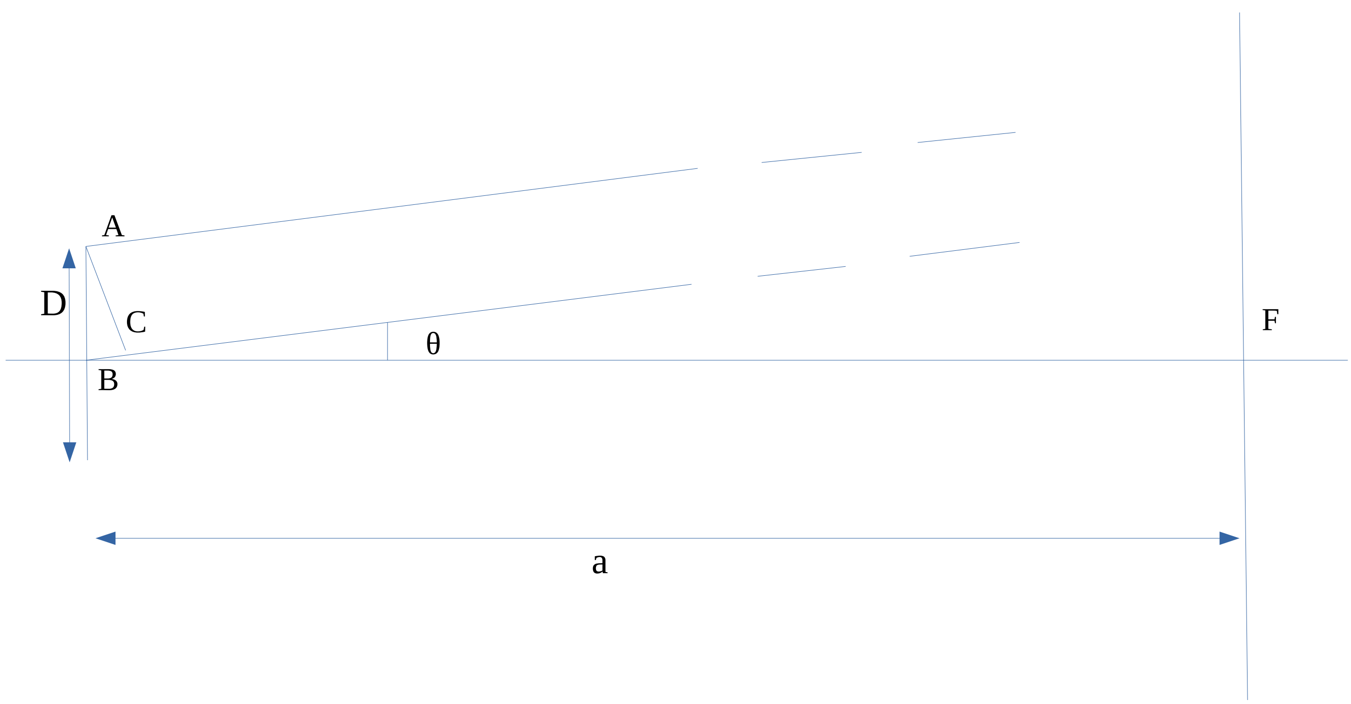
1) To explain semi-analytically and in reasonable depth, how diffraction leads to the side lobes and the main lobe (of width ~ λ/D), which is responsible for determining the field of view of a parabolic antenna.

Consider a one dimensional straight aperture representing the radiation emitting surface of an antenna. Every point on the antenna is assumed to emit radio waves of equal electric field magnitude. In addition, for astronomical purposes we are considering the far field pattern of the antenna. Hence the kind of diffraction we are dealing with is Fraunhofer diffraction in which the individual radio wave contributed by each point on the antenna to a given point on the screen can be considered parallel. Let the physical aperture size be *D* and the distance between the aperture and the screen where we are analysing the radiation pattern be *a*.



