. Under these conditions, the cutoff frequency at which a mode is supported is given by

$$\nu_{mn} = kc = \frac{c}{2}\sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2}, \tag{4.38}$$

where c is the speed of light. Frequencies bellow the cutoff are exponentially attenuated as the wave travels along the waveguide. A rectangular waveguide can propagate both electric and magnetic fields in this way and are referred to as Transverse Magnetic (TM_{mn}) and Transverse Electric (TE_{mn}) depending on which field components is perpendicular to the direction of energy propagation. Solutions to Maxwell's equations require that $m \neq 0$ and $n \neq 0$ for TM modes, and that at least one of m or n is not zero for TE modes.

The rectangular waveguides used in this work possess the geometry $a = b = 0.59\,\text{mm}$. By evaluation Eq. 4.38 for all permutations of $m, n \in [0, 20]$, with separate accounting for TM and TE modes, and taking the cumulative sum, the number of active modes as a function of optical frequency was computed. Multiplying the number of modes by the wavelength squared corresponding to these optical frequencies gives the theoretical multimoded throughput. These results are shown in Fig. 4.12. The number of modes increases with frequency in a roughly quadratic fashion while wavelength squared has the reciprocal de-

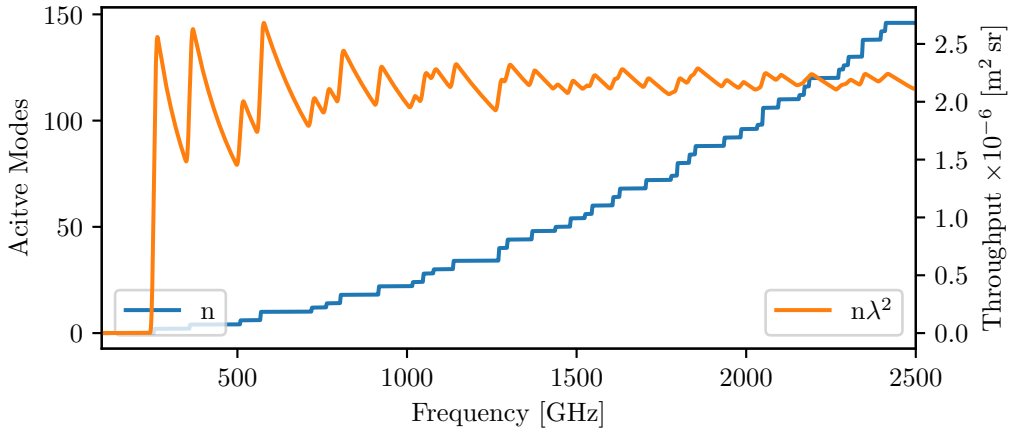


Figure 4.12: Modal analysis of the feed horn waveguide. The frequency dependence of active modes supported by the waveguide is shown in blue and associated with the left vertical axis. Multiplying the number of active modes by the wavelength of light squared associated with a particular frequency gives the multimodal throughput of the waveguide. This throughput curve is shown in orange and is associated with the right vertical axis. Although rapidly varying at low frequencies, throughput is roughly constant with frequency.

pendence. As such, their product, while varying significantly at low frequencies, is largely constant.

The effective transmission spectrum is a product of multiple independent factors associated with different mechanism which attenuate radiation as it propagates from the source to the detector. The effective transmission spectrum can decomposed into

$$\mathcal{T}_{eff} = \mathcal{T}_{atm} \cdot \mathcal{T}_{bs} \cdot \mathcal{T}_{win} \cdot \mathcal{T}_{stg} \cdot \mathcal{T}_{loss} \ , \tag{4.39}$$

where $\mathcal{T}_{atm}$ is atmospheric transmission, $\mathcal{T}_{bs}$ is defined to be equivalent to the beamsplitter efficiency discussed in § 3.2.3 (Eq. 3.2), $\mathcal{T}_{win}$ is associated with the cold filter within the cryostat window, $\mathcal{T}_{stg}$ is the low-pass filter associated with the detector frequency response and chosen stage speed (see Fig. 4.11), and $\mathcal{T}_{loss}$ includes loss due to mirrors and half the optical power being lost to the second output port of the interferometer. Any losses not accounted for (i.e., feed horn reflections) are incorporated into the spectral response estimate.

Atmospheric transmission was simulated using the *am* atmospheric model software [125]. Parameters input to this simulation included environmental variables such as pressure $\sim$900 mbar, temperature 21.5 °C, and relative humidity 14%. These values were taken from the laboratory environment logs recorded during the time when relevant data was acquired. The simulation also included the path length through which light travels, which was set to 5.4 m and is approximately the average length through either arm of the interferometer, from the source to the front of the detector window, during a full spectral stage scan. The only molecular species assumed within the atmospheric model was H_2O.

The transmission spectrum of the cryostat window cold filter was obtained by analyzing a picture provided by QMC Instruments. Loss associated with the mirrors is estimated by assuming a constant mirror reflectivity of 0.97, which is comparable to measured values for aluminum at Far-Infrared (FIR) wavelengths [109]. With a total of eight mirror reflections in each arm of the interferometer in FTS mode, including the collimating and focusing mirrors, $\mathcal{T}_{loss} = 0.97^8 \times 0.5 \approx 0.4$. The factor of 0.5 is included to account for the fact that half of the optical power is lost to the second output port of the interferometer.

The various transmission factors are shown in the top panel of Fig. 4.13. The bottom panel shows the experimental spectral density obtained when observing a 1,200 °C thermal source. This spectrum is bandwidth normalized using the 2 THz instrument bandwidth. Also shown is the effective spectral density obtained by multiplying an appropriate Planck function by the multimoded throughput and effective transmission efficiency spectrum (see Eq. 4.37). Note the high level of agreement in positions of the atmospheric absorption lines between the experimental and simulated spectra over the entire instrument bandwidth. The experimental spectrum expresses a jagged profile that is not represented in any of the transmission factors and is plausibly associated with the multimoded throughput (see Fig. 4.12). The observed peaks do not agree perfectly with the peaks in the model throughput curve. This may be explained by the fact that the geometry at the end of the feed horn is only approximately a rectangular waveguide.

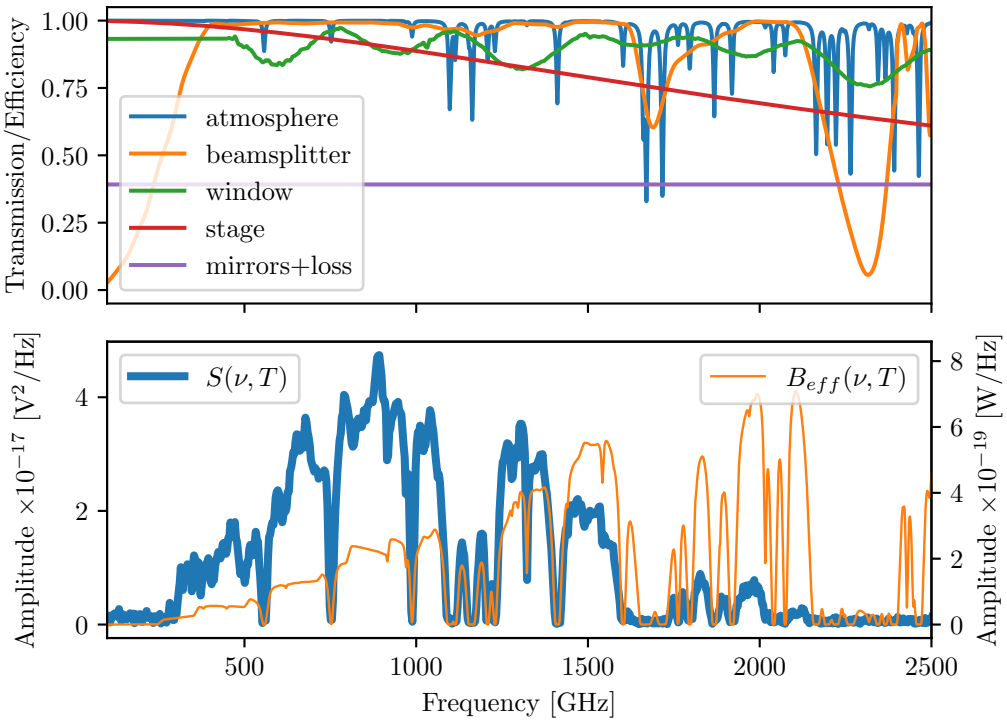


Figure 4.13: Experimental and effective spectral density. The top panel shows the various transmission factors included in Eq. 4.39. The bottom panel shows the experimental spectral density in blue associated with the left vertical axis, and modeled effective spectral density in orange associated with the right vertical axis. Note the difference in units and the high level of agreement in the position of atmospheric absorption features between the two spectra.

The spectral response is obtained by dividing the experimental spectral density by the modeled effective spectral density, blue and orange curves in bottom panel of Fig. 4.13, respectively. A more accurate approach, which cancels any constant signal that may enter the second input port of the FTS, is to use a differential method given by,

$$\mathcal{R}(\nu) = \frac{S(\nu, T_2) - S(\nu, T_1)}{B_{eff}(\nu, T_2) - B_{eff}(\nu, T_1)} \qquad \left[\frac{\mathrm{V}^2}{\mathrm{W}}\right] . \tag{4.40}$$

Calculating the quantities in this equation proceeds in the same way as described above. The spectral response curves obtained using thermal sources at $T_2 = 1,200\,^{\circ}\mathrm{C}$ and $T_1 = 900\,^{\circ}\mathrm{C}$ are shown in Fig. 4.14. It can be observed that the detector response decreases rapidly with frequency. This may be due to loss of absorption efficiency resulting from the geometry of

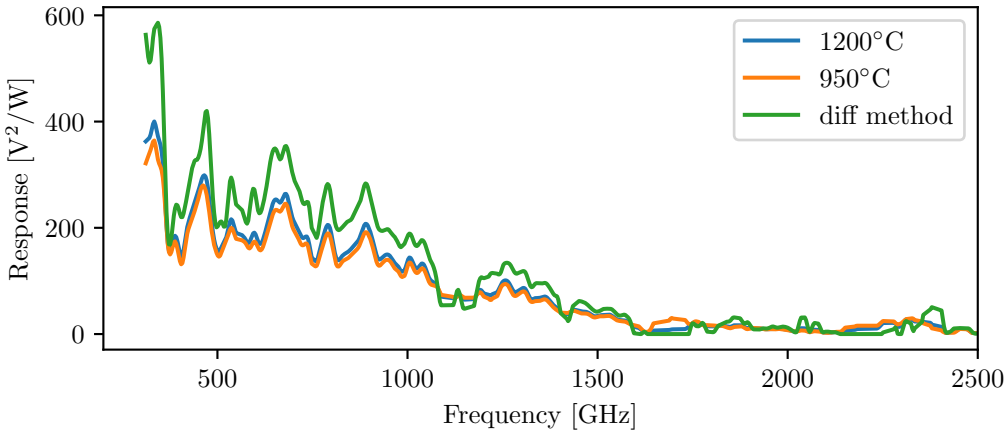


Figure 4.14: Detector spectral response. The curves show the spectral response calculated from observations of a thermal source at two temperature. The differential result is also included which cancels any constant signal entering through the second input port of the FTS interferometer. The spectral response evidently decreases rapidly with frequency.

the bolometer integrating cavity. Alternatively, the spectral response curve appears somewhat reciprocal to the number of active modes curve (see Fig. 4.12). Each unique mode propagates down the waveguide at a slightly different speed due to the slightly different angle at which light corresponding to those modes enter the waveguide. When multiple modes are present, their fields may add incoherently reducing detector sensitivity.

4.4 Summary

In this chapter, the detector system used in this work was discussed in detail. These detectors belong to the class of TES bolometers which detect incident optical power by measuring the change in temperature of an absorbing element that result when the incident optical power is converted to heat. The TES element expresses a rapid change in electrical resistance as a function of temperature and can thus be used as a sensitive thermometer. Conventional TES bolometers use an electrical circuit with a voltage biased TES element, and the change in resistance is inferred from a change in current passing through the element. Temperature of the TES element is regulated primarily though negative electrothermal feedback, and through this process, the TES is maintained at a constant temperature

with relatively small fluctuations. Conventional bolometers have been used extensively and their theoretical foundation is well developed. This theory was reviewed in order to communicate relevant thermodynamic processes, primary sources of noise, and bolometer architecture. However, the detectors used in this work use a significantly different operating principle than conventional bolometers. These bolometers use current biased TES elements and heater resistors that are thermally coupled to the TES absorber assemblies. In the system under study, temperature of the TES is regulated by varying the voltage applied across the heater element and is used as the detector signal. When optical power is incident on the detector, the amount of power the heater must add to the system to maintain a particular temperature decreases. In this way, the detector signal is proportional to the inverse of the incident optical power. The heater voltage is moderated by a PID controller which seeks to adjusted the heater voltage in such a way that the process variable, in this case the voltage measured across the TES resistor, maintains a constant setpoint with each iteration of the PID loop.

The detector system consists of a 5×5 array of feed horn coupled detectors time division multiplexed over four channels. The active phase of a detector during the multiplex loop is called a readout cycle. The readout cycle is parameterized in terms of how much time it takes to complete a cycle, how long the heater element is active within the cycle, and what TES voltage values read during this cycle are used as input to the PID controller. The heater is active during only a fraction of a readout cycle. During this time, the temperature of the TES varies rapidly, and is one of the key differences with respect to conventional bolometers. Another notable difference is that the detectors used in this work are sensitive due to the precise changes in heater voltage required to maintain a particular TES voltage. That is, signal variations are small but precise, while conventional bolometers operate as amplifiers where small changes in incident optical power produce large changes in detector signal.

Detector performance strongly depends on the chosen PID and detector readout pa-

rameters. The parameters used in this work are almost the same as the default parameters provided by QMC Instruments, except the PID setpoint has been modified so that it corresponds to the point on the TES transition curve which varies most rapidly with changes in temperature. With these parameters, various performance metrics were measured for the A0 detector at the center of the detector array. Important figures of merit include the NEP $= 7.1 \pm 0.2\,\mathrm{pW}/\sqrt{\mathrm{Hz}}$ and detector frequency response $174 \pm 4\,\mathrm{Hz}$. Also studied was the noise spectrum and distribution, the multimoded throughput associated with the feed horn waveguide geometry, the transmission spectrum of the interferometer system, and the spectral response of the detector.

Chapter 5

Observations

In this chapter, a detailed overview of experimental procedures, data calibration, and data processing for the purpose of image synthesis using the Double-Fourier Interferometer (DFI) testbed is presented. Sec. 5.1 describes the observation process including all of the measurements that are required to produce a complete dataset with the necessary interferometer and calibration measurements. Sec. 5.2 describes the various calibration steps that are required to remove instrument systematics such that recorded interferograms can be compared to the ideal DFI response, and subsequently analyzed using DFI theory. In Sec. 5.3, a step-by-step procedure for extracting complex visibilities from interferometric data is presented. This is followed by a discussion on aperture synthesis, the process by which complex visibilities are used to synthesize dirty images. This includes a discussion on deconvolution techniques used to reduce distortions resulting from incomplete *uv*-plane sampling.

5.1 Observation Process

In the context of the DFI tesbed used in this work, a complete dataset consist of DFI mode and Fourier Transform Spectrometer (FTS) mode interferograms, in addition to photometric calibration measurements required to compensate for the different intensity through each arm of the interferometer. It is also instructive to discuss various aspects of observation preparation and source plane configuration to better understand the expected performance and limitation of the system.

5.1.1 Observation Preparation

Source Position

The source is placed on the focal plane of the collimating mirror. In general, the "sources" must conform to a model where each point of the source can be represented as a point source that emits light as a spherical wave. Put another way, light emitted from any point on the source must intersect all points of the collimating mirror. If this does not occur, light from that point cannot propagate through both arms of the spatial interferometer and interference is not possible. In the case of the photomixer or Virginia Diodes (VDI) lasers, it is the back focal point of the emitter that is typically co-located with the focal plane. Applying the Huygens-Fresnel principle [8], every point on a wavefront is itself a source of secondary spherical waves, implying that if the back focal point is not on the focal plane, some cross section of the source's emission cone will be. As viewed by the interferometer, this virtual source will then be extended with the same shape as the cross section. This statement is approximately true for small cross sections. For relatively large cross sections, the inclination factor associated with the secondary waves prevents them from illuminating the full collimating mirror. This is why it is not generally possible to mask parts of this cross section area in order to generate arbitrary intensity distributions. The mask will cast shadows on the collimating mirror, violating a necessary condition for spatial interferometry.

The emitting surface of a thermal source can be modelled as a collection of point sources emitting spherical waves. Ideally, this surface would be co-located with the focal plane of the collimating mirror, and an arbitrary source intensity distribution would be constructed by masking parts of the emission surface. In principle, this approach is possible, but in order to achieve a sufficiently bright thermal surface in a safe and efficient way, a thermal cavity is required. As such, the best that can be done is masking off parts of the exit aperture of the thermal cavity. This strategy proves to be sufficient if the emission surface of the thermal source is uniform and all points of the collimating mirror can see this surface when

looking through the aperture mask. The limits under which this requirement is satisfied are discussed in § 3.2.4. In preparation for an interferometer observation, an aperture mask is placed within a mask holder and positioned to be aligned with the focal plane of the collimating mirror. The Pegasus thermal source is then positioned such that its aperture is as close as possible to the mask. Horizontal and vertical positions of the black body are then optimized for mirror illumination. Optimal illumination can be verified by looking at the collimating mirror from the collimated side and observing the mask at different positions along the mirror surface.

With the thermal source operating at temperatures of $\sim 1,000^{\circ}$C, it is not surprising that the mask material gets rather hot, and even charring the material in some cases. Not only does this damage the masks, but it also makes the shape of the source ambiguous since the mask material itself will radiate like a black body. To address this issue, a repurposed computer fan was placed ~ 20 cm away from the mask and angled such that the fan does not obstruct the light path between source and mirror, and such that turbulent air is directed away from the rest of the instrument. Due to the excellent thermal properties of the mask material, this solution was found to reduce the temperature of the mirror facing surface of the mask to room temperature, and the thermal source facing side showed no sign of damage after prolonged use. The turbulence induced in front of the source could in principle degrade the quality of the observation. However, the length of turbulent air the light must pass through is quite small, and any random phase variations between the two arms of the DFI should average out over many scans. These phase variations would be common to both arms of the FTS and thus have little effect on observations in that mode.

Aperture Masks

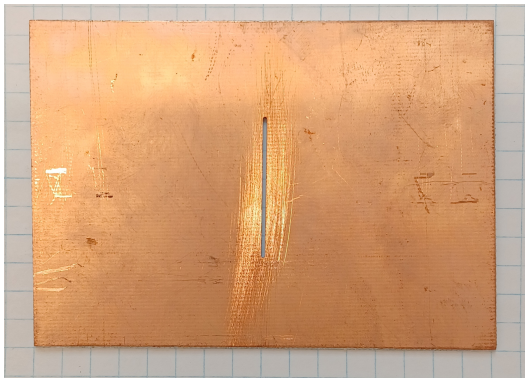
Aperture masks are derived from commercial circuit board material. The core substrate is a laminate of continuous woven fiber glass with epoxy resin. This substrate is flame retardant, a thermal insulator, and maintains its shape and stiffness at high temperatures.

Both surfaces of the substrate are metalized with copper, which has the typical properties of being a thermal conductor with relatively low emissivity. This combination is ideal for the aperture masks. The metal surface facing the source reflects back most light helping to thermalized the cavity while not efficiently transferring heat to the mask. When this side inevitably does heat up, heat is not efficiently transferred though the substrate due to its insulating properties. On the mirror facing side, the low emissivity helps to reduce emission of light from the mask and the high thermal conductivity facilitates cooling by conducting heat to air particles provided by the cooling fan. The relatively high thermal tolerances and mass production of the base material makes these masks safe, cheap, and durable.

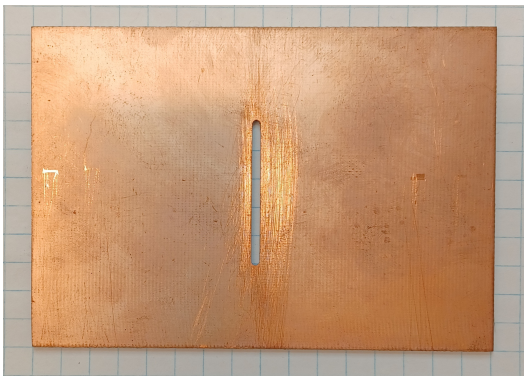
Arbitrary shapes are cut out of this material using a three axis mill. The aperture masks used in this work are shown in Fig. 5.1. The top four are effectively one dimensional sources where the relevant source variation is along the horizontal axis. These masks are convenient since they do not require baseline rotation, or alternatively, source rotation in order to reconstruct an image from interferometer measurements. The bottom mask in Fig. 5.1 has spatial variations in the horizontal and vertical directions, and baseline rotation is required for image reconstruction. Although simple in structure, observations of these masks are relatively easy to model, and are sufficient to validate the core aspects of the DFI technique.

5.1.2 Data Acquisition

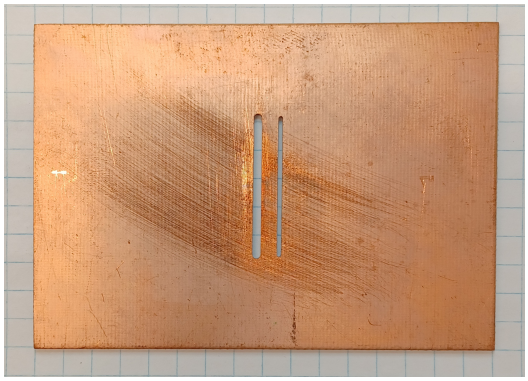
A full observation consists of DFI, FTS, and photometric calibration measurements. Each of these measurements start by homing the translation stages, and for interferometer measurements, the spectral stage is configured to output a Position Synchronized Output (PSO) pulse every 15 μm of stage travel corresponding to 2.5 THz optical frequency Nyquist sampling. Although the order of presentation is DFI then FTS measurement, the observation can be performed equally in reverse order. A comprehensive dataset consists of the



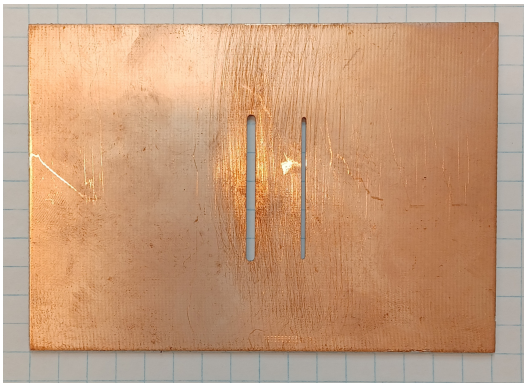
(a) 1 mm slit.



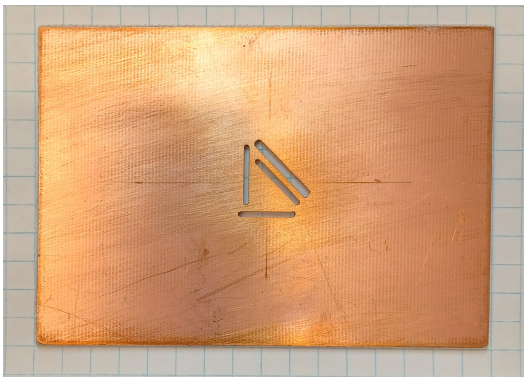
(b) 2 mm slit.



(c) 2 mm and 1 mm slit separated by 5 mm.



(d) 2 mm and 1 mm slit separated by 12 mm.



(e) 2D mask composed of 1 mm and 2 mm slits.

Figure 5.1: Aperture masks used to demonstrate the DFI observation technique.

items discussed in this section, however, some of these items may be omitted depending on the objective of the observation.

DFI Measurement

During a typical DFI observation, the spatial stage moves the dynamic aperture to its first position, setting the initial baseline. Typically, this is the minimum 135 mm baseline. The spectral stage then moves the double rooftop mirror assembly to its appropriate initial position. This position depends on the commanded maximum optical path difference. Typically, the stage is positioned at 2 mm mechanical (8 mm optical), before the phase center. The phase center is initially estimated with ray tracing, and is further refined through observation (see § 5.2.2). The spectral stage is then scanned symmetrically about the phase center at a mechanical speed of $6.7\,\mathrm{mm/s}$. This speed was chosen to be as fast as the stage could travel before resulting in detector signal oscillations, from relevant optical frequencies, above the cutoff frequency of the detector system (see § 4.3.3). Once the stage travels its commanded distance and stops, it is immediately commanded to travel in the opposite direction for the same distance at the same speed with no delay. Scans performed in the first direction are designated as forward scans, and when performed in the other direction, are designated as reverse scans. Short scan distances with minimal delays were chosen to minimize observation time given the large number of scans that need to be averaged to achieve sufficient Signal to Noise Ratio (SNR). The number of forward and reverse scans recorded for each baseline is constant within an observation and is chosen based on the source being observed. For bright sources with modest spatial coherence, like the photomixer or VDI lasers, the number of scans is relatively low, typically between $1-5$. However, for observations of the masked Pegasus thermal source, the number of scans is high, typically between $1,000-4,000$.

Once all spectral scans are performed for a given baseline, the spectral stage will have returned to its initial position. The spatial stage is then moved to set a larger baseline. In-

creasing the baseline reduces the optical path associated with the dynamic aperture, and to compensate for the resulting shift in the phase center, the spectral stage is moved towards the static aperture by one-quarter of the increase in baseline. Spectral scans are then repeated as previously described. The number of baselines measured depends on the observation goals but are typically ten uniformly spaced baseline lengths between $135 - 310\,\text{mm}$. After completing all spectral scans, photometric measurements are conducted as described later in this section.

FTS Measurements

Once DFI mode measurements are complete, the system is configured to operate in FTS mode. This is accomplished by inserting a beamsplitter in front of the dynamic aperture and a flat mirror in front of the static aperture (see Fig. 3.1). These components require some tip-tilt alignment, but due to the lack of spatial coherence effects, the alignment accuracy is somewhat relaxed compared to DFI mode. Before inserting these components, alignment is facilitated by back propagating visible lasers from the beamsplitter along the optical axis. The laser associated with the static aperture is placed just in front of the beamsplitter, while the laser associated with the dynamic aperture is placed between the last two mirrors such that both lasers are pointed upwards with respect to Fig. 2.14. The dynamic aperture is then moved to its nominal position which was previously optimized to align with the center of the FTS beamsplitter. The position and orientation of the lasers are adjusted so that the beam intersects the centers of all the flat mirrors and propagate to the previously measured nominal focal point of the collimating mirror (see § 3.2.5). These procedures are facilitated using 1 mm pinhole apertures which condition the laser beams and can be used to check laser height and alignment with respect to the optics table grid. Typically these pinholes are placed near the collimating mirror and the lasers are adjusted to intersect these pinholes. The spectral stage is then moved to both ends of its allowed travel to ensure intersection does not change with optical path. Light passing through these pinholes and

reaching the collimating mirror focal point over a range of optical paths ensures the lasers are sufficiently aligned. The FTS beamsplitter and mirror are then inserted and their tip-til degrees of freedom are optimized. The laser associated with the dynamic aperture is only transmitted through the beamsplitter and its propagation to the source plane is largely unaffected by the FTS components. As such, the inserted mirror need only be adjusted such that the laser associated with the static aperture intersect the same point on the beamsplitter as the transmitted laser. The beamplitter tip-tilt degrees of freedom are then adjusted to ensure the reflected and transmitted lasers take the same path to the source plane. Since the transmitted laser should already be aligned to the center of the beamsplitter, and the reflected laser is aligned to the transmitted laser, tip-tilt variations should not change the position of either laser on the beamsplitter.

Once the system is aligned, the observation takes place in the same way as a single baseline observation in DFI mode. In principle, the orientation of the source should not affect the FTS measurement, assuming the source is centered on and confined to the beam of the detector. As such, with both one and two dimensional aperture masks, only one FTS measurement is necessary. After recording interferometric data, photometric calibration measurements are made.

Photometric Calibration Measurement

Photometric calibration observations are required to account for the power imbalance between the two arms of the interferometer. This imbalance results from imperfections in the optical system, be it alignment errors or quality of the optics, and also from the geometric properties of the collimating mirror. As discussed in Ap. B, the power in each arm depends on the source intensity and the inverse square of the distance between the source and the point on the collimating mirror associated with the input apertures. Since the static, dynamic, and FTS apertures generally sample different regions of the aperture plane, power imbalance is expected for all observations. In DFI mode, the amplitude of the response is

proportional to the geometric mean of the intensity through both arms Eq. B.11. If the intensity through the dynamic aperture is baseline dependent, then variations of fringe amplitude are affected both by intensity variations and spatial coherence. Thus intensity variations must be corrected to accurately estimate fringe visibility. Similarly, if FTS spectra are to be used for normalization when obtaining complex visibility estimates, the spectrum must be corrected to make it comparable with DFI measurements.

It is assumed that the variations in intensity though any aperture are frequency independent such that the power spectrum obtained with any two aperture positions obey the relation $I_1(k) = aI_2(k)$. With this assumption, calibration measurements only require an estimate of the total power reaching the detector through an aperture. To measure this, a mechanical chopper is placed in front of the source and configured to rotate at a frequency of 5 Hz. Care is taken to ensure the source can be seen through the chopper blades from the perspective of all points on the collimating mirror. The relatively low chopper frequency is chosen to minimize any low-pass filter effects associated with the detector frequency response (see § 4.3.3). The chopper controller sends out a pulse every period of the chopper which is input to the digital channel of the detector. A DFI or FTS measurement is then initiated. These measurements are nearly identical to the process described above with the exceptions that time it takes to perform scans with the spectral stage are limited to ~ 30 s, and one arm of the interferometer is blocked by a sheet of metal. The measurement is then repeated with the other arm of the interferometer blocked. In this way, measurements are taken for each aperture position in DFI mode and both the static and dynamic arms in FTS mode.

For each aperture, data are processed by averaging the resulting waveforms using the trigger signal from the chopper controller to align the individual waveforms. The peak-to-peak amplitude is used as an estimate of intensity through each interferometer arm.

5.2 Calibration

Calibration in the context of this work refers to processes that remove systematic errors from measurements such that they can be reasonably approximated by ideal theoretical models. Ideal instrument response equations for the DFI and FTS modes of observation are derived in Ap. B. More general expressions which allow for potential sources of systematic error are presented below. For DFI mode, the modified equation is,

$$I_{DFI}(\vec{B},z) = \int_{-\infty}^{\infty} I_{12}(k)|\gamma_0(\vec{B},k)|\cos[2\pi k(az+\delta z)+\phi_0(\vec{B},k)+\phi_{DFI}(k)]\,dk\,. \qquad (5.1)$$

The complex visibility amplitude $|\gamma_0(\vec{B},k)|$, and phase $\phi_0(\vec{B},k)$, are the observable quantities of interest, and are baseline $\vec{B}$, and frequency k, dependent.

Optical Path Difference (OPD) z, maybe subject to a systematic scaling error a and offset δz. Scaling error affects the size of OPD steps between interferogram samples. This can then affect the instrument bandwidth, and potentially spectral resolution, resulting in line position and line width errors. Measurements show a high level of agreement between atmospheric absorption models and observed spectra (see § 4.3.4). Experiments with VDI sources (see § 6.2) also suggest a line position accuracy of 50 MHz, a factor of four lower than the best possible spectral resolution of the instrument. As such, errors due to the OPD scaling factor are considered negligible and do not need correction. A systematic offset in OPD implies an error in the DFI phase center. The phase center was initially estimated using geometric models and ray tracing, but this approach is expected to be accurate only to a few millimeters of OPD, and a more accurate estimate of δz is required. A frequency dependent phase term $\phi_{DFI}(k)$, is used to represent phase shifts resulting from the beamsplitter ϕ_{bs}, in addition to any other sources of systematic phase error. If not corrected, these phase effects will show up in the complex visibility phase producing erroneous results. Finally, there is

the term representing the source intensity distribution,

$$I_{12}(k) = \int_{\Omega} \xi(\vec{\theta},k)\sqrt{I_1(\vec{\theta},k)I_2(\vec{\theta},k)}\,d\Omega\,, \tag{5.2}$$

where $\xi(\vec{\theta},k)$ represents the source plane sensitivity of the detector (i.e., the detector beam), $I_i(\vec{\theta},k)$ are the intensities obtained at the detector resulting from light emitted from point sources located at $\vec{\theta}$ propagating through each arm of the interferometer, and the integral is taken over the source Ω. Due to the geometry of the collimating mirror, it is generally expected that $I_1(\vec{\theta},k) \neq I_2(\vec{\theta},k)$, and the relationship between these quantities may also be affected by instrument alignment error and the positional dependence of the collimating mirror surface quality. The source plane sensitivity may reduce the apparent brightness of sources as a function of their position within a detector beam and should also be corrected.

Another potential source of systematic error is the baseline of the DFI, controlled by the spectral stage. Although no calibration process is proposed, the baselines have been physically measured with an absolute accuracy of 2 mm and a relative accuracy of 0.1 mm. These measurements were performed with a ruler. The precision of the absolute baseline measurement is limited by the ability to determine the centers of the aperture mirrors, while the relative accuracy is limited by the ability to read the ruler scale. The relative accuracy is expected to be significantly better, on the order of tens of nanometers, similar to that of the spectral stage which was discussed in detail in § 3.2.2.

An expression similar to Eq. 5.1 can be written for the instrument response in FTS mode,

$$I_{FTS}(z) = \int_{-\infty}^{\infty} I_0(k)\cos[2\pi k(az+\delta z)+\phi_{FTS}(k)]\,dk\,. \tag{5.3}$$

The quantities in this equation have the same interpretation as Eq. 5.1. Of principle concern is the term related to the source intensity $I_0(k)$ which is applied to the DFI response as a normalization function in order to obtain complex visibilities. Due to the relatively high SNR of FTS measurements compared to DFI measurements, the absolute value of the FTS

response may be used which is independent of phase. As such, systematic errors in phase or in quantities which affect the phase can be ignored.

The rest of this section is dedicated to a detailed discussion of how the phase, OPD offset, and source intensity are calibrated.

5.2.1 Phase

The ideal FTS response consists of a superposition of cosine functions and is expected to produce a real symmetric function. From the symmetry relations of Fourier transforms (see § 2.2.1), the Fourier transform of the response is expected to be real and symmetric. However, the instrument may be subject to uncompensated dispersion typically resulting from frequency dependent variations in the beamsplitter and detector electronics. Various sources of random error are also present in the system which combine to produce a response that deviates from its symmetric ideal. As such, in general $\phi_{FTS}(k) \neq 0$ and is composed of systematic and random components. As discussed in § 2.4.1, a non-zero phase shift applied to a cosine function results in a Fourier transform consisting of a delta function rotated in the complex plane by the same phase value. As such, phase errors in the FTS response result in Fourier transforms with complex values. Fortunately, the symmetries present within the ideal FTS response mean that, in principle, any phase errors present in any measurement can be corrected without reference to information external to the measurement.

The ideal DFI response also contains phase contributions resulting from the same sources of systematic and random error. However, the response also contains the phase associated with the complex visibility which gives information about the source intensity distribution. In particular, it effectively encodes positional information about the source which is of critical importance if the source is displaced from the optical axis or is not symmetric. The complex visibility phase is also baseline dependent, further complicating matters. Since there is no consistent symmetry in the DFI response, it is not generally possible to correct each measurement without reference to external information and a standard phase correc-

tion product is required which removes only instrumental effects.

Instrumental phase effects can be decomposed into linear $\phi_L(k)$, non-linear $\phi_{NL}(k)$, and random ϕ_R components,

$$\phi_{DFI}(k) = \phi_L(k) + \phi_{NL}(k) + \phi_R \, . \tag{5.4}$$

Linear phase contributions take the form, $\phi_L(k) = mk$, where m is a constant. In connection with Eq. 5.1, it can be observed that the OPD offset δz produces a phase shift with a linear dependency and slope $2\pi\delta z$. Recall also that a point source offset from the optical axis by an angle θ parallel to the baseline axis B will result in an OPD shift $\delta z = B\theta$. That is, a source offset from the optical axis produces a linear phase shift. This is effectively what is meant by the statement "complex visibility phase encoding source position." This results in a situation where an erroneous OPD offset δz is expressed in the same way as the complex visibility phase. The phase correction process must then be careful not to confuse the two.

In a more comprehensive sense, the complex visibility phase is produced by the combined effects of all the point sources that are imagined to constitute an extended source. If the source possesses significant structure, the resulting phase is more complicated and may be difficult to model. Even if the source structure is simple, if it is near the edge of the detector beam the effective shape of the source can become more complicated. It is also the case that a point source expresses perfect spatial coherence resulting in a DFI response with frequency components that are negligibly attenuated by visibility effects while an extended source may have significant attenuation. Frequency components with low amplitude have low SNR and obtaining an accurate phase measurement for these components is more difficult. Taking account of these effects, phase correction is most productively performed when observing an on-axis point-like source. These conditions are difficult to obtain in practice. The next best option is to use a symmetric source that is as small as possible situated near the optical axis, and to acquire data using the smallest baseline of the interferometer so as to minimize visibility effects and to keep the phase characteristics as simple as possible. For

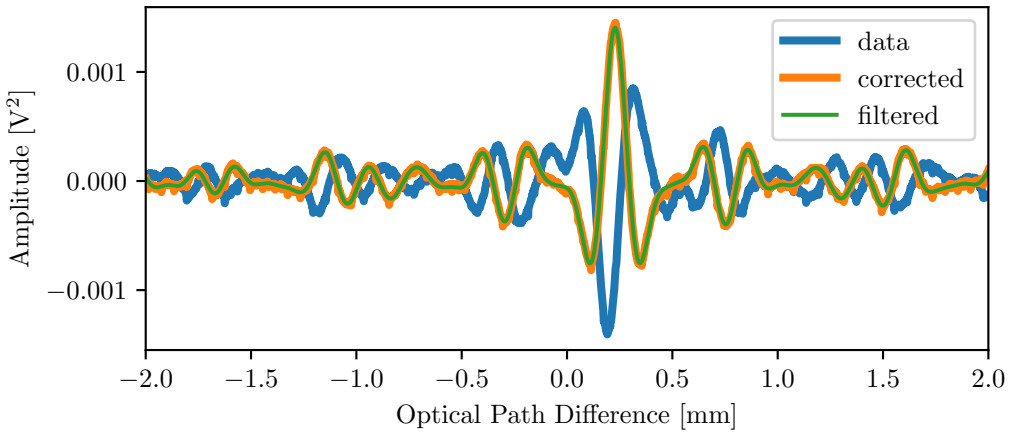


Figure 5.2: Phase corrected interferogram. The blue curve is an averaged interferogram obtained when observing a 1 mm , approximately on-axis, slit backlit by a $1,100\,^{\circ}$C black body, using a 135 mm baseline. The orange curve shows the same interferogram after phase correction, which results in the green curve after passing through at $300-2,200$ GHz optical frequency band-pass filter. Note the high degree of symmetry after correcting for non-linear phase errors.

this work, a 1 mm slit illuminated by a $1,100\,^{\circ}$C black body was used, and the observation was conducted with a 135 mm baseline. Although the slit position was optimized to be as close the optical axis as possible, this step is not required as long as the slit is not near the edge of the detector beam. The data used to calculated the phase correction product, along with the corrected interferogram, is shown in Fig. 5.2.

The phase correction product is obtained in the following way. A point on the interferogram is taken as an estimate for Zero Path Difference (ZPD) of the source, which is not in general the phase center of the instrument. Due to the approximately anti-symmetric nature of the unprocessed interferogram, the initial estimate of ZPD is the sampled point nearest the zero crossing along the axis of anti-symmetry. The interferogram is then truncated on one side so as to follow the sampling requirements of the Discrete Fourier Transform (DFT) (see § 2.3.2). The Fourier transform is then calculated producing real and imaginary components which are subsequently used to calculate the phase using Eq. 2.50. In Python, the calculated phase is bound between $\pm\pi$ resulting in discontinuities when the phase exceeds these boundaries. The "unwrapped" phase is shown in the top panel of Fig. 5.3.

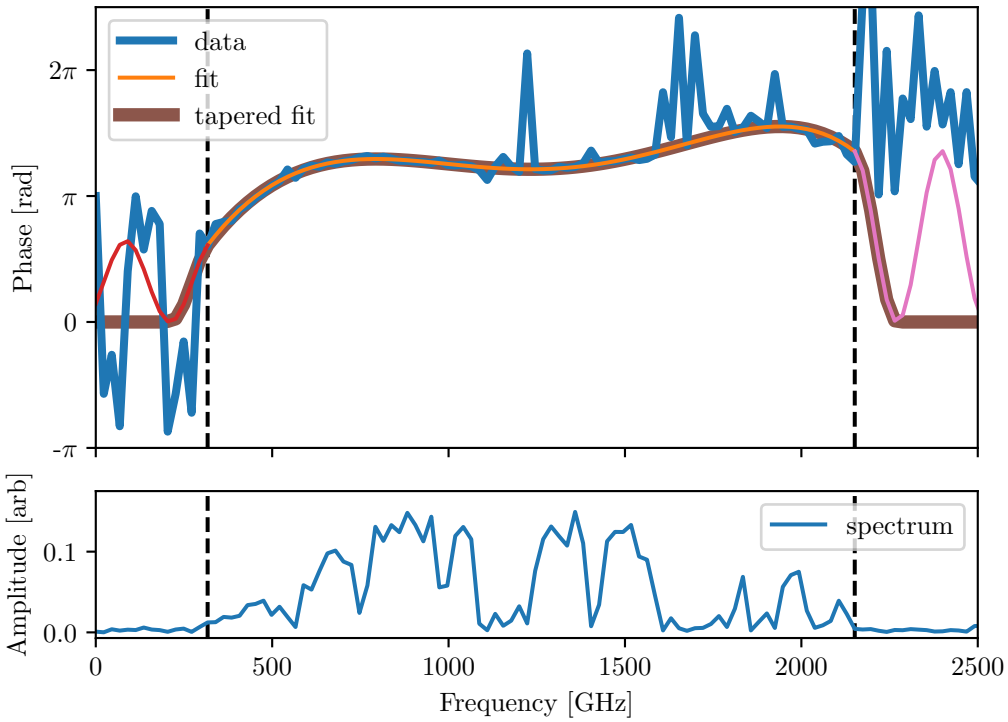


Figure 5.3: Polynomial fit to non-linear phase. The blue curves in the top and bottom panels represent the calculated phase and amplitude spectra of the data from Fig. 5.2, respectively. The phase data residing within the instrument bandwidth (vertical dashed lines), is fit to a fifth degree polynomial, which provided a good result without over fitting the noise. Each data point in the phase spectrum is weighted by its corresponding amplitude in the amplitude spectrum. Note the fit matches the data quite well in regions with high amplitude. The fit is tapered to zero outside of the instrument bandwidth using sinusoids that satisfy boundary conditions such that their value and slope are equal to the polynomial's value and slope at the point where they intersect at 300 GHz and 2,200 GHz.

In addition to calculating the phase, the amplitude spectrum is also calculated using Eq. 2.51. The phase data is then fit using a fifth degree polynomial and the least-squares regression algorithm is supplied with weights corresponding to the amplitude spectrum. This operation ensures that samples with high SNR are given proportionately more consideration when comparing the model polynomial to data. It can be observed in Fig. 5.3 that phase data with low corresponding spectral amplitudes tend to deviate from the otherwise smooth phase curve. The frequency range that is fit is limited to the instrument bandwidth of $\sim 300 - 2,200$ GHz, as indicated by the horizontal dashed lines.

To avoid correction for the linear phase effects associated with the complex visibility,

an attempt is made to minimized the linear term in the resulting polynomial fit. The extent to which a linear term is needed is influenced by which point is chosen as ZPD in the uncorrected interferogram data. As such, twenty points on either side of the initial ZPD estimate are tested, and the one which minimizes the linear term is ultimately used. To this end, it is important to expand the polynomial fit around the center of the instrument bandwidth. A quadratic term, for example, looks very much like a line far away from the origin. If the polynomial is expanded around 0 GHz, then the linear term serves largely to correct the effects of other terms in the polynomial and does not accurately reflect the effects of complex visibility phase or ZPD offset.

Once the desired polynomial is obtained, it must be tapered to zero outside of the instrument bandwidth. To avoid ringing (see § 2.4.6), the polynomial is smoothly tapered to zero using a cosine function. The cosines are joined to the polynomial by imposing the boundary conditions that both functions must have the same value and first derivative at the joining point. This results in two linearly independent equations for a cosine function with three variables: amplitude, phase offset, and frequency. Frequency is taken as a free parameter which can be adjusted to control how rapidly the tapper tends to zero. In this work, the taper frequency was chosen to have an equivalent "wavelength" in frequency space of 250 GHz.

