Microphone-Array Signal Processing

José A. Apolinário Jr. and Marcello L. R. de Campos

{apolin},{mcampos}@ieee.org

LPS COPPE/Poli/UFRJ

IME – Lab. Processamento Digital de Sinais

Outline

1. Introduction and Fundamentals
2. Sensor Arrays and Spatial Filtering
3. Optimal Beamforming
4. Adaptive Beamforming
5. DoA Estimation with Microphone Arrays

2. Sensor Arrays and Spatial Filtering

2.1 Wavenumber-Frequency Space

Space-time Fourier Transform

The four-dimensional Fourier transform of the space-time signal $s(\mathbf{x}, t)$ is given by

$$S(\mathbf{k}, \omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(\mathbf{x}, t) e^{-j(\omega t - \mathbf{k}^T \mathbf{x})} d\mathbf{x} dt$$

temporal frequency

spatial frequency: wavenumber

Space-time Fourier Transform

The four-dimensional Fourier transform of the space-time signal $s(\mathbf{x}, t)$ is given by

$$S(\mathbf{k}, \omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(\mathbf{x}, t) e^{-j(\omega t - \mathbf{k}^T \mathbf{x})} d\mathbf{x}\, dt$$

$$s(\mathbf{x}, t) = \frac{1}{(2\pi)^4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} S(\mathbf{k}, \omega) e^{j(\omega t - \mathbf{k}^T \mathbf{x})} d\mathbf{k}\, d\omega$$

Space-time Fourier Transform

We have already concluded that if the space-time signal is a propagating waveform such that $s(\mathbf{x}, t) = s(t - \boldsymbol{\alpha}_0^T \mathbf{x})$, then its Fourier transform is equal to

$$S(\mathbf{k}, \omega) = S(\omega)\delta(\mathbf{k} - \omega\boldsymbol{\alpha}_0)$$

Remember the nonperiodic propagating wave Fourier transform?

This means that $s(\mathbf{x}, t)$ only has energy along the direction of $\mathbf{k} = \mathbf{k}_0 = \omega\boldsymbol{\alpha}_0$ in the wavenumber-frequency space.

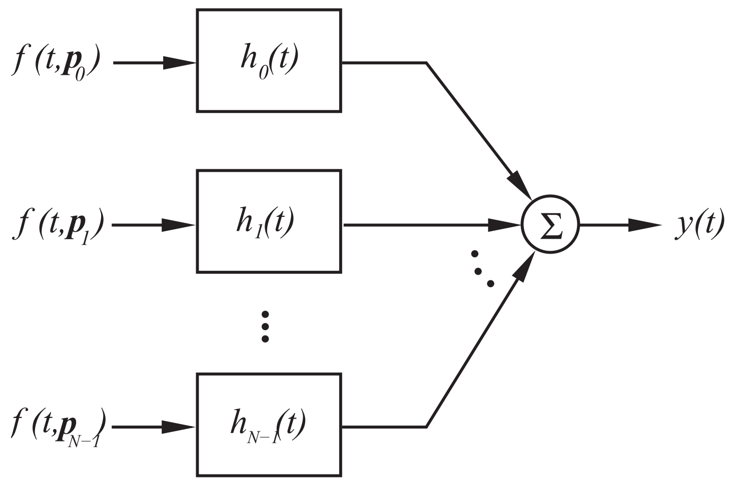
2.2 Frequency-Wavenumber (WN) Response and Beam patterns (BP)

Signals at the sensors

- An array is a set of N (isotropic) sensors located at positions $\boldsymbol{p}_n, n = 0, 1, \cdots, N-1$
- The sensors spatially sample the signal field at locations $\boldsymbol{p}_n$
- At the sensors, the set of N signals are denoted by

$$\boldsymbol{f}(t, \boldsymbol{p}) = \begin{bmatrix} f(t, \boldsymbol{p}_0) \\ f(t, \boldsymbol{p}_1) \\ \vdots \\ f(t, \boldsymbol{p}_{N-1}) \end{bmatrix}$$

