Fourier Transform Spectrometers

Astronomy 525

Lecture 12

Outline

Fourier Transform Spectrometers

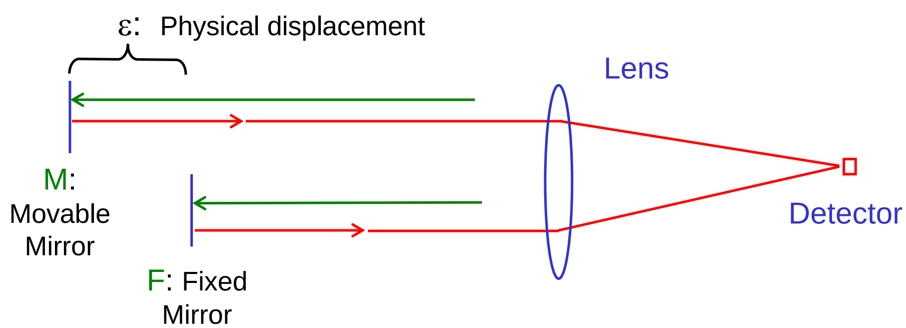
- Basic Principles
- The Interferogram
- Recovering the Spectrum
- The Fellgett Advantage
- Resolving Power
- Michelson Interferometer
- Lamellar Grating Spectrometer
- Apodization

- References: Harwit, M. and Sloan, N.J.A. 1979 Hadamard Transform Optics (New York: Academic Press); Stacey, G.J. 1985 PhD Thesis Cornell University

Basic Principles

- Fourier transform spectrometers interfere (typically) two beams of light by changing the phase delay between the beams.
- This *autocorrelation function*, often called *the interferogram* encodes the spectral properties of the beam, and can be Fourier transformed to achieve the power spectrum.
- Under certain conditions, an FTS can yield much improved signal-to-noise ratio over a mono-chromator due to the *Fellgett advantage*.
- We will show that the resolving power is simply the twice the path length of the interferogram, divided by the wavelength of interest.

The Interferogram: I



Simple case: Two beam interferometer (This is a simple form of a Lamellar grating interferometer)

Incoming beam, electric field amplitude, A
Mirror M moves a physical displacement, ε

The Interferogram: II

- The interference pattern for a monochromatic line (laser line) is a sine wave.
- For a bichromatic source, there will be beating between the two frequencies.
- In general, there will be a continuum of wavelengths, so that the interference pattern will be complex.
- However, for all cases, the "white light" fringe has unity amplitude, and the deepest trough will be just beyond the white light fringe.
- Note that for both cases, the average intensity observed is ½

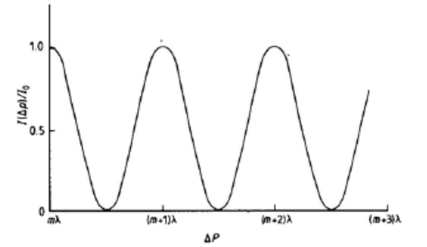


Figure 4.1.20 Variation of fringe intensity with mirror position in a Michelson interferometer.

Figure 4.1.21 Output of a Michelson interferometer observing a bichromatic source.

The Interferogram : III

The optical path delay is twice the physical displacement, so that the return beam lags by a phase $\phi = 4\pi\nu\varepsilon$, where ν is the frequency (wavenumbers, [cm^{-1}]).

At zero path, the return amplitude from M or F is equal to A/2. At arbitrary path difference, ε, the sum return amplitude is given by:

$$A(\varepsilon) = A/2 \cdot [1 + \exp(4\pi i\nu\varepsilon)]$$

So that after recombining the beams at the detector, the intensity is given by:

$$\eta(\varepsilon) = A(\varepsilon)^* A(\varepsilon) = A^2/2 \cdot [1 + \cos(4\pi\nu\varepsilon)]$$

The Interferogram : IV

Suppose the incident wave front is a superposition of many wavenumbers, ν, each with its own intensity. Then the total intensity in the frequency interval ν to $\nu + d\nu$ may be written $\psi(\nu)d\nu$ so that:

$$\eta(\varepsilon) = \tfrac{1}{2}\int_o^\infty \psi(\nu) \cdot [1 + \cos(4\pi\nu\varepsilon)]d\nu$$

$$= \tfrac{1}{2}\psi_o + \tfrac{1}{2}\int_o^\infty \psi(\nu) \cdot \cos(4\pi\nu\varepsilon)d\nu$$