The resulting phase function is smooth and can be accurately interpolated onto a more or less dense frequency axis as needed. If an anti-symmetric interferogram is desired, $\pi/2$ can be added to the phase function. This option may be useful for comparing model interferograms to data since the data is somewhat more naturally anti-symmetric. However, symmetric correction is more useful for aperture synthesis as the anti-symmetry would need to be removed before complex visibility extraction anyway. In either case, the phase function is then inverted and mirrored onto the negative frequency axis, preserving the symmetry conditions required for a fully real interferogram.

In principle, the phase function obtained in this way should account for all instrumental

phase effects that cannot be interpreted as resulting from an OPD offset. As such, it should be applicable to any DFI observation. To help validate this assumption, the phase function for a 215 mm observation of the same source was calculated and found to be equivalent below 1.7 THz. Above this frequency, the amplitude spectrum was too low due to visibility effects to calculated accurate phase values.

From the translation property of Fourier transforms (or more explicitly Eq. 2.48),

$$\mathcal{F}\left\{2\int_0^\infty I(k)\cos[2\pi kz+\phi_{inst}(k)]\,dk\right\} = e^{i\phi_{inst}(k)}\,\mathcal{F}\left\{\int_0^\infty I(k)\cos[2\pi kz]\,dk\right\}, \qquad (5.5)$$

where $\phi_{inst}(k)$ represents instrumental phase shift. The work up until now has produced a model for the instrumental phase shift where $\phi_{mod}(k) \approx \phi_{inst}(k)$, at least over the instrument bandwidth. Phase correction is then performed in the frequency domain using

$$e^{-i\phi_{mod}(k)}e^{i\phi_{inst}(k)}\,\mathcal{F}\left\{\int_0^\infty I(k)\cos[2\pi kz]\,dk\right\} \approx \mathcal{F}\left\{\int_0^\infty I(k)\cos[2\pi kz]\,dk\right\}, \qquad (5.6)$$

and by use of the multiplication property of Fourier transforms, phase correction in the interferogram domain is performed using the convolution,

$$\mathcal{F}^{-1}\left\{e^{-i\phi_{mod}(k)}\right\} * \int_0^\infty I(k)\cos[2\pi kz+\phi_{inst}(k)]\,dk \approx \int_0^\infty I(k)\cos[2\pi kz]\,dk\,. \qquad (5.7)$$

The approximate equality is due largely to the fact that the polynomial that models the phase does not have enough degrees of freedom to accurately represent the noise in the phase spectrum.

Phase correction applied in the signal domain is show in Fig. 5.2. The resulting interferogram shows a high degree of symmetry with slight variations corresponding to signal frequencies outside of the instrument bandwidth.

5.2.2 Phase Center

The next step in calibration involves identifying the instrument phase center which represents ZPD for an on-axis source. After the phase correction process, ZPD for an off-axis source is affected by the linear phase terms

$$\begin{aligned}\phi_{offset}(k) &= \phi_0(\vec{B},k) + \phi_{\delta z}(k) + \phi_{res}(k) \\ &= 2\pi k\vec{B}\cdot\vec{\theta} + 2\pi k\,\delta z + 2\pi km\,.\end{aligned} \tag{5.8}$$

The first term on the right-hand side describes the phase shift associated with visibility effects. If the source is extended and symmetric instead of point-like, then $\vec{\theta}$ represents the geometric center of the source. The second and third terms describe the phase shift associated with a systematic OPD offset δz, and any residual linear phase error present in the instrument that was intentionally ignored during the nominal phase correction, respectively. From the phase shift property, each phase term manifests as on offset along the OPD axis,

$$z_{offset} = \vec{B}\cdot\vec{\theta} + \delta z + m\,, \tag{5.9}$$

which notably are all constant with the exception of the first term which is dependent on baseline. By estimating z_{offset} when $\vec{B} = 0$, the constant terms can be measured and removed.

When a source is symmetric, the resulting shape of the DFI response is dominated by temporal coherence effects. With a symmetric phase correction, the response is approximately a sinc function and ZPD for that source can be easily estimated by the peak of the sinc function. Although the DFI response at zero baseline cannot be physically measured, the response at multiple baselines can be measured, and given the linear relationship of Eq. 5.9, regression techniques can be used to estimate the constant term as the y-intercept of a linear equation. This procedure is demonstrated in Fig. 5.4 for a 1 mm slit. It can be observed in the upper panel that as the baseline increases, the peak of the sinc like DFI

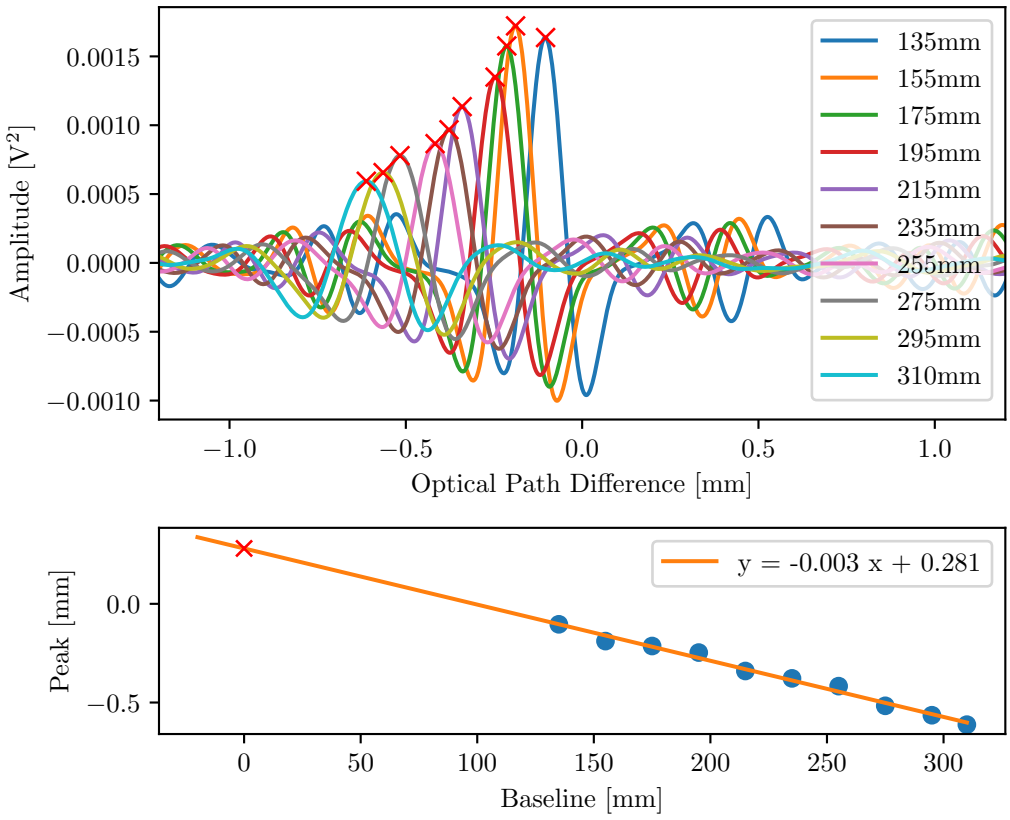


Figure 5.4: Phase center calibration. The top panel shows the phase corrected DFI response obtained at multiple baselines when viewing a 1 mm source. Peaks of the sinc-like functions are identified and plotted as a function of baseline in the lower panel. These data are fit using a linear equation producing a y-intercept of 0.28 ± 0.03 mm. This value is identified as an uncorrected OPD offset in connection with Eq. 5.9, and is subtracted from the OPD axis as part of the calibration procedures.

response moves further away from the phase center. The OPD of the identified peaks is plotted as a function of baseline in the lower panel, showing a linear relationship. The data are fit using a linear equation resulting in a y-intercept of 0.28 ± 0.03 mm. Apparently, initial estimates from geometric modelling was quite accurate, but this analysis has further constrained the phase center down to $\sim 30\mu$m. Applying this correction simply consists of subtracting the offset from the OPD axis that was constructed before the correction was determined.

Repeating this process with the slit at various positions on the source plane results in offsets which are positive and correlated with the distance between the slit and optical axis.

The calculated offset is also roughly equivalent for different detectors when these detectors can see the same slit. This demonstrates a seemingly asymmetric source dependent term not captured by Eq. 5.9. The most compelling explanation is aberrations associated with the collimating and focusing optics which are both asymmetric and source plane dependent (see Fig. 3.14).

5.2.3 Photometric

Intensity

Photometric calibration measurements are part of a complete observation and have been discussed in detail in § 5.1.2. In summary, the source is modulated using a mechanical chopper while one arm of the interferometer is blocked. The resulting peak-to-peak amplitude recorded by the detector gives an estimate of the intensity that propagates through the unobstructed arm of the interferometer

$$I_i = \int_{\Omega} \int_{-\infty}^{\infty} \xi(\vec{\theta}, k)\, I_i(\vec{\theta}, k)\, dk\, d\Omega \,. \tag{5.10}$$

This process is conducted once for the static aperture and at every position of the dynamic aperture that was used when performing a DFI observation. This process is also conducted for both the static and dynamic aperture when performing an FTS observation.

In connection with Eqs. 5.1 – 5.3, the response is normalized by dividing it by $\sqrt{I_1 I_2}$, the geometric mean of the intensity through both arms of the interferometer. This calibration is valid under the conditions that the source plane sensitivity $\xi(\vec{\theta}, k)$ is the same for both arms of the interferometer, and that the spectral intensities through both arms of the interferometer are related to the source spectral intensity by a scalar multiple $I_i(k) = a_i I(k)$. After applying the correction, the FTS response has the form

$$
\begin{aligned}
I_{FTS}(z) &= \frac{\int_{-\infty}^{\infty} \sqrt{I_1(k) I_2(k)} \cos(2\pi k z)\, dk}{\sqrt{\int_{-\infty}^{\infty} I_1(k)\, dk \times \int_{-\infty}^{\infty} I_2(k)\, dk}} \\
&= \frac{\sqrt{a_1 a_2} \int_{-\infty}^{\infty} I(k) \cos(2\pi k z)\, dk}{\sqrt{a_1 a_2} \int_{-\infty}^{\infty} I(k)\, dk} \\
&= \frac{\int_{-\infty}^{\infty} I(k) \cos(2\pi k z)\, dk}{\int_{-\infty}^{\infty} I(k)\, dk}\,.
\end{aligned}
\tag{5.11}
$$

A similar result is obtained for the DFI response,

$$
I_{DFI}(\vec{B}, z) = \frac{\int_{-\infty}^{\infty} I(k) |\gamma_0(\vec{B}, k)| \cos(2\pi k z + \phi_0(\vec{B}, k))\, dk}{\int_{-\infty}^{\infty} I(k)\, dk}\,. \tag{5.12}
$$

This calibration is generally needed since the intensity through the dynamic aperture varies as a function of position resulting from the geometry of the collimating mirror (see Ap. B). If not corrected, this intensity variation would be interpreted as resulting from the amplitude of the complex visibility $|\gamma_0(\vec{B}, k)|$. Furthermore, the intensity through each arm of the interferometer in FTS mode is generally different than in DFI mode for the same reason. If FTS spectra are to be used to normalize DFI measurements, they must be made comparable though this calibration. An additional benefit of this calibration is that it helps to correct for intensity variations due to aliment error or other imperfections in the optical system.

Beam

All the previous calibration procedures are required to extract accurate complex visibility estimates from measured interferograms. The beam correction is applied as the last step of image reconstruction and accounts for variations in the source plane sensitivity of the detector. Effectively, the reconstructed image is divided by $\xi(\vec{\theta}, k)$ which for this work is assumed to be independent of frequency and aperture position. Based on the analysis

presented in Sec. 3.4, these assumptions are reasonably valid for the central detector and potentially its nearest neighbours.

The data used to estimate $\xi(\vec{\theta})$ is the 135 mm baseline measurement shown in Fig. 3.18. For this observation, a 1 mm slit was moved across the source plane in 2 mm increments. The source intensity is then estimated through the modulation procedure summarized previously in this subsection. The intensity curve is then normalized so that its maximum value is one. Since the data is relatively smooth, the result can be interpolated onto whatever source plane axis is required. Due to the somewhat imprecise slit mask mounting solution used in this work, the position of the source on the source plane is only known to ~ 0.5 mm accuracy. Similarly, the position of the phase center on the source plane is not well constrained and has an uncertainty of ~ 2 mm. As such, the beam correction is somewhat dubious, but also has little effect over a ~ 10 mm range where the sensitivity curve is relatively flat.

5.3 Data Processing

5.3.1 Parsing and Averaging

After a full observation is complete, the raw timeline data is parsed into individual scans. This process is facilitated using the PSO signal sent out by the spectral stage and recorded on the detector system's digital channel. The stage is configured to send out a pulse every 0.015 mm of mechanical travel, which corresponds to the sampling interval required to obtain a 2.5 THz optical Nyquist frequency. The pulse itself is a change of state that lasts for a commanded amount of time and is chosen so as to achieve an approximately 50:50 duty cycle when the stage is travelling at a constant 6.7 mm/s. Signal on the digital channel thus appear as a sequence of uniform and closely spaced pulses when the stage is moving, interrupted by extended regions with no pulses when the stage is not moving. The PSO is in a high state by default, so a scan is identified by the first falling edge and the last rising edge within a sequence of regular pulses. Note that this method discards some of the interferogram data, both at the start of the scan before the stage has travelled one

PSO interval, and at the end where there is some uncertainty about whether or not the stage is even moving. Forward scans always precede reverse scans, and by keeping track of the scan number, the scan direction can be later determined by whether it is even or odd. In rare occasions, the detector system encounters an error which breaks the regular pattern of PSO pulses. When this occurs, the scan parsing algorithm may combine or discard scans and some manual intervention may be required to determine which scan numbers correspond to which direction. The problem is not discussed in detail in this work. In general, raw data packets are exported every $\sim 1.2\,\mathrm{s}$. Errors can be identified by deviations from the expected file sizes, and based on when they occur, the corrupted scan number can be estimated. Whether this scan was combined with another scan or discarded can be determined by inspecting the size of resulting scan files and noting whether there are any missing scans. This approach is fairly simple for a small number of errors, which is usually the case since detector system errors are rare.

Once the timeline data is parsed into scans, the forward and reverse scans are averaged separately. To ensure only the constant speed portions of the scans are used, data contained within the first 15 PSO pulses are discarded, and then a constant number of data points afterwards are used. In this way, scans can be easily and accurately averaged as demonstrated in § 3.2.2. Fig. 5.5 shows how the SNR of the interferogram improves with averaging. To get the highest SNR possible with the available data, the average forward and average reverse scans are averaged together. Combining forward and reverse scans can be challenging. As previously discussed, the end of a scan is ambiguous and aligning the end of a forward scan with the start of a reverse scan is unreliable. Forward and reverse scans may not sample the same OPD positions and interpolation of noisy data is generally bad practice. Additionally, the time constant of the detector results in a delay between variations in optical signal and detector signal, and thus perfect alignment with respect to OPD may produce sub-optimal results. To overcome these challenges in a simple and effective way, scan alignment is determined by correlating the interferograms and determining what offset between the two is

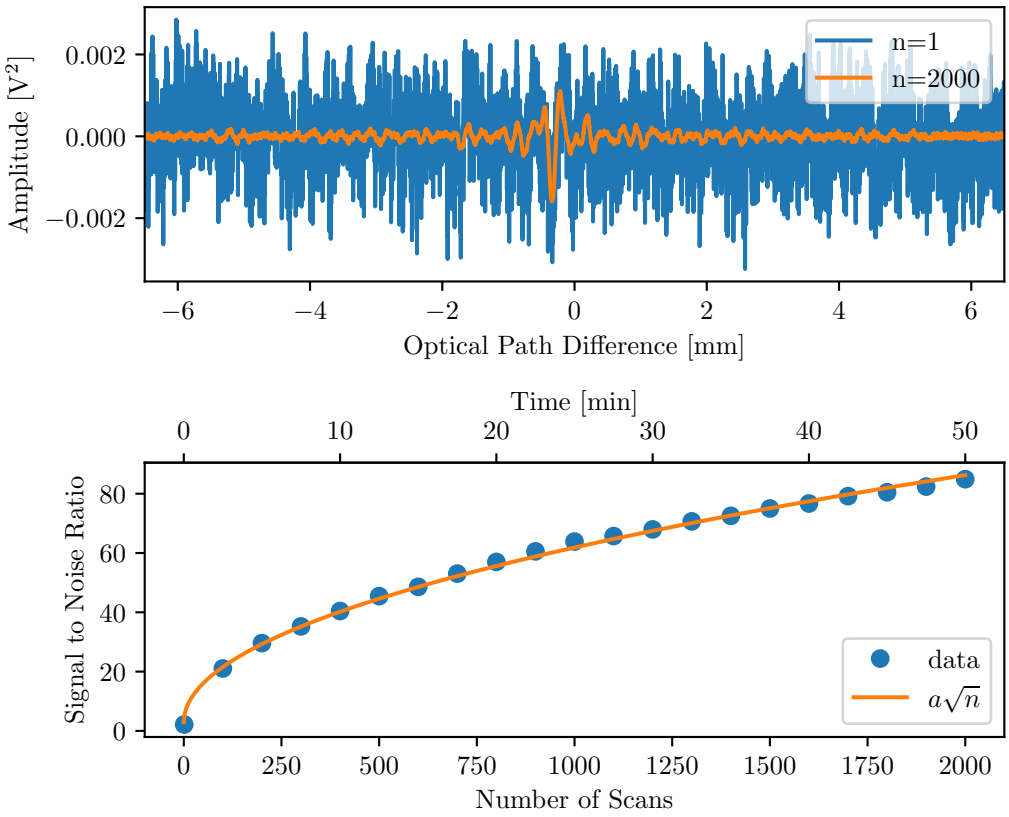


Figure 5.5: Interferogram averaging. The top panel shows an interferogram obtained after one scan (blue), and after averaging 2,000 interferograms (orange). A substantial increase in SNR is evident. The bottom panel shows how SNR varies with n, the number of scans used for averaging. Results show the expected $\sqrt{n}$ scaling, and it appears the trend will continue for greater n, allowing for even more SNR improvements if additional observation time is acceptable.

required to obtain maximum correlation.

Returning to Fig. 5.5, the top panel shows results obtained when averaging 2,000 forward scans. Notably, there is a low amplitude sinusoid that dominates the signal at around 6 mm with a wavelength of 0.06 mm. Converted to stage mechanical path length, this wavelength corresponds to 0.015 mm of stage travel and correlates exactly with the PSO signal. This indicates there is some cross-talk between the digital channel and the detector channels within the detector system. Fortunately, this signal shows up at 2.5 THz in the optical spectrum, as expected, and is sufficiently far away from the 2.2 THz upper limit of the instrument band and can simply be ignored. The bottom panel shows how the SNR varies

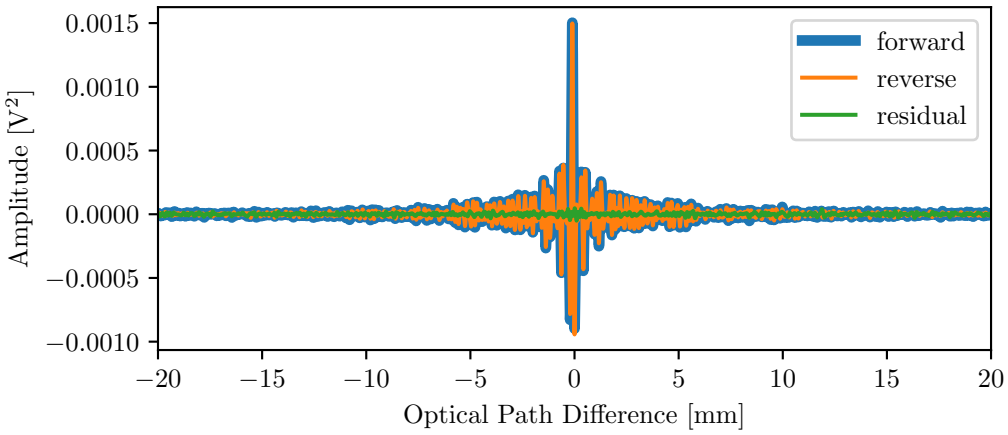


Figure 5.6: Productive scan range test. The blue and orange curves represent FTS interferograms expected to have similar noise properties to that which can be obtained with 2,000 scans in DFI mode. The green curve is the residual, calculated as the difference between the forward and reverse scan, and represents the noise present in the interferograms. Interferograms can be recorded productively while the signal is above the noise. From inspection, the orange signal is obscured by the green noise at around ±15 mm. As such, it is likely that the DFI scan range can be productively extended to this limit when using 2,000 scans.

with the number of scans used and expresses a $\sqrt{n}$ scaling that is expected for a signal that is constant in time with ergodic noise. The amplitude in the SNR calculation is constant and determined by the peak-to-peak value of the central interferogram feature that results from averaging all forward and reverse scans. The noise of the SNR is calculated as the standard deviation of the difference between forward and reverse scans for a particular number of scans averaged. This analysis was performed with data obtained from a 135 mm baseline viewing a 1 mm source. At longer baselines, and for more extended sources, the visibility of fringes are generally lower and the amplitude is decreased. That is, the analysis presented here is nearly a best case scenario, and under different conditions, the SNR is expected to be worse.

The nominal scan range of about ±6.5 mm OPD was primarily chosen to reduce observation time. Indeed, almost an hour of integration time is required to achieve the modest SNR signal shown in Fig. 5.5, and a full DFI observation for one baseline orientation can take >14 hrs, which does not included processing time. However, it is instructive to

consider what maximum scan range can be used productively with $2,000$ averaged scans. Fig. 5.6 shows an FTS interferogram with noise properties comparable to what can be obtained with $2,000$ DFI scans. Forward and reverse scans are shown and have been subjected to a band-pass filter to remove out of band noise and the spurious PSO component. The residual is also shown, and is calculated as the difference between forward and reverse scans. The residual is a good representation of the noise present in the interferograms, and the data is considered useful if the interferogram signal is greater than the noise. After this point, mostly noise is contributed to the spectrum, degrading the spectrum quality. From inspection, it is reasonable to conclude that the DFI scan range can be productively extended to around $\pm 15\,\mathrm{mm}$, where the orange curve is fully obscured by the green curve, when averaging $2,000$ scans.

After obtaining the average interferogram, the phase, photometric, and ZPD corrections are applied, as discussed in Sec. 5.2. The order in which these corrections are applied is not important.

5.3.2 Complex Visibility Extraction

The main observable quantity of interest when using the DFI instrument in this work is the complex visibility. Note that for a more practical space interferometer, the Mutual Coherence Function (MCF), which is the non-normalized complex visibility, may be desired (see Ap. A and § 2.5.2). The MCF preserves the relative intensity differences of the source intensity distribution between different spectral components. However, if the source is spectrally uniform, spectral components can be normalized and combined in a way that extends *uv*-plane coverage resulting in higher quality image reconstruction at the cost of spectral imaging. The DFI instrument used in this work is quite limited in its *uv* coverage due to the economical size of the collimating mirror, and working with complex visibilities is generally required. However, in some cases, spectral imaging can still be demonstrated, albeit with lower spectral resolution than nominally provided by the DFI (see § 6.3.2).

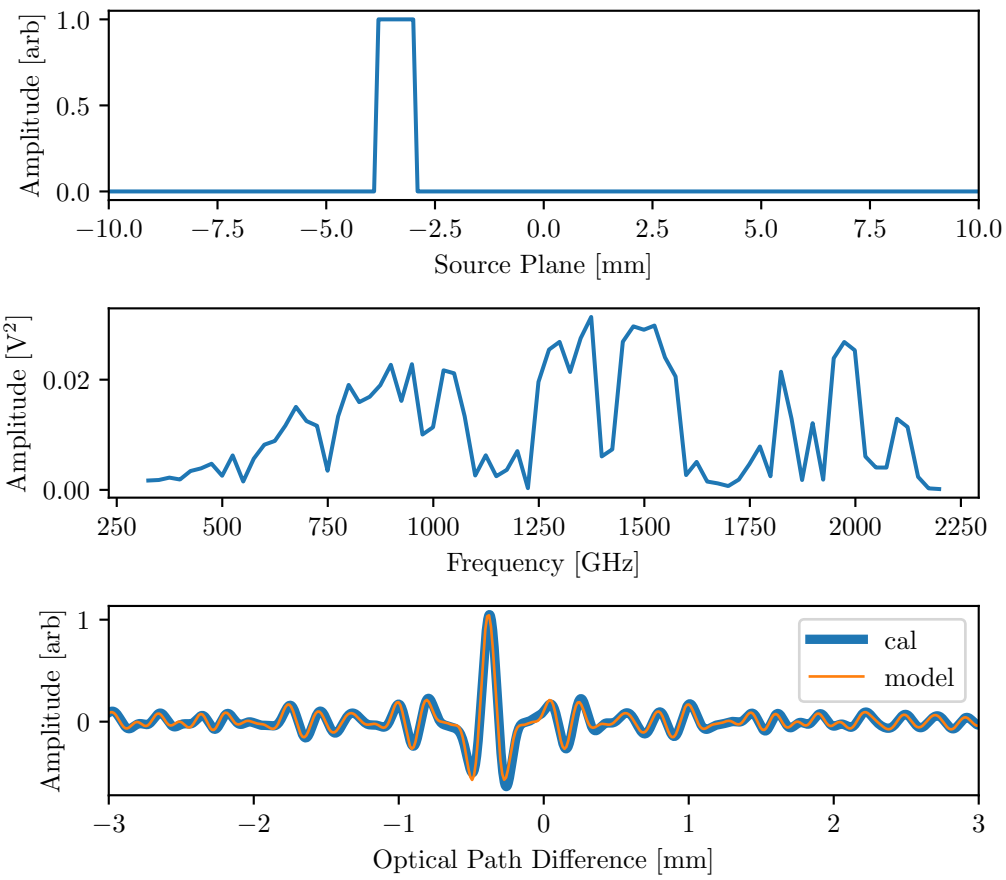


Figure 5.7: Interferogram simulation for an off-axis ~ 1 mm slit illuminated by a $1,100\,^{\circ}$C thermal source. The top panel shows the source geometry modelled as a 0.7 mm slit with a -3.4 mm offset. The middle panel is the spectrum resulting from an FTS measurement of the 1 mm aperture mask back-lit by the Pegasus thermal source. The orange curve in the bottom panel is the simulated DFI interferogram for a 135 mm baseline. The blue curve is the observed interferogram under the same conditions after phase correction and band-pass filtering.

For the remainder of this chapter, processing will be demonstrated with reference to a test case dataset obtained when observing a 1 mm slit offset from the optical axis by ~ 3 mm. To provide confidence in the processed results, data is presented alongside simulated results for the same source geometry using ideal, noise free, models. An explicit expression for computing interferograms which include the spatial and spectral structure of the source is given by Eq. B.10. An example evaluation is shown in Fig. 5.7 which demonstrates excellent agreement between model and data.

After applying the nonlinear phase, photometric, and phase center corrections, the DFI

response can be written as,

$$I_{DFI}(\vec{B},z) = A\int_{-\infty}^{\infty} I(k)|\gamma_0(\vec{B},k)|\cos(2\pi kz+\phi_0(\vec{B},k))\,dk \tag{5.13}$$

Similarly, the FTS response can be written as,

$$I_{FTS}(z) = A\int_{-\infty}^{\infty} I(k)\cos(2\pi kz+\phi_{FTS}(k))\,dk\,, \tag{5.14}$$

where $A = 1/\int_{-\infty}^{\infty} I(k)\,dk$ is the reciprocal of the integrated spectral intensity (see Eq. 5.11 and 5.12). Taking the Fourier transform of Eqs. 5.13 and 5.14 gives,

$$\mathcal{F}\left\{I_{DFI}(\vec{B},z)\right\} = \frac{AI(k)}{2}|\gamma_0(\vec{B},k)|e^{i\phi_0(\vec{B},k)}\,, \tag{5.15}$$

and

$$|\mathcal{F}\left\{I_{FTS}(z)\right\}| = \frac{AI(k)}{2}\,, \tag{5.16}$$

where the delta functions associated with the cosine functions have been suppressed for simplicity but without loss of accuracy. The magnitude of the FTS transform is computed to remove the FTS phase (see Eq. 2.51). FTS phase can also be removed through a phase calibration procedure like the one described in § 5.2.1, which can reduce the noise in the FTS spectrum by a factor of $\sqrt{2}$. However, interferograms recorded in FTS mode are typically of significantly higher SNR than DFI mode interferograms due to the greater number of scans averaged and the absence of reduced fringe visibility from spatial coherence effects. Since the FTS spectrum will be used to normalize the DFI spectrum, noise in the resulting complex visibility is dominated by noise in the DFI spectrum. Thus, the modest noise reduction that comes from a more detailed FTS phase correction renders the process unnecessary. The complex and absolute value of the spectra obtained according to Eqs. 5.15 and 5.16 when applied to the test case dataset are shown in Fig. 5.8. The decrease in DFI spectral amplitude compared to the FTS spectral amplitude at higher frequencies results

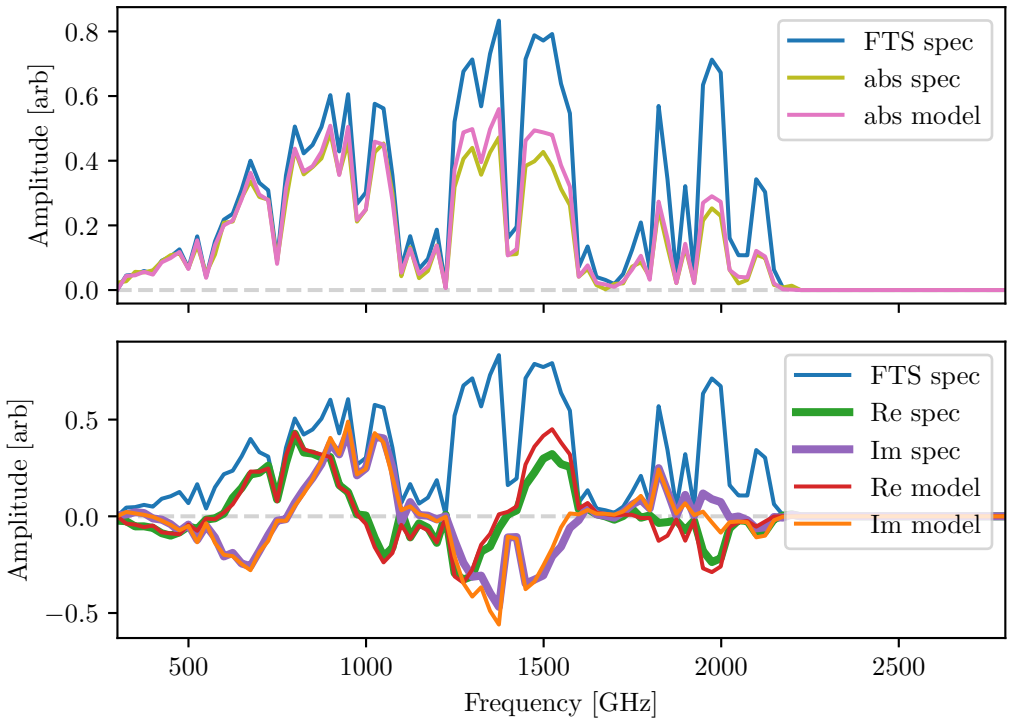


Figure 5.8: Amplitude and complex spectrum for the test case dataset. The top panel shows the FTS spectrum (blue), and the absolute value of the DFI spectrum using a 135 mm baseline. Experimental results are shown in yellow, with simulated results shown in pink. The difference in amplitude between the FTS and DFI spectrum is a result of decreased fringe visibility at higher frequencies. The bottom panel shows the same fully real FTS spectrum, in addition to the complex DFI spectrum obtained from the Fourier transform of the interferogram in Fig. 5.7. Real and imaginary experimental results are shown in green and purple, respectively. Real and imaginary simulated results are shown in red and orange, respectively. The complex nature of the DFI spectrum results from the complex visibility phase.

from decreased fringe visibility due to spatial coherence effects. The complex nature of the DFI spectrum results from the complex visibility phase.

The complex visibility is recovered as the ratio Eqs. 5.15 and 5.16,

$$\gamma_0(\vec{B},k) = |\gamma_0(\vec{B},k)|e^{i\phi_0(\vec{B},k)} = \frac{\mathcal{F}\left\{I_{DFI}(\vec{B},z)\right\}}{|\mathcal{F}\left\{I_{FTS}(z)\right\}|}\,. \tag{5.17}$$

The result of this process applied to the testcase dataset is shown in Fig. 5.9. It is worth emphasizing that this process results in simultaneous measurements of complex visibility

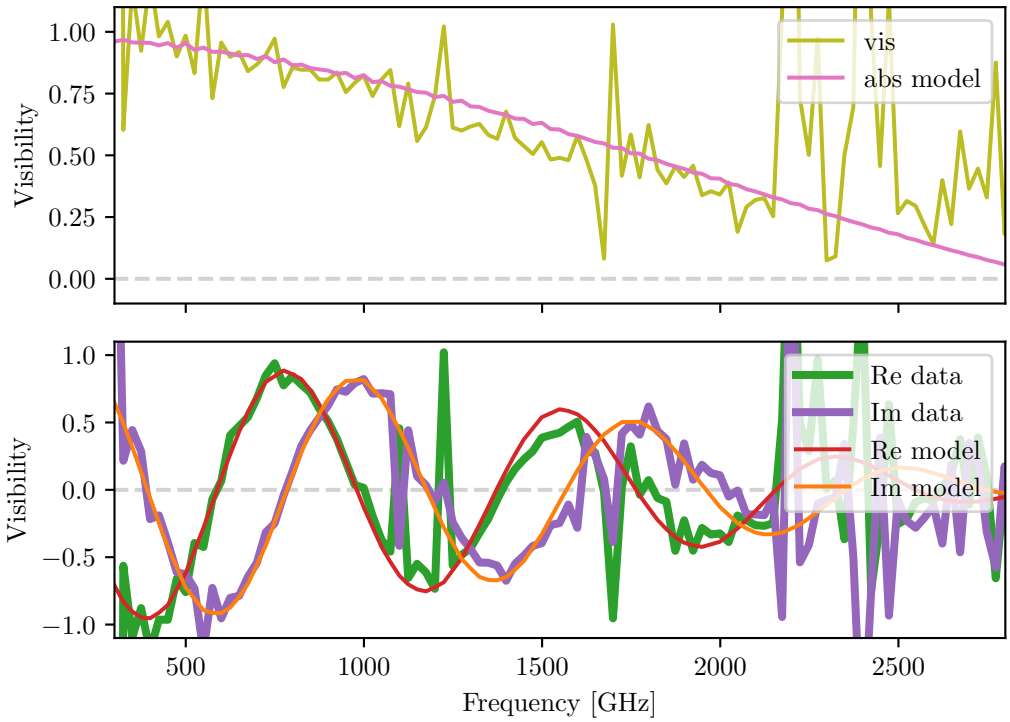


Figure 5.9: Frequency dependent complex visibility. The bottom panel shows the complex visibilities obtained as the ratio of the complex DFI and real FTS spectrum from the bottom panel of Fig. 5.8. The top panel shows the absolute value of the complex visibilities from the bottom panel, which can be equivalently computed as the ratio of the absolute value of the DFI spectrum and the FTS spectrum shown in the top panel of Fig. 5.8. The color scheme used in this figure is the same as that used in Fig. 5.8.

at each frequency over the entire bandwidth of the instrument. This is the unique aspect of the DFI method that distinguishes it from traditional methods used in radio astronomy. The frequency dependent visibility measurements over a broad spectral range allows for the production of high spatial resolution spectral images.

5.3.3 Aperture Synthesis

Once complex visibilities are obtained, the subsequent processing steps required to generate images follow the well developed procedures developed in the context of radio astronomy [126]. The first step is to project the frequency dependent visibilities onto the so called

"*uv*-plane" using the mapping

$$\begin{aligned} \frac{B_x}{\lambda} &= \frac{|\vec{B}|\cos(\theta_{\vec{B}})}{\lambda} \rightarrow u \\ \frac{B_y}{\lambda} &= \frac{|\vec{B}|\sin(\theta_{\vec{B}})}{\lambda} \rightarrow v \,, \end{aligned} \tag{5.18}$$

where $\theta_{\vec{B}}$ is the rotation of the baseline vector $\vec{B}$, and λ is the wavelength of light. The *uv* coordinates describe how many wavelengths can fit along the orthogonal components of the baseline vector B_x and B_y, thus the *uv*-plane can be thought of as a space defined in terms of spatial frequencies. When the source has an intensity distribution that is frequency dependent, a separate *uv*-plane should be constructed for each frequency. If the source is spectrally uniform within some bandwidth, the *uv* measurements from all frequencies with in that bandwidth may be combined into a single *uv*-plane. An example of this procedure is shown in Fig. 5.10 for the test case observation. In order to exclude unreliable measurements, complex visibilities are only used when their associated amplitude in the FTS spectrum is $\geq 5\%$ of the maximum value of the FTS spectrum. This requirement avoids both low SNR data, and data that may be divergent due to the low value of the denominator when using the FTS spectrum for normalization. The intuition behind aperture synthesis is that an arbitrary two dimensional signal can be constructed from a superposition of two-dimensional sinusoids with various amplitudes, spatial frequencies, and phases, in the same way an arbitrary one dimensional signal can be constructed using a Fourier series (see Eq. 2.5). Once the *uv*-plane is populated, the source intensity distribution is recovered through Fourier inversion.

Complex visibility distributions in the *uv*-plane will be denoted by $V(u,v)$. In practice, this distribution is sampled discretely based on the sampled baselines and optical frequencies. A general sampling function can be written as

$$S(u,v) = \sum_{n=1}^{N} \delta(u-u_n, v-v_n) \,, \tag{5.19}$$

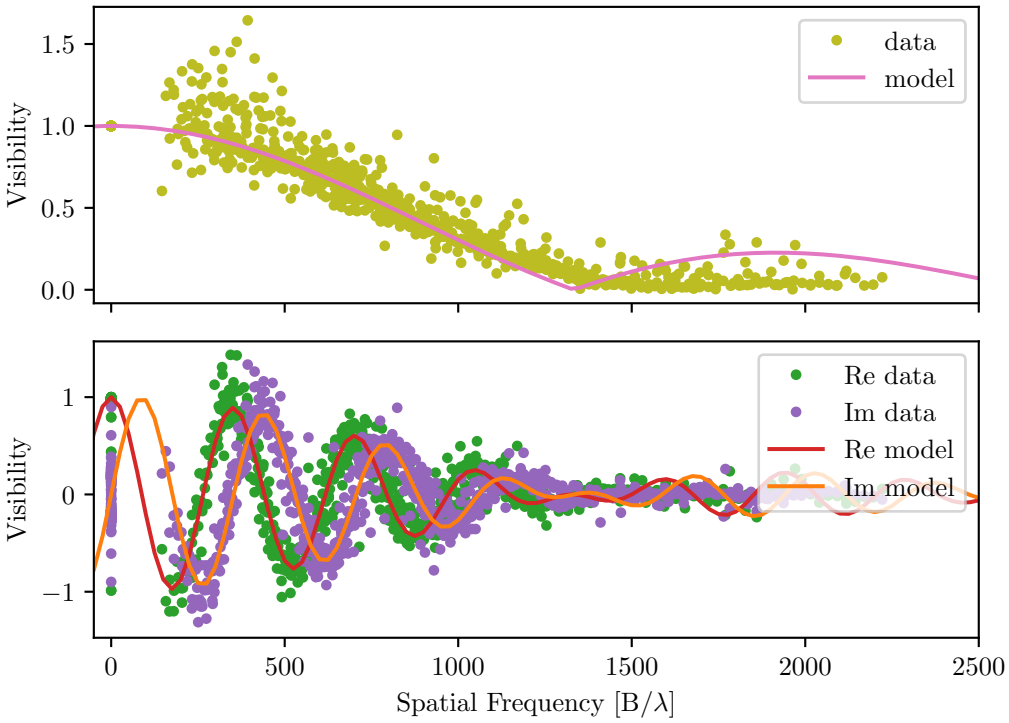


Figure 5.10: Complex visibility samples. This figure represents a one dimensional slice of the *uv*-plane and is populated by *uv* samples from multiple frequencies using the mapping of Eq. 5.18. To ensure high quality samples, visibilities are only taken from frequencies with corresponding amplitudes in the FTS spectrum that are $\geq 5\%$ of the maximum amplitude in the FTS spectrum. The bottom panel shows complex visibilities, and the top panels shows the absolute value of these complex visibilities. Note that the complex visibilities express hermitian symmetry: negative spatial frequencies are obtained as the complex conjugate of the positive spatial frequencies, or equivalently, the real components are symmetric while the imaginary components are anti-symmetric. The color scheme used in this figure is the same as that used in Fig. 5.8.

where (u_n, v_n) are sampled *uv* coordinates, which, when applied to the visibility distribution, produce a set of sampled visibilities

$$V_s(u_n, v_n) = S(u,v)V(u,v) \ . \tag{5.20}$$

Rarely, if ever, are these samples uniformly distributed on a rectangular gird. In most cases the sampling is sparse with significant density variations within the *uv*-plane. As such, the Fast Fourier Transform (FFT) cannot be used to perform Fourier inversion. Instead, the direct Fourier transform must be computed though brute force evaluation of

$$I_{DFT}(\theta_x, \theta_y) = \frac{1}{N} \sum_{n=1}^{N} V_s(u_n, v_n) e^{-i2\pi(u_n\theta_x + v_n\theta_y)} \ , \tag{5.21}$$

where (θ_x, θ_y) are arbitrary angular coordinates in the source plane. If the N samples in the *uv*-plane are transformed into N samples in the reciprocal domain, the computational complexity of Eq. 5.21 scales as N^4. This task rapidly becomes prohibitively expensive with increasing N and a faster method that approximates I_{DFT} is desirable.

Every strategy for efficient computation of Eq. 5.21 attempts to leverage the FFT. The most common approach is to construct a regularly sampled rectangular grid from the irregularly sampled data. Interpolation is an obvious option, but this does not take advantage of the statistical benefits of the densely sampled regions of the *uv*-plane, and is sensitive to noise. Another approach that overcomes these short comings is the convolution method.

As part of the convolution method, the sampled visibilities may be multiplied by various weighting factors. These weights may account for effects like the reliability of the sample, the local sampling density, or some enforced taper function. The purpose of these weights is to obtain a synthesized beam with desirable side-lobe characteristics. The tapering weight, for example, serves the same purpose as an apodization function in FTS (see 2.4.6). The local density weighting, calculated as the inverse of the number of sample points within some neighbourhood, is an attractive option when the sampling density varies slowly, but

can result in discontinuities otherwise. In this work, the visibility samples are not weighted.

The discrete convolution operation is written as,

$$C * VS = \sum_{n=1}^{N} C(u_n - u, v_n - v) V_s(u_n, v_n) \, , \tag{5.22}$$

where C is some convolution kernel. A suitable kernel should be symmetric and tend to zero away from its origin. In this work, a Gaussian function is used,

$$C_{Gauss}(u,v) = \exp\left[-\frac{1}{2\sigma^2} \left[(u-u_0)^2 + (v-v_0)^2 \right] \right] \, , \tag{5.23}$$

which is centered at (u_0, v_0) with a standard deviation σ along both orthogonal axes. In practice, convolution is only evaluated at the uniformly distributed resampling positions. The resampling function is given by

$$R(u,v) = \sum_{i=-M/2}^{M/2} \sum_{j=-M/2}^{M/2} \delta(u - i\Delta u, v - j\Delta v) \tag{5.24}$$

where i, j are evaluated over the desired M number of points in the resampled grid. The pixel size in the uv-plane is defined by Δu and Δv. When choosing M, Δu, and Δv, it is important to recall the sampling conditions of DFTs (see § 2.3.2). In particular, $\Delta\theta = 1/M\Delta u$ and $\pm\theta_{max} = 1/2\Delta u$. Applying Eq. 5.24 to Eq. 5.22 produces the regridding by convolution equation,

$$V_r(u_m, v_m) = R \times (C * SV) \, , \tag{5.25}$$

where (u_m, v_m) are uniformly spaced coordinates in the uv-plane obtained from the delta functions centered at $(i\Delta u, j\Delta v)$ from Eq. 5.24. In practice, the convolution is evaluated only at the resampled coordinates and is calculated as a weighted sum of the surrounding irregular points. This abstracted process is represented visually in Fig. 5.11. Usually the calculation is limited to the irregular points within a small neighbourhood around the resampled point which reduces computational complexity. There is no generally accepted rule for

determining the width of the convolution kernel. In this work, the Gaussian kernel width is constrained to $\sigma = \Delta u = \Delta v$, and is chosen so that the convolution produces a smooth function. From testing various options, a value of $\sigma = 25$ was chosen and consistently applied to all observations.

