Back-ends

J. R. Fisher

NRAO, P. O. Box 2, Green Bank, WV 24944, rfisher@nrao.edu

Abstract.

1. Introduction

Radio telescopes have traditionally been divided into three subsystems: the antenna, R.F. and I.F. electronics (front-end), and the "back-end." The first two are common to all types of observing, but a back-end is usually specific to continuum, spectroscopic, or pulsar measurements. Signal processing and astronomical information extraction that are unique to a particular observing type typically reside in the back-end.

Nearly all astronomical observations can be summaried as the measurement of radiation intensity as a function of five independent variables: two position coordinates (e.g. right ascension and declination), time, frequency, and polarization. All other parameters (distance, pulsar period, radial velocity, etc.) are derived quantities. The back-end is directly involved in the measurement of intensity as a function of time, frequency and polarization, and it is also responsible for the coordination of calibration and other switching signals with data integrators.

2. System Overview

Figure 1 shows the relationship of the back-end to the rest of the telescope system. The association of intensity with sky coordinates is usually done between the telescope control system and the data recording computers and, hence, does not directly involve the back-end.

A very simplified back-end is drawn in Figure 1 to show only how the signal intensity accumulators are synchronized with the calibration and front-end switches. A reference load switch is shown, but it could also be a beam switch, polarization switch, or something else. In this illustration the four accumulators could be sequenced as shown in Table 1. The details of the sampler and processor block will be discussed in later sections. We assume here only that the output of this block is an intensity value proportional to the system noise power plus power from the radio source in the antenna beam. The accumulators are set to integrate the signal for a chosen amount of time, often through many switch cycles, and then their values are recorded with the antenna position at the center of the integration in the output data file. The data file usually contains many
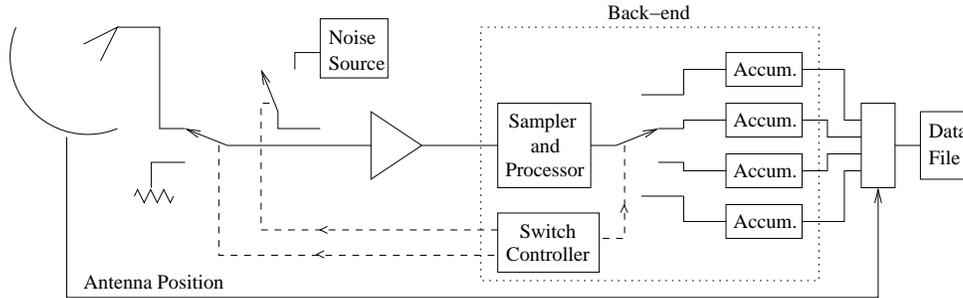
Figure 1.   The back-end's place in the telescope system.

successive accumulator integrations. Interpretation of the accumulator values is then the task of the data analysis software.

Table 1.   Example accumulator, calibrator, and front-end switching sequence in Figure 1.

Active Accumulator	Calibration Noise Source	Load Switch
First	On	Antenna
Second	Off	Antenna
Third	On	Load
Fourth	Off	Load

Some front-ends and some types of observing do not use a reference switch so the switching cycle would use only two accumulators for calibration noise source on and off. If the noise source were left on or off for the duration of a scan, only one accumulator would be used. Any precise synchronization that is required between the back-end and other parts of the system is usually driven by the timing circuitry in the back-end.

Since many electronic switches take a finite time to change states or may generate transient noise during the state change, it is best to avoid accumulating data during a switch transition. Most back-ends provide for a "blanking" interval at the beginning of each switch transition when all accumulators are inactive. The length of the blanking interval depends on the type of switch being used, but it is typically on the order of tens of microseconds to a few milliseconds. If the switch involves a physical movement with significant inertia, such as a mirror or reference load motion, the switching command to the front-end might be sent slightly in advance of the accumulator switch change to maximize the useful accumulation time on each switch state and minimize to required blanking time.