Consider radio waves from one end of the aperture *A* and from the center *B* to a point on screen at θ degrees from BF. Waves from A and B are considered parallel and AC is the perpendicular dropped from A on the ray from B. To know the resultant EM radiation at this point on the screen from A and B, we need to know the path length difference for the 2 radio waves. Since they are parallel to each other, the path length difference is equal to length of BC. The angle BAC is equal to θ and so the length of BC is equal to $(D \sin\theta)/2$. We are dealing with small angles so this becomes $(D\,\theta/2)$.

Now suppose we have a point P between A and B. We can then choose a point D/2 distance below P and the path length difference of the waves from these two points will again be $(D\,\theta/2)$. In this way we can choose every point on the antenna in a pair of points separated by D/2.

At D, $\theta = 0$ length of BC is zero and waves from every point of antenna is in phase and so we get a peak in intensity at D. If BC is equal to $\pm$ *$\lambda/2$* for some point on screen, the two waves will be totally opposite in phase and will cancel each other completely. But then every such pair of points on the antenna will cancel each other completely and we will have no radiation that point on the screen. The angle at which this happens is-

$$(D\,\alpha)/2 = \pm \lambda/2$$
$$\alpha = \pm\lambda/D$$

This is first minimum in the diffraction pattern. Beyond α, we will have more maximums corresponding to path length difference of $n\lambda$ and points of no radiation corresponding to $(2n+1)\lambda/2$. The intensity at D will be the highest and it will fall monotonically on either side for all maximums.
Hence in the region between $-\alpha$ and $+\alpha$ the antenna radiates energy at a much higher power as compared to other directions. In other words, it has a much higher *gain* in this direction and thus (by reciprocity) it can also *receive* signals much more efficiently in this direction. This region is called the ‘main lobe’ of the radiation pattern of the antenna and its size is proportional to α which is equal to λ/D.

2) To explain semi-analytically and in reasonable depth, how using an array of parabolic dishes in a radio interferometer improves the overall resolution of the telescope from (λ/D) to (λ/B), where B is the baseline of the array.

The most basic interferometer is made of only two antennas separated by a distance. An array of parabolic dishes can be effectively thought of as a large number of such interferometric pairs. So consider only two antennas separated by a distance b and both pointing in the direction θ relative to the vertical direction as shown in the figure.

