'Seeing' radio waves

How can we 'see' radio waves? **André Young** and **David Davidson** explain the engineering components for the Square Kilometre Array.

Interferometry

Before reading the rest of this article, find 'From theory to practice', QUEST 8(3) 2012 by Oleg Smirnov.

In that article, Oleg explains that *interferometry* is a way to massively improve the optical resolution of radio telescopes. To summarise, interferometry is a method by which radio or light waves that are received at two different locations are combined and the resulting *interference pattern* is measured. Careful measurement of this pattern allows us to achieve an effective resolution that is determined by the distance between the two locations – the *baseline*. In the 1950s Sir Michael Ryle and his group at the Unversity of Cambridge developed a technique called *aperture synthesis* that used the principle of interferometry to combine multiple radio dishes into a single virtual telescope.

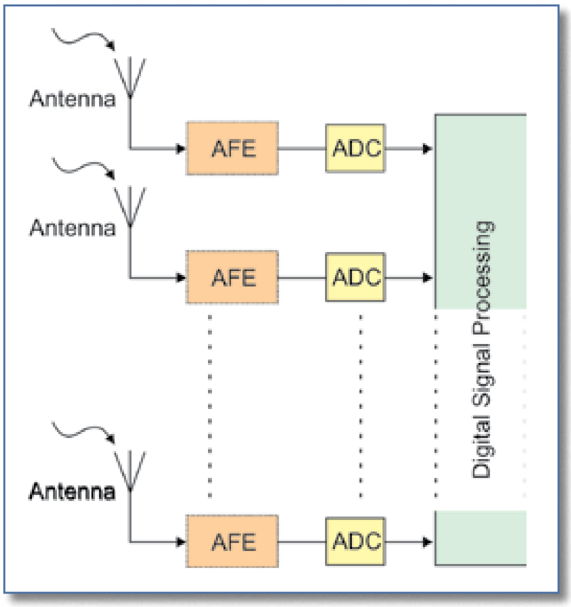


Figure 1: A simplified diagram of a general interferometer.

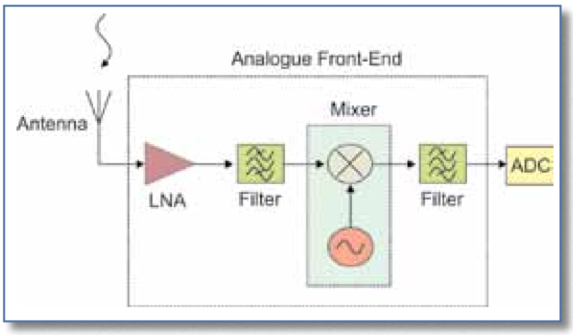


Figure 2: The components that make up the analogue front-end.

The Square Kilometre Array is planned to be the largest radio interferometer array ever built. Needless to say, there is a lot of engineering that goes into the realisation of such an ambitious project, from the design of the antennas that convert the incoming cosmic radiation into electronic signals, to the construction of the necessary infrastructure such as roads that allow access to the site. To cover all these aspects in one article is an impossible task. Instead, we will focus mainly on one small aspect of the big picture, called the *analogue front-end*, or AFE.

Figure 1 is a simplified diagram of a general interferometer.

You can think of an interferometer array as a large number of antennas that pick up radio waves – most likely radio dishes – standing in a remote location. These are only the first components in a very long receiver chain through which the radio waves propagate until they eventually reach the digital back-end. It is at this digital back-end that analogue signals are converted to digital signals by a component aptly called an *analogue-to-digital converter*, or ADC. This allows further digital processing to be done at later stages, such as combining signals of different antenna stations in the interferometer array to produce images of the radio sky. Between the antenna and the ADC is a series of components that ensure the correct transfer of the analogue signals put out by the antennas to the input of the ADC. This section is called the *analogue front-end* – *analogue* because that is the nature of the signal during this stage, and *front-end* because this is at the front of the receiver chain.

Figure 2 presents a more detailed view of the components that typically make up the analogue front-end. The first component is called a *low-noise amplifier* (LNA). This starts by amplifying the typically very small signal that is received by the antenna. This is then followed by a *filter*, which removes certain unnecessary or unwanted components that are present in the signal, then a *mixer* which transforms the signal frequency spectrum to allow simpler processing at later stages, and finally another filter stage to remove any unwanted signal components that may have sneaked in at earlier stages. After this stage the signal is about ready for digitisation and is passed on to the ADC. We will now discuss each of these stages in a bit more detail.

Antenna

Radio astronomy is the way that we 'see' celestial radio sources – that is celestial bodies that produce electromagnetic radiation within the *radio frequency*, or RF band of the electromagnetic spectrum. This frequency band extends roughly from 3 kilohertz up to 300 gigahertz. Unlike optical frequencies (around 400–800 terahertz, more than a 1 000 times higher than the highest RF frequencies) in which we are able to actually see electromagnetic radiation as visible light with our eyes, we cannot see RF radiation. So how exactly do we 'see' radio sources? This is where the *antenna* comes in – it transforms the incident electromagnetic waves into an electrical *signal* – current flowing through the terminals attached to the antenna – which we are able to manipulate, measure and analyse.