Where $\psi_o \equiv \int_o^\infty \psi(\nu)d\nu$ is the intensity at zero path (constructive interference at all wavelengths). The *interferogram* consists of a modulated part superposed on a fixed unmodulated (D.C.) term. The physically interesting part is the modulated term:

$$\eta'(\varepsilon) = \tfrac{1}{2}\int_o^\infty \psi(\nu) \cdot \cos(4\pi\nu\varepsilon)d\nu$$

Since cosine and $\Psi(\nu)$ are symmetric:

$$\eta'(\varepsilon) = \tfrac{1}{4}\int_{-\infty}^\infty \psi(\nu) \cdot e^{(4\pi i\nu\varepsilon)}d\nu$$

The Interferogram : V

Or, defining $x \equiv 2\varepsilon$ we have:

$$\eta'(\varepsilon) = \tfrac{1}{4}\int_{-\infty}^\infty \psi(\nu) \cdot e^{(i2\pi\nu x)}d\nu$$

The integral is the Fourier transform of $\psi(\nu), \equiv \boldsymbol{\Psi}(x)$, so:

$$\eta'(\varepsilon) = \tfrac{1}{4}\boldsymbol{\Psi}(2\varepsilon) \qquad \text{or,}$$

$$\boldsymbol{\Psi}(2\varepsilon) = 4\eta'(\varepsilon)$$

So, the (spatial lag function) $\eta'(\varepsilon)$ at the detector is proportional to the Fourier transform of the incident waveform (frequency function) $\psi(\nu)$:

$$\int_{-\infty}^\infty \boldsymbol{\Psi}(2\varepsilon) \cdot e^{(-4\pi i\nu\varepsilon)}\,2d\varepsilon = 4\int_{-\infty}^\infty \eta'(\varepsilon) \cdot e^{(-4\pi i\nu\varepsilon)}2d\varepsilon$$

$$\Rightarrow \psi(\nu) = 8\int_{-\infty}^\infty \eta'(\varepsilon) \cdot e^{(-4\pi i\nu\varepsilon)}d\varepsilon$$

Or, again appealing to symmetry:

$$\boxed{\psi(\nu) = 16\int_o^\infty \eta'(\varepsilon) \cdot \cos(4\pi\nu\varepsilon)d\varepsilon}$$

Recovering the Spectrum: I

To precisely recover $\psi(\nu)$, we must sample $\eta(\varepsilon)$ from zero path difference to infinity continuously, and then perform the cosine transform of $\eta'(\varepsilon)$.

In practice, it is impossible to sample the entire interferogram. Instead, $\eta(\varepsilon)$ is sampled at n evenly spaced discrete points Δ apart along our total travel of $(n - 1)\Delta$. The step size, Δ is chosen so that:

$$\Delta = 1/[4(\nu_{max} - \nu_{min})$$

where ν_{max} is the highest frequency to be sampled and ν_{min} is the lowest sampled frequency

Usually, $\nu_{min} = 0$ so that the step size relationship reduces to:

$$\Delta = 1/(4\nu_{max})$$

This, you will recognize as an expression of the Nyquist sampling theorem if your recall that the physical path difference, ε and Δ are half the optical path difference.

Recovering the Spectrum: II

The interferogram, thus discretely sampled, may be transformed to form an estimate, $\hat{\psi}(\nu)$, of the true spectrum $\psi(\nu)$:

$$\hat{\psi}_1(\nu) = 16 \sum_{r=0}^{n-1} \eta'(r\Delta) \cos(4\pi r\Delta\nu)\cdot\Delta$$

We wish to evaluate $\psi(\nu)$, at the n discrete frequencies ν_s given by:

$$\nu_s = \nu_{min} + s \cdot\delta$$

for s = 0, 1, 2, 3,…, n-1 within the spectral bandwidth δ:

$$\delta \equiv \{(\nu_{max} - \nu_{min})/((n-1)\} = 1/\{4(n-1)\Delta\}$$

[bin width in cm^{-1}]