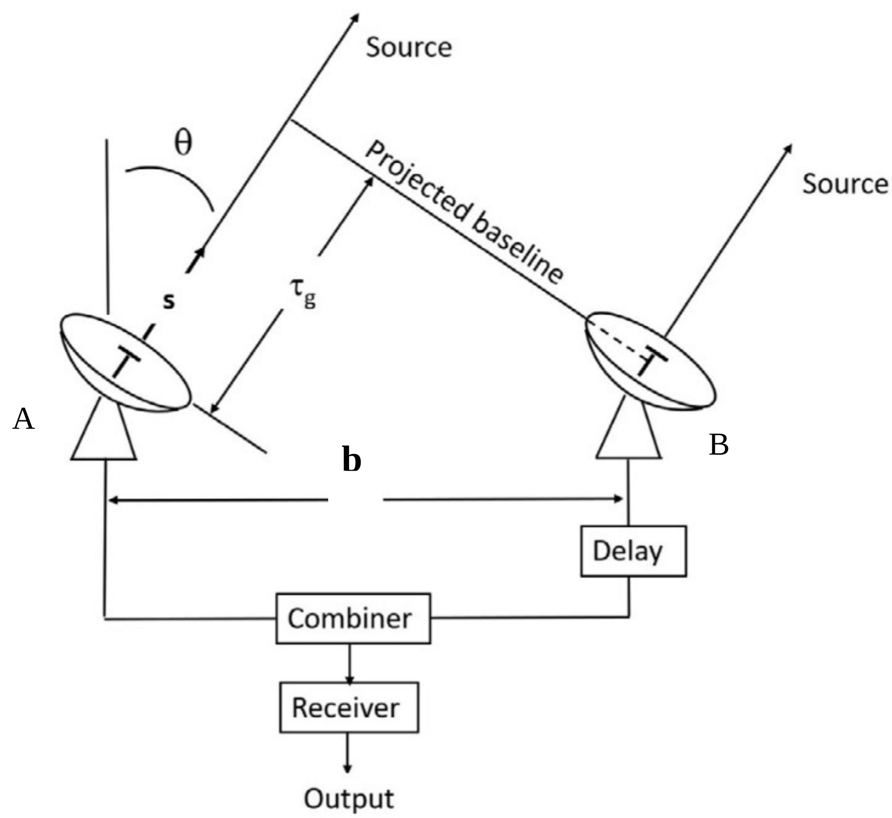


Figure 1
Image source – An Introduction to Radio Astronomy – Bernard F. Burke, Francis Graham Smith, Peter N. Wilkinson

Since the source is very far away, we can assume the approaching wavefronts to be planar. If that is so, then in the situation given in the diagram the antenna B receives the signal first and there is delay after which the antenna A receives. This is called *geometric delay* and is denoted by $\tau_{g.}$ This delay causes a phase difference between the response of the 2 antennas. From the diagram its clear that $\tau_g = (b \sin\theta)/c$ and the phase difference caused by this is-

$$\Delta\phi = 2\pi \nu \tau_g$$
$$\Delta\phi = 2\pi b \sin\theta / \lambda$$

The electrical signals produced by the individual antennas are brought together into a *combiner.* The combiner multiples the 2 signals and takes a time-average, in which case it is called *correlator*. Let the response of antenna A be $V_1 \cos \omega t$ and of antenna B be $V_2 \cos \omega(t+\tau_g)$. The output of the correlator will be-

$$R_{AB}(\tau_g) = \langle V_1 \cos \omega t * V_2 \cos \omega(t+\tau_g) \rangle \quad (\langle\rangle \text{ refers to time average})$$

which simplifies to-

$$R_{AB}(\tau_g) \propto \cos \omega\tau_g$$

Thus R_{AB} is a fringe pattern wrt τ_g which is function θ. To get the fringe width close to $\theta = 0^o$ -

$$\omega b (\sin(\theta+\Delta\theta) - \sin(\theta))/\lambda = 2\pi$$

Assuming θ to be small--

$$\omega b \Delta\theta/\lambda = 2\pi$$
$$\rightarrow \Delta\theta \text{ (fringe width)} = \lambda/b$$