Antennas come in a variety of forms, ranging from a simple monopole antenna (a straight piece of wire, typically found on an FM radio) to parabolic reflector antennas (such as those used to receive satellite television). Each antenna is designed for a particular application. One of the factors that influences the size of the antenna is the wavelength at the frequency for which it is designed – the lower the frequency the larger the antenna. For example, a monopole is roughly one quarter of a wavelength in length, that is about 75 cm at 100 MHz (more or less the centre

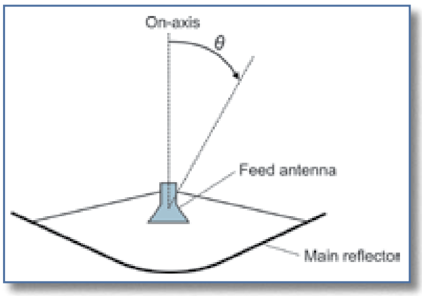
A close up of the electronics in the analogue front-end. Image: SKA South Africa/Nick van der Leek

of the frequency band in which FM radio stations transmit) and about nine times as short at 900 MHz (one of the frequencies used in cellular phones). For this reason cellular phones with built-in FM radios typically need headphones to be plugged in – they use the earphone cable itself as an antenna. This is because the antennas built into the phone for cellular communication are too small to receive FM radio transmissions.

In applications where very weak signals need to be detected, as is the case in radio astronomy, we use *high-gain* antennas. A widely used design, which falls in this category, is the parabolic reflector antenna, which consists of a large reflective surface – the *main reflector* – and a smaller antenna which is called the *feed antenna*. If you point the reflector antenna in a direction that means that a distant radio source is in line with the main reflector axis, the electromagnetic wave incident on the reflector surface is reflected so that all the power is focused at the focal point of the parabolic reflector. The feed antenna then absorbs this power at the focal point of the parabolic reflector. The main advantage of such a high-gain antenna is that waves incident from other directions (e.g. radio sources at other positions in the sky where we are not looking) are suppressed very strongly in the antenna output signal. An important property of an antenna, called the *gain pattern*, describes how well signals from various directions are received.

The ideal gain pattern of a typical parabolic reflector antenna is shown in Figure 3.

The reflector has a diameter of 13.5 m and the gain pattern is for a frequency of 1 GHz. In this figure we see that waves arriving from the direction $\theta = 0°$ (this direction is referred to as *on-axis*) are received with the maximum gain of the antenna, and as the direction of arrival changes (*off-axis*) the gain rapidly decreases. An important measure of the gain pattern of the antenna is the *half-power beamwidth* (HPBW) which indicates how far off-axis a source has to be positioned to reduce the received power by half. A simple formula using the wavelength λ and reflector diameter d can be used to estimate this figure for the antenna whose gain pattern is shown in Figure 3b: HPBW $\approx (180°/\pi) \times (\lambda/d) =$



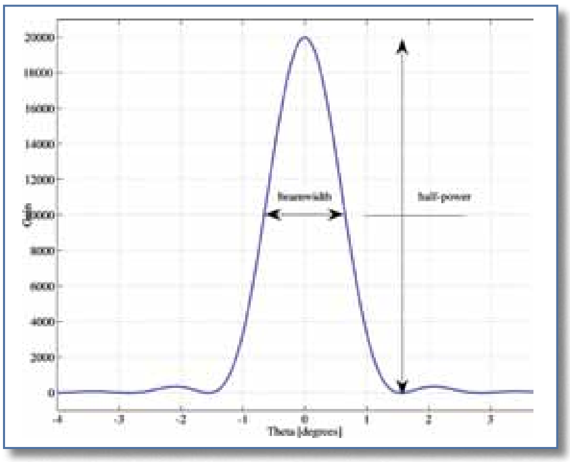


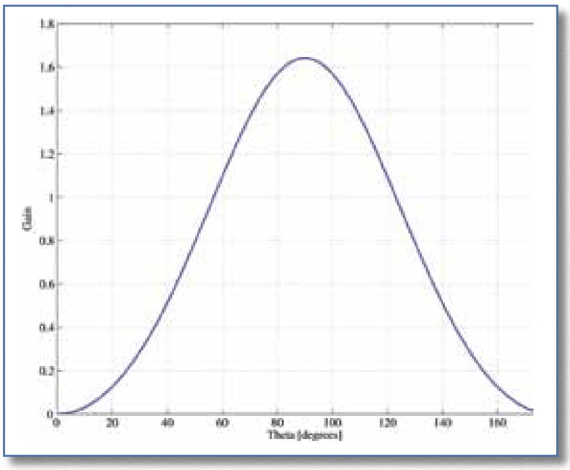
Figure 3: a (above): How electromagnetic waves are collected at the focal point of the parabolic reflector; b (right): The ideal gain pattern of a typical parabolic reflector antenna.

$(180°/\pi) \times 0.3/13.5 = 1.27°$

This means that if we point the antenna towards a source and measure the power we receive, and then point the antenna HPBW/2 = 0.635 degrees away from the source (why only half of the HPBW?) the received power would be halved.

> Satellite television transmissions are centred at roughly 12 GHz and a typical dish used in this application has a diameter of about 1 m. Use this information to estimate the HPBW for such a system. This should give you an idea of how accurately the dish should be orientated to successfully receive the transmission.