Now that visibilities have been projected onto a regular grid, the inverse Fourier transform can be computed efficiently with the inverse FFT which gives,

$$\tilde{I}(\theta_x, \theta_y) = \mathcal{F}^{-1}\{R\} * \left[\mathcal{F}^{-1}\{C\}\left(\mathcal{F}^{-1}\{S\} * \mathcal{F}^{-1}\{V\}\right)\right] , \tag{5.26}$$

where the multiplication/convolution property of Fourier transforms has been used. The resampling function makes $\tilde{I}$ periodic as described in Sec. 2.3, which is why Δu and Δv must be chosen carefully to avoid aliasing. Generally they should be selected such that the resulting $\pm\theta_{max}$ is comfortably larger than the instrument Field of View (FOV). Eq. 5.26 is a quantity referred to as the dirty image which may express significant distortion due to sparse sampling in the *uv*-plane. It is further distorted by the convolution process, which is easily compensated by dividing by the inverse transform of the convolution kernel $c = \mathcal{F}^{-1}\{C\}$,

$$\tilde{I}_c(\theta_x, \theta_y) = \frac{\tilde{I}}{c} . \tag{5.27}$$

This calculation diverges at points where $c \to 0$, as such, it is customary to set low values to one, and simply ignore image points within this region. Note that the inverse transform of the kernel is subject to the scaling property of Fourier transforms. If the kernel is too wide, its transform will tend to zero rapidly and limit the FOV of the reconstructed image.

The dirty beam is obtained in the same way described by the preceding discussion but with $V_s(u_n, v_n) = 1$, which is simply the sampling function S,

$$\tilde{b}(\theta_x, \theta_y) = \mathcal{F}^{-1}\{R\} * \left[\mathcal{F}^{-1}\{C\}\,\mathcal{F}^{-1}\{S\}\right] , \tag{5.28}$$

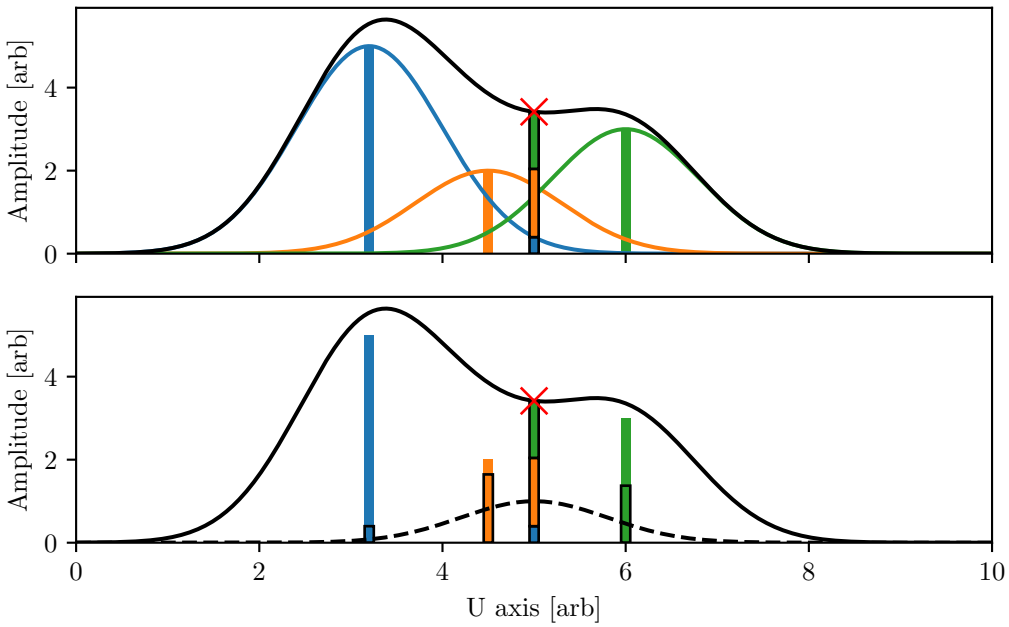


Figure 5.11: Regridding by convolution algorithm in one-dimension. Three irregularly sampled visibility points SV, real or imaginary, are represented by vertical bars. The top panel shows the convolution process as it is typically understood where the irregular samples are multiplied by a convolution kernel (Gaussian in this case), and the sum of these profiles gives the full convolution (black curve; $C * SV$). Black outlined vertical bars indicate the contribution to the convolution from the various irregular samples. This strategy is not easily performed on irregular data since irregular positions are not generally sampled by the regular convolution axis. The bottom panel shows an alternative approach, where a Gaussian with unit amplitude (dashed curve) is centered on the resampled coordinate. The height of the Gaussian where it intersects with irregular samples, which can be computed analytically, is multiplied by the heights of the irregular samples resulting in the black outlined vertical bars. The sum of the heights of these outlined bars give the convolution value at the resampled coordinate $R(C * SV)$. The process is thus equivalent to a weighted sum of the irregular points with weights obtained from the convolution kernel.

followed by the same kernel correction,

$$\tilde{b}_c(\theta_x, \theta_y) = \frac{\tilde{b}}{c} \,. \tag{5.29}$$

Additionally, the dirty beam is normalized by dividing it by its maximum value. This same value is used to normalized the dirty image. After this step, the dirty image is a close approximation of the result that would be obtained using the direct Fourier transform Eq. 5.21. Fig. 5.12 provides a visual summary of the regridding by convolution procedure described in this section.

5.3.4 Deconvolution

The distortions in the dirty image result from incomplete *uv*-plane sampling. Sampling in the *uv*-plane can be modelled as multiplying the fully sampled *uv*-plane by a masking function which is one at sampled points and zero everywhere else (i.e., the sampling function S, from Eq. 5.19). In the reciprocal image plane, this procedure corresponds to a convolution of the source with the inverse Fourier transform of the masking function. This concept is discussed in more detail in § 2.5.1. The inverse transform of the masking function gives the dirty beam, so the dirty image is a convolution of the dirty beam with the source. Removing distortions from the dirty image is then a deconvolution problem.

There are two classes of deconvolution algorithms commonly used in radio astronomy. The first employs the maximum entropy method [127, 103], which casts the problem in terms of an optimization problem. Many variations exit, but all such algorithms attempt to fill in missing *uv* samples while minimally affecting the sampled points. Missing data is inserted in a way that maximizes "entropy," which is not related to physical or information entropy. The resulting clean image is regarded as the most probably intensity distribution that is consistent with measured data. This optimization conspires to produce images that are maximally smooth, and as such, produces poor results for images with point sources that result in discontinuous variations. Conversely, this technique works particularly well

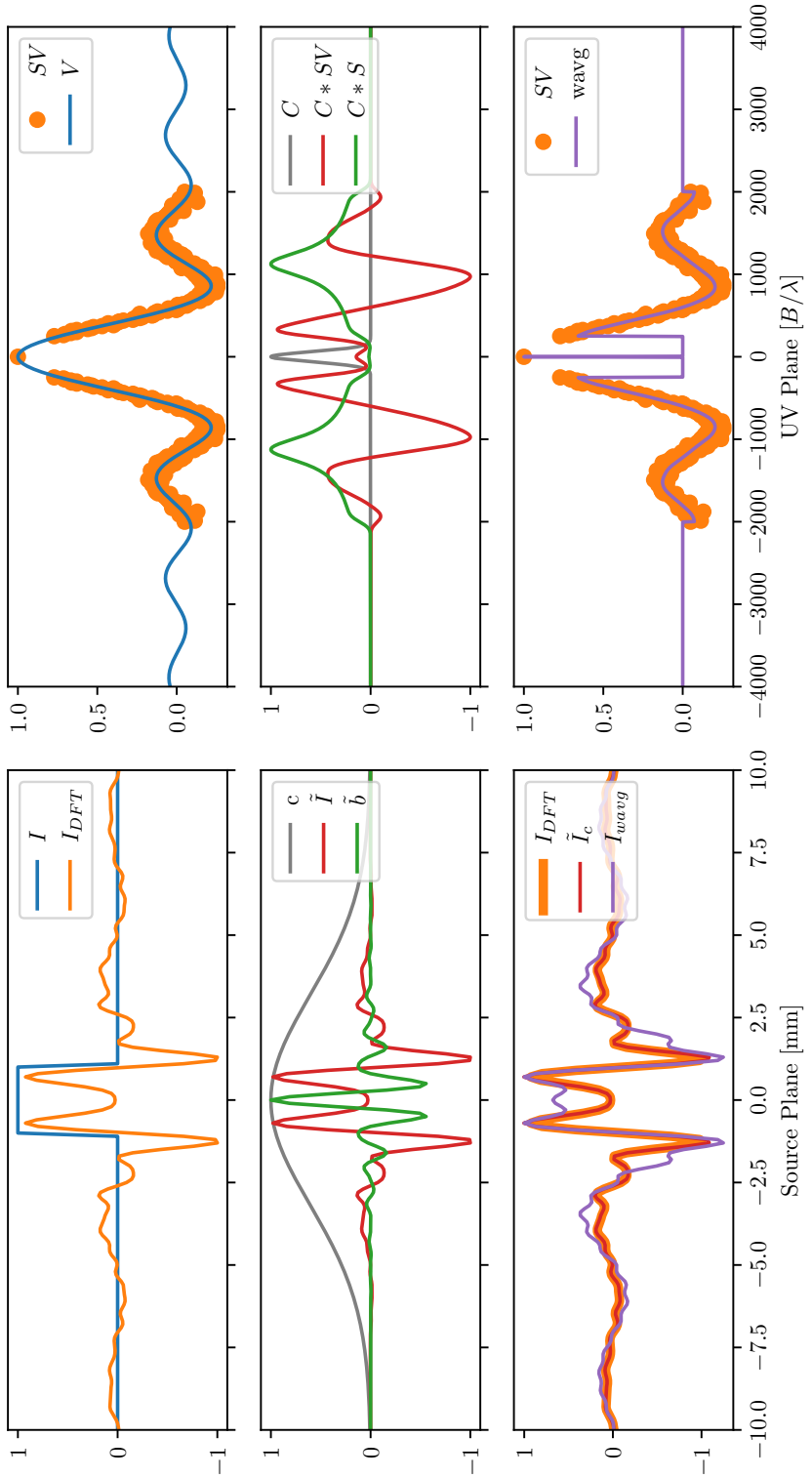


Figure 5.12: Gridding by convolution demonstration for a one-dimensional source. The source with intensity distribution I is real and symmetric so that its Fourier transform is real and symmetric and only the real component of the complex visibility V need be considered. The left column shows source plane quantities, and the right column shows uv-plane quantities. First row: direct Fourier transform I_{DFT} of irregular samples SV. Second row: uncorrected dirty image $\tilde{I}$, dirty beam $\tilde{b}$ resulting from the inverse transform of the convolution of the kernel C with the irregular samples and sampling function, respectively. Third row: corrected dirty image $\tilde{I}_c$ obtained by dividing $\tilde{I}$ by the inverse transform of the kernel c, compared to I_{DFT}. The inverse transform I_{wavg} of weighed average $wavg = (C * SV)/(C * S)$, which closely matches V, is also shown for comparison.

on extended sources and tends to converge on a final result faster than the alternative deconvolution class when the number of *uv* samples is large. The maximum entropy method can incorporate a priori information, like a low resolution image, which helps bias the optimization towards a particular result. Additionally, the algorithm enforces positive flux in the image domain which imposes a real physical constraint on the clean image.

An alternative approach is the "clean" method [102, 128], which employs a more intuitive strategy. The peak intensity and its associated coordinates are identified within the dirty image. A scaled version of the dirty beam, centered on this coordinate, is subtracted from the dirty image, where the scaling factor is the peak intensity multiplied by some gain. The value of the scaling factor and the coordinate of peak intensity is recorded for future use, and is called a "clean component." This process is repeated until a stopping condition is reached which may be a maximum number of iterations or when the peak intensity in the residual dirty image is below some threshold. When the stopping condition is met, clean components are added to an empty image plane and convolved with a synthetic clean beam. The residual dirty image is then added to the convolved clean components producing the final clean image. Typically, the clean beam is constructed by fitting a Gaussian to the main lobe of the dirty beam. The clean method performs particularly well on images with point-like sources and is generally easier to implement than the maximum entropy method. The clean method also converges faster than alternative deconvolution methods when the image sizes are small. For these reasons, the clean method is used in this work. All deconvolution algorithms are sensitive to their input parameters, and there is no consensus for determining optimal parameters. As such, the gain and number of iterations is selected on a case by case basis with the goal of minimizing sidelobes in the reconstructed image. An example of the clean method applied to the test case dataset is shown in Fig. 5.13.

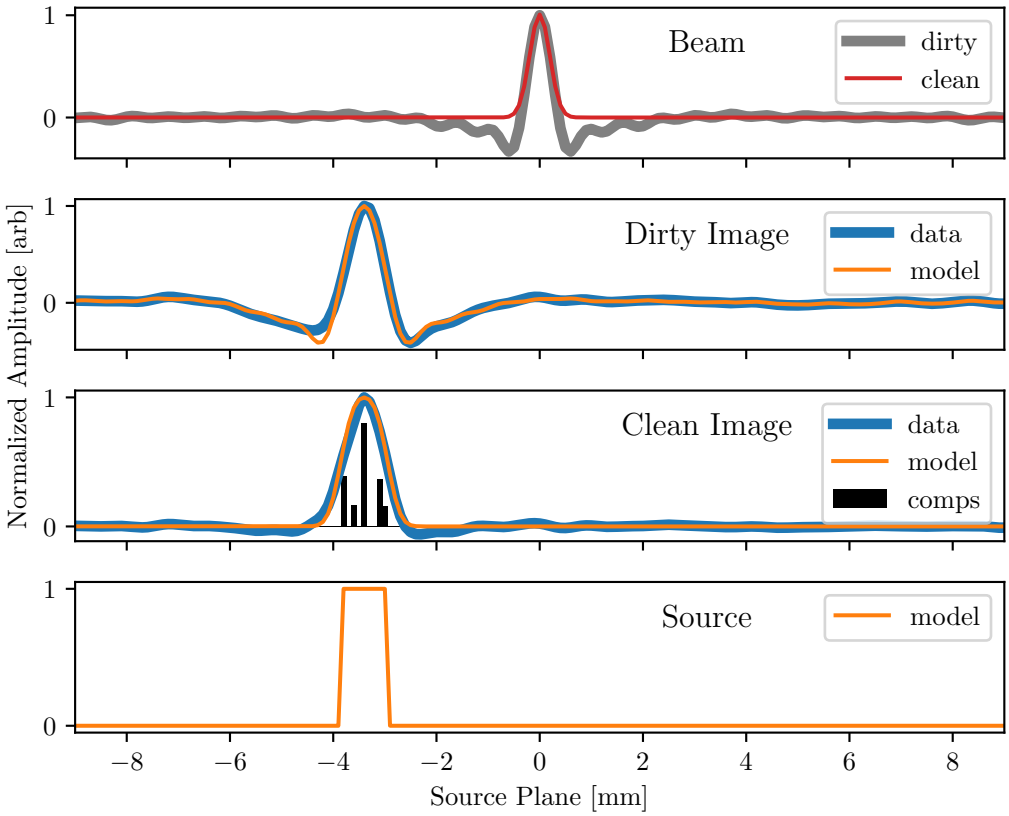


Figure 5.13: Demonstration of image cleaning. The top panel shows the dirty (gray; Eq. 5.29) and synthesized clean beam (red) generated by fitting a Gaussian to the top 50% for of the central lobe of the dirty beam. The second panel shows the dirty image resulting from data processing (blue) and a simulated dirty image (orange) obtained by convolving the dirty beam with the model source. The third panel shows clean components (black bars) resulting from the cleaning process, which when convolved with the clean beam and added to the clean residual, produce the clean image (blue). A simulated clean image (orange), produced by convolving the model source with the clean beam, demonstrates the best case scenario when using the clean method. The bottom panel shows the model source which the clean image is meant to approximate.

5.4 Summary

In this chapter, the details of performing testbed DFI observations and data processing were presented. The discussion on observations included information on the initial setup of the instrument in order to obtain the best results. The core measurements include DFI interferograms recorded at multiple baselines and FTS interferograms required for normalization and the extraction of complex visibility samples. Thousands of interferograms were averaged in order to bring the SNR up to a reasonable value and a typical observation took in excess of 20 hours, conducted over two working days, not including processing time.

In order to fully leverage the dataset, photometric observations were required to calibrate the intensity through each arm of the interferometer. If not corrected, baseline dependent intensity variations would be indistinguishable from variations in fringe visibility and would produce spurious visibility measurements. It is also necessary to correct for instrumental phase errors resulting from dispersion in the optical components and non-linear effects of the detector electronics. To this end, a standard non-linear phase correction product was obtained and applied to all future DFI measurements. The residual linear phase error was corrected by observing the shift of interferograms along the OPD axis, as a function of baseline, associated with a small symmetric off-axis source. The OPD shift at zero baseline can be identified as a residual linear phase error. When this shift is subtracted from the OPD axis, this point corresponds to the phase center of the instrument. A beam correction was also discussed, which compensates for source plane dependent detector sensitivity and can attenuate source intensity at the edges of the beam.

After the calibrations procedures, complex visibility extraction was discussed and demonstrated. A unique feature of the DFI technique is that the instrument is sensitive to a broad bandwidth, and complex visibilities associated with each sampled frequency can be obtained. This allows for image synthesis at each sampled frequency, resulting in the production of spectral images with enhanced spatial resolution. However, the limited baseline coverage of the testbed instrument required complex visibility samples from multiple fre-

quencies to be combined in order to extend *uv* coverage and produce higher quality images. This process was demonstrated, and the conditions under which such a process is valid were discussed. Once complex visibilities are obtained, the well developed techniques of aperture synthesis can be employed to reconstruct an image of the source plane. One challenge involves projecting the generally non-uniformly sampled visibility data onto a uniformly sampled grid so that fast computation of the inverse Fourier transform is possible. The regridding through convolution technique, used in this work, was explained and demonstrated. Due to the generally sparse sampling of the *uv*-plane, reconstructed images are often distorted. Removing these distortion is effectively a deconvolution problem. Deconvolution strategies were discussed, and the "clean" method used in this work was demonstrated. As such, this chapter presented an end-to-end description for the acquisition and processing of DFI data for the production of high spatial resolution images, and was supported by examples with experimental data.

Chapter 6

Results

In this chapter, the results from various Double-Fourier Interferometer (DFI) experiments are presented. The measurements in Sec. 6.1 demonstrate how source extent affects the visibility of fringes when observed at different baseline lengths. It is also shown how source extent can be inferred from visibility curves when some assumptions can be made about the source geometry. The measurements in Sec. 6.2 demonstrate how the phase of the complex visibility changes as a function of baseline and source position. These results can also be used to validate the spectrometer accuracy of the testbed DFI system. Taken together, these experiments independently probe the two aspects of complex visibility: amplitude and phase, and are intended to give a clear demonstration of how these quantities relate to source structure.

Subsequent experiments exploit complex visibility as a whole and are oriented towards demonstrating image reconstruction from interferometric measurements. These experiments are intended to independently verify the various imaging requirements outlined in Sec. 1.4. Sec. 6.3 includes observations of one dimensional sources which are further divided into three categories. The first category is spectrally uniform sources imaged using a single detector, which verifies that aperture synthesis techniques can be applied to DFI data, and that the DFI instrument is capable of resolving structure much smaller than the size of the detector beam. The second category is nonuniform spectral sources, where the source has a different structure at different optical bandwidths. By performing aperture synthesis within these bands separately, it is shown that the DFI technique is capable of

providing high spatial resolution spectral images. The third category involves source structure that is extended such that the full source can only be observed when using multiple detectors. These are the conditions that describe wide field imaging. Multi detector aperture synthesis is demonstrated, and experimental results provide insight into the processing steps required to exploit multiple detectors. Finally, Sec. 6.4 demonstrates aperture synthesis when observing a two dimensional source, which requires measurements from multiple baseline/source orientations. The details of how to fill the *uv*-plane when using arbitrary baseline lengths, orientations, and telescope pointings are discussed, in addition to how this process is extended to wide field applications.

6.1 Photomixer Observations

The photomixer source, described in § 3.2.4, provides a convenient target for studying the effects of source extent on fringe visibility. Due to the monochromatic nature of the emitted radiation, the resulting interference pattern is well defined by a sinusoid and removes a layer of complexity associated with broadband sources.

For these observations, the divergence point of the photomixer was placed at foal point of the collimating mirror. The back focal point of the photomixer emitter is not known with high precision, and as discussed in § 3.2.5, the focal point of the collimating mirror is only constrained to $\pm 2\,\mathrm{mm}$. To achieve the best results, the emitter was manually adjusted so as to maximize the sum, and minimize the difference, of the intensity through both arms of the interferometer. Intensity was estimated by electronically modulating the photomixer output and measuring the peak-to-peak amplitude of the the resulting signal when either arm of the interferometer was blocked. Position was iteratively adjusted at different dynamic aperture positions until a position was found that resulted in strong, and nearly equivalent, signal through the static aperture and all position of the dynamic aperture. Unfortunately, this result did not hold over a broad range of optical frequencies, suggesting that the back focal point and/or beam characteristics of the emitter are frequency dependent. For this

reason, observations were carried out in the frequency range $\sim 840 - 860\,\mathrm{GHz}$, at $5\,\mathrm{GHz}$ intervals, and alignment was performed at $850\,\mathrm{GHz}$. This range was chosen for its lack of atmospheric absorption (see Fig. 4.13), high detector sensitivity (see Fig. 4.14), and high photomixer output power.

At each of the sampled frequencies, interferograms were obtained using interferometer baselines from $135 - 310\,\mathrm{mm}$ in increments of $5\,\mathrm{mm}$. A subset of these interferograms when viewing radiation with frequency $\nu \approx 855\,\mathrm{GHz}$ is shown in Fig. 6.1. The resulting interferograms are well described by the monochromatic DFI response Eq. 2.62, which has the functional dependence

$$I_{DFI}(B,z) = I(k)|\gamma_0(B,k)|\sin(2\pi kz) \ , \tag{6.1}$$

where B is the baseline, z is Optical Path Difference (OPD), and $k = 1/\lambda = \nu/c$ is the wavenumber of the radiation with wavelength λ and speed of light c. The main quantity of interest is $|\gamma_0(B,k)|$, the absolute value of the complex visibility. By definition, this quantity is $1 \geq |\gamma_0| \geq 0$, and serves to modulate the sinusoidal interference pattern. This modulation can be observed in the top panel of Fig. 6.1 where $|\gamma_0|$ decreases with increasing baseline from 135 to $175\,\mathrm{mm}$.

The amplitude of the sinusoid can be extracted directly from the interference pattern by measuring the peak-to-peak intensity. However, this approach is only possible for a monochromatic source, and since the optical frequency is also of interest, this approach can be somewhat cumbersome. A more general strategy is to take the Fourier transform of the interferogram which produces a complex valued sinc function centered on the optical frequency. Taking the absolute value of this result gives the absolute value of the sinc function with amplitude $I(k)|\gamma_0(B,k)|/2$. An example of this process is shown in the bottom panel of Fig. 6.1. By fitting these profiles with an appropriate function, the optical frequency and amplitude of the interference pattern can be obtained with high precision.

Taking the absolute value of the spectrum removes any phase information encoded in

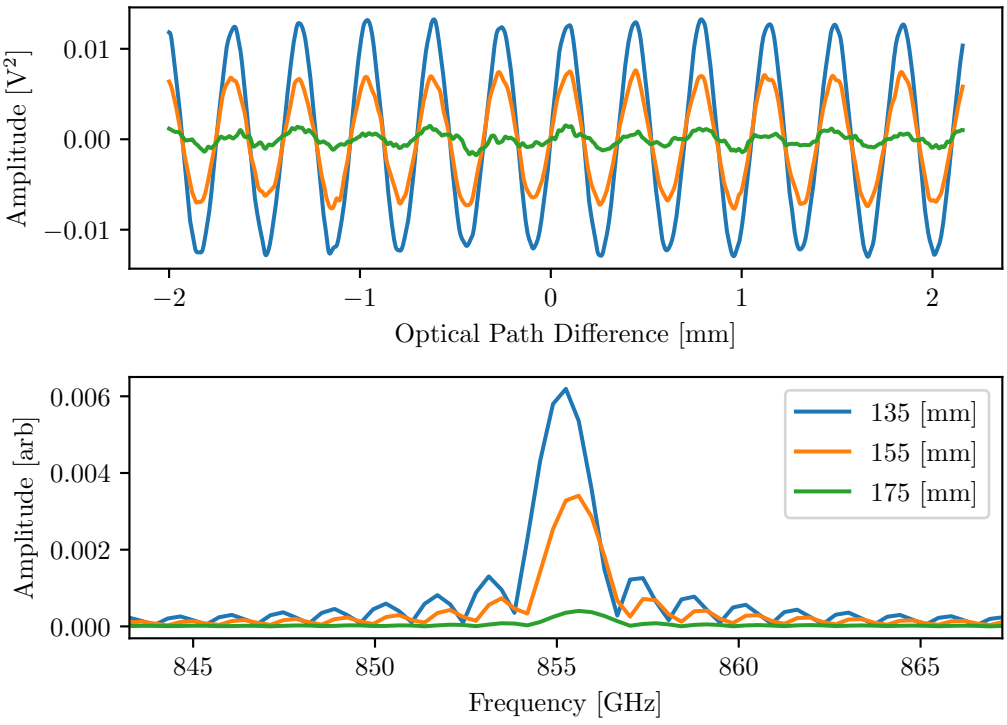


Figure 6.1: Baseline effect on fringe visibility. The top panel shows the measured monochromatic interferogram when observing the photomixer at various interferometer baselines. It is observed that the contrast or visibility of fringes decreases with increased baseline. The amplitude of the sinusoidal interferogram is proportional to the fringe visibility and can be obtained by computing the absolute value of its Fourier transform, as shown in the bottom panel. The amplitude of the resulting spectral feature is equal to half the amplitude of the interferogram. Drift in the spectral feature is attributed to thermal drift in the photomixer's distributed feedback lasers.

the complex values, but in some cases this may not be problematic, particularly if some assumptions can be made about the source geometry. In this case, the photomixer emitter consists of a log spiral antenna coupled to a 10 mm diameter circular lens. With these considerations, it is perhaps reasonable to approximate the source as a circle. From the van Cittert-Zernike theorem, the expected visibility profile is,

$$|\gamma_0(B)| \propto \left| \frac{J_1(2\pi\theta_r B/\lambda)}{2\pi\theta_r B/\lambda} \right| , \tag{6.2}$$

which is the absolute value of the Fourier transform of a circle, where J_1 is the first order Bessel function of the first kind and θ_r is the angular radius of the circle. Note that source plane angles can be converted to spatial extent using Eq. 3.18.

The absolute values of the sinc functions for all the observed interferograms were obtained through model fitting. These results are shown in Fig. 6.2. The dashed curves show the best fit for the data to Eq. 6.2. Both the cure and the data have been scaled by a normalization factor such that the model is unity at $B = 0$ so that the curve is equivalent to the visibilities. This factor was obtained from the fitting process. Although there are some deviations, the model function represents the data reasonably well. Furthermore, the fitted angular radius of the model is equivalent to 4.6 ± 0.1 mm on the source plane, which is consistent with the 10 mm diameter photomixer emitter lens. From these experiments, the relationship between fringe visibility and source extent has been demonstrated in addition to how visibilities can be extracted from interferograms and how such information can be used to infer the structure of the source through model fitting.

6.2 Virginia Diodes Observations

Due to their high power and relatively high spatial coherence, the Virginia Diodes (VDI) sources, provided by Blue Sky Spectroscopy and discussed in § 3.2.4, are convenient for investigating how source position and interferometer baseline affect the phase of the DFI

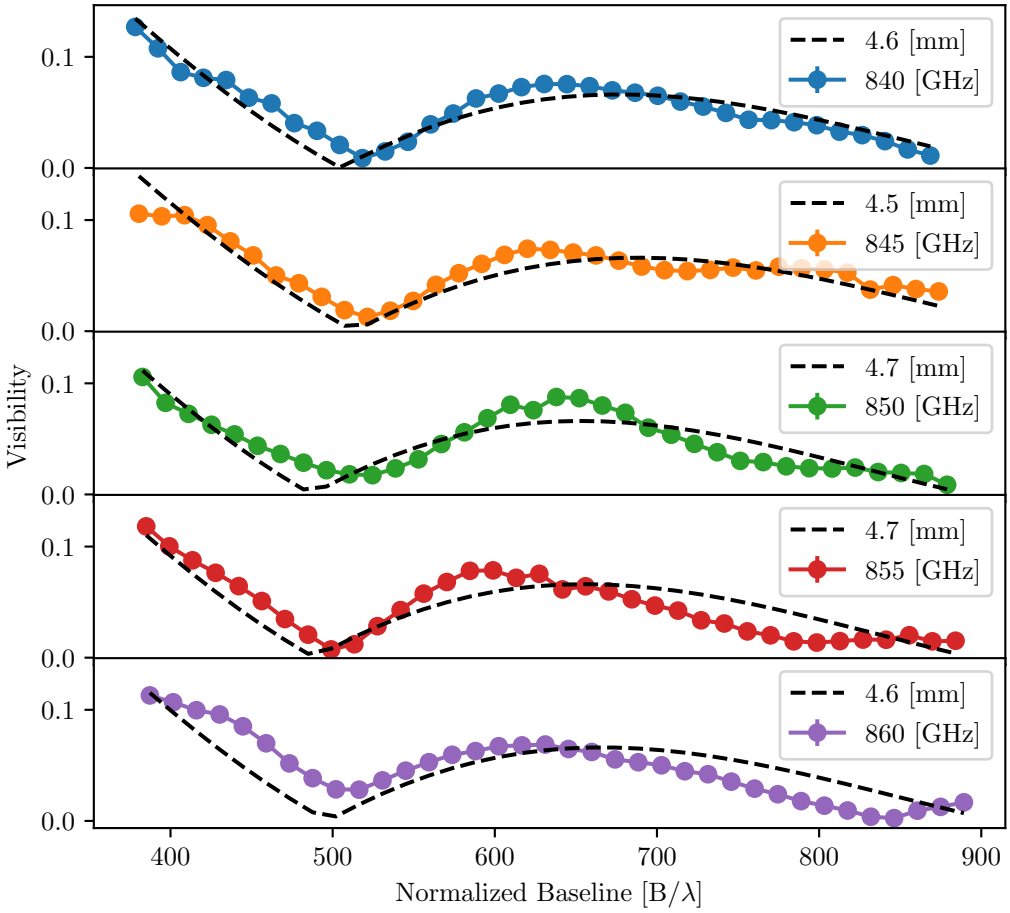


Figure 6.2: Visibility curves obtained when viewing the photomixer at different monochromatic frequencies. The curves are well modelled by the expected visibility profile given by Eq. 6.2. Best fit models are given by the black dashed curve with the best fit radius of the assumed circular source indicated in the legend. The discrepancy between data and model may be due to imperfect assumptions related to the source geometry.

response. Additionally, the high temporal coherence and precision with which the source frequencies can be controlled makes these sources ideal for some aspects of spectral characterization of the DFI instrument.

For these experiments, two VDI source were used in addition to a polarizer (see Fig. 6.3). One source was position on the focal plane of the collimating mirror (i.e., the nominal source plane). The polarizer was placed $\sim 200\,\mathrm{mm}$ in front of this source with its polarization axis at 45° with respect to the normal vector of the optics table. A second source was positioned at the reflected focal point of the collimating mirror. The source which emits light that is transmitted through the polarizer to the collimating mirror maintained a constant frequency $\nu_T = 320.001\,\mathrm{GHz}$, but its horizontal position on the source plane was allowed to vary. The source which emits light that is reflected off the beamsplitter to the collimating mirror had a frequency ν_R which is varied, but its position was held constant. Configured in this way, transmitted and reflected light have an electric field with the same polarization, and the sources can be made to effectively overlap on the source plane. The sources were initially aligned by back propagating a visible laser through the interferometer system. The position of the reflected source was then optimized for maximum throughput in the same way as the photomixer described in Sec. 6.1. The position of the transmitted source was made to align with the effective position of the reflected source, where alignment accuracy was determined by inspection of measured interferograms, as will be detailed later in this section. Finally, the output power of each source was tuned to be approximately equal through both the static and dynamic aperture at a 135 mm baseline.

When observing either the reflected or transmitted source individually, the resulting interferogram is well described by the monochromatic DFI response Eq. 2.62, with the functional dependence

$$I_{DFI}(B,z) = I(k)\sin\left[2\pi k(z - B\theta)\right] \ , \tag{6.3}$$

where the variables are defined in the same way as Eq. 6.1, and θ is the angular position of the source. Comparing this equation to Eq. 6.1, note that the fringe visibility is set to

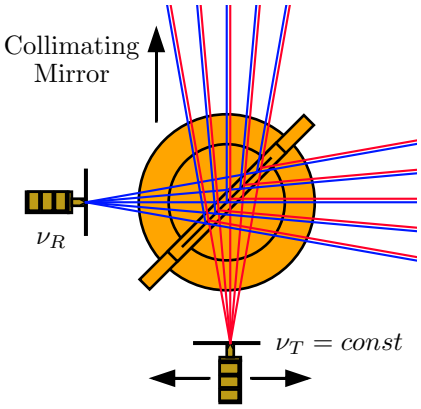


Figure 6.3: VDI experimental setup. One source is aligned with the nominal focal plane of the collimating mirror and emits radiation that is transmitted through the beamsplitter. This transmitted source maintains a constant frequency, but its position along the focal plane can be varied. A second source is placed at the reflected focal point of the collimating mirror. This reflected source is operated at a constant position, but its frequency can be varied. After the beamsplitter, light from the two sources have orthogonal polarization states.

unity, as is characteristic of a spatially coherent point-like source, and an extra term has been added to the argument of the sin function which is the phase of the complex visibility $\phi_0(\vec{B},k) = -2\pi kB\theta$, associated with a point-like source. Although the VDI sources are not point sources, the physical extent of their emitter is small and the wavelength at which they are operated is relatively large, so the point-like description is approximately true.

The signal obtained when observing both sources simultaneously can be approximated by summing the responses associated with two monochromatic sources,

$$\begin{aligned} I_{VDI}(B,z) &= I_1 \sin\left[2\pi k_1(z - B\theta_1)\right] + I_2 \sin\left[2\pi k_2(z - B\theta_2)\right] \\ &\approx (I_1 + I_2)\cos\left[2\pi z(k_1 - k_2) - \alpha\right] \sin\left[2\pi z(k_1 + k_2) - \beta\right] \ , \end{aligned} \tag{6.4}$$

where the subscripts denote variables from different sources and,

$$\alpha = 2\pi B(k_1\theta_1 - k_2\theta_2) \tag{6.5}$$

and

$$\beta = 2\pi B(k_1\theta_1 + k_2\theta_2) \tag{6.6}$$

are constants for an individual interferogram. The second line in Eq. 6.4 is obtained using the sum-to-product trigonometric identity, and the approximation assumes the intensities I_1 and I_2 are equivalent. The full response is then a high frequency sine wave modulated by low frequency cosine function. The envelope of the response is the most apparent structure when inspecting interferograms and is phase shifted by α which maps to an OPD shift $\delta_z = B(k_1\theta_1 - k_2\theta_2)/(k_1 - k_2)$. Fourier transforming the composite interferogram results in a complex spectrum with complex sinc features centred on the optical frequencies of the sources. Complex visibility phase can then be computed for each source using

$$\phi_0(B,k) = -2\pi kB\theta = \arctan\left(\frac{S_{Im}(k)}{S_{Re}(k)}\right) , \tag{6.7}$$

where $S_{Re}(k)$ and $S_{Im}(k)$ are the real and imaginary components of the complex spectrum. For this experiment, the frequencies $\nu = ck$ and amplitudes of these components are determined by fitting a sinc function to the data. Monochromatic and composite interferograms are shown in Fig. 6.4 in addition to the complex spectrum associated with the composite interferogram.

Alignment of the transmitted source with the position of the reflected source was assessed with reference to the envelope of the composite interferogram. Inspecting α in Eq. 6.5, it is clear that the envelope is independent of baseline when $k_1\theta_1 = k_2\theta_2$, and if $k_1 \approx k_2$, this condition reduces to $\theta_1 \approx \theta_2$. The reflected source frequency was set to 323.773 GHz, representing a 1.2% difference with respect to the transmitted source frequency. The position of the transmitted source was then adjusted so that the interferogram envelope was independent of baseline. The end result is presented in the left panel of Fig. 6.5. This figure also shows the complex spectra, obtained from the Fourier transform of the interferograms. Positions of the sources with respect to the optical axis can be com-

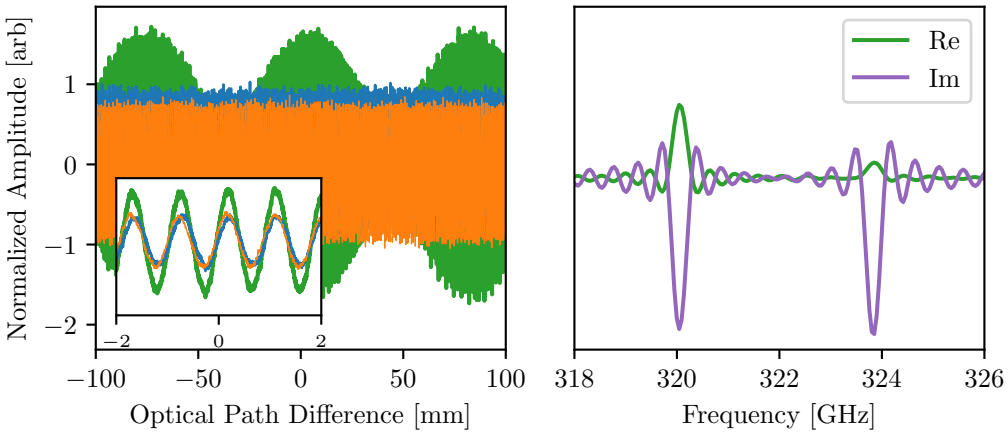


Figure 6.4: Example VDI response and complex spectrum. The left panel shows the sinusoidal monochromatic response associated with the reflected (blue) and transmitted (orange) sources, with a vertical offset for clarity. Both sources observed simultaneously produces the green curve, referred to as the composite response, which shows a prominent cosine modulating envelope. The inset shows a close up of the ZPD region. The right panel shoes the complex spectrum of the composite interferogram with real (green) and imaginary (purple) sinc features centered on the source frequencies of the sources.

puted from the baseline dependent phase of each source. With reference to Eq. 6.7, the phase can be represented as a linear equation with slope $-2\pi k\theta$ and independent variable B. Calculating the baseline dependent phases, and fitting linear equations to the data, results in a slope which correspond to positions -1.4 ± 0.1 mm for the reflected source and -1.3 ± 0.1 mm for the transmitted source.

When the source positions do not align, the envelope of the composite interferogram is expected to shift as a function of baseline. To validate this hypothesis, the transmitted source was moved 1.8 ± 0.2 mm to the left of its reference position, which corresponds to a positive increase with respect to the interferometer coordinate system (see Fig. B.1). Interferograms and complex spectra from measurements in this configuration are shown in Fig. 6.6. It can be observed that the interferogram envelope slowly drifts to the left with increasing baseline. Furthermore, the complex spectra associated with the stationary reflected source at ~ 324 GHz is equivalent to complex spectra associated with the same source in the previous configuration (see Fig. 6.5). However, the phase characteristics of

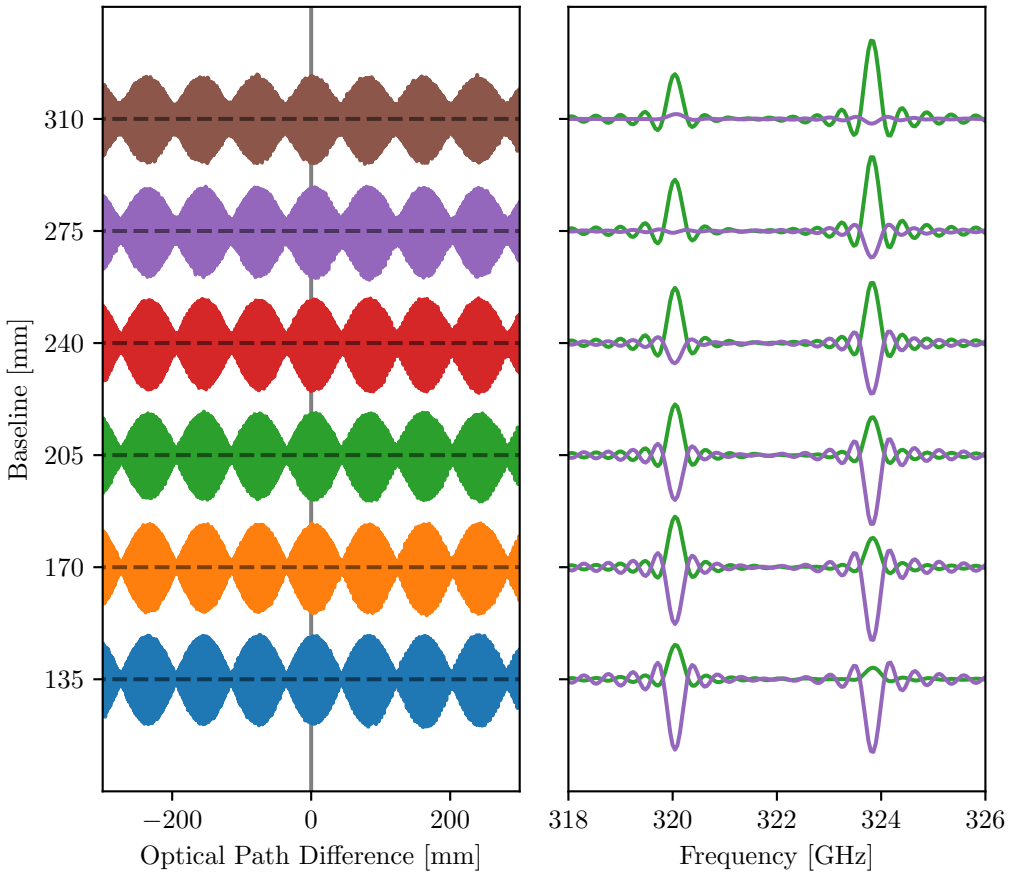


Figure 6.5: Results obtained at different baselines when the VDI sources overlap. The left panel shows the composite interferograms at different baselines, and it can be observed that the positions of the envelopes are nearly constant. The right panel shows the resulting complex spectra. By calculating the phase associated with each source at each baseline and fitting a linear equation, the resulting slope can be used to estimate source position. In this case, the reflected source position is -1.4 ± 0.1 mm, and the transmitted source position is -1.3 ± 0.1 mm, which are approximately equal, as expected.

the transmitted source have changed. Fitting a linear equation to the baseline dependent phase of the transmitted source in the same way described in the previous paragraph reveals a source position of 0.5 ± 0.1 mm, corresponding to a 1.8 mm shift in source position, as expected. The results of the linear fitting procedure are shown in Fig. 6.7

An important feature of Figs. 6.5 and 6.6 to emphasize is how the phase of the complex spectra varies with baseline. The functional dependence is given by Eq. 6.7, which is linear with baseline, as previously mentioned. Invoking the translation property of Fourier transforms, this linear phase in the spectral domain causes a translation in the interferogram domain. As such, when a source is at a fixed point offset from the optical axis, the interferogram will be increasing shifted from the phase center as the baseline increases. Analogously, when observing at a fixed baseline, the interferogram will be increasingly shifted from the phase center when the source position is increasingly offset from the optical axis. This effect is demonstrated in Fig. 6.8 which shows interferograms and complex spectra for observations with a constant baseline, where the transmitted source is offset from its reference position in increments of ~ 2 mm. It is observed that the phase of the reflected source at ~ 324 GHz is constant, since the baseline and position are fixed. However, the phase of the transmitted source is changing with source position at increments of $\sim -\pi/2$. With reference to Eq. 6.7, this observed phase increment corresponds to a source position increment of ~ 2 mm, as expected.