R_{AB} looks like this (D is the aperture size of the intividual antenna) -

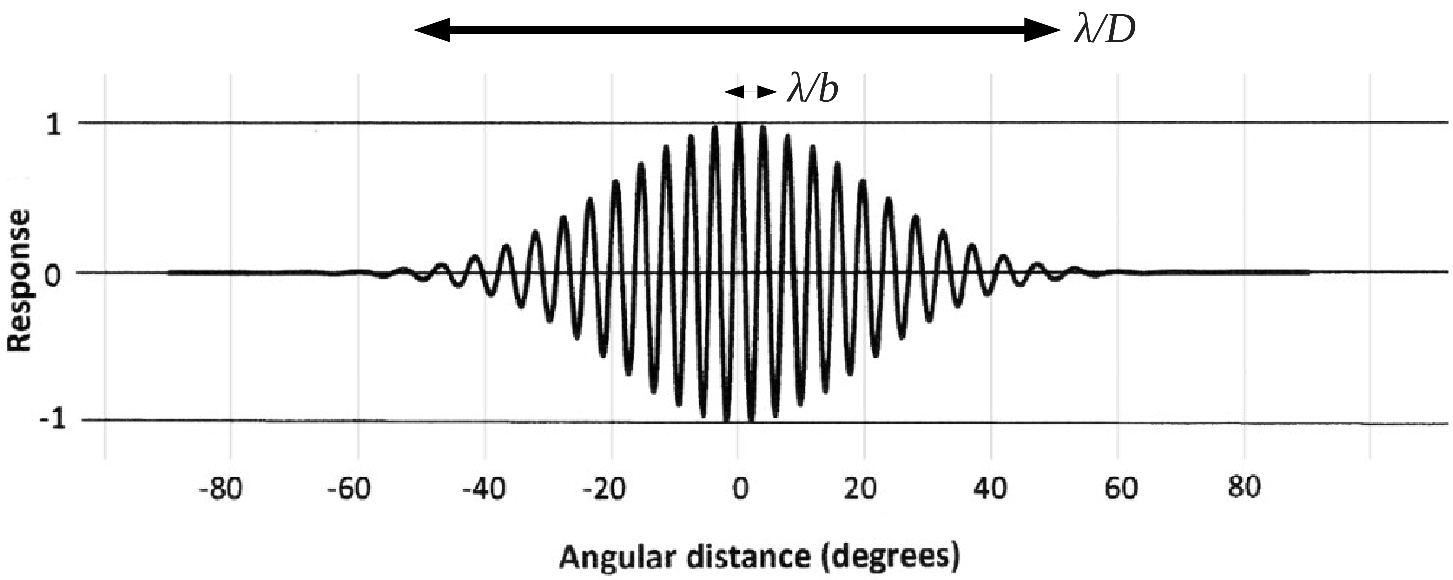


Figure 2 *Image source – An Introduction to Radio Astronomy – Bernard F. Burke, Francis Graham Smith, Peter N. Wilkinson*

But this is also the diffraction pattern of a point source on a single antenna of aperture size *b*. Hence if we measure the interferometer's response as Earth's rotation takes the point source along all the values of θ, our observations would be equivalent to if we had a single parabolic dish of aperture size *b*. This is the idea behind 'aperture synthesis' where we use Earth's rotation to generate an array's reception pattern as close to that of a single large antenna.

For resolution of the interferometer, consider another 2 point sources Aand B at an angular separation of say ϕ. Let ϕ be along the direction parallel to the baseline b. Then all the responses of the interferometer for A and B as Earth rotates will have the same form as figure 2 above except that the fringe patterns for the 2 sources will be shifted by the angle ϕ. Applying the Rayleigh Criterion for angular resolution, we see that that ϕ needs to be at least equal to the fringe width in order for the interferometer to resolve the 2 sources. The fringe width is equal to λ/b hence the resolution of the interferometer is λ/b.

3) In terms of the signal's wavelength, what is the maximum size of the deformations and irregularities in the antenna that can be used to focus the signal at the feeds efficiently?

For an ideally shaped parabolic antenna dish, the RF waves falling and reflecting off every point of the dish arrive at the focus in phase which leads to maximum strength of the antenna response. But if the parabolic surface has deformities and deviations from an ideal parabola, then there will be phase differences between the waves reaching the focus from different points of the dish. Interference between those waves will reduce the overall amplitude and thus the power collected by the antenna. For ex, considering only 2 points on the surface having a deformity of $\lambda/16$ above and below the ideal parabola respectively. The waves from these 2 points will reach the focus with a path difference of $\lambda/8$ which corresponds to phase difference of $\pm 45^0$. The superposed wave will have a reduced amplitude by .707 which translates to reduction in power by about .5.