Now consider the size of a reflector antenna which would give comparable performance at the much lower frequency of 100 MHz and which would mean that the SKA would still be able to function. The frequency is ten times smaller than 1 GHz, which means that the size of the antenna needs to be ten times larger – that is a diameter of 135 m! Building such a large reflector is not impossible. In fact the world's largest single-reflector telescope at the Arecibo Observatory in Puerto Rico has a diameter of 305 m. The problem is that building an interferometer array on the scale required for the SKA means that thousands such antennas would need to be built. ▹▹



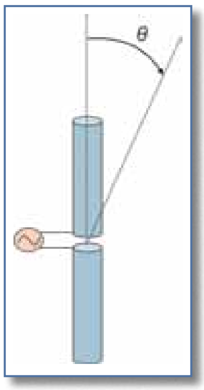
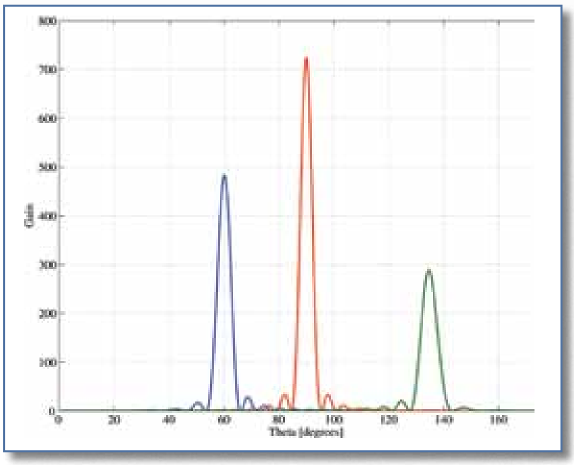


Figure 4: a (right): A dipole antenna; b (above): The radiation pattern of a single dipole antenna.

Figure 5: a (right): Dipoles arranged in a linear array; b (below): The pattern that results from adding up the signals received by each of the antennas in the array.



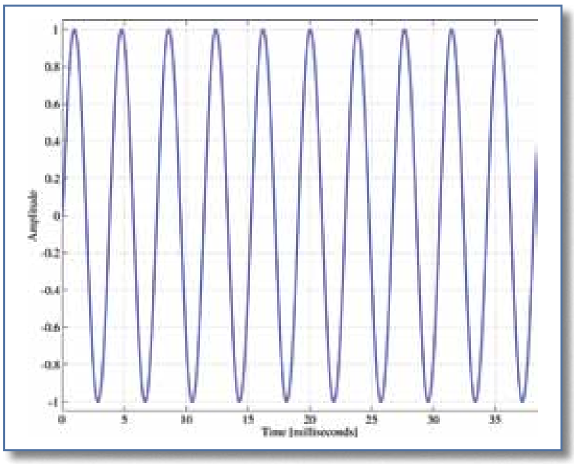


Figure 6a: The sinusoid produced when we plot the amplitude of a signal of 262 Hz as a function of time.

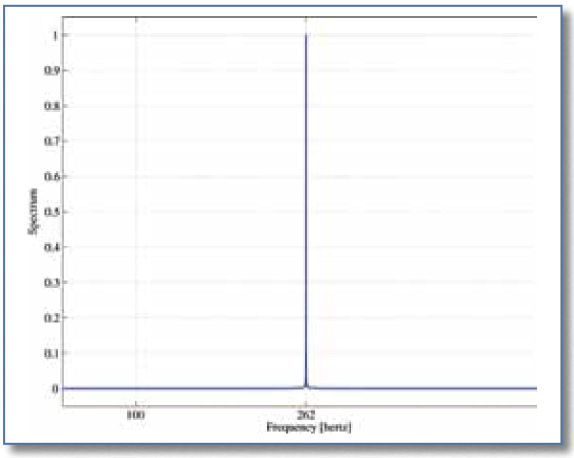


Figure 6b: The frequency spectrum of a signal of 262 Hz.

In this case another type of 'antenna', called an *antenna array* is particularly useful. Instead of building one large antenna, the signals of a bunch of smaller antennas are combined in a specific way so that they interfere constructively (add up) for waves incident from the directions of interest, and interfere destructively (cancel out) for waves incident from other directions. This not only improves the total antenna gain, but it also allows the gain pattern of the antenna array to be changed by changing the way the signals received by each of the antennas in the array are added up. Figure 4b shows the radiation pattern of a single dipole antenna (Figure 4a).

A dipole is very much like a monopole, except that it has two arms, as shown in Figure 4a. Notice that the gain of this antenna is a couple of orders of magnitude smaller than that of a large reflector antenna. Similarly the beamwidth of this smaller antenna is also much larger.

Now note how the radiation pattern changes when we use an array of dipoles, in this case a total of 21 dipoles arranged in a linear array, as shown in Figure 5a. Each of the plots shows the pattern that results from a different way of adding up the signals received by each of the antennas in the array (Figure 5b).

The gain is much higher and the beamwidth much narrower. Furthermore, we are able to *scan* the antenna electronically, that is to change the look-direction by simply changing the electronics in the receiver. In some cases this offers a major advantage over dish antennas, which require the entire reflector structure to be repositioned to change the look-direction.