3.   Samplers and Detectors

The output of the R.F./I.F. section of the receiver is a band-limited voltage. By that we mean that the receiver noise and cosmic signals have been restricted in frequency by bandpass or lowpass filters and the output of the last amplifier is a voltage proportional to the amplitude (not intensity) of the combined signals. The average value of this voltage is zero. To measure the power of the receiver output this voltage must be squared. The average value of this squared signal is not zero, and it may be accumulated for long periods of time to obtain a better estimate of its average value.

Because all cosmic signals are noise-like with no *a priori* knowledge of their emitted phases, the only measure of these signals available to astronomers is the temporal and spatial nature of their intensity or power. However, there are many signal processing operations that must be performed on the voltage output from the front-end, such as further band limiting, spectral decomposition, and correlation with other copies of the same signal, before the final squaring is done to measure the power. We will discuss a few of these signal processing operations in later sections. The important point here is to keep in mind the difference between the voltage and power signal domains. Voltages can add coherently. Powers add only statistically.

3.1.   Square-law detector

The simplest telescope output is the direct squaring of the voltage from the front-end. This constitutes what is often called a continuum receiver. A natural power detector is a calorimetric device where the signal is used to heat a resistor, and the power is proportional to the resistor's temperature rise. This is the principle used by a bolometer. Unfortunately, the response time of such a device is too slow for many radio telescope measurements. The most frequently used alternative is a semiconductor diode whose exponential voltage-current transfer characteristic is used to approximate a square-law detector over a limited signal power range. Its output voltage is proportional to the square of its input voltage, and the output voltage may be further processed as a power quantity.

Figure 2 shows three methods for averaging or accumulating the output of a square-law detector. The first method uses a lowpass filter on the output of the detector. The effective integration time is approximately equal to the inverse of the bandwidth of this filter ($\tau \approx (1/B)$) and the analog-to-digital converter is expected to sample the filter output slightly faster than this interval. Each integration-time selection requires a different filter and sampling time.

The second method in Figure 2 uses just one lowpass filter and A/D sampling rate consistent with the shortest integration time required. Longer integrations are performed by accumulating the samples in an arithmetic logic unit. This requires considerably less hardware for a wide selection of integration times. One disadvantage of both this and especially the first method is that for wide pre-detection filter bandwidths the A/D converter must have sufficient amplitude resolution to resolve the lowpass filter output noise on top of a large DC offset plus adequate dynamic range for signals stronger than the system noise. A/D output words at least as wide as 16-bits may be required, which can be rather expensive.

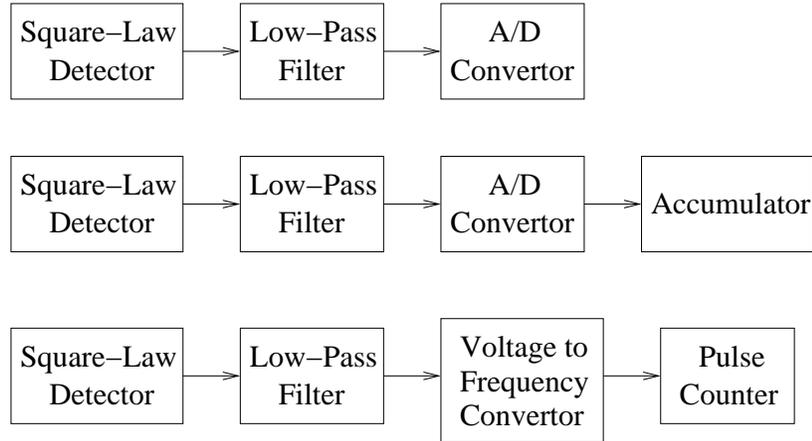Power Detection and Accumulation

Figure 2.   Three methods for power averaging following a square-law detector.

A third method of square-law detector output accumulation is shown at the bottom of Figure 2. The low-pass filter output drives a voltage-to-frequency converter, running at a nominal frequency of about one megahertz, which in turn drives a pulse counter. The pulse counting interval is equal to the chosen integration time. This scheme has the property that the amplitude resolution is proportional to the integration time. Long integration times and wide pre-detection bandwidths are easily accommodated. For a 1 MHz pulse rate and an integration time of 0.1 seconds the amplitude resolution is one part in 10^5, which is equivalent to a 20-bit A/D, assuming a factor of ten headroom for strong signals.

