Note for Interferometer Array

Sehighs Aug. 2025

Contents

Notation Summary

Symbol	Meaning
$\mathbf{A}(\hat{\mathbf{r}}, \nu)$	Voltage primary beam (vector)
$A_p(\theta, \nu)$	Power primary beam ($\sim \lvert\mathbf{A}\rvert^2$)
$I(\theta, \nu)$	True sky brightness
$I'(\theta, \nu)$	Observed brightness ($= A_p I$)
$\mathbf{b} = (b_x, b_y, b_z)$	Baseline vector [m]
$\mathbf{u} = \mathbf{b}/\lambda = (u, v, w)$	Dimensionless baseline
$\theta = (l, m)$	Direction cosines
$V(\mathbf{b}, \nu)$	Visibility (complex)
η	Fourier dual to ν [Hz^{-1}]

1 Basic

1.1 Single–Dish Response and Primary Beam

For single dish, The resolution θ_{dish} of the resulting map is then roughly determined by the characteristic size D of the dish, where

$$\theta_{dish} = \frac{\lambda}{D} \tag{B.1}$$

and λ is the wavelength of observation. In interferometer case, The correlated data (termed the visibility) from each antenna pair of the interferometer roughly probes a Fourier mode that corresponds to the **characteristic angular scale**

$$\theta_{int} = \frac{\lambda}{b} \tag{B.2}$$

where b is the the distance between the two antennas that are being correlated. We will talk about its detail in Sec. 1.2 & 1.4.

A single dish receives the sky electric field as

$$\mathbf{E}(\mathbf{x}, t) = \iint \boldsymbol{\epsilon}(\hat{\mathbf{r}}, \nu)\, e^{i2\pi\nu t}\, d\Omega\, \mathrm{d}\nu. \tag{B.3}$$

ϵ is an electric-field spectral amplitude with units of $\mathrm{V\,m^{-1}\,sr^{-1}\,Hz^{-1}}$. The antenna response in below $\mathbf{A}(\hat{\mathbf{r}}, \nu)$ is more commonly called its primary beam or often just its beam. Note the primary beam here talk about is **vector electric field/voltage primary beam**. Hereafter we denote the electric field/voltage time stream of a single antenna by R(t):

$$R(t) = \alpha \iint \mathbf{A}(\hat{\mathbf{r}}, \nu)\, \boldsymbol{\epsilon}(\hat{\mathbf{r}}, \nu)\, e^{i2\pi\nu t}\, d\Omega\, \mathrm{d}\nu, \tag{B.4}$$

where α is a constant gain factor. For a physical interpretation of the beam we can think of the amplitude of the vector as telling us how sensitive we are in each direction on the sky; the orientation of the vector tells us what orientation of the incoming electric field we are sensitive to (i.e., the polarization); finally, the complex phase tells us about any relative time delay with which we receive the emission.

Let us analyze the unit of $R(t)$ in Eq. (B.4) : After the integration over solid angle and frequency we obtain a **electric field/voltage time series** $R(t)$ measured in electric field/volts (V) or (E). Throughout the rest of this section we use

$$R_i(t) \quad (i = 1, 2)$$

to denote the voltage output of antenna i. In the next subsection, we write E for its **electric field/voltage amplitude** at the receiver input, so that $R_i(t) = E\cos[\omega(t - \tau_i)]$ is consistent with the general definition above.

1.2 Visibility

Here we introduce the basic of Interferometer array. Two antennas (labeled 1 and 2) are separated by a baseline vector $\mathbf{b}$ of length b. A monochromatic plane wave of angular frequency ω arrives from direction $\hat{\mathbf{s}}$ (Note here we treat it as a unit vector), making an angle A with the baseline (see Fig. 1).

Fig. 1: Sketch of two-element interferometer geometry.

The geometric time delay between the two antennas is

$$\tau = \frac{\mathbf{b} \cdot \hat{\mathbf{s}}}{c} = \frac{b \sin A}{c}. \tag{B.5}$$

Assume a wave of angular frequency ω arriving at both antennas. For clarity, we first set the vector voltage primary beam $\mathbf{A}(\hat{\mathbf{r}}, \nu) \equiv 1$ and derive the two–element visibility under an ideal, frequency–independent beam. The received signals can be modeled as:

$$\begin{aligned} R_1(t) &= E \cos(\omega t) \\ R_2(t) &= E \cos[\omega(t - \tau)] \end{aligned} \tag{B.6}$$

The cross-correlation of the two signals, averaged over a time much longer than $1/\omega$, picks out the fringe term:

$$R_c(t) = \langle R_1(t) \cdot R_2(t) \rangle = \langle E^2 \cos(\omega t) \cos[\omega(t - \tau)] \rangle \tag{B.7}$$

Applying the trigonometric identity:

$$\cos A \cos B = \frac{1}{2}[\cos(A - B) + \cos(A + B)]$$

We get:

$$R_c(t) = \frac{E^2}{2} \langle [\cos(\omega\tau) + \cos(2\omega t - \omega\tau)] \rangle \tag{B.8}$$

Taking the time average over a period eliminates the fast oscillating term:

$$R_c(t) = \frac{E^2}{2} \cos(\omega\tau) \sim P \cos(\omega\tau) \tag{B.9}$$

where the fast oscillating term $\cos[\omega(2t - \tau)]$ averages to zero, and $P \propto E^2$ is the received power. This is the real part of the correlation output.

For an extended brightness distribution $I(\theta) \equiv P(\theta)$ (if gain α and $\mathbf{A}(\hat{\mathbf{r}}, \nu) \equiv 1$ and no noise, the output signal is real specific intensity/power, but in non-ideal case, we must replace $P(\theta)$ by $I'(\theta) \equiv A_p(\boldsymbol{\theta}, \nu)\, I(\boldsymbol{\theta}, \nu)$. This is a product in image (θ) space; equivalently, in Fourier (u, v) space it corresponds to a convolution of the sky intensity with the **power primary**

beam. The term A_p accounts for the fact that the antennas of an interferometer do not have equal sensitivity to all parts of the sky.) on the celestial sphere, parameterize direction cosines.

Note the $\theta \equiv (\theta_x, \theta_y) \equiv (l, m)$ is the coordinate of flat-sky (image plane) and uv-plane we will define later is the Fourier dual of it.

$$\hat{\mathbf{s}} = \hat{\mathbf{s}}_0 + \theta, \quad \hat{\mathbf{s}}_0 \perp \theta, \quad \mathrm{d}^2\theta \text{ the differential solid angle.} \tag{B.10}$$

Define the outputs by the function of distribution θ:

$$R_c(\mathbf{b}) = \iint P(\theta)\, \cos\!\Big[\frac{\omega}{c}\,\mathbf{b}\cdot\hat{\mathbf{s}}\Big]\, \mathrm{d}^2\theta \tag{B.11}$$

Here R_c is the *beam response* (real fringe pattern).

To extract the imaginary part, introduce a 90° phase shift to one of the signals:

$$\begin{aligned} R_1(t) &= E\sin(\omega t) \\ R_2(t) &= E\cos[\omega(t-\tau)] \end{aligned} \tag{B.12}$$

Similarly:

$$R_s(t) = \langle R_1(t)\cdot R_2(t)\rangle = \frac{E^2}{2}\sin(\omega\tau) \;\sim\; P\sin(\omega\tau) \tag{B.13}$$

Combining both in a complex form:

$$\begin{aligned} V(\mathbf{b}) = R_c(\mathbf{b}) \;-\; i\,R_s(\mathbf{b}) &= \iint P(\theta)\, \exp\!\Big[-i\,\frac{\omega}{c}\,\mathbf{b}\cdot\hat{\mathbf{s}}\Big]\, \mathrm{d}^2\theta \\ &= \exp\!\Big[-i\,\frac{\omega}{c}\,\mathbf{b}\cdot\hat{\mathbf{s}}_0\Big] \iint P(\theta)\, \exp\!\Big[-i\,\frac{\omega}{c}\,\mathbf{b}\cdot\theta\Big]\, \mathrm{d}^2\theta \\ &\xrightarrow{\text{calculate by hardware}} \iint P(\theta)\, \exp\!\Big[-i\,\frac{\omega}{c}\,\mathbf{b}\cdot\theta\Big]\, \mathrm{d}^2\theta \end{aligned} \tag{B.14}$$

This is the definition of the **visibility function**.

1.3 Amplitude & Phase of Visibility

From $V = R_c - iR_s$, define:

$$\begin{aligned} A(\mathbf{b}) &= |V| = \sqrt{R_c^2 + R_s^2}\,, && \text{(amplitude)} \\ \phi(\mathbf{b}) &= \arg V = \tan^{-1}\!\big(R_s/R_c\big)\,. && \text{(phase)} \end{aligned} \tag{B.15}$$

The amplitude encodes the spatial coherence (beam pattern), while the phase encodes astrometric information.

1.4 Baseline in Wavelength Units

Review the B.14, we rewrite it as below: the fundamental relation for visibility:

$$\begin{aligned} V(\mathbf{b}) &= \iint I'(\theta) \exp\left(-i\frac{\omega}{c}\mathbf{b}\cdot\theta\right) \mathrm{d}^2\theta \\ V(\mathbf{b}) &= \iint I'(\theta) \exp\left(-i\, 2\pi\mathbf{b}\cdot\theta/\lambda\right) \mathrm{d}^2\theta \end{aligned} \tag{B.16}$$

To describe the system, we introduce a **dimensionless coordinates** system (u, v, w), where the w-axis is perpendicular to the plane of the sky (i.e., parallel to the reference direction s_0, or we can say Line of Sight, or redshift).

In the image plane (θ flat-sky plane) and uv-plane (Fourier dual of θ) :

- l and m are the direction cosines toward the image.
- u, v are the spatial frequency coordinates is the Fourier dual of l, m, and we define $\mathbf{u} \equiv (u, v)$ *or* (u, v, w).

The direction cosines are defined as:

$$\begin{aligned} l &= \cos(\alpha) \\ m &= \cos(\beta) \\ n &= \cos(A) = \sqrt{1 - l^2 - m^2} \sim 1 \end{aligned} \tag{B.17}$$

where A is small angle from the reference direction, and l, m are very small (small field). The sky and reference vectors are:

$$\begin{aligned} \mathbf{s_0} &= (0, 0, 1) \\ \mathbf{s} &= (l, m, n) \end{aligned} \tag{B.18}$$

Thus, the deviation vector is:

$$\theta = \mathbf{s} - \mathbf{s_0} \quad \Rightarrow \quad \theta = (l, m, 0) \tag{B.19}$$

where we assume $l, m \ll 1$.

The baseline vector $\mathbf{b}$ is expressed as:

$$\mathbf{b} = \lambda(u, v, w) = \left(\frac{2\pi c}{\omega}\right)(u, v, w) \tag{B.20}$$

where λ (its unit) is the observed wavelength. I don't want to just present this formula without explanation like many book and review, now we see B.16. The complex visibility $V(\mathbf{b})$ in Equation B.16 contains a phase term:

$$\phi = -2\pi\frac{\mathbf{b}\cdot\theta}{\lambda} \tag{B.21}$$

This phase is the **core physical observable** in interferometry. To interpret it, note:

1. **Physical Meaning**: One can imaging that a single-dish antenna can be regarded as a special baseline with $\mathbf{b} = 0$ and it records the **auto-correlation** of its $R(t)$, therefore carries no phase information. In contrast, an antenna pair measures the **cross-correlated signal** which preserves the relative phase between antennas. Although the geometric phase term is $\mathbf{b} \cdot \theta/\lambda$, the phase arises only when $\mathbf{b} \neq 0$, hence it is natural to factor out the baseline–to–wavelength ratio $\mathbf{b}/\lambda$.