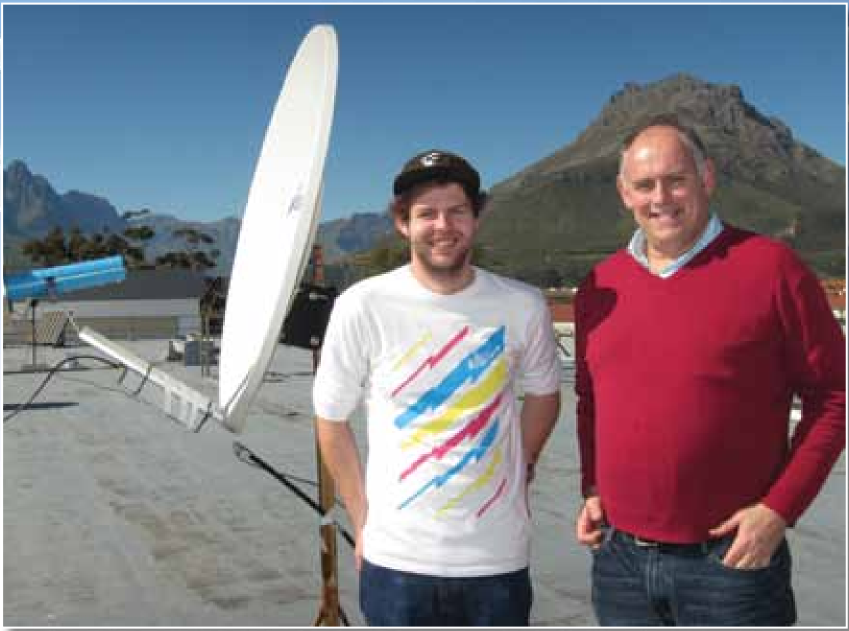
Low-noise amplifier

The next component that follows the antenna is a *low-noise amplifier*, or LNA, which amplifies the signal output by the antenna. Remember that the signal received by the antenna in a radio telescope is extremely small, originating from a source which is typically many light years away! Usually these LNAs are placed very close to the antenna terminals so that the cable between the antenna and the LNA is as short as possible. In order to understand the reason behind this, we first need to understand a little more about how 'noise' enters and affects the behaviour of our receiver system.

In any receiving system we distinguish between the desired part of what we receive (we simply call this the *signal*), and the undesired part of what we receive (we call this the *noise*). Imagine you are trying to listen to a friend talking on the phone while you are standing in a room full of people talking loudly – the friend's voice that you hear on your phone would be the signal, and all other sounds that you hear count as noise. In situations like these a useful measure is the *signal-to-noise ratio*, or SNR. This is simply the ratio of signal power to noise power. Suppose you were equipped with a device with which you could measure the power of the sound that you hear. You could then measure the signal power by moving to a quiet room and measuring the power in the sound of the friend's voice produced by the phone speaker alone. Similarly you could move to the noisy room, mute the phone speaker, and then measure the noise power as the power of the sound produced by the people talking loudly. The ratio of these powers (signal power divided by noise power) gives the SNR. Of course, the higher the SNR, the more clearly the signal can be identified and interpreted. If the noise in your room is too loud you are not able to hear your friend on the phone clearly.

Can you think of a way to better hear your friend speaking without leaving the noisy room? In other words, how can you increase the SNR? Turning up the volume of the phone speaker would certainly help, because the friend's voice would become louder while the noise in your room would remain the same (assuming the people in the room continue to talk at the same volume). The solution can also be understood by considering what happens to the SNR – increasing the volume increases the signal power (it amplifies the signal) without affecting the noise power (the noise in your room has nothing to do with what is happening in the phone). Increased signal power and constant noise power equals increased SNR.

Suppose now that you move to a quiet room so that you only hear the phone speaker, and also suppose that your friend on the phone moves



Juan-Pierre Jansen van Rensburg (MSc,completed 2012) and David Davidson with the one dish of two-element Stellenbosch University Experimental Interferometer in the background.

Image: David Davidson

to a room full of people talking loudly. Once again you will hear a combination of signal (your friend's voice) and noise (the voices of other people in your friend's room). What would happen now if you were to turn up the volume of the speaker? You increase the volume of the voice of your friend (signal power), but you also increase – by the same factor – the volume of the voices of the other people in your friend's room (noise power). Therefore the SNR does not increase. So what is different in this case?

In the previous scenario only the signal (friend's voice) was processed by the amplifier (phone volume control), and the noise (voices of other people in your room) only entered the system after the signal was amplified. But this time the noise (voices of other people in friend's room) entered the system *before* the amplifier (phone volume control) so that both the signal *and* the noise were amplified. The point is this: if the noise enters the system *before* the amplifier, the SNR is unaffected by the amplifier; if the noise enters the system *after* the amplifier, the amplifier increases the SNR.

Now in a receiver system such as a radio telescope, we need to deal with *external* as well as *internal* noise. The external noise enters the system the same way that the signal does, that is, as electromagnetic waves incident on the antenna. Sources which produce this noise include cosmic radio sources which we are not looking at, satellites in the region of the sky where we are observing, and terrestrial communication transmitters. Internal noise, on the other hand, is produced by the receiving system itself. The fact is, any practical electronic component will produce a certain amount of noise, even something as simple as a cable. This means that as we move along the signal path from the terminals of the antenna (where we already have a certain SNR – the ratio of signal power to external noise power), more and more internal noise is added by each component in the receiver system, so that the SNR may decrease as we move down the receiver chain. In order to minimise this effect that internal noise has on the SNR, it is necessary to add an amplifier *before* internal noise is introduced by any other components (of course, the amplifier itself should also add as little noise as possible, which explains the "low-noise" part of the LNA). This means placing the amplifier as close to the antenna terminals as possible, so that the amount of noise power added by the cable between the antenna and LNA is as small as possible.

Filters

Before we can understand the function of a filter we first need to understand the concept of a *frequency spectrum*. Almost any signal that can be represented as a function of time can also be represented as a collection of different frequencies. Depending on the nature of the signal, some frequencies will be more prominent than others. Consider for example playing the note middle C on a piano.