Array output



$$y(t) = \sum_{n=0}^{N-1} \int_{-\infty}^{\infty} h_n(t-\tau) f_n(\tau, \boldsymbol{p}_n) d\tau$$

$$= \int_{-\infty}^{\infty} \boldsymbol{h}^T(t-\tau) \boldsymbol{f}(\tau, \boldsymbol{p}) d\tau$$

where $\boldsymbol{h}(t) = [h_o(t) \;\; h_1(t) \;\; \cdots \;\; h_{N-1}(t)]^T$

In the frequency domain, . . .

$$Y(\omega) = \int_{-\infty}^{\infty} y(t) e^{-j\omega t} dt$$
$$= \boldsymbol{H}^T(\omega)\boldsymbol{F}(\omega)$$

where

$$\boldsymbol{H}(\omega) = \int_{-\infty}^{\infty} \boldsymbol{h}(t) e^{-j\omega t} dt$$
$$\boldsymbol{F}(\omega) = \boldsymbol{F}(\omega, \boldsymbol{p}) = \int_{-\infty}^{\infty} \boldsymbol{f}(t, \boldsymbol{p}) e^{-j\omega t} dt$$

Plane wave propagating . . .

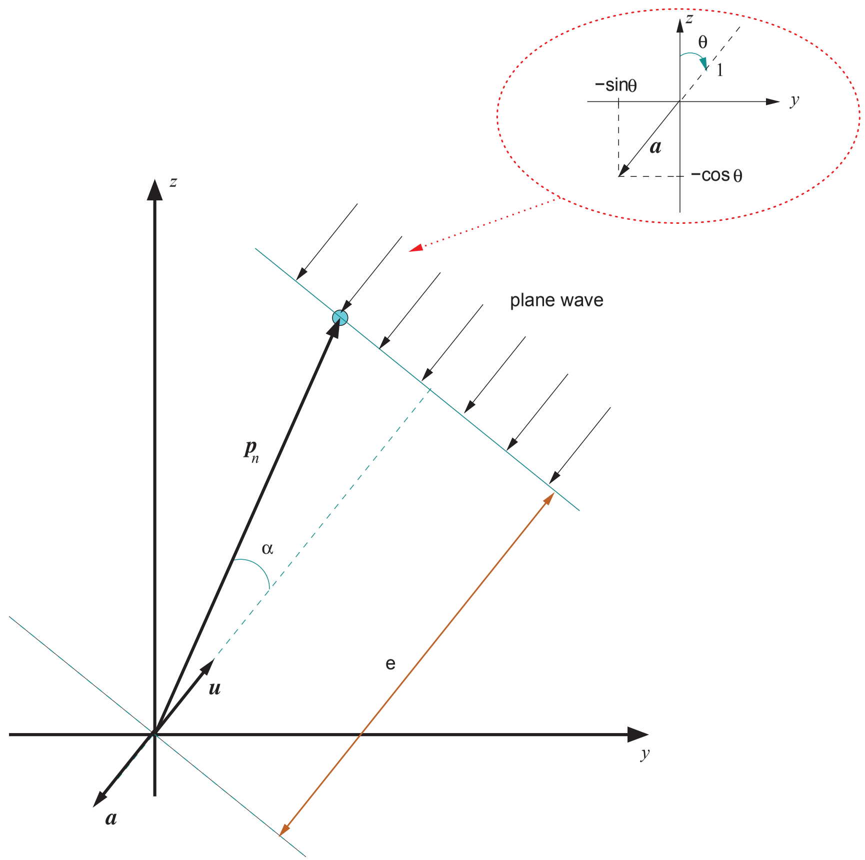
- Consider a plane wave propagating in the direction of vector $\boldsymbol{a}$:

$$\boldsymbol{a} = \begin{bmatrix} -sin\theta cos\phi \\ -sin\theta sin\phi \\ -cos\theta \end{bmatrix}$$