The estimator of the spectral power $\hat{\psi}(\nu_s)$, in the frequency range $\nu_{min} \leq \nu_s \leq \nu_{max}$ is then:

$$\hat{\psi}(\nu_s) = \hat{\psi}_1(\nu_s)\,\delta = \hat{\psi}_1(\nu_s)/[4(n-1)\Delta],$$

[total power in a spectral bin; *the observable*]

or:

$$\hat{\psi}(\nu_s) = 4/(n-1) \cdot \sum_{r=0}^{n-1} \eta'(r\Delta) \cos(4\pi r\Delta\nu_s)$$

Recovering the Spectrum: III

We may formally write this as the matrix equation:

$$\hat{\psi} = A \cdot \eta$$

Where:

$$\hat{\psi} = \begin{bmatrix} \hat{\psi}(\nu_{min}) \\ \hat{\psi}(\nu_{min} + \delta) \\ \hat{\psi}(\nu_{min} + 2\delta) \\ \vdots \\ \hat{\psi}(\nu_{max}) \end{bmatrix} \qquad \eta = \begin{bmatrix} \eta(0) \\ \eta(\Delta) \\ \eta(2\Delta) \\ \vdots \\ \eta((n-1)\Delta) \end{bmatrix}$$

and A is an $n \times m$ matrix whose elements a_{sr} are given by:

$$a_{sr} = 4/(n-1) \cdot \cos[4\pi r\Delta(\nu_{min} + s\cdot\delta)] \text{ for } s,r = 0, 1, \ldots n-1.$$

The Fellgett Advantage: I

The factor 4/(n-1) in: $\hat{\psi}(\nu_s) = 4/(n-1) \cdot \sum \eta'(r\Delta) \cos(4\pi r\Delta\nu_s)$, is very important as it leads to the Fellgett (or multiplex) advantage. This advantage is realized if:

1. The noise in the system is independent of the photon flux
2. The noise of each measurement e_i has zero mean: $\langle e_i \rangle = 0$. Where the brackets indicate the time averaged expectation value
3. The noise between measurements is uncorrelated, i.e , $\langle e_i \cdot e_j \rangle = 0$
4. The noise has a variance, defined by $\langle (e_j)^2 \rangle = \sigma^2$.

Under these circumstances, the expectation value of the estimator, $\hat{\psi}$ is given by: $\langle \hat{\psi} \rangle = \psi(\nu)$ So that:

$$\hat{\psi} - \psi = A \cdot e \Rightarrow \hat{\psi}(\nu_s) - \psi(\nu_s) = a_{s,0}e_0 + a_{s,1}e_1 \ldots a_{s,n-1}e_{n-1}$$

(e is the noise column vector)

The Fellgett Advantage: II

If we define, $e(\nu_s)$, as the variance of the spectral estimator for frequency ν_s:

$$e(\nu_s) \equiv \langle[\hat{\psi}(\nu_s) - \psi(\nu_s)]^2\rangle = \sigma^2(a^2_{s,0} + a^2_{s,1} \dots a^2_{s,n-1}) = \sigma^2 \sum_r a^2_{s,r}$$

The average mean square error, e, is then given by:

$$e \equiv (1/n) \cdot \sum_s e(\nu_s) = \sigma^2/n \cdot \sum_s \sum_r a^2_{s,r}$$

$$= 16\sigma^2/[n(n-1)^2] \cdot \sum_s \sum_r \cos^2[4\pi r\Delta(\nu_{min} + s\cdot\delta)]$$

Half-angle formula

$$= 8n\sigma^2/(n-1)^2 + 8\sigma^2/[n(n-1)^2] \cdot \sum_s \sum_r \cos[8\pi r\Delta(\nu_{min} + s\cdot\delta)]$$

$$\Rightarrow e \sim 8\,\sigma^2/n \text{ for large } n$$

If we measure a variance, σ^2 for each sample $\eta(\varepsilon)$ then we will measure a variance in the transform spectrum that is $(8/n)^{1/2}$ smaller. This Fellgett advantage is quite powerful for detector noise limited systems.