2. **Spatial Frequency Interpretation**: The phase represents a **spatial wave** on the sky:
$$\phi = -2\pi \left(\frac{\mathbf{b}}{\lambda}\right) \cdot \theta \tag{B.22}$$
where $\dfrac{\mathbf{b}}{\lambda}$ defines the *spatial frequency* (cycles per radian), and θ defines the *angular position*.
   - $u = b_x/\lambda$: Spatial frequency along RA (cycles/rad)
   - $v = b_y/\lambda$: Spatial frequency along Dec (cycles/rad)

3. **Dimensional Analysis**: $\mathbf{b}$ (baseline) has units of length $[L]$, λ (wavelength) has units $[L]$, and $\theta = (l, m)$ are dimensionless. $\implies$ The combination $\dfrac{\mathbf{b}}{\lambda}$ is **dimensionless**.

4. **Mathematical Necessity for Fourier Duality**: For $V(\mathbf{b})$ to be the Fourier transform of $I'(\theta)$, the exponent must satisfy:
$$-i2\pi \underbrace{\left(\frac{\mathbf{b}}{\lambda}\right)}_{\text{Fourier domain}} \cdot \underbrace{\theta}_{\text{Image domain}} \tag{B.23}$$
This **forces** the definition:
$$\boxed{\mathbf{u} \equiv \frac{\mathbf{b}}{\lambda} = (u, v, w)} \tag{B.24}$$

We can write
$$\mathbf{b} \cdot \theta \longrightarrow \left(\frac{2\pi c}{\omega}\right)(lu + mv), \tag{B.25}$$
review the Eq. B.16, we have
$$\begin{aligned} V(\mathbf{b}) &= \iint I'(\theta) \exp\left(-i\, 2\pi\, \mathbf{b} \cdot \theta/\lambda\right)\, \mathrm{d}^2\theta \\ V(u, v) &= \iint I'(l, m) \exp\left[-i2\pi(lu + mv)\right]\, \mathrm{dl}\, \mathrm{dm} \end{aligned} \tag{B.26}$$

We can see that the 2D Fourier transform naturally emerges. It's FT:

$$I'(l,m) = \iint V(u,v) \exp\left[i2\pi(lu+mv)\right] \, du \, dv \tag{B.27}$$

Okay! Time to calculate the visibility including **power primary beam**! We can write

$$V(\mathbf{b},\nu) = \iint I(\theta,\nu) \, A_p(\theta,\nu) \, e^{-i\,2\pi\,\mathbf{b}\cdot\theta/\lambda} \, \mathrm{d}^2\theta, \tag{B.28}$$

just using:

$$\boxed{I'(\boldsymbol{\theta},\nu) \equiv A_p(\boldsymbol{\theta},\nu) \, I(\boldsymbol{\theta},\nu)} \tag{B.29}$$

The **convolution theorem**

$$\boxed{\mathrm{FT}\left\{f(\theta)\,g(\theta)\right\}(u) = \left[\mathrm{FT}\left\{f\right\} * \mathrm{FT}\left\{g\right\}\right](u) = \int d^2u' \, \tilde{f}(u') \, \tilde{g}(u-u')}$$

tells us that the Fourier transform of a product is equal to the convolution of its respective Fourier transforms. Be regarded as a function on the θ–plane. By Eq. B.29, we can see its 2-D Fourier transform is just Eq. B.26, rewrite:

$$V(\mathbf{b},\nu) = \widetilde{I}'(\mathbf{u},\nu) = \int d^2\theta \, I'(\boldsymbol{\theta},\nu) \, e^{-i2\pi\mathbf{u}\cdot\theta}, \qquad \mathbf{u} \equiv (u,v) \equiv \mathbf{b}/\lambda. \tag{B.30}$$

Because we have defined the $\mathbf{u} \equiv (u,v)$ is the Fourier dual of $\theta \equiv (l,m)$, we can write FT of I and A_p as below:

$$\begin{aligned} \widetilde{I}(\mathbf{u},\nu) &\equiv \int_{-\infty}^{\infty} d^2\theta \; I(\boldsymbol{\theta},\nu) \, e^{-i\,2\pi\mathbf{u}\cdot\boldsymbol{\theta}}, \\ \widetilde{A}_p(\mathbf{u},\nu) &\equiv \int_{-\infty}^{\infty} d^2\theta \; A_p(\theta,\nu) \, e^{-i2\pi\mathbf{u}\cdot\theta}. \end{aligned} \tag{B.31}$$

Finally we get:

$$\begin{aligned} V(\mathbf{b},\nu) &= \int d^2u' \, \widetilde{I}(\mathbf{u}',\nu) \, \widetilde{A}_p\Big(\mathbf{u} - \mathbf{u}',\nu\Big). \\ V(\mathbf{b},\nu) &= \int d^2u' \, \widetilde{I}(\mathbf{u}',\nu) \, \widetilde{A}_p\Big(\frac{\mathbf{b}\nu}{c} - \mathbf{u}',\nu\Big). \end{aligned} \tag{B.32}$$

It is the visibility function in uv-plane, which shows that rather than precisely probing $\mathbf{u} = \mathbf{b}\nu/c$, the visibility probes a smeared out footprint around that mode. However, if the primary beam is relatively broad, then this footprint is relatively localized (δ function), and our basic intuition remains accurate. Under this approximation, each baseline probes a particular Fourier mode, and considering all the baselines of an interferometer (i.e., all possible antenna pairings) then provides information about multiple Fourier modes of the sky. These Fourier modes can then be Fourier transformed back into the image domain to provide a map of the sky. The more unique baselines there are in an interferometer array, the closer our map will be to a true image of the sky, and the longer the baselines of an array, the higher the resolution of our map. Note that as the Earth rotates, baseline vectors rotate relative to the sky, and thus they rotate through the uv plane (i.e., a two-dimensional plane with Cartesian axes given by the two components of $\mathbf{u}$).

1.5 Visibility with Frequency Space

We can assume the A_p very broad (set $A_p = 1$ in θ-plane, it's Fourier transform will be a δ function, we can ignore it as bellow). If we take the Fourier transform of Eq. B.32 along the line of sight, we end up with

$$\begin{aligned}
\widetilde{V}(\mathbf{b},\eta) &= \int \mathrm{d}\nu\, V(\mathbf{b},\nu)e^{-i2\pi\eta\nu} \\
\widetilde{V}(\mathbf{b},\eta) &= \int \mathrm{d}^2u\,\mathrm{d}\nu\, \widetilde{I}(\mathbf{u},\nu)\widetilde{A}_p(\mathbf{b}\nu_0/c-\mathbf{u},\nu_0) \\
&\approx \int \mathrm{d}\nu\, \widetilde{I}(\mathbf{b}\nu_0/c,\nu)e^{-i2\pi\eta\nu} \qquad \text{(B.33)} \\
&= \int d^2\theta\,\mathrm{d}\nu\, I(\boldsymbol{\theta},\nu)\, e^{-i\,2\pi(u\cdot\theta+\eta\nu)}|_{\nu=\nu_0} \\
&= I(\frac{\mathbf{b}\nu_0}{c},\eta),
\end{aligned}$$

the δ power primary beam also provides the central frequency ν_0 of the observation, note we also can write $\mathbf{u}\cdot\theta+\eta\nu$ as $lu+mv+\eta\nu$. Combine equations to give

$$\widetilde{V}(\mathbf{b},\eta) \propto \int \mathrm{d}^2\theta\,\mathrm{d}\nu\, I(\theta,\nu)e^{-i2\pi(\eta\nu+\mathbf{b}\cdot\theta\nu_0/c)}. \qquad \text{(B.34)}$$

We have

$$\widetilde{I}(\mathbf{k}_\perp,k_\parallel) \equiv \int_{-\infty}^{\infty} \mathrm{d}^2r_\perp dr_\parallel\, I(\mathbf{r}_\perp,r_\parallel)e^{-i(\mathbf{k}_\perp\cdot\mathbf{r}_\perp+k_\parallel r_\parallel)} \qquad \text{(B.35)}$$

We introduce the Cosmological Geometry X, Y, X^2Y here, X^2Y is a scalar translating observed units to cosmological distances, X converts from angles on the sky to transverse distances, Y from bandwidth to line-of-sight distance, these relation as below:

$$\begin{aligned}
X(z) &\equiv D_\mathrm{c}(z) = \int_0^z \frac{c\,\mathrm{d}z}{H_0E(z)}, \\
Y(z) &\equiv \frac{\partial D_c}{\partial\nu} = \frac{\partial D_c}{\partial z}\frac{\mathrm{d}z}{\mathrm{d}\nu} = \frac{c}{H_0E(z)}\frac{1+z}{\nu_0} = \frac{c\,(1+z)^2}{\nu_{21}\,H_0\,E(z)}, \qquad \text{(B.36)} \\
E(z) &= \sqrt{\Omega_m(1+z)^3+\Omega_\Lambda}.
\end{aligned}$$

Follow Eq. B.36, we can write:

$$\begin{aligned}\mathbf{r}_\perp &= X(z)\theta = D_c\theta \\ \Delta r_\parallel &= Y(z)\Delta\nu = \frac{\partial D_c}{\partial \nu}\Delta\nu = \frac{c(1+z)^2}{H_0\nu_{21}E(z)}\Delta\nu\end{aligned} \tag{B.37}$$

So we have

$$\widetilde{V}(\mathbf{b},\eta) \propto \int \mathrm{d}^3r\ I(\theta,\nu)exp[-i2\pi\left(\frac{H_0\nu_{21}E(z)}{c(1+z)^2}r_\parallel + \frac{\mathbf{b}\cdot\mathbf{r}_\perp\nu_0}{cD_c}\right)], \tag{B.38}$$

$$\begin{aligned}\mathbf{k}_\perp &= \frac{2\pi\nu_0\mathbf{b}}{cD_c} = \frac{2\pi\mathbf{u}}{D_c}; \\ k_\parallel &= \frac{2\pi\nu_{21}H_0E(z)}{c(1+z)^2}\eta\,.\end{aligned} \tag{B.39}$$

Or we can write:

$$\begin{array}{ll} 2\pi\,\mathbf{b}\cdot\theta\,\nu_0/c = \mathbf{k}_\perp\cdot\mathbf{r}_\perp & 2\pi\eta\,\nu = k_\parallel r_\parallel \\ \longrightarrow\ \mathbf{k}_\perp = 2\pi\,\mathbf{b}\cdot\theta\,\nu_0/\mathbf{r}_\perp & \longrightarrow\ k_\parallel = 2\pi\eta\,\nu/r_\parallel \\ \longrightarrow\ \mathbf{k}_\perp = \dfrac{2\pi\nu_0\,\mathbf{b}}{cD_c} = \dfrac{2\pi\mathbf{u}}{D_c}\,; & \longrightarrow\ k_\parallel = \dfrac{2\pi\eta}{Y(z)} = \dfrac{2\pi\nu_{21}H_0E(z)}{c(1+z)^2}\,\eta\,. \end{array} \tag{B.40}$$

We can see the $\mathbf{k}_\perp \propto \mathbf{b}$, it is because $\mathbf{b} \propto \mathbf{u}$, and $\mathbf{u}$ is the Fourier dual of sky plane θ, so we have $\mathbf{u} \propto \dfrac{1}{\theta}$ and proper distance also in direct proportion to θ, so as proper distance Fourier dual, $\mathbf{k}_\perp \propto \dfrac{1}{\theta} \propto \mathbf{u} \propto \mathbf{b}$. For η, likes k, the Fourier dual of proper distance, is $\propto \dfrac{1}{\nu}$, like

$$k = (2\pi/L)(i_x, i_y, i_j), i = 0\ to\ n-1 \tag{B.41}$$

we also have

$$\eta = (2\pi/\Delta\nu)(i) \tag{B.42}$$

Here i integers from 0 to n, in the real-world observation, the DFT mode number n determined by channel number.