- If $f(t)$ is the signal that would be received at the origin, then:

$$\boldsymbol{f}(t, \boldsymbol{p}) = \begin{bmatrix} f(t - \tau_0) \\ f(t - \tau_1) \\ \vdots \\ f(t - \tau_{N-1}) \end{bmatrix}$$

Plane wave (assuming $\phi = 90^o$)



$e = ?$

$e = c\tau_n$

$\Rightarrow \tau_n = \frac{e}{c}$ BUT

$$e = \|\mathbf{p}_n\| cos(\alpha) = \underbrace{\|\mathbf{u}\|}_{=1} \|\mathbf{p}_n\| cos(\alpha)$$

$$\therefore \tau_n = -\frac{\mathbf{u}^T \mathbf{p}_n}{c} = \frac{\mathbf{a}^T \mathbf{p}_n}{c}$$

$_n$ is the time since the plane wave hits the sensor at location $\mathbf{p}_n$ until it reaches point $(0, 0)$.

Back to the frequency domain

- Then, we have:

$$\boldsymbol{F}(\omega) = \begin{bmatrix} \int_{-\infty}^{\infty} e^{-j\omega t} f(t-\tau_0)dt \\ \int_{-\infty}^{\infty} e^{-j\omega t} f(t-\tau_1)dt \\ \vdots \\ \int_{-\infty}^{\infty} e^{-j\omega t} f(t-\tau_{N-1})dt \end{bmatrix}$$

$$= \begin{bmatrix} e^{-j\omega\tau_0} \\ e^{-j\omega\tau_1} \\ \vdots \\ e^{-j\omega\tau_{N-1}} \end{bmatrix} F(\omega)$$

Definition of Wavenumber

- For plane waves propagating in a locally homogeneous medium:

$$\boldsymbol{k} = \frac{\omega}{c}\boldsymbol{a} = \frac{2\pi}{c/f}\boldsymbol{a} = \frac{2\pi}{\lambda}\boldsymbol{a} = -\frac{2\pi}{\lambda}\boldsymbol{u}$$

Wavenumber Vector ("spatial frequency")

- Note that $|\boldsymbol{k}| = \frac{2\pi}{\lambda}$
- Therefore

$$\omega\tau_n = \frac{\omega}{c}\boldsymbol{a}^T\boldsymbol{p}_n = \boldsymbol{k}^T\boldsymbol{p}_n$$

Array Manifold Vector

- And we have

$$\boldsymbol{F}(\omega) = \begin{bmatrix} e^{-j\boldsymbol{k}^T\boldsymbol{p}_0} \\ e^{-j\boldsymbol{k}^T\boldsymbol{p}_1} \\ \vdots \\ e^{-j\boldsymbol{k}^T\boldsymbol{p}_{N-1}} \end{bmatrix} F(\omega) = F(\omega)\boldsymbol{v}_{\boldsymbol{k}}(\boldsymbol{k})$$

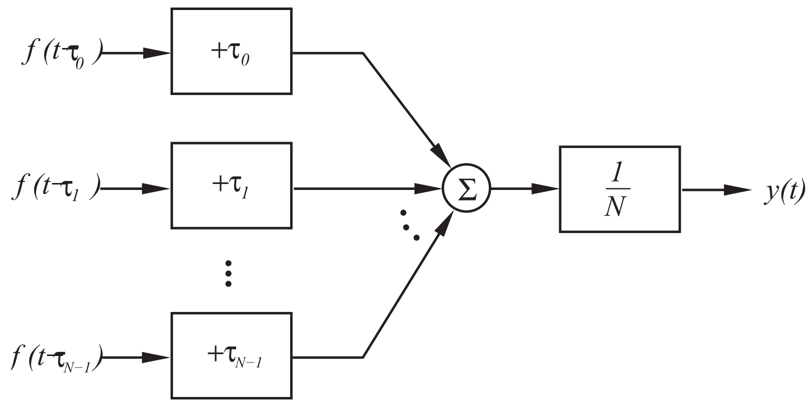
Array Manifold Vector

- In this particular example, we can use $h_n(t) = \frac{1}{N}\delta(t+\tau_n)$ such that

$$y(t) = f(t)$$

Following, we have the delay-and-sum beamformer.

