Sensores espectroscópicos e modelos de regressão aplicados na análise de solos

Aula 6 – Redes Neurais Convolucionais - parte 2

Me. José Vinícius Ribeiro

Universidade Estadual de Londrina

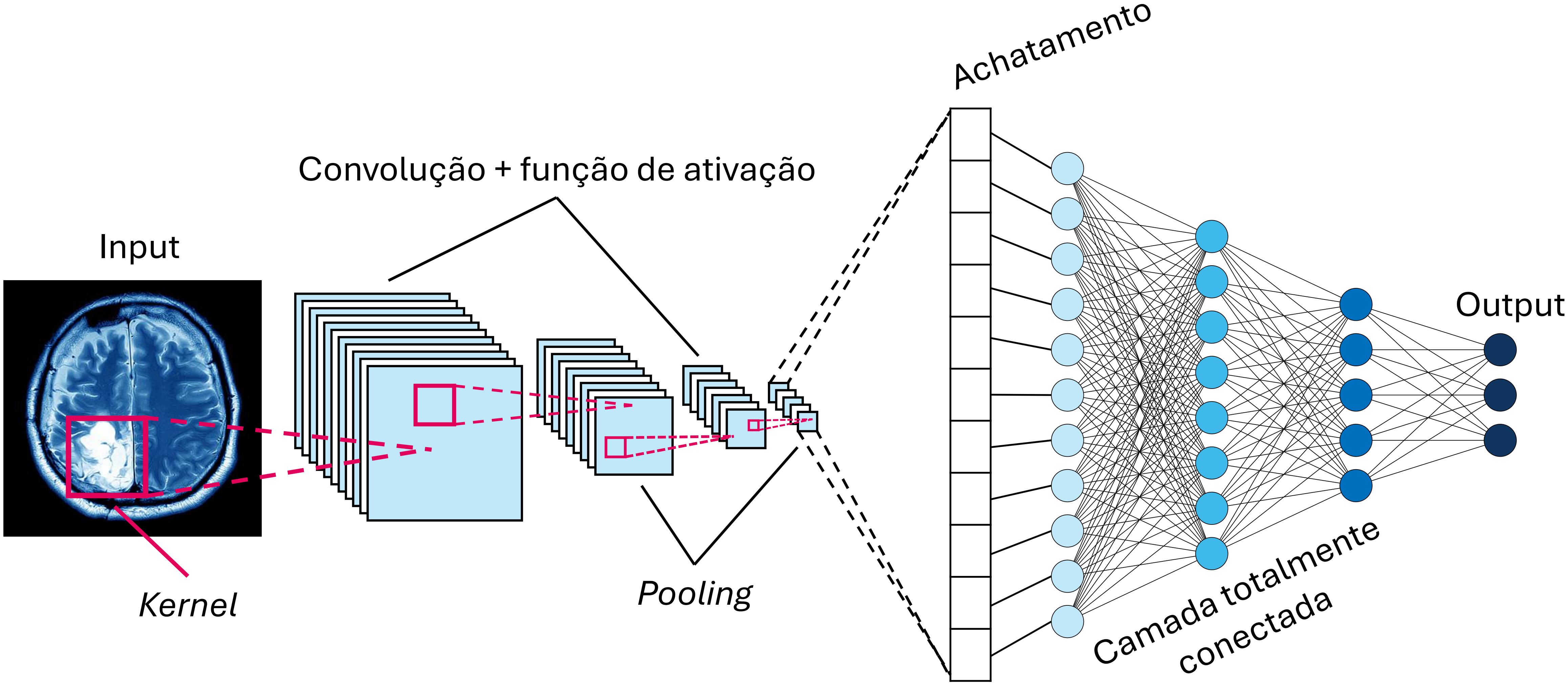
PÓS GRADUAÇÃO FÍSICA UEL

SUMÁRIO

- Intuição por traz da Transformada de Fourier
- Transformada discreta de Fourier
- Transformada de Fourier de tempo curto
- CNN 2D – espectrogramas
- Prática no python (google colab)

Relembrando a arquitetura básica

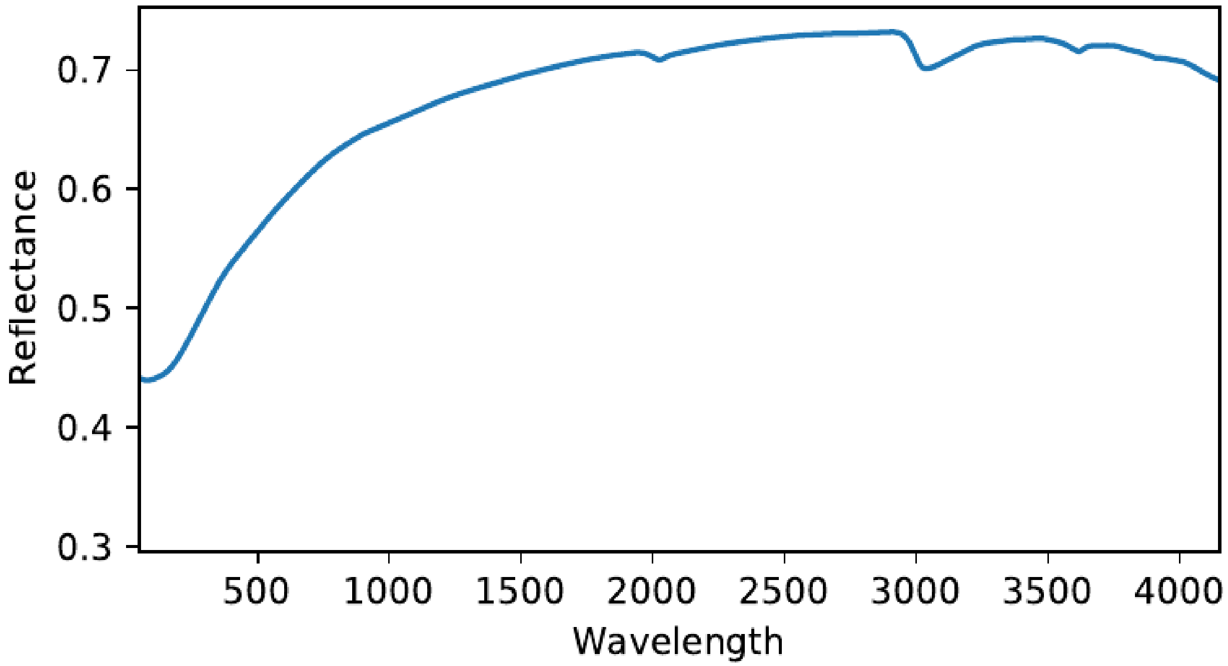
Redes Neurais Convolucionais (CNN)

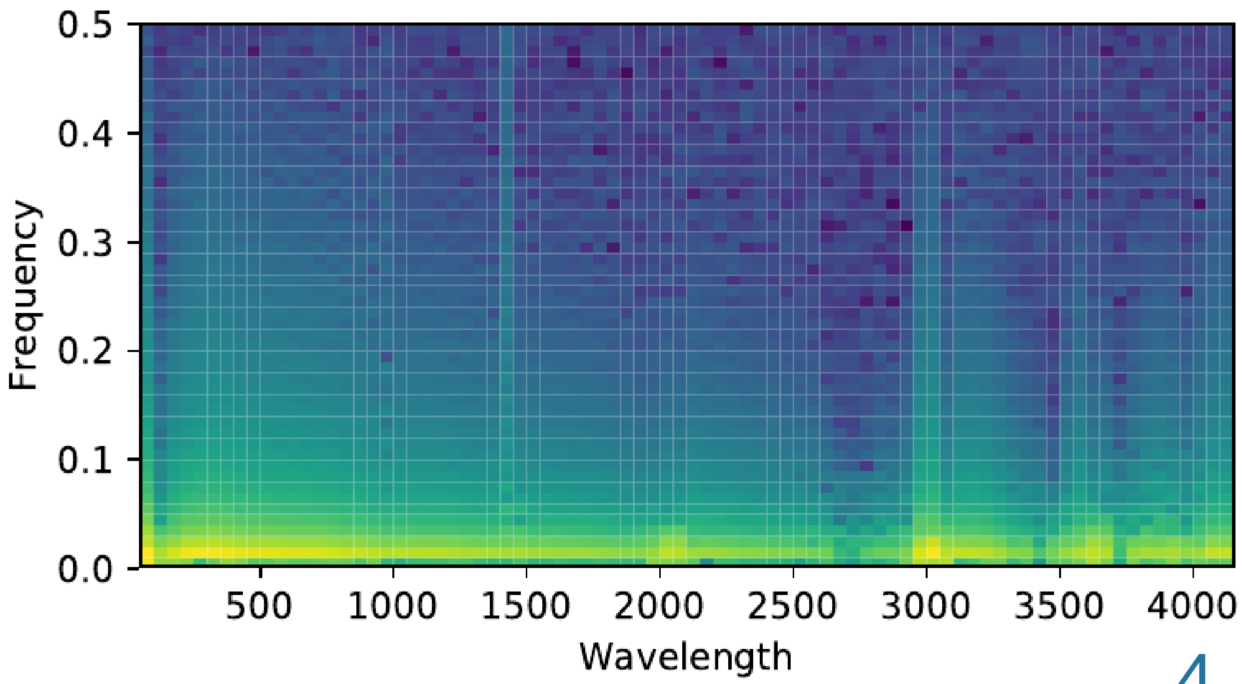


Motivação

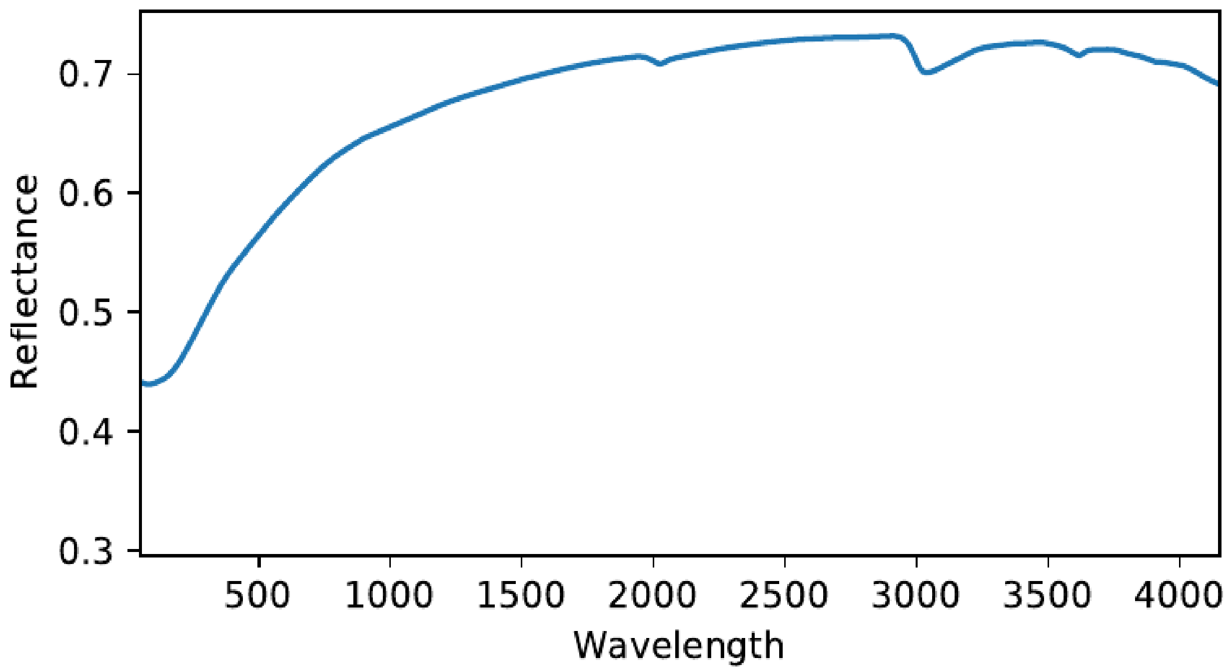
Como extrair todo o potencial da camada de convolução de uma CNN no contexto de dados espectrais 1D?