In summary, or convenience to the derive Power Spectrum, or for consensus, we write the visibility form of (l, m, ν) and (u, v, η), and keep $A_p(l, m, \nu)$:

$$\widetilde{V}(u,v,\eta) = \int \mathrm{d}l dm \mathrm{d}\nu\ I(l,m,\nu)A_p(l,m,\nu)\exp[-i2\pi(lu+mv+\eta\nu)], \tag{B.43}$$

$$\widetilde{V}(u,v,\eta) = \frac{1}{X^2Y}\int \mathrm{d}^3r\ I(\mathbf{r})A_p(\mathbf{r})\exp(-i\mathbf{k}\cdot\mathbf{r}), \tag{B.44}$$

$$\begin{aligned}\mathbf{r} &= (r_x, r_y, r_\parallel) = (Xl, Xm, Y\nu) \\ 2\pi(u,v,\eta) &= (Xk_x, Xk_y, Yk_\parallel)\end{aligned} \tag{B.45}$$

1.6 Sampling & Imaging

Review the power primary beam:

$$I(l,m) = \frac{I'(l,m)}{A_p(l,m)}, \quad A_p(l,m) \sim G(l,m) : \text{approximately 2D Gaussian} \tag{B.46}$$

$$I'(l,m) = A_p(l,m)\, I(l,m) = \iint V(u,v) \exp\left[i2\pi(lu+mv)\right]\, du\, dv \tag{B.47}$$

During a single observation by an interferometric array, each baseline corresponds to a point on the (u,v) plane at a certain moment, and as time progresses, it "draws" a track.

Since the number of antennas is finite, the sampling of the $V(u,v)$ function is incomplete in practice. Thus, we can define a sampling function $S(u,v)$:

$$S(u,v) = \sum_{k=1}^{M} \delta(u-u_k, v-v_k) \tag{B.48}$$

where (u_k, v_k) are the sampled points.

In real observations, we cannot obtain a fully sampled $V(u,v)$. Instead, we only obtain $V(u,v)S(u,v)$, and taking the inverse Fourier transform gives the **Dirty Image**:

$$I^D(l,m) = \iint V(u,v)\, S(u,v)\, \exp[i2\pi(lu+mv)]\, du\, dv \tag{B.49}$$

Note the difference compared to:

$$I'(l,m) = \iint V(u,v) \exp[i2\pi(lu+mv)]\, du\, dv \tag{B.50}$$

Specifically,

$$I^D(l,m) = \mathrm{FT}^{-1}\{V(u,v)S(u,v)\} \tag{B.51}$$

$$I'(l,m) = \mathrm{FT}^{-1}\{V(u,v)\} \tag{B.52}$$

The Fourier transform of the sampling function $S(u,v)$ is called the **Dirty Beam**:

$$B(l,m) = \iint S(u,v) \exp[i2\pi(lu+mv)]\, du\, dv \tag{B.53}$$

$$B(l,m) = \mathrm{FT}^{-1}\{S(u,v)\} \tag{B.54}$$

Summary of key Fourier transforms:

$$
\begin{aligned}
I'(l,m) &= \iint V(u,v)\exp[i2\pi(lu+mv)]\,du\,dv \\
B(l,m) &= \iint S(u,v)\exp[i2\pi(lu+mv)]\,du\,dv \\
I^D(l,m) &= \iint V(u,v)S(u,v)\exp[i2\pi(lu+mv)]\,du\,dv
\end{aligned}
\tag{B.55}
$$

Thus, we have the key relationship:

$$
I^D(l,m) = I'(l,m) * B(l,m) \tag{B.56}
$$

where "$*$" denotes convolution. That is, the Dirty Image is the convolution of the True Image and the Dirty Beam. The Dirty Beam $B(l,m)$ acts effectively as the **PSF** (Point Spread Function).

2 Road to Power Spectrum

For the power spectrum derivation, we adopt a Fourier transform normalization convention that is consistent with that used in theoretical models (and is standard in cosmological work), but which differs from that used in radio astronomy. With respect to the brightness temperature in a pixel of the sky plane/frequency data cube, $T(\mathbf{x})$, and its Fourier dual $\widetilde{T}(\mathbf{k})$, this convention yields:

$$
\begin{aligned}
\widetilde{T}(\mathbf{k}) &= \frac{1}{V}\int T(\mathbf{r})\,e^{-i\mathbf{k}\cdot\mathbf{r}}\,\mathrm{d}^3\mathbf{r} \\
T(\mathbf{r}) &= \frac{V}{(2\pi)^3}\int \widetilde{T}(\mathbf{k})\,e^{i\mathbf{k}\cdot\mathbf{r}}\,\mathrm{d}^3\mathbf{k}
\end{aligned}
\tag{P.1}
$$

Here, V refers to the volume of the observed data cube and $\mathbf{r}$ is a 3D vector that indicates direction on the sky and depth (the frequency dimension) within the field. Likewise, k is a 3D wave-vector with projection $\mathbf{k}_\perp \equiv (k_x, k_y)$ in the plane of the sky, and $k_\parallel$ along the line-of-sight (frequency) direction.

It follows from this convention that an estimate of the power spectrum is given by

$$
\begin{aligned}
\widehat{P}(\mathbf{k}) &\equiv \left\langle\, |\widetilde{T}(\mathbf{k})|^2 \right\rangle = \int \widehat{\xi}(\mathbf{r})\,e^{-i\mathbf{k}\cdot\mathbf{r}}\,\mathrm{d}^3\mathbf{r} \\
\widehat{\xi}(\mathbf{r}') &= \frac{1}{V}\int T(\mathbf{r})\,T(\mathbf{r}+\mathbf{r}')\,\mathrm{d}^3\mathbf{r}
\end{aligned}
\tag{P.2}
$$

2.1 Response of Power Spectrum

We keep the ***radio interferometry convention*** in this section, rewrite the visibility:

$$\widetilde{V}(u,v,\eta) = \int \mathrm{d}^3r \; I(l,m,\nu)A_p(l,m,\nu)\exp[-i2\pi(lu+mv+\eta\nu)], \tag{P.3}$$

Squaring both sides and using $I = 2k_B T/\lambda^2$, with λ being the mean wavelength over the sub-band used in the Fourier transform, we have:

$$\begin{aligned} \widetilde{V}^2(u,v,\eta) \;=\; & \left(\frac{2k_\mathrm{B}}{\lambda^2}\right)^2 \int \mathrm{d}l\,\mathrm{d}m\,\mathrm{d}\nu\,\mathrm{d}l'\,\mathrm{d}m'\,\mathrm{d}\nu' \\ & \times A_p(l,m,\nu)\,T(l,m,\nu)\,A_p(l',m',\nu')\,T(l',m',\nu') \\ & \times e^{-i2\pi\,[u(l-l')+v(m-m')+\eta(\nu-\nu')]} \,. \end{aligned} \tag{P.4}$$

Drawing $A_p(l,m,\nu)$ into the bounds of the integral yields:

$$\begin{aligned} \widetilde{V}^2(u,v,\eta) \;=\; & \left(\frac{2k_\mathrm{B}}{\lambda^2}\right)^2 \int_{(0,0,0)}^{(\theta,\theta,B)} \mathrm{d}l\,\mathrm{d}m\,\mathrm{d}\nu \int_{(0,0,0)}^{(\theta,\theta,B)} \mathrm{d}l'\,\mathrm{d}m'\,\mathrm{d}\nu' \\ & \times T(l,m,\nu)\,T(l',m',\nu') \\ & \times e^{-i2\pi\,[u(l-l')+v(m-m')+\eta(\nu-\nu')]} \,. \end{aligned} \tag{P.5}$$

where $\theta \equiv \sqrt{\Omega}$, for primary beam field of view Ω because of the integration $\Omega B \equiv \int_0^\theta dl \int^\theta dm \int_0^B \mathrm{d}\nu = \theta^2 B$. Changing variables so that $(l_r, m_r, \nu_r) = (l-l', m-m', \nu-\nu')$:

$$\begin{aligned} \widetilde{V}^2(u,v,\eta) \;=\; & \left(\frac{2k_\mathrm{B}}{\lambda^2}\right)^2 \\ & \Bigg[\int_{(-\theta,-\theta,-B)}^{(0,0,0)} \mathrm{d}l_r\,\mathrm{d}m_r\,\mathrm{d}\nu_r \int_{(0,0,0)}^{(\theta+l_r,\,\theta+m_r,\,B+\nu_r)} \mathrm{d}l\,\mathrm{d}m\,\mathrm{d}\nu \\ & + \int_{(0,0,0)}^{(\theta,\theta,B)} \mathrm{d}l_r\,\mathrm{d}m_r\,\mathrm{d}\nu_r \int_{(l_r,\,m_r,\,\nu_r)}^{(\theta,\theta,B)} \mathrm{d}l\,\mathrm{d}m\,\mathrm{d}\nu\Bigg] \\ & \times T(l,m,\nu)\,T(l-l_r,\,m-m_r,\,\nu-\nu_r) \\ & \times e^{-i2\pi\,[u\,l_r+v\,m_r+\eta\,\nu_r]} \,. \end{aligned} \tag{P.6}$$

Eq. (P.6) comes from Eq. (P.5):

- (l,m,ν) are sky plane (image plane) coordinates inside the primary beam;
- (l_r, m_r, ν_r) denote the displacement between two points where brightness temperatures will be multiplied;

- the two points encodes a same pair with Fourier coordinates (u, v, η).

Consider the simple top-hat window power primary beam:

$$A_p(l, m, \nu) = \begin{cases} 1, & |l| \leq \theta, \quad |m| \leq \theta, \quad |\nu| \leq B, \\ 0, & \text{otherwise.} \end{cases} \tag{P.7}$$

because $|l_r|, |m_r| \ll \theta, \; |\nu_r| \ll B$, we have

$$\begin{aligned} A_p(l, m, \nu) &\sim A_p(l - l_r, m - m_r, \nu - \nu_r) \\ A_p(l, m, \nu) &\sim A_p(l, m, \nu) A_p(l - l_r, m - m_r, \nu - \nu_r) \, , \end{aligned} \tag{P.8}$$

so

$$\begin{aligned} &\int_{(l_r, m_r, \nu_r)}^{(\theta, \theta, B)} \mathrm{d}l \, \mathrm{d}m \, \mathrm{d}\nu \\ &\qquad \times T(l, m, \nu) \, T(l - l_r, \, m - m_r, \, \nu - \nu_r) \, e^{-i2\pi \, [u \, l_r + v \, m_r + \eta \, \nu_r]} \\ &= \int_{(0,0,0)}^{(\theta + l_r, \, \theta + m_r, \, B + \nu_r)} \mathrm{d}l \, \mathrm{d}m \, \mathrm{d}\nu \\ &\qquad \times T(l, m, \nu) \, T(l - l_r, \, m - m_r, \, \nu - \nu_r) \, e^{-i2\pi \, [u \, l_r + v \, m_r + \eta \, \nu_r]} \\ &= e^{-i2\pi \, [u \, l_r + v \, m_r + \eta \, \nu_r]} \\ &\qquad \times \int_{(0,0,0)}^{(\theta, \theta, B)} \mathrm{d}l \, \mathrm{d}m \, \mathrm{d}\nu \; T(l, m, \nu) \, T(l - l_r, \, m - m_r, \, \nu - \nu_r) \\ &= \Omega B \times \widehat{\xi}_{21}(l_r, m_r, \nu_r) \, e^{-i2\pi \, [u \, l_r + v \, m_r + \eta \, \nu_r]}. \end{aligned} \tag{P.9}$$