Delay-and-sum Beamformer

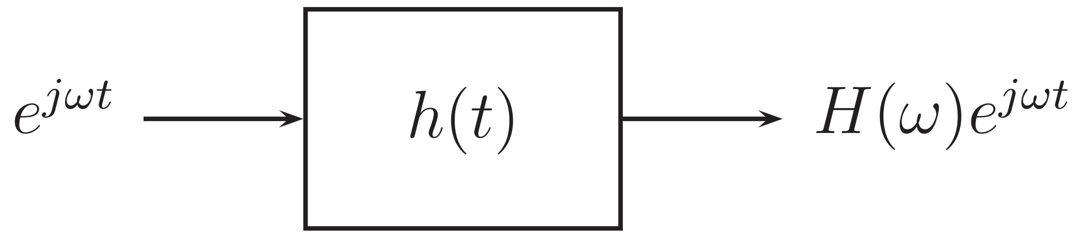


- A common delay is added in each channel to make the operations physically realizable
- Since $\mathcal{F}\left\{h_n(t)\right\} = \mathcal{F}\left\{\frac{1}{N}\delta(t+\tau_n)\right\} = e^{j\omega\tau_n}$
- We can write

Array Manifold Vector

$$\boldsymbol{H}^T(\omega) = \frac{1}{N}\boldsymbol{v}_{\boldsymbol{k}}^H(\boldsymbol{k})$$

LTI System



- Space-time signals (base functions):

$$f_n(t, \boldsymbol{p}) = e^{j\omega(t-\tau_n)} = e^{j(\omega t - \boldsymbol{k}^T \boldsymbol{p}_n)}$$

Note that $\omega\tau_n = \boldsymbol{k}^T \boldsymbol{p}_n$

- $\therefore \ \boldsymbol{f}(t, \boldsymbol{p}) = e^{j\omega t} \boldsymbol{v}_{\boldsymbol{k}}(\boldsymbol{k})$

Frequency-Wavenumber Response Function

- The response of the array to this plane wave is:

$$y(t, \boldsymbol{k}) = \boldsymbol{H}^T(\omega)\boldsymbol{v}_{\boldsymbol{k}}(\boldsymbol{k})e^{j\omega t}$$

- After taking the Fourier transform, we have:

$$Y(\omega, \boldsymbol{k}) = \boldsymbol{H}^T(\omega)\boldsymbol{v}_{\boldsymbol{k}}(\boldsymbol{k})$$

- And we define the Frequency-Wavenumber Response Function:

$$\Upsilon(\omega, \boldsymbol{k}) \triangleq \boldsymbol{H}^T(\omega)\boldsymbol{v}_{\boldsymbol{k}}(\boldsymbol{k})$$

Upsilon

$\Upsilon(\omega, \boldsymbol{k})$ describes the complex gain of an array to an input plane wave with wavenumber $\boldsymbol{k}$ and temporal frequency ω.

Beam Pattern and Bandpass Signal

- BEAM PATTERN is the Frequency Wavenumber Response Function evaluated versus the direction:

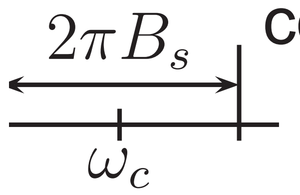
$$B(\omega : \theta, \phi) = \Upsilon(\omega, \boldsymbol{k})$$

Note that $\boldsymbol{k} = \frac{2\pi}{\lambda}\boldsymbol{a}(\theta, \phi)$, and $\boldsymbol{a}$ is the unit vector with spherical coordinates angles θ and ϕ