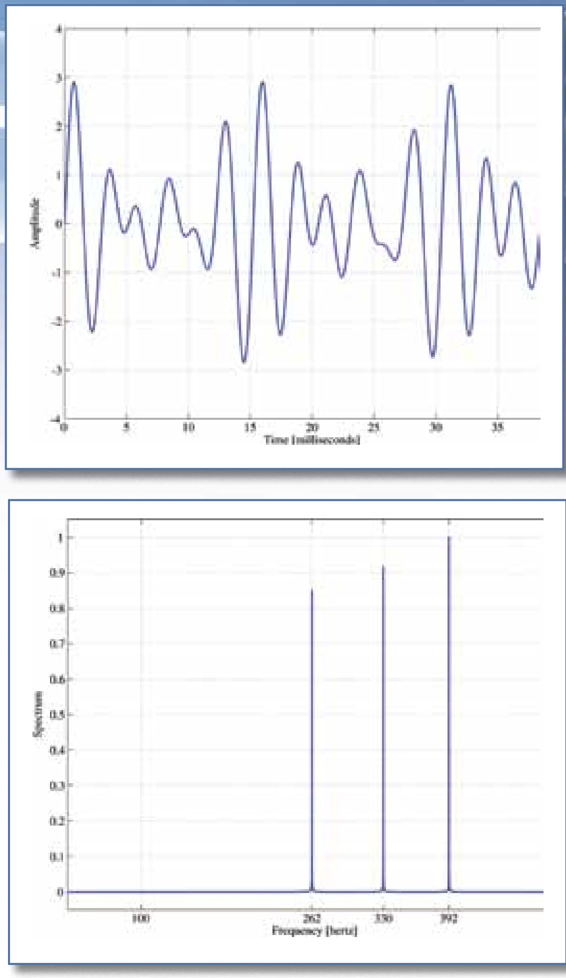


Figure 7: a (top): The signal produced by the three-note chord – C4, E4 and G4; b (above): The frequency spectrum of this signal.

David Davidson (left) and current MSc student Nicholas Thompson (right) with the analogue front end built by Juan-Pierre Jansen van Rensburg as part of his MSc. Image: SKA South Africa/Nick van der Leek

The sound that you would hear has a frequency of approximately 262 Hz. Now suppose the sound is recorded (transformed into an electronic signal) and we plot the amplitude of this signal as a function of time. The result is a sinusoid which we can express mathematically as $x(t) = \sin(2\pi \times 262t)$ and is shown in Figure 6a.

Using this time function and a mathematical tool called the *Fourier transform*, we are then able to represent the same signal as a function of frequency. This frequency ▹▹

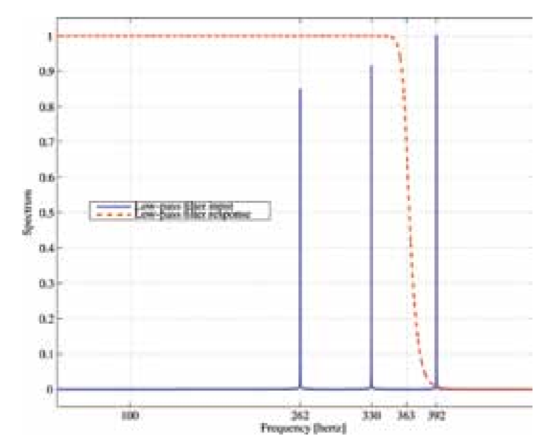


Figure 8a: The spectrum of the three-note chord with its low-pass filter response.

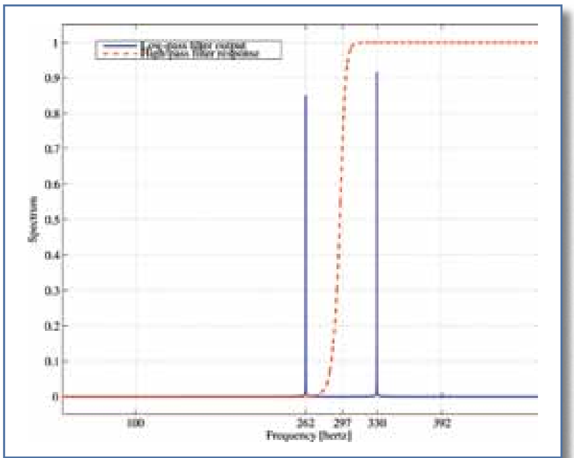


Figure 8b: The spectrum where the highest tone (392 Hz) has been removed, and the high-pass filter response.

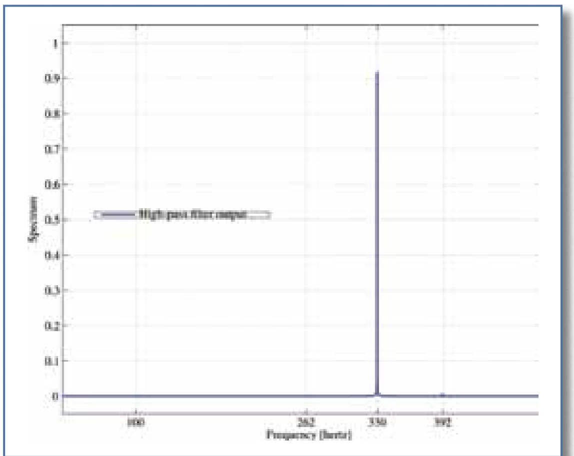


Figure 8c: The output of a high-pass filter – the lowest tone (262 Hz) has been removed from the signal so only E4 is left.