Finally we get:

$$\begin{aligned} \widetilde{V}_{21}^2(u, v, \eta) \; = \; \left(\frac{2k_{\mathrm{B}}}{\lambda^2} \right)^2 \Omega B \int_{(-\theta, -\theta, -B)}^{(\theta, \theta, B)} \mathrm{d}l_r \, \mathrm{d}m_r \, \mathrm{d}\nu_r \\ \times \, \widehat{\xi}_{21}(l_r, m_r, \nu_r) \, e^{-i2\pi \, [u \, l_r + v \, m_r + \eta \, \nu_r]} \; . \end{aligned} \tag{P.10}$$

Use the relation $\mathbf{r} = (Xl, Xm, Y\nu)$, take $\mathrm{d}l_r \, \mathrm{d}m_r \, \mathrm{d}\nu_r \longrightarrow \dfrac{1}{X^2 Y} \mathrm{d}X l_r \, \mathrm{d}X m_r \, \mathrm{d}Y \nu_r \longrightarrow \dfrac{1}{X^2 Y} \mathrm{d}^3\mathbf{r}$, we get:

$$\widetilde{V}_{21}^2(u, v, \eta) \; = \; \left(\frac{2k_{\mathrm{B}}}{\lambda^2} \right)^2 \frac{\Omega B}{X^2 Y} \int_{(-X\theta, -X\theta, -YB)}^{(X\theta, X\theta, YB)} \mathrm{d}^3\mathbf{r} \;\; \widehat{\xi}_{21}(\mathbf{r}) \, e^{-i\mathbf{k} \cdot \mathbf{r}} \; . \tag{P.11}$$

We should note the term $\widehat{\xi}_{21}(\mathbf{r})$ in fact is the Fourier transform of $\xi_{21}(\mathbf{r})$ with a A_p window, here is top-hat window, in Fourier space it becomes the convolution of the Fourier transform of these functions:

$$\widetilde{V}_{21}^2(u, v, \eta) \; = \; \left(\frac{2k_{\mathrm{B}}}{\lambda^2} \right)^2 \frac{\Omega B}{X^2 Y} \; \widehat{P}_{21}(\mathbf{k}) \, * A_p(X\theta k_x, X\theta k_y, YBk_{\|}) \; . \tag{P.12}$$

Assume A_p is very broad, we can ignore the term, finally:

$$\begin{aligned}\widetilde{V}_{21}^{2}(u,v,\eta) &= \left(\frac{2k_{\rm B}}{\lambda^2}\right)^2 \frac{\Omega B}{X^2 Y}\, \widehat{P}_{21}(\mathbf{k}) \\ &= \left(\frac{2k_{\rm B}}{\lambda^2}\right)^2 \frac{\Omega B}{X^2 Y}\, \frac{2\pi^2}{k^3} \widehat{\Delta}_{21}(\mathbf{k})\ .\end{aligned} \tag{P.13}$$

It is the single baseline response.

2.2 General derivation & Review

In the previous subsection, we assume the power primary beam A_p pattern is top hat. Consequently we get the relation: $A_p(l,m,\nu) \sim A_p(l,m,\nu)A_p(l-l_r, m-m_r, \nu-\nu_r) \sim A_p^2(l,m,\nu)$, and used in Eq. P.9 Now we follow parsons 2014 to give a more general derivation. Because we re-derive the result from first principles, we spell out every relation we use so that this subsection is a self-contained, **fully detailed** exposition. Simply: In previous derivation/parsons 2012, we set A_P to 1, and the contribution of this term is just give a limit of integral, like $V(u,v,\eta)$, it should be integral of square of the term, $\int A_p^2(l,m,\nu)$. So the term Ω in Eq. P.13 should be replaced as $\Omega_{pp} = \int A_p^2(l,m,\nu)$. Now we give the detailed derivation begin with Eq. B.43:

$$\widetilde{V}(u,v,\eta) = \int \mathrm{d}^2\theta\, \mathrm{d}\nu\ I(\theta,\nu)A_p(\theta,\nu)\exp\{-i2\pi[(\theta\cdot\mathbf{u})+\eta\nu]\}, \tag{P.14}$$

$$\widetilde{V}(u,v,\eta) = \int \mathrm{dl}\,\mathrm{dm}\,\mathrm{d}\nu\ I(l,m,\nu)A_p(l,m,\nu)\exp[-i2\pi(lu+mv+\eta\nu)], \tag{P.15}$$

$$\widetilde{V}(u,v,\eta) = \frac{1}{X^2Y}\int \mathrm{d}^3r\ I(\mathbf{r})A_p(\mathbf{r})\exp(-i\mathbf{k}\cdot\mathbf{r}), \tag{P.16}$$

The translation of the two equation is just use the relation:

$$\begin{aligned} 2\pi\mathbf{u}\cdot\theta &= \mathbf{k}_\perp\cdot\mathbf{r}_\perp, \\ 2\pi\eta\nu &= k_\parallel r_\parallel, \\ \mathbf{r} = (r_x, r_y, r_\parallel) &= (Xl, Xm, Y\nu). \end{aligned} \tag{P.17}$$

Using the convolution theorem, review the Fourier convention:

$$\widetilde{F}(\mathbf{k}) = \frac{1}{V}\int \mathrm{d}^3r\, F(\mathbf{r})\, e^{-i\mathbf{k}\cdot\mathbf{r}}, \qquad F(\mathbf{r}) = V\int \frac{\mathrm{d}^3k}{(2\pi)^3}\, \widetilde{F}(\mathbf{k})\, e^{+i\mathbf{k}\cdot\mathbf{r}}\ , \tag{P.18}$$

the the visibility as follow:

$$\widetilde{V}(u,v,\eta) = \frac{V}{X^2Y}\int \frac{\mathrm{d}^3k'}{(2\pi)^3}\ I(\mathbf{k}')A_p(\mathbf{k}-\mathbf{k}'), \tag{P.19}$$

The result of multiplying two complex visibility is:

$$
\begin{aligned}
\widetilde{V}^2(u,v,\eta) &= \frac{V^2}{(X^2Y)^2}\int \frac{\mathrm{d}^3k'}{(2\pi)^3}\frac{\mathrm{d}^3k''}{(2\pi)^3}\; I(\mathbf{k}')I^*(\mathbf{k}'')A_p(\mathbf{k}-\mathbf{k}')\; A_p^*(\mathbf{k}-\mathbf{k}''),\\
\widetilde{V}^2(u,v,\eta) &=(\frac{2k_B}{\lambda^2})^2\frac{V^2}{(X^2Y)^2}\int \frac{\mathrm{d}^3k'}{(2\pi)^3}\frac{\mathrm{d}^3k''}{(2\pi)^3}\; T(\mathbf{k}')T^*(\mathbf{k}'')A_p(\mathbf{k}-\mathbf{k}')\; A_p^*(\mathbf{k}-\mathbf{k}''),
\end{aligned}
\tag{P.20}
$$

Now we handle the term TT^*, consider the cosmology observation, we will treat it as $< TT^* >$:

$$
< T(k')T^*(k'') >= \int \mathrm{d}^3x\mathrm{d}^3y\; < T(x)T(y) > e^{-ik'x}e^{ik''y}. \tag{P.21}
$$

Changing the variable $r = x - y$ so that $\mathrm{d}^3r = \mathrm{d}^3x$ and $x = r + y$, we get:

$$
\begin{aligned}
< T(k')T^*(k'') > &= \int \mathrm{d}^3r \int \mathrm{d}^3y\, \xi(r)e^{-ik'(r+y)}e^{ik''y}\\
&= \int \mathrm{d}^3r\, \xi(r)e^{-ik'r}\int \mathrm{d}^3y\, e^{i(k'-k'')y}\\
&= \widehat{P}(k')\frac{(2\pi)^3}{V}\delta(k'-k'').
\end{aligned}
\tag{P.22}
$$

Combine the Eq. P.20 and this equation, we get:

$$
\begin{aligned}
\widetilde{V}^2(u,v,\eta) =&(\frac{2k_B}{\lambda^2})^2\frac{V^2}{(X^2Y)^2}\int \frac{\mathrm{d}^3k'}{(2\pi)^3}\frac{\mathrm{d}^3k''}{(2\pi)^3}\; \widehat{P}(k')\frac{(2\pi)^3}{V}\delta(k'-k'')\\
&\times A_p(\mathbf{k}-\mathbf{k}')\; A_p^*(\mathbf{k}-\mathbf{k}''),\\
=&(\frac{2k_B}{\lambda^2})^2\frac{V}{(X^2Y)^2}\int \frac{\mathrm{d}^3k'}{(2\pi)^3}\; \widehat{P}(k')A_p(\mathbf{k}-\mathbf{k}')\\
&\times \int \mathrm{d}^3k''\delta(k'-k'')A_p^*(\mathbf{k}-\mathbf{k}'')\;,\\
=&(\frac{2k_B}{\lambda^2})^2\frac{V}{(X^2Y)^2}\int \frac{\mathrm{d}^3k'}{(2\pi)^3}\; \widehat{P}(k')A_p(\mathbf{k}-\mathbf{k}')A_p^*(\mathbf{k}-\mathbf{k}'),\\
=&(\frac{2k_B}{\lambda^2})^2\frac{V}{(X^2Y)^2}\int \frac{\mathrm{d}^3k'}{(2\pi)^3}\; \widehat{P}(k')|A_p(\mathbf{k}-\mathbf{k}')|^2\,.
\end{aligned}
\tag{P.23}
$$

In fact, this is our final result, **we should not remove the term $\widehat{P}$ from the integral!** In HERA paper, they convolution the $P_(k)$ and window function W, here W is partially come from A_p^2.

Continue the theory analysis, the k of A_p^2 is narrow enough, and the $P(k)$ will smooth enough in the range, so we remove the term from integral, we get:

$$
\widetilde{V}^2(u,v,\eta) = (\frac{2k_B}{\lambda^2})^2\frac{V}{(X^2Y)^2}\widehat{P}(k)\int \frac{\mathrm{d}^3k'}{(2\pi)^3}\; |A_p(\mathbf{k}-\mathbf{k}')|^2\,. \tag{P.24}
$$

Changing the variable $\mathbf{k} - \mathbf{k}' = \mathbf{q}$ again, and using Parseval's theorem, we have:

$$\begin{aligned}\widetilde{V}^2(u,v,\eta) =&(\frac{2k_B}{\lambda^2})^2\frac{V}{(X^2Y)^2}\widehat{P}(k)\int\frac{\mathrm{d}^3q}{(2\pi)^3}\,|A_p(\mathbf{q})|^2\\ =&(\frac{2k_B}{\lambda^2})^2\frac{1}{(X^2Y)^2}\widehat{P}(k)\int\mathrm{d}^3r\,|A_p(\mathbf{r})|^2\\ =&(\frac{2k_B}{\lambda^2})^2\frac{1}{X^2Y}\widehat{P}(k)\int\mathrm{dld}m\mathrm{d}\nu\,|A_p(l,m,\nu)|^2\\ =&(\frac{2k_B}{\lambda^2})^2\frac{B}{X^2Y}\widehat{P}(k)\int\mathrm{dld}m\,|A_p(l,m)|^2\,.\end{aligned} \tag{P.25}$$

Adopt $\Omega_{pp} = \int \mathrm{dld}m\,|A_p(l,m)|^2$ and $\Omega_p = \int \mathrm{dld}m\,A_p(l,m)$, The final result is:

$$\begin{aligned}\widetilde{V}^2_{21}(u,v,\eta) &= \left(\frac{2k_\mathrm{B}}{\lambda^2}\right)^2\frac{B}{X^2Y}\,\widehat{P}_{21}(\mathbf{k})\int\mathrm{dld}m\,|A_p(l,m)|^2\\ &= \left(\frac{2k_\mathrm{B}}{\lambda^2}\right)^2\frac{B}{X^2Y}\,\frac{2\pi^2}{k^3}\widehat{\Delta}_{21}(\mathbf{k})\int\mathrm{dld}m\,|A_p(l,m)|^2\\ &= \left(\frac{2k_\mathrm{B}}{\lambda^2}\right)^2\frac{\Omega_{pp}\,B}{X^2Y}\,\frac{2\pi^2}{k^3}\widehat{\Delta}_{21}(\mathbf{k})\,.\end{aligned} \tag{P.26}$$