- Let's write a bandpass signal:

$$f(t, \boldsymbol{p}_n) = \sqrt{2}Re\{\tilde{f}(t, \boldsymbol{p}_n)e^{j\omega_c t}\}, n = 0, 1, \cdots, N-1$$

- ω_c corresponds to the carrier frequency and the complex envelope $\tilde{f}(t, \boldsymbol{p}_n)$ is bandlimited to the region



$$|\underbrace{\omega - \omega_c}_{\omega_L}| \leq 2\pi B_s/2$$

Bandlimited and Narrowband Signals

- Bandlimited plane wave:
  $f(t, \boldsymbol{p}_n) = \sqrt{2} Re\{\tilde{f}(t - \tau_n) e^{j\omega_c (t - \tau_n)}\}, n = 0, 1, \cdots, N - 1$

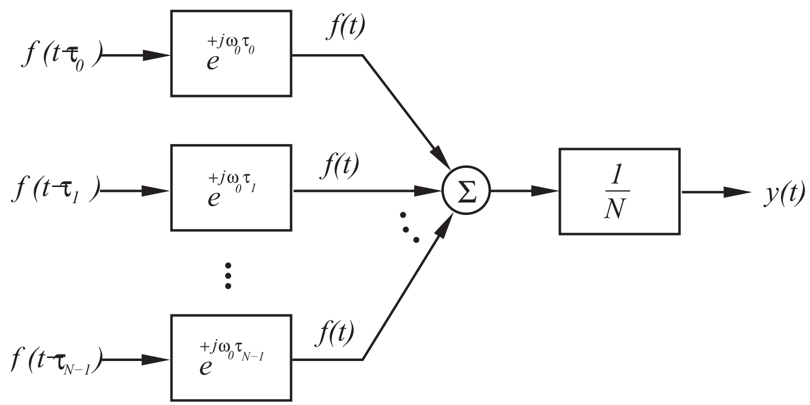
- Maximum travel time (ΔT_{max}) across the (linear) array: travel time between the two sensors at the extremities (signal arriving along the end-fire)

- Assuming the origin is at the array's center of gravity:
  $\sum_{n=0}^{N-1} \boldsymbol{p}_n = 0 \Rightarrow \tau_n \leq \Delta T_{max}$

- In Narrowband (NB) signals, $B_s \Delta T_{max} \ll 1$
  $\Rightarrow \tilde{f}(t - \tau_n) \simeq \tilde{f}(t)$ and
  $f(t, \boldsymbol{p}_n) = \sqrt{2} Re\{\tilde{f}(t) e^{-j\omega_c \tau_n} e^{j\omega_c t}\}$

Phased-Array

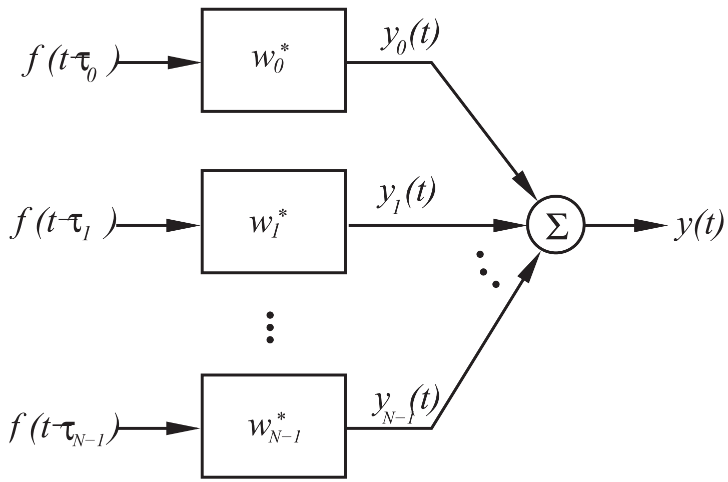
- For NB signals, the delay is approximated by a phase-shift:
  $\Rightarrow$ delay&sum beamformer $\equiv$ PHASED ARRAY



- The phased array can be implemented adjusting the gain and phase to achieve a desired beam pattern