The preceding demonstrations are intended to emphasize the relationship between interferometer baseline and source position on phase in the complex spectrum. Note that in these demonstrations, the phase under study is associated with the phase of the complex visibility Eq. 6.7, and as such, these demonstrations show how complex visibility phase reveals information about source position. In these experiments, two sources have been used. It is clear that the phase associated with each source position is independent of the other, as indicated by Fig. 6.8, so the distribution of sources on the source plane can be obtained with these kinds of observations. Similarly, it has been shown that the spectral components

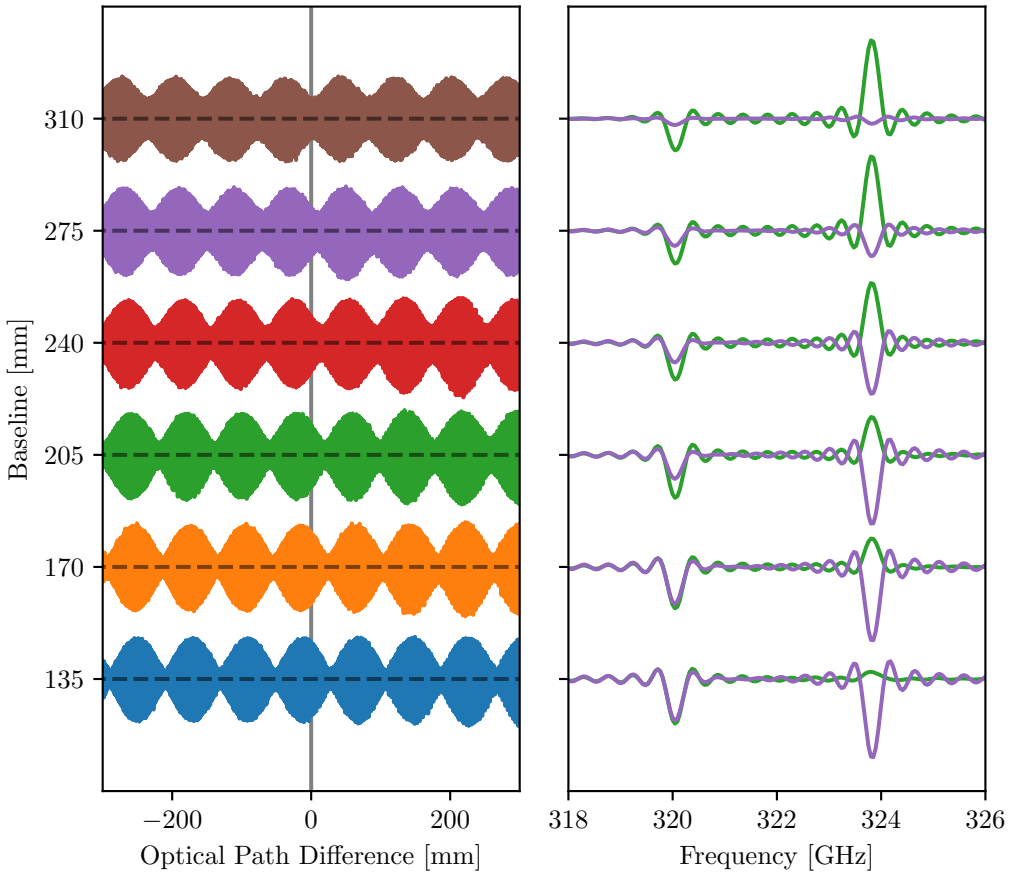


Figure 6.6: Results obtained at different baselines when the VDI sources do not overlap. The left panel shows the composite interferograms at different baselines, and it is observed that the envelope of the composite interferogram slowly drifts to the left as the baseline increases. The right panel shows the complex spectra associated with each interferogram. The phase properties of the stationary/reflected source at $\sim 324\,\mathrm{GHz}$ have not changed from the previous experiment, but the phase properties of the translated/transmitted source at $\sim 320\,\mathrm{GHz}$ have. Estimating source position from the phase gives $-1.5 \pm 0.1\,\mathrm{mm}$ for the reflected source, and $0.5 \pm 0.1\,\mathrm{mm}$ for the transmitted source.

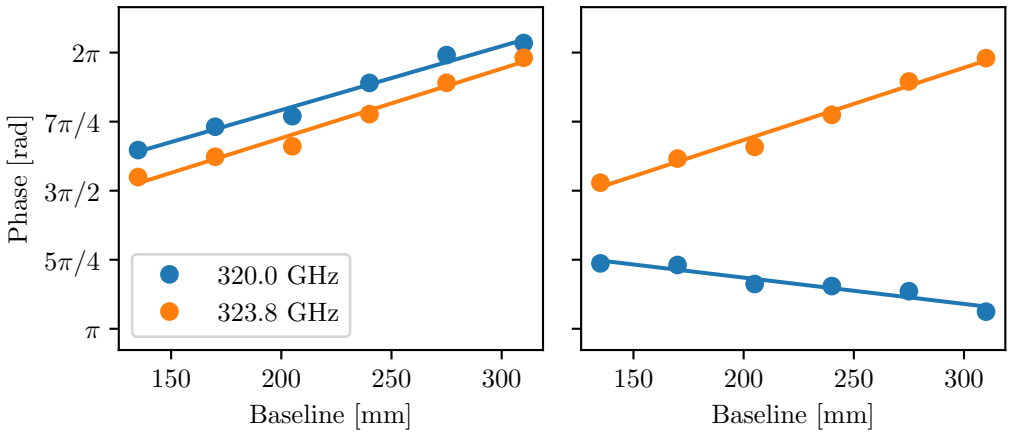


Figure 6.7: Linear fit to baseline dependent phase for the transmitted (blue) and reflected (orange) sources with the transmitted source at different positions. The left panel shows data for when the two sources overlapped. The nearly equivalent slopes indicate a common position on the source plane. The right panel shows data for when the sources do not overlap. The distinctly different slopes indicate different positions on the source plane.

of the complex spectrum can be treated independently. With the discussion of complex visibility phase in this section, combined with the discussion of complex visibility amplitude in Sec. 6.1, it should be clear how complex visibility as a whole relates to the source intensity distribution and how the DFI technique can, in principle, be used to generate spectral images.

Due to the high precision of the VDI sources, these results can be used to validate the performance of the spectrometer. By fitting sinc functions to the spectral features in the previous figures, and using the centers of these fitted functions as estimates for the feature frequencies, the source frequencies were measured to be $320.048 \pm 0.006\,\text{GHz}$ for the transmitted source and $323.824 \pm 0.004\,\text{GHz}$ for the reflected source. These are compared to the expected frequencies of $320.001\,\text{GHz}$ and $323.773\,\text{GHz}$ for the transmitted and reflected sources, respectively. The discrepancy represents a $0.05\,\text{GHz}$ systematic error with a $0.005\,\text{GHz}$ random error. Since the systematic error is four times lower than the best possible spectral resolution of the instrument, and more than three orders of magnitude lower than the spectral resolution under nominal observing conditions, the spectral performance is considered more than adequate.

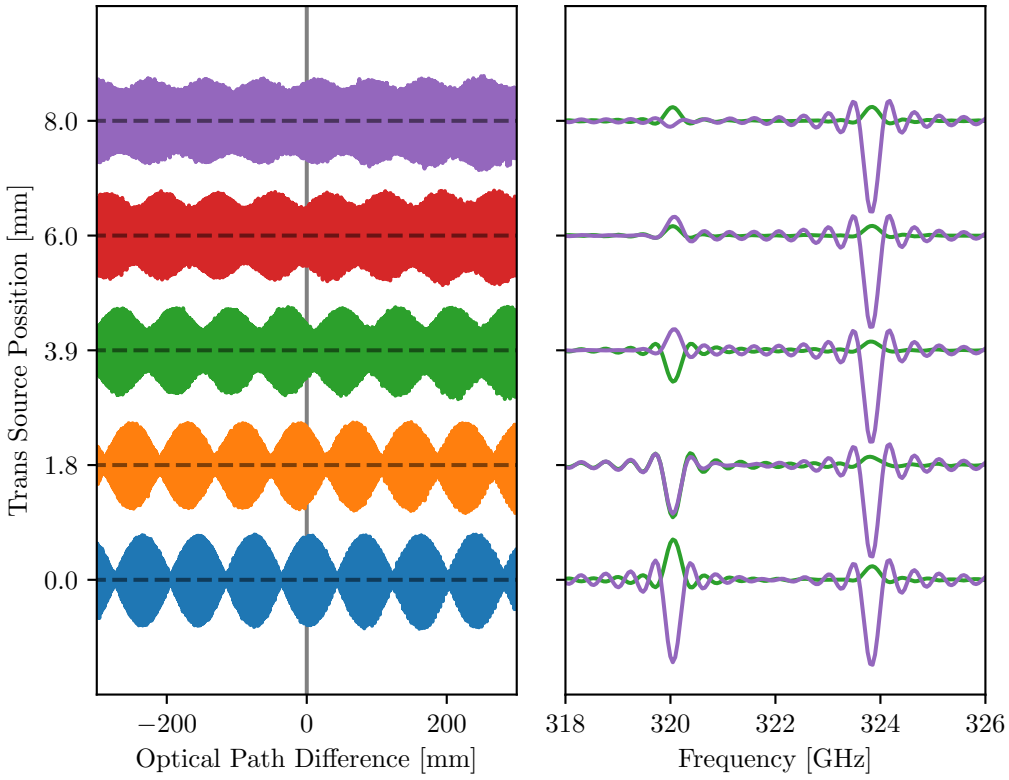


Figure 6.8: Phase dependence on source position. The left panel shows composite interferograms at a constant 135 mm baseline where the transmitted source is translated in increments of $\sim$2 mm. It can be observed that the envelope of the interferogram drifts to the left. The right panel shows the complex spectra associated with the interferograms. The phase properties of the reflected source are constant since the source position and interferometer baseline are stationary. The phase properties of the transmitted source, however, are variable and change in increments of $\Delta\phi \approx 90^\circ$ which corresponds to translation increment of $\Delta x \approx 2.0$ mm, as expected. Note the decreased amplitude of the transmitted source is due to the source moving out of the detector beam, and results in decreased contrast in the interferogram envelope.

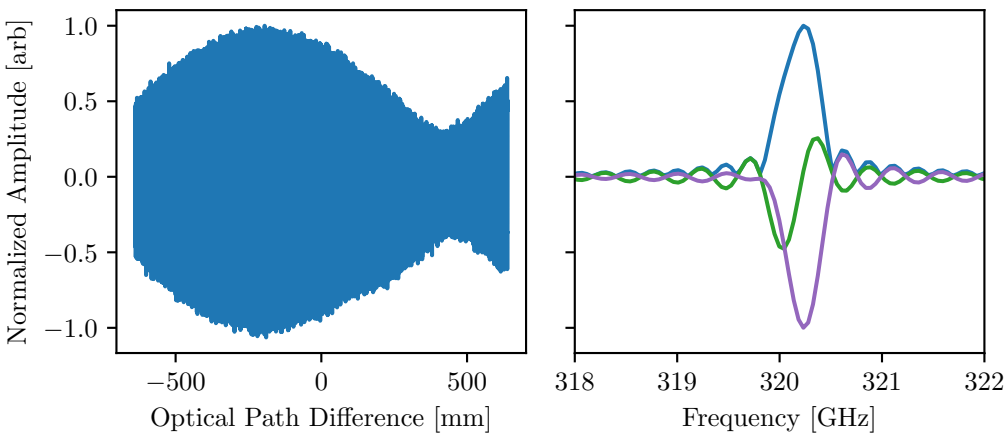


Figure 6.9: Demonstration of spectral resolving capabilities of the DFI technique. The VDI sources were configured to have a frequency difference of $\sim$0.235 GHz which is considered resolved with a 640 mm maximum OPD. The left panel shows the resulting interferogram, and expresses the characteristic single period interferogram envelope for two resolved spectral feature. The right panel shows the associated complex spectrum and power spectrum given by the blue curve. The blue curve is similar to what would be seen by an Fourier Transform Spectrometer (FTS), and it is difficult to distinguish the two features. Under certain conditions, however, the complex nature of the DFI spectrum make the two features easy to resolve.

As an interesting demonstration of the spectral resolving capabilities of the DFI, the transmitted source frequency was set to 320.236 GHz. The frequency difference between the transmitted and reflected source was then $\sim$0.235 GHz. Resolving these two sources would require a maximum OPD of 640 mm (see Eq. 2.54). Furthermore, the complex spectra of the transmitted and reflected sources are expected to be $\pi/2$ out of phase when separated by 2 mm and observed with a baseline of 135 mm. The interferogram and complex spectrum obtained under these conditions is shown in Fig. 6.9. It is observed that the envelope of the interferogram goes though one complete cycle over the measured OPD, which is characteristics of an interferogram with two resolved spectral features. This statement is verified by inspecting the cosine term of Eq. 6.4 and noting that an envelope with wavelength $2z_{max}$ is expressed as $2\pi z_{max}\Delta k = \pi$, which is consistent with the spectral resolution equation of an FTS, $\Delta k = 1/2z_{max}$. The absolute value of the complex spectrum is shown in blue in the right panel of Fig. 6.9. This curve is similar to the absolute value of the real

spectrum that would be obtained using an FTS. Even though the spectral features are considered resolved, the lower amplitude of the lower frequency feature makes it quite difficult to distinguish the two features by inspecting the blue curve. However, the $\pi/2$ complex visibility phase difference resulting from the DFI observing conditions results in the feature associated with the transmitted source being mostly real, while the feature associated with the reflected source is mostly imaginary. Thus, under certain conditions, features that are difficult to resolve with an FTS are easily distinguish with a DFI.

6.3 One Dimensional Source

The remaining observations discussed in this chapter were obtained and processed following the procedures discussed in Ch. 5. The sources consist of various aperture masks backlit by the Pegasus thermal source. Aperture masks are constructed by cutting slits of various widths with various orientations into the mask material. Slits are used because they can be cut with high precision, have well defined and easily modelled visibility curves, and they provide a sensitivity advantage. A spatial interferometer is only sensitive to spatial variations parallel to the baseline vector. In effect, the interferometer sees the two dimensional source plane, multiplied by the detector sensitivity, as if it was projected onto a one dimension space represented as a line that is extended parallel to the baseline vector. When using a slit mask with structure that is constant along the perpendicular direction, this projection increases throughput in proportion to the perpendicular extent of the slit within the detector beam without changing the source structure along the parallel direction. This effect helps to overcome the sensitivity constraints of the detector system and is the driving motivation for an optical design with a <1 magnification which increases the detector beam size, and subsequently, the perpendicular extent of the slit that the detector is sensitive to. A slit mask is effectively a one dimensional source since there is no variation along one axis. As such, they are best used when studying single baseline orientation observations. Fortunately, generalizing the observation and processing steps required for a single orientation to

multiple orientations is, in principle, quite simple (see Sec. 6.4). Validating the DFI technique at a single baseline orientation is thus foundational to developing more complicated observation strategies.

In this section, results are presented for single baseline orientation observations when viewing one dimensional slit masks. For each observation, data is compared to ideal noise free models at each relevant step. These results are generally communicated though a multi panel figure which includes panels for measured data, processed data, and aperture synthesis. Due to the limited space in each figure, legends are excluded. However, the color scheme used in each panel is consistent, and can be referenced to Fig. 6.10. Each panel is analogous to a corresponding figure previously presented in Ch. 5 where they are discussed in detail. The top left panel shows phase corrected band-pass filtered interferograms. The top right panel shows the FTS spectrum for the observation and complex spectra associated with each interferogram. The center panel shows complex visibility samples. The panels bellow this show the dirty image, clean image with clean components, and the source model that is assumed when calculating simulated results. When appropriate, experimental data is presented alongside ideal models, and the qualitative agreement between the two is used to demonstrate the consistency between experiment and theory.

The model source was constructed so as to accurately represent the source being observed with minor manual adjustment in width and position to improve agreement between data and simulation. Position is effectively a free parameter since the aperture mask mounting strategy was unsophisticated and the position of the slit relative to the optical axis is limited to the precision of manual adjustment in the absence of a well defined reference (i.e., a few millimeters). Model slit width consistently needed to be made smaller than the actual source. The reason for this is left as a discussion topic in the next section.

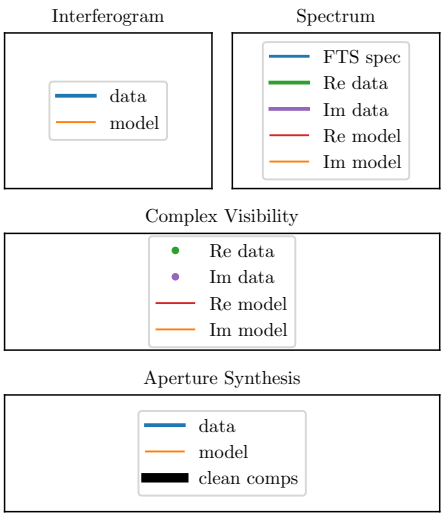


Figure 6.10: Multi-panel figure legend. The color scheme used in the results figures presented throughout this chapter are consistent with those indicated in this figure.

6.3.1 Uniform Spectrum

The observations presented in this subsection consist of slit masks back lit by the Pegasus thermal source, and data is derived from the central A0 detector of the detector array. Although the relative intensity at each optical frequency may be different, the shape of the souce should be constant at all frequencies. As such, these experiments are described as having a uniform spectrum. Results when observing a 1 mm slit offset from the optical axis by $\sim$3 mm are shown in Fig. 6.11. Results when observing a 1 mm and 2 mm slit separated by 5 mm are shown in Fig. 6.12.

In general, there is excellent agreement between experimental and model results at every stage of processing. It is clear that the DFI instrument can extend the spatial interferometry techniques. In particular, the DFI can accurately reconstruct the shape and distribution of sources much smaller than the detector beam. Note that in the case of the double slit experiment, light from both slits is integrated onto the same detector, and the slits would be impossible to distinguish with non-interferometric imaging techniques. Although agreement between data and model is quite good, a few deviations are evident which can be fully understood as instrumental or experimental error.

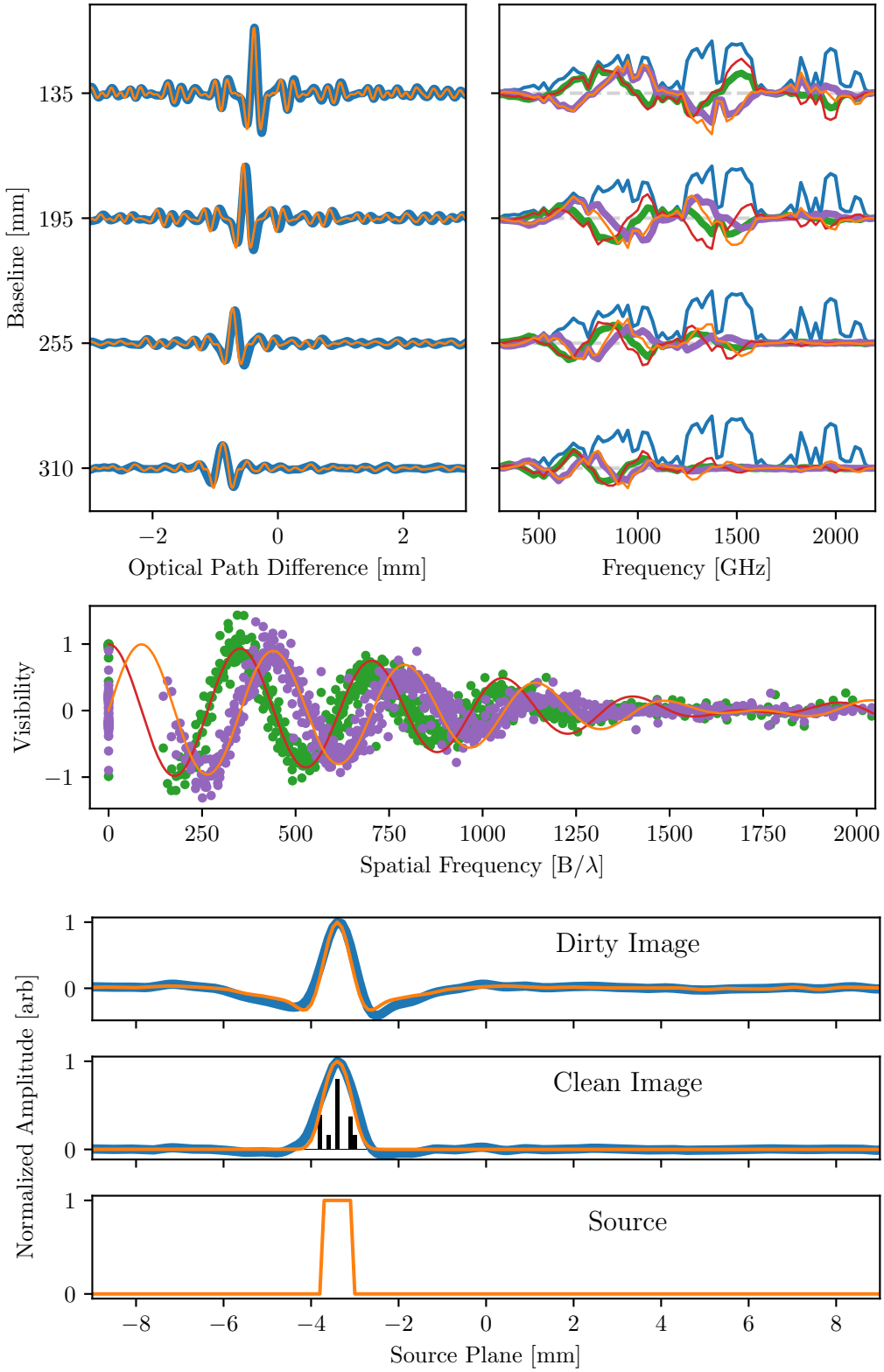


Figure 6.11: Results for an offset 1 mm slit. Model is a 0.7 mm slit centered at −3.4 mm.

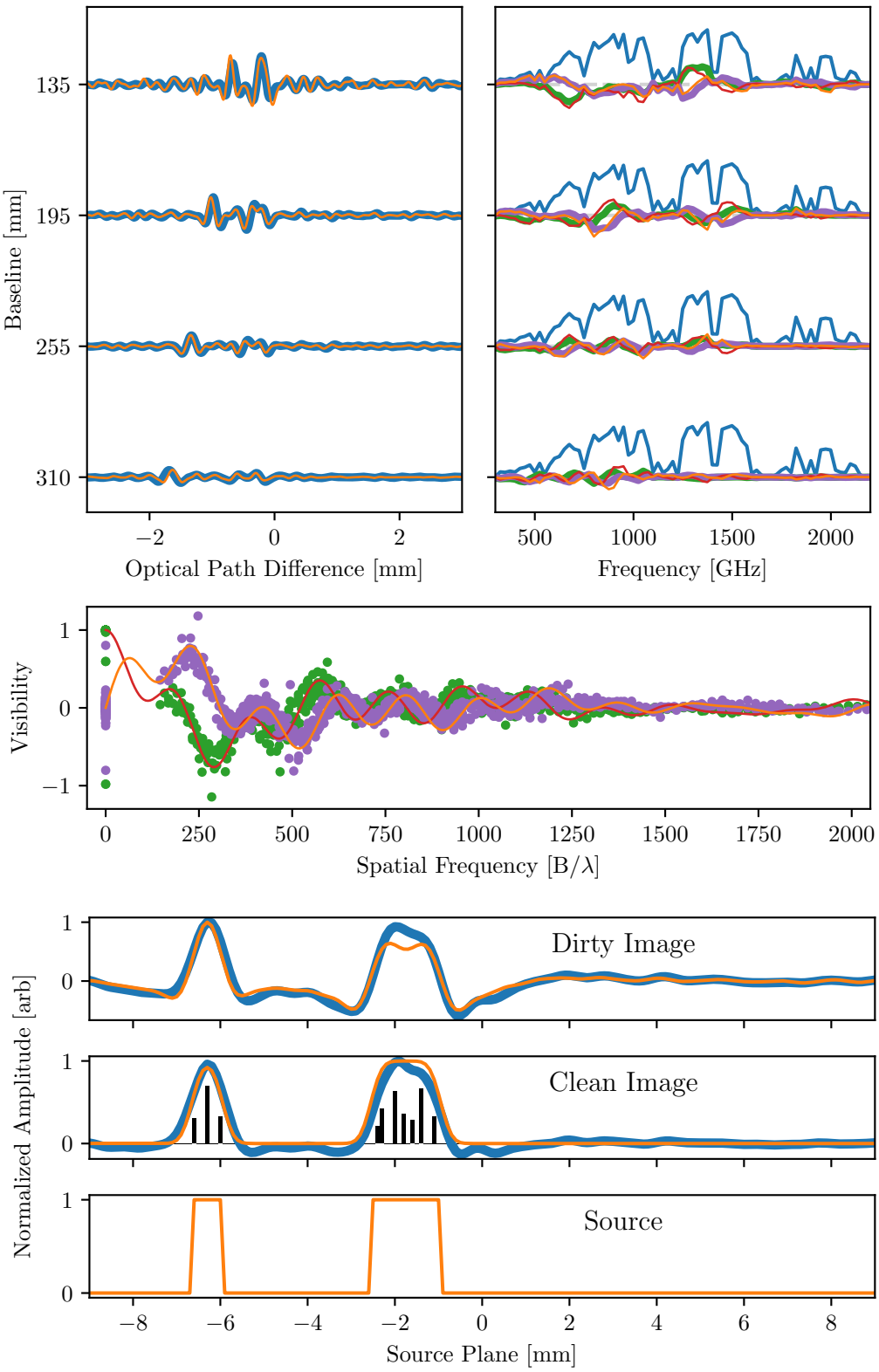


Figure 6.12: Results for a 1 mm and 2 mm slit separated by 5 mm. Model is a 0.7 mm and 1.6 mm slit separated by 4.5 mm.

Model source size must consistently be assumed to be smaller than the actual slit width in order to improve agreement between data and models. This can be explained by the geometry of the aperture mask with respect to the collimating mirror. The masking material has thickness $t = 2\,\text{mm}$. When the source of width w is viewed at an angle θ, the apparent source size is

$$s = w\cos(\theta) - t\tan(\theta)\ . \tag{6.8}$$

The first term represents the projected width of the source when viewed at an angle, while the second term accounts for the additional reduction in apparent source size that results from the finite thickness of the masking material. The apparent size of the source is a monotonically decreasing function of viewing angle. For example, a $w = 1\,\text{mm}$ slit viewed at an angle of $\theta = 8^\circ$, has an apparent size $s \approx 0.7\,\text{mm}$. From the perspective of the static aperture, which views the source through the collimating mirror at a similar angle, the source always appears to be smaller than it actually is. Since spatial interference can only occur with source points that overlap from the perspective of both apertures, measured interferograms are limited by the smallest apparent size between the two apertures of the spatial interferometer.

Baseline dependent phase errors result from a slight OPD offset of the interferogram. As such, the phase errors have a linear dependence and can be amended at a particular baseline by assuming a sub-mm shift in source position. As can be observed in Fig. 6.11, the complex spectrum is in good agreement with simulations at $B = 135\,\text{mm}$. As the baseline increases to $B = 195\,\text{mm}$, the agreement with simulation decreases as a result of linear phase error. As the baseline continues to increase to $B = 310\,\text{mm}$, agreement between data and simulation improves. This baseline dependence indicates that simply assuming a different source position will not improve the overall agreement. A potential explanation is the errors in curvature associated with the collimating and focusing mirrors. When characterizing the collimating mirror, parallel rays originating from the aperture plane were focused onto the nominal focal point and a $\pm 1\,\text{mm}$ horizontal spread was observed (see Fig. 3.10). Reversing

the rays from focal point to aperture plane would then result in a ± 1 mm apparent shift in source position when viewed from different position in the aperture plane. Inspecting Fig. 3.10, it is notable that the difference between the horizontal positions of rays focused by the outer edges of the collimating mirror is small. If the shift of apparent source position is a result of the collimating mirror curvature, this fact would help to explain why there is agreement between results at small and large baselines.

Another notable deviation from theory is the loss of fringe visibility at higher optical frequencies. A specific example would be at $\sim 1,400$ GHz for the 310 mm baseline observation in Fig. 6.11. Although the expected visibility is low, there seems to be a statistically significant absence of spectral amplitude in the data. However, visibility is at the expected level in the same spectral region for the $B = 135$ mm data, so whatever is causing visibility loss is expected to be baseline and frequency dependent.

Visibility loss can result from differential wavefront error between the two beams of the interferometer, amplitude mismatch, OPD errors, and alignment errors in the form of partial beam overlap and tip-tilt angular differences between the beams when they are combined [129, 130]. Visibility loss can also be due to polarization mismatch, however, the DFI used in this work is not sensitive to polarization so these effects are not relevant here. Of the various contributions listed, differential wavefront, OPD, and tip-tilt alignment errors have an explicit wavelength dependence. Fractional visibility loss can be expressed as

$$\begin{aligned} V_{loss} &= 1 - V_{WFE} \cdot V_{OPD} \cdot V_{tt} \\ &= 1 - \left[e^{-(2\pi\sigma_{WFE}/\lambda)^2} \cdot e^{-(2\pi\sigma_{OPD}/\lambda)^2} \cdot 2\frac{J_1(\pi D\sigma_{tt}/\lambda)}{\pi D\sigma_{tt}/\lambda} \right] , \end{aligned} \tag{6.9}$$

where σ_{WFE}, σ_{OPD}, and σ_{tt} are Root Mean Squared (RMS) deviations associated with differential wave front, OPD, and tip-tilt alignment errors, respectively. The tip-tilt term assumes circular apertures with diameter D. It is clear from Eq. 6.9 that for a given error source, visibility generally deceases with increased frequency. Beam overlap is defined as the fractional area of one interferometer arm beam that overlaps with the other. Although

this is not explicitly a function of frequency, the beam sizes may be frequency dependent and fractional overlap errors related to alignment will have a greater effect on small beams associated with high frequencies than larger beams at lower frequencies. In § 3.4.2, it was found that the model beam sensitivity profile has a slight frequency dependence which may result in a frequency dependent fractional beam overlap error. Worse alignment at larger baselines is expected to be the cause of decreased visibility at larger baselines and at higher frequencies.

6.3.2 Nonuniform Spectrum

The baselines accessible with the testbed DFI are too limited to fill the *uv*-plane sufficiently for the production of monochromatic images at each optical frequency sampled by the DFI. This is why emphasis has been placed on extracting complex visibilities and combining the results of multiple frequencies to increase the sampling in the *uv*-plane. However, this strategy is applicable only when the source is spectrally uniform, which implies that, apart from a scaling factor, the source has the same intensity distribution at all frequencies. This condition is not met in many interesting cases, and analyzing data in this way ignores one of the most attractive features of the DFI technique: the ability to produce spectral images.

To demonstrate spectral imaging capabilities with the testbed system, it is still necessary to combine data from multiple frequencies to appreciably fill the *uv*-plane. In effect, the source spectrum needs to be segmented into broad spectral bands with the source expressing different spatial structure in each band. Such a scene was manufactured using the double slit aperture mask by placing a low-pass THz filter over the 2 mm slit. The filter material consists of a capacitive metal mesh grid with a 1 THz cutoff frequency. The transmission curve, measured using the interferometer in FTS mode, is shown in Fig. 6.13. In this configuration, the source appear as two slits at frequencies $\lesssim 1$ THz, and appears as one slit at frequencies $\gtrsim 1$ THz.

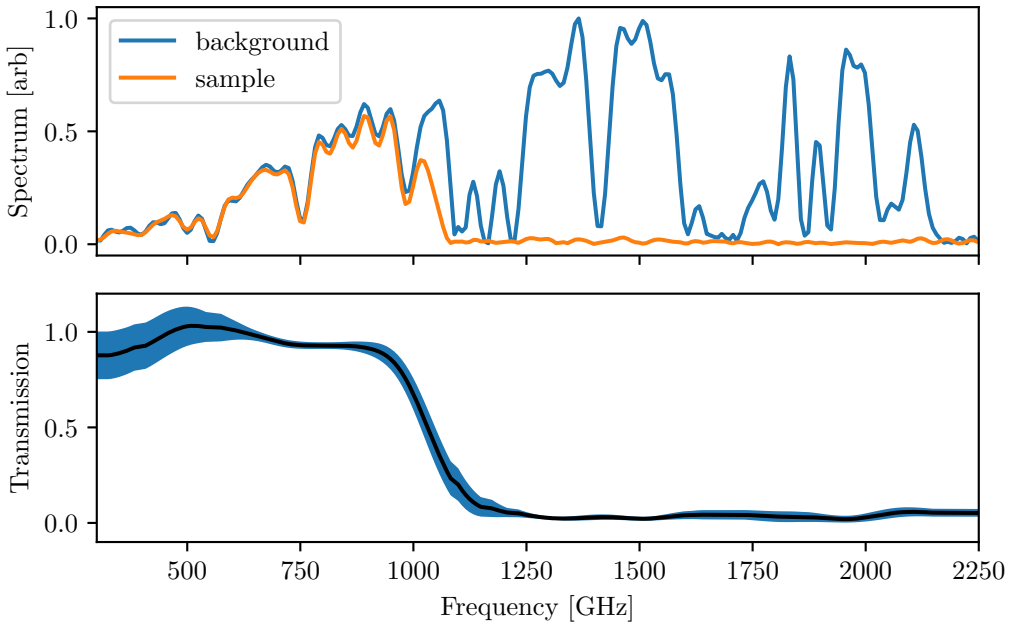


Figure 6.13: Low-pass filter transmission curve. The top panel shows the background spectrum (blue) obtained when viewing an unfiltered slit, back lit by a thermal source. The sample spectrum (orange) was obtained when viewing the slit masked by the low-pass filter material. The transmission curve in the lower panel (black) is computed as the ratio of the sample and the background spectra. Uncertainty in transmission is conveyed by the blue shaded region.

Results for the full instrument bandwidth are shown in Fig. 6.14. Since the source is not spectrally uniform, it is only sensible to show interferograms, and complex spectra for the whole bandwidth. For these quantities, there is a high level of agreement between data and model. Observed deviations result from the same effects previously discussed in § 6.3.1. Results for the lower band are shown in Fig. 6.15. In this case, only complex visibility data derived from complex spectral components between $300 - 1,000\,\mathrm{GHz}$ are considered. In this spectral range, the source is expected to appear as a 1 mm and 2 mm slit separated by 5 mm. Thus, a similar geometry is used when constructing the model and is in excellent agreement with the data.

Results for the upper band are obtained using complex visibilities derived from complex spectral components between $1,200 - 2,200\,\mathrm{GHz}$. In this range, transmission through the low-pass filter is negligible and the source is expected to appear as a single 2 mm slit at the

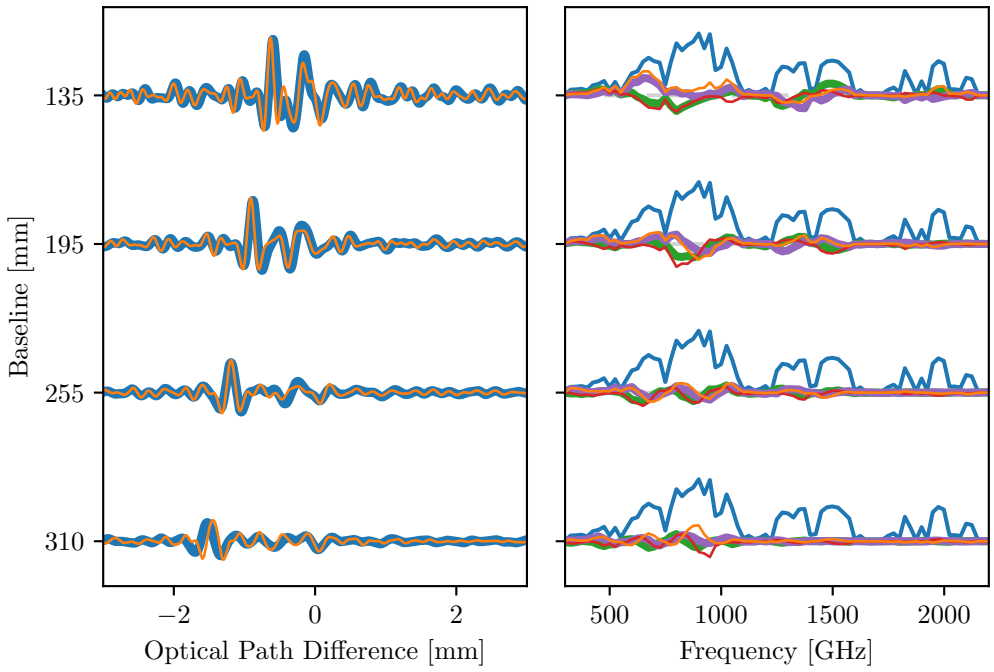


Figure 6.14: Full bandwidth results for a 1 mm and 2 mm slit separated by 5 mm, with the 2 mm slit masked by the 1 THz low-pass material. For modelling, the 2 mm slit is assigned the "sample" spectrum and the 1 mm slit is assigned the "background" spectrum (see Fig. 6.13). Additional model details are described in Figs. 6.15 and 6.16.

same position as the 2 mm slit used in the lower band model. Data and simulated results with these constraints is shown in Fig. 6.16. Complex visibility amplitudes are observed to be systematically lower than the model, which is consistent with previous observations, where visibilities from higher frequency spectral components tend to express lower visibility resulting from the higher alignment precision required for interference. The cleaning procedure performed poorly in this case, which is reasonable considering the absence of low spatial frequency *uv* components. However, the agreement between the measured and simulated dirty images is substantial, and is a compelling indication that experimental results conform to theoretical expectations.

Although this is a particularly simple example with only two spectral channels, the experiment presented here clearly demonstrates that the DFI is sensitive to the spectral structure of the source. In a more general situation with greater baseline sampling, the same principles and processing techniques can be used to construct spectral images with

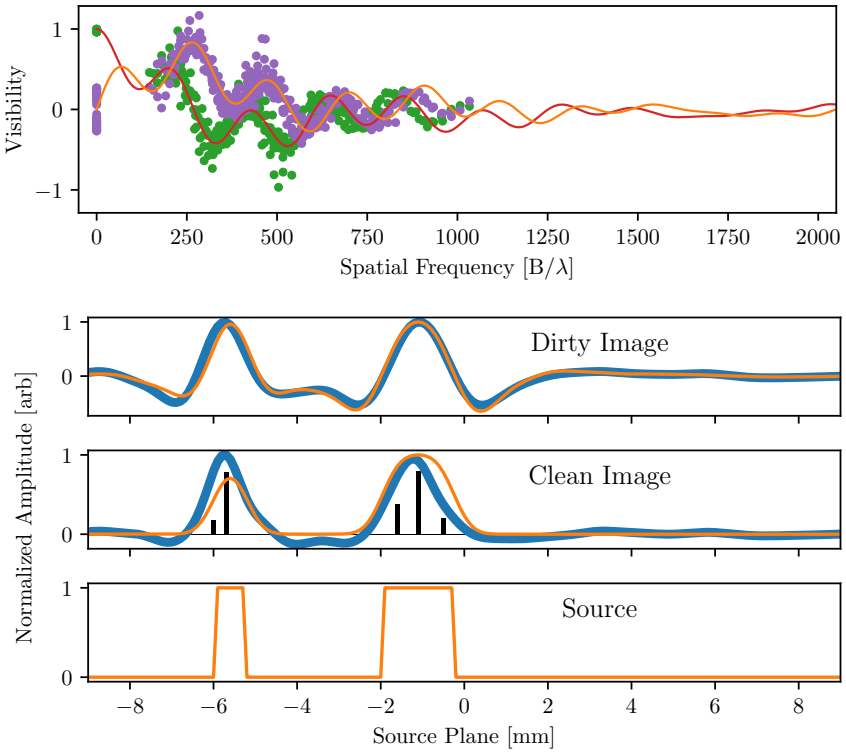


Figure 6.15: Lower band results. The model is a 1.6 mm slit and a 0.7 mm slit separated by 4.5 mm.

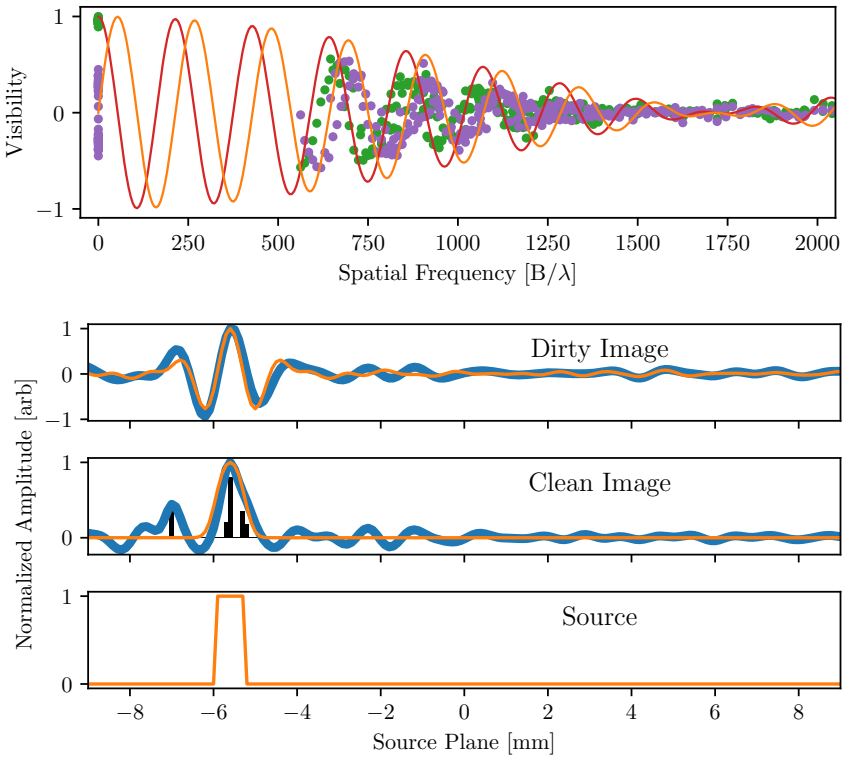


Figure 6.16: Upper band results. The model is a 0.7 mm slit centered at −5.6 mm.

each layer of the image corresponding to a sampled spectral component in the DFI complex spectrum.

6.3.3 Wide Field

Wide field observations are facilitated by the use of multi-pixel detector arrays where the beam associated with each detector in the array is centered on a different position of the source plane. Such a situation is achieved by the optical system used in this work (see Fig. 3.16). To demonstrate wide field imaging, a slit mask was used where a 2 mm and a 1 mm slit were separated by 12 mm. The separation between slits is sufficient for each slit to be approximately centred within the beams of adjacent detectors. For this experiment, the 2 mm slit was placed within the A0 detector beam, and the beam was sufficiently small that it was not sensitive to the 1 mm slit. Similarly, the 1 mm slit was placed within the beam of the A3 detector which had no sensitivity to the 2 mm slit.

Measurements taken with the A3 detector were processed in the exact same way as measurements taken with the A0 detector. Results for the A0 detector are shown in Fig. 6.17 and results for the A3 detector are shown in Fig. 6.18. In both cases, data is well represented by simulation using the assumed source geometry. Combining the source plane models from the two detectors with summation produces a composite model consisting of a 1.4 mm slit, from the A0 detector, and a 0.7 mm slit, from the A3 detector, separated by 13.0 mm. That is, the wide field source can be constructed though straight forward summation of the clean images from both detectors with no additional processing steps.

Due to the linear property of Fourier transforms, results from different detectors can be combined though summation in any of the interferogram, spectral, *uv*-plane, or spatial domains. This fact is demonstrated in Fig. 6.19 which shows interferograms that were obtained when viewing a 1 mm slit that was positioned such that adjacent detectors could view the slit simultaneously. It can be observed that, apart from a scaling factor and some asymmetry, the interferograms obtained from different detectors when viewing the same source

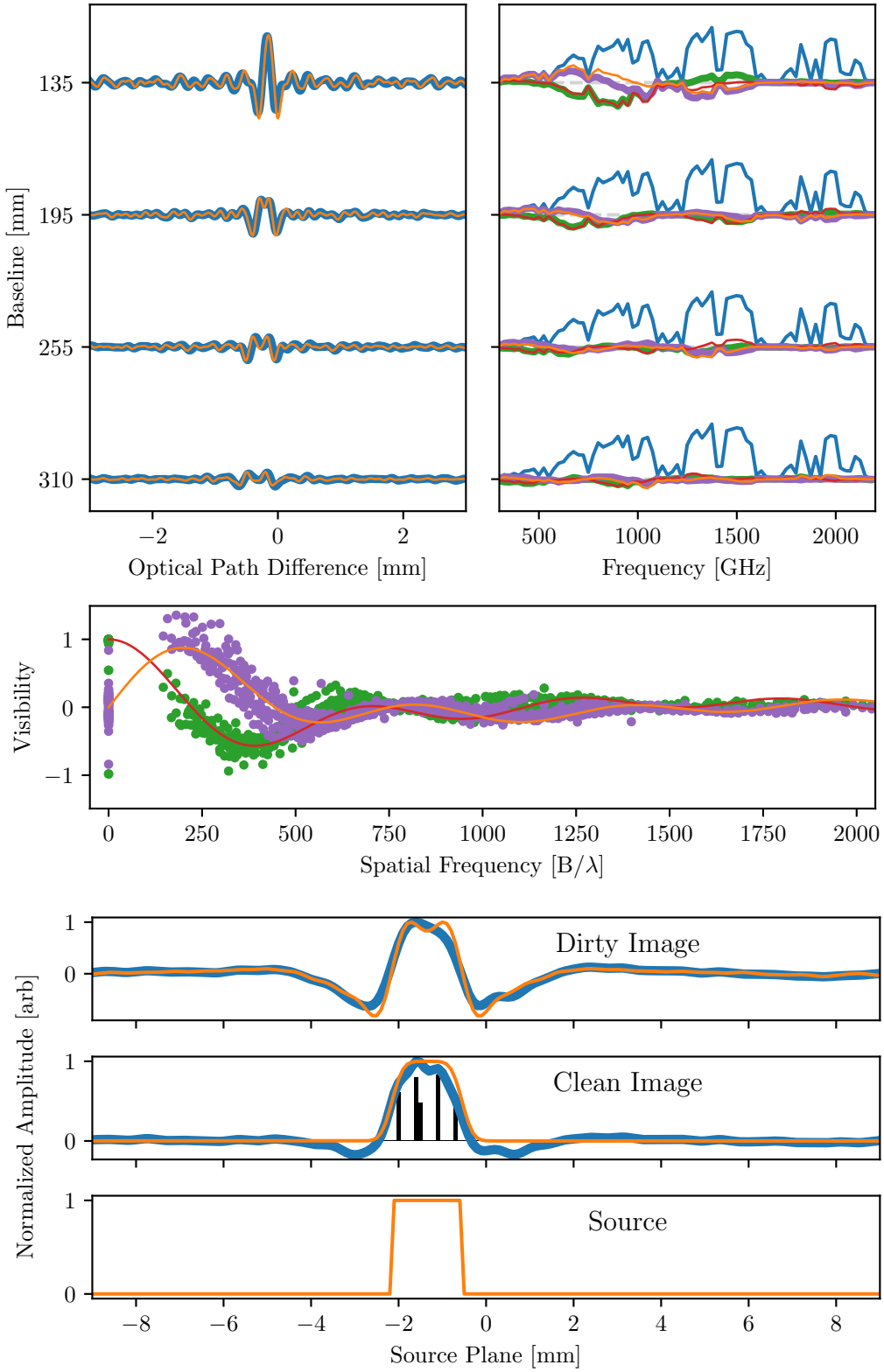


Figure 6.17: Wide field A0 results. Model is a 1.6 mm slit centered at −1.4 mm.