2.3 Sensitivity of Power Spectrum

Starting from the power spectral density of white thermal noise $P(\nu) = k_\mathrm{B}T_\mathrm{sys}\,[\mathrm{W\,Hz^{-1}}]$, the total noise power in a band $\Delta\nu$ is $P = k_\mathrm{B}T_\mathrm{sys}\Delta\nu$. Nyquist sampling requires two independent real samples per Hz, so in an integration time τ we obtain $N_\mathrm{samp} = 2\Delta\nu\tau$ independent samples. Averaging reduces the rms by $\sqrt{N_\mathrm{samp}}$ because of the **Incoherent Measurement** property we will talk about later, so, for a channel width $\Delta\nu$ and integration time τ with system temperature T_{sys} observation:

$$\sigma_T = \frac{T_\mathrm{sys}}{\sqrt{2\Delta\nu\,\tau}} \qquad \text{unit [T]}\,. \tag{P.27}$$

For flux

$$V_\mathrm{rms} = \frac{2k_B\,T_\mathrm{sys}}{A_e\,\sqrt{2\Delta\nu\,\tau}} \quad \text{unit [Jy]} \quad \text{(Morales 2005 Eq. 5)} \tag{P.28}$$

Use antenna theorem $A_e\Omega_p = \lambda^2$ (very simple $\Omega_p \sim \theta^2 \sim (\frac{\lambda}{D})^2$, $A_e \sim \mathrm{d}^2$), we have:

$$V_\mathrm{rms} = \frac{2k_B}{\lambda^2}\,T_\mathrm{sys}\,\Omega_p\frac{1}{\sqrt{2\Delta\nu\,\tau}}. \tag{P.29}$$

In real observation using DFT we have $\widetilde{V}(\eta) = \sum_i V(\nu_i)\,\Delta\nu\, e^{2\pi i\eta\nu_i}$, for white noise, the value independent on phase, we can ignore the phase term $e^{2\pi i\eta\nu_i}$, we have $N = B/\Delta\nu$ measurements, the rms of the sum is $\sqrt{N}$ times $V_{\rm rms}$:

$$\widetilde{V}_N = V_{\rm rms}\sqrt{\frac{B}{\Delta\nu}}\Delta\nu = \frac{2k_B T_{\rm sys}}{\lambda^2}\,\frac{\Omega_p\,\Delta\nu}{\Delta\nu}\sqrt{\frac{B}{2\tau}}\,. \tag{P.30}$$

Consider the relation $T_{rms}\sqrt{2B\tau} = T_{sys}$, the counts of measurement same to the section start, it just is $\sqrt{N} = \sqrt{B/\Delta\nu}\sqrt{2\Delta\nu\tau}$, we get:

$$\begin{aligned}\widetilde{V}_N &= \frac{2k_B}{\lambda^2}\,T_{\rm rms}\,\Omega_p\,B \qquad \text{(Parsons 2012 Eq. 14)}\;;\\ \widetilde{V}_N^2 &= (\frac{2k_B}{\lambda^2})^2\,T_{\rm sys}^2\,\frac{\Omega_p^2\,B}{2\tau}\;.\end{aligned} \tag{P.31}$$

Combine Eq. P.13 and Eq. P.31, we get:

$$\begin{aligned}\Delta_N^2(\mathbf{k}) &= X^2 Y\frac{k^3}{2\pi^2}\frac{\Omega}{2\tau}T_{sys}^2\\ \Omega = \frac{\Omega_p^2}{\Omega_p p} &= \frac{[\int {\rm d}l{\rm d}m\, A_p(l,m)]^2}{\int {\rm d}l{\rm d}m\,|A_p(l,m)|^2}\end{aligned} \tag{P.32}$$

2.3.1 coherent or not

In 21cmSense, we can see the sensitivity written as:

$$\Delta(k) = \frac{D_c^2 c(1+z)^2}{\nu_{21}H_0E(z)}\frac{\Omega_p^2}{\Omega_{pp}}T_{sys}^2\frac{1}{2t_{\rm total} * \sqrt{nLSTbins}}\frac{k^3}{2\pi^2}\,. \tag{P.33}$$

Why the term 2τ become $2t_{\rm total} * \sqrt{nLSTbins}$? In fact, it is just from the incoherent measurement.

We imagine a very long time observation in which the interferometer scans the same **field/chunk** of sky every night. In practice, the array periodically forms (in hardware) and stores the corresponding data products; let each dump span a **time duration** T, while the instrument integration time is τ. For an incoherent measurement, we absorb the effect of repeated measurements into the **effective observing time**, so the time-dependent factor in the sensitivity is $\sqrt{2\tau}\sqrt{\frac{T}{\tau}} = \sqrt{2T}$. For an array with many baselines, in the T we can treat observations from baselines of the same length as repeated measurements, absorbing the number of such baselines $N_{\rm total}$ into the time term, we get $\sqrt{2T}\sqrt{N_{\rm total}} = \sqrt{2T_{\rm total}}$, this is the finally time term in the one field/chunk. In different days, we can also treat the observation from same field of a day as repeated measurements, absorbing again the number of such days $N_{\rm day}$ into the time term, we get $\sqrt{2T_{\rm total}}\sqrt{N_{\rm day}} = \sqrt{2t_{\rm total}}$, the true different occurs at the different fields/chunks, we need to calculate the square visibility first and than average the fields/chunks, but why??? Or I should say I can get the same fields/chunks average because the baselines are probing the same sky, but why we can average different fields/chunks in square visibility? Anyway, assuming the observing time per day is t_{day}, and the counts of repeated measurements in square visibility is $\frac{t_{\text{day}}}{T} = nLSTbins$, note the LST bins is just means the field/chunk of sky. In summary, the time term in square visibility is $2t_{\text{total}}\sqrt{nLSTbins}$.

Last, we copy the intepretation of 21cmSense here:

The three parameters all have to do with how much time we are observing for. The `integration_time` specifies the number of seconds that the telescope integrates observations to yield a single snapshot. All data within this time frame is averaged coherently (per baseline). In addition, observations falling within a single uv cell and within an LST-bin are averaged together coherently. By default, an LST-bin is considered to be the length of time it takes for a point on the sky to travel through the FWHM of the beam. This is an approximation. by setting `observation_duration` (in minutes). The interpretation is that within this time-frame, the sky has not changed significantly, and that if a baseline stays in the same uv cell (remember, its uv co-ordinate changes with time), it should be averaged coherently. The `time_per_day` then effectively gives the number of such LST-bins that are observed during the night. Different LST-bins are averaged incoherently, even in the same uv cell. Finally, `n_days` gives the total number of days for which these kinds of observations are averaged. Each day, the same sky appears again, and these are assumed to be able to be added coherently.

Quantity	Expression	Units
σ_T	$T_{\text{sys}}/\sqrt{2\Delta\nu\tau}$	K
V_{rms}	$\dfrac{2k_{\text{B}}T_{\text{sys}}}{A_e\sqrt{2\Delta\nu\tau}}$	Jy
$\tilde{V}_N$	$(2k_{\text{B}}/\lambda^2)\, T_{\text{rms}}\, \Omega_p B$	Jy Hz
$\Delta^2_N(k)$	$X^2 Y\, \dfrac{k^3}{2\pi^2}\, \dfrac{\Omega}{\tau}\, T^2_{\text{sys}}$	mK^2

3 Foreground Wedge

We work in the flat-sky limit. We have defined $D_c(\nu)$ is the comoving distance at frequency ν, and $\mathbf{r}_\perp = D_c\,\theta$. For a single baseline $\mathbf{b}$, the frequency-domain visibility is

$$V(b,\nu) = \int \mathrm{d}^2 r_\perp\, I(\mathbf{r}_\perp,\nu) A_p(\mathbf{r}_\perp,\nu) \exp\left[-i2\pi\,\nu\,\frac{\mathbf{b}\cdot\mathbf{r}_\perp}{c\,D_c(\nu)}\right]. \tag{F.1}$$

Like the handle of Eq. P.24, we assume $A_p(\mathbf{r}_\perp, \nu) \approx A_p(\mathbf{r}_\perp)$. Review the Eq. B.28 and Eq. B.33, we have:

$$\begin{aligned}
\widetilde{V}(\mathbf{b},\eta) &= \int \mathrm{d}\nu\, V(\mathbf{b},\nu) e^{-i2\pi\eta\nu} \\
V(\mathbf{b},\nu) &= \iint I(\theta,\nu)\, A_p(\theta,\nu)\, e^{-i2\pi\nu\, \mathbf{b}\cdot\theta/c}\, \mathrm{d}^2\theta \\
\widetilde{V}(\mathbf{b},\eta) &= \int \mathrm{d}^2\theta \mathrm{d}\nu\, I(\theta,\nu)\, A_p(\theta,\nu)\, e^{-i2\pi\nu\,(\mathbf{b}\cdot\theta/c+\eta)} \\
&= \frac{1}{X^2}\int \mathrm{d}^2 r_\perp \mathrm{d}\nu\, I(\mathbf{r}_\perp,\nu)\, A_p(\mathbf{r}_\perp)\, e^{-i2\pi\nu\,(\mathbf{b}\cdot\mathbf{r}_\perp/cD_c+\eta)}\,,
\end{aligned} \tag{F.2}$$

and review the Eq. B.37, we should re-note here $\Delta r_\parallel$ is corresponding to the narrow bandwidth $\Delta\nu$, it is the volume in frequency. We have $r_\parallel = -Y\nu$. Our Fourier convention is:

$$\widetilde{F}(\mathbf{k}) = \frac{1}{V}\int \mathrm{d}^3 r\, F(\mathbf{r})\, e^{-i\mathbf{k}\cdot\mathbf{r}}, \qquad F(\mathbf{r}) = V\int \frac{\mathrm{d}^3 k}{(2\pi)^3}\, \widetilde{F}(\mathbf{k})\, e^{+i\mathbf{k}\cdot\mathbf{r}}\,, \tag{F.3}$$

we get:

$$\begin{aligned}
\widetilde{V}(\mathbf{b},\eta) &= \frac{1}{X^2}\int \mathrm{d}^2 r_\perp \mathrm{d}\nu\, I(\mathbf{r}_\perp,\nu)\, A_p(\mathbf{r}_\perp)\, e^{-i2\pi\nu\,(\mathbf{b}\cdot\mathbf{r}_\perp/cD_c+\eta)} \\
&= \frac{V}{X^2}\int \frac{d^3k}{(2\pi)^3}\, I(k)\, e^{i(\mathbf{k}_\perp\cdot\mathbf{r}_\perp + k_\parallel r_\parallel)} \int \mathrm{d}^2 r_\perp \mathrm{d}\nu\, A_p(\mathbf{r}_\perp)\, e^{-i2\pi\nu\,(\mathbf{b}\cdot\mathbf{r}_\perp/cD_c+\eta)} \\
&= \frac{V}{X^2}\int \frac{d^3k}{(2\pi)^3}\, I(k) \int \mathrm{d}^2 r_\perp \mathrm{d}\nu\, e^{i(\mathbf{k}_\perp\cdot\mathbf{r}_\perp)} A_p(\mathbf{r}_\perp)\, e^{ik_\parallel Y\nu} e^{-i2\pi\nu\,(\mathbf{b}\cdot\mathbf{r}_\perp/cD_c+\eta)} \\
&= \frac{V}{X^2}\int \frac{d^3k}{(2\pi)^3}\, I(k) \int \mathrm{d}^2 r_\perp e^{i(\mathbf{k}_\perp\cdot\mathbf{r}_\perp)} A_p(\mathbf{r}_\perp) \int \mathrm{d}\nu\ \exp[-i(\frac{2\pi\nu\,(\mathbf{b}\cdot\mathbf{r}_\perp)}{cD_c} + \eta - \frac{Yk_\parallel\nu}{2\pi})],
\end{aligned} \tag{F.4}$$