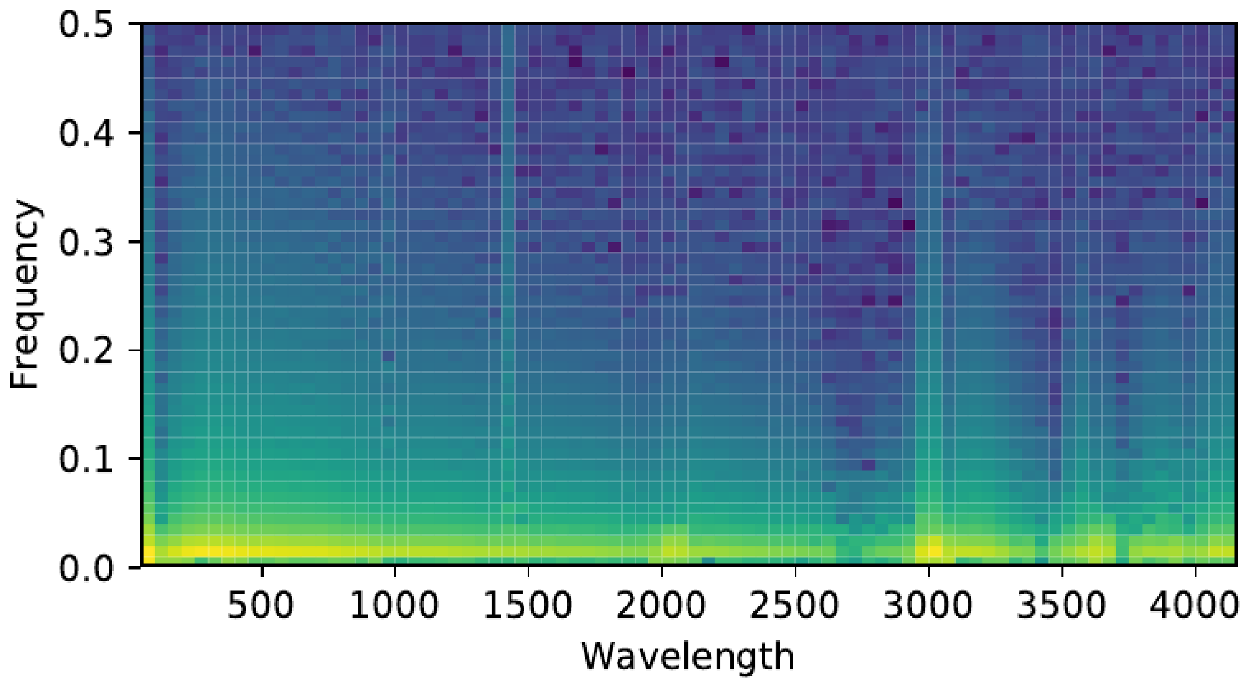
Uma possível saída é tratar os espectros como séries temporais e extrair a sua representação no domínio das frequências, gerando os espectrogramas (2D). Isso pode ser feito por meio da Transformada de Fourier de Tempo Curto





Espectrograma





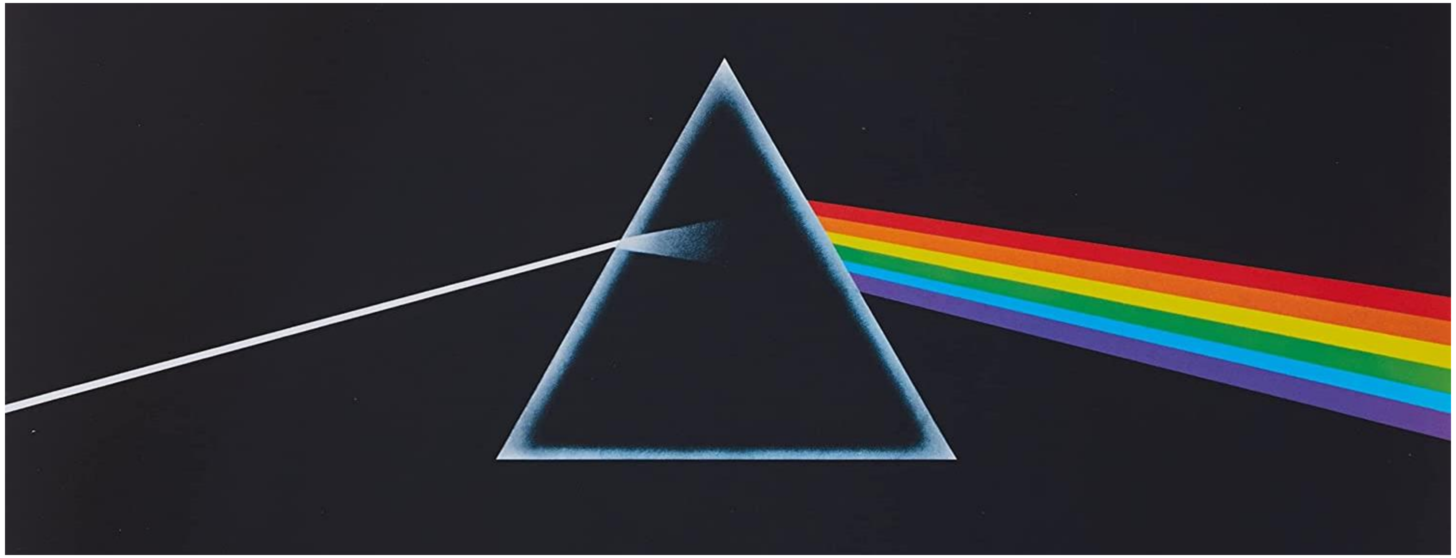
Eixo X: variáveis (energia/comprimento de onda)
Eixo Y: diferentes frequências que produziram os sinais ao longo do espectro
Intensidades das cores (eixo Z): Amplitude das frequências (intensidade)

Intuição por traz da Transformada de Fourier

Intuição por traz da Transformada de Fourier

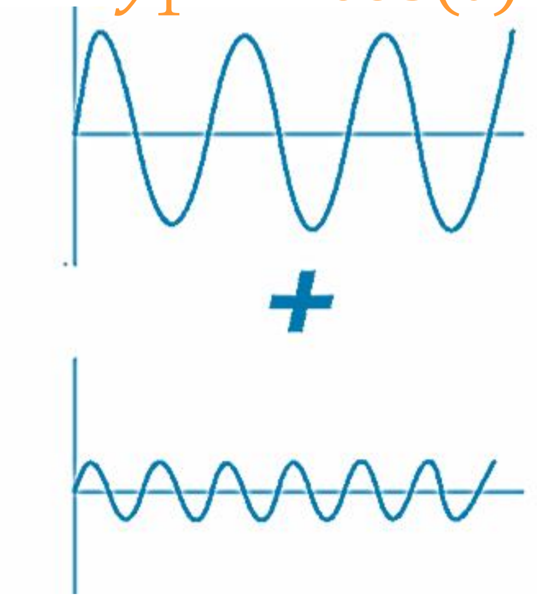
A ideia por traz da transformada de Fourier (FT) é fornecer uma ferramenta matemática capaz de decompor sinais (espectros) complexos em seus constituintes mais básicos (frequências fundamentais)

Como analogia, podemos pensar no processo de decomposição da luz por um prisma



Intuição por traz da Transformada de Fourier

$y_1 = 2\cos(t)$

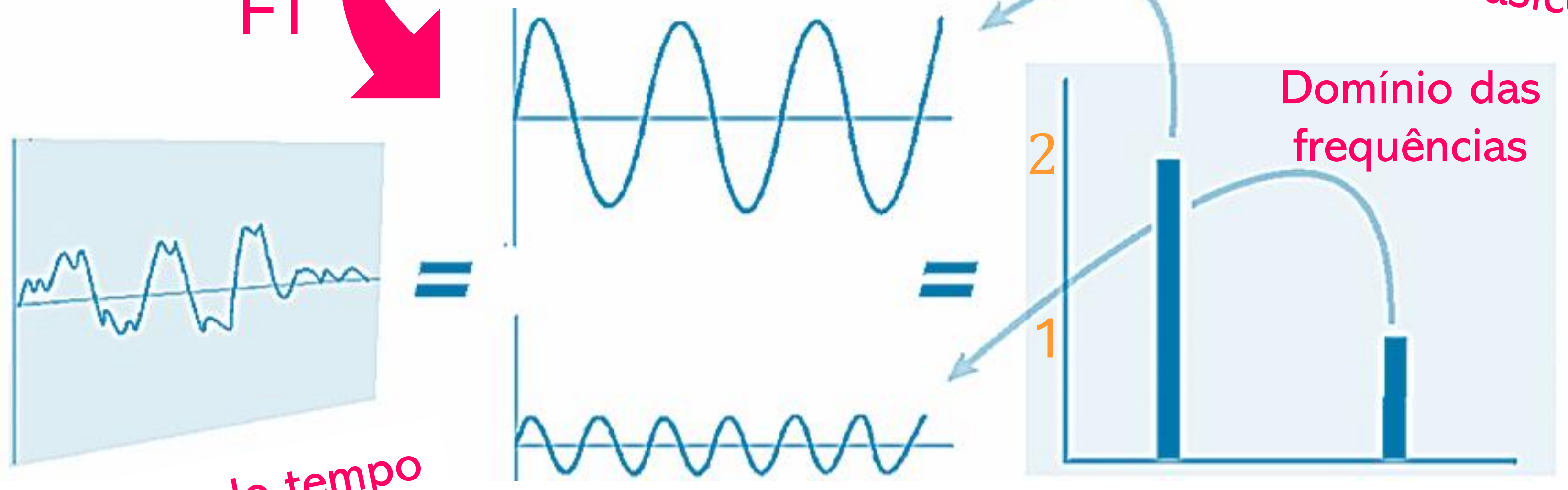


$y_2 = 1\cos(2t)$

$y_3 = 1\cos(2t) + 2\cos(t) =$

Domínio do tempo

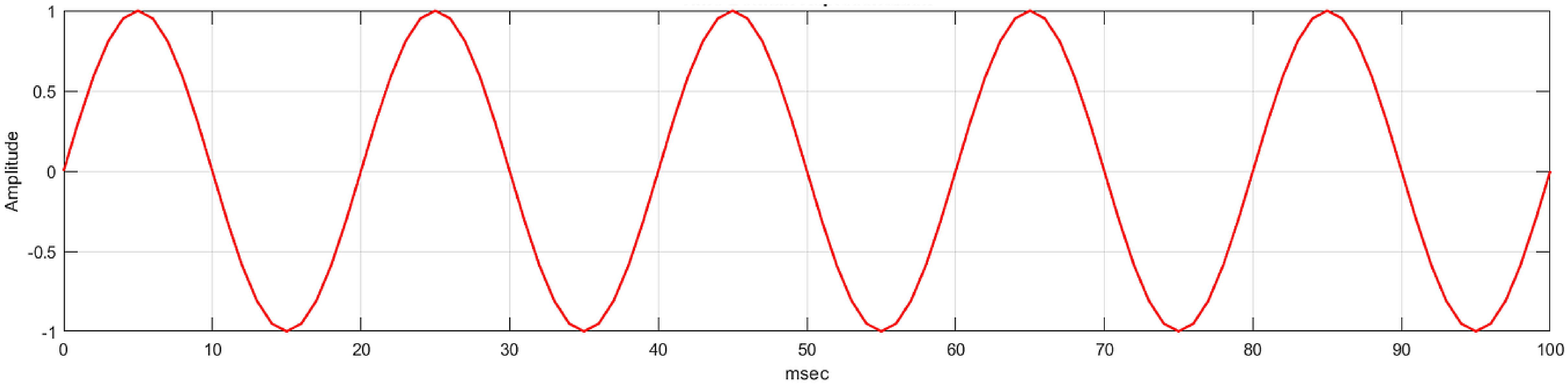
FT



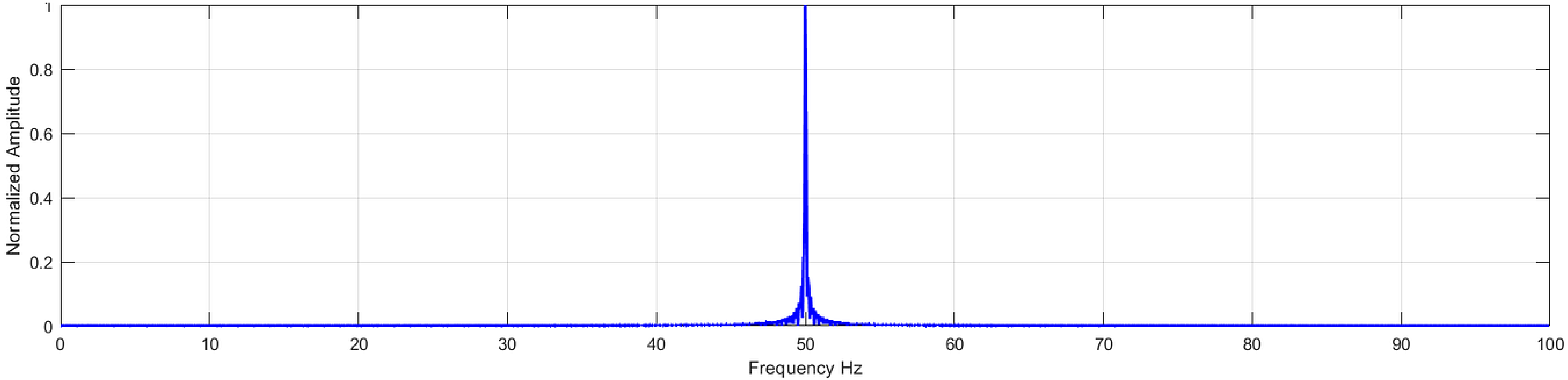
Domínio do tempo

Frequência: número de oscilações/repetições de um evento em um dado intervalo de tempo