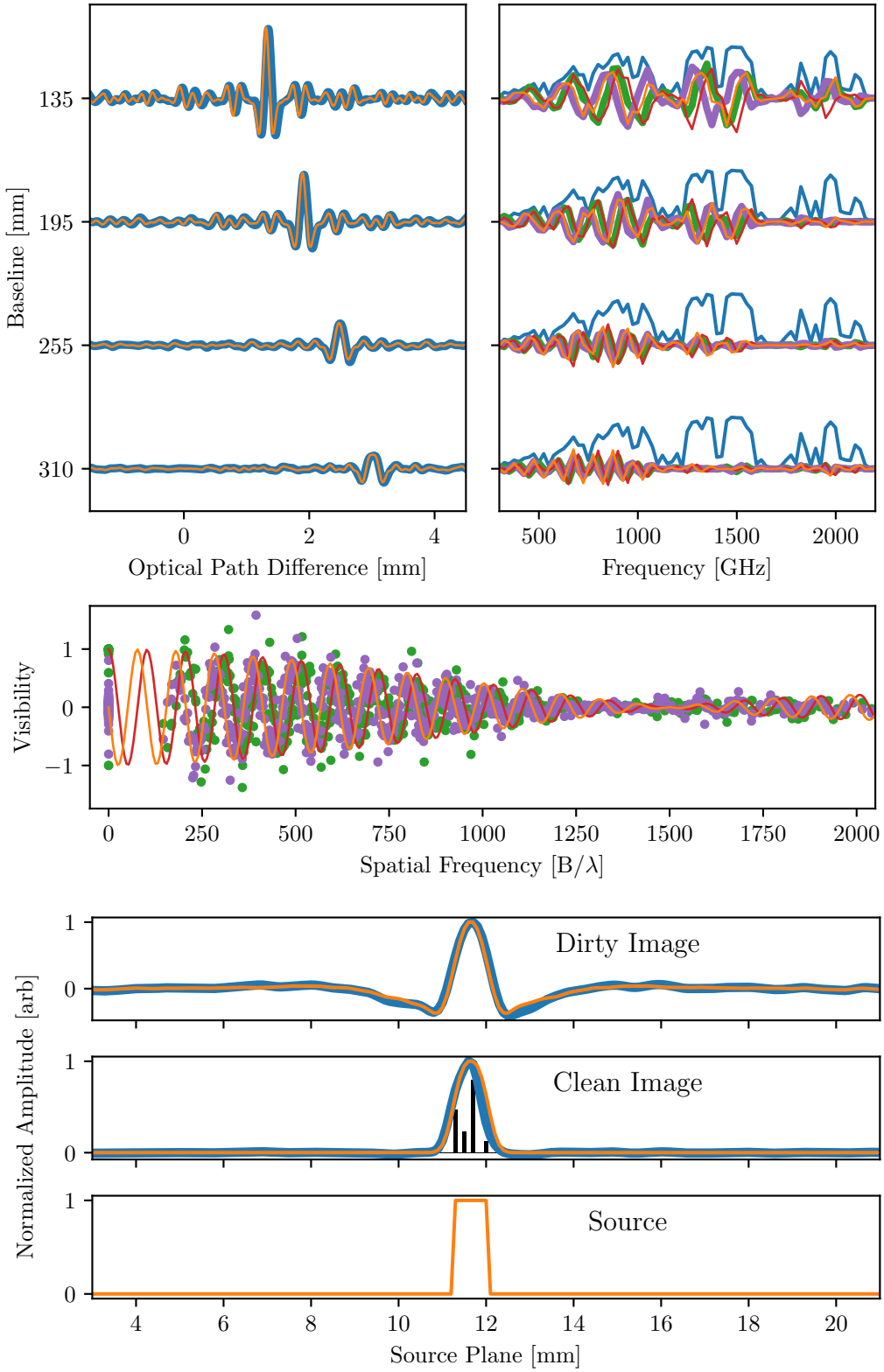


Figure 6.18: Wide field A3 results. Model is a 0.7 mm slit centered at 11.6 mm.

are identical. As such, it is appropriate to simply add the two signals in wide field applications. Although each detector may have their own systematics requiring detector specific phase, beam, or sensitivity calibration products, these results are a compelling demonstration that detector specific processing for the purpose of visibility extraction and *uv*-plane population is not required.

The differences between interferograms in Fig. 6.19 can be explained by the position of the slit with respect to the detector beams. The slit is nominally a rectangle, but when the slit is near the edge of the beam, the side nearest the edge of the beam is attenuated. From the perspective of the detector, the rectangular slit may then appear as something like a trapezoid. The asymmetry of the apparent source geometry produces an asymmetric interferogram as can be seen for the A0 detector in the top panel of Fig. 6.19. For the adjacent A3 detector, the other side of the slit is attenuated, and the asymmetry is reflected. The difference in amplitude of the interferograms is explained by the beam sensitivity profile. In both cases shown in Fig. 6.19, the slit was paced closer to the center of the A0 detector than the adjacent detector, resulting in better sensitivity and higher amplitude for the A0 interferogram.

A further caveat must be noted. When performing nominal phase center calibration on the A0 detector it was determined that a 0.28 ± 0.03 mm offset was required for the OPD axis (see § 5.2.2). Repeating this analysis on the A0 data obtained for Fig. 6.19 gave a different result. The same analysis was performed on the A3 and C2 detector data. This analysis gave OPD corrections of 0.15 ± 0.03 mm for A0/A3 and 0.42 ± 0.02 mm for A0/C2. Evidently, the phase center correction depends on the position of the source used during the calibration, but given a particular source position, the correction is the same for adjacent, and potentially all, detectors. This discrepancy was not anticipated from theory and is attributed to imperfections in the optical system resulting from the collimating mirror's aberrations. It is emphasized that the OPD correction is asymmetric about the optical axis which is characteristic of modelled aberrations (see Fig. 3.14), and if this correction was

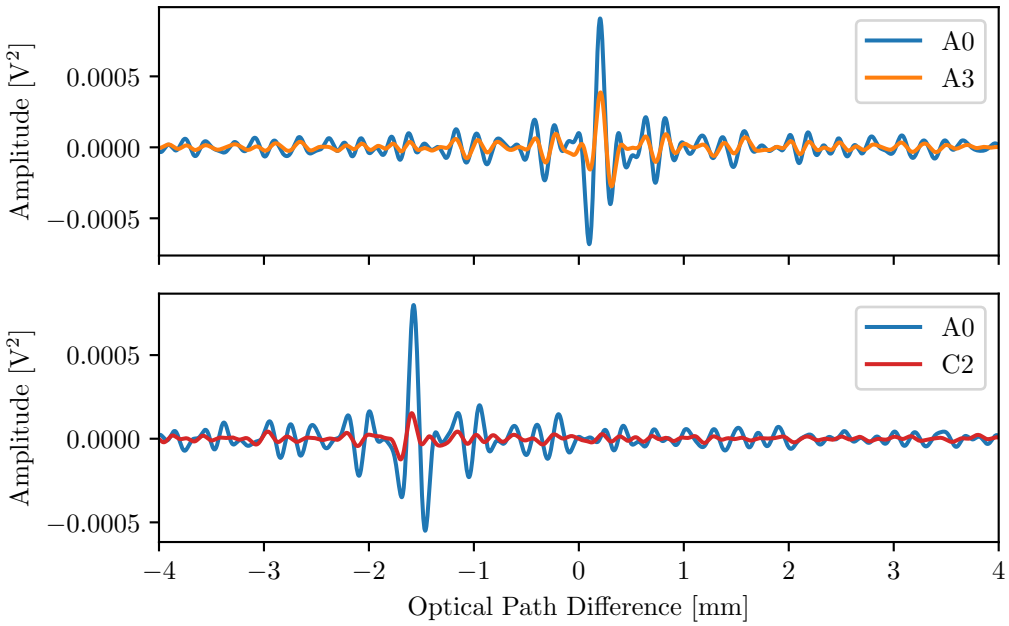


Figure 6.19: Wide field source overlap interferograms. Interferograms were obtained using a 135 mm baseline and by placing a 1 mm slit in the Field of View (FOV) of adjacent detectors. The interferograms within each panel are nearly identical apart from a scaling factor and some asymmetry, both resulting from the slit being near the edges of the detector beams. The equivalence of interferograms when viewing the same source suggests detector specific processing is not needed for off-axis detectors in wide field imaging. Interferograms from different detectors can simply be added.

a necessary aspect of the ideal instrument, the correction would be symmetric about the optical axis.

When conducting wide field observations with a DFI, it is generally necessary to account for well known obliquity effects of an FTS instrument [100, 105]. In particular, when viewing an off-axis source that is at an angle α with respect to the optical axis, the OPD is modified to

$$z' = z\cos(\alpha)\,, \tag{6.10}$$

where z is the OPD for an on-axis source. This obliquity factor typically increases instrument bandwidth while reciprocally decreases spectral resolution. Although these effects can become significant, in this work, the cosine correction is on the order of 10^{-5} and is not considered with nominal processing.

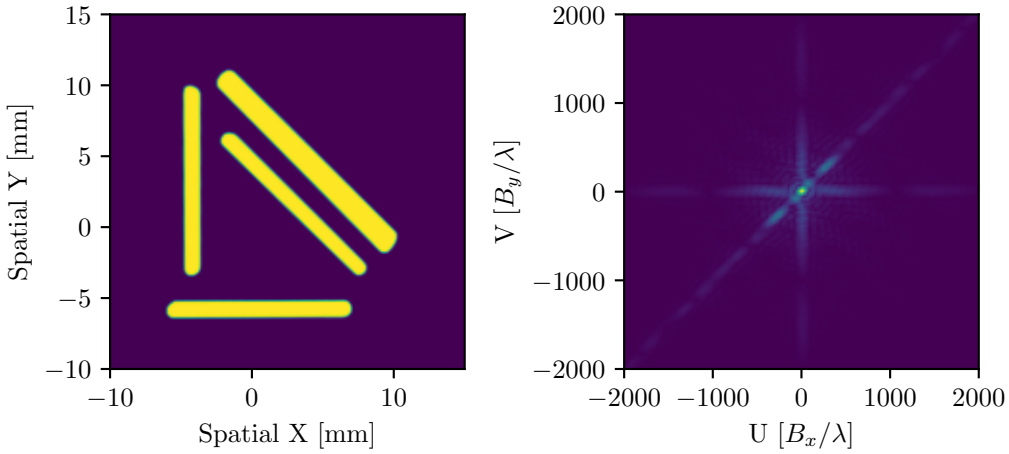


Figure 6.20: Simulated two dimensional slit mask and associated power spectrum. The left panel shows an approximate model for the two dimensional slit mask used in this experiment. The left panel shows the absolute value of the Fourier transform of the simulated source, and represents the energy distribution in the *uv*-plane. It is observed that most of the energy is along the 0°, 45°, and 90° lines that intersect the origin. Sampling these regions thus provides the most information and has the lowest sensitivity requirements.

6.4 Two Dimensional Source

The experiments discussed so far have been conducted using a constant baseline orientation. Since the DFI is only sensitive to spatial variation on the source plane along the direction parallel to the baseline vector, this configuration is only suitable for one dimensional sources. In order to probe two dimensional spatial variations, the baseline orientation must be changed. In the case of a real telescope, this is accomplished by physical rotation of the instrument by an angle $\theta_{\vec{B}}$. With the testbed system, however, baseline rotation is not practical, and instead, the aperture mask is rotated by $\theta_s = -\theta_{\vec{B}}$.

The two dimensional aperture mask used for these measurements was shown in Fig. 5.1, with an approximate model shown in the left panel of Fig. 6.20. The target consists of a variety of slits at 0°, 45°, and 90°. The right panel of Fig. 6.20 shows the absolute value of the Fourier transform of the target model. It is evident that most of the energy in the *uv*-plane resides along lines, rotated by the previously mentioned angles, that intersect the origin. As such, baseline orientations that sample these axes are expected to be most productive and least demanding in terms of Signal to Noise Ratio (SNR) requirements.

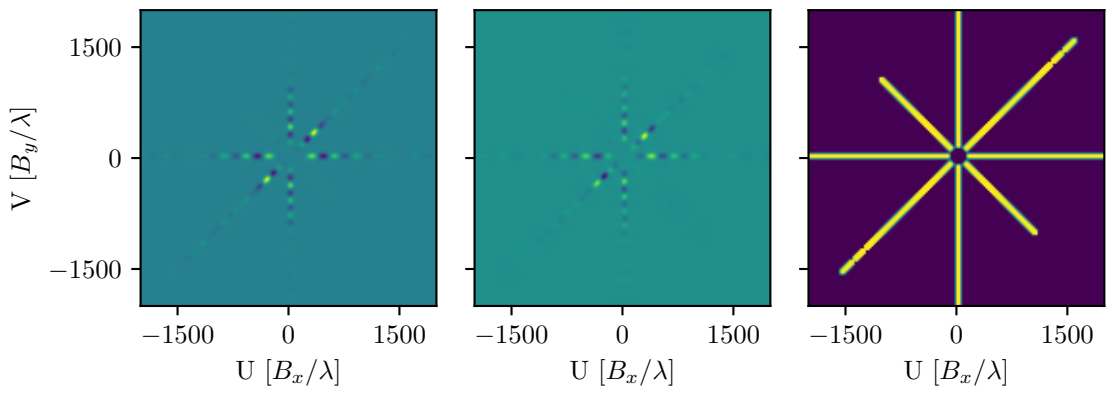


Figure 6.21: Two dimensional *uv* sampling. The left and center panels show the real and complex *uv* samples after regridding, respectively. The right panel shoes the *uv*-plane coverage for this experiment.

Observations and data processing were conducted in the same way as described in Ch. 5. The main point to emphasize is that, before regridding, complex visibility samples were projected onto the *uv*-plane using the typical *uv* basis mapping in Eq. 5.18. In previous observations the baseline angle was always zero, which is not the case here. Real and imaginary *uv* components after regridding, for the sampled baseline orientations, are shown in Fig. 6.21 with the right panel of the figure showing the *uv* coverage for this experiment. In addition to the baseline rotation angles previously mentioned, the data presented includes measurements taken at $\theta_{\vec{B}} = 135^\circ$. The resulting *uv* samples are of such low amplitude that they barely register on the complex *uv*-plane. This result was expected based on the model *uv*-plane in Fig. 6.20.

The synthesized dirty image is shown in left panel of Fig. 6.22. Due to the extended nature of the source, the clean algorithm performed poorly and the results are not shown. Although the dirty image fails to distinguish the length of each slit, the positions, widths, and orientations of the slits are well represented. In the right panel of Fig. 6.22 the dirty image is superimposed on top of a picture of the aperture masked used in this experiment. It is clear that the bright bands are an accurate representation of the source structure.

In order to improve image quality and distinguish the lengths of the slits, more *uv*-coverage is necessary. The full dataset for the two dimensional source experiment includes

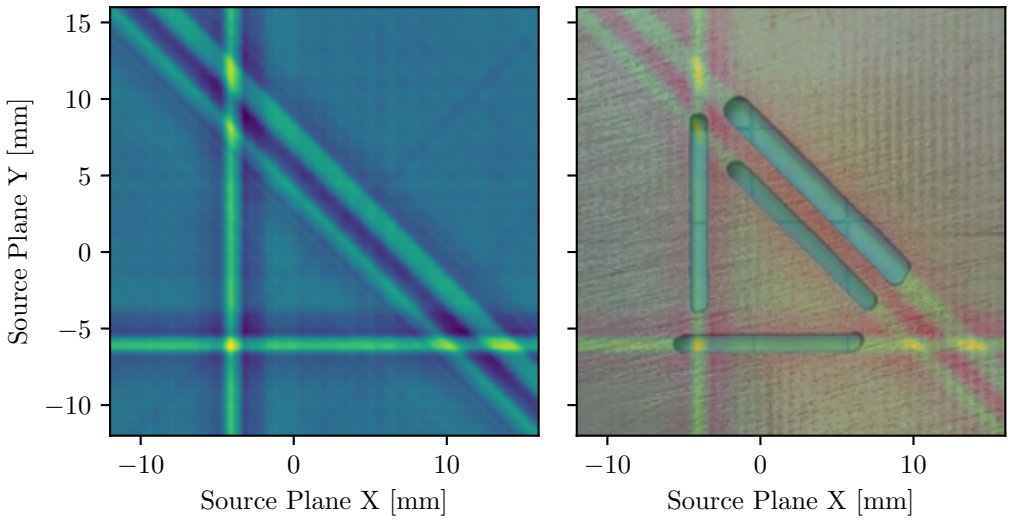


Figure 6.22: Two dimensional image reconstruction. The left panel shows the resulting dirty image, which is compared to a picture of the slit mask used in this experiment. Although the lengths of the slits are not evident in the dirty image, the widths, orientations, and relative positions are well represented by the dirty image.

baseline rotations between $0^\circ - 162^\circ$ at intervals of 18°. Unfortunately, including these results did not improve the quality of the image and only succeeded in increasing the noise. However, it should be possible to improve image quality with this system, even accounting for low SNR. A dirty image constructed by applying the same *uv* coverage used in this experiment to the model *uv*-plane is shown in the center panel of Fig. 6.23. This can be compared to the experimental image shown in the left panel. Both images are very similar in quality which suggests the experimental results are consistent with theory and the image obtained is precisely what would be expected with the given *uv* coverage. The right panel is a synthesized dirty image which has been constructed with *uv* components residing within an annulus defined by $150 \leq \sqrt{u^2+v^2} \leq 1,700$, and is a *uv* space easily accessible by the testbed DFI. Additionally, Gaussian noise with a standard deviation of 0.2 has been applied to this complex visibility space, which is comparable to the complex visibility noise measured with previous experiments in this chapter. Under these conditions, the slits are fully imaged in simulation.

In order to recreate such results in the future, it is necessary to develop an improved

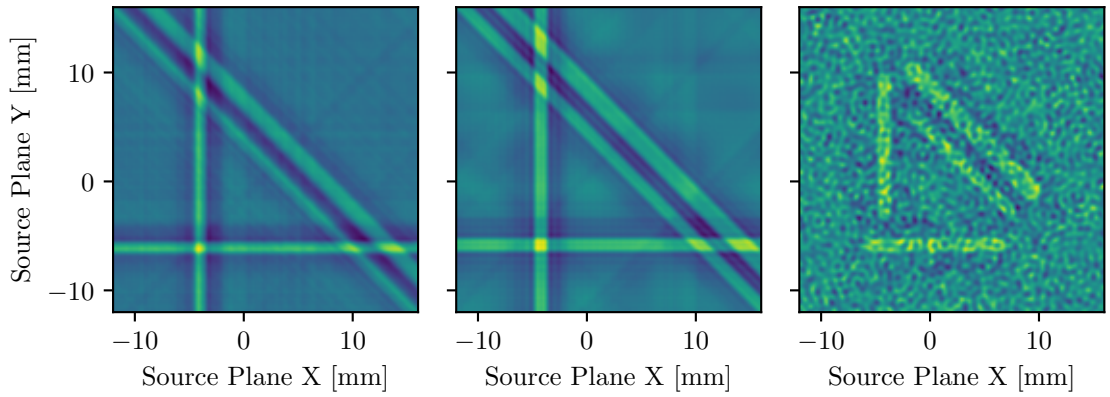


Figure 6.23: Two dimensional dirty image compared to model. The left panel shows the experimental dirty image, and the center panel shows the model dirty image with the same *uv*-plane sampling. Both images are similar, suggesting that experimental results for two dimensional sources agree with theory. The right panel shows a theoretical image reconstruction with full *uv* sampling within the annular region $150 \leq \sqrt{u^2+v^2} \leq 1,700$, which is accessible by the testbed DFI. Complex visibility samples have been subjected to Gaussian noise with a standard deviation of 0.2, similar to the visibility noise observed in this work. Even with this distortion, the source is well resolved, and such performance is expected to be possible with improved slit mask rotation infrastructure.

aperture mask mounting strategy where the rotation of the mask can be controlled, ideally though automation, to high precision. Part of the reason why adding additional baseline orientations to the experimental dataset did not improve image quality was because the rotation was only accurate to within $\pm 3^\circ$ and it is suspected that the position of the aperture mask was shifted during rotation. With these errors, spatial frequencies added to the *uv*-plane cannot be expected to contribute to the image in a productive way. Using the heuristic criteria $\lambda/10 \geq r\,\sigma_{\theta_{\vec{B}}}$, which states that the error in the source position a distance r from the optical axis as a function of baseline rotation should be less than $\lambda/10$. For the central detector with beam extent $r \approx 10\,\text{mm}$ and using the highest frequency of light in the instrument bandwidth $\nu = 2,200\,\text{GHz}$, the required rotational precision is $\sigma_{\theta_{\vec{B}}} \leq 0.1^\circ$, which must be further constrained for wide field observations. However, greater rotational precision may be necessary in some applications (e.g., $\lambda/100$ criteria). Another benefit of an automated mounting solution is that it facilitates more uniform *uv* coverage. With manual rotation, the only practical *uv* sampling is lines that extend out of the origin of the *uv*-plane,

as in Fig. 6.21. This inevitably leaves large gaps further away from the origin, which may, as a whole, contribute significant information to the source intensity distribution.

Although wide field observations were not performed at multiple baseline orientations, the discussion in § 6.3.3 is easily generalized to two dimensional sources. The key insight from § 6.3.3 was that data can be combined from multiple detectors through summation in any of the interferogram, spectral, visibility, or image domains. This statement applies equally to all baseline orientations. The crucially important difference is that complex visibility samples must be correctly mapped to the *uv*-plane using Eq. 5.18. A potential complication arises when there is no overlap between detector beams on the source plane. Such sub-Nyquist sampling leaves gaps in the source plane that are typically amended by slight adjustments to the telescope pointing. Such pointing variations change the phase center of the instrument, which subsequently affect the phase of the measured complex visibilities. Measured interferograms can be corrected for pointing by applying an appropriate baseline dependent offset to the OPD axis. However, this is not the best option since the required shift may not be an integer multiple of the OPD increments and would require interpolation in order to put all observations onto a common OPD axis. A better solution is to use the translation property of Fourier transforms. That is, translation in the interferogram domain can be accomplished in the spectral or visibility domain by applying an appropriate linear phase shift through simple multiplication to each sample point. In this way, a common *uv*-plane can be accurately populated with data obtained from multiple detectors, baseline lengths and orientations, and telescope pointings. Once a dirty or clean image has been synthesized from these measurements, it is still necessary to correct for the source plane sensitivity. This will require an accurate model of each detector beam for the full detector array. If the instrument is configured such that each detector has constant pointing for the full duration of the observation, as was the case for the one dimensional targets in this work, the image can be divided by the sensitivity map. If the detector pointing changes during the observation, either from rotation or translation of the instrument, a composite sensitivity

map must be constructed by iteratively multiplying the source plane by the nominal beam sensitivity map rotated or translated by the same amount as the instrument. The image is then corrected by dividing it by the composite sensitivity map.

6.5 Summary

In this chapter, aperture synthesis was demonstrated using the testbed DFI system for a variety of interesting situations. The nature of complex visibility amplitude and phase were demonstrated independently through monochromatic photomixer and VDI source observations. These experiments revealed how source extent and position are related to visibility amplitude and phase, respectively. Aperture synthesis was performed separately for sources that were spectrally uniform, spectrally nonuniform, and extended over an FOV greater than could be observed with a single detector. These results demonstrated the enhanced spatial resolution, spectral imaging, and wide field extension of the DFI technique, respectively. Observations of two dimensional sources were also presented with discussions on how to populate the *uv*-plane for arbitrary baseline length, orientation, and instrument pointing. Although the individual experiments performed were simple test cases, taken as a whole, they demonstrated all the critical observing requirements for a more sophisticated space based instrument.

In general, experimental results were found to be in excellent agreement with noise free models constructed using the ideal theoretical DFI response. Slight deviation were observed including systematic under estimates of the source size, loss of visibility at higher frequencies for measurements using larger baselines, and slight baseline dependent shifts in recorded interferograms. These deviations are well explained by the geometry of the aperture masks with respect to the collimating mirror, increased alignment error at larger baselines, and deviations in collimating mirror curvature, respectively. For wide field observations, aberrations resulting from the collimating and focusing mirrors produced source plane dependent shifts in the interferometer phase center. Dirty images obtained when

viewing a two dimensional source are in good agreement with simulated results using the same *uv*-plane sampling. Although increasing *uv* coverage did not significantly improve image quality, it is expected that a higher precision aperture mask mounting and rotation solution will provide a significant improvement, even when assuming substantial noise in the measured complex visibility samples. In summary, the testbed DFI system produced the expected theoretical result under a variety of unique test cases, and the slight deviations observed are explained by experimental error associated with imperfections of the optical system.

Chapter 7

Conclusion

Although the Double-Fourier Interferometer (DFI) technique has been previously studied by other research groups, a comprehensive demonstration of its observing capabilities has yet to be provided. In particular, it must be shown that complex visibilities can be obtained using the instrument, and that these complex visibilities can be successfully used in standard aperture synthesis techniques to produce images with enhanced spatial resolution. It must be shown that the instrument is sensitive to spectral variations in the source and is capable of producing spectral images with enhanced spatial resolution. Finally, it must be shown that the technique can be extended to wide field applications which exploit an array of detectors sensitive to different regions of the source plane.

7.1 Thesis Summary

The primary result of this thesis was the successful independent demonstration of the key observational capabilities of the DFI technique. More broadly this work serves as a convenient reference for the theoretical foundation and practical implementation of a DFI system. This was achieved with a presentation of Fourier analysis in which was derived the relevant theorems and transforms necessary for the interpretation and analysis of interferometric data. Additional emphasis was placed on the discretization of signals and how this sampling affects the Fourier transform. These tools were subsequently applied in the context of Fourier Transform Spectrometer (FTS) and spatial interferometry to develop an understanding of the instrument response and how it relates to the sources being observed.

Much of this thesis was dedicated to providing a detailed report on the design and characterization of the particular DFI system used in this work. The design was initially guided by geometric modeling and ray tracing simulations in order to meet specific observation requirements. This design was validated with empirical measurements and found to be in good agreement with expectations. Relevant figures of merit were summarized in Tab. 3.1. Each component of the system was quantitatively characterized to ensure sufficient quality such that the overall performance of the system would be capable of combining the electric field from both arms of the interferometer coherently over the full instrument bandwidth. Critically, the precision and repeatability of the linear translation stages were found to be suitable for averaging thousands of scans, which was required to achieve appreciable Signal to Noise Ratio (SNR) in the measured interferograms.

The detector system used in this work consists of a 5×5 array of feed horn coupled Transition Edge Sensor (TES) bolometers that are read out using time division multiplexing over four channels. Unlike conventional bolometers, in this case, the TES element temperature is regulated by a thermally coupled heater resistor. The voltage applied across the resistor is used as the detector output and is controlled by a Proportional-Integral–Derivative (PID) controller that tries to maintain a particular voltage drop across the current biased TES element, which increases with TES temperature. Increased optical loading reduces the heater voltage required to maintain the thermal balance, and thus the detector output is inversely proportional to the incident optical power. Figures of merit for these detectors including Noise Equivalent Power (NEP), frequency response, and spectral response were measured experimentally. Also included was a detailed discussion of the detector readout, the configurable detector parameters, and multimoded throughput of the feed horn waveguides.

Discussion of the interferometer system was followed by a detailed explanation of how observations were conducted and how data was processed. Full datasets consist of FTS and DFI interferometric measurements and photometric calibration measurements at each aper-

ture position. Data calibration applies corrections for intensity variations between the two arms of the interferometer, removes non-linear instrumental phase, registers interferograms to the spatial interferometer phase center, and corrects for variations in sensitivity due to the detector beam profile. Complex visibility values were extracted by normalizing the generally complex DFI spectra by the fully real FTS spectrum. This frequency dependent complex visibility extraction process is the critical aspect of the DFI instrument that allows for the synthesis of spectral images with enhanced spatial resolution. The processing steps for aperture synthesis were discussed with emphasis on how irregularly sampled data can be projected onto a uniform grid for efficient computation of the inverse Fourier transform.

Finally, the results for various DFI observations were presented and discussed. Photomixer observations were used to demonstrate how the complex visibility amplitude varies as a function of source extent and baseline. Virginia Diodes (VDI) observations were used to demonstrate how the complex visibility phase varies as a function of source position and baseline. These two datasets independently verify the two aspects of complex visibility and help to communicate how this quantity is sensitive to the source intensity distribution. The rest of the observations independently demonstrate the key observational capabilities of the DFI technique using aperture masks with various geometries backlit by a thermal source. It was successfully demonstrated that aperture synthesis techniques can be applied to DFI measurements producing one and two dimensional images that are sensitive to spatial features much smaller than an individual detector beam. Masking part of the scene with a low-pass optical filter imposed spectral variation on the source plane. It was demonstrated that the DFI instrument was sensitive to this spectral variation, producing images with different and accurate source intensity distributions on either side of the low-pass filter cutoff frequency. Image reconstruction was also demonstrated for an extended source with structure larger than the size of an individual detector beam. This was followed by an evidence based discussion on how to combine visibility samples in the *uv*-plane for measurements with arbitrary baseline orientation and instrument pointing.

The results summarized in the previous paragraph were found to be in excellent agreement with theory, providing compelling evidence that the DFI technique can fulfill the needs of the Far-Infrared (FIR) community. The small deviations from expected results are well explained by experimental errors attributed to imperfect optical alignment, and distortions resulting from the geometry of the instrument collimating and focusing optics.

7.2 Future Work

The logical next step with respect to experimental validation of the DFI technique is to demonstrate all of the relevant observing capabilities simultaneously instead of independently. Extending observations to sources with more complicated spatial and spectral structure should also be a priority. Such a task is largely infeasible with the testbed system used in this work as it is currently configured. The accessible baselines are too restrictive, the detector sensitivity is too low, and distortion due to aberrations are too high. These limitations are precisely why this work has been structured around spectrally uniform, one dimensional slit masks, imaged using only the central detector. Only one of these factors was allowed to be varied at a time, and only in a restricted way. Bellow are discussed a few ways in which the system can potentially be modified to accommodate more complicated observations.

Generating more complicated scenes at FIR wavelengths in itself is challenging. One options is to cut out more complicated shapes into the aperture masks and overlay a greater variety of low-pass, high-pass, and band-pass filters. However, this limits the shapes which can be used. An annulus, for example, would not be possible. Additionally, acquiring the necessary filters might be difficult. Alternatively, a scene constructed from a collection of wires heated to different temperatures may be viable. This option avoids the issues with using masking material with finite thickness, and allows for wider emission cones. However, rotating such an apparatus is somewhat more challenging and the spectral variation contrast will be lower. For example, the percent difference in intensity between an 800 °C and

$1,200\,^{\circ}$C thermal source at 1 THz is only 27%, which may be difficult to unambiguously distinguish with the present detector sensitivity.

On the topic of sources, a better source mounting strategy is needed. The unsophisticated solution used in this work is subject to source position and source rotation precision limited by visual alignment and manual adjustment. This lack of precision become significant when imaging two dimensional sources where sub-degree and sub-mm rotation and translation errors, respectively, prevent spatial frequency samples from adding productively to the synthesized image.

Instrument alignment can be improved. Comparing the intensity through each arm of the interferometer to the peak of the white light fringe in FTS mode suggests a $\sim 20\%$ reduction in fringe visibility. Since FTS interferograms are ideally insensitive to spatial coherence effects, the loss of visibility is attributed to alignment error. Aligning an optical system with up to 18 optical components, each with multiple degrees of freedom, with respect to invisible light, is inevitably going to be a challenge. However, nontrivial improvement seems possible.

Extending baseline coverage requires a larger collimating mirror. Obtaining a large mirror is an expensive proposition but cost can be reduced by using an on-axis spherical mirror instead of an off-axis parabolic mirror. An issue with using a larger collimating mirror for a given focal length is that the emission angle of the source must be increased so as to fully illuminate the mirror. As discussed in § 6.3.1, the apparent source size decreases with viewing angle, and as discussed in § 3.2.4, the emission angle of the backlit aperture mask is limited. Thus, accommodating a larger mirror will require an increase in the focal length of the mirror, which will decrease optical throughput. Integrating a larger on-axis (or slightly off-axis) spherical mirror with a larger focal length into the DFI system may require adjustment of the optical layout, and both source and collimating mirror may not be accommodated by a single optics table.

Increasing the unvignetted Field of View (FOV) and reducing the effect of aberrations

will be important for future wide field observations. This will allow for the periphery detectors to be used productively. A simple way to accomplish this is to decrease the minification of the optical system which can be accomplished most easily by increasing the focal length of the focusing mirror (see Eq. 3.13). For example, if the focal length is increased by a factor of three, the nine central detectors of the detector array will be sensitive to the same region of the source plane as the single central detector currently is. Distortions due to the imaging optics over all nine detectors would then be expected to be similar to the distortion currently observed over the central detector. Unfortunately, since throughput is proportional to the detector FOV, this modification will have a sensitivity penalty.

Extending baseline coverage and more productive detector utilization, using the suggested modifications, both reduce throughput and SNR when the current system SNR is already very low. SNR can be improved with better alignment and averaging significantly more interferograms. There is hope for improved detector performance as well. Although not discussed in Ch. 4, a survey of detector parameters reveals that an increase in SNR by a factor of four, and a cutoff frequency increase to 300 Hz is possible. Using the current optical design, this improvement implies that sixteen times fewer scans are required to obtained the same quality of results, and scans can be acquired $\sim$40% faster.

Improvements may also be possible with a different observing mode. In this work, the "rapid scan" approach was used where the spectral stage is moved at a constant speed and data is continuously recorded as a function of Optical Path Difference (OPD). The stage speed in combination with the $\sim$13 kHz detector sampling rate results in an optical Nyquist frequency of 72 THz, roughly 34 times larger the highest frequency in the instrument bandwidth. Alternatively, the "step and integrate" strategy integrates signal at specified OPD intervals which can be configured so as to more closely match the Nyquist frequency to the instrument bandwidth. The increases bandwidth efficiency may reduce the observing time required to obtain results of similar quality. However, the step and integrate method is sen-

sitive to low frequency noise which is typically overcome by modulating the optical signal and feeding the analogue detector signal to a lock-in amplifier. The detector signal in this work is intrinsically digitized and cannot be processed in this way. An analogous approach is to modulated the optical signal, average the detector signal over many wave forms, and use the peak-to-peak amplitude as an estimate for the interferogram measurement. The time requried for a full observation can be calculated as

$$t_{obs} = \frac{L}{\Delta L} \times \left(\frac{n_{samp}}{f_{mod}} + t_{move} \right) . \tag{7.1}$$

The number of samples taken at a particular OPD is $n_{samp} = 4,000$, and is given by the number of scan in rapid scan mode, or the number of averaged waveforms in step and integrate mode. The frequency at which the optical signal is modulated is $f_{mod} = 100\,\mathrm{Hz}$, which should be well below the instrument cutoff frequency. The time to move the stage between OPD samples is $t_{move} = 0.5\,\mathrm{s}$. The distance the stage moves over the entire observation is $L = 4\,\mathrm{mm}$, which occurs in intervals of $\Delta L = 0.015\,\mathrm{mm}$. The values provided are those required to make the step and integrate mode comparable the rapid scan observations obtained in this work with a 2.5 THz Nyquist frequency. Evaluating Eq. 7.1 with these values gives an observation time of $\sim 3\,\mathrm{hr}$, which is about 30% longer than the rapid scan mode. However, there is some freedom to adjust the modulation frequency parameter which could significantly reduce observation time, and the way in which the modulated waveforms are analyzed may provided improved SNR not captured by this discussion. As such, this observing mode merits closer investigation and may provide new revelations.

7.3 Final Remarks

FIR astronomy is of critical importance for understanding the formation and evolution of galaxies, stars, and planets. During the earliest stages of formation, theses objects are invariably obscured by gas and dust which scatters or absorbs light with wavelengths less

than or comparable to the size of dust grains $\lesssim 1\,\mu$m. FIR radiation consists of light with wavelengths that are above this threshold and can pass largely unobstructed through these denser regions of the Interstellar Medium (ISM). FIR astronomy thus provides the ability to directly observe the structure and distribution of objects embedded within these regions.

Within the FIR band resides a wealth of ionic, atomic, and molecular emission features. Ionic and atomic features result from fine structure lines which are the product of electron spin state transitions. A number of lines including [CI], [CII], [OI], [OIII], and [NII] have rest frame emission within the FIR and probe the neutral and warm interstellar medium. Highly ionized species, produced in the vicinity of Active Galactic Nuclei (AGN), have fine structure lines with rest frame frequencies in close spectral proximity to the FIR including [ArII], [SIV], [NeV], and [OIV]. These lines are shifted into the FIR at redshifts $z \gtrsim 2$ during the peak of star formation and AGN activity. Molecular lines result from transitions in rotation and vibration states with species such as CO, HD, OH, H_2O, and NH_3 expressing many emission features within the FIR band. These features are associated with molecular gas and probe both the cold quiescent phase, and regions where gas is excited by feedback processes. Additionally, the FIR band corresponds to the peak emission of thermal radiation for sources at temperatures between $\sim 10 - 100\,$K. Mapping the distribution of interstellar dust, typically found at these temperatures, is most productively conducted in the FIR. As such, the FIR probes the full range of physical conditions within the ISM and is often the most productive band when studying the chemistry and physical processes within many astronomical structures.

Despite the merits of FIR astronomy, this region remains relatively unexplored compared to the optical and radio bands which surround it. This is primarily due to three reasons. First, although the ability to observe interstellar water is one of the features of FIR astronomy, the abundance of water in earth's atmosphere results in low transmission of FIR radiation to earth's surface. As such, observations must generally be conducted in space. Second, FIR detector technology is comparatively immature and lacks the de-

sired sensitivity. Only within the past decade have background limited detector arrays been feasible, and fully leveraging this sensitivity requires telescopes that are cooled to $\sim 4\,\mathrm{K}$. Third, the spatial resolution of diffraction limited imaging systems is proportional to the wavelength of light being observed. FIR wavelengths are of such length that it is wildly impractical to build a single aperture system of sufficient size to achieve the sub-arcsecond spatial resolution necessary to resolve the astronomical targets of interest.

With the proposed spatial resolution: active star forming regions can be distinguished from AGN during the period of reionization to better understand their coevolution; prestellar cores of massive stars can be resolved in the local universe under metallicity conditions similar to those expected during the peak of star formation; stellar environments in the local galaxy can be resolved to 1 AU allowing the search for signatures of protoplanets within protoplanetary disks; mapping the dynamics, chemistry, and the distribution of water in protoplanetary disks, providing unprecedented insight into where and how habitable planets form. These observations require not only higher spatial resolution, but also access to the full FIR band. The most promising instrument for meeting such requirements is a DFI which combines the spatial resolution enhancements of a spatial interferometer with the spectral imaging capabilities of an FTS. This work has demonstrated the key observing capabilities of a DFI, and provides compelling experimental support for future space-based DFI missions like Far-InfraRed Interferometer (FIRI), Space Interferometer for Cosmic Evolution (SPICE), and Submillimeter Probe of the Evolution of Cosmic Structure (SPECS).

Bibliography

[1] National Research Council, Division on Engineering, Physical Sciences, Space Studies Board, Board on Physics, and Astrophysics Survey Committee. *Astronomy and Astrophysics in the New Millennium: Panel Reports*. National Academies Press.

[2] David Leisawitz, Charles Baker, Amy Barger, Dominic Benford, Andrew Blain, Rob Boyle, Richard Broderick, Jason Budinoff, John Carpenter, Richard Caverly, Phil Chen, Steve Cooley, Christine Cottingham, Julie Crooke, Dave DiPietro, Mike DiPirro, Michael Femiano, Art Ferrer, Jacqueline Fischer, Jonathan P. Gardner, Lou Hallock, Kenny Harris, Kate Hartman, Martin Harwit, Lynne Hillenbrand, Tupper Hyde, Drew Jones, Jim Kellogg, Alan Kogut, Marc Kuchner, Bill Lawson, Javier Lecha, Maria Lecha, Amy Mainzer, Jim Mannion, Anthony Martino, Paul Mason, John Mather, Gibran McDonald, Rick Mills, Lee Mundy, Stan Ollendorf, Joe Pellicciotti, Dave Quinn, Kirk Rhee, Stephen Rinehart, Tim Sauerwine, Robert Silverberg, Terry Smith, Gordon Stacey, H. Philip Stahl, Johannes Staguhn, Steve Tompkins, June Tveekrem, Sheila Wall, and Mark Wilson. The space infrared interferometric telescope (SPIRIT): High-resolution imaging and spectroscopy in the far-infrared. 40(5):689–703.

[3] Frank Helmich and Rob Ivison. FIRI - a far-infrared interferometer. 23(1):245–276.

[4] Duncan Farrah, Kimberly Ennico Smith, David Ardila, Charles M. Bradford, Michael Dipirro, Carl Ferkinhoff, Jason Glenn, Paul Goldsmith, David Leisawitz, Thomas Nikola, Naseem Rangwala, Stephen A. Rinehart, Johannes Staguhn, Michael Zemcov, Jonas Zmuidzinas, James Bartlett, Sean Carey, William J. Fischer, Julia Kamenetzky, Jeyhan Kartaltepe, Mark Lacy, Dariusz C. Lis, Lisa Locke, Enrique Lopez-Rodriguez, Meredith MacGregor, Elisabeth Mills, S. Harvey Moseley, Eric J. Murphy, Alan Rhodes, Matt Richter, Dimitra Rigopoulou, David Sanders, Ravi Sankrit, Giorgio Savini, John-David Smith, and Sabrina Stierwalt. Review: Far-infrared instrumentation and technology development for the next decade. 5(2):1.

[5] Pau Amaro-Seoane, Heather Audley, Stanislav Babak, John Baker, Enrico Barausse, Peter Bender, Emanuele Berti, Pierre Binetruy, Michael Born, Daniele Bortoluzzi, Jordan Camp, Chiara Caprini, Vitor Cardoso, Monica Colpi, John Conklin, Neil Cornish, Curt Cutler, Karsten Danzmann, Rita Dolesi, Luigi Ferraioli, Valerio Ferroni, Ewan Fitzsimons, Jonathan Gair, Lluis Gesa Bote, Domenico Giardini, Ferran Gibert, Catia Grimani, Hubert Halloin, Gerhard Heinzel, Thomas Hertog, Martin Hewitson, Kelly Holley-Bockelmann, Daniel Hollington, Mauro Hueller, Henri Inchauspe, Philippe Jetzer, Nikos Karnesis, Christian Killow, Antoine Klein, Bill Klipstein, Natalia Korsakova, Shane L. Larson, Jeffrey Livas, Ivan Lloro, Nary Man, Davor

Mance, Joseph Martino, Ignacio Mateos, Kirk McKenzie, Sean T. McWilliams, Cole Miller, Guido Mueller, Germano Nardini, Gijs Nelemans, Miquel Nofrarias, Antoine Petiteau, Paolo Pivato, Eric Plagnol, Ed Porter, Jens Reiche, David Robertson, Norna Robertson, Elena Rossi, Giuliana Russano, Bernard Schutz, Alberto Sesana, David Shoemaker, Jacob Slutsky, Carlos F. Sopuerta, Tim Sumner, Nicola Tamanini, Ira Thorpe, Michael Troebs, Michele Vallisneri, Alberto Vecchio, Daniele Vetrugno, Stefano Vitale, Marta Volonteri, Gudrun Wanner, Harry Ward, Peter Wass, William Weber, John Ziemer, and Peter Zweifel. Laser interferometer space antenna.