and

$$\begin{aligned}
&\frac{2\pi\nu\,(\mathbf{b}\cdot\mathbf{r}_\perp)}{cD_c} + \eta - \frac{Yk_\parallel\nu}{2\pi} = 0 \\
&\mathbf{r}_\perp = -\frac{cD_c}{2\pi\mathbf{b}}(Yk_\parallel + 2\pi\eta)\,,
\end{aligned} \tag{F.5}$$

so we get:

$$\begin{aligned}
\widetilde{V}(\mathbf{b},\eta) &= \frac{V}{X^2}\int \frac{\mathrm{d}^3k}{(2\pi)^3}\, I(k) \int \mathrm{d}^2 r_\perp e^{i(\mathbf{k}_\perp\cdot\mathbf{r}_\perp)} A_p(\mathbf{r}_\perp)\, \delta_D(\frac{2\pi\nu\,(\mathbf{b}\cdot\mathbf{r}_\perp)}{cD_c} + \eta - \frac{Yk_\parallel\nu}{2\pi}) \\
&= \frac{V}{X^2}\int \frac{\mathrm{d}^3k}{(2\pi)^3}\, I(k) \int \mathrm{d}^2 r_\perp e^{i(\mathbf{k}_\perp\cdot\mathbf{r}_\perp)} A_p[-\frac{cD_c}{2\pi\mathbf{b}}(Yk_\parallel + 2\pi\eta)]\,,
\end{aligned} \tag{F.6}$$

Now we calculate the power spectrum:

$$
\begin{aligned}
\widetilde{V}^2(\mathbf{b},\eta) =&(\frac{2k_B}{\lambda^2})^2\frac{V^2}{(X^2)^2}\int\frac{\mathrm{d}^3k'}{(2\pi)^3}\frac{\mathrm{d}^3k''}{(2\pi)^3}T(k')T^*(k'') \\
&\times\Big|\int \mathrm{d}^2r_\perp e^{i(\mathbf{k}_\perp\cdot\mathbf{r}_\perp)}A_p[-\frac{cD_c}{2\pi\mathbf{b}}(Yk_\parallel+2\pi\eta)]\Big|^2 \\
=&(\frac{2k_B}{\lambda^2})^2\frac{V^2}{(X^2)^2}\int\frac{\mathrm{d}^3k'}{(2\pi)^3}\frac{\mathrm{d}^3k''}{(2\pi)^3}\widehat{P}(k')\frac{(2\pi)^3}{V}\delta(k'-k'') \\
&\times\Big|\int \mathrm{d}^2r_\perp e^{i(\mathbf{k}_\perp\cdot\mathbf{r}_\perp)}A_p[-\frac{cD_c}{2\pi\mathbf{b}}(Yk_\parallel+2\pi\eta)]\Big|^2 \\
=&(\frac{2k_B}{\lambda^2})^2\frac{V}{(X^2)^2}\int\frac{\mathrm{d}^3k''}{(2\pi)^3}\widehat{P}(k'')\Big|\int \mathrm{d}^2r_\perp e^{i(\mathbf{k}_\perp\cdot\mathbf{r}_\perp)}A_p[-\frac{cD_c}{2\pi\mathbf{b}}(Yk_\parallel+2\pi\eta)]\Big|^2,
\end{aligned}
\tag{F.7}
$$

because $A_p(\mathbf{r}) = A_p(-\mathbf{r})$, we get:

$$
\widetilde{V}^2(\mathbf{b},\eta) = (\frac{2k_B}{\lambda^2})^2\frac{V}{(X^2)^2}\int\frac{\mathrm{d}^3k}{(2\pi)^3}\widehat{P}(k)\Big|\int \mathrm{d}^2r_\perp e^{i(\mathbf{k}_\perp\cdot\mathbf{r}_\perp)}A_p[\frac{cD_c}{2\pi\mathbf{b}}(Yk_\parallel+2\pi\eta)]\Big|^2. \tag{F.8}
$$

As a simple toy model for foreground emission, suppose that $P_f(k) \propto P(k_\perp)\delta(k_\parallel)$, and the square visibility we observed have the relation: $\widetilde{V}^2(\mathbf{b},\eta) \propto \widehat{P}(k)\int A_p^2[\frac{cD_c}{2\pi\mathbf{b}}(Yk_\parallel+2\pi\eta)]$, we combine the δ_D function dominate foreground power spectrum and observation power spectrum, we get:

$$
P_f(k) \propto \int A_p^2[\frac{cD_c}{\mathbf{b}}\eta]. \tag{F.9}
$$

We know the relation $\eta = \frac{Yk_\parallel}{2\pi}$, $\mathbf{k}_\perp = \frac{2\pi\nu_0\mathbf{b}}{cD_c}$ and the Field of View is $D_c\theta$, we get:

$$
\begin{aligned}
&P_f(k) \propto \int A_p^2[\frac{cD_c}{\mathbf{b}}\eta] \propto A_p^2(\frac{Y\nu_0k_\parallel}{\mathbf{k}_\perp}) \\
&\frac{Y\nu_0k_\parallel}{\mathbf{k}_\perp} = D_c\theta = X\theta \\
&\frac{Y\nu_0}{X}\frac{k_\parallel}{\mathbf{k}_\perp} = \theta \qquad\qquad \text{beautiful relation!} \\
&\qquad\quad \frac{k_\parallel}{\mathbf{k}_\perp}\frac{c(1+z)^2\nu_0}{H_0D_cE(z)\nu_{21}} = \theta \\
&k_\parallel = \mathbf{k}_\perp\frac{H_0D_cE(z)\theta}{c(1+z)} \qquad\qquad \text{Foreground Wedge}.
\end{aligned}
\tag{F.10}
$$

We note that the foreground wedge assumes the foreground contamination is a δ-function in $k_\parallel$ and the primary beam is frequency-independent over a narrow band. But in reality, foregrounds are not "clean" and the primary beam is inevitably chromatic. Consequently, in practical simulations we typically introduce a intercept k_{buffer} to the wedge boundary to ensure the **foreground avoidance.**

$$k_{\parallel} > \frac{X\theta}{Y\nu_0}\mathbf{k}_{\perp} + k_{\text{buffer}} \tag{F.11}$$

4 Why Gaussian Beam?

It seems to naturally define a beam which pattern is gaussian, but really? In this section we will explain why we use gaussian beam.

First, as we all know the resolution of a telescope is defined by its diameter, more detail, its come from the size of Airy disk. The Airy disk is come from Fraunhofer diffraction, given a circular aperture of diameter D in the Fraunhofer limit. Define:

$$x \equiv k\,\frac{D}{2}\,\sin\theta \simeq \frac{\pi D}{\lambda}\,\theta \qquad (\theta \ll 1), \tag{4.1}$$

and write the **voltage** pattern and the **power** (primary beam) as:

$$E(\theta) = \frac{2J_1(x)}{x}, \qquad A_p(\theta) = \left[\frac{2J_1(x)}{x}\right]^2. \tag{4.2}$$

The Bessel function's expansion is:

$$J_1(x) = \sum_{m=0}^{\infty} (-1)^m \frac{x^{\,2m+1}}{2^{\,2m+1}\, m!\,(m+1)!} = \frac{x}{2} - \frac{x^3}{16} + \frac{x^5}{384} - \frac{x^7}{18432} + \cdots. \tag{4.3}$$

So the **voltage** ratio expands as:

$$\frac{2J_1(x)}{x} = 1 - \frac{x^2}{8} + \frac{x^4}{192} - \frac{x^6}{9216} + \mathcal{O}(x^8). \tag{4.4}$$

Airy power (primary beam) expansion
Squaring (4.4) we have:

$$\begin{aligned} A_p(\theta) &= \left(1 - \frac{x^2}{8} + \frac{x^4}{192} - \frac{x^6}{9216} + \cdots\right)^2 \\ &= 1 - \frac{x^2}{4} + \frac{5}{192}\,x^4 - \frac{7}{4608}\,x^6 + \mathcal{O}(x^8). \end{aligned} \tag{4.5}$$

Changing variable $x = (\pi D/\lambda)\theta$ we have:

$$A_p(\theta) = 1 - \frac{\pi^2}{4}\left(\frac{D}{\lambda}\right)^2 \theta^2 + \cdots. \tag{4.6}$$

Gaussian matching

Because the Gaussian profile seems like the Airy pattern and it is easy to calculate, so we assume a Gaussian pattern can be expansion to:

$$G(\theta) = \exp\left(-\frac{\theta^2}{2\sigma^2}\right) = 1 - \frac{\theta^2}{2\sigma^2} + \frac{\theta^4}{8\sigma^4} - \cdots . \tag{4.7}$$

Matching the second term of Eq. (4.6), we can give σ:

$$\frac{1}{2\sigma^2} = \frac{\pi^2}{4}\left(\frac{D}{\lambda}\right)^2 \quad \Longrightarrow \quad \boxed{\sigma = \frac{\sqrt{2}}{\pi}\frac{\lambda}{D} \simeq 0.45\,\frac{\lambda}{D}} \tag{4.8}$$

and therefore

$$\mathrm{FWHM_G} = 2\sqrt{2\ln 2}\,\sigma \simeq 1.06\,\frac{\lambda}{D} = 2.35\sigma \propto \lambda. \tag{4.9}$$

Now we calculate the Ω_p and Ω_{pp}, $d\Omega = \sin\theta\, d\theta\, d\phi \simeq \theta\, d\theta\, d\phi$:

$$\begin{aligned}\Omega_p &= \int G(\theta)\,\mathrm{d}\Omega = \int_0^{2\pi} \mathrm{d}\phi \int_0^{\infty} \theta\, e^{-\theta^2/(2\sigma^2)}\,\mathrm{d}\theta \\ &\xrightarrow{u=\theta^2/(2\sigma^2)} 2\pi \int_0^{\infty} e^{-u}\,\sigma^2\,\mathrm{d}u \qquad \big(\mathrm{d}u = \frac{\theta}{\sigma^2}\mathrm{d}\theta \Rightarrow \theta d\theta = \sigma^2 \mathrm{d}u\big) \\ &= 2\pi\sigma^2\Big[-e^{-u}\Big]_0^{\infty} = 2\pi\sigma^2 \approx 1.133\,\mathrm{FWHM}^2.\end{aligned} \tag{4.10}$$

$$\begin{aligned}\Omega_{pp} &= \int G^2(\theta)\,\mathrm{d}\Omega = \int_0^{2\pi} \mathrm{d}\phi \int_0^{\infty} \theta\, e^{-\theta^2/\sigma^2}\,\mathrm{d}\theta \\ &\xrightarrow{v=\theta^2/\sigma^2} 2\pi \int_0^{\infty} e^{-v}\,\frac{\sigma^2}{2}\,\mathrm{d}v \qquad \big(\mathrm{d}v = \frac{2\theta}{\sigma^2}\mathrm{d}\theta \Rightarrow \theta\mathrm{d}\theta = \frac{\sigma^2}{2}\mathrm{d}v\big) \\ &= \pi\sigma^2 = \frac{1}{2}\,\Omega_P.\end{aligned} \tag{4.11}$$

5 Polarzation

to add (video of SKA & 21cm summer school from Xiaohui Sun)