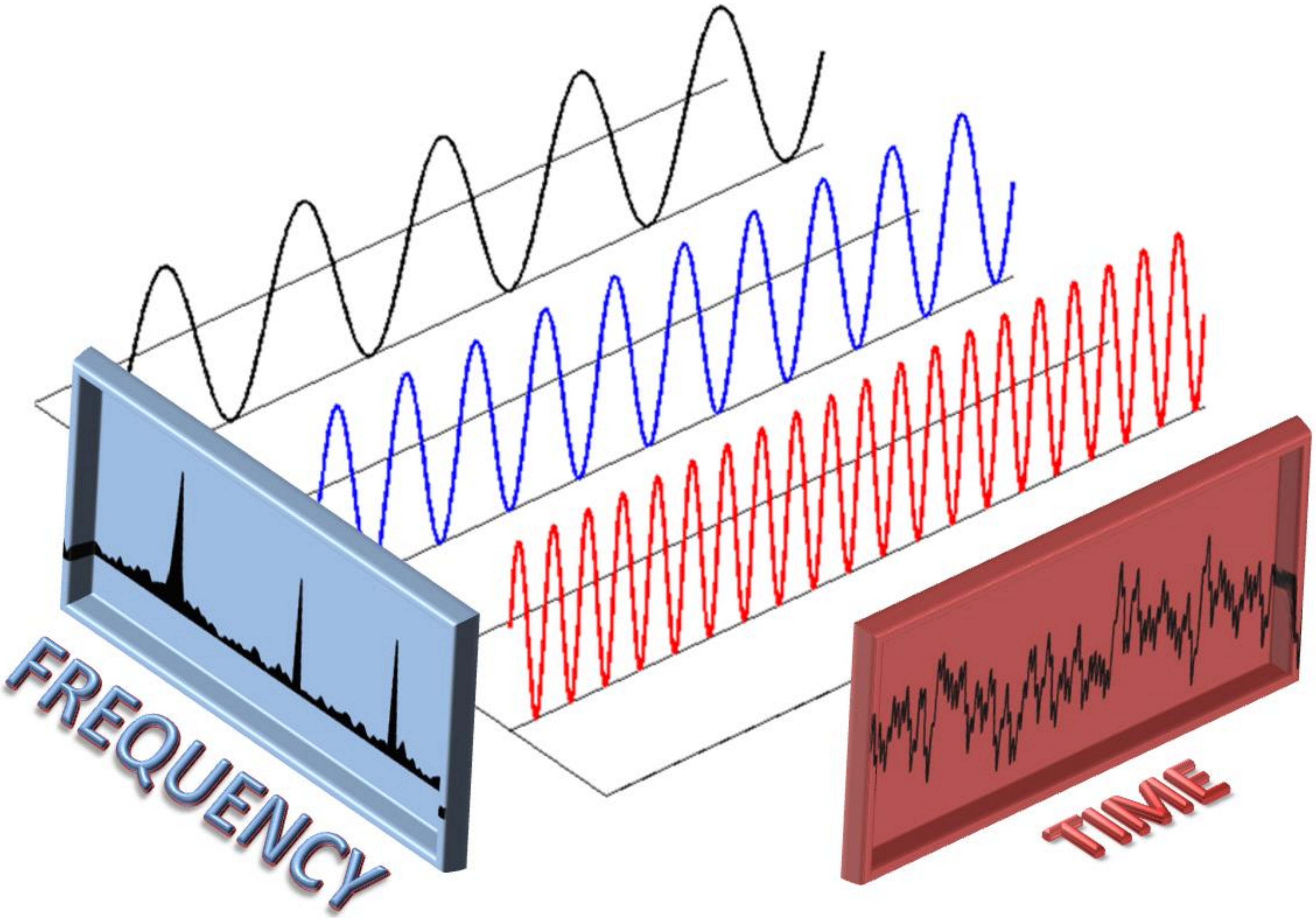
Domínio temporal (original)



Domínio das frequências



Domínio temporal (original) × Domínio das frequências



Intuição por traz da Transformada de Fourier

A ideia é usar o fato de que qualquer sinal pode ser decomposto em uma soma de senos e cossenos (série de Fourier) com amplitudes e frequências específicas (fundamentais). A FT, portanto, extrai essas frequências fundamentais contabilizando os coeficientes da série de Fourier

Matematicamente, a TF de uma série temporal continua $x(t)$ é:

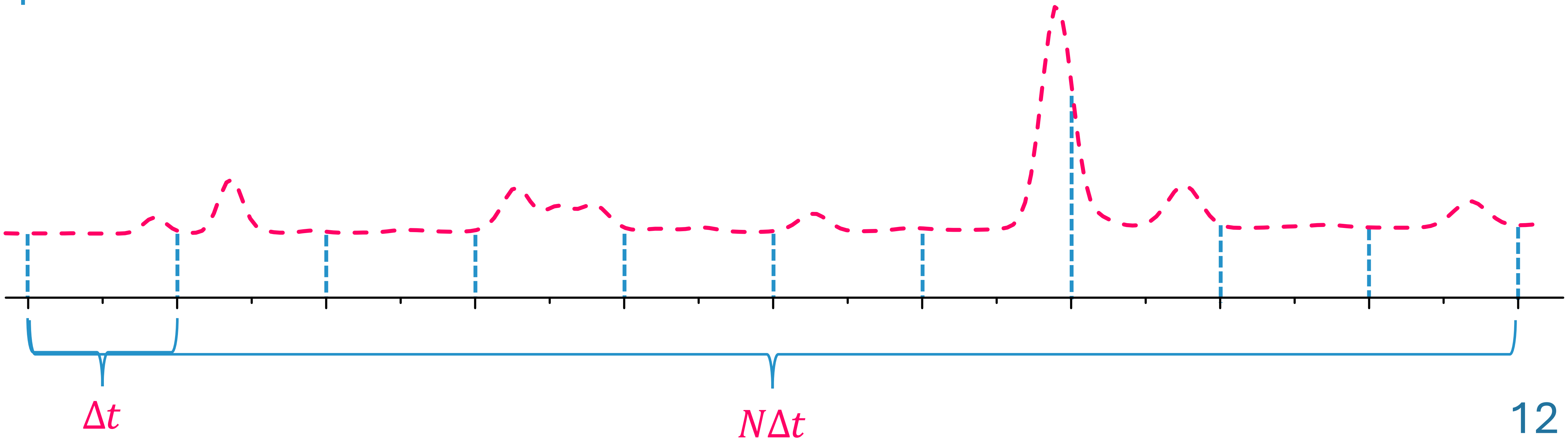
$$x(f) = \int_{-\infty}^{\infty} x(t)e^{-i2\pi f t}dt = \int_{-\infty}^{\infty} x(t)[\cos(2\pi f t) - i sen(2\pi f t)]\, dt$$

Uma série temporal é a representação de uma sequência de observações tomadas em intervalos de tempo periódicos. Ex: medida da temperatura em um determinado local, de hora em hora. No nosso caso o "tempo" é a energia/comp de onda e temos a representação das contagens x intervalos de energia/comp de onda

Transformada discreta de Fourier

$$x(f) = \int_{-\infty}^{\infty} x(t)e^{-i2\pi ft}dt = \int_{-\infty}^{\infty} x(t)[\cos(2\pi ft) - i\,sen(2\pi ft)]\,dt$$

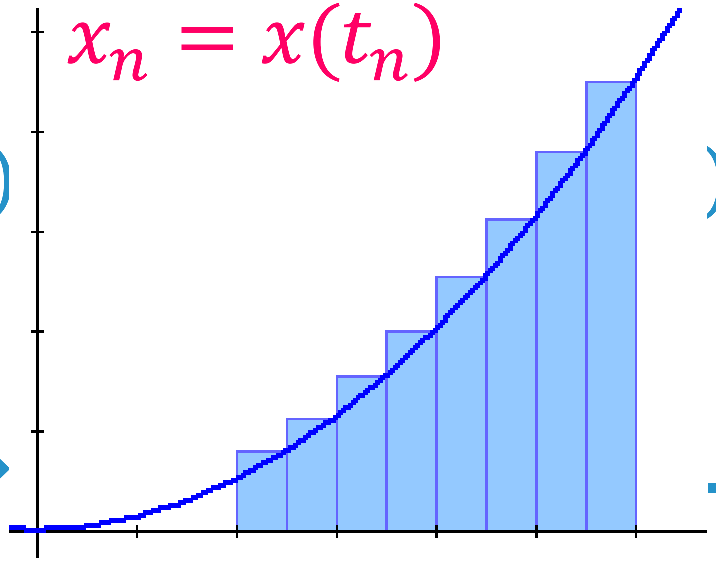
No contínuo, a TF integra $x(t)$ sobre todo t. Entretanto, nossos dados são discretos no tempo (energia/comp de onda). Em outras palavras, nosso $x(t)$ é discreto nos valores de t. Por isso precisamos da versão discreta da FT.



Transformada discreta de Fourier

$$x(f) = \int_{-\infty}^{\infty} x(t)e^{-i2\pi f t}dt = \int_{-\infty}^{\infty} x(t)[\cos(2\pi f t) - i sen(2\pi f t)]\, dt$$

No caso discreto, podemos discretizar o tempo em N intervalos Δt. Com amostragens em instantes $t_n = n\Delta t$ só teremos valores pontuais $x_n = x(t_n)$. Logo, para cada amostra x_n:



$$x(f) = \sum_{n=0}^{N-1} x_n e^{-i2\pi f n\Delta t} \quad ; n = 0, 1, \dots, N-1$$

$$x(k) = \sum_{n=0}^{N-1} x_n e^{-i2\pi\frac{kn}{N}}$$

Transformada discreta de Fourier

$$x(f) = \int_{-\infty}^{\infty} x(t) e^{-i2\pi f t} dt$$

Caso contínuo

$$x(k) = \sum_{n=0}^{N-1} x_n e^{-i2\pi \frac{kn}{N}} \qquad f_k = \frac{k}{N\Delta t}$$

Caso discreto $t_n = n\Delta t; \quad n, k = 0,1, \ldots N-1$

Como resultado, teremos as frequências f ou f_k com as quais $x(t)$ ou $x_n = x(t_n)$ podem ser decompostos, gerando os numeros complexos $x(f)$ ou $x(k)$ dos quais as intensidades (ou amplitudes) das frequências podem ser extraídas. Por exemplo, para o caso discrete temos:

$$x(k) = \sum_{n=0}^{N-1} x_n \cos\left(2\pi \frac{kn}{N}\right) - i \sum_{n=0}^{N-1} x_n \operatorname{sen}\left(2\pi \frac{kn}{N}\right)$$

$$Z = a + ib$$

$$a = Re(Z)$$
$$b = Im(Z)$$

Transformada discreta de Fourier

$$x(f) = \int_{-\infty}^{\infty} x(t)e^{-i2\pi f t}\,dt$$

Caso contínuo

$$x(k) = \sum_{n=0}^{N-1} x_n e^{-i2\pi\frac{kn}{N}} \qquad f_k = \frac{k}{N\Delta t}$$