[6] Andrew W. Blain, Ian Smail, R. J. Ivison, J. P. Kneib, and David T. Frayer. Submillimeter galaxies. 369(2):111–176.

[7] A. G. G. M. Tielens. *The Physics and Chemistry of the Interstellar Medium*. Cambridge University Press.

[8] Eugene Hecht. *Optics*. Pearson, 5th edition edition.

[9] Hendrik Linz, Henrik Beuther, Maryvonne Gerin, Javier R. Goicoechea, Frank Helmich, Oliver Krause, Yao Liu, Sergio Molinari, Volker Ossenkopf-Okada, Jorge Pineda, Marc Sauvage, Eva Schinnerer, Floris van der Tak, Martina Wiedner, Jerome Amiaux, Divya Bhatia, Luisa Buinhas, Gilles Durand, Roger Förstner, Urs Graf, and Matthias Lezius. Bringing high spatial resolution to the far-infrared: A giant leap for astrophysics. 51(3):661–697.

[10] P. Mazumdar, F. Wyrowski, D. Colombo, J. S. Urquhart, M. A. Thompson, and K. M. Menten. High-resolution LAsMA ^{12} CO and ^{13} CO observation of the g305 giant molecular cloud complex: I. feedback on the molecular gas. 650:A164.

[11] Peter F Bernath. *Spectra of atoms and molecules*. Oxford university press.

[12] L. Page, G. Hinshaw, E. Komatsu, M. R. Nolta, D. N. Spergel, C. L. Bennett, C. Barnes, R. Bean, O. Dore, J. Dunkley, M. Halpern, R. S. Hill, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, H. V. Peiris, G. S. Tucker, L. Verde, J. L. Weiland, E. Wollack, and E. L. Wright. Three-year *Wilkinson Microwave Anisotropy Probe* (*WMAP*) observations: Polarization analysis. 170(2):335–376.

[13] Piero Madau and Mark Dickinson. Cosmic star-formation history. 52(1):415–486.

[14] A Moullet, T Kataria, D Lis, S Unwin, Y Hasegawa, E Mills, C Battersby, A Roc, and M Meixner. PRIMA general observer science book.

[15] N. Rando, D. Lumb, M. Bavdaz, D. Martin, and T. Peacock. Space science applications of cryogenic detectors. 522(1):62–68.

[16] Kent D Irwin and Gene C Hilton. Transition-edge sensors. pages 63–150.

[17] K. D. Irwin. An application of electrothermal feedback for high resolution cryogenic particle detection. 66(15):1998–2000.

[18] John Clarke and Alex I Braginski. *The SQUID Handbook: Fundamentals and Technology of SQUIDs and SQUID Systems*, volume 1. Wiley-Vch.

[19] Peter K. Day, Henry G. LeDuc, Benjamin A. Mazin, Anastasios Vayonakis, and Jonas Zmuidzinas. A broadband superconducting detector suitable for use in large arrays. 425(6960):817–821.

[20] M. D. Shaw, J. Bueno, P. Day, C. M. Bradford, and P. M. Echternach. Quantum capacitance detector: A pair-breaking radiation detector based on the single cooper-pair box. 79(14):144511.

[21] P. M. Echternach, B. J. Pepper, T. Reck, and C. M. Bradford. Single photon detection of 1.5 THz radiation with the quantum capacitance detector. 2(1):90–97.

[22] Jonathan P. Gardner, John C. Mather, Mark Clampin, Rene Doyon, Matthew A. Greenhouse, Heidi B. Hammel, John B. Hutchings, Peter Jakobsen, Simon J. Lilly, Knox S. Long, Jonathan I. Lunine, Mark J. McCaughrean, Matt Mountain, John Nella, George H. Rieke, Marcia J. Rieke, Hans-Walter Rix, Eric P. Smith, George Sonneborn, Massimo Stiavelli, H. S. Stockman, Rogier A. Windhorst, and Gillian S. Wright. The james webb space telescope. 123(4):485–606.

[23] The infrared astronomical satellite (IRAS) mission. 278.

[24] G. H. Rieke. History of infrared telescopes and astronomy. 25(1):125–141.

[25] Carol J Lonsdale, Duncan Farrah, and Harding E Smith. *Ultraluminous infrared galaxies*. Springer.

[26] H. H. Aumann. IRAS OBSERVATIONS OF MATTER AROUND NEARBY STARS. 97(596):885.

[27] MF Kessler, JA Steinz, ME Anderegg, J Clavel, G Drechsel, P Estaria, J Faelker, JR Riedinger, A Robson, BG Taylor, and others. The infrared space observatory (ISO) mission. 315:L27–L31.

[28] M. W. Werner, T. L. Roellig, F. J. Low, G. H. Rieke, M. Rieke, W. F. Hoffmann, E. Young, J. R. Houck, B. Brandl, G. G. Fazio, J. L. Hora, R. D. Gehrz, G. Helou, B. T. Soifer, J. Stauffer, J. Keene, P. Eisenhardt, D. Gallagher, T. N. Gautier, W. Irace, C. R. Lawrence, L. Simmons, J. E. Van Cleve, M. Jura, E. L. Wright, and D. P. Cruikshank. The spitzer space telescope mission. 154(1):1.

[29] G. L. Pilbratt, J. R. Riedinger, T. Passvogel, G. Crone, D. Doyle, U. Gageur, A. M. Heras, C. Jewell, L. Metcalfe, S. Ott, and M. Schmidt. Herschel space observatory - an ESA facility for far-infrared and submillimetre astronomy. 518:L1.

[30] F. P. Helmich, E. F. van Dishoeck, J. H. Black, T. de Graauw, D. A. Beintema, A. M. Heras, F. Lahuis, P. W. Morris, and E. A. Valentijn. Detection of hot, abundant water toward AFGL 2591. 315:L173–L176.

[31] D. Lutz, H. Feuchtgruber, R. Genzel, D. Kunze, D. Rigopoulou, H. W. W. Spoon, C. M. Wright, E. Egami, R. Katterloher, E. Sturm, E. Wieprecht, A. Sternberg, A. F. M. Moorwood, and T. de Graauw. SWS observations of the galactic center. 315:L269–L272.

[32] O. Laurent, I. F. Mirabel, V. Charmandaris, P. Gallais, S. C. Madden, M. Sauvage, L. Vigroux, and C. Cesarsky. Mid-infrared diagnostics to distinguish AGNs from starbursts. 359:887–899.

[33] Drake Deming, Sara Seager, L. Jeremy Richardson, and Joseph Harrington. Infrared radiation from an extrasolar planet. 434:740–743.

[34] Ph André, A. Men'shchikov, S. Bontemps, V. Könyves, F. Motte, N. Schneider, P. Didelon, V. Minier, P. Saraceno, D. Ward-Thompson, J. Di Francesco, G. White, S. Molinari, L. Testi, A. Abergel, M. Griffin, Th Henning, P. Royer, B. Merín, R. Vavrek, M. Attard, D. Arzoumanian, C. D. Wilson, P. Ade, H. Aussel, J.-P. Baluteau, M. Benedettini, J.-Ph Bernard, J. a. D. L. Blommaert, L. Cambrésy, P. Cox, A. Di Giorgio, P. Hargrave, M. Hennemann, M. Huang, J. Kirk, O. Krause, R. Launhardt, S. Leeks, J. Le Pennec, J. Z. Li, P. G. Martin, A. Maury, G. Olofsson, A. Omont, N. Peretto, S. Pezzuto, T. Prusti, H. Roussel, D. Russeil, M. Sauvage, B. Sibthorpe, A. Sicilia-Aguilar, L. Spinoglio, C. Waelkens, A. Woodcraft, and A. Zavagno. From filamentary clouds to prestellar cores to the stellar IMF: Initial highlights from the herschel gould belt survey. 518:L102.

[35] P. A. Oesch, G. Brammer, P. G. van Dokkum, G. D. Illingworth, R. J. Bouwens, I. Labbé, M. Franx, I. Momcheva, M. L. N. Ashby, G. G. Fazio, V. Gonzalez, B. Holden, D. Magee, R. E. Skelton, R. Smit, L. R. Spitler, M. Trenti, and S. P. Willner. A REMARKABLY LUMINOUS GALAXY AT z = 11.1 MEASURED WITH HUBBLE SPACE TELESCOPE GRISM SPECTROSCOPY. 819(2):129.

[36] Luisa Buinhas, Eloi Ferrer-Gil, and Roger Forstner. IRASSI: InfraRed astronomy satellite swarm interferometry — mission concept and description. In *2016 IEEE Aerospace Conference*, pages 1–20.

[37] S. Hocuk, L. Szűcs, P. Caselli, S. Cazaux, M. Spaans, and G. B. Esplugues. Parameterizing the interstellar dust temperature. 604:A58.

[38] G Savini. Far infra-red space interferometer critical assessment: Scientific definition and technology development for the next generation THz space interferometer.

[39] K. L. Luhman, P. R. Allen, C. Espaillat, L. Hartmann, and N. Calvet. THE DISK POPULATION OF THE TAURUS STAR-FORMING REGION. 186(1):111–174.

[40] B. A. Wilking, M. Gagné, and L. E. Allen. *Star Formation in the Rho Ophiuchi Molecular Cloud*, volume 5.

[41] August Muench, Konstantin Getman, Lynne Hillenbrand, and Thomas Preibisch. Star formation in the orion nebula i: stellar content.

[42] Frank H Shu. Self-similar collapse of isothermal spheres and star formation. 214:488–497.

[43] Fred C. Adams and Philip C. Myers. Modes of multiple star formation. 553(2):744.

[44] L. Ilsedore Cleeves. Zooming in on the chemistry of protoplanetary disks with ALMA. 13:57–68.

[45] Alwyn Wootten and A. Richard Thompson. The atacama large millimeter/submillimeter array. 97(8):1463–1471.

[46] JB Lovell, L Trapman, J Bergner, KD Doney, D Leisawitz, MK McClure, LG Mundy, J Najita, JA Sturm, D Wilner, and others. Tracing the physical and astrochemical assembly of planetary systems in the far-infrared at high angular resolution.

[47] T. Birnstiel, M. Fang, and A. Johansen. Dust evolution and the formation of planetesimals. 205(1):41–75.

[48] An introduction to galaxies and cosmology.

[49] N. Jarosik, C. L. Bennett, J. Dunkley, B. Gold, M. R. Greason, M. Halpern, R. S. Hill, G. Hinshaw, A. Kogut, E. Komatsu, D. Larson, M. Limon, S. S. Meyer, M. R. Nolta, N. Odegard, L. Page, K. M. Smith, D. N. Spergel, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. SEVEN-YEAR *WILKINSON MICROWAVE ANISOTROPY PROBE* (*WMAP*) OBSERVATIONS: SKY MAPS, SYSTEMATIC ERRORS, AND BASIC RESULTS. 192(2):14.

[50] P. a. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, N. Bartolo, E. Battaner, R. Battye, K. Benabed, A. Benoît, A. Benoit-Lévy, J.-P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, A. Bonaldi, L. Bonavera, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, R. C. Butler, E. Calabrese, J.-F. Cardoso, A. Catalano, A. Challinor, A. Chamballu, R.-R. Chary, H. C. Chiang, J. Chluba, P. R. Christensen, S. Church, D. L. Clements, S. Colombi, L. P. L. Colombo, C. Combet, A. Coulais, B. P. Crill, A. Curto, F. Cuttaia, L. Danese, R. D. Davies, R. J. Davis, P. de Bernardis, A. de Rosa, G. de Zotti, J. Delabrouille, F.-X. Désert, E. Di Valentino, C. Dickinson, J. M. Diego, K. Dolag, H. Dole, S. Donzelli, O. Doré, M. Douspis, A. Ducout, J. Dunkley, X. Dupac, G. Efstathiou, F. Elsner, T. A. Enßlin, H. K. Eriksen, M. Farhang, J. Fergusson, F. Finelli, O. Forni, M. Frailis, A. A. Fraisse, E. Franceschi, A. Frejsel, S. Galeotta, S. Galli, K. Ganga, C. Gauthier, M. Gerbino, T. Ghosh, M. Giard, Y. Giraud-Héraud, E. Giusarma, E. Gjerløw, J. González-Nuevo, K. M. Górski, S. Gratton, A. Gregorio, A. Gruppuso, J. E. Gudmundsson, J. Hamann, F. K. Hansen, D. Hanson, D. L. Harrison, G. Helou, S. Henrot-Versillé, C. Hernández-Monteagudo, D. Herranz, S. R. Hildebrandt, E. Hivon, M. Hobson, W. A. Holmes, A. Hornstrup, W. Hovest, Z. Huang, K. M. Huffenberger, G. Hurier, A. H. Jaffe, T. R. Jaffe, W. C. Jones, M. Juvela, E. Keihänen, R. Keskitalo, T. S.

Kisner, R. Kneissl, J. Knoche, L. Knox, M. Kunz, H. Kurki-Suonio, G. Lagache, A. Lähteenmäki, J.-M. Lamarre, A. Lasenby, M. Lattanzi, C. R. Lawrence, J. P. Leahy, R. Leonardi, J. Lesgourgues, F. Levrier, A. Lewis, M. Liguori, P. B. Lilje, M. Linden-Vørnle, M. López-Caniego, P. M. Lubin, J. F. Macías-Pérez, G. Maggio, D. Maino, N. Mandolesi, A. Mangilli, A. Marchini, M. Maris, P. G. Martin, M. Martinelli, E. Martínez-González, S. Masi, S. Matarrese, P. McGehee, P. R. Meinhold, A. Melchiorri, J.-B. Melin, L. Mendes, A. Mennella, M. Migliaccio, M. Millea, S. Mitra, M.-A. Miville-Deschênes, A. Moneti, L. Montier, G. Morgante, D. Mortlock, A. Moss, D. Munshi, J. A. Murphy, P. Naselsky, F. Nati, P. Natoli, C. B. Netterfield, H. U. Nørgaard-Nielsen, F. Noviello, D. Novikov, I. Novikov, C. A. Oxborrow, F. Paci, L. Pagano, F. Pajot, R. Paladini, D. Paoletti, B. Partridge, F. Pasian, G. Patanchon, T. J. Pearson, O. Perdereau, L. Perotto, F. Perrotta, V. Pettorino, F. Piacentini, M. Piat, E. Pierpaoli, D. Pietrobon, S. Plaszczynski, E. Pointecouteau, G. Polenta, L. Popa, G. W. Pratt, G. Prézeau, S. Prunet, J.-L. Puget, J. P. Rachen, W. T. Reach, R. Rebolo, M. Reinecke, M. Remazeilles, C. Renault, A. Renzi, I. Ristorcelli, G. Rocha, C. Rosset, M. Rossetti, G. Roudier, B. Rouillé d'Orfeuil, M. Rowan-Robinson, J. A. Rubiño-Martín, B. Rusholme, N. Said, V. Salvatelli, L. Salvati, M. Sandri, D. Santos, M. Savelainen, G. Savini, D. Scott, M. D. Seiffert, P. Serra, E. P. S. Shellard, L. D. Spencer, M. Spinelli, V. Stolyarov, R. Stompor, R. Sudiwala, R. Sunyaev, D. Sutton, A.-S. Suur-Uski, J.-F. Sygnet, J. A. Tauber, L. Terenzi, L. Toffolatti, M. Tomasi, M. Tristram, T. Trombetti, M. Tucci, J. Tuovinen, M. Türler, G. Umana, L. Valenziano, J. Valiviita, F. Van Tent, P. Vielva, F. Villa, L. A. Wade, B. D. Wandelt, I. K. Wehus, M. White, S. D. M. White, A. Wilkinson, D. Yvon, A. Zacchei, and A. Zonca. Planck 2015 results - XIII. cosmological parameters. 594:A13. Publisher: EDP Sciences.

[51] Volker Bromm and Richard B. Larson. The first stars. 42:79–118.

[52] F. Bertoldi, C. L. Carilli, P. Cox, X. Fan, M. A. Strauss, A. Beelen, A. Omont, and R. Zylka. Dust emission from the most distant quasars. 406(3):L55–L58.

[53] Takaya Nozawa, Takashi Kozasa, Hideyuki Umeda, Keiichi Maeda, and Ken'ichi Nomoto. Dust in the early universe: Dust formation in the ejecta of population III supernovae. 598(2):785. Publisher: IOP Publishing.

[54] David Merritt and Laura Ferrarese. Relationship of black holes to bulges. Issue: arXiv:astro-ph/0107134.

[55] H. W. W. Spoon, D. Farrah, V. Lebouteiller, E. González-Alfonso, J. Bernard-Salas, T. Urrutia, D. Rigopoulou, M. S. Westmoquette, H. A. Smith, J. Afonso, C. Pearson, D. Cormier, A. Efstathiou, C. Borys, A. Verma, M. Etxaluze, and D. L. Clements. DIAGNOSTICS OF AGN-DRIVEN MOLECULAR OUTFLOWS IN ULIRGs FROM HERSCHEL-PACS OBSERVATIONS OF OH AT 119 um. 775(2):127.

[56] Volker Beckmann and Chris R Shrader. The AGN phenomenon: open issues.

[57] Albert A. Michelson and Edward W. Morley. LVIII. on the relative motion of the earth and the luminiferous Æther. 24(151):449–463.

[58] Zehnder Ludwig. Ein neuer interferenzrefraktor. 11:275–285.

[59] Albert A. Michelson. I. on the application of interference methods to astronomical measurements. 30(182):1–21.

[60] Albert Abraham Michelson. On the application of interference methods to astronomical measurements. 51:257.

[61] P. H van Cittert. Die wahrscheinliche schwingungsverteilung in einer von einer lichtquelle direkt oder mittels einer linse beleuchteten ebene. 1(1):201–210.

[62] F. Zernike. The concept of degree of coherence and its application to optical problems. 5(8):785–795.

[63] H. H. Hopkins and George Paget Thomson. The concept of partial coherence in optics. 208(1093):263–277.

[64] E Wolf and Max Born. A macroscopic theory of interference and diffraction of light from finite sources II. fields with a spectral range of arbitrary width. 230(1181):246–265.

[65] L. L. McCready, Joseph Lade Pawsey, Ruby Payne-Scott, and Albert Cherbury David Rivett. Solar radiation at radio frequencies and its relation to sunspots. 190(1022):357–375.

[66] M. Ryle and D. D. Vonberg. An investigation of radio-frequency radiation from the sun. 193(1032):98–120.

[67] M. A. Johnson, A. L. Betz, and C. H. Townes. 10-um heterodyne stellar interferometer. 33(27):1617–1620.

[68] JE Baldwin, MG Beckett, RC Boysen, D Burns, DF Buscher, GC Cox, CA Haniff, CD Mackay, NS Nightingale, J Rogers, and others. The first images from an optical aperture synthesis array: mapping of capella with COAST at two epochs. 306:L13.

[69] Kazuyoshi Itoh and Yoshihiro Ohtsuka. Fourier-transform spectral imaging: retrieval of source information from three-dimensional spatial coherence. 3(1):94–100.

[70] J. M. Mariotti and S. T. Ridgway. Double fourier spatio-spectral interferometry - combining high spectral and high spatial resolution in the near infrared. 195:350–363.

[71] David T. Leisawitz, Brad J. Frey, Douglas B. Leviton, Anthony J. Martino, William L. Maynard, Lee G. Mundy, Stephen A. Rinehart, Stacy H. Teng, and Xiaolei Zhang. Wide-field imaging interferometry testbed i: purpose, testbed design, data, and synthesis algorithms. In *Interferometry in Space*, volume 4852, pages 255–267. SPIE.

[72] Stephen A. Rinehart, Brad J. Frey, David T. Leisawitz, Douglas B. Leviton, Anthony J. Martino, William L. Maynard, Lee G. Mundy, Stacy H. Teng, and Xiaolei Zhang. Wide-field imaging interferometry testbed II: implementation, performance, and plans. In *Interferometry in Space*, volume 4852, pages 674–684. SPIE.

[73] Douglas B. Leviton, Brad J. Frey, David T. Leisawitz, Anthony J. Martino, William L. Maynard, Lee G. Mundy, Stephen A. Rinehart, Stacy H. Teng, and Xiaolei Zhang. Wide-field imaging interferometry testbed 3: metrology subsystem. In *Interferometry in Space*, volume 4852, pages 827–838. SPIE.

[74] S. Rinehart, D. Leisawitz, B. Frey, R. Lyon, S. Maher, and N. Memarsadeghi. The wide-field imaging interferometry testbed (WIIT): recent progress and results. In *Optical and Infrared Interferometry*, volume 7013, pages 920–929. SPIE.

[75] Izumi S. Ohta, Makoto Hattori, and Hiroshi Matsuo. Development of a multi-fourier-transform interferometer: fundamentals. 45(12):2576–2585.

[76] Izumi S. Ohta, Makoto Hattori, and Hiroshi Matsuo. Development of a multi-fourier-transform interferometer: imaging experiments in millimeter and submillimeter wave bands. 46(15):2881–2892.

[77] Izumi S. Ohta, Makoto Hattori, Yuji Chinone, Yuan Luo, Yoshihiro Hamaji, Junichi Takahashi, Hiroshi Matsuo, and Nario Kuno. Astronomical testing observation in multi-fourier transform interferometer: Aperture synthesis technique and CMB. In *2007 Joint 32nd International Conference on Infrared and Millimeter Waves and the 15th International Conference on Terahertz Electronics*, pages 335–336.

[78] William F. Grainger, Roser Juanola-Parramon, Peter A. R. Ade, Matt Griffin, Flo Liggins, Enzo Pascale, Giorgio Savini, and Bruce Swinyard. Demonstration of spectral and spatial interferometry at THz frequencies. 51(12):2202–2211.

[79] Eri Kato, Hiroshi Shibai, Mitsunobu Kawada, Masanao Narita, Taro Matsuo, Atsushi Ohkubo, Miki Suzuki, Tetsuo Kanoh, Koudai Yamamoto, and Fite Team. Far-infrared interferometric telescope experiment : I. interferometer optics. 7:47–53.

[80] Ayana Sasaki, Hiroshi Shibai, Misato Fukagawa, Takahiro Sumi, Tesuo Kanoh, Kodai Yamamoto, Yusuke Itoh, Naoki Akiyama, Atsushi Terano, Yukako Aimi, Yoshihiro Kuwada, Mihoko Konishi, and Masanao Narita. Far-infrared interferometric telescope experiment: Optical adjustment system. 4(2):179–183.

[81] S. A. Rinehart, M. Rizzo, D. J. Benford, D. J. Fixsen, T. J. Veach, A. Dhabal, D. T. Leisawitz, L. G. Mundy, R. F. Silverberg, R. K. Barry, J. G. Staguhn, R. Barclay, J. E. Mentzell, M. Griffin, P. A. R. Ade, E. Pascale, G. Klemencic, G. Savini, and R. Juanola-Parramon. The balloon experimental twin telescope for infrared interferometry (BETTII): An experiment for high angular resolution in the far-infrared. 126(941):660.

[82] S. A. Rinehart, A. Dhabal, D. Fixsen, R. Juanola-Parramon, D. Leisawitz, S. Maher, J. E. Mentzell, L. Mundy, M. Rizzo, H. Sampler, E. Sharp, R. Silverberg, M. Casalprim Torres, T. Veach, J. Vila Hernández de Lorenzo, P. a. R. Ade, C. Tucker, E. Pascale, and G. Savini. The balloon experimental twin telescope for infrared interferometry (BETTII): first flight. In *Ground-based and Airborne Telescopes VII*, volume 10700, pages 121–133. SPIE.

[83] David Leisawitz, John C. Mather, S. Harvey Moseley, and Xiaolei Zhang. The submillimeter probe of the evolution of cosmic structure (SPECS). 269(0):563–567.

[84] David Leisawitz and William Oegerle. Community plan for far-infared/submillimeter space astronomy.

[85] J Clavel. ESA's cosmic vision 2015-2025 program. In *Astronomical Data Analysis Software and Systems XVIII*, volume 411, page 89.

[86] Martin Harwit, George Helou, Lee Armus, C Matt Bradford, Paul F Goldsmith, Michael Hauser, David Leisawitz, Daniel F Lester, George Rieke, and Stephen A Rinehart. Far-infrared/submillimeter astronomy from space: Tracking an evolving universe and the emergence of life. 2008.

[87] M Sauvage, RJ Ivison, and F Helmich. Sub-arcsecond far-infrared space observatory: a science imperative.

[88] Jeremy Scott and Adam Sundberg. A spatial-spectral interferometer testbed for the far-infrared. Canadian Astronomical Society.

[89] J. Scott, L. D. Spencer, and Chris S. Benson. Optical design of a spatial-spectral fourier transform interferometer for the far-infrared. In *OSA Optical Sensors and Sensing Congress 2021 (AIS, FTS, HISE, SENSORS, ES) (2021), paper FM5C.7*, page FM5C.7. Optica Publishing Group.

[90] Jeremy Scott, Chris Benson, and Locke Spencer. Phase calibration of a spatial-spectral fourier transform interferometer testbed. Optica Publishing Group.

[91] Jeremy Scott, Chris Benson, and Locke Spencer. Design, operation, and characterization of a laboratory spatial-spectral fourier transform interferometer. International Symposium on Space Terahertz Technology.

[92] Locke D. Spencer, Rashmi V. Sudiwala, Chris S. Benson, Jeremy P. Scott, Manimeldura Dinula De Silva, Tyrone Jones, James Cox, Ken Wood, and Peter A. R. Ade. Optimization of a cryogenic transition-edge sensor detector array for far-infrared astrophysics. In *Space Telescopes and Instrumentation 2022: Optical, Infrared, and Millimeter Wave*, volume 12180, pages 1042–1059. SPIE.

[93] Jeremy Scott, Chris Benson, and Locke Spencer. Design, operation, and characterization of a laboratory spatial-spectral fourier transform interferometer. 13(6):636–644.

[94] Jean Baptiste Joseph Fourier. *Théorie analytique de la chaleur*. Gauthier-Villars.

[95] K. F. Riley, M. P. Hobson, and S. J. Bence. *Mathematical Methods for Physics and Engineering: A Comprehensive Guide*. Cambridge University Press, 3rd edition.

[96] David W. Kammler. *A First Course in Fourier Analysis*. Cambridge University Press, 2nd edition.

[97] Claude Elwood Shannon. Communication in the presence of noise. 37(1):10–21. Publisher: IEEE.

[98] James W. Cooley and John W. Tukey. An algorithm for the machine calculation of complex fourier series. 19(90):297–301.

[99] David A Naylor, Brad G Gom, Trevor R Fulton, Margaret K Tahic, and Gary R Davis. Increased efficiency through undersampling in fourier transform spectroscopy. In *Fourier Transform Spectroscopy*, page FTuD14. Optica Publishing Group.

[100] Sumner P. Davis, Mark C. Abrams, and James W. Brault. *Fourier Transform Spectrometry*. Academic Press, 1st edition edition.

[101] A. Richard Thompson, James M. Moran, and George W. Swenson. *Interferometry and Synthesis in Radio Astronomy*. Astronomy and Astrophysics Library. Springer International Publishing.

[102] J. A. Högbom. Aperture synthesis with a non-regular distribution of interferometer baselines. 15:417.

[103] T. J. Cornwell and K. F. Evans. A simple maximum entropy deconvolution algorithm. 143:77–83.

[104] Peter A. R. Ade, Giampaolo Pisano, Carole Tucker, and Samuel Weaver. A review of metal mesh filters. In *Millimeter and Submillimeter Detectors and Instrumentation for Astronomy III*, volume 6275, pages 248–262. SPIE.

[105] Locke Dean Spencer and University of Lethbridge Faculty of Arts and Science. Imaging fourier transform spectroscopy from a space based platform : the herschel/SPIRE fourier transform spectrometer.

[106] Anselm J. Deninger, A. Roggenbuck, S. Schindler, and S. Preu. 2.75 THz tuning with a triple-DFB laser system at 1550 nm and InGaAs photomixers. 36(3):269–277.

[107] Orazio Svelto. *Principles of Lasers*. Springer/Sci-Tech/Trade, 5th edition edition.

[108] A.V. Raisanen. Frequency multipliers for millimeter and submillimeter wavelengths. 80(11):1842–1852.

[109] Mira Naftaly and Richard Dudley. Terahertz reflectivities of metal-coated mirrors. 50(19):3201.

[110] John E. Greivenkamp. *Field Guide to Geometrical Optics*. SPIE Publications.

[111] Max Born and Emil Wolf. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, 7 edition.

[112] Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual*. CreateSpace.

[113] Charles R. Harris, K. Jarrod Millman, Stéfan J. van der Walt, Ralf Gommers, Pauli Virtanen, David Cournapeau, Eric Wieser, Julian Taylor, Sebastian Berg, Nathaniel J. Smith, Robert Kern, Matti Picus, Stephan Hoyer, Marten H. van Kerkwijk, Matthew Brett, Allan Haldane, Jaime Fernández del Río, Mark Wiebe, Pearu Peterson, Pierre Gérard-Marchant, Kevin Sheppard, Tyler Reddy, Warren Weckesser, Hameer Abbasi, Christoph Gohlke, and Travis E. Oliphant. Array programming with NumPy. 585(7825):357–362.

[114] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental algorithms for scientific computing in python. 17:261–272.

[115] Thomas P. Robitaille, Erik J. Tollerud, Perry Greenfield, Michael Droettboom, Erik Bray, Tom Aldcroft, Matt Davis, Adam Ginsburg, Adrian M. Price-Whelan, Wolfgang E. Kerzendorf, Alexander Conley, Neil Crighton, Kyle Barbary, Demitri Muna, Henry Ferguson, Frédéric Grollier, Madhura M. Parikh, Prasanth H. Nair, Hans M. Günther, Christoph Deil, Julien Woillez, Simon Conseil, Roban Kramer, James E. H. Turner, Leo Singer, Ryan Fox, Benjamin A. Weaver, Victor Zabalza, Zachary I. Edwards, K. Azalee Bostroem, D. J. Burke, Andrew R. Casey, Steven M. Crawford, Nadia Dencheva, Justin Ely, Tim Jenness, Kathleen Labrie, Pey Lian Lim, Francesco Pierfederici, Andrew Pontzen, Andy Ptak, Brian Refsdal, Mathieu Servillat, and Ole Streicher. Astropy: A community python package for astronomy. 558:A33.

[116] J. D. Hunter. Matplotlib: A 2d graphics environment. 9(3):90–95.

[117] Paul F. Goldsmith. Radio telescopes and measurements at radio wavelengths. 278:45–79.

[118] P. D. Mauskopf. Transition edge sensors and kinetic inductance detectors in astronomical instruments. 130(990):082001.

[119] Charles Kittel and Herbert Kroemer. *Thermal Physics*. W. H. Freeman, 2nd edition.

[120] J. M. Lamarre. Photon noise in photometric instruments at far-infrared and submillimeter wavelengths. 25(6):870–876. Publisher: Optica Publishing Group.

[121] L. D. Landau and E. M. Lifshitz. *Statistical Physics: Volume 5*. Butterworth-Heinemann, 3rd edition.

[122] J. N. Ullom, W. B. Doriese, G. C. Hilton, J. A. Beall, S. Deiker, W. D. Duncan, L. Ferreira, K. D. Irwin, C. D. Reintsema, and L. R. Vale. Characterization and reduction of unexplained noise in superconducting transition-edge sensors. 84(21):4206–4208.

[123] Edward F. Kuester. *Theory of Waveguides and Transmission Lines*. CRC Press, 1st edition edition.

[124] Larry D. Paarmann. *Design and Analysis of Analog Filters: A Signal Processing Perspective*. Springer/Sci-Tech/Trade, softcover reprint of the original 1st ed. 2003 edition edition.

[125] Scott Paine. The am atmospheric model. Version Number: 13.0.

[126] Greg B Taylor, Chris Luke Carilli, and Richard A Perley. Synthesis imaging in radio astronomy II. 180.

[127] S. F. Gull and J. Skilling. Maximum entropy method in image processing. 131(6):646–659. Publisher: IET Digital Library.

[128] B. G. Clark. An efficient implementation of the algorithm 'CLEAN'. 89:377. ADS Bibcode: 1980A&A....89..377C.

[129] Principles of long baseline stellar interferometry: Course notes from the 1999 michelson summer school august 15-19, 1999.

[130] Rizzo Maxime. BETTII: A PATHFINDER FOR HIGH ANGULAR RESOLUTION OBSERVATIONS OF STAR-FORMING REGIONS IN THE FAR-INFRARED.

[131] Nicholas M. Elias Ii, Martin Harwit, David Leisawitz, and Stephen A. Rinehart. The mathematics of double-fourier interferometers. 657(2):1178–1200.

[132] B. I. Bleaney and B. Bleaney. *Electricity and Magnetism, Volume 1: Third edition*. Oxford Classic Texts in the Physical Sciences. Oxford University Press.

Appendix A

van Cittert-Zernike theorem

The theoretical foundation for combining interference measurements to synthesize images is provided by the van Cittert-Zernike theorem. This equation relates the source intensity distribution to a quantity called the Mutual Coherence Function (MCF). The MCF describes the coherence of electric fields at two points in space under the condition that these points are a great distance from the sources that produce the electric fields. An appropriate interferometer instrument is capable of measuring the MCF, and with enough samples, the source intensity distribution can be reconstructed.

Although there are many notable contributions to the principles of interferometry and the coherence of light which provide the framework for the van Cittert-Zernike theorem, the first formal description of the theorem was provided by Pieter Hendrik van Cittert [61], and separately Frits Zernike [62]. van Cittert's approach was based explicitly on statistical calculations related to a perfect optical system with a rectangular aperture. Zernike generalized these results by considering the Young's double slit experiment and how sources of finite extent affect fringe visibility. Further developments were made by Harold Horace Hopkins [63] who extended the theory to more general optical systems and showed how the complex visibility relates to image formation using a fairly simple formalism. The most general result was obtained by Emil Wolf who extended the theorem to sources of arbitrary spectral width [64]. Slight variations in terminology are used by the authors mentioned above, and in this work the terminology used by Wolf is employed. It is interesting to note that in the absence of a generalized theory of mutual coherence, Albert A. Michelson was separately applying the principles spatial and temporal coherence almost half a century earlier with his famous stellar spatial interferometer experiment [59] and the Michelson-Morley spectral interferometer [57].

In this appendix, the MCF is derived through relation to a general source intensity distribution following a framework similar to [64]. It is also shown how the MCF can be measured with an appropriate interferometer instrument. An intuitive description of what the MCF represents and how it manifests within interferometer measurements are also discussed.

A.1 Deriving the Mutual Coherence Function

To derive the relationship between the MCF and a source intensity distribution, the geometry shown in Fig. A.1 will be used. A planar source Σ is embedded in the source plane and described using coordinates $\vec{s}$. The aperture plane is parallel to and separated

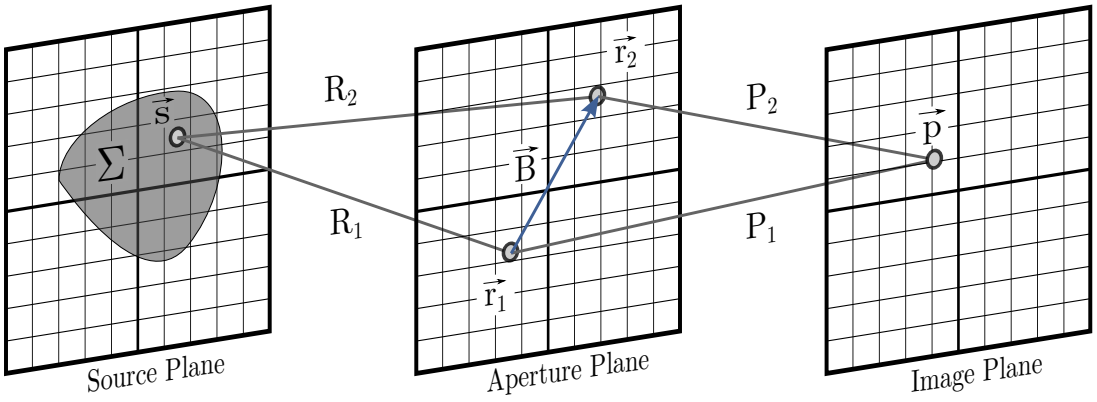


Figure A.1: Coordinate system used in deriving the van Cittert-Zernike theorem. Although not well represented by the figure, the parallel planes are assumed to be separated by large distances such that pairs of light rays passing through the aperture plane are approximately parallel and the angles of the rays with respect to the plane normal vectors are small. Positions in the source, aperture, and image plane are represented with Cartesian coordinate vectors $\vec{s}$, $\vec{r}$, and $\vec{p}$, respectively. The baseline vector $\vec{B}$ is the displacement vector from $\vec{r}_1$ to $\vec{r}_2$. Quantities R_1, R_2, P_1, and P_2 are scalar distance between the indicated points from one plane to the next.

from the source plane by a distance large enough that spherical waves emitted from points in the source plane can be approximated as plane waves when incident on the aperture plane. Positions in the aperture plane are described by coordinates $\vec{r}$, and the separation vector between points where the electric field is considered is called the baseline $\vec{B} = \vec{r}_2 - \vec{r}_1$. Similarly, the image plane is parallel to and separated from the aperture plane by a large distance, with positions described using coordinates $\vec{p}$. Coordinate systems listed above are all Cartesian with length units. Scalar quantities R_1, R_2, P_1, and P_2 represent the distances between considered points.

The generally complex MCF is given by,

$$\begin{aligned}\Gamma(\vec{r}_1,\vec{r}_2,\tau) &= \langle E(\vec{r}_1,t+\tau)E^*(\vec{r}_2,t)\rangle \\ &= \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} E(\vec{r}_1,t+\tau)E^*(\vec{r}_2,t)\,dt\ ,\end{aligned} \tag{A.1}$$

and represents the time averaged correlation of the generally complex random variable E, considered at two points in space $\vec{r}_1$ and $\vec{r}_2$, with $E(\vec{r}_1)$ being considered at a time τ later than $E(\vec{r}_2)$. In the context of electromagnetic fields, E can be regarded as one orthogonal component of the electric field for linearly polarized light. In effect, the MCF describes the temporal and spatial coherence of light. In the special case where $\vec{r}_1 = \vec{r}_2 = \vec{r}$ and $\tau = 0$, Eq. A.1 is proportional to the intensity at $\vec{r}$,

$$\Gamma(\vec{r},\vec{r},0) = |E(\vec{r})|^2 = \frac{I(\vec{r})}{Z}\ . \tag{A.2}$$

The proportionality constant Z, is the impedance of the propagation medium $\sqrt{\mu/\varepsilon}$, where μ and ε are the permeability and permittivity of the propagation medium, respectively. In the context of this appendix, the proportionality constant will be suppressed for simplicity, since the main results of interest obtained in this appendix are normalized quantities. The normalized MCF is obtained as,

$$\gamma(\vec{r}_1,\vec{r}_2,\tau) = \frac{\Gamma(\vec{r}_1,\vec{r}_2,\tau)}{\Gamma(\vec{r}_1,\vec{r}_1,0)\Gamma(\vec{r}_2,\vec{r}_2,0)} = \frac{\Gamma(\vec{r}_1,\vec{r}_2,\tau)}{\sqrt{I(\vec{r}_1)I(\vec{r}_2)}}\,, \tag{A.3}$$

with $0 \leq |\gamma(\vec{r}_1,\vec{r}_2,\tau)| \leq 1$ by way of the Cauchy–Schwarz inequality [95]. For reasons that will be evident in Sec. A.2, $\gamma(\vec{r}_1,\vec{r}_2,\tau)$ is called the complex visibility.

At a particular point in the aperture plane $\vec{r}$, the time-dependent electric field can be expanded as a Fourier series,

$$E(\vec{r},t) = \int_{-\infty}^{\infty} \hat{E}(\vec{r},\nu) e^{i2\pi\nu t}\, d\nu\,, \tag{A.4}$$

where $\hat{E}(\vec{r},\nu)$ is the spectral component of the field at frequency ν. Although negative frequencies are not physical, their inclusion allows for a more direct use of Fourier transforms and does not affect the derivation if it is understood that for real $E(\vec{r},t)$, the spectral components will obey the condition $\hat{E}^*(\vec{r},\nu) = \hat{E}(\vec{r},-\nu)$. Inserting Eq. A.4 into Eq. A.1 and using the correlation theorem (Sec. C.1), the MCF can be written in terms of spectral components,

$$\Gamma(\vec{r}_1,\vec{r}_2,\tau) = \lim_{T\to\infty} \frac{1}{2T} \int_{-\infty}^{\infty} \hat{E}(\vec{r}_1,\nu)\hat{E}^*(\vec{r}_2,\nu) e^{i2\pi\nu\tau}\, d\nu\,. \tag{A.5}$$

The source is modelled as a collection of statistically independent point sources with differential area $d\Sigma = ds_x\, ds_y$. Computing the total electric field in the aperture plane is obtained as the superposition of all point source contributions

$$E(\vec{r},t) = \int_{\Sigma} E(\vec{r},t;\vec{s})\, d\Sigma\,, \tag{A.6}$$

where $E(\vec{r},t;\vec{s})$ is the electric field at $\vec{r}$ due to a point source on the source plane at $\vec{s}$. Substituting Eq. A.6 into Eq. A.1 and using the same correlation theorem that was used to obtain Eq. A.5, the MCF is modified to,

$$\Gamma(\vec{r}_1,\vec{r}_2,\tau) = \int_{\Sigma}\int_{\Sigma} \Gamma(\vec{r}_1,\vec{r}_2,\tau;\vec{s}_1,\vec{s}_2)\, d\Sigma_1\, d\Sigma_2\,, \tag{A.7}$$

with

$$\Gamma(\vec{r}_1,\vec{r}_2,\tau;\vec{s}_1,\vec{s}_2) = \lim_{T\to\infty} \frac{1}{2T} \int_{-\infty}^{\infty} \hat{E}(\vec{r}_1,\nu;\vec{s}_1)\hat{E}^*(\vec{r}_2,\nu:\vec{s}_2) e^{i2\pi\nu\tau}\, d\nu\,, \tag{A.8}$$

and the source integrals explore the source plane independently. Since it is assumed that the source is composed of statistically independent point sources, $\Gamma(\vec{r}_1,\vec{r}_2,\tau;\vec{s}_1,\vec{s}_2) = 0$ when $\vec{s}_1 \neq \vec{s}_2$. Consequently, the first integral in Eq. A.7 selects for the condition $\vec{s}_1 = \vec{s}_2$.

Light emitted from source points propagate as spherical waves. If the medium between the source and aperture plane is homogeneous, the electric field at the aperture plane is

given by

$$\hat{E}(\vec{r},\nu;\vec{s}) = \frac{\hat{a}(\vec{s},\nu)}{R} e^{i2\pi\nu R/c}\,, \tag{A.9}$$

where $\hat{a}(\vec{s},\nu)$ is the source electric field amplitude, c is the speed of light, and $R = |\vec{r}-\vec{s}|$ is the distance between the source point and the position in the aperture plane. Inserting Eq. A.9 into Eq. A.7 and noting the selection criteria mentioned above, the MCF becomes

$$\Gamma(\vec{r}_1,\vec{r}_2,\tau) = \int_\Sigma \int_{-\infty}^{\infty} \frac{I(\vec{s},\nu)}{R_1 R_2} e^{i2\pi\nu\left(\frac{R_1-R_2}{c}+\tau\right)}\, d\nu\, d\Sigma\,, \tag{A.10}$$

where

$$I(\vec{s},\nu) = \lim_{T\to\infty} \frac{1}{2T}\, \hat{a}(\vec{s},\nu)\hat{a}^*(\vec{s},\nu) \tag{A.11}$$

is the source spectral intensity distribution.

A new set of angular coordinates is introduced as $\vec{\theta} = \vec{s}/R$, where R is the distance between the source and aperture planes. Since the distance between planes is large compared to the separation between aperture plane positions, the denominator of Eq. A.10 can be approximated as $R_1 \approx R_2 = R$. The large distance between planes also implies that the angular coordinates are small, admitting the small angle approximation and neglecting obliquity factors associated with the field propagation. Additionally, $\Delta R = R_1 - R_2 \approx \vec{B}\cdot\vec{\theta}$, with $\vec{B} = \vec{r}_2 - \vec{r}_1$, as is illustrated in Fig. A.2 for a one-dimension case. Using the substitutions described, the final form of the MCF is

$$\Gamma(\vec{B},\tau) = \int_\Omega \int_{-\infty}^{\infty} I(\vec{\theta},\nu) e^{-i2\pi\vec{B}\cdot\vec{\theta}/\lambda}\, e^{i2\pi\nu\tau}\, d\nu\, d\Omega\,, \tag{A.12}$$

with complex visibility

$$\gamma(\vec{B},\tau) = \frac{\Gamma(\vec{B},\tau)}{\int_\Omega \int_{-\infty}^{\infty} I(\vec{\theta},\nu)\, d\nu\, d\Omega}\,, \tag{A.13}$$

where $\lambda = c/\nu$ is the wavelength of the electric field, and the integration is performed over the source in terms of solid angle $d\Omega = d\Sigma/R^2 = d\theta_x\, d\theta_y$. Integration can be extended to the entire source plane without loss of generality if contributions to the integral at large angles are negligible. This is typically the case with real detectors due to the rapid drop-off in sensitivity away from the optical axis. Since the baseline vector could have been drawn from $\vec{r}_2$ to $\vec{r}_1$ (Fig. A.1), and the time delay τ could have been applied to $E(\vec{r}_2)$ (Eq. A.1), the sign of the exponents in Eq A.12 are arbitrary and have been chosen to be consistent with Ch. 2.