3.2.   Voltage sampling

To retain the spectral and possibly the phase information in the signal from the front-end for further processing, the voltage signal must be sampled. This is done with a fast A/D converter connected directly to the front-end amplifier output. To prevent aliasing of the spectral information the A/D sampling rate must be twice the bandwidth of the RF signal, i.e. a 100 MHz bandwidth signal must be sampled at 200 megasamples per second. The first stages of digital processing following the A/D must handle the same data rate.

To construct the widest bandwidth digital spectrometers and signal cor-relators at an affordable cost most machines have taken advantage of the fact that the signal from a radio astronomy receiver is almost always dominated by system noise with the astronomical signal adding a only small fraction to the to-tal. This permits the use of an extremely simple sampling device and subsequent processing arithmetic with only one or two bits of input voltage resolution. Such a sampler would severely distort a non-noise-like signal, producing many har-
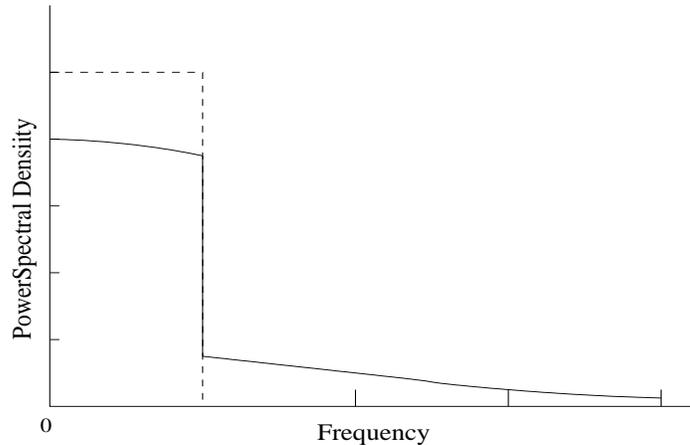
Figure 3.   Output spectrum of a one-bit (two-level) sampler with a band-limited, white noise input whose spectrum is shown by the dashed line. This figure is adapted from Figure 8.3 of Thompson, Moran, and Swenson (2001).

monics and cross-products of the frequencies present in the spectrum. However, when random noise is the dominant signal component, all of the harmonics and cross-products are converted into broadband noise that can usually be tolerated in an astronomical measurement.

Figure 3 shows the spectral content of the output of a one-bit (two-level) voltage sampler. The dashed line shows the band-limited input noise spectrum. The power lost from the input noise frequency range is due to harmonics and cross-products being generated by the sampler quantization, as can be seen by the long spectral tail extending to higher frequencies. This spectral tail will be aliased back into the input frequency range, if the sampler is clocked at twice the bandwidth of the input spectrum. The harmonic power loss and aliasing results in the sensitivity of a one-bit sampler being only 64% ($2/\pi$) of a perfect sampler. A faster sampling clock frequency will reduce the aliased noise power and improve the sensitivity somewhat. Using more sampling levels also helps. Most spectrometers and correlators in use today use two-bit (three-level) samplers as a compromise between sensitivity loss and hardware complexity.

As you might surmise, a one-bit sampler loses all amplitude information. This must be restored by measuring the total spectral noise power with a separate square-law detector to renormalize the sampled spectrum. A two-bit sampler retains the amplitude information in the statistics of the three sampler level counts over a limited amplitude range, so a separate power detector is not necessarily required. However, for best sensitivity and accurate spectral power recovery the input level to a two-bit sampler must be well controlled and carefully monitored.

At the time that this article is being written we are entering a transition to multi-bit samplers that will no doubt result in many receiver functions currently implemented with analog electronics (mixers, local oscillators, filters, etc.) being taken over by digital signal processors. The enormous increase in digital com-

putational speed, financed by the personal computer revolution, makes this all possible at frequencies and bandwidths of interest to radio astronomy. Multi-bit samplers increase sensitivity and amplitude dynamic range, but they also allow higher performance filters and much more stable signal processing in the digital domain.