Caso discreto $t_n = n\Delta t; \quad n, k = 0,1, \ldots N-1$

Como resultado, teremos as frequências f ou f_k com as quais $x(t)$ ou $x_n = x(t_n)$ podem ser decompostos, gerando os numeros complexos $x(f)$ ou $x(k)$ dos quais as intensidades (ou amplitudes) das frequências podem ser extraídas. Por exemplo, para o caso discrete temos:

$$|Z| = \sqrt{a^2 + b^2} \qquad I_k = |x(k)| = \sqrt{\left[\sum_{n=0}^{N-1} x_n \cos\left(2\pi\frac{kn}{N}\right)\right]^2 + \left[\sum_{n=0}^{N-1} x_n \mathrm{sen}\left(2\pi\frac{kn}{N}\right)\right]^2}$$

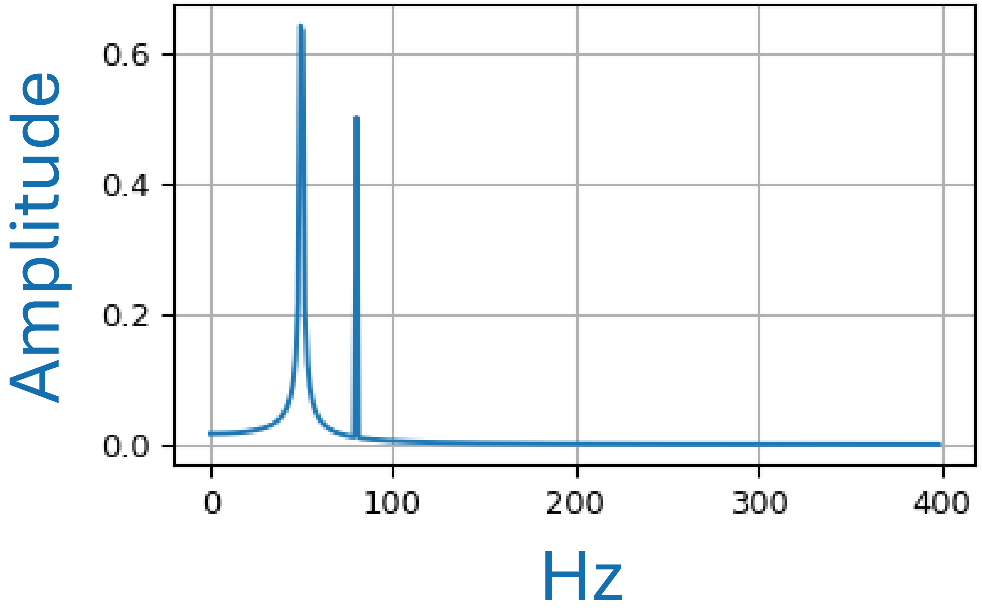
Transformada discreta de Fourier

$$x(k) = \sum_{n=0}^{N-1} x(t_n) e^{-i2\pi\frac{kn}{N}}$$

$$f_k = \frac{k}{N\Delta t} \qquad I_k = |x(k)|$$

$$t_n = n\Delta t; \quad n, k = 0, 1, \ldots N-1$$

Portanto, teremos um vetor 1D cujos coeficientes são amplitudes de frequências (I_k) presentes no sinal original em função das f_k frequências



Simetria hermitiana

Por outro lado, devido a natureza subjacente dos números complexos, existe uma redundância nas respostas da FFT. É a chamada simetria hermitiana

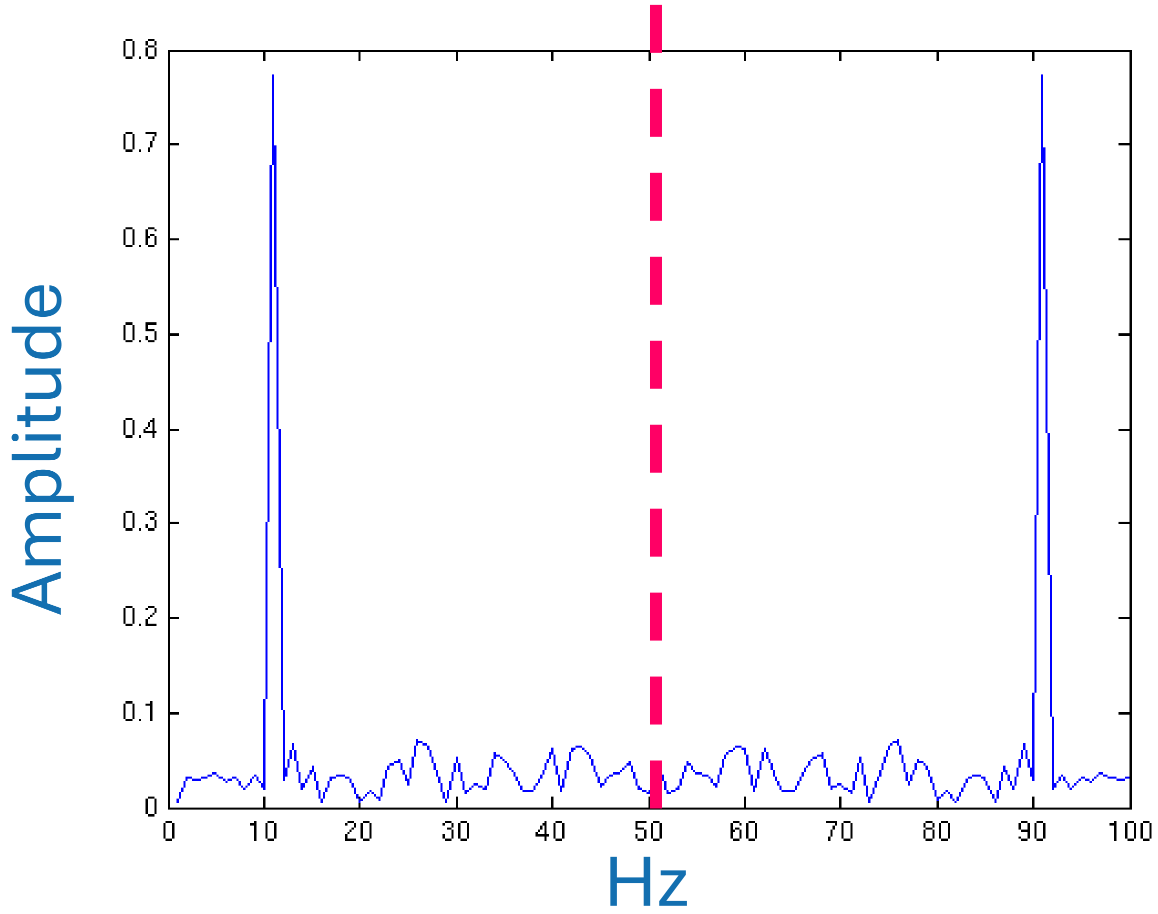
$$x(N-k) = \sum_{n=0}^{N-1} x(t_n) e^{-i\frac{2\pi n}{N}(N-k)} = \sum_{n=0}^{N-1} x(t_n) e^{-i\frac{2\pi n}{N}N} e^{i\frac{2\pi n}{N}k}$$

$$x(N-k) = \sum_{n=0}^{N-1} x(t_n) e^{-i2\pi n} e^{i\frac{2\pi n}{N}k} \qquad e^{-i2\pi n} = cos\,(2\pi n) - isen(2\pi n) = 1$$

$$x(N-k) = \sum_{n=0}^{N-1} x(t_n)\, e^{i\frac{2\pi n}{N}k} = \overline{x(k)} \qquad \overline{x(k)} = \sum_{n=0}^{N-1} x(t_n)\, e^{+i\frac{2\pi n}{N}k}$$

Simetria hermitiana

Portanto $x(N-k)=\overline{x(k)}$, isso acarreta que em algum k os valores de frequência começarão a assumir valores equivalentes aos complexo conjugados daqueles que vieram anteriormente. Como a amplitude de frequência não é afetada pelo sinal da parte imaginária de $x(k)$, existirá um k "**espelho**" que repetirá valores



Simetria hermitiana

$$f_k = \frac{k}{N\Delta t} \qquad k = 0,1, \dots N-1 \qquad f_{N-1} = \frac{(N-1)}{N\Delta t}$$

$$x(k+pN) = \sum_{n=0}^{N-1} x(t_n) e^{-i\frac{2\pi n}{N}(k+pN)} \qquad p = 0,1,2,3 \dots$$

$$x(k+pN) = \sum_{n=0}^{N-1} x(t_n) e^{-i\frac{2\pi n}{N}k} e^{-i\frac{2\pi n}{N}Np}$$

$$e^{-i2\pi np} = cos\,(2\pi np) - isen(2\pi np)$$

$$e^{-i2\pi np} = 1$$

$$x(k+pN) = \sum_{n=0}^{N-1} x(t_n) e^{-i\frac{2\pi n}{N}k} = x(k)$$

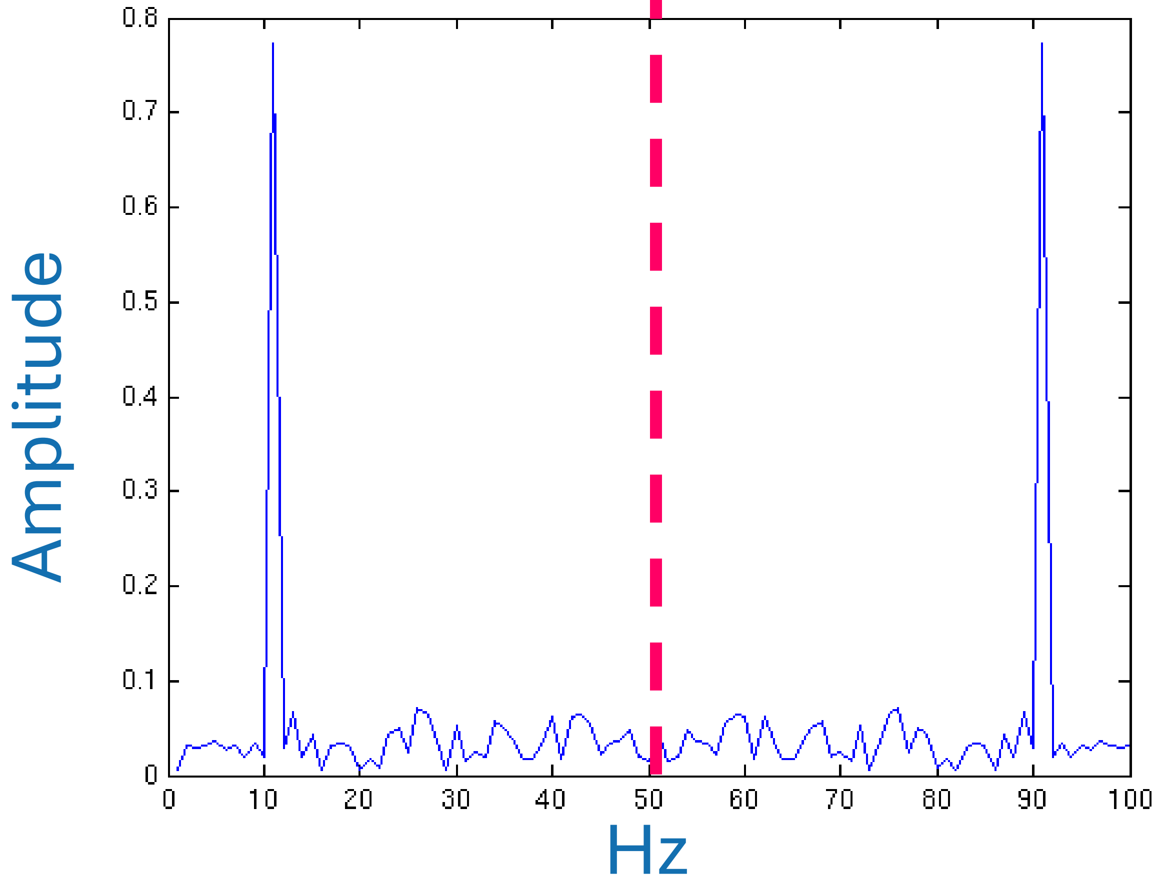
Ou seja, a FFT é periódica com relação a múltiplos inteiros de N