Eq. A.12 is the van Cittert-Zernike theorem generalized to sources with arbitrary spectral width. The key insight is that the MCF is related to the source spectral intensity distribution through a two-dimensional Fourier transform in space and an inverse Fourier transform in frequency. The restricted van Cittert-Zernike theorem is obtain by setting $\tau = 0$ and considering only one spectral frequency $\nu = \nu_0$,

$$\Gamma_0(\vec{B}) = \int_\Omega I(\vec{\theta},\nu_0) e^{-i2\pi\vec{B}\cdot\vec{\theta}/\lambda_0}\, d\Omega\,, \tag{A.14}$$

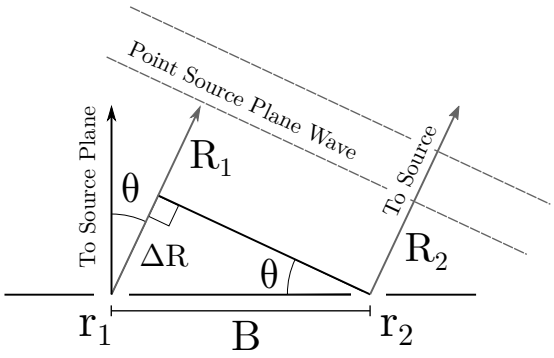


Figure A.2: One dimensional aperture geometry with respect to the source plane. The optical path difference $\Delta R = R_1 - R_2$ is related to the source angular position $\theta = s/R$ and baseline length $B = r_2 - r_1$ by $\Delta R = B\sin(\theta) \approx B\theta$ for small angles. Note that θ is in the same direction as B and this relation is equivalent to $\Delta R = \vec{B}\cdot\vec{\theta}$. In the general case where $\vec{\theta}$ has a component perpendicular to $\vec{B}$, if the incoming wave if a plane wave, only the projection of $\vec{\theta}$ on $\vec{B}$ contributes to the optical path difference.

with complex visibility,

$$\gamma_0(\vec{B}) = \frac{\Gamma_0(\vec{B})}{\int_\Omega I(\vec{\theta},\nu_0)\,d\Omega}\,. \tag{A.15}$$

Conceptually, one can imaging building up $\Gamma(\vec{B},\tau)$ as a series of monochromatic $\Gamma_0(\vec{B})$ slices. It is interesting to note that $\Gamma_0(\vec{B})$ is equivalent to the Fraunhofer diffraction pattern associated with an aperture having the same shape and illumination as the intensity distribution $I(\vec{\theta},\nu_0)$. This fact is a useful tool when considering the form of the MCF given a particular source spectral intensity distribution.

Often the MCF is written using a wavelength normalized baseline basis set,

$$\begin{aligned} u &= \frac{B_x}{\lambda} \\ v &= \frac{B_y}{\lambda}\,, \end{aligned} \tag{A.16}$$

and sampling the MCF is referred to as "filling the uv plane." When filling the uv plane, the absolute position of the baseline does not matter, only the magnitude and orientation. This is because the statistical properties of the incident plane waves are invariant to translation, as is evident from Fig. A.2. Additionally, the MCF possesses Hermitian symmetry [95],

$$\Gamma(-\vec{B},-\tau) = \Gamma^*(\vec{B},\tau)\,. \tag{A.17}$$

In particular for the restricted case, $\Gamma_0(-\vec{B}) = \Gamma_0^*(\vec{B})$, this suggests only half of the uv plane need be sampled for image reconstruction. Similarly, $\Gamma(0,-\tau) = \Gamma^*(0,\tau)$, and the spectral content of the source can be obtained using only either negative or positive time delays.

A.2 Measuring the Mutual Coherence Function

To understand how the MCF can be measured through experiments, the propagation of the MCF through space must first be considered. The electric field at a point $\vec{p}$ on the image plane due to the electric field distribution on the aperture plane $\hat{E}(\vec{r},\nu)$ (see Fig. A.1), is given by

$$\hat{E}(\vec{p},\nu) = \int_{\mathcal{A}} \Lambda \hat{E}(\vec{r},\nu) e^{i2\pi\nu P/c}\, d\mathcal{A}\ , \tag{A.18}$$

where $\mathcal{A}$ denotes the aperture plane and Λ is a general function that describes the propagation of the electric field from $\vec{r}$ to $\vec{p}$. For a spherical wave $\Lambda = 1/P$, where P is the distance between $\vec{r}$ and $\vec{p}$. If the electric field is efficiently coupled to a waveguide, the electric field amplitude is roughly constant with propagation distance and $\Lambda \approx 1$. Using Eq. A.18 in connection with Eq. A.7, the MCF at the image plane is,

$$\Gamma(\vec{p}_1,\vec{p}_2,\tau) = \int_{\mathcal{A}}\int_{\mathcal{A}} \left[\lim_{T\to\infty} \frac{1}{2T} \int_{-\infty}^{\infty} \hat{E}(\vec{r}_1,\nu)\hat{E}^*(\vec{r}_2,\nu) e^{i2\pi\nu\left(\frac{P_1-P_2}{c}+\tau\right)}\, d\nu \right] \Lambda_1\Lambda_2^*\, d\mathcal{A}_1\, d\mathcal{A}_2\ . \tag{A.19}$$

The expression in square parenthesis can be identified as the MCF between the electric fields at two points in the aperture plane Eq. A.5, and Eq. A.19 can be written more compactly as

$$\Gamma(\vec{p}_1,\vec{p}_2,\tau) = \int_{\mathcal{A}}\int_{\mathcal{A}} \Gamma\left(\vec{r}_1,\vec{r}_2,\left(\frac{P_1-P_2}{c}+\tau\right)\right) \Lambda_1\Lambda_2^*\, d\mathcal{A}_1\, d\mathcal{A}_2\ , \tag{A.20}$$

which relates the MCF in the image plane to the MCF in the aperture plane.

If the two points in the image plane are made to overlap $\vec{p}_1 = \vec{p}_2 = \vec{p}$ and the time delay is set to $\tau = 0$, Eq. A.20 in connection with Eq. A.2 gives the intensity at $\vec{p}$,

$$I(\vec{p}) = \int_{\mathcal{A}}\int_{\mathcal{A}} \Gamma\left(\vec{r}_1,\vec{r}_2,\frac{P_1-P_2}{c}\right) \Lambda_1\Lambda_2^*\, d\mathcal{A}_1\, d\mathcal{A}_2\ . \tag{A.21}$$

Note that the situation just described is the functional equivalent of an intensity interferometer where the electric fields from points $\vec{r}_1$ and $\vec{r}_2$ in the aperture plane are combined and the resulting intensity is obtained.

Consider the situation where a screen is placed across the aperture plane with two small opening at $\vec{r}_1$ and $\vec{r}_2$ (Fig. A.1). Taking the integral of Eq. A.21 on the side facing the image plane gives,

$$\begin{aligned} I(\vec{p}) = {} & I(\vec{r}_1)\gamma(\vec{r}_1,\vec{r}_1,0)|\Lambda_1|^2\, d\mathcal{A}_1^2 + I(\vec{r}_2)\gamma(\vec{r}_2,\vec{r}_2,0)|\Lambda_2|^2\, d\mathcal{A}_2^2 \\ & + \sqrt{I(\vec{r}_1)I(\vec{r}_2)}\gamma(\vec{r}_1,\vec{r}_2,z/c)\Lambda_1\Lambda_2^*\, d\mathcal{A}_1\, d\mathcal{A}_2 \\ & + \sqrt{I(\vec{r}_2)I(\vec{r}_1)}\gamma(\vec{r}_2,\vec{r}_1,-z/c)\Lambda_2\Lambda_1^*\, d\mathcal{A}_2\, d\mathcal{A}_1\ , \end{aligned} \tag{A.22}$$

where the normalization in Eq. A.3 is used, and $z = P_1 - P_2$ is the optical path difference between the two openings on the aperture plane and the point at which the fields from these openings are combined on the image plane.

From Eq. A.2 and Eq. A.3, it is clear $\gamma(\vec{r},\vec{r},0)=1$. Additionally,

$$I(\vec{p};\vec{r}) = I(\vec{r})|\Lambda|^2\, d\mathcal{A}^2 \tag{A.23}$$

is the intensity at a point in the image plane due to the intensity at a point in the aperture plane. From the property of Hermitian symmetry of the MCF, $\gamma(\vec{r}_1,\vec{r}_2,z/c)=\gamma(\vec{r}_2,\vec{r}_1,-z/c)$, and Eq. A.22 can be written as,

$$I(\vec{p}) = I(\vec{p};\vec{r}_1)+I(\vec{p};\vec{r}_2)+2\sqrt{I(\vec{p};\vec{r}_1)I(\vec{p};\vec{r}_2)}\gamma(\vec{r}_1,\vec{r}_2,z/c)\;. \tag{A.24}$$

To summarize the discussion so far, the electric fields from two small regions in the aperture plane have been propagated to and combined at a point in the aperture plane, potentially with a non-zero optical path difference, with Eq. A.24 describing the resulting intensity. Propagating and combining electric fields from separate regions of space and measuring the resulting intensity is precisely the function of a spatial interferometer with direct detection. The measured intensity of a spatial interferometer is thus proportional to the complex visibility, which when related to the van Cittert-Zernike theorem Eq. A.12, gives information about the source spectral intensity distribution.

To see the connection between Eq. A.24 and the response of a spatial interferometer more clearly, the complex visibility is written using polar notation for a monochromatic source,

$$\gamma(\vec{B},z/c) = \frac{\int_\Omega I(\vec{\theta},\nu_0)e^{-i2\pi\vec{B}\cdot\vec{\theta}/\lambda_0}\,d\Omega}{\int_\Omega I(\vec{\theta},\nu_0)\,d\Omega}e^{i2\pi z/\lambda_0} = |\gamma_0(\vec{B})|e^{i[\phi_0(\vec{B})+2\pi z/\lambda_0]}\;, \tag{A.25}$$

where this result is obtained from Eq. A.13 and Eq. A.15. The amplitude of the complex visibility is $|\gamma_0(\vec{B})|$, and its associated phase is $\phi_0(\vec{B})$. A spatial interferometer can only measure real intensity, thus by using Euler's formula $re^{i\phi}=r[\cos(\phi)+i\sin(\phi)]$ and taking the real component, Eq. A.24 becomes,

$$I(\vec{p},z) = I(\vec{p};\vec{r}_1)+I(\vec{p};\vec{r}_2)+2\sqrt{I(\vec{p};\vec{r}_1)I(\vec{p};\vec{r}_2)}|\gamma_0(\vec{B})|\cos(\phi_0(\vec{B})+2\pi z/\lambda_0)\;. \tag{A.26}$$

Evidently, for a monochromatic source, a spatial interferometer produces a sinusoidal intensity pattern as a function of optical path difference z. The pattern is modulated by the complex visibility amplitude $|\gamma_0(\vec{B})|$, and has a phase offset given by $\phi_0(\vec{B})$. An explicit dependence on z has been added since this is the independent variable over which an intensity pattern is measured using an interferometer. This intensity pattern is usually called an interference pattern or interferogram.

In the typical case where $I(\vec{p};\vec{r}_1)=I(\vec{p};\vec{r}_2)=I_0$, the maximum and minimum intensity of the interference pattern in Eq. A.26 is,

$$\begin{aligned} I_{max} &= 2I_0(1+|\gamma_0(\vec{B})|)\;,\\ I_{min} &= 2I_0(1-|\gamma_0(\vec{B})|)\;. \end{aligned} \tag{A.27}$$

The contrast or visibility of fringes within an interference pattern is computed as

$$V = \frac{I_{max} - I_{min}}{I_{max} + I_{min}} = |\gamma_0(\vec{B})| \,, \tag{A.28}$$

and is a quantitative description of the quality of the interference pattern, or the degree of coherence between the electric fields being combined to generate the interference pattern. With this relation, the reason for naming γ the complex visibility is clear.

Consider the interference pattern produced when observing a monochromatic point source located at $\vec{\theta}_0$. In such a case $|\gamma_0(\vec{B})| = 1$ and, ignoring the DC terms, the interference pattern is

$$I_{point}(\vec{p}, z) = 2I_0 \cos(-2\pi(\vec{B} \cdot \vec{\theta}_0 - z)/\lambda_0) \,, \tag{A.29}$$

where the substitution $\phi_0(\vec{B}) = -2\pi\vec{B} \cdot \vec{\theta}_0/\lambda_0$ has been used and is obtained from Eq. A.15. If the point source is on the optical axis $\vec{\theta}_0 = 0$, the resulting interference pattern is centred on $z = 0$. However, if the point source is offset from the optical axis, the interference patter in centred on $z = \vec{B} \cdot \vec{\theta}_0$. These observations suggest that $\phi_0(\vec{B})$ is related to the position of sources on the source plane. Although this statement is explicitly true for point sources, more complicated extended source distributions, which can be modelled as collections of point sources, can have a more complicated phase dependence.

To obtain the interference pattern for a uniform monochromatic line source centered on the optical axis and of linear extent θ_0, Eq. A.29 is integrated with respect to source coordinates,

$$\begin{aligned} I_{line}(z) &= \frac{2I_{total}}{\theta_0} \int_{-\theta_0/2}^{\theta_0/2} \cos(-2\pi(B\theta - z)/\lambda_0)\, d\theta \\ &= 2I_{total} \mathrm{sinc}(\pi B\theta_0/\lambda_0) \cos(2\pi z/\lambda_0) \\ &= 2I_{total} |\gamma_0(B)| \cos(2\pi z/\lambda_0) \,. \end{aligned} \tag{A.30}$$

The intensity of individual point sources is given by I_{total}/θ_0, where I_{total} is the total integrated intensity of the extended line source. The visibility $|\gamma_0(B)| = \mathrm{sinc}(\pi B\theta_0/\lambda_0)$ can be obtained analytically either by direct integration, or in connection with the van Cittert-Zernike theorem, by taking the Fourier transform of the rectangle function which represents the source intensity distribution under consideration. However, the point to be emphasized here is how the visibility coefficient emerges through the sum of interference patterns produced by independent point sources which compose an extended source. This behaviour is demonstrated in Fig. A.3. Since electric fields of the elements making up the source are statistically independent, their electric fields do not combine in such a way to preferentially produce constructive or destructive interference. A consequence of this is that the net interference pattern is the sum of individual point source interference patterns with unit visibility, which produces fringes of decreasing contrast dependent on the source extent.

For a source of finite spectral width, the interference pattern at a particular baseline orientation for an arbitrary source is obtained by integrating Eq. A.26 over all frequencies,

$$I(\vec{p}, z) = \int_{-\infty}^{\infty} 2I(\nu)|\gamma(\vec{B}, \nu)| \cos(\phi(\vec{B}, \nu) + 2\pi z\nu/c)\, d\nu \,, \tag{A.31}$$

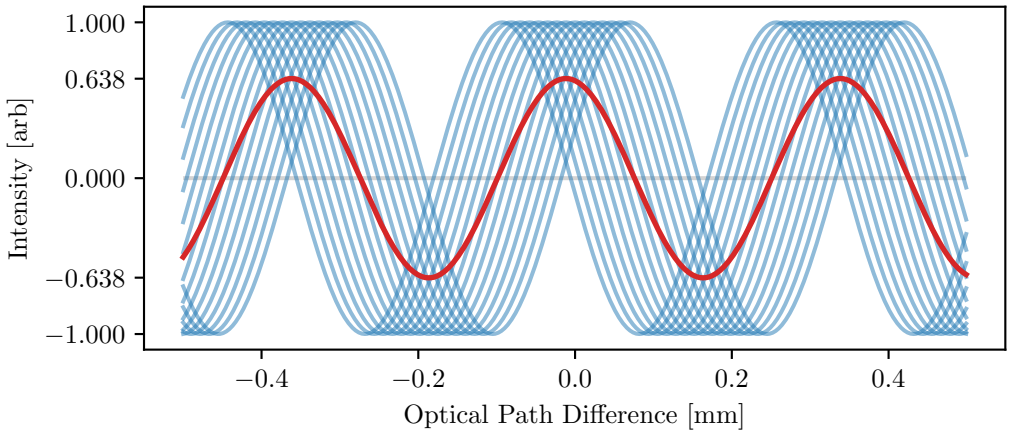


Figure A.3: Mean subtracted interference pattern for an extended line source of width $\theta_0 = 3\,\text{arcmin}$, observed using a baseline $B = 300\,\text{mm}$, at wavelength $\lambda_0 = 353\,\mu\text{m}$. Blue curves show interference patterns with unit visibility associated with the individual point sources that make up the extended source. The red curve shows that average value of the blue curves, and has a peak amplitude $\text{sinc}(\pi B\theta_0/\lambda_0) = 0.638$. Apart from a scaling factor, the red curve represents the interference pattern that would be obtained using a spatial interferometer. The decreased amplitude of the red curve indicates the reduced visibility of the total interference pattern.

where the DC term has been omitted. The frequency dependent complex visibilities can be efficiently extracted from the interference pattern by performing a Fourier transform. Fourier components obtained in this way will generally have real and imaginary contributions with magnitude and phase,

$$\begin{aligned} I(\nu)|\gamma(\vec{B},\nu)| &= \sqrt{S_{Re}^2(\nu) + S_{Im}^2(\nu)} \\ \phi(\vec{B},\nu) &= \arctan\left(\frac{S_{Im}(\nu)}{S_{Re}(\nu)}\right) , \end{aligned} \tag{A.32}$$

respectively, where $S_{Re}(\nu)$ and $S_{Im}(\nu)$ are the real and imaginary components.

To extract $|\gamma(\vec{B},\nu)|$ from the Fourier components, a normalization factor is required. If the sampled points in the aperture plane are made to overlap $\vec{r}_1 = \vec{r}_2 \rightarrow \vec{B} = 0$, the complex visibility $\gamma_0(0) = 1$, as is evident from Eq. A.15. Under these conditions, Eq. A.31 becomes

$$I(\vec{p},z) = \int_{-\infty}^{\infty} 2I(\nu)\cos(2\pi z\nu/c)\,d\nu\,, \tag{A.33}$$

and after taking the Fourier transform, the Fourier components are real with magnitude $S_{Re}(\nu) = I(\nu)$, which is the required normalization factor needed to extract $|\gamma(\vec{B},\nu)|$ from Eq. A.32. Note the modification proposed is in effect to convert the spatial interferometer into a Fourier transform spectrometer which probes exclusively the temporal coherence of the electric fields.

Appendix B

Interferometer Response

In this appendix, the ideal interferometer responses are derived for the Double-Fourier Interferometer (DFI) and Fourier Transform Spectrometer (FTS) observation modes. The results obtained provide first order equations suitable for modelling sources located relatively close to the optical axis. Diffraction effects resulting from finite aperture size are discussed in addition to phase characteristics appropriate for deriving the instrument response.

B.1 Double-Fourier Interferometer Response

A diagram showing the optical layout of the DFI used in this work is shown in Fig. B.1. A detailed discussion of the interferometer design is presented in Ch. 3. In this appendix, only the optical paths are considered for the purpose of deriving an equation for the interferometer response.

Consider a point source, located at the focal point of the off-axis parabolic collimating mirror, that emits an oscillating electrical disturbance of amplitude $E_0(k)$, as a spherical wave. When incident on the collimating mirror, the electric field is given by,

$$\begin{aligned} E_1(k) &= \frac{E_0(k)}{R_1} e^{i(\kappa R_1 - \omega t)} \\ E_2(k) &= \frac{E_0(k)}{R_2} e^{i(\kappa R_2 - \omega t)} \ , \end{aligned} \tag{B.1}$$

where R_i is the distance between the source and the point of intersection on the primary mirror, and t is time. The subscripts identify the different beams of the interferometer, where 1 and 2 represent the upper/red and lower/blue beam, respectively. The field is assumed to be monochromatic with wavenumber $\kappa = 2\pi k = 2\pi/\lambda$ and angular frequency $\omega = 2\pi\nu$, where λ and ν represent the wavelength and frequency of the field, respectively.

After incidence with the collimating mirror, the spherical wave is modified to a plane wave propagating towards the aperture plane, and the electric field amplitude maintains a constant value as it propagates through the remainder of the system. When the electric field interacts with an optical component, the field is modified in phase and amplitude. At the

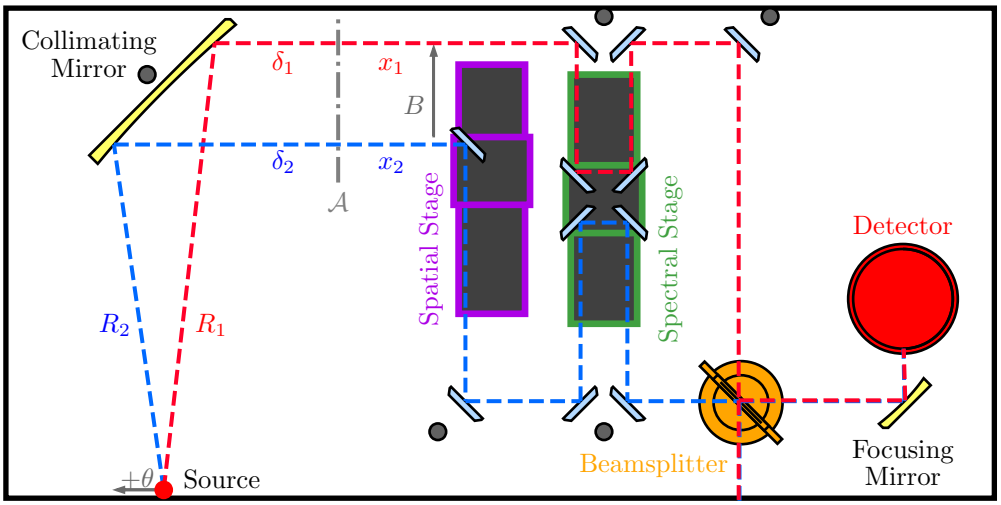


Figure B.1: Diagram of the interferometer in DFI mode.

aperture plane $\mathcal{A}$, the electric fields are,

$$\begin{aligned} E_1(k) &= r_M \frac{E_0(k)}{R_1} e^{i[\kappa(R_1+\delta_1)-\omega t+\phi_M]} \\ E_2(k) &= r_M \frac{E_0(k)}{R_2} e^{i[\kappa(R_2+\delta_2)-\omega t+\phi_M]} \ , \end{aligned} \tag{B.2}$$

where $r_M \approx 0.97$ is the mirror reflection amplitude coefficient [109], and ϕ_M is the phase shift on mirror reflection. Due to boundary continuity conditions, $\phi_M = \pi$. The electric field phase is further modified by δ_i, which describes the distance between the point of reflection on the collimating mirror and the aperture plane. From the directrix based definition of parabolas, $R_i + \delta_i = constant$, indicating that all light from the source on the focal point incident on the aperture plane has the same phase. The position of the aperture plane is arbitrary, and for convenience, the position is chosen such that $\kappa(R_i + \delta_i) + \phi_M$ is an integer multiple of 2π. With this simplification, Eq. B.2 is,

$$\begin{aligned} E_1(k) &= r_M \frac{E_0(k)}{R_1} e^{-i\omega t} \\ E_2(k) &= r_M \frac{E_0(k)}{R_2} e^{-i\omega t} \ , \end{aligned} \tag{B.3}$$

and these equations are interpreted as the fields entering the DFI.

The discussion so far has assumed a point source is on the optical axis. For a point source displaced from the optical axis by an angle $\vec{\theta}$, the distance each beam must travel before intersecting the aperture plane changes. If $|\vec{\theta} f_{col}| \ll R_i$, where f_{col} is the focal length of the collimating mirror, the resulting change in field amplitudes is negligible, but results in a relative path length difference of $\vec{B} \cdot \vec{\theta}$ between the two beams (see Fig. A.2). With these

considerations Eq. B.3 can be approximated as,

$$\begin{aligned} E_1(\vec{\theta},\vec{B},k) &\approx r_M \frac{E_0(\vec{\theta},k)}{R_1} e^{-i\omega t} \\ E_2(\vec{\theta},\vec{B},k) &\approx r_M \frac{E_0(\vec{\theta},k)}{R_2} e^{-i[\omega t+\kappa\vec{B}\cdot\vec{\theta}]} \ , \end{aligned} \tag{B.4}$$

Any absolute phase offset can be removed by adjusting the position of the aperture plane.

Following the optical paths for each beam, the electric fields at the detector plane are,

$$\begin{aligned} E_1^{det}(\vec{\theta},\vec{B},k) &= r_M^6 r_{bs} E_1(\vec{\theta},\vec{B},k) e^{i[\kappa x_1-\omega\tau+6\phi_M+\phi_{bs}]} \\ E_2^{det}(\vec{\theta},\vec{B},k) &= r_M^7 t_{bs} E_2(\vec{\theta},\vec{B},k) e^{i[\kappa x_2-\omega\tau+7\phi_M]} \ , \end{aligned} \tag{B.5}$$

where τ is the time in advance of t. All mirrors in the interferometer are manufactured from the same aluminum alloy and are assumed to possess the same reflection amplitude coefficient. Reflection and transmission amplitude coefficients of the beamsplitter have been measured, and assuming lossless interaction, have values $r_{bs} \approx t_{bs} \approx 1/\sqrt{2}$. Phase shifts associated with reflection and transmission of the beamsplitter have not been measured directly. However, relative measurements and energy conservation constraints discussed below support the conclusion that the relative phase shift between reflection and transmission is $\phi_{bs} = \pm\pi/2+\varepsilon$, where the sign depends on which side of the beamsplitter is used, and ε is a frequency dependent perturbation that is independent of what side of the beamsplitter is used. There is a phase shift associated with electric fields as they propagate along the optical paths x_1 and x_2, from the aperture plane to the detector plane. Note that there is parity in the number of reflections in both arms of the interferometer, which is required for the images of the source to overlap with the same orientation at the detector plane.

The electric field at the detector plane is the sum on the two fields in Eq. B.5. However, the detector system used in this work is a direct detector that is sensitive to the integrated intensity. The time averaged intensity incident on the detector is,

$$\begin{aligned} I_{DFI}(\vec{\theta},\vec{B},k) &= \left\langle \left| E_1^{det}(\vec{\theta},\vec{B},k) + E_2^{det}(\vec{\theta},\vec{B},k) \right|^2 \right\rangle \\ &= r_M^{12} r_{bs}^2 |E_1(\vec{\theta},\vec{B},k)|^2 + r_M^{14} t_{bs}^2 |E_2(\vec{\theta},\vec{B},k)|^2 \\ &\quad - 2 r_M^{13} r_{bs} t_{bs} e^{i[\kappa(x_2-x_1)-\phi_{bs}]} \left\langle E_1^*(\vec{\theta},\vec{B},k) E_2(\vec{\theta},\vec{B},k) \right\rangle \end{aligned} \tag{B.6}$$

where in the last term, the negative sign comes from an absorbed e^{ϕ_M}, and the factor of two comes from the hermitian symmetry of $\left\langle E_1^{det\,*}(\vec{\theta},\vec{B},k) E_2^{det}(\vec{\theta},\vec{B},k) \right\rangle$. In practice, the detector signal is mean subtracted and the meaningful signal is given by,

$$I_{DFI}(\vec{\theta},\vec{B},k,z) = -2\sqrt{I_1(\vec{\theta},k) I_2(\vec{\theta},k)} e^{-i2\pi k\vec{B}\cdot\vec{\theta}} e^{i2\pi kz} e^{-i\phi_{bs}} \ , \tag{B.7}$$

where

$$\begin{aligned} I_1(\vec{\theta},k) &= r_M^{12} r_{bs}^2 |E_1(\vec{\theta},\vec{B},k)|^2 \\ I_2(\vec{\theta},k) &= r_M^{14} t_{bs}^2 |E_2(\vec{\theta},\vec{B},k)|^2 \,, \end{aligned} \tag{B.8}$$

and $z = x_2 - x_1$ is the optical path difference between the two beams resulting from instrumental effects. Note the intensities in Eq. B.8 have no explicit dependence on $\vec{B}$, however, R_2 does change as the spatial stage position is varied. This notation is chosen since the relevant information that results from variations in $\vec{B}$ is encoded in the relative phase of the electric fields and is lost when calculating the magnitude squared. The fact that intensity depends on aperture position is still implicitly denoted by subscripts.

Eq. B.7 describes the complex response of the DFI to a single point source at position $\vec{\theta}$ and frequency $k = \nu/c$. The full response is obtained by integrating over the source plane and spectral components,

$$\begin{aligned} I_{DFI}(\vec{B},z) &= -2e^{-i\phi_{bs}} \int_\Omega \int_0^\infty \sqrt{I_1(\vec{\theta},k) I_2(\vec{\theta},k)} e^{-i2\pi k \vec{B}\cdot\vec{\theta}} e^{i2\pi k z} \, dk \, d\Omega \\ &= -e^{-i\phi_{bs}} \int_\Omega \int_{-\infty}^\infty \sqrt{I_1(\vec{\theta},k) I_2(\vec{\theta},k)} e^{-i2\pi k \vec{B}\cdot\vec{\theta}} e^{i2\pi k z} \, dk \, d\Omega \,, \end{aligned} \tag{B.9}$$

where $d\Omega = d\theta_x d\theta_y$ denotes the source plane differential in terms of solid angle. The source spectrum is physically real and consists of frequencies $k \geq 0$, however, by imposing symmetry, integration can be extended to $\pm\infty$, absorbing a factor of two, as presented in the second line. By noting that the intensity coefficient is proportional to the source intensity through Eqs. B.8 and B.4, it can be observed that the DFI response is proportional to the Mutual Coherence Function (MCF) Eq. A.12. Since the interferometer measures only real intensity, the measurable signal is,

$$\begin{aligned} I_{DFI}(\vec{B},z) &= -\int_\Omega \int_{-\infty}^\infty \sqrt{I_1(\vec{\theta},k) I_2(\vec{\theta},k)} \cos(2\pi k z - 2\pi k \vec{B}\cdot\vec{\theta} - \phi_{bs}) \, dk \, d\Omega \\ &= -\int_{-\infty}^\infty \sqrt{I_1(k) I_2(k)} |\gamma_0(\vec{B},k)| \cos(2\pi k z + \phi_0(\vec{B},k) - \phi_{bs}) \, dk \,, \end{aligned} \tag{B.10}$$

where, $|\gamma_0(\vec{B},k)|$ and $\phi_0(\vec{B},k)$ represent the amplitude and phase of the complex visibility, respectively (see Eq. A.15).

The beamsplitter relative phase shift takes a value depending on orientation and is generally frequency dependent. For modelling purposes, the phase is set to $\phi_{bs} = \pi/2$ and is reasonably consistent with results, although the observed relative phase is typically less than this value. Using trigonometric identities, the cosine function becomes a sine function. For reasons that are explained in Ch. 4, the detector signal is inverted such that an increase in incident power produces a lower signal. Accounting for these effects, the most practical

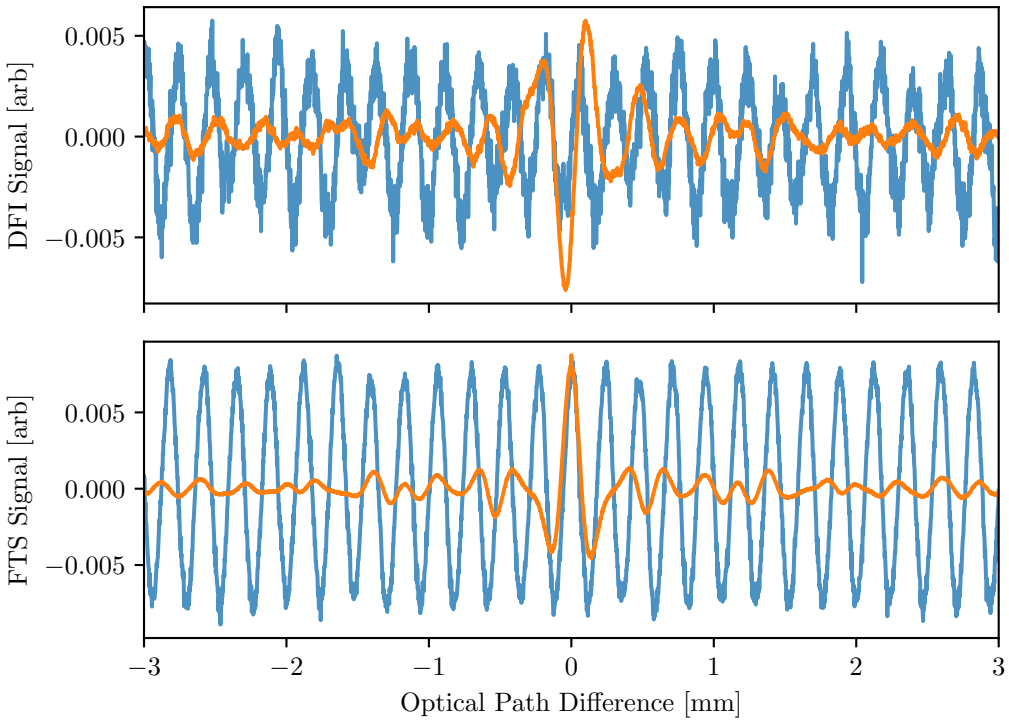


Figure B.2: Interferometer response in DFI (top) and FTS (bottom) observation modes when observing monochromatic (blue) and polychromatic (orange) light. DFI interferograms express a anti-symmetry about ZPD, characteristic of a signal generated from sine functions, while FTS interferograms express symmetry characteristic of signals generated from cosine functions. Polychromatic responses are normalized to their respective monochromatic response for clarity. The decreased amplitude of the DFI monochromatic response compared to FTS is a result of visibility effects.

expression for the DFI response used in this work is,

$$\begin{aligned} I_{DFI}(\vec{B},z) &= \int_{\Omega}\int_{-\infty}^{\infty} I_{12}(\vec{\theta},k)\sin\left(2\pi kz - 2\pi k\vec{B}\cdot\vec{\theta}\right)\,dk\,d\Omega \\ &= \int_{-\infty}^{\infty} I_{12}(k)|\gamma_0(\vec{B},k)|\sin\left(2\pi kz+\phi_0(\vec{B},k)\right)\,dk\,, \end{aligned} \tag{B.11}$$

where $I_{12}(\vec{\theta},k) = \sqrt{I_1(\vec{\theta},k)I_2(\vec{\theta},k)}$ is the geometric mean of the power spectrum of the light at the detector plane, which is proportional to the source power spectrum. Evidently, the response of the DFI is a superposition of sine waves with amplitudes and frequencies proportional to the intensity and frequency of the incident electric fields. The sinusoids are modulated by the visibility $|\gamma_0(\vec{B},k)|$ and shifted along the Optical Path Difference (OPD) axis by the phase of the complex visibility $\phi_0(B,k)$.

Diffraction Effects

So far, the discussion has implicitly assumed the interferometer apertures are infinite in extent. In reality, as the electric field produced by a points source, modelled as a plane wave, passes through the aperture, diffraction occurs as a result of the finite boundary of the aperture. The wave proceeds to spread out as it propagates through the optical system, potentially interacting with other apertures in a similar way, until it reaches the focusing optics where the field is focused onto some finite region of the detector plane. The intensity distribution that is produced on the detector plane is called the Point Spread Function (PSF) and its finite extent is an expression of the diffraction limit of imaging optics [111].

Diffraction effects can be incorporated into the modelling process by assuming each point on the surface of the aperture is a point source which emits a spherical wave. The intensity at a particular point on the detector plane is then the square of the integral of the electric field amplitude taken over the surface of the aperture, effectively the same approach that was used to obtain Eq. B.9 with respect to the source plane. The end result is that the PSF is given by the square of the Fourier transform of the electric field amplitude distribution on the aperture plane. The signal recorded a the detector is then the convolution of the source plane with the PSF, producing the diffraction limited image on the source plane, integrated over the area of the detector surface. Although this approach is present in the literature [131], the process can be carried out in the opposite direction, from detector plane to source plane which seems to be more common when modelling existing telescopes [117].

In the more typical treatment, it is imagined that a point on the detector plane is replaced by a point source. The electric field is reverse propagated through the optical system, diffracting as a result of interaction with the apertures, and is then projected onto the source plane. The source plane intensity distribution resulting from this process is called the Beam Spread Function (BSF). Much like the PSF, the BSF is obtained as the square of the Fourier transform of the field amplitude distribution on the aperture plane. The sensitivity distribution of the detector is then given by the convolution of the BSF with the detector area. This produces a quantity, which when normalized such that the maximum value in the sensitivity distribution is one, is called the beam or normalized power pattern in radio astronomy [117]. The beam effectively describes how well the detector couples to different regions of the source plane. Multiplying the source intensity distribution by the detector beam gives the effective source intensity distribution that the detector sees,

$$I_{det}(\vec{\theta},k) = \xi(\vec{\theta},k) \times I(\vec{\theta},k) \,, \tag{B.12}$$

where $\xi(\vec{\theta},k)$ is the generally frequency dependent beam function. The beam function appropriate for this work is obtained in § 3.4.2. For a well designed and aligned spatial interferometer, the electric field distributions resulting from a point source on the detector plane propagating through each aperture will have the same shape and be aligned such that $I_{12}(\vec{\theta},k)$ can be replaced by $I_{det}(\vec{\theta},k)$ in Eq. B.11.

From the van Cittert-Zernike theorem Eq. A.14, the spatial frequency components of the source intensity distribution are obtained as the Fourier transform of the source intensity distribution. As a result of the multiplication property of Fourier transforms § C.1, the

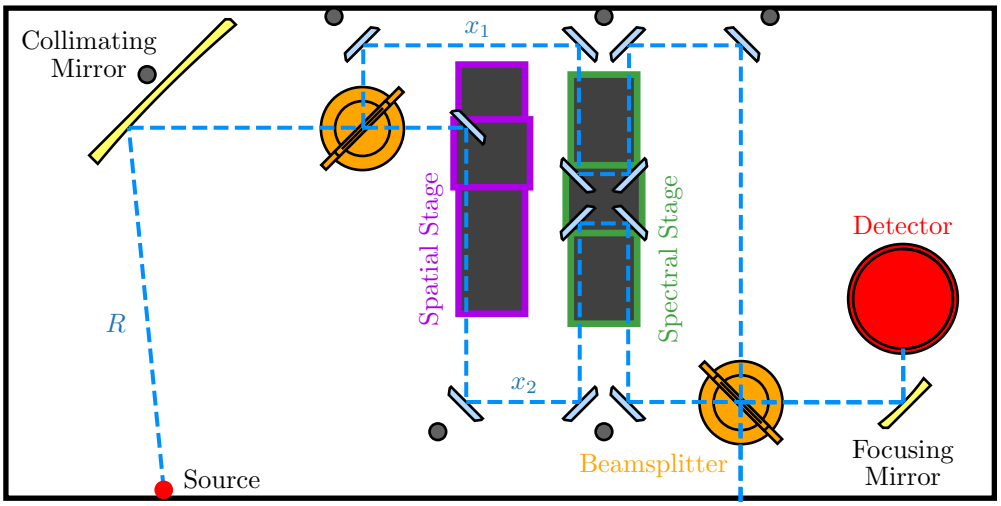


Figure B.3: Diagram of the interferometer in FTS mode.

spatial frequency components obtained from the detector measurements are,

$$\mathcal{F}\left\{I_{det}(\vec{\theta},k)\right\} = \hat{\xi}(u,v) * \Gamma_0(u,v)\,, \tag{B.13}$$

where $\hat{\xi}(u,v)$ is the Fourier transform of the beam function, and $u,v = \vec{B}k$ coordinates are used. When sampling a particular point in the uv-plane in the presence of diffraction, the value obtained is actually a weighted sum of the values in a region defined by $\hat{\xi}(u,v)$ and centered on the nominal sample coordinate. This result has a natural Fourier analysis interpretation. Just like the sampling frequency of a time dependent signal determines the bandwidth in frequency space (see § 2.3.2), sampling interval in the uv-plane determines the extent of the reconstructed image on the source plane. By truncating the source plane with the beam function, the separation between statistically independent uv samples has been increased.

B.2 Spectral Interferometer Response

In order to extract visibilities from Eq. B.11 it is necessary to determine the normalization factor $I_{12}(k)$. Accomplishing this requires a measurement that removes the effects of the visibilities. As shown in Ap. A, when the input aperture of the interferometer are made to overlap, producing a zero-baseline configuration $\vec{B} = 0$, the complex visibility is guaranteed to be unity. Although physically overlapping the input apertures is impossible using the DFI configuration shown in Fig. B.1, the interferometer can be modified to use a single input aperture by introducing an additional beamsplitter and mirror. This configuration is shown in Fig. B.3 and is functionally equivalent to a standard FTS.

Analyzing this configuration proceeds in much the same way as Sec. B.1. A spherical wave emitted from a source at the focal point of the collimating mirror is converted to a plane wave on reflection that propagates to the first beamsplitter which functions as the input aperture of the interferometer. The electric field at the aperture plane just before the

beamsplitter is,

$$E(\vec{\theta},k) \approx r_M \frac{E_0(\vec{\theta},k)}{R} e^{-[i\omega t+\phi_{\vec{\theta}}]} , \tag{B.14}$$

where the position of the aperture plane has been chosen such that the phase shift associated with the collimating mirror reflection and wave propagation for an on-axis point source is an integer multiple of 2π. Variables in Eq. B.14 have the same interpretation as Eq. B.1. Point sources displaced from the optical axis by an angle $\vec{\theta}$ will have a different amplitude, but if $|\vec{\theta} f_{col}| \ll R$, the effect will be negligible. A position dependent phase shift is also included $\phi_{\vec{\theta}}$ which is zero for on-axis sources.

Tracing the beams through both arms of the interferometer gives the following electric fields at the detector,

$$\begin{aligned} E_1^{det}(\vec{\theta},k) &= r_M^7 r_{bs}^2 E(\vec{\theta},k) e^{i[\kappa x_1-\omega\tau+7\phi_M+\phi_{bs1}+\phi_{bs2}]} \\ E_2^{det}(\vec{\theta},k) &= r_M^7 t_{bs}^2 E(\vec{\theta},k) e^{i[\kappa x_2-\omega\tau+7\phi_M]} , \end{aligned} \tag{B.15}$$

where the variables have the same interpretation as those in Eq. B.5. There are an odd number of reflections in both beams, which indicates the beams should overlap with the same orientation when incident on the detector. Note that in this case, there are two beamsplitter relative phase terms to account for the two beamsplitter interactions. It is also the case that the beams are not balanced in the sense that one arm transmits through the beamsplitters twice while the other arm reflects off the beamsplitters twice instead of one reflection and one transmission each.

The intensity at the detector is computed as,

$$\begin{aligned} I_{FTS}(\vec{\theta},k) &= \left\langle \left| E_1^{det}(\vec{\theta},k) + E_2^{det}(\vec{\theta},k) \right|^2 \right\rangle \\ &= r_M^{14} r_{bs}^4 |E(\vec{\theta},k)|^2 + r_M^{14} t_{bs}^4 |E(\vec{\theta},k)|^2 \\ &\quad + 2 r_m^{14} r_{bs}^2 t_{bs}^2 e^{i[\kappa(x_2-x_1)-\phi_{bs1}-\phi_{bs2}]} |E(\vec{\theta},k)|^2 . \end{aligned} \tag{B.16}$$

In practice, the detector signal is mean subtracted and the useful signal is given by,

$$I_{FTS}(\vec{\theta},k,z) = 2\sqrt{I_1(\vec{\theta},k) I_2(\vec{\theta},k)} e^{i[2\pi k z-\phi_{bs1}-\phi_{bs2}]} , \tag{B.17}$$

where

$$\begin{aligned} I_1(\vec{\theta},k) &= r_M^{14} r_{bs}^4 |E(\vec{\theta},k)|^2 \\ I_2(\vec{\theta},k) &= r_M^{14} t_{bs}^4 |E(\vec{\theta},k)|^2 \end{aligned} \tag{B.18}$$

are the intensities incident on the detector from each beam of the interferometer, and $z = x_2 - x_1$ is the OPD between the two beams. Using the assumption $r_{bs} = t_{bs}$, which is the intended relationship for a 50:50 intensity beamsplitter, $I_1(\vec{\theta},k) = I_2(\vec{\theta},k) = I_0(\vec{\theta},k)$.