4.   Spectrometers and the Fourier Transform

Nearly all spectrometers in radio astronomy use the Fourier transform principle of spectral decomposition. They can have one of two forms, either a direct Fourier transform of the temporal series of voltage samples, $V(t)$:

$$A(f) = \int V(t)e^{i2\pi ft}dt, \tag{1}$$

$$P(f) = A^2(f), \tag{2}$$

or a transform of the autocorrelation function, $ACF(\Delta t)$,

$$P(f) = \int ACF(\Delta t)e^{i2\pi f\Delta t}d\Delta t, \tag{3}$$

where $A(f)$ is the voltage frequency spectrum of $V(t)$, $P(f)$ is the power spectrum, and Δt is the autocorrelation function delay. The relationship between the voltage samples and the autocorrelation function is

$$ACF(\Delta t) = \int V(t)V(t + \Delta t)dt. \tag{4}$$

The advantage of an autocorrelation spectrometer is that the relatively complex computation of the Fourier transform needs to be performed only after that data have been averaged for a considerable length of time as is implied by the integral in Equation 4. The direct Fourier transform of the voltage samples must be done at the rate dictated by the sampling speed. For example, if we wanted to compute a 10-MHz, 1024-point spectrum averaged over one second, the direct transform method requires 10,000 transforms while the autocorrelation method requires only one. The one transform per second can be done in a general purpose computer while the much faster transform must be implemented in more specialized hardware. For one- and two-bit samplers the hardware implementation of an autocorrelator is relatively inexpensive.

The hardware block diagram of an autocorrelator is shown in Figure 4. With a one-bit sampler the Δt delay chain is a simple binary shift register, the multipliers are binary 'exclusive or' gates, and the accumulators are simple counters. The linear architecture of the autocorrelator also fits well into a hardware design. A two-bit sampler requires somewhat more hardware complexity, but only by a factor of three or four. A multi-bit sampler requires a multi-bit shift register, an arithmetic logic unit for each multiplier and accumulator and multi-bit memory associated with each accumulator.

The direct Fourier transform spectrometer is inherently a multi-bit arithmetic device because the complex exponential term in Equation 1 must be represented with sufficient precision, typically 12 to 16 bits, to prevent the generation
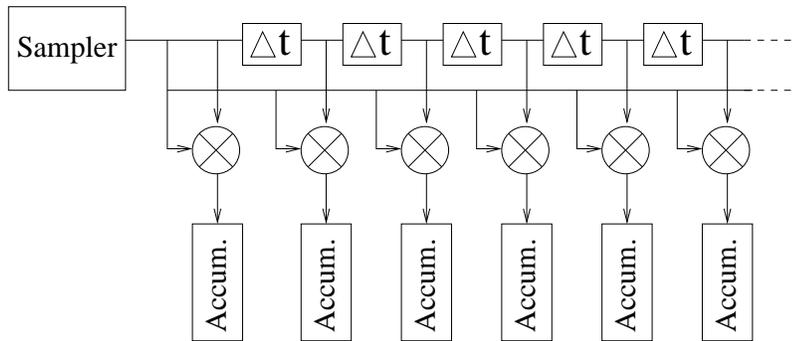
Figure 4.   Hardware block diagram of an autocorrelator.

of spurious products due to phase and amplitude errors in the computational process. Small input word lengths offer no advantage. However, the Fourier transform, in its fast fourier transform (FFT) implementation (Cooley & Tukey, 1965), does have an inherent computational advantage over the autocorrelation calculation. The number of arithmetic operations per computed spectrum in an FFT is proportional to $NLog_2(N)$ while an autocorrelator requires N^2 operations, where N is the number of points in the spectrum. For large N this is a considerable advantage.