The Fellgett Advantage: III

Intuitive Understanding of the Fellgett Advantage

Lets compare an FTS with a Fabry-Perot interferometer (a monochrometer)

The FTS samples the whole spectrum, consisting of m independent resolution elements at once. A F-P must scan over these m resolution elements, so it takes *m times longer to obtain a spectrum*.

For the FTS, we must sample 2m times (Nyquist). Also, on average, half the signal is thrown away in the encoding process, since the average value of the autocorrelation function is ½:

$$\eta(\varepsilon) = \tfrac{1}{2}\,\psi_o + \tfrac{1}{2}\int_o^\infty \psi(\nu) \cdot \cos(4\pi\nu\varepsilon)d\nu$$

The Fellgett Advantage: IV

In a given total integration time, t, then the FTS system will yield a SNR ratio of:

$$SNR_{FTS} = S/(2\sigma) \cdot (t/2)^{1/2}$$

where S = signal per resolution element

σ = noise (assumed independent of signal or background)

Every spectral element is sampled ½ the time so total time/element is $t_{element} = t/(2m)*m = t/2$

For the FP system:

$$SNR_{FP} = S/\sigma \cdot (t/m)^{1/2}$$

Every spectral element gets $t_{element} = t/m$ integration time.

Therefore, the ratio is:

$$SNR_{FTS}/SNR_{FP} = (m/8)^{1/2}$$

Notice that this hold if and only if we are not background limited!! If you are background limited, the Fellgett advantage vanishes. In fact, you then *loose* due to encoding loses etc.

The Fellgett Advantage: An Example

In the early stages of the Cornell airborne FTS system, we were detector noise limited. We did "see" a Fellgett advantage. Later on we were able to clean up our electronics to the point that this advantage vanished! (good news, and bad news…)

FTS systems are typically less useful in background limited situations.

1. Often we do not care to sample the spectrum over a broad range ⇒ you can win rapidly with a monochrometer
2. Fourier transform spectra (interferograms) take time. Your system must be stable over this time, since you are always encoding the entire spectrum. Practical difficulties, such as changing airmass, loss of the source, or changing system gain can therefore be a disaster for an FTS.
3. However, in space, one is often not background limited, so an FTS might well be an excellent choice, e.g. the SPIRE FTS on the Herschel Telescope.

Resolving Power: I

The maximum resolving power of an FTS system is determined by the longest optical path length.:

$$R = 2\Delta(n-1)/\lambda$$

which is entirely analogous to the grating situation.

If we don't know where zero path (the white light fringe) is, then we must sample our interferogram at both positive and negative paths, and perform both a sine and cosine transform.

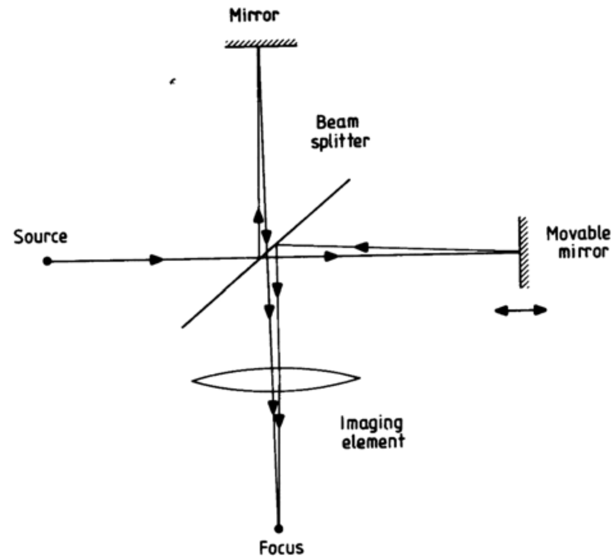
Resolving Power: II

However, if one know a priori that the interferogram has begun at zero path, then the sine transform will vanish as the interferogram must be symmetric: $\eta(\varepsilon) = \eta(-\varepsilon)$. We may then fold the interferogram about zero path to mathematically obtain the two sided interferogram.