The full complex response of the detector is obtained by integrating over the source

plane and spectral components,

$$\begin{aligned} I_{FTS}(z) &= \int_{\Omega}\int_{-\infty}^{\infty} I_0(\vec{\theta},k)e^{i[2\pi kz-\phi_{bs1}-\phi_{bs2}]}\,dk\,d\Omega \\ &= \int_{-\infty}^{\infty} I_0(k)e^{i[2\pi kz-\phi_{bs1}-\phi_{bs2}]}\,dk\,, \end{aligned} \tag{B.19}$$

where spectral symmetry has been imposed, extending the integration limits to $\pm\infty$ and absorbing a factor of two. The detector measures only real intensity, thus the detector output is obtained by taking the real part of Eq. B.19,

$$I_{FTS}(z) = \int_{-\infty}^{\infty} I_0(k)\cos(2\pi kz-\phi_{bs1}-\phi_{bs2})\,dk\,. \tag{B.20}$$

As discussed in Sec. B.3, the difference between relative beamsplitter phase shifts can take the values $\phi_{bs1}-\phi_{bs2}\in\{0,\pm\pi\}$ depending on relative orientation. Results obtained in this work are consistent with $\pm\pi$. Additionally, the detector response is such that increased incident power produces a lower signal. Accounting for these effects, the interferometer response in FTS mode is,

$$I_{DFI}(z) = \int_{-\infty}^{\infty} I_0(k)\cos(2\pi kz)\,dk\,, \tag{B.21}$$

where $I_0(k)$ is proportional to the power spectrum of the light entering the FTS which is proportional to the source power spectrum.

Note that Eq. B.21 is independent of the complex visibility. In the ideal case $2I_0(k)=I(k)$ from Eq. B.11, and the FTS response is used as to extract the normalization constant required to determine the visibility $|\gamma(B,k)|$. However, in general $R_1R_2\neq R^2$ (see Figs. B.1 and B.3). This complication is a product of the collimating optics used to simulate a distant source, and would not be present with an astronomical instrument. In any case, the FTS configuration still provides a measurement with unit visibility that can be used to accurately estimate the source power spectrum which, in principle, can be used to calibrate visibility measurements.

B.3 Beamsplitter Phase

Constraints can be placed on the phase properties of beamsplitters by analyzing the system using conservation of energy. Consider the situation shown in Fig. B.4 where electric fields E_1 and E_2 enter the beamsplitter, producing fields E_3 and E_4 on the opposite sides. If the energy contained in the electric fields is conserved,

$$|E_1|^2+|E_2|^2=|E_3|^2+|E_4|^2\,. \tag{B.22}$$

The electric field E_3 is produced by the sum of the reflected component of E_1 and the transmitted component of E_2. A similar construction exists for E_4, resulting in the equations,

$$\begin{aligned} E_3 &= re^{i\phi_{1r}}E_1 + te^{i\phi_{1t}}E_2 \\ E_4 &= re^{i\phi_{2r}}E_2 + te^{i\phi_{2t}}E_1 \,, \end{aligned} \tag{B.23}$$

where r and t are taken to be real numbers and represent the reflection and transmission coefficients of the beamsplitter, respectively. The angles ϕ_{ir} and ϕ_{it} represent the phase shifts resulting from reflected and transmitted components of opposing sides of the beamsplitter. Noting that $|E|^2 = E^*E$, Eq. B.23 in connection with Eq. B.22 gives,

$$\begin{aligned} |E_3|^2 + |E_4|^2 = (r^2+t^2)(|E_1|^2 + |E_2|^2) \\ + rtE_1^*E_2(e^{i\phi_2} + e^{-i\phi_1}) + rtE_2^*E_1(e^{i\phi_1} + e^{-i\phi_2}) \,, \end{aligned} \tag{B.24}$$

where

$$\phi_i = \phi_{ir} - \phi_{it} \,, \tag{B.25}$$

are the relative phase shifts between reflected and transmitted components of opposing sides of the beamsplitter. For a lossless system $r^2 + t^2 = 1$, and Eq. B.22 holds when the last two terms on the right-hand side of Eq. B.24 cancel. If the electric field amplitudes are real, as is usually the case, this condition requires,

$$e^{i\phi_1} + e^{-i\phi_1} + e^{i\phi_2} + e^{-i\phi_2} = 4\cos\left(\frac{\phi_1+\phi_2}{2}\right)\cos\left(\frac{\phi_1-\phi_2}{2}\right) = 0 \,, \tag{B.26}$$

which is true when,

$$\begin{aligned} \phi_1 + \phi_2 &= \pm n\pi \\ &\text{or} \\ \phi_1 - \phi_2 &= \pm n\pi \,, \end{aligned} \tag{B.27}$$

where n is an odd integer. Given the agreement between data and modelled responses for DFI and FTS observing modes, the relative phase shifts for the beamsplitters are consistent with,

$$\begin{aligned} \phi_1 &= +\pi/2 + \varepsilon \\ \phi_2 &= -\pi/2 + \varepsilon \,, \end{aligned} \tag{B.28}$$

where ε is a phase perturbation that is generally frequency dependent and is affected by the phase shift on transmission. The sign associated with the relative phase shift is attributed to the reflection properties of metal mesh beamsplitters.

To model the phase shifts on reflection, it is assumed that one side of the beamsplitter is dominated by an inductive interaction while the other side is capacitive. This assumption is based on the fact that the metal mesh beamsplitters used in this work are fabricated as layers of inductive and capacitive conductor grids. Reflection is modelled as light propagating along a lossless transmission line which are terminated by reactive elements [132]. Under

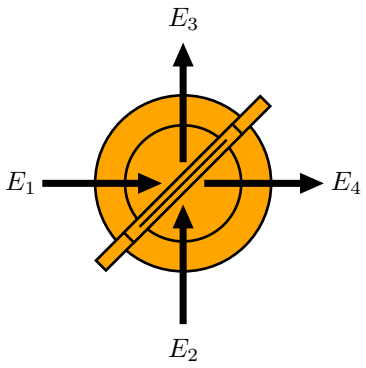


Figure B.4: Diagram showing the orientation of electric fields with respect to the beamplitter when deriving phase constraints using conservation of energy.

this framework, reflected voltage signals, analogous to electric field signals, acquire a phase shift of,

$$\phi_r = 2\arctan\left(\frac{Z_0}{X}\right) , \tag{B.29}$$

where Z_0 is the characteristic impedance of the transmission line, free space in the case of light, and $X = X_L + X_C$ is the reactance, with $X_L = \omega L$ and $X_C = -1/\omega C$ the inductive and capacitive reactance which terminates the line, respectively. It can be seen that $|X| = Z_0$ results in a phase shift

$$\begin{aligned} \phi_L &= +\pi/2 \\ \phi_C &= -\pi/2 , \end{aligned} \tag{B.30}$$

when termination is purely inductive or capacitive, respectively. Eqs. B.30 referrer to the phase shifts on reflection and are only approximations as either side of the beamsplitter is expected to have both inductive and capacitive contributions.

Fig. B.5 provides some insight about the phase properties of the beamsplitters. The blue curve in the top panel shows data obtained when observing an 850 GHz source in DFI mode. The red and orange curves are obtained by placing the beamsplitter in front of the dynamic aperture, oriented such that light is incident on the S and W side of the beamsplitter, respectively. S and W are designations given to each side of the beamsplitter, used purely for reference, and were chosen because these notations were already present on the beamsplitter when obtained from the manufacturer. Evidently, the beamsplitter introduces a constant phase shift on transmission with both orientations. That is, the phase shift on transmission preserves Eq. B.28.

The bottom panel of Fig. B.5 shows interferograms obtained in FTS mode where the input beamsplitter is oriented in two ways. From Eq. B.28 in connection with Eq. B.20, the

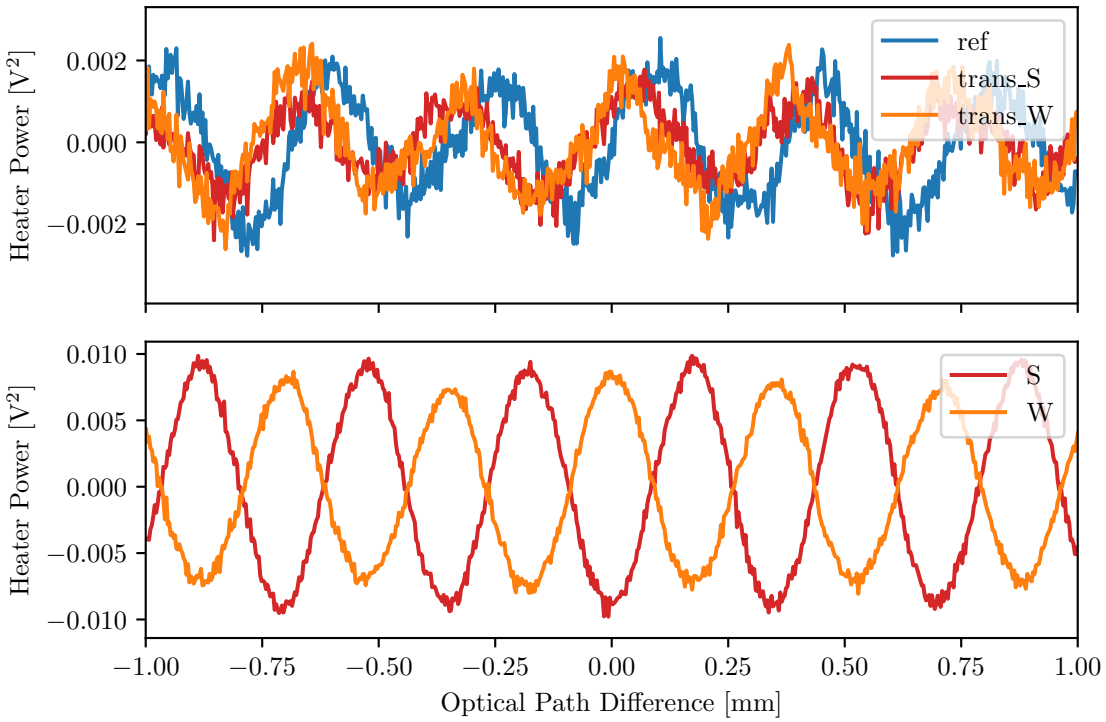


Figure B.5: Beamsplitter phase relationships. The top panel shows data obtained in DFI mode (blue) with a beamsplitter inserted in front of the dynamic aperture, oriented such that light is incident on either the S (red) or W (orange) side. The curves show a phase shift resulting from transmission, independent of orientation. The bottom panel shows data obtained in FTS mode where the input beamsplitter is oriented such that light is incident on the S (red) or W (orange) side. The curves show a phase difference of $\sim\pi$, which is consistent with energy conservation constraints and the beamsplitter relative phase, as described in the text.

possible beamsplitter orientation combinations produce,

$$\begin{aligned} \phi_1 - \phi_1 = \phi_2 - \phi_2 = 0 \\ |\phi_2 - \phi_1| = |\phi_1 - \phi_2| = \pi\,. \end{aligned} \tag{B.31}$$

As such, flipping one of the beamsplitters in FTS mode is expected to induce a $\pm\pi$ phase shift in the resulting interferogram. This result is observed in Fig. B.5.

Appendix C

Fourier Analysis Proofs

In this appendix, proofs for the Fourier properties and transforms used in this work are derived. The Fourier transform,

$$\mathcal{F}\left[f(x)\right] \equiv \int_{-\infty}^{\infty} f(x) e^{-i2\pi kx}\, dx = \hat{f}(k)\ , \tag{C.1}$$

and inverse Fourier transform,

$$\mathcal{F}^{-1}\left[\hat{f}(k)\right] \equiv \int_{-\infty}^{\infty} \hat{f}(k) e^{i2\pi kx}\, dk = f(x)\ , \tag{C.2}$$

are repeated for convenience. The independent variables x and k are reciprocal pairs typically representing either position and spatial frequency, or time and temporal frequency, respectively.

C.1 Properties

Duality

Consider the definition of the inverse Fourier transform Eq. C.2 and apply the substitution $x \to -x$,

$$\begin{aligned} f(-x) &= \int_{-\infty}^{\infty} \hat{f}(k) e^{-i2\pi kx}\, dk \\ &= \mathcal{F}\left[\hat{f}(k)\right]\ . \end{aligned} \tag{C.3}$$

Interchanging the variables x and k gives the result,

$$\mathcal{F}\left[\hat{f}(x)\right] = f(-k)\ . \tag{C.4}$$

Effectively, if a function $f(x)$ has the Fourier transform $\hat{f}(k)$, the Fourier transform of $\hat{f}(x)$ is $f(-k)$. The utility of this result comes from the fact that if you can find the Fourier transform of a given function, the Fourier transform of the result is obtained as well. For example, as demonstrated in Sec. C.2, the Fourier transform of the rectangle function is a sinc function, and as a result of the duality property, the Fourier transform of the sinc function is the rectangle function.

Linearity

The Fourier transform of a weighted sum of functions is obtained as,

$$\begin{aligned}\mathcal{F}\left[af(x)+b(gx)\right] &= \int_{-\infty}^{\infty}\left[af(x)+bg(x)\right]e^{-i2\pi kx}\,dx \\ &= \int_{-\infty}^{\infty} af(x)e^{-i2\pi kx}\,dx + \int_{-\infty}^{\infty} bg(x)e^{-i2\pi kx}\,dx \\ &= a\int_{-\infty}^{\infty} f(x)e^{-i2\pi kx}\,dx + b\int_{-\infty}^{\infty} g(x)e^{-i2\pi kx}\,dx \\ &= a\hat{f}(k)+b\hat{g}(k)\ ,\end{aligned} \tag{C.5}$$

where a and b are constants. As such, the Fourier transform of the sum of functions is the sum of their respective Fourier transforms. Any multiplicative scaling factors applied to a function is also applied to its Fourier transform.

Scaling

When the independent variable of a function is scaled by some nonzero real constant a, the Fourier transform is

$$\mathcal{F}\left[f(ax)\right] = \int_{-\infty}^{\infty} f(ax)e^{-i2\pi kx}\,dx\ . \tag{C.6}$$

Using the substitution,

$$\begin{aligned} u = ax &\rightarrow x = u/a \\ du = a\,dx &\rightarrow dx = du/a\ ,\end{aligned} \tag{C.7}$$

the Fourier transform becomes,

$$\frac{1}{a}\int_{-\infty}^{\infty} f(u)e^{-i2\pi ku/a}\,du\ . \tag{C.8}$$

In the case where $a > 0$, Eq. C.8 is unchanged. If $a < 0$, then negative signs in expressions involving u/a and its derivatives cancel since $u \propto a$. As a consequence, the only change is $f(u) \rightarrow f(-u)$, and the resulting equation that cover both cases is,

$$\mathcal{F}\left[f(ax)\right] = \frac{1}{|a|}\hat{f}\left(\frac{k}{a}\right)\ . \tag{C.9}$$

The scaling property expresses the reciprocal nature of Fourier transforms. If a function, for example a Gaussian, is widened by a factor of a, its Fourier transform will be narrowed and increased in maximum amplitude by the same factor of a.

Translation

When a function is translated along the independent axis by the value a, the Fourier transform is

$$\mathcal{F}\left[f(x-a)\right] = \int_{-\infty}^{\infty} f(x-a)e^{-i2\pi kx}\,dx\,. \tag{C.10}$$

Using the substitution

$$\begin{aligned} u = x-a \rightarrow x = u+a \\ du = dx \rightarrow dx = du\,, \end{aligned} \tag{C.11}$$

the Fourier transform becomes,

$$\begin{aligned} \mathcal{F}\left[f(x-a)\right] &= \int_{-\infty}^{\infty} f(u)e^{-i2\pi k(u+a)}\,du \\ &= \int_{-\infty}^{\infty} f(u)e^{-i2\pi ku}e^{-i2\pi ka}\,du \\ &= e^{-i2\pi ka}\hat{f}(k)\,. \end{aligned} \tag{C.12}$$

Translating the function by a results in a Fourier transform of the original function multiplied by a complex exponential. This result can be interpreted as the Fourier transform of the original function with a linear phase shift of slope $-a$.

Phase Shift

The phase shift property is the inverse of the translation property. If a function is multiplied by a linear phase term, the Fourier transform is,

$$\begin{aligned} \mathcal{F}\left[e^{i2\pi ax}f(x)\right] &= \int_{-\infty}^{\infty} f(x)e^{i2\pi ax}e^{-i2\pi kx}\,dx \\ &= \int_{-\infty}^{\infty} f(x)e^{-i2\pi x(k-a)}\,dx \\ &= \hat{f}(k-a)\,. \end{aligned} \tag{C.13}$$

Applying a linear phase shift with slope a to a function thus result in a Fourier transform consisting of the Fourier transform of the original function shifted by a.

Multiplication

When a function is the product of two functions $f(x)$ and $g(x)$ the Fourier transform is,

$$\begin{aligned}\mathcal{F}\left[f(x)g(x)\right] &= \int_{-\infty}^{\infty} f(x)g(x)e^{-i2\pi kx}\,dx \\ &= \int_{-\infty}^{\infty} f(x)\left[\int_{-\infty}^{\infty} \hat{g}(a)e^{i2\pi ax}\,da\right] e^{-i2\pi kx}\,dx \\ &= \int_{-\infty}^{\infty} \hat{g}(a)\left[\int_{-\infty}^{\infty} f(x)e^{-i2\pi x(k-a)}\,dx\right] da \\ &= \int_{-\infty}^{\infty} \hat{g}(a)\hat{f}(k-a)\,da\,.\end{aligned} \tag{C.14}$$

In the second line, $g(x)$ is replaced by its inverse Fourier transform equivalent (Eq. C.2). The final line of Eq. C.14 can be identified as the convolution (see Eq. C.16), of the functions $\hat{f}(a)$ and $\hat{g}(a)$ which are the Fourier transforms of $f(x)$ and $g(x)$, respectively, and a is an arbitrary variable with the same units as k. With this consideration, the multiplication property can be written,

$$\mathcal{F}\left[f(x)g(x)\right] = \left[\hat{f} * \hat{g}\right](k)\,, \tag{C.15}$$

where $*$ indicates the convolution operator.

Convolution

The convolution property is the inverse of the multiplication property. The convolution of two functions is defined as,

$$\left[f * g\right](x) = \int_{-\infty}^{\infty} f(a)g(x-a)\,da \tag{C.16}$$

where a is an arbitrary variable with the same units as x. If a function results from the convolution of two functions, the Fourier transform is,

$$\begin{aligned}\mathcal{F}\left[\left[f * g\right](x)\right] &= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(a)g(x-a)\,da\right] e^{-i2\pi kx}\,dx \\ &= \int_{-\infty}^{\infty} f(a)\left[\int_{-\infty}^{\infty} g(x-a)e^{-i2\pi kx}\,dx\right] da \\ &= \hat{g}(k)\int_{-\infty}^{\infty} f(a)e^{-i2\pi ak}\,da \\ &= \hat{g}(k)\hat{f}(k)\,,\end{aligned} \tag{C.17}$$

where the third line is obtained using the translation property Eq. C.12. This result shows the Fourier transform of the convolution of two functions is the product of the Fourier transforms of the two functions.

Correlation

Closely related to the convolution property is the correlation property. The correlation of two functions is defined as,

$$[f \star g](x) = \int_{-\infty}^{\infty} f^*(a) g(a+x)\, da \,, \tag{C.18}$$

where a is an arbitrary variable with the same units as x, and the superscript * indicates the complex conjugate. If a function results from the correlation of two functions, the Fourier transform is,

$$\begin{aligned}
\mathcal{F}\left[[f \star g](x)\right] &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} f^*(a) g(x+a)\, da \right] e^{-i2\pi kx}\, dx \\
&= \int_{-\infty}^{\infty} f^*(a) \left[\int_{-\infty}^{\infty} g(x+a) e^{-i2\pi kx}\, dx \right] da \\
&= \hat{g}(k) \int_{-\infty}^{\infty} f^*(a) e^{i2\pi ak}\, da \\
&= \hat{g}(k) \left[\int_{-\infty}^{\infty} f(a) e^{-i2\pi ak}\, da \right]^* \\
&= \hat{g}(k) \hat{f}^*(k) \,.
\end{aligned} \tag{C.19}$$

Again, the third line is obtained using the translation property Eq. C.12. This result shows the Fourier transform of the correlation of two functions is the product of the Fourier transform of one of the functions and the complex conjugate of the Fourier transform of the other function. Which transform the complex conjugate is applied to is arbitrary.

C.2 Functions and Distributions

Dirac Delta

The Dirac delta function $\delta(x)$, or simply delta function, is a mathematical abstraction with the property that it is zero everywhere, except at the origin where it is infinite. It is further constrained to satisfy the identity,

$$\int_{-\infty}^{\infty} \delta(x)\, dx = 1 \,. \tag{C.20}$$

A practical application of the delta function is to use it as a filter,

$$\int_{-\infty}^{\infty} f(x) \delta(x-a)\, dx = f(a) \,. \tag{C.21}$$

The sum of a collection of evenly spaced delta functions is called a Dirac comb function,

$$\text{Ш}_L(x) = \sum_{n=-\infty}^{\infty} \delta(x - nL) \,, \tag{C.22}$$

where L is the spacing between delta functions.

The Fourier transform of the delta function is easily computed using the filter property,

$$\mathcal{F}\left[\delta(x-a)\right] = \int_{-\infty}^{\infty} \delta(x-a) e^{-i2\pi kx} dx = e^{-i2\pi ka} \,. \tag{C.23}$$

Conversely, the Fourier transform of the complex exponential is,

$$\begin{aligned} \mathcal{F}\left[e^{i2\pi ax}\right] &= \mathcal{F}\left\{\mathcal{F}\left[\delta(k-a)\right]^*\right\} \\ &= \mathcal{F}\left\{\mathcal{F}^{-1}\left[\delta(k-a)\right]\right\} \\ &= \delta(k-a) \,. \end{aligned} \tag{C.24}$$

To determine the Fourier transform of the comb function, note that Eq. C.22 is periodic in x with period L. As such it can be expanded as a Fourier series. The complex coefficient, Eq. 2.15, is computed as,

$$c_n = \frac{1}{L} \int_{-L/2}^{L/2} \sum_{m=-\infty}^{\infty} \delta(x-mL) e^{-i2\pi k_n x} dx = \frac{1}{L} \,, \tag{C.25}$$

Since the integral is evaluated over the range $[-L/2, L/2]$, the only delta function that contributes is the one at $x = 0$. The Fourier series Eq. 2.14, is then,

$$\text{Ш}_L(x) = \frac{1}{L} \sum_{n=-\infty}^{\infty} e^{i2\pi k_n x} \,. \tag{C.26}$$

Taking the Fourier transform of this equation, and by using Eq. C.24, the Fourier transform of the Dirac comb is,

$$\mathcal{F}\left[\text{Ш}_L(x)\right] = \frac{1}{L} \sum_{n=-\infty}^{\infty} \delta\left(k - \frac{n}{L}\right) \,, \tag{C.27}$$

which is itself a Dirac comb.

Sinusoid

The Fourier transform of a cosine function is obtained as,

$$\begin{aligned} \mathcal{F}\left[A\cos(2\pi k_0 x)\right] &= A \int_{-\infty}^{\infty} \cos(2\pi k_0 x) e^{i2\pi kx} dx \\ &= \frac{A}{2} \int_{-\infty}^{\infty} \left[e^{i2\pi k_0 x} + e^{-i2\pi k_0 x}\right] e^{i2\pi kx} dx \\ &= \frac{A}{2} \left[\delta(k-k_0) + \delta(k+k_o)\right] \,, \end{aligned} \tag{C.28}$$

where the second line employs the exponential representation of the cosine function, and the third line follows from Eq. C.24. Using a similar procedure, the Fourier transform of a sine function is,

$$\mathcal{F}\left[A\sin(2\pi k_0 x)\right] = \frac{A}{2i} \left[\delta(k-k_0) - \delta(k+k_o)\right] \,. \tag{C.29}$$

Rectangle

The rectangle function is defined as,

$$\Pi\left(\frac{x}{a}\right) = \begin{cases} 0 & \text{if } |x| > \frac{a}{2} \\ 1 & \text{if } |x| \leq \frac{a}{2} \end{cases}, \tag{C.30}$$

where a is a real and positive constant.

Taking the Fourier transform of the rectangle function gives,

$$\begin{aligned} \mathcal{F}\left[\Pi\left(\frac{x}{a}\right)\right] &= \int_{-\infty}^{\infty} \Pi\left(\frac{x}{a}\right) e^{-i2\pi kx}\, dx \\ &= \int_{-a/2}^{a/2} e^{-i2\pi kx}\, dx \\ &= \left[\frac{-e^{i2\pi kx}}{i2\pi k}\right]_{-a/2}^{a/2} \\ &= \frac{a\sin(\pi ak)}{\pi ak} \\ &= a\,\mathrm{sinc}(ak)\,, \end{aligned} \tag{C.31}$$

where

$$\mathrm{sinc}(ax) = \frac{\sin(\pi ax)}{\pi ax}\,, \tag{C.32}$$

is the normalized sinc function.

Sinc

Having previously shown that the Fourier transform of the rectangle function is a sinc function (Eq. C.31), the Fourier transform of the sinc function is easily obtained with application of the duality property (Eq. C.4):

$$\begin{aligned} \mathcal{F}\left[\mathrm{sinc}(bx)\right] &= \frac{1}{b}\Pi\left(\frac{-k}{b}\right) \\ &= \frac{1}{b}\Pi\left(\frac{k}{b}\right)\,, \end{aligned} \tag{C.33}$$

where the second line is obtained as a consequence of the even symmetry of the rectangle function. Using the substitution $a = 1/b$, Eq. C.33 can be written as,

$$\mathcal{F}\left[\mathrm{sinc}\left(\frac{x}{a}\right)\right] = a\Pi(ak)\,. \tag{C.34}$$

The Fourier transform of the sinc function is the rectangle function.

Step Function Decay

The truncated exponential decay function has the form,

$$f(x) = u(x)e^{-ax}\,, \tag{C.35}$$

where a is a real constant, and

$$u(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \geq 0\,, \end{cases} \tag{C.36}$$

is the Heaviside step function. Eq. C.35 is useful for modelling the impulse response of a detector system.

The Fourier transform of Eq. C.35 can be obtained as,

$$\begin{aligned} \mathcal{F}\left[u(x)e^{-ax}\right] &= \int_{-\infty}^{\infty} u(x)e^{-ax}e^{-i2\pi kx}\,dx \\ &= \int_{0}^{\infty} e^{-(a+i2\pi k)x}\,dx \\ &= \left[\frac{-e^{-(a+i2\pi k)x}}{a+i2\pi k}\right]_0^{\infty} \\ &= \frac{1}{a+i2\pi k}\,. \end{aligned} \tag{C.37}$$

This result can be associated with the frequency response of a detector system, and provides a quantitative way of relating the detector response to its frequency response.

Gaussian

The Fourier transform of a Gaussian function is obtained by evaluating the integral,

$$\begin{aligned} \mathcal{F}\left[e^{-ax^2}\right] &= \int_{-\infty}^{\infty} e^{-ax^2}e^{-i2\pi kx}\,dx \\ &= \int_{-\infty}^{\infty} e^{-(ax^2+i2\pi kx)}\,dx \\ &= e^{-\frac{(\pi k)^2}{a}} \int_{-\infty}^{\infty} e^{-(\sqrt{a}x+i\pi k/\sqrt{a})^2}\,dx\,. \end{aligned} \tag{C.38}$$

The last line is obtained by completing the square of the expression in the exponential in the second line,

$$-(ax^2 + i2\pi kx) = -\left(\sqrt{a}x + \frac{i\pi k}{\sqrt{a}}\right)^2 - \frac{(\pi k)^2}{a}\,. \tag{C.39}$$

Using the substitution,

$$\begin{aligned} u &= \sqrt{a}x + \frac{i\pi k}{\sqrt{a}} \\ du &= \sqrt{a}\,dx \rightarrow dx = \frac{du}{\sqrt{a}}\,, \end{aligned} \tag{C.40}$$

Eq. C.38 becomes,

$$\begin{aligned} \mathcal{F}\left[e^{-ax^2}\right] &= \frac{e^{-\frac{(\pi k)^2}{a}}}{\sqrt{a}} \int_{-\infty}^{\infty} e^{-u^2}\,du \\ &= \sqrt{\frac{\pi}{a}} e^{-\frac{(\pi k)^2}{a}}\,, \end{aligned} \tag{C.41}$$

where the well known integration of a Gaussian function $\int_{-\infty}^{\infty} e^{-u^2}\,du = \sqrt{\pi}$, is used in the last line. It can be observed that the Fourier transform of a Gaussian function is another Gaussian function.

Appendix D

Technical Drawings

In this appendix, technical drawing for main interferometer components are presented.

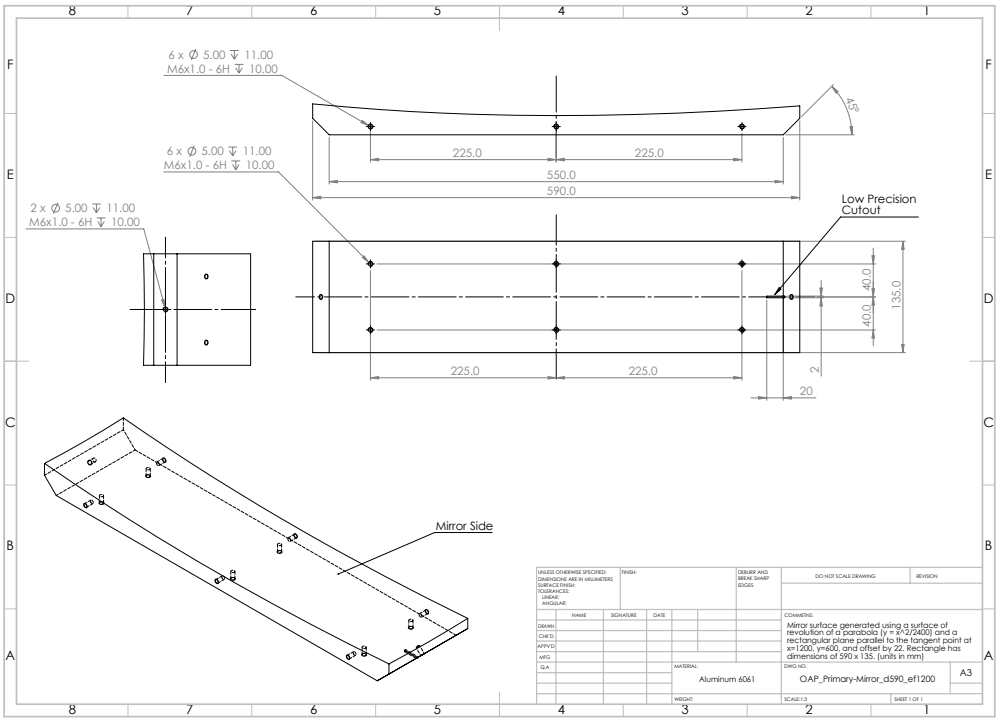


Figure D.1: Off-axis parabolic collimating mirror with effective an effective focal length of $1,200\,\text{mm}$.

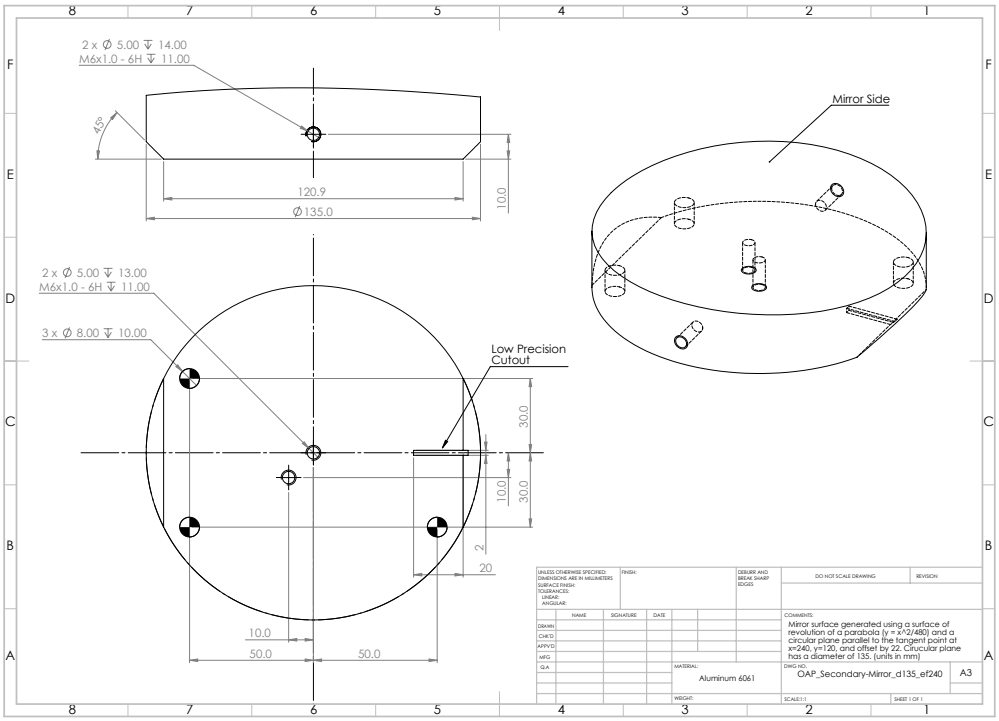


Figure D.2: Off-axis parabolic focusing mirror with effective an effective focal length of 240 mm.

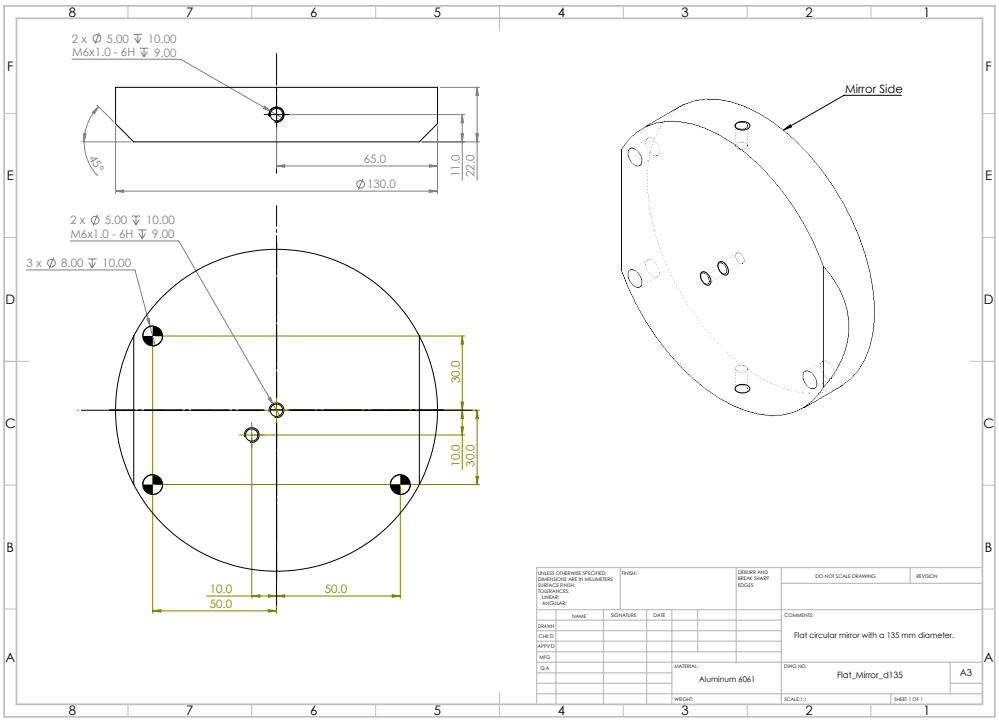


Figure D.3: Circular flat mirror which serves as input aperture and optical relay in the interferometer.

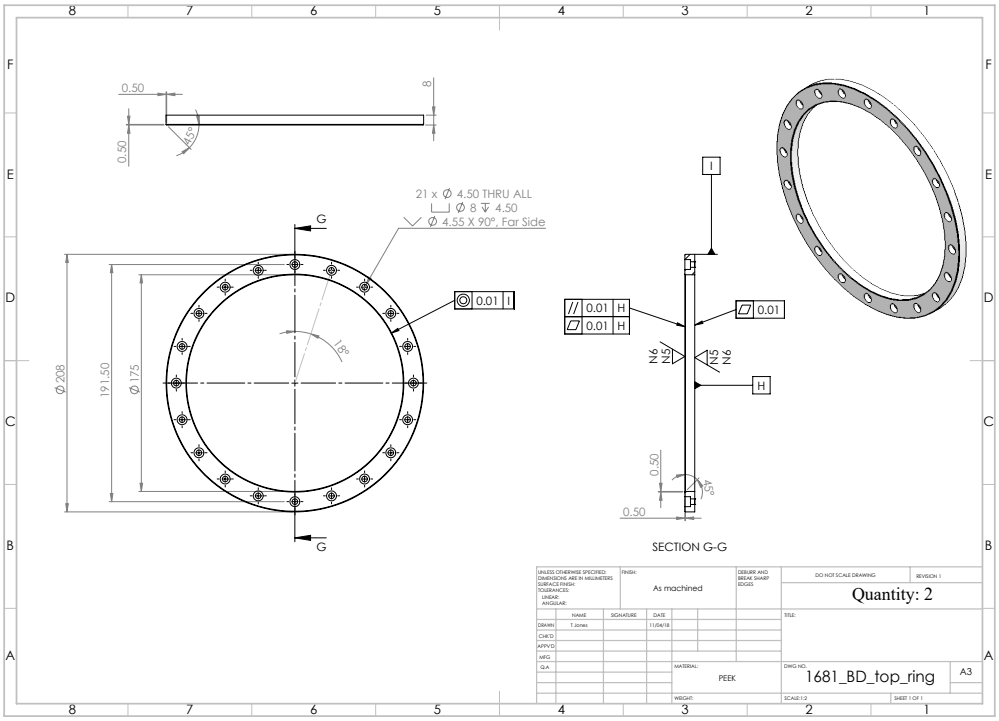


Figure D.4: Beamsplitter ring used as part of the beamsplitter assembly. Drawing provided by QMC Instruments.

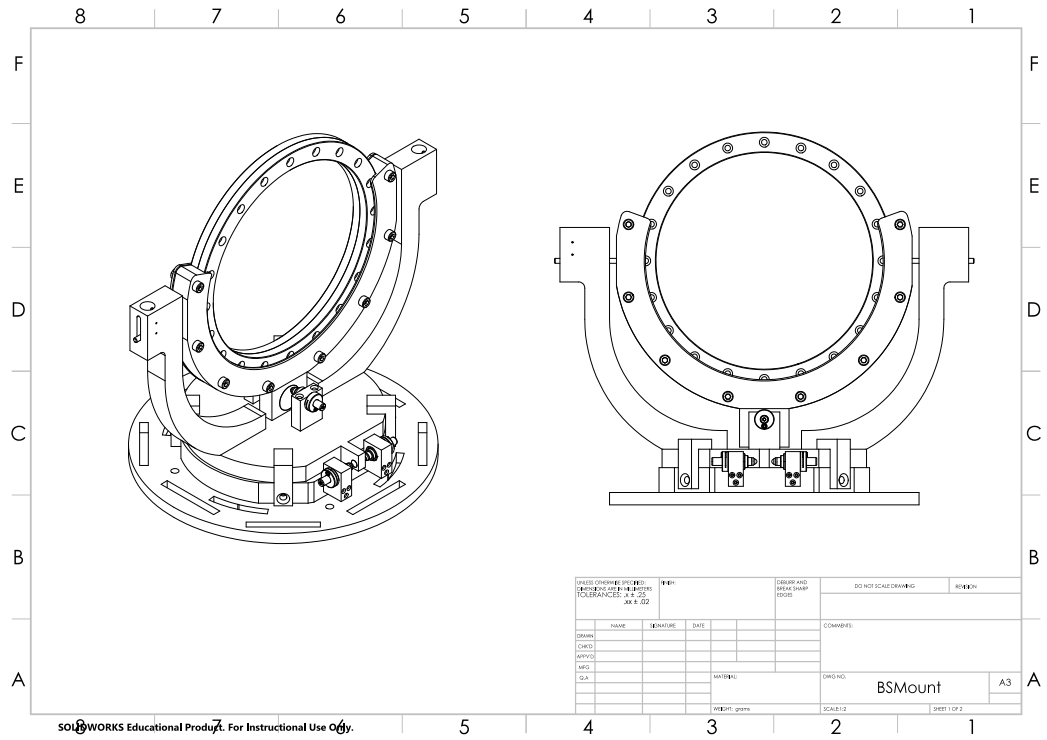


Figure D.5: Beamsplitter assembly. Drawing provided by Adam Sundberg.

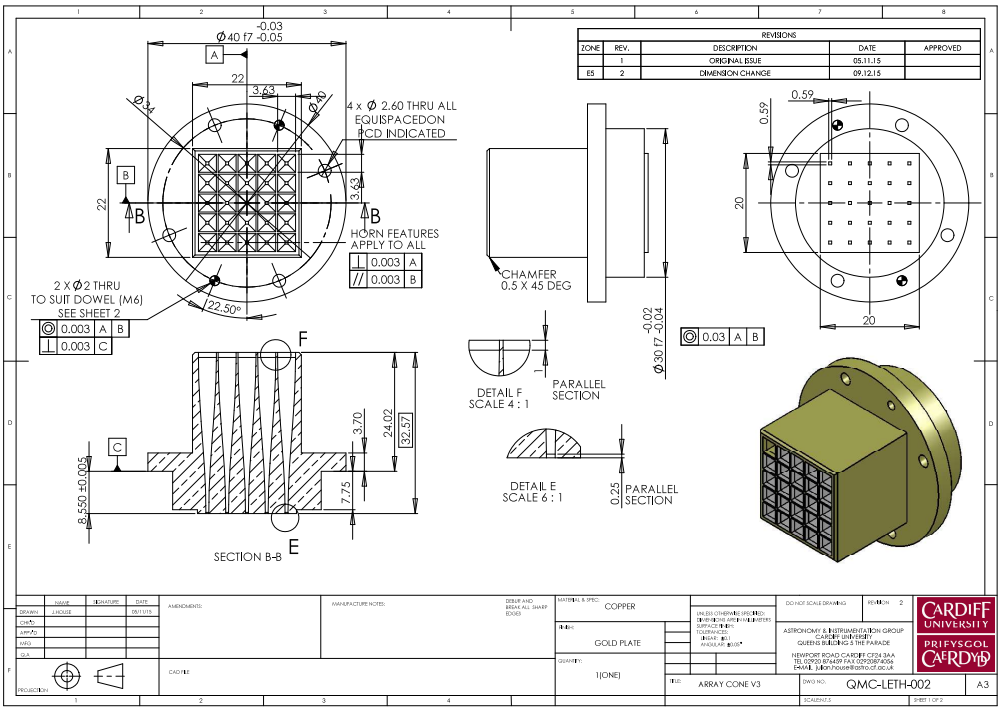


Figure D.6: Detector feed horn array. Drawing provided by QMC Instruments.

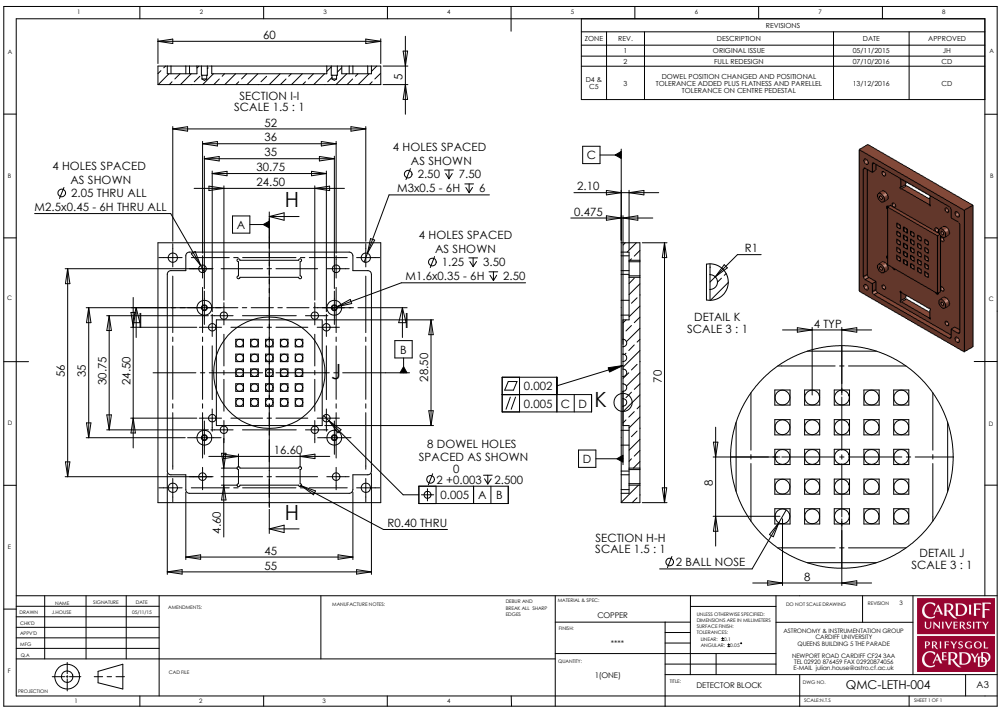


Figure D.7: Detector block. Drawing provided by QMC Instruments.