6 21cmSense

6.1 Foundation of 21cmSense

1. Given the antenna locations of an interferometer, compute the uv coordinates of all baselines.
2. Segement a uv plane into discrete cells, each of which has roughly the same width as $\tilde{A}_p$ [*i.e. the Fourier-Transform of the beam*]. Roughly speaking, this width is the inverse of the instantaneous field of view of an

antenna. Define a three-dimensional $uv\eta$ space, extending the uv plane into a third dimension of Fourier space. Each cell in this dimension should have an extent of $1/B$, where B is the frequency bandwidth of the observation.
3. Simulate the movement of baselines through the uv cells over a sidereal day of observations (or over however much time the telescope is being operated per sidereal day). Record how much time t_u each baseline spends in each uv cell.
4. Assuming a roughly constant noise contribution across frequency channels, the uncertainty (standard deviation) from measurement noise for a given $uv\eta$ cell is given by

$$\frac{D_c^2 c(1+z)^2}{\nu_{21} H_0 E(z)} \frac{\Omega_p^2}{\Omega_{pp}} \frac{T_{\rm sys}^2}{2t_u} \tag{4.1}$$

where

$$\Omega_p \equiv \int d\Omega A_p(\hat{\mathbf{r}}); \quad \Omega_{pp} \equiv \int d\Omega A_p^2(\hat{\mathbf{r}}) \tag{4.2}$$

and $T_{\rm sys}$ is the **system temperature**, which is the sum of the sky brightness temperature and the reciever temperature. This is the uncertainty *for one baseline*. If multiple baselines spend time in a given uv cell, the integration times of each baseline must be added together before being inserted in the above equation. **1.** Divide the uncertainty computed above by the total number of sidereal days $t_{\rm days}$ of observation. Note that the standard deviation scales $t_{\rm days}^{-1}$ rather than $t_{\rm days}^{-1/2}$ because visibility measurements repeat every day and thus can be coherently averaged prior to the squaring step of forming a power spectrum. **2.** Associate each $uv\eta$ cell with its location in **k** space using

$$\mathbf{k}_\perp = \frac{2\pi\mathbf{u}}{D_c}, \tag{4.3}$$

$$k_{||} = \frac{2\pi\nu_{21} H_0 E(z)}{c(1+z)^2}\eta \tag{4.4}$$

7. For high signal-to-noise measurements it is necessary to account for cosmic variance which contributes a standard deviation of $P(\mathbf{k})$ (i.e. the cosmological signal itself) to every $uv\eta$ cell. This is added to the instrumental uncertainty.

8. Assuming that different cells in the 3D **k** space are averaged together with an inverse variance weighting to obtain $P(k_\perp, k_{||})$ or $P(k)$, the corresponding errors can be combined in inverse quadrature. In other words, if $\epsilon(\mathbf{k})$ is the standard deviation for a **k**-space cell, then the final averaged error bars are given by $1/\sqrt{\sum \epsilon^{-2}}$, where the sum is over all the **k**-space cells that contribute to a particular bin in $(k_\perp, k_{||})$ or k.

6.2 Code & Theroy

In this section we will see how 21cmSense work by examine the source code.

6.2.1 Beam

We only read the Gaussian beam in this section, the simply source codes are:

```python
class GaussianBeam(PrimaryBeam):


    def wavelength(self) -> un.Quantity[un.m]:
        """The wavelength of the observation."""
        return (cnst.c / self.frequency).to("m")

    def dish_size_in_lambda(self) -> float:
        """The dish size in units of wavelengths."""
        return (self.dish_size / (cnst.c / self.frequency)).to("").value
```

Here `dish_size_in_lambda` returns D/λ, it is the wavenumber in the diameter, is also corresponding to the $1/\sigma$, see Eq. 4.8. The 21cmSense use `dish_size_in_lambda` as the "appropriate resolution" for a uv cell as below, the code `uv_resolution`. Intuitively, the primary-beam main lobe (Airy disk) angular width scales as $\sim \lambda/D$, so its Fourier width scales as $\sim D/\lambda$. Therefore, setting the bin size Δu at the level of D/λ effectively matches the uv cell size to the width of $\sim A_p$, allowing measurements within the same uv cell to be coherently (just repeated measurement before square) averaged.

```python
    def uv_resolution(self) -> un.Quantity[1 / un.radian]:
        """The appropriate resolution of a UV cell given the beam size."""
        return self.dish_size_in_lambda

    def area(self) -> un.Quantity[un.steradian]:
        """The integral of the beam over angle, in sr."""
        return 1.13 * self.fwhm**2

    def width(self) -> un.Quantity[un.radian]:
        """The width of the beam (i.e. sigma), in radians."""
        return un.rad * 0.45 / self.dish_size_in_lambda

    def fwhm(self) -> un.Quantity[un.radians]:
        """The full-width half maximum of the beam."""
        return 2.35 * self.width

    def sq_area(self) -> un.Quantity[un.steradian]:
```

```
        """The integral of the squared beam, in sr."""
        return self.area / 2

    def first_null(self) -> un.Quantity[un.radians]:
        """The angle of the first null of the beam.
        .. note:: The Gaussian beam has no null, and in this case we use the first null for an airy disk.
        """
        return un.rad * 1.22 / self.dish_size_in_lambda
```

Here we only note the `width`, it is just σ as same as the Eq. 4.8 we derived. Other codes is clearly so we will not continue to read.

6.2.2 Power Spectrum

Sensitivity from code is class `PowerSpectrum` and function `calculate_sensitivity_2d`. Here we only confirm the thermal noise, we read the codes step by step:

```
    def k1d(self) -> tp.Wavenumber:
        """1D array of wavenumbers for which sensitivities will be generated."""
        delta = (
            conv.dk_deta(
                self.observation.redshift,
                self.cosmo,
                approximate=self.observation.use_approximate_cosmo,
            )
            / self.observation.bandwidth
        )
        dv = delta.value
        return np.arange(dv, dv * self.observation.n_channels, dv) * delta.unit
```

The true k_{1d} should be $\sqrt{k_\perp^2 + k_\parallel^2}$ in every different baseline (different $k_\perp$), for all baseline the $k_\parallel$ sequence are same. But the $k_\perp$ are too small compare to the $k_\parallel$, even the shortest baseline, to the k_{1d} is dominated by $k_\parallel$. Specifically, 21cmSense computes k_{1d} sequence for each $k_\perp$. It then cross-references this with the $k_\parallel$ sequence, determines for each k_{1d} which $k_\perp$ is most closely, and assigns it to that $k_\parallel$ bin.

```
    def power_normalisation(self, k: tp.Wavenumber) -> float:

        return (
            self.X2Y
            * self.observation.observatory.beam.b_eff
            * self.observation.bandwidth
            * k**3
            / (2 * np.pi**2)
        ).to_value("")
```

```
    def Trms(self) -> un.Quantity[un.K]:

        out = np.ones(self.total_integration_time.shape) * np.inf * self.Tsys.unit
        mask = self.total_integration_time > 0
        out[mask] = self.Tsys.to("K") / np.sqrt(
            2 * self.bandwidth * self.total_integration_time[mask]
        ).to("")
        return out

    def thermal_noise(
        self, k_par: tp.Wavenumber, k_perp: tp.Wavenumber, trms: tp.Temperature
    ) -> tp.Delta:

        k = np.sqrt(k_par**2 + k_perp**2)
        scalar = self.power_normalisation(k)

        return scalar * trms.to("mK") ** 2
```

The `power_normalisation` is just $\frac{D_c^2 c(1+z)^2}{\nu_{21} H_0 E(z)} \frac{\Omega_p^2}{\Omega_{pp}} \frac{k^3}{2\pi^2}$, `Trms` give us $T_{\rm sys}/\sqrt{2Bt_{\rm total}}$, so the `thermal_noise` gives us the result $\frac{D_c^2 c(1+z)^2}{\nu_{21} H_0 E(z)} \frac{\Omega_p^2}{\Omega_{pp}} \frac{k^3}{2\pi^2} \frac{T_{\rm sys}^2 B}{2Bt_{\rm total}}$, let B disappear, we found we only need times the result with $1/\sqrt{nLSTbins}$, we will see it soon.

6.2.3 Wedge & Buffer

Continue to read codes within class `PowerSpectrum`, but we start a new section to explain how the foreground wedge and horizon buffer work.

```
    def horizon_limit(self, umag: float) -> tp.Wavenumber:
        horizon = (
            conv.dk_deta(
                self.observation.redshift,
                self.cosmo,
                approximate=self.observation.use_approximate_cosmo,
            )
            * umag
            / self.observation.frequency
        )
        # calculate horizon limit for baseline of length umag
        if self.foreground_model in ["moderate", "pessimistic"]:
            return horizon + self.horizon_buffer
        elif self.foreground_model in ["optimistic"]:
            return horizon * np.sin(self.observation.observatory.beam.first_null / 2)
```

Here we transform it to equation:

$$\mathrm{Hor} = \frac{\mathrm{d}k}{\mathrm{d}\eta}\frac{\mathbf{u}}{\nu_0} = \frac{2\pi\mathrm{d}\nu}{\mathrm{d}D_c}\frac{X\mathbf{k}_\perp}{2\pi\nu_0} = \frac{X\mathbf{k}_\perp}{Y\nu_0} \tag{4.5}$$

Combine the code `_nsamples_2d` presented below, we can see the wedge is:

$$\begin{aligned} \text{moderate/pessimistic}: k_\parallel &> \frac{X\mathbf{k}_\perp}{Y\nu_0}\cdot 1 + k_{\text{buffer}} = \frac{X\mathbf{k}_\perp}{Y\nu_0}\sin\theta\Big|_{\theta=\frac{\pi}{2}} + k_{\text{buffer}}\ ; \\ \text{optimistic}: k_\parallel &> \frac{X\mathbf{k}_\perp}{Y\nu_0}\sin\frac{\theta_{\text{Airy}}}{2}\ . \end{aligned} \tag{4.6}$$

Here moderate/pessimistic case is consider the worst case that the beam can trace the light from horizon.

```
def _nsamples_2d(
    self,
) -> dict[str, dict[tp.Wavenumber, un.Quantity[1 / un.mK**4]]]:

    sense = {"sample": {}}

    nonzero = np.where(self.uv_coverage > 0)
        u, v = self.observation.ugrid[iu], self.observation.ugrid[iv]
        trms = self.observation.Trms[iv, iu]

        umag = np.sqrt(u**2 + v**2)
        k_perp = umag * conv.dk_du(
            self.observation.redshift,
            self.cosmo,
            approximate=self.observation.use_approximate_cosmo,
        )

        hor = self.horizon_limit(umag)

        if k_perp not in sense["thermal"]:
            sense["thermal"][k_perp] = (
                np.zeros(len(self.observation.kparallel)) / un.mK**4
            )

        # Exclude parallel modes dominated by foregrounds
        kpars = self.observation.kparallel[self.observation.kparallel >= hor]

        if not len(kpars):
            continue

        start = np.where(self.observation.kparallel >= hor)[0][0]
```

```python
            n_inds = (self.observation.kparallel.size - 1) // 2 + 1
            inds = np.arange(start=start, stop=n_inds)

            thermal = self.thermal_noise(kpars, k_perp, trms)
            sample = self.sample_noise(kpars, k_perp)

            # The following assumes that the power spectra are averaged with inverse
            # variance weighting.
            t = 1.0 / thermal**2
            sense["thermal"][k_perp][inds] += t
            sense["thermal"][k_perp][-inds] += t

        return sense
```

```python
    def calculate_sensitivity_2d(
        self, thermal: bool = True, sample: bool = True
    ) -> dict[tp.Wavenumber, tp.Delta]:

        if thermal:
            sense = self._nsamples_2d["thermal"]
        final_sense = {}
        for k_perp in sense.keys():
            mask = sense[k_perp] > 0
            if self.systematics_mask is not None:
                mask &= self.systematics_mask(k_perp, self.observation.kparallel)

            if not np.any(mask):
                continue

            final_sense[k_perp] = np.inf * np.ones(len(mask)) * un.mK**2
            if thermal:
                total_std = thermal_std = 1 / np.sqrt(
                    self._nsamples_2d["thermal"][k_perp][mask]
                    * self.observation.n_lst_bins
                )
            final_sense[k_perp][mask] = total_std

        return final_sense
```