For the same economic reasons that mixers, filters, and other traditionally analog functions are now possible in digital hardware, the direct FFT spectrometer is now feasible with bandwidths approaching a GHz. This will very likely increase the role of FFT devices in future back-end designs. The details of a hardware version of an FFT are somewhat more than can be covered in this article. For more information see one of a number of digital signal processing texts, such as Rabiner & Gold (1975).

Notice in Equation 1 that the output of a direct Fourier transform spectrometer, $A(f)$, is a complex voltage spectrum as opposed to a real-valued power spectrum, $P(f)$, from the autocorrelation function in Equation 3. Each frequency channel of the voltage spectrum is exactly like a voltage sample of a much narrower passband, and it may be used in exactly the same way for further signal processing. All phase information is lost in the power spectrum so, in many cases, the autocorrelation operation is necessarily the final step in any real-time signal processing chain.

Alert readers will notice that I have casually glossed over the difference between real and complex voltage samples when dealing the the Fourier transform spectrometer. This is a key issue that must be addressed in the FFT design, but it is not a fundamental limitation. One can build complex (sine and cosine) voltage samplers or add a stage to the FFT computation to account for the fact that real samples have been fed to the inherently complex FFT operation.

5.   Signal Correlation

Any radio signal, even random noise, is always correlated with itself. This fact is exploited in the autocorrelation spectrometer. Two copies of a signal may travel through different antenna and receiver electronics, such as two orthogonal polarizations or two adjacent focal plane antenna elements. These two copies are also correlated.

When we say that signals are correlated we mean that their time-averaged vector product is not zero. Mathematically, this is the average of one signal times the complex conjugate of the other.

$$C = \langle V_1(t) V_2^*(t) \rangle \tag{5}$$

Since $V_1(t)$ and $V_2(t)$ are copies of the same signal,

$$V_1(t) = aV e^{i2\pi ft}, V_2(t) = bV e^{i2\pi ft + \phi}, \tag{6}$$

and

$$C = abV^2 e^{i\phi}. \tag{7}$$

The correlation product, C, is a vector with units of power and a phase equal to the phase difference between the two copies of the signal.

A good example of the utility of this correlation value in single dish work is for the measurement of polarization. The self and cross (correlation) products of the signals from two orthogonal feed polarization outputs contain all there is to know about a polarized radio source. A few specific cases of fully polarized radiation using orthogonal linear feed outputs are

    Linear pol'n aligned with 'b' output........ $a = 0.000, b = 1.000, \phi = n/a$
    Linear pol'n at 45 degrees to 'a' and 'b'.. $a = 0.707, b = 0.707, \phi = 0$
    Circular pol'n... $a = 0.707, b = 0.707, \phi = 90$

These values are normalized to the total signal power and phase referenced to the feed output. The receiver electronics will modify the relative phase and amplitude of the two signal paths, and these receiver effects need to be calibrated away. A partially polarized signal will produce only partial correlation between the two feed polarization outputs. In general one needs to solve for the correlated an uncorrelated components of the receiver outputs, but the basic idea is the same. A much more complete treatment of this subject will be covered elsewhere in this compendium.

6.   Pulsar Processing

Pulsar observations are some of the most demanding in terms of back-end signal processing because high resolution is required in both the time and frequency domains. Pulsar pulses are dispersed in the sense that a pulse arrives at high frequencies earlier than at low frequencies as is illustrated schematically in Figure 5.

The most common method for removing pulse dispersion is to bin the power samples in both time and frequency, as shown by the grid in Figure 5, and advance the lower frequency channel outputs to match the arrival times of the
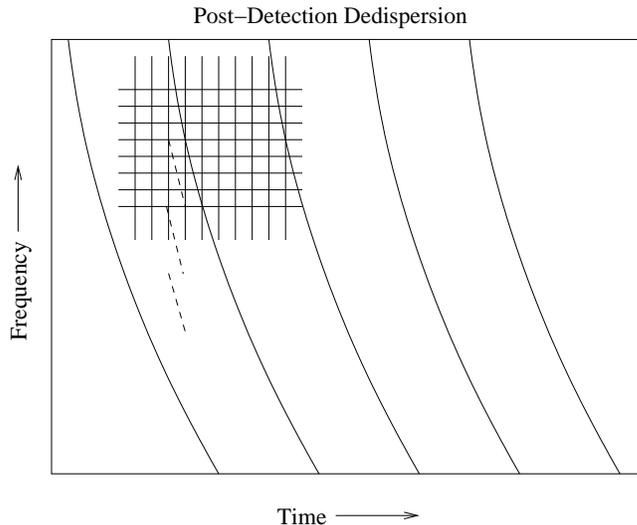
Post–Detection Dedispersion