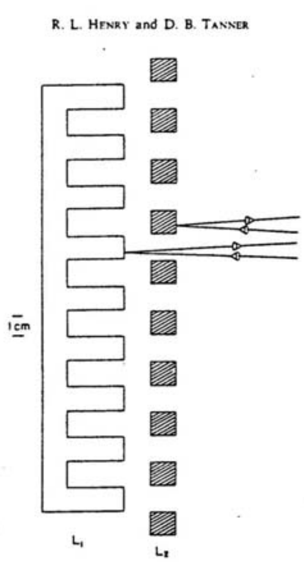
Zero path is found by measuring an interference fringe at two separate wavelengths. If they coincide, this must be the zero path (or the 2 λ's are integer multiples of each other). To eliminate this possibility, one measures at a third wavelength.

Michelson Interferometer

The most common form of FTS is the 2-beam Michelson interferometer. It features rather good mixing of the beams, but has has a serious disadvantage in that 50% of the light is reflected back out the entrance window.

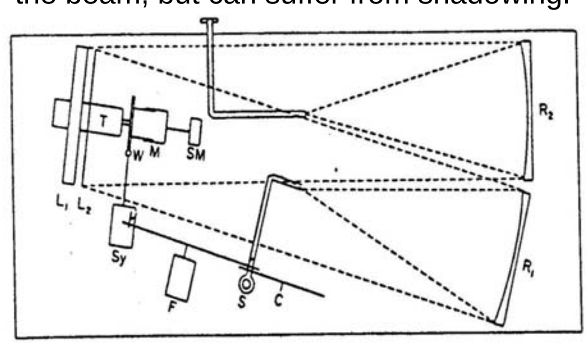


The Lamellar Grating Spectrometer: I



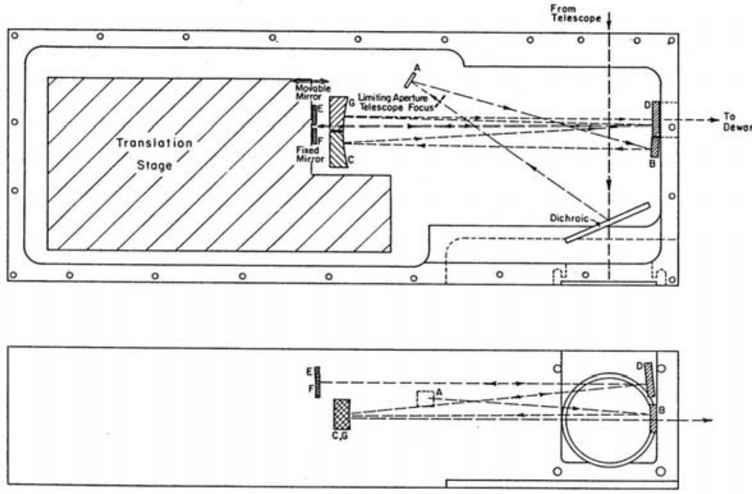
A lamellar grating spectrometer is another version of an FTS.

Lamellar grating spectrometers fully utilization the beam, but can suffer from shadowing.



. 2. Plan of the interferometer. The important parts are labeled thus: S: source, C: chopper, filter wheel with selsyn drive, R_1: collimating mirror, L_1: fixed and L_2: moveable grating, focusing mirror, T: slide, M: micrometer, SM, stepping motor, W worm gear, and Sy: selsyn drive for symmetrizer. The entrance and exit light-pipes are shown.

The Lamellar Grating Spectrometer: II



The Cornell lamellar grating spectrometer used on the KAO. Just two mirrors reduces complexity. Maximum path length of 5 cm, for a resolving power of 0.1 cm^{-1} operating from 60 to 200 μm (166 to 50 cm^{-1}). Latter incarnation employed a 15 cm path stage.

The Lamellar Grating Spectrometer: III

The Nyquist Theorem requires that we sample an interferogram at a rate equal to twice the highest unaliased frequency of interest.

In other words, we must sample $\eta(\varepsilon)$ every $1/(4\nu_{max})$ cm. For the Cornell FTS, at a path difference of 15 cm at 150 μm, we would need to sample 4000 times.