Simetria hermitiana

Logo, se a FFT é periódica com relação a múltiplos inteiros de N, e tem simetria hermitiana, o eixo de simetria só pode ocorrer na frequência cujo $k = N/2$

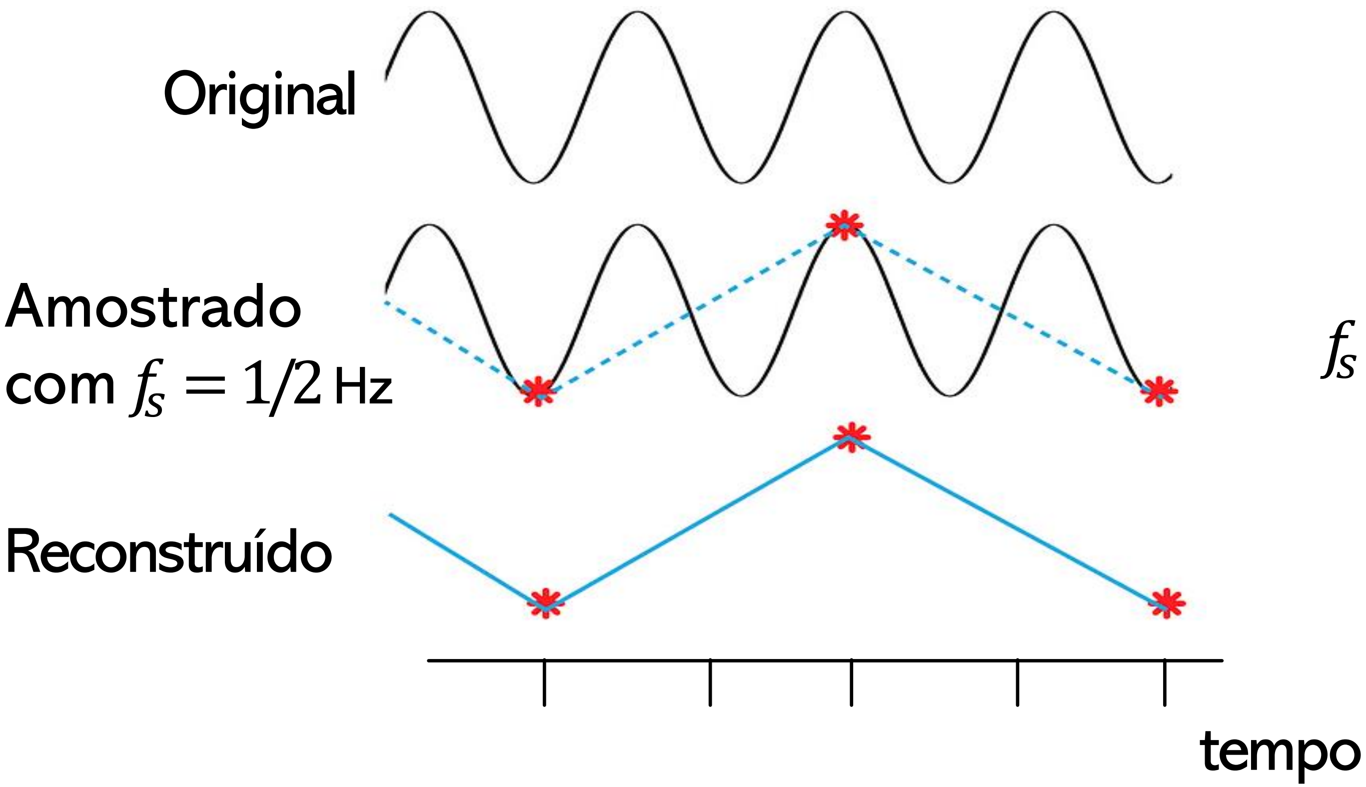
$$f_{N/2} = \frac{1}{N\Delta t}\frac{N}{2} = \frac{1}{2\Delta t}$$

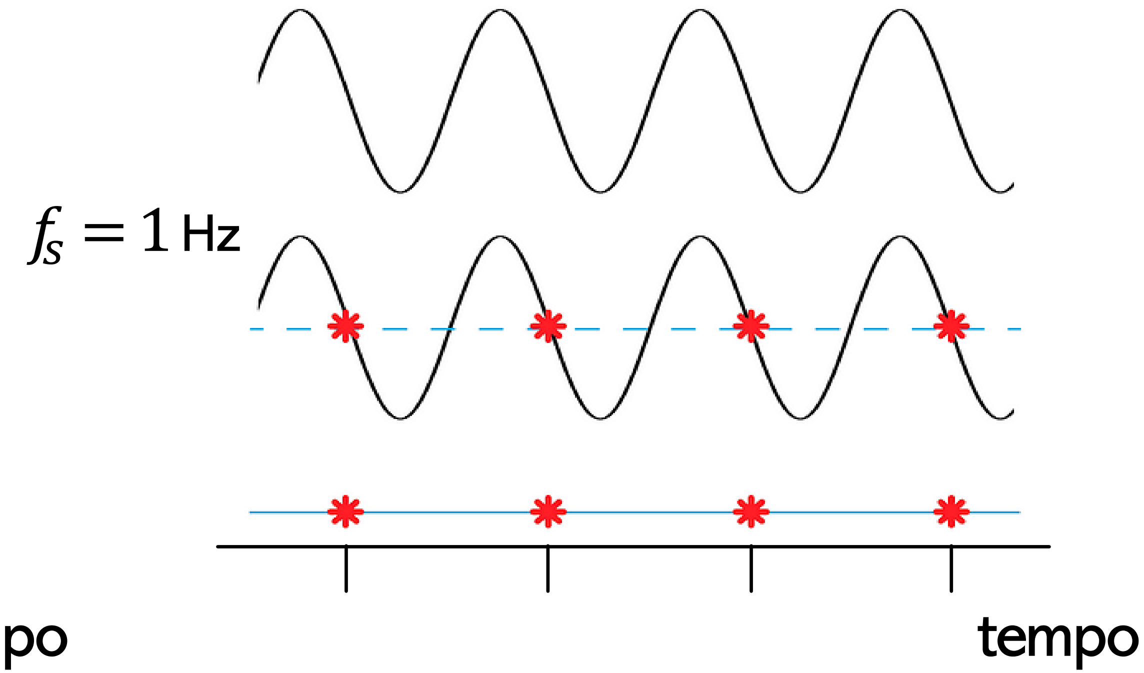
Essa é a chamada frequência (ou taxa) de *Nyquist*



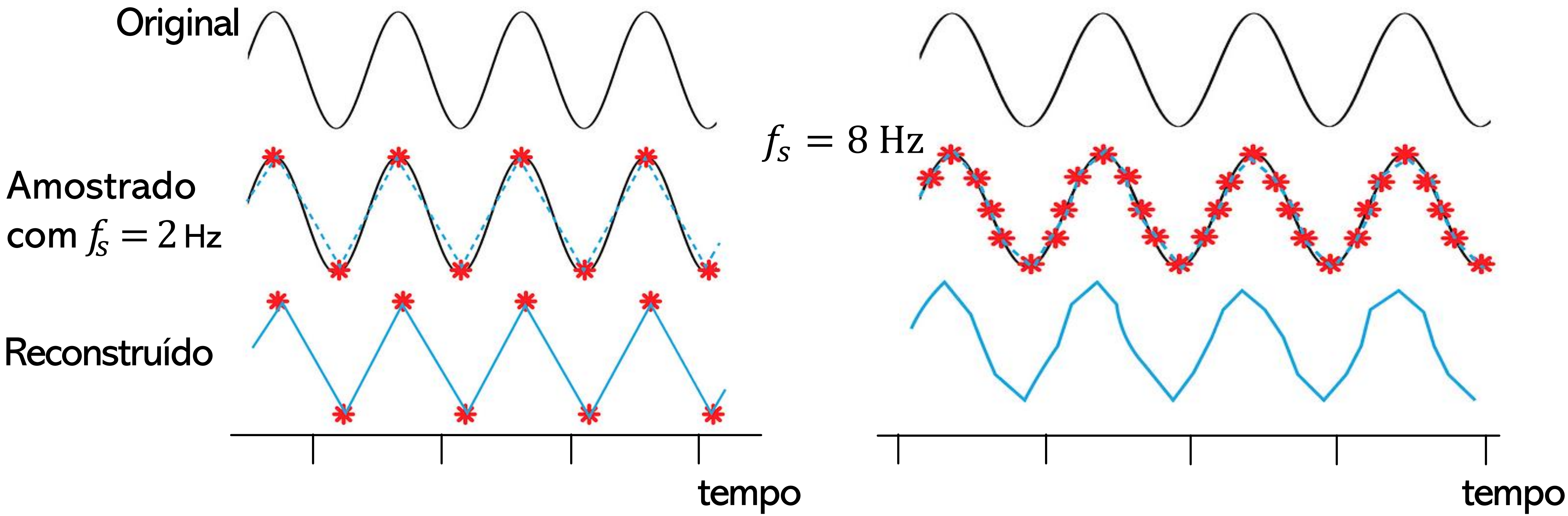
Frequência de amostragem

Definindo $f_s \equiv 1/\Delta t$ como a frequência de amostragem (*sampling frequency),* ou seja, o número de amostras coletadas por unidade de tempo (variáveis = energia/comp de onda), temos um parâmetro relacionado com a **resolução** com a qual o sinal base vai ser reconstruído discretamente



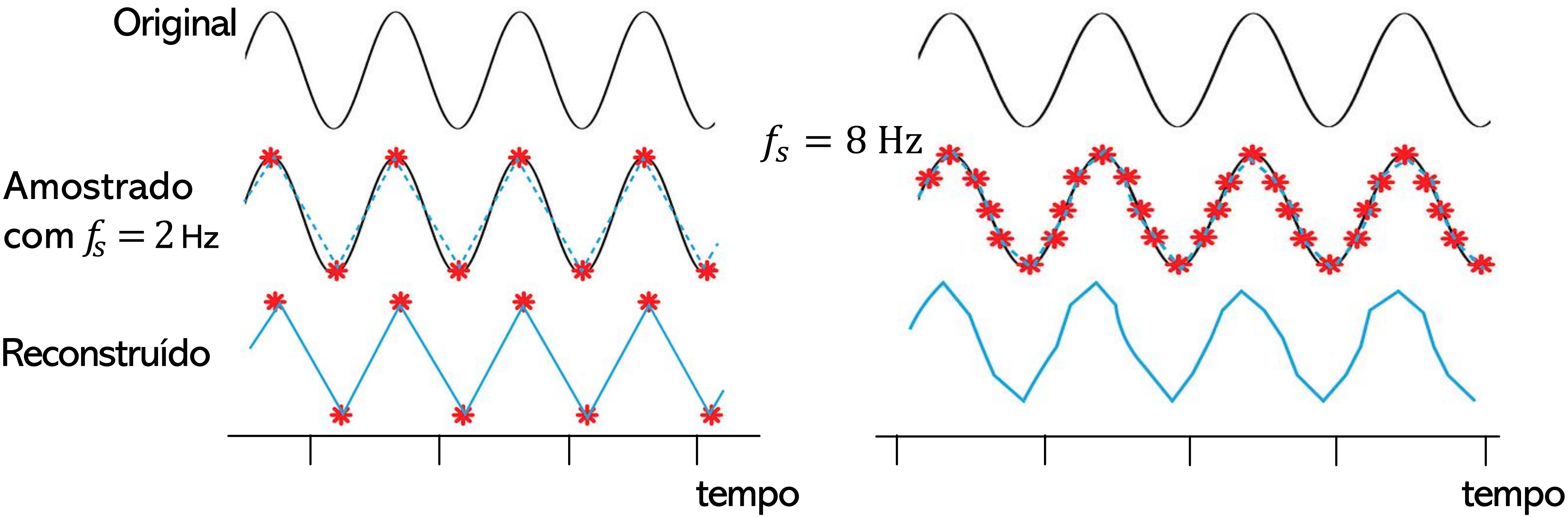


Transformada discreta de Fourier



Logo, a taxa de amostragem é um valor crucial para a FT reconstruir o sinal adequadamente, e assim extrair as frequências da melhor forma possível. O valor mínimo recomendado é $2 \times frequência\, original.$ Ex: para sinais gerados via 60 Hz, o mínimo para uma boa reconstrução via FT é um f_s de 120 Hz

Transformada discreta de Fourier



Logo, $f_s \geq 2 \times f_0$ para uma boa reconstrução do sinal e consequentemente bons resultados via FT (livres de distorções). Para sinais com multiplas frequências (como os espectros), devemos considerer a maior frequência entre todas como o f_0

Transformada Rápida de Fourier

A Transformada Rápida de Fourier (FFT) é um algoritmo eficiente para calcular a Transformada Discreta de Fourier (DFT), permitindo a análise de sinais no domínio da frequência com menor complexidade computacional e maior rapidez

A ideia do algoritmo baseia-se no chamado método de dobramentos sucessivo

a Transformada discreta de Fourier de n pontos da maneira ingênua, usando a definição, leva operações aritméticas de $O(N^2)$, enquanto uma Transformada rápida de Fourier pode computar a mesma Transformada discreta de Fourier em apenas $O\,(N * logN)$

Os resultados obtidos seja via transformada original seja via transformada rápida são essencialmente os mesmos, diferindo muito pouco

FFT de Tempo Curto

A Transformada de Fourier tradicional decompõe um sinal em suas componentes de frequência, assumindo que essas componentes são constantes ao longo do tempo (sinal estacionário). No entanto, muitos sinais do mundo real são não estacionários, ou seja, suas propriedades espectrais variam ao longo do tempo ou da energia/comprimento de onda (XRF, vis-NIR, etc...)