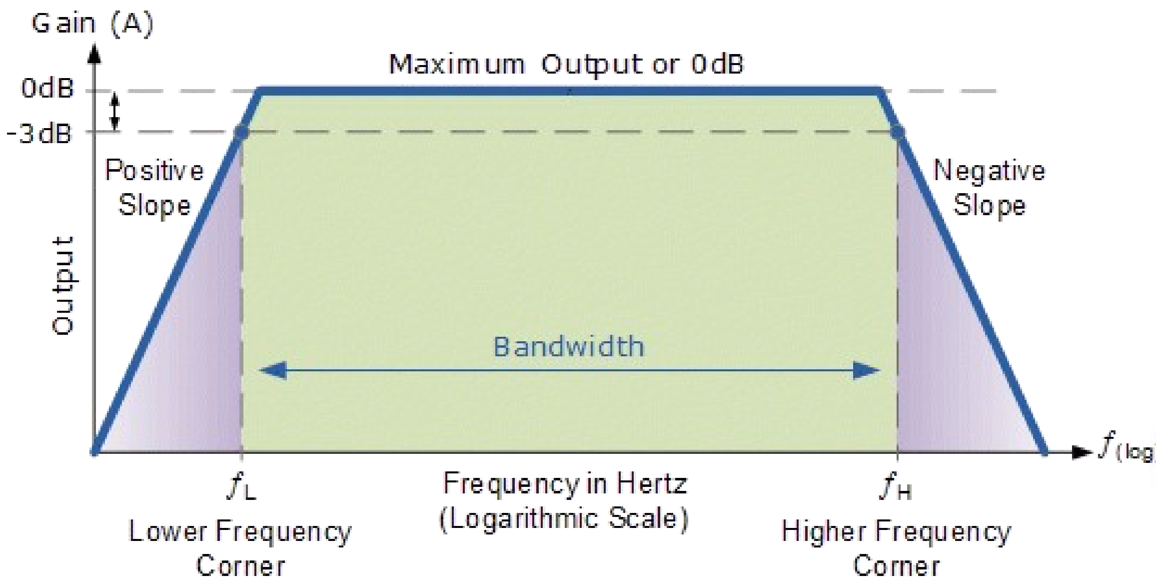
The Ruze equation provides a relationship between the effective aperture area of an ideal parabolic dish (A_0) and that of an dish with root-mean-square deviation of δz (A_δ) . The equation is-

$$A_\delta = A_0 \exp\{-(4\pi\ \delta z/\lambda)^2\}$$

The ratio A_δ / A_0 approaches 1 as δz gets smaller and smaller. The value of $\lambda/20$ for δz is considered sufficiently efficient with a A_δ / A_0 ratio of .67.

4) Apart from the broader limitations set by the dish and the feeds, which other parts of the signal processing line can limit the range of frequencies the telescope can be used for?

The RF low-noise amplifier used as the first element in the signal processing line receives signals of all the frequencies the antenna is passing down to it. But amplifiers have a particular ‘frequency response’ ie their gain is dependent on the frequency of the input signal. They amplify signals only in a particular range of frequencies (called their bandwidth) and their gain falls more than the acceptable level outside this range, so they cant be used much outside their bandwidth. An example of the frequency response of an amplifier is given below-



Source - www.electronics-tutorials.ws

The bandwidth of the amplifier is quantified as the difference between those 2 frequencies where the gain becomes half of the maximum gain of the amplifier. Half of the max gain would correspond to a decrease in 3dB of gain.

After the mixer reduces the frequency of the signal to an Intermediate Frequency (IF) and the IF frequency is amplified, the signal is passed through a band-pass filter at the backend, which selects a narrow range of frequencies. The mixer has to adjust the Local

Oscillator (LO) frequency in order to down-convert the desired RF frequencies to the range of the band-pass filter, but having a single filter limits the overall band-width the RF signal can have.

Thus the RF-amplifier and the band-pass filter are two of the ways by which the bandwidth of the radio telescope is limited.

5) What are the various ways in which electric circuits (in instruments like amplifiers etc.) are accompanied by additional noise?

There are different kinds of noises arising from different reasons. Some are-

1) Thermal Noise – Electrons in any circuit above absolute zero temperature have random motions in random directions, the extent of which increases with temperature. When a voltage is applied in the circuit, the net electric field drives the net flow of electrons in one direction, but the individual electrons are still bouncing off each other and having a velocity component in random directions. This causes small fluctuations in the current as compared to the current at 0 K. If the applied voltage is very small (ie if the radio signal is very weak), these random fluctuations can even dominate the current caused by the voltage.