Problems:

1. Lots of samples - very inefficient, little time per point, and lots of "move" time – during a 1 hour KAO flight leg, we might expect to obtain only 1 spectrum and it would have ~ 1/2 second integration time per point. This could be disasterous when we "loose track" etc.
2. Lots of samples to transform – no longer a problem

The Lamellar Grating Spectrometer: IV

To circumvent these problems, *and especially to lower the thermal background*, this system employed a moderate resolution grating spectrometer to act as an order sorter for the FTS.

The grating order sorter had a bandpass of ~ 0.9 cm^{-1} wide at 150 μm ($\Delta\nu_{spect}$ ~ 0.9 cm^{-1}). Therefore, we need only sample fast enough to ensure that we sample all of the frequencies between ν_{min} and ν_{max} where $\nu_{min} < \nu_{spec} < \nu_{max}$ This is ensured by choosing the step size, Δ such that:

$$\Delta = 1/(4(\nu_{max} - \nu_{min})$$

For $\nu_{max} - \nu_{min} = \Delta\nu_{spec}$ the required step size is lessened to $\Delta = 0.278$ cm, so that to sample the full 15 cm path requires only 54 samples. We would use 64 samples, to easily implement a fast Fourier transform (FFT).

The Lamellar Grating Spectrometer: V

Practical Problems

There were three problems in implementing this design that reduced sensitivity.

1. Loss of coherence at long path.
2. Imperfect interference at zero path

   For a sufficiently narrow bandpass, at zero path, the beams should recombine to form a sinusoidally varying interferogram of maximum intensity P, and minimum intensity, V (= 0). This is the ideal case, where efficiency $E \equiv (P-V)/P = 1$. For laboratory sources, we obtained E = 83% (P:V – 6:1)
3. "Ringing" in the Fourier transform due to finite lag path.

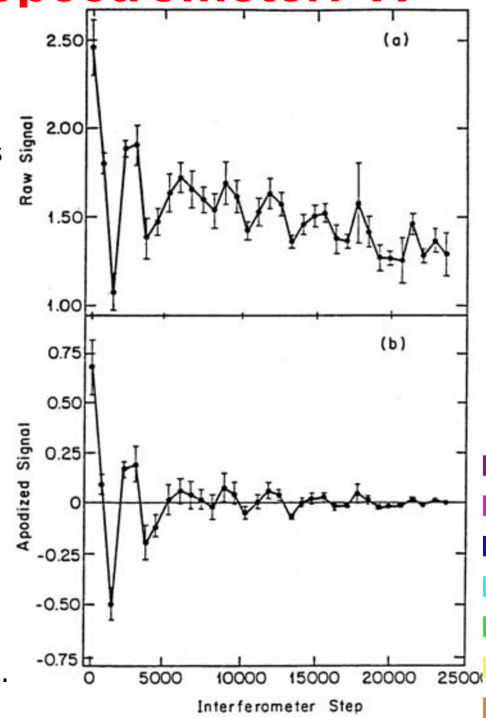
The Lamellar Grating Spectrometer: VI

1. Loss of coherence at long path.

If the beam is not perfectly collimated, and/or the path of the translation stage is not totally parallel to the beams, the beams will recombine poorly at large paths. (Assuming that we optimize at short paths.) This then leads to a fall off in the power: "droopy shoulders".

This problem is mathematically ameliorated by fitting, and subtracting a linear slope to the data

The problem was eliminated entirely through the introduction of corner cube mirrors in the final incarnation of the FTS.

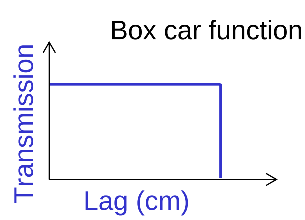


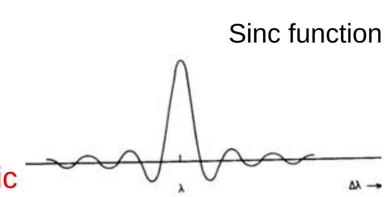
Apodization: I

3. Ringing in the Fourier transform due to finite lag path.

This ringing is a natural artifact of the finite lag path. The finite lag path is identical to multiplying the interferogram by a box-car function.

The resulting transform is the convolution of the transforms of the the source intensity with the box-car function. (recall that the FT of a product of two functions is the convolution of the FT of each function).