The Fourier transform

Joseph Fourier.
Image: Wikimedia Commons

The Fourier transform, developed in the 19th century by French mathematician Joseph Fourier, lies at the heart of all interferometry. This transform is a technique for representing any continuous signal by a sum of waves of different frequency, called Fourier components. If you've ever seen a spectrum analyser on a sound engineer's console (or just on a particularly advanced stereo), you have essentially seen a Fourier transform at work. The set of Fourier components is, in some sense, completely equivalent to the original signal. If we can measure the Fourier components, we can perform an inverse transform to recover the exact original signal.

Audio signals are one-dimensional, but an image of the sky can also be thought of as a signal – a two-dimensional (2D) one. A 2D Fourier transform turns it into a sum of 2D 'spatial waves', with large objects in the image associated with large ('low-frequency') waves, and small objects with small ('high-frequency') waves.

Now here's the interesting thing. If you were to somehow measure the signal on the surface of a telescope's mirror (or in the lens aperture of a camera) – before it is focused on the detector – you would find a Fourier transform of whatever the telescope or camera is pointed at. When the mirror (or lens) then focuses this signal on the detector, it is actually, just by its very nature, performing an inverse Fourier transform. As you read this article, the optical system of your eyes is also continuously performing forward-and-inverse Fourier transforms, in order to form an image of the page on your retina. Fourier transforms are literally everywhere.

Now remember that an interferometer works by connecting radio telescopes into a 'synthetic aperture', and measuring the signal in this aperture. This means that an interferometer actually measures the Fourier components corresponding to an image of the sky (rather than measuring an image of the sky directly). We can then do an inverse Fourier transform in software, and thus recover the original image. In some sense, an interferometer is like the 'front half' of a lens – with the rear half, the one responsible for the focusing, replaced by a computer that does inverse Fourier transforms.

From: From theory to practice – Oleg Smirnov, Quest *8(3) 2012.*

representation is called the *frequency spectrum* of the signal and is shown in Figure 6b. Note that the spectrum is zero at almost all frequencies, except for a spike at 262 Hz. This is because the signal is a sinusoid with that same single frequency.

Now consider what happens if we play a three-note chord on the piano by simultaneously playing C4 (another name for middle C, 262 Hz), E4 (a slightly higher tone, 330 Hz), and G4 (even higher tone, 392 Hz). The sound that is produced is a combination of three sinusoidal waves, each with a different frequency. If we again record the sound and plot the signal as a function of time the result would look like Figure 7a. The mathematical expression for this function is $x(t)= \sin(2\pi \times 262t) + \sin(2\pi \times 330t) + \sin(2\pi \times 392t)$. Accordingly the frequency spectrum of this signal has three spikes at the three frequencies corresponding to the notes played, as shown in Figure 7b.

The frequency spectra of signals are not always as simple as the examples shown here. One signal which is of particular interest is called *white noise* and is often used as a mathematical model for a noise in the analysis of a system. The spectrum for a white noise signal is a horizontal line – it contains all frequencies. An interesting question is how would such a signal sound? The hissing sound that a radio makes when it is not tuned in to a station is more or less how white noise would sound.

We now have enough knowledge to understand the purpose of a filter. Suppose you have a recording of the three-note chord played on a piano and you want to listen to the sound of one particular note only, for example E4. What you need in this application is a device that removes the unwanted frequencies (those of the C4 and G4 notes) from the signal – this is called a *filter*.

Filters come in many varieties. One classification of filters is based on which frequencies it allows to pass through. For example, a filter which removes all frequencies above a certain threshold (called the *cut-off frequency*) is called a *low-pass filter*, because all the lower frequencies are passed through. Similarly a *high-pass filter* will remove all frequencies below the cut-off frequency and allows higher frequencies to pass through.

Now let us apply filtering to the three-note chord example. Suppose we first apply a low-pass filter with a cut-off frequency of 363 Hz.

In Figure 8a the spectrum of the recorded three-note chord signal (input to the low-pass filter) is shown along with the filter response. The *filter response* is a visualisation which helps to determine what the frequency spectrum of the output would look like for a given input. The output is determined by multiplying the input spectrum with the filter response. For low frequencies the filter response is equal to one, so the output spectrum is simply equal to the input spectrum over that range of frequencies. At

The SKA research team at Stellenbosch. The authors are 2nd from the right (AY) and back row, 2nd from left (DBD). Image: SKA South Africa/Nick van der Leek

higher frequencies the filter response is equal to zero, so multiplication with any input spectrum produces an output spectrum that is equal to zero over that range of frequencies.

The output of this filter is shown in Figure 8b, where the highest tone (392 Hz) is seen to have been removed successfully. All that is left in this signal are the frequencies corresponding to the lowest and middle tones – if we had a recording of a two-note chord containing C4 and E4, this is what the spectrum of that signal would look like.

Now suppose we use the output of the low-pass filter as input to a high-pass filter with a cut-off frequency of 297 Hz and with the response shown in Figure 8b. Multiplying the filter response with the input signal produces zero at low frequencies and the input spectrum at higher frequencies. The output of this high-pass filter is shown in Figure 8c, where the lowest tone (262 Hz) is also seen to have been removed from the signal – the result of our two-stage filtering process is that we only have the frequency of the E4 note left in the signal.

In effect we have realised a *band-pass filter*, which only allows frequencies within a specified frequency band to pass through (in this case frequencies within the range 297–363 Hz), by cascading a low-pass filter and high-pass filter. This is not typically how band-pass filters are implemented in practice, but the underlying principles are the same.

Now that we have the frequency spectrum of our filtered signal, what does the filtered signal look like as a function of time? To answer this question we use the *inverse Fourier transform*. Just as the Fourier transform was used to represent a function of time as a function of frequency, the inverse Fourier transform is used to perform the reverse operation. Applying this inverse transform to the spectrum in Figure 8c gives the time function shown in Figure 9. Of course, the signal is exactly what we expect to see – a single sinusoidal wave, as would be produced if we recorded the sound of the E4 note played on its own.

ADC (analogue-to-digital converter)

Before the need for the mixer can be explained we first need to look ahead towards the following component in the signal chain, that is, the *analogue-to-digital converter* or ADC. This device does exactly what its name says – it converts the analogue signal at its input to a digital signal at its output. One way to think of this device is to think of it as measuring the strength of the signal at its input, and then recording the numerical value in computer memory. Of course, the input signal changes constantly and what we actually need is the signal as a function of time. So instead of measuring the signal once and outputting a single number, this device has to measure the signal repeatedly and every time output the value of the last measurement – this is called *sampling*.

One question that comes to mind is how often should the ADC measure the input signal? In other words, what should the interval between consecutive measurements be? This is called the *sampling frequency* of the ADC, and it turns out that if we want to keep all the information that is carried in the input signal the sampling frequency should be at

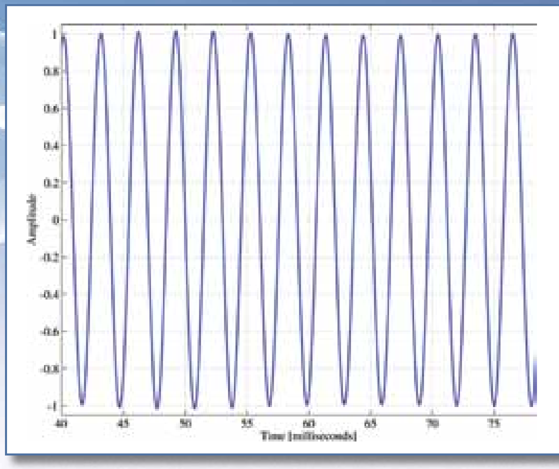
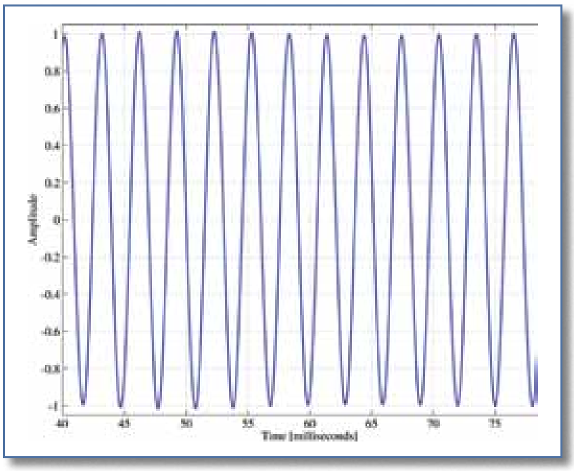


Figure 9: The result of applying the inverse transform to the spectrum in Figure 8c – a single sinusoidal wave.



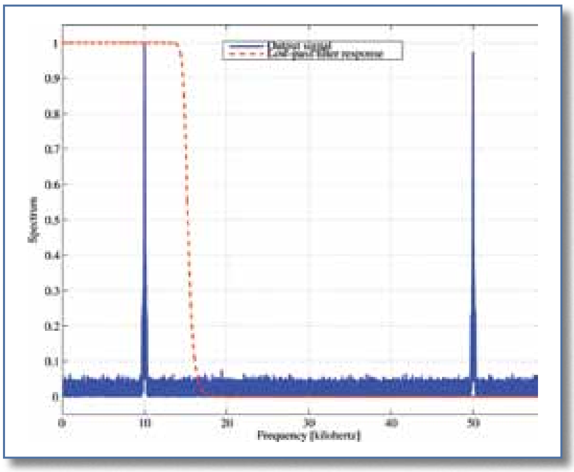


Figure 10: a (top): The input signal with a mixer frequency of 20 Hz; b (above): The output signal with a low-pass filter response.

least double the highest frequency present in the signal spectrum (this is called the *Nyquist frequency*). For example, the input signal in Figure 7b has frequency components at 262, 330, and 392 Hz. If we want to sample this signal without loss of information, then we need a sampling frequency that is at least 2 × 392 Hz = 784 Hz. That is, we need to take a measurement every 1.276 milliseconds. Suppose we are not interested in the higher frequency component and we filtered this out using the low-pass ▷▷

The maths behind the mixer

The mathematics behind the mixer is relatively simple and uses the sum and difference formulas for sine and cosine functions.

Let's suppose our input signal is a cosine with frequency f_1 and multiply it by a cosine of frequency f_2. So the input is $x(t) = \cos(2\pi f_1 t)$ and the output is: $y(t) = \cos(2\pi f_1 t) \times \cos(2\pi f_2 t)$.

We now perform a simple trick often used in situations like these, adding and subtracting the same thing from our expression, which does not change its value, but which helps us to look at it in a different way. So let us add $\frac{1}{2}\sin(2\pi f_1 t)\sin(2\pi f_2 t)$ and subtract that same quantity from $y(t)$, we then have: $y(t) = \cos(2\pi f_1 t)\cos(2\pi f_2 t) + \frac{1}{2}\sin(2\pi f_1 t)\sin(f_2 t) - \frac{1}{2}\sin(2\pi f_1 t)\sin(2\pi f_2 t)$

We write the product of cosine functions as a sum of two similar terms, just to rearrange the expression to give:

$$y(t) = \frac{1}{2}\cos(2\pi f_1 t)\cos(2\pi f_2 t) + \frac{1}{2}\sin(2\pi f_1 t)\sin(f_2 t) + \frac{1}{2}\cos(2\pi f_1 t)\cos(2\pi f_2 t) - \frac{1}{2}\sin(2\pi f_1 t)\sin(2\pi f_2 t)$$

The first two terms on the right are the difference formula for the cosine function, and the last two terms are the sum formula for the cosine function. So we can write:

$$y(t) = \frac{1}{2}\cos(2\pi(f_2 - f_1)t) + \frac{1}{2}\cos(2\pi(f_2 + f_1)t)$$

The output is now the sum of two cosines, one with a frequency $(f_2 + f_1)$ and one with a frequency $(f_2 - f_1)$.

If we use a low-pass filter with a cut-off frequency just above $(f_2 - f_1)$ then we can remove the higher frequency component, so that we just have: $y(t) = \frac{1}{2}\cos(2\pi(f_2 - f_1)t)$.

filter. The signal spectrum then looks like in Figure 8b, and the highest frequency is 330 Hz. To keep all the information in this signal intact we then only need to sample at 660 Hz – that is we need to take a measurement every 1.515 milliseconds. This means that the ADC can operate at a slower pace for lower frequency signals, a useful result to which we will return in the next section.

Another important property of the ADC is the *resolution*, expressed as the number of *bits* (binary digits) used in the number representing the measured value. The more bits available to represent the measurement, the higher the quality of the measurement. One can think of this as being the equivalent of the number of decimal digits used in a calculation. Say for example you want to calculate the circumference of a circle using the formula *circumference* = π × *diameter*. Since π is an irrational number we can only use an approximation of its true value in a numerical calculation. If we are only able to use one decimal digit we would use the value π ~ 3.1. Using two decimal digits, we have π ~ 3.14, and so on. More digits equals higher accuracy – the same principle applies to the number of binary digits that the ADC is able to use in representing the signal strength as a number.

The sampling frequency and resolution can be combined into a single figure of merit for the ADC, called the *bitrate*, which is measured in bits-per-second, or bps. This figure is simply the product of the sampling frequency and resolution – if every sample produces x bits, and the ADC obtains y amount of samples per second, then the output is $x \times y$ bits per second in total. The higher the bitrate, the higher the quality of the digital signal.

If you listen to digitally stored music on a computer (e.g. *.mp3 or *.wma files) you may notice that some players indicate the bitrate of the track that is playing. If you have access to encoder software, play around with the available bitrate options. Try to see if you can hear the effects of using very low bitrates.

Mixer

We now know that the frequency spectrum of the signal passed to the ADC determines at what rate the ADC should output samples of this signal. However, there is a practical limit to the number of samples which the ADC is able to process in a given time. This in turn places a limit on the highest frequency which should be passed to the ADC. Suppose we now have a signal that contains frequencies above the maximum frequency that is allowed to be fed to the ADC. This is similar to trying to listen to a dog whistle that works at 30 kHz – the frequency is just too high to be processed by the receiver (in this case the human ear, which is only able to hear up to about 20 kHz). One way to proceed is to filter out these high frequencies before passing the signal on, but this means we would lose the information that is carried in these high frequencies. A better solution is possible using a mixer in a process called *frequency multiplying* for reasons which will become clear in a moment.

To explain how this device works, we again need to represent our input signal as a function of time, let us call it $x(t)$. The mixer multiplies this signal with a sinusoidal function of a specific frequency, so that the output is $y(t) = x(t) \times \cos(2\pi f t)$. Does the name *frequency multiplier* make sense now? At the output of the mixer two signals are then produced, one signal which has the same spectrum as $x(t)$ only shifted down by f hertz, and another shifted up by f hertz. Consider again the example of the sound produced by the dog whistle. Suppose we have a recording device which is able to record this sound perfectly (works at all frequencies), transforming the sound waves into an electronic signal. We then pass this signal through a mixer, which multiplies the input with a 20 kHz sinusoidal wave, and then plays that output signal through a speaker system which again perfectly transforms the electronic signal into sound waves (works at all frequencies). The sound produced by the speaker would contain a 10 kHz (30 – 20 kHz) component as well as a 50 kHz (30 + 20 kHz) component. Each of these components contain all the information in the original signal, so we are allowed to filter out the one which is more convenient. In this case we might apply a low-pass filter so that we keep the component which we are able to hear. This is illustrated in Figure 10a and b.

In conclusion

This is only one aspect of the engineering components required for the SKA. In fact, there are still numerous questions about the design and construction of the SKA that have not yet even been answered. The nature of science and engineering is such that the more questions we seek to answer, the more questions we land up asking. But without the electrical engineering back-end, the SKA would not be able to answer the big questions you have learnt something about in the previous three articles. ❑

David Bruce Davidson received BEng, BEng (Hons), and MEng degrees from the University of Pretoria, South Africa, in 1982, 1983, and 1986 respectively, and the PhD degree from the University of Stellenbosch in 1991. In 2011, he was appointed to the South African Research Chair in Electromagnetic Systems and EMI Mitigation for SKA. His main research interest through most of his career has been computational electromagnetics. Recently, his interests have expanded to include engineering electromagnetics for radio astronomy.

André Young received the BEng degree in electrical and electronic engineering with computer science, MScEng degree in electronic engineering, and PhD in electronic engineering in 2005, 2007, and 2013, respectively, from the University of Stellenbosch. Currently he holds a post-doctoral appointment at Stellenbosch University, where his research focuses on calibration and phased array feed systems in radio astronomy.