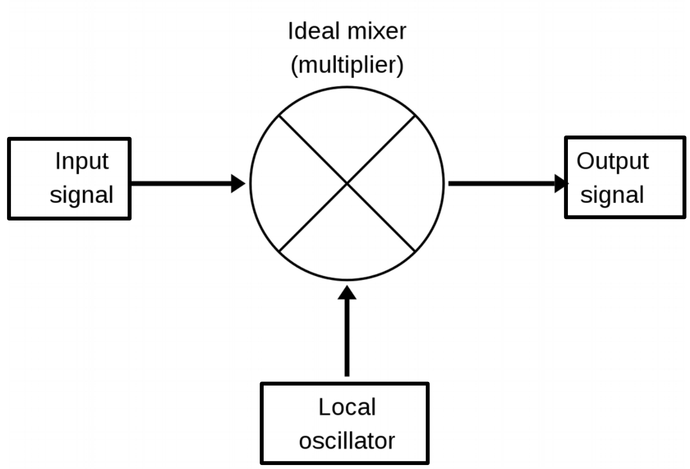
2) Shot Noise – When a current is passing through a region in the circuit, the number of electrons passing though that region in a second will on average be the same (ie the current value itself) but it wont be the *exact* same always. The electrons comprising the current are discrete packets of charge and not a continuous flow like that of a fluid. So the number of electrons passing through a cross-section will fluctuate. This causes fluctuations in the current which is seen as noise. The strength of these fluctuations relative to the signal will be high when the current is low and the number of electrons involved is small.

3) Flicker noise or 1/f noise – arises due to the fluctuations in the properties of the material making the resistors etc in the circuit. Its called 1/f because the noise power is inversely proportional to the frequency of the current. And so this noise dominates at low frequencies and thermal and shot noise dominate at higher frequencies.

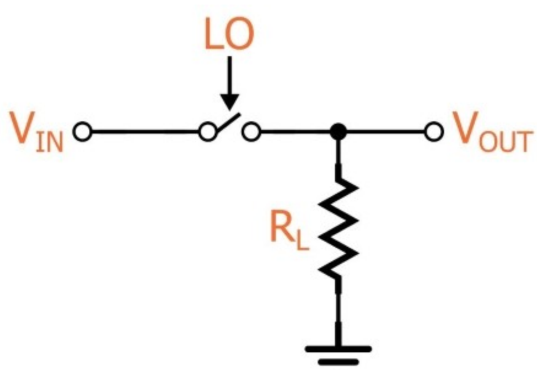
4) Avalanche Noise – In a reverse-biased p-n junction diode, there exists a high voltage at which the electrons from the 'leakage current' crossing the junction gain sufficient energy to knock off other electrons in the diode, resulting in a cascade. The current suddenly increasing, a large number of electrons flow to the other side reducing the previously high barrier potential. The cascade stops and with time the barrier potential again becomes high to start the cycle again. The exact moment at which the avalanche happens again and with what intensity is random by nature. This manifests as noise when using diodes.

6) To explain the input-output functionality of the mixer and how it does its job of combining 2 frequencies and giving their sum and difference as output frequencies.