Finally, we can see the sensitivity indeed is

$$\Delta(k) = \frac{D_c^2 c(1+z)^2}{\nu_{21} H_0 E(z)} \frac{\Omega_p^2}{\Omega_{pp}} T_{sys}^2 \frac{1}{2t_{\text{total}} * \sqrt{nLSTbins}} \frac{k^3}{2\pi^2}\,. \tag{4.7}$$

6.3 Partice

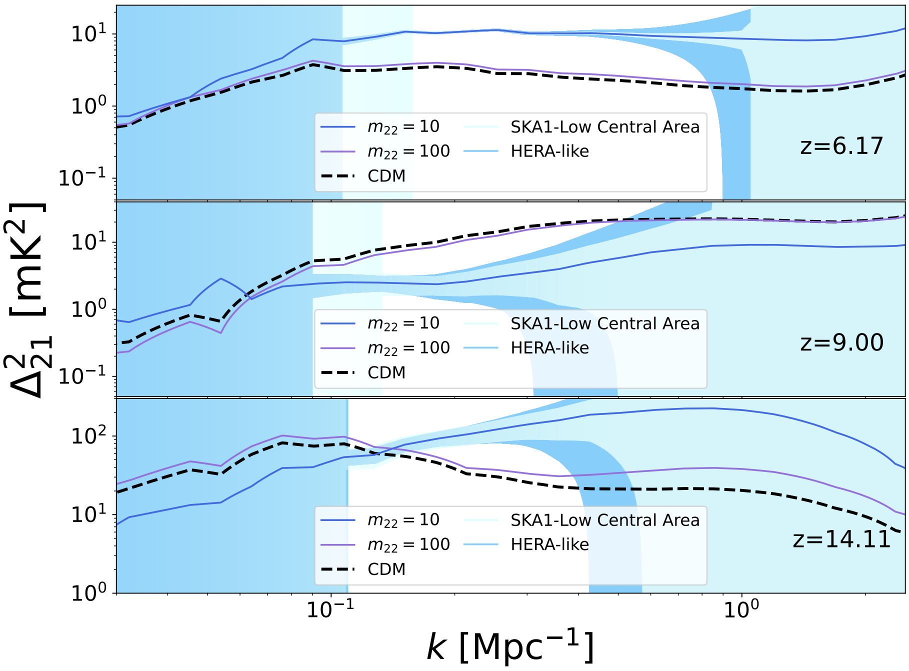

Figure 1: The relationship between wavenumber k and the power spectrum at different redshifts. The light blue and dark blue shaded regions represent the expected 1σ noise levels for SKA1-Low central area and HERA radio arrays respectively with separate predictions for $m_{22} = 10$.

A Damn you! Fourier factor

Review the Fourier convention we used:

$$\widetilde{F}(\mathbf{k}) = \frac{1}{V}\int \mathrm{d}^3r\, F(\mathbf{r})\, e^{-i\mathbf{k}\cdot\mathbf{r}}, \qquad F(\mathbf{r}) = V\int \frac{\mathrm{d}^3k}{(2\pi)^3}\, \widetilde{F}(\mathbf{k})\, e^{+i\mathbf{k}\cdot\mathbf{r}}\ , \tag{A.1}$$

Now we confirm the convolution theorem and the Parseval's theorem can carry only one Fourier factor.

A useful identity consistent with Eq. (A.1) is the Dirac–Kronecker symbol:

$$\int_V \mathrm{d}^3r\, e^{i(\mathbf{k}-\mathbf{k}')\mathbf{r}} \;=\; (2\pi)^3\, \delta_D^{(3)}(\mathbf{k}-\mathbf{k}')\, \frac{1}{V} \;=\; \frac{(2\pi)^3}{V}\, \delta_D(\mathbf{k}-\mathbf{k}'). \tag{A.2}$$

Convolution theorem

Let $\widetilde{V}(\mathbf{k}) \equiv \int \mathrm{d}^3r\, A(\mathbf{r})T(\mathbf{r})\, e^{-i\mathbf{k}\cdot\mathbf{r}}$. Expanding only one term using the inverse transform in Eq. (A.1) gives

$$\begin{aligned}
\widetilde{V}(\mathbf{k}) &= \int \mathrm{d}^3 r\, A(\mathbf{r})\, e^{-i\mathbf{k}\cdot\mathbf{r}} \left[V \int \frac{\mathrm{d}^3 k'}{(2\pi)^3}\, \widetilde{T}(\mathbf{k}')\, e^{+i\mathbf{k}'\cdot\mathbf{r}} \right] \\
&= V \int \frac{\mathrm{d}^3 k'}{(2\pi)^3}\, \widetilde{T}(\mathbf{k}') \underbrace{\int \mathrm{d}^3 r\, A(\mathbf{r})\, e^{-i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}}}_{\widetilde{A}(\mathbf{k}-\mathbf{k}')} \\
&= \boxed{V \int \frac{\mathrm{d}^3 k'}{(2\pi)^3}\, \widetilde{A}(\mathbf{k}-\mathbf{k}')\, \widetilde{T}(\mathbf{k}')}.
\end{aligned} \tag{A.3}$$

Thus with our convention, **the product $\to$ convolution carries exactly one Fourier factor V.**

Another view (expand both and use (A.2)**):**

$$\begin{aligned}
\widetilde{V}(\mathbf{k}) &= \int \mathrm{d}^3 r \left[V \int \frac{\mathrm{d}^3 k'}{(2\pi)^3} \widetilde{A}(\mathbf{k}') e^{+i\mathbf{k}'\cdot\mathbf{r}} \right] \left[V \int \frac{\mathrm{d}^3 k''}{(2\pi)^3} \widetilde{T}(\mathbf{k}'') e^{+i\mathbf{k}''\cdot\mathbf{r}} \right] e^{-i\mathbf{k}\cdot\mathbf{r}} \\
&= V^2 \int \frac{\mathrm{d}^3 k'}{(2\pi)^3} \frac{\mathrm{d}^3 k''}{(2\pi)^3}\, \widetilde{A}(\mathbf{k}')\widetilde{T}(\mathbf{k}'') \underbrace{\int_V \mathrm{d}^3 r\, e^{i(\mathbf{k}'+\mathbf{k}''-\mathbf{k})\cdot\mathbf{r}}}_{\frac{(2\pi)^3}{V}\, \delta_D(\mathbf{k}'+\mathbf{k}''-\mathbf{k})} \\
&= V^2 \int \frac{\mathrm{d}^3 k'}{(2\pi)^3} \frac{\mathrm{d}^3 k''}{(2\pi)^3}\, \widetilde{A}(\mathbf{k}')\widetilde{T}(\mathbf{k}'')(2\pi)^3 \frac{1}{V}\, \delta_D(\mathbf{k}'+\mathbf{k}''-\mathbf{k}) \\
&= V \int \frac{\mathrm{d}^3 k''}{(2\pi)^3} \mathrm{d}^3 k'\, \widetilde{A}(\mathbf{k}')\widetilde{T}(\mathbf{k}'')\, \delta_D(\mathbf{k}'+\mathbf{k}''-\mathbf{k}) \\
&= V \int \frac{\mathrm{d}^3 k''}{(2\pi)^3} \mathrm{d}^3 k'\, \widetilde{A}(\mathbf{k}')\widetilde{T}(\mathbf{k}'')\, \delta_D(\mathbf{k}'-(\mathbf{k}-\mathbf{k}'')) \\
&= V \int \frac{\mathrm{d}^3 k''}{(2\pi)^3} \widetilde{T}(\mathbf{k}'') \int \mathrm{d}^3 k'\, \widetilde{A}(\mathbf{k}')\, \delta_D(\mathbf{k}'-(\mathbf{k}-\mathbf{k}'')) \\
&= V \int \frac{\mathrm{d}^3 k''}{(2\pi)^3}\, \widetilde{T}(\mathbf{k}'')\, \widetilde{A}(\mathbf{k}-\mathbf{k}''),
\end{aligned} \tag{A.4}$$

again leaving a **single** factor V.

Parseval's theorem

For given two fields f and g,

$$\begin{aligned}
\int \mathrm{d}^3 r\, f(\mathbf{r})\, g^*(\mathbf{r}) &= \int \mathrm{d}^3 r \left[V \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\, \widetilde{f}(\mathbf{k})\, e^{+i\mathbf{k}\cdot\mathbf{r}} \right] g^*(\mathbf{r}) \\
&= V \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\, \widetilde{f}(\mathbf{k}) \underbrace{\int \mathrm{d}^3 r\, g^*(\mathbf{r})\, e^{+i\mathbf{k}\cdot\mathbf{r}}}_{[\widetilde{g}(\mathbf{k})]^*} \\
&= \boxed{V \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\, \widetilde{f}(\mathbf{k})\, \widetilde{g}^*(\mathbf{k})}.
\end{aligned} \tag{A.5}$$

Setting $g = f$ we get:

$$\frac{1}{V}\int \mathrm{d}^3 r\, |f(\mathbf{r})|^2 = \int \frac{\mathrm{d}^3 k}{(2\pi)^3}\, |\widetilde{f}(\mathbf{k})|^2. \tag{A.6}$$

B 2PCF with $k = 0$

As we know, the definition of the two-point correlation function in 21cm Cosmology should be $\delta\overline{T}^2\langle\delta\delta\rangle$. But we use $\widehat{P}(\mathbf{k}) \equiv \langle|\delta\widetilde{T}(\mathbf{k})|^2\rangle$(see the symbol definition below), what is the difference? This section we will clarify some thing of 2PCF.

We begin with the symbol; note that the brightness temperature is δT, the average brightness temperature is $\delta\overline{T}$, the over-brightness temperature is $\delta T - \delta\overline{T}$, the dimensionless over-brightness temperature is $\delta \equiv \dfrac{\delta T - \delta\overline{T}}{\delta\overline{T}}$, we will see the difference of the three cases 2PCF P_{TT}, $P_{\delta T}$ and $P_{\delta\delta}$. First we calculate the second case $P_{\delta T}$:

$$\begin{aligned} P_{\delta T}(k) =& \langle(\delta T(k') - \delta\overline{T})(\delta T(k'') - \delta\overline{T})\rangle = \langle\delta T(k')\delta T(k'')\rangle - \langle\delta\overline{T}\delta\overline{T}\rangle \\ =& P_{TT}(k) - P_{TT}(k=0)\,. \end{aligned} \tag{B.1}$$

So the relation of P_{TT} and $P_{\delta T}$ is very simple: To some extent they are equivalent, especially for power spectra in radio observations—because we do not observe the zero-baseline $k = 0$ mode. And

$$\begin{aligned} P_{\delta\delta}(k) =& \langle\frac{\delta T(k') - \delta\overline{T}}{\delta\overline{T}}\frac{\delta T(k'') - \delta\overline{T}}{\delta\overline{T}}\rangle = \langle(\delta T(k') - \delta\overline{T})(\delta T(k'') - \delta\overline{T})\rangle/\delta\overline{T}^2 \\ =& P_{\delta T}(k)/\delta\overline{T}^2 \\ P_{\delta T}(k) =& \delta\overline{T}^2 P_{\delta\delta}(k)\,, \end{aligned} \tag{B.2}$$

so all thing have clarified. More detail see my Note for 2PCF.