Figure 5.   Schematic illustration of pulsar dispersion. The curved lines represent the loci of peak pulse radiation. The rectangular grid indicates how a pulsar spectrometer bins the time and frequency samples. The short dashed lines show how a pulsar may be dedispersed with post-detection sample shifts.

higher frequency channels as shown by the dashed line segments. This works well for low and moderate dispersions where the time and frequency bin sizes may be adjusted to provide as much resolution as needed to resolve the time structure within one pulse. However, for short period and high dispersion pulsars one runs into a fundamental limit that the transit time across one frequency bin can be shorter than the response time of the frequency channel that is set by its inverse bandwidth ($\tau \approx 1/B$).

This time resolution limitation may be overcome with pre-detection, sometimes called coherent, dedispersion as illustrated in Figure 6. In this case the broadband voltage time samples are transformed into the frequency domain with an FFT the same way as one would do for a direct Fourier transform spectrometer. Instead of squaring the output channel voltages to generate power samples in each time and frequency bin, the complex voltages are given a phase offset with progressively larger offset values across the frequency band. A linear phase gradient as a function of frequency is exactly equivalent to a simple delay in the time domain. A non-linear phase gradient is equivalent to dispersion or dedispersion according to its sign. The phase-modified spectrum is then transformed back into the time domain to produce a dedispersed time series with high pulse time resolution. An equivalent pre-detecton dedisperser can be constructed with a delay-domain digital filter instead of the back-to-back FFT's.

Maybe the most demanding type of observing in terms of signal and data processing load is a search for pulsars. Since the none of the parameters of the pulsars (sky position, period, and dispersion measure) are known ahead of time, spectra must be recorded continuously at about 200 microsecond intervals
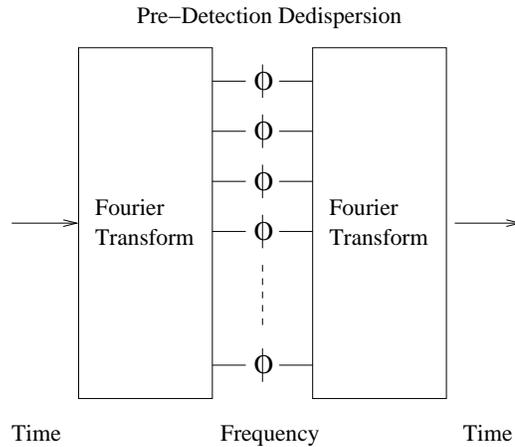
Pre–Detection Dedispersion

Figure 6.   A pre-detection pulse dedispersion scheme for achieving higher time resolution on short period and high dispersion pulsars.

for the duration of the survey. These data are then searched in pulse period and dispersion measure at each position on the sky with very capable off-line computers. If the spectra are 1000 channels of 4-bit data, one hour of observing requires 9 gigabytes of storage, and the data rate is 2.5 megabytes per second.

References

Cooley, J. W. & Tukey, J. W. 1965, "An Algorithm for the Machine Calculation of Complex Fourier Series," Mathematics of Computation, vol. 19, pp297-301

Rabiner, L. W. & Gold, B. 1975, Theory and Application of Digital Signal Processing, Prentice Hall (Englewood Cliffs, N.J.)

Thompson, A. R., Moran, J. M., & Swenson, G. W., Jr. 2001, Interferometry and Synthesis in Radio Astronomy, John Wiley & Sons, Inc., New York