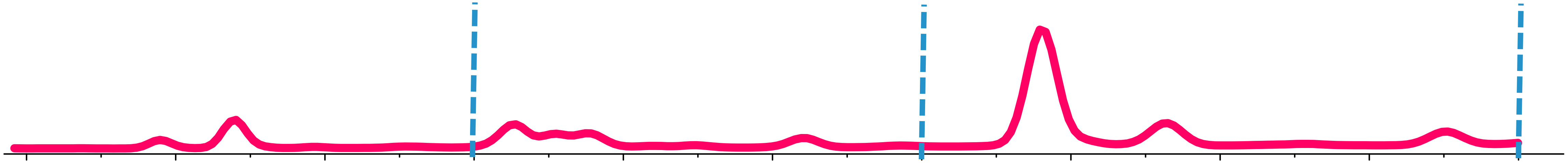
Para esse tipo de dataset, a transformada discreta (ou rápida) de Fourier de Tempo Curto (STFFT) é mais adequada.

A ideia é particionar o espectro original em intervalos menores, em janelas de "tempo curto", de forma que os sinais possam ser considerado constante no seu interior, e aplicar a FFT em cada um deles. Aqui temos o conceito de janela deslizante (semelhante ao processo de convolução unidimensional das redes neurais convolucionais)

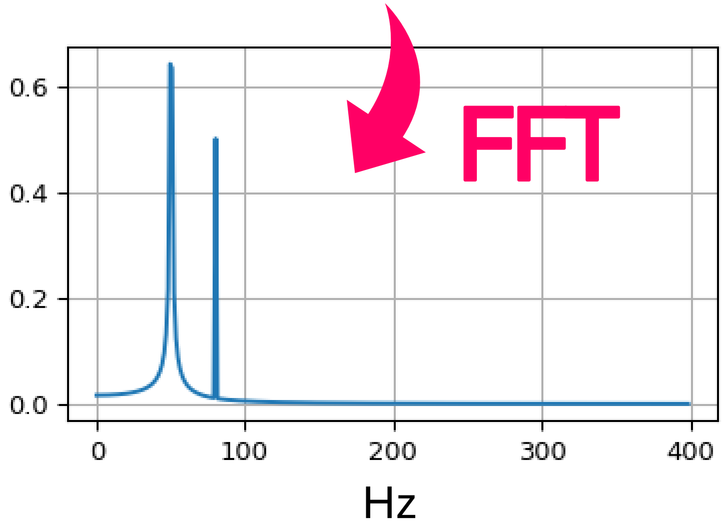
FFT de Tempo Curto

Em outras palavras, enquanto a FFT nos fornece QUAIS são as frequências base que construíram o nosso sinal, não da pra saber QUANDO. Por isso precisamos quebrar a FFT em pequenos intervalos de tempo (energia/comp de onda) para saber em QUAIS intervalos tal frequencai estava MAIS ou MENOS presente, permitindo assim avaliar a sua VARIAÇÃO.

Portanto, consideramos o sinal original em vários segmentos (ou frames) e aplicamos a FFT localmente.

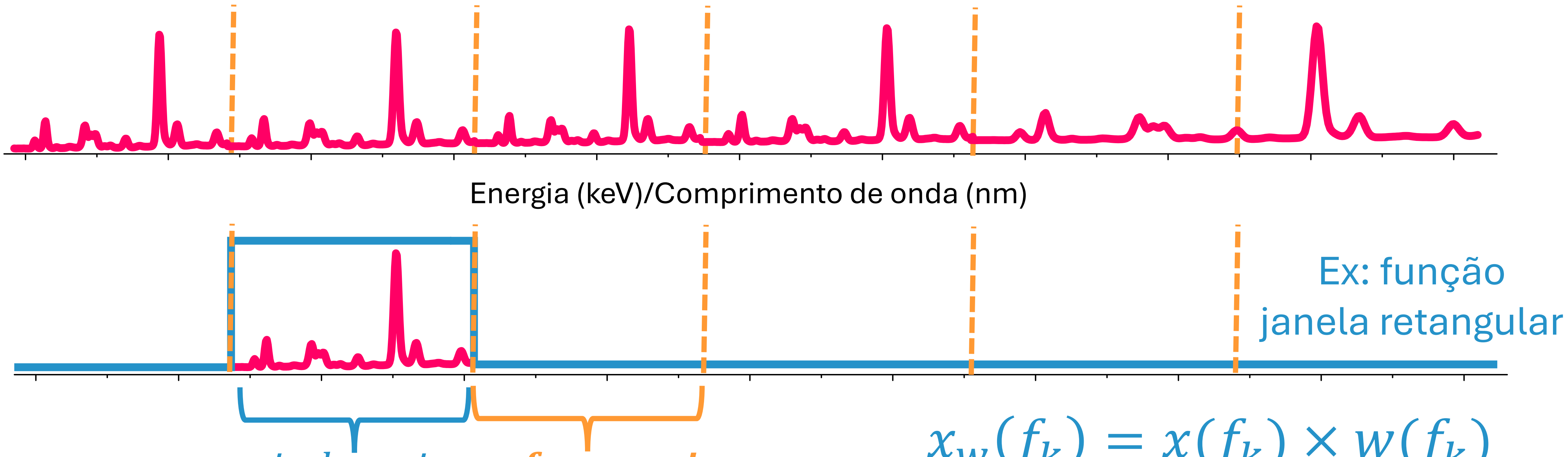


Energia (keV)/Comprimento de onda (nm)