NB Beamformers

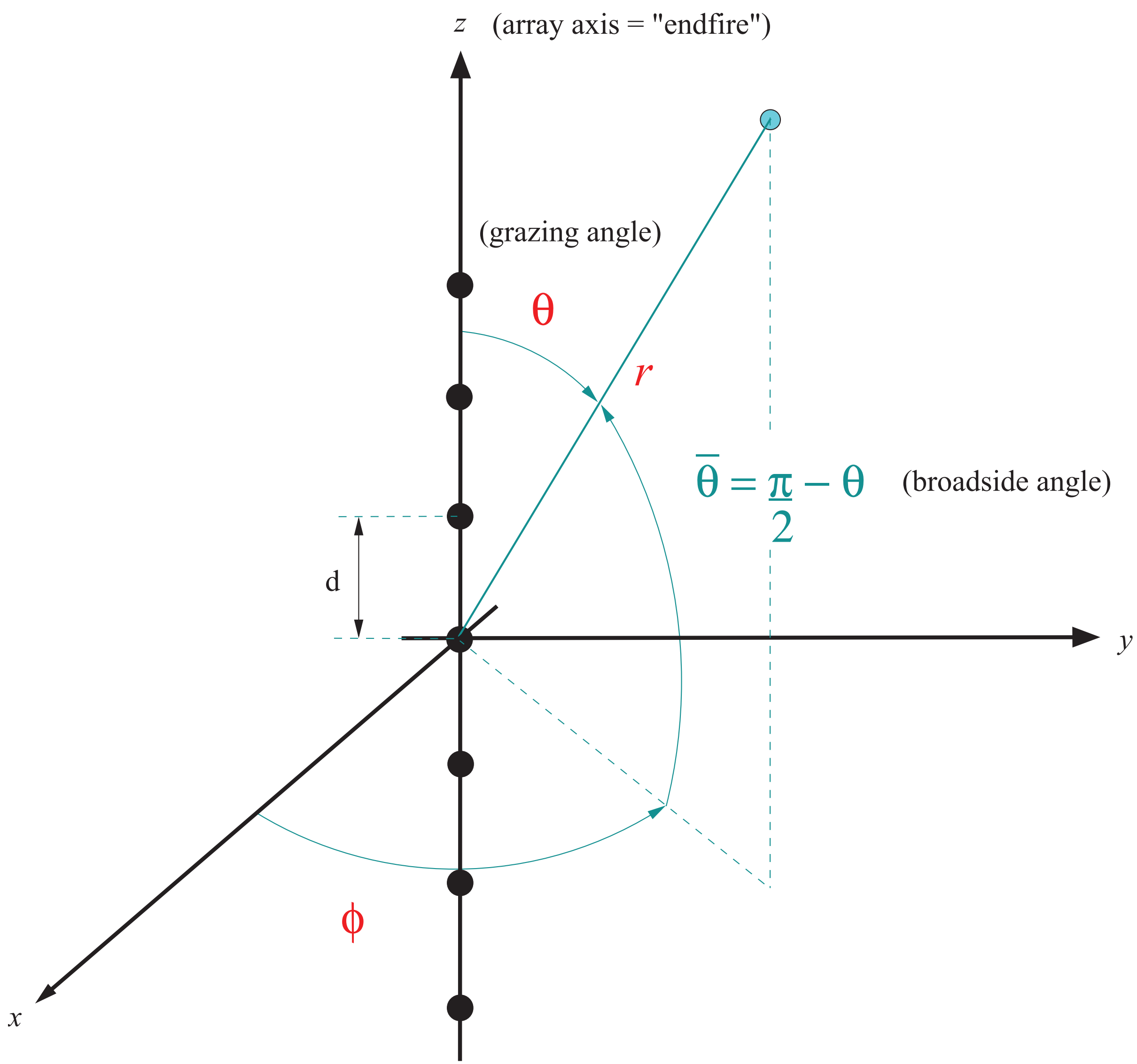
- In narrowband beamformers: $y(t, \boldsymbol{k}) = \boldsymbol{w}^H \boldsymbol{v}_{\boldsymbol{k}}(\boldsymbol{k}) e^{j\omega t}$