The transform of the box-car function is the sinc function: sin(x)/x. Therefore, for a laser line, whose FT is a delta function, we expect the basic instrumental profile to be the sinc function.

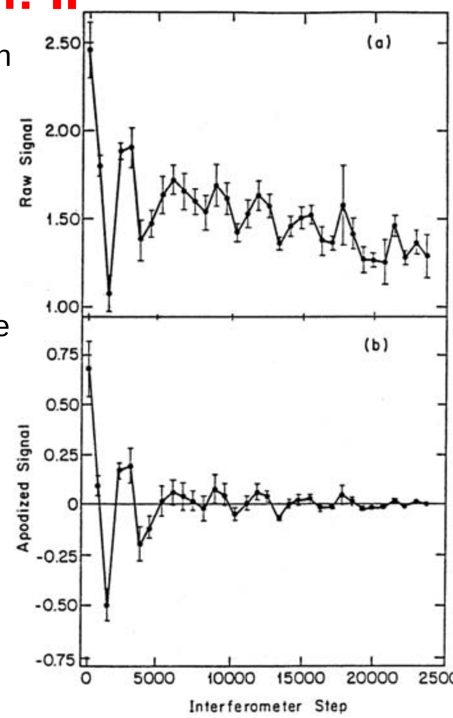
Apodization: II

To minimize this “ringing”, a weighting function is often applied to the interferogram to bring the intensity down continuously to zero a the end.

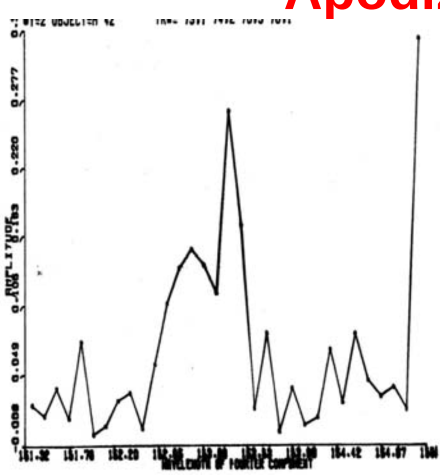
This process is called “apodization”.

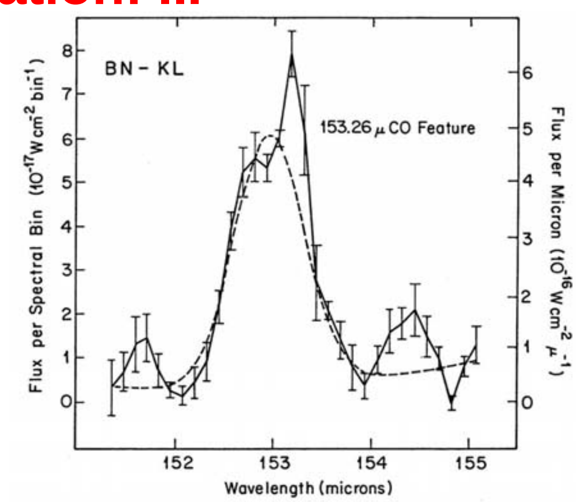
It will necessarily degrade the resolution of the spectrum (since the highest frequencies are given less weight), but greatly reduces artifacts in the spectrum.

There are several forms to the apodization window. For the interferogram at left, we applied a simple triangle (Bartlett) window (after removing the droopy shoulders!)



Apodization: III





Apodization with a triangle function.

(left) The FT of the raw spectrum on the previous viewgraph.

(right) The FT of the apodized (and droopy shoulder removed) spectrum.

Figure 2.4. Fourier transform spectrum of the BN-KL region of Orion. This spectrum is the Fourier transform of Figure 2.3b, following reflection about zero path. The fitted Gaussian represents the grating instrumental profile. The feature at 153.3μm is the first astronomical detection of the J = 17→16 rotational transition of shocked ^{12}CO (Stacey et al. 1982).