A técnica de janelamento é aplicada a cada segmento antes de passar pela FFT

Técnica de janelamento



$\mathit{window\ size} = \mathbf{\mathit{frame\ size}}$

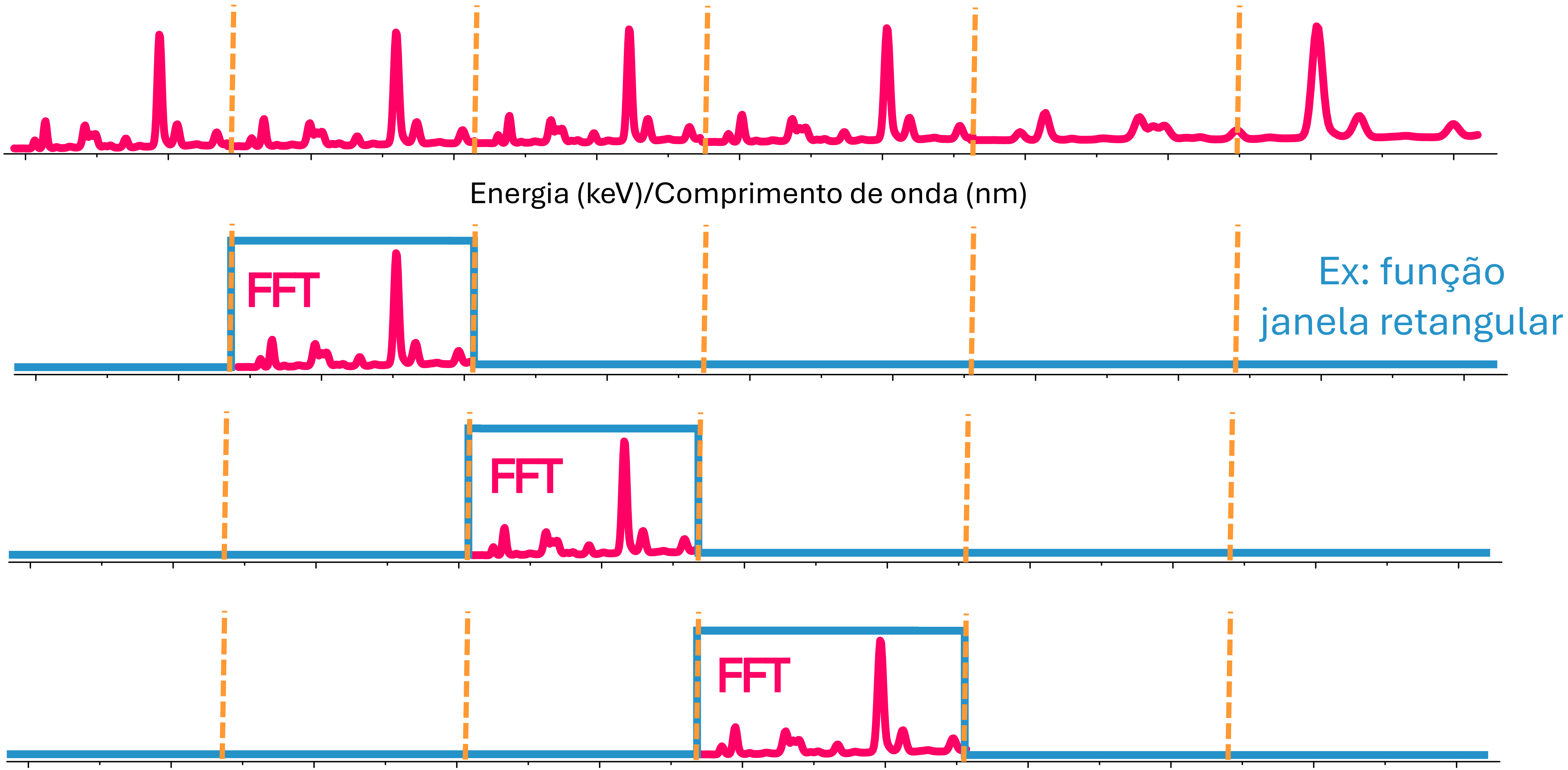
$$x_w(f_k) = x(f_k) \times w(f_k)$$

$$w(f_k) = 0\ se\ |f_k| > window\ size;\ w(f_k) = 1\ se |f_k| < window\ size$$

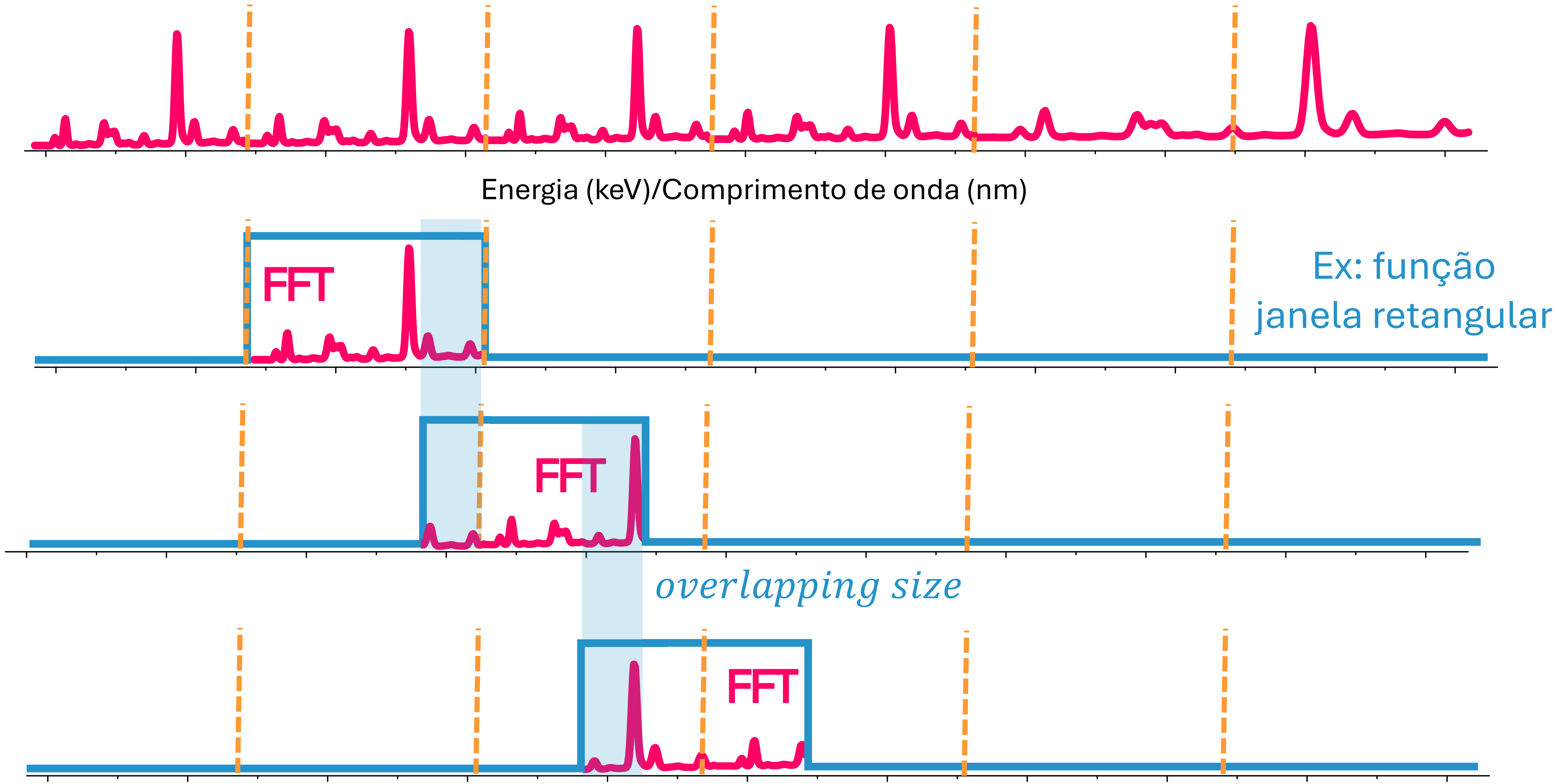
O *window size* representa o numero de canais que vão ser amostrados na janela

Já o ***frame size*** é o numero de canais que vão ser amostrados em cada segmento passado para a FFT de tempo curto (ele é definido primeiro). Usualmente eles são iguais

Técnica de janelamento

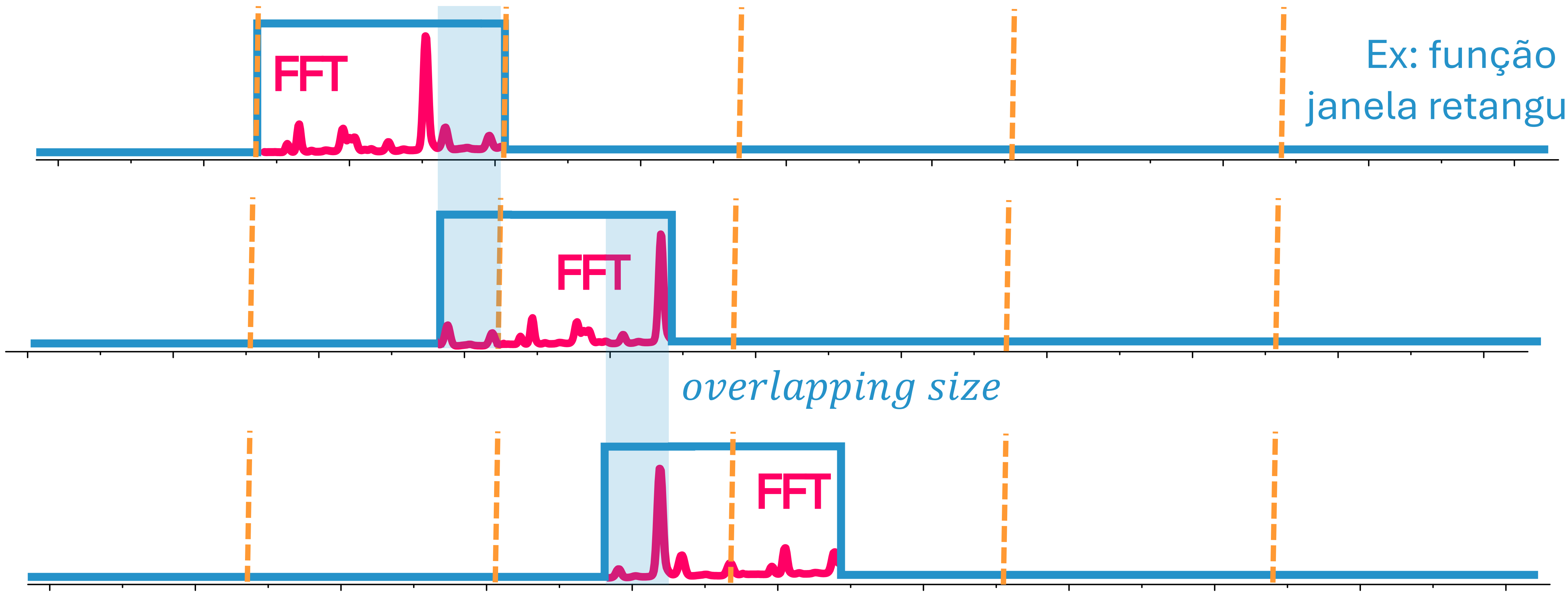


Técnica de janelamento - sobreposição



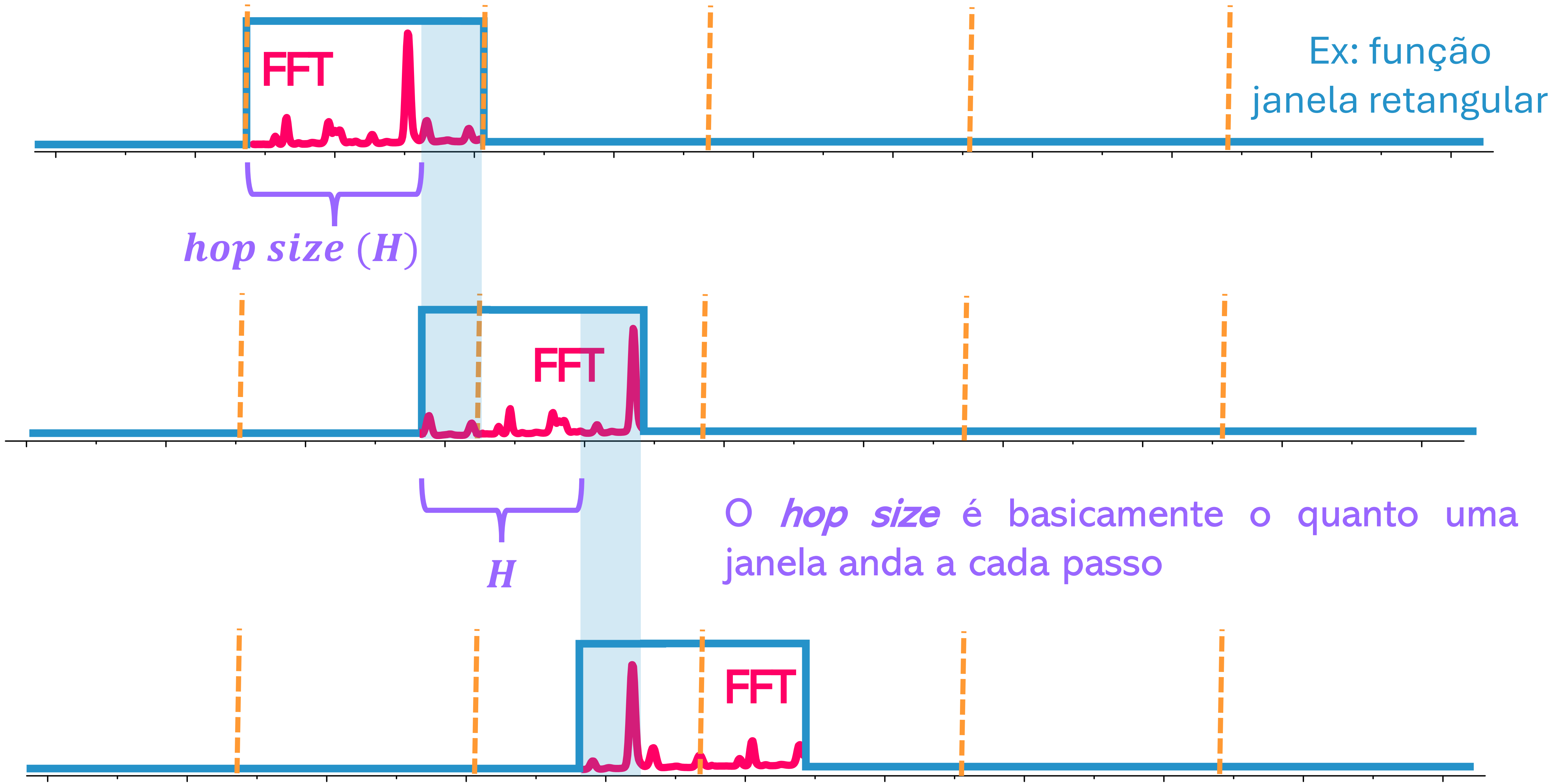
Técnica de janelamento - sobreposição

A sobreposição ajuda a capturar transições suaves no sinal entre uma janela e outra. Ela é usada para garantir que não hajam “buracos” na cobertura do espectro quando a janela “desliza”



Ex: função janela retangular

Técnica de janelamento - sobreposição



O *hop size* é basicamente o quanto uma janela anda a cada passo

De FFT para STFFT

Transformada Discreta (sinal estacionário)

$$x(k) = \sum_{n=0}^{N-1} x(n) e^{-i2\pi n \frac{k}{N}}$$

$t_n = n\Delta t; \quad n, k = 0,1, \dots N-1$

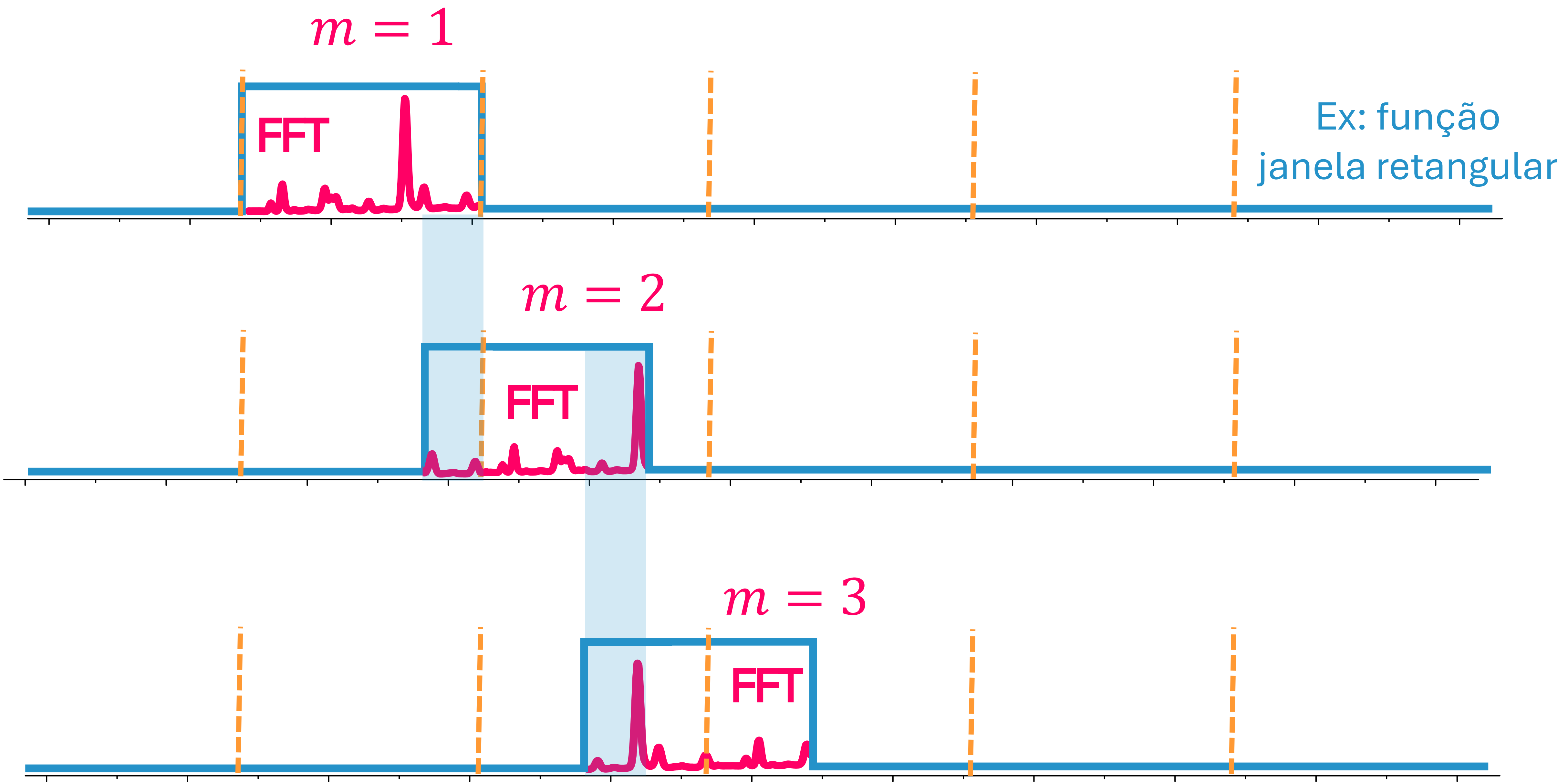
$x_n = x(t_n) = x(n)$

Transformada Discreta de Tempo Curto (sinal não-estacionário)

$$x(k, m) = \sum_{n=0}^{N-1} x(n + mH) w(n) e^{-i2\pi n \frac{k}{N}}$$

O m representa o frame (segmento) no qual a transformada está atuando no momento

De FFT para STFFT



De FFT para STFFT

Transformada Discreta (sinal estacionário)

$$x(k) = \sum_{n=0}^{N-1} x(n) e^{-i2\pi n \frac{k}{N}}$$

$$t_n = n\Delta t; \quad n, k = 0,1, \dots N-1$$

$$x_n = x(t_n) = x(n)$$

Transformada Discreta de Tempo Curto (sinal não-estacionário)

$$x(k, m) = \sum_{n=0}^{N-1} x(n + mH) w(n) e^{-i2\pi n \frac{k f_s}{N}}$$

A resposta da transformada vai será a k –ésima amplitude de frequência do m –ésimo frame (segmento)

De FFT para STFFT

Transformada Discreta
(sinal estacionário)

$$x(k) = \sum_{n=0}^{N-1} x(n) e^{-i2\pi n \frac{k}{N}}$$

$$t_n = n\Delta t; \quad n, k = 0,1, \dots N-1$$

$$x_n = x(t_n) = x(n)$$

Transformada Discreta de Tempo Curto
(sinal não-estacionário)

$$x(k,m) = \sum_{n=0}^{N-1} x(n + mH) w(n) e^{-i2\pi n \frac{k f_s}{N}}$$

O N na transformada de tempo curto representa todas as amostras existentes em CADA frame (segmento), não mais todos os pontos do dadaset. Já o n denota o n-ésimo ponto existente neste frame

De FFT para STFFT

Transformada Discreta
(sinal estacionário)

$$x(k) = \sum_{n=0}^{N-1} x(n) e^{-i2\pi n \frac{k}{N}}$$

$$t_n = n\Delta t; \quad n, k = 0,1, \ldots N-1$$

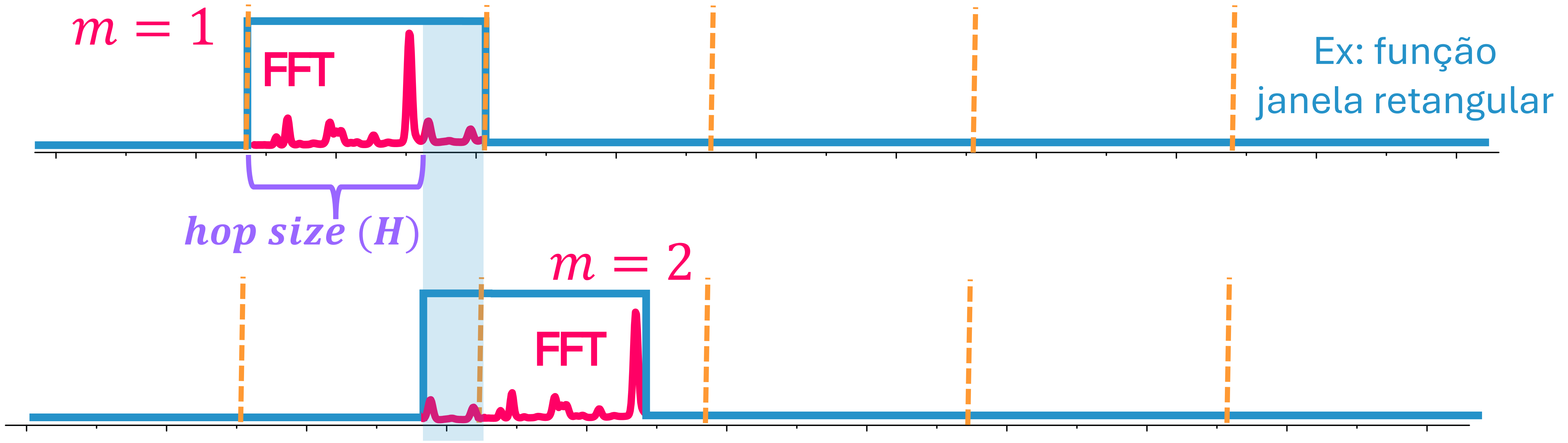
$$x_n = x(t_n) = x(n)$$

Transformada Discreta de Tempo Curto
(sinal não-estacionário)

$$x(k, m) = \sum_{n=0}^{N-1} x(n + mH) w(n) e^{-i2\pi n \frac{k f_s}{N}}$$

Já o termo $x(n + mH)$ indica que estamos somando sobre todos os pontos (n) do sinal presentes no m-ésimo frame (segmento) específico

Técnica de janelamento - sobreposição



$$x(k,m) = \sum_{n=0}^{N-1} x(n+mH)w(n)e^{-i2\pi n\frac{kf_s}{N}}$$

Perceba que se n denota os pontos coletados dentro de cada frame (segmento), mH indica a primeira amostra em cada frame. Logo, a medida que m varia temos a ideia de passo deslizante sobre os frames no qual vamos somando sobre todos os n pontos existentes

De FFT para STFFT

Transformada Discreta (sinal estacionário)

$$x(k) = \sum_{n=0}^{N-1} x(n)e^{-i2\pi n\frac{k}{N}}$$

$$t_n = n\Delta t; \quad n, k = 0,1, \dots N-1$$

$$x_n = x(t_n) = x(n)$$

Transformada Discreta de Tempo Curto (sinal não-estacionário)

$$x(k,m) = \sum_{n=0}^{N-1} x(n+mH)w(n)e^{-i2\pi n\frac{kf_s}{N}}$$

O termo $w(n)$ corresponde a janela que será aplicada em cada frame, englobando todos n-pontos presentes

FFT de Tempo Curto - Janelas

A janela retangular é só para exemplificar. Na verdade as principais janelas utilizadas são mais sofisticadas:

$$w(n) = 0.5 - 0.5cos\left(\frac{2\pi n}{N-1}\right)$$

$$0 \leq n \leq N-1; \; w(0) = w(N-1) = 0$$

Janela de Hann

$$w(n) = 0.54 - 0.46cos\left(\frac{2\pi n}{N-1}\right)$$

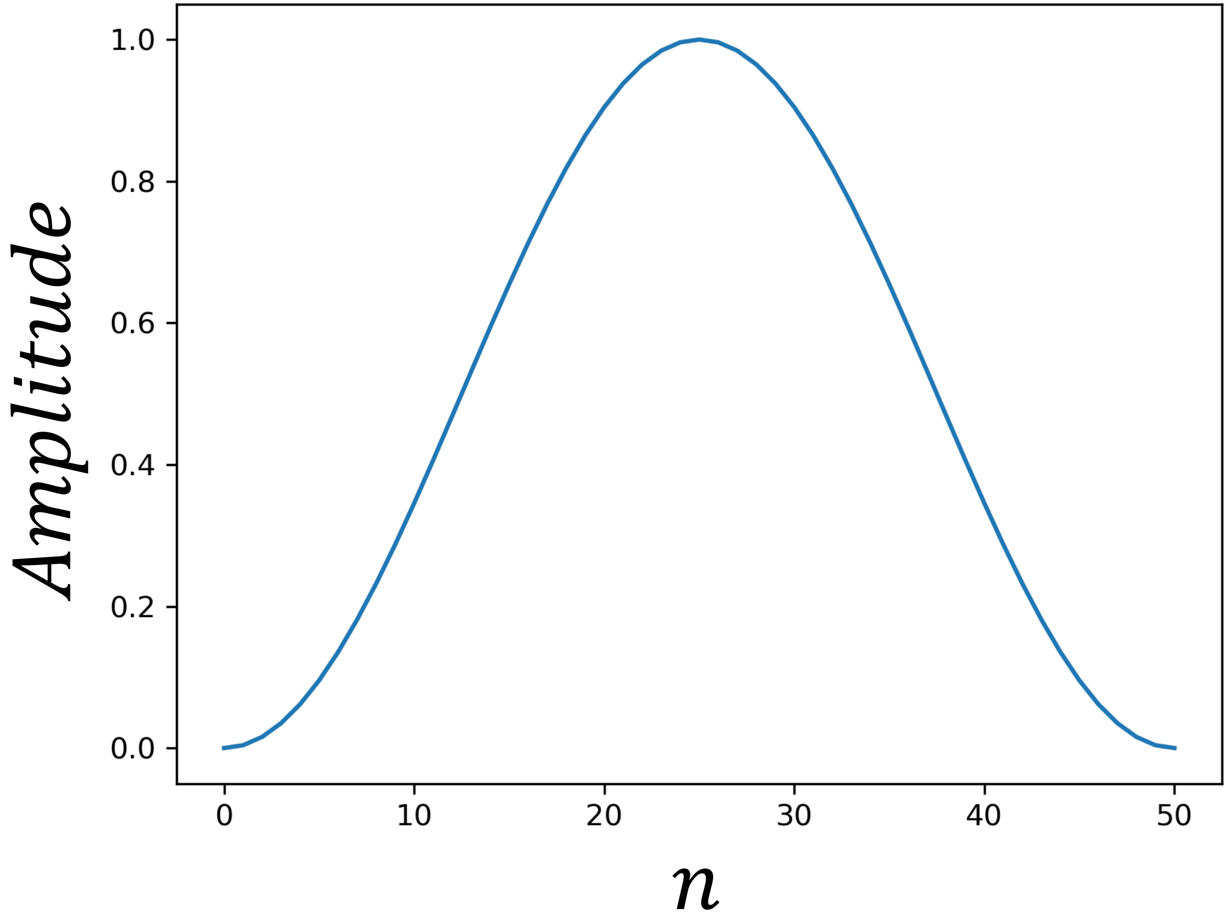
$$0 \leq n \leq N-1; \;\; w(0) = w(N-1) \neq 0.$$

Janela de Hamming