- $\Upsilon(\omega, \boldsymbol{k}) = \underbrace{\boldsymbol{w}^H}_{\boldsymbol{H}^T(\omega)} \boldsymbol{v}_{\boldsymbol{k}}(\boldsymbol{k})$

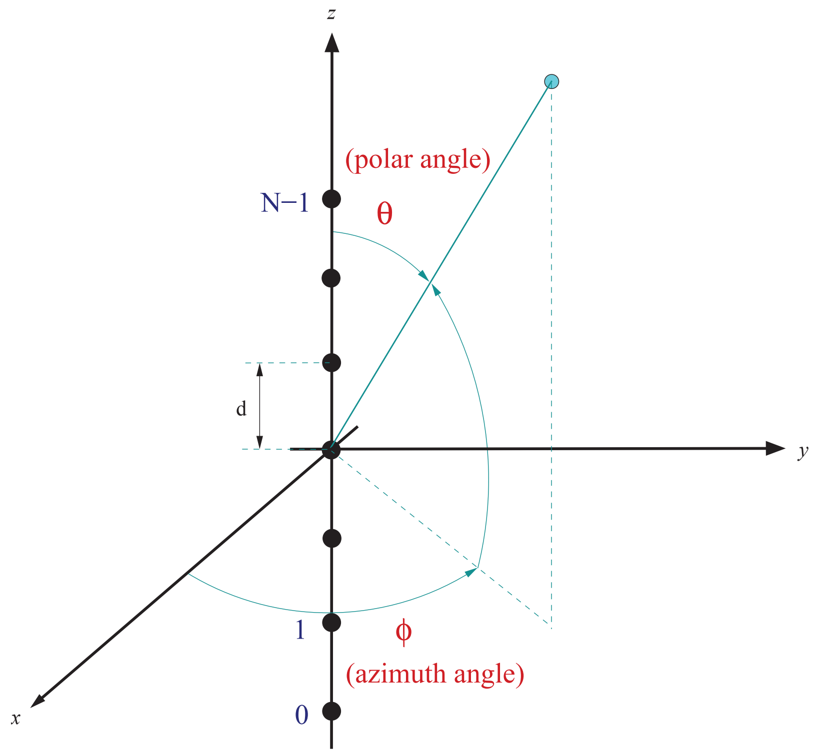
2.3 Uniform Linear Arrays (ULA)

Uniformly Spaced Linear Arrays



ULA

- An ULA along axis z:



- Location of the elements:

$$\begin{cases} p_{zn} = (n - \frac{N-1}{2})d, \text{ for } n = 0, 1, \cdots, N-1 \\ p_{xn} = p_{yn} = 0 \end{cases}$$

- Therefore, $\boldsymbol{p}_n = \begin{bmatrix} 0 \\ 0 \\ (n - \frac{N-1}{2})d \end{bmatrix}$

ULA

- Array manifold vector:

$$\boldsymbol{v}_{\boldsymbol{k}}(\boldsymbol{k}) = \begin{bmatrix} e^{-j\boldsymbol{k}^T\boldsymbol{p}_0} & e^{-j\boldsymbol{k}^T\boldsymbol{p}_1} & \ldots & e^{-j\boldsymbol{k}^T\boldsymbol{p}_{N-1}} \end{bmatrix}^T$$

$$\therefore \boldsymbol{v}_{\boldsymbol{k}}(\boldsymbol{k}) = \boldsymbol{v}_{\boldsymbol{k}}(k_z) = \begin{bmatrix} e^{+j\frac{(N-1)}{2}k_z d} \\ e^{+j(\frac{N-1}{2}-1)k_z d} \\ \vdots \\ e^{-j(\frac{N-1}{2})k_z d} \end{bmatrix}$$