Apodization: IV

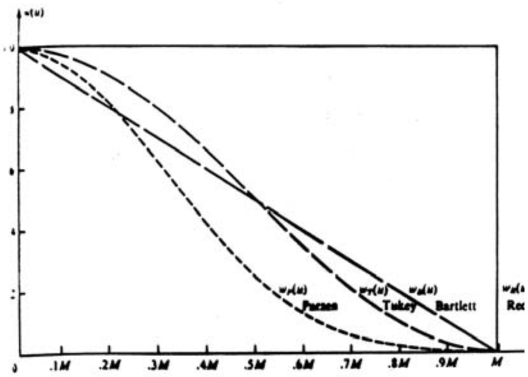
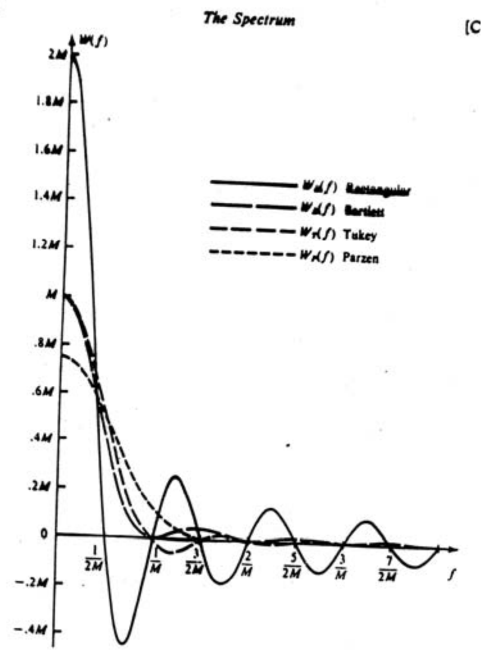


FIG. 6.12: Some common lag windows



Other common apodization functions (left) and the resulting instrument profiles (right)

Apodization: V

TABLE 6.5: Lag and spectral windows

Description	Lag window	Spectral window
rectangular	$w_R(u) = \begin{cases} 1, & \lvert u\rvert \leq M \\ 0, & \lvert u\rvert > M \end{cases}$	$W_R(f) = 2M\left(\frac{\sin 2\pi f M}{2\pi f M}\right), \quad -\infty \leq f \leq \infty$
Bartlett	$w_B(u) = \begin{cases} 1 - \frac{\lvert u\rvert}{M}, & \lvert u\rvert \leq M \\ 0, & \lvert u\rvert > M \end{cases}$	$W_B(f) = M\left(\frac{\sin \pi f M}{\pi f M}\right)^2, \quad -\infty \leq f \leq \infty$
Tukey	$w_T(u) = \begin{cases} \frac{1}{2}\left(1 + \cos\frac{\pi u}{M}\right), & \lvert u\rvert \leq M \\ 0, & \lvert u\rvert > M \end{cases}$	$W_T(f) = M\left\{\frac{\sin 2\pi f M}{2\pi f M} + \frac{1}{2}\frac{\sin 2\pi M(f + \frac{1}{2}M)}{2\pi M(f + \frac{1}{2}M)} + \frac{1}{2}\frac{\sin 2\pi M(f - \frac{1}{2}M)}{2\pi M(f - \frac{1}{2}M)}\right\} = M\left(\frac{\sin 2\pi f M}{2\pi f M}\right)\left(\frac{1}{1 - (2fM)^2}\right), \quad -\infty \leq f \leq \infty$
Parzen	$w_P(u) = \begin{cases} 1 - 6\left(\frac{u}{M}\right)^2 + 6\left(\frac{\lvert u\rvert}{M}\right)^3, & \lvert u\rvert \leq \frac{M}{2} \\ 2\left(1 - \frac{\lvert u\rvert}{M}\right)^3, & \frac{M}{2} < \lvert u\rvert \leq M \\ 0, & \lvert u\rvert > M \end{cases}$	$W_P(f) = \frac{3}{4}M\left(\frac{\sin \pi f M/2}{\pi f M/2}\right)^4, \quad -\infty \leq f \leq \infty$

Progress on Cornell FT Spectrometer

Progress with resolving power and sensitivity with the Cornell FTS on the KAO.

1. First detection of CO 16 → 15 rotational transition from Orion BN-KL
2. First detection of OH rotational lines near the CO line – this required higher sensitivity
3. Clearly resolved OH lines: longer translation stage (15 cm), hence higher resolving power.