The most basic layout of a mixer is given below-

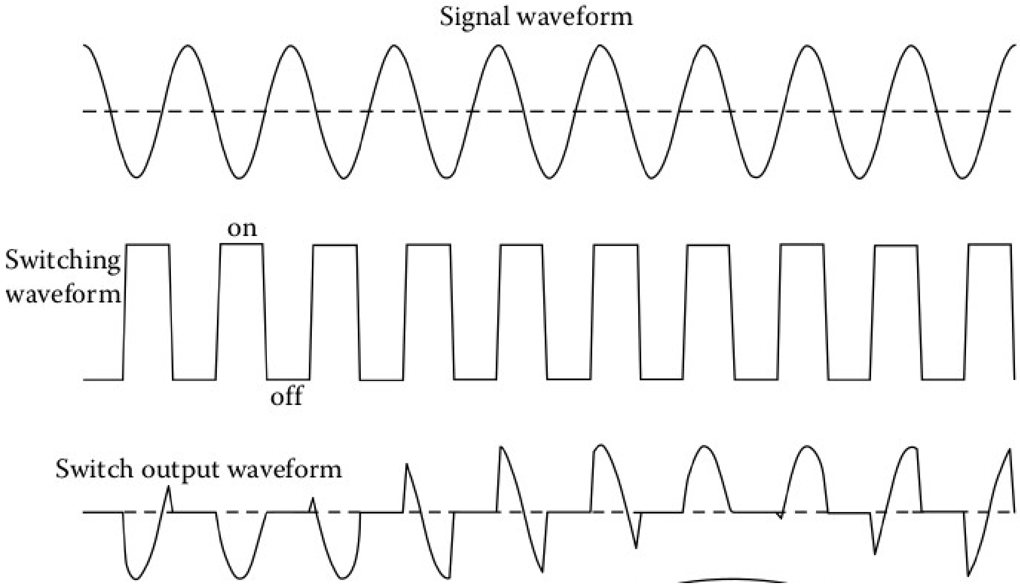


The mixer combines the input signal with another signal created by the local oscillator to return an output comprising 2 frequencies- the sum and difference of the input and LO frequency.
One of the ways to achieve this (called 'switching mixer') is to have the signal from the LO as a square wave of frequency close (but not equal to) the input frequency. This square wave can drive a switch which controls the passage of the input signal. The arrangement can be represented like-



Source - www.allaboutcircuits.com

The effect would be to remove parts of the signal according to the LO frequency and the V_{out} would like a very irregular signal as shown below-



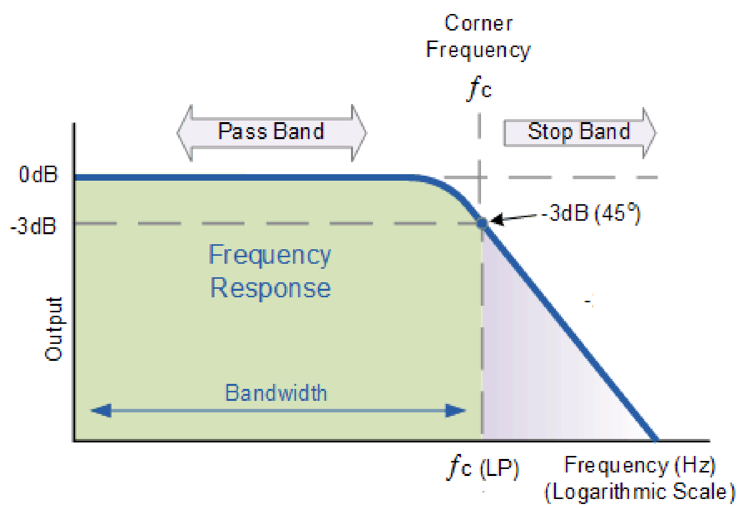
Source - Fundamentals of Radio Astronomy: Observational Methods

This irregular signal is composed of the (Input – LO) frequency and the (Input + LO) frequency. However, it also has various harmonics which are removed further in the signal processing line.

7) How does the mixer filter out one of the two frequencies it generates?

As mentioned above, the output of the switching mixer is a very irregular signal composed of a large number of frequencies, one of which is the difference between the input frequency and the LO frequency. This is also the smallest frequency present in the output so to filter it from the rest of the output, the signal is passed through a *low-pass filter.*

A low-pass filter only allows frequencies *below* a set cut-off and rejects all frequencies above it. The frequency response of a low-pass filter looks like-



Source - http://technlab.blogspot.com/

Hence, the low-pass filter retains only the (Input – LO) frequency which is called the Intermediate Frequency (IF). If there are some other unwanted frequencies left in the mixer output, they are filtered out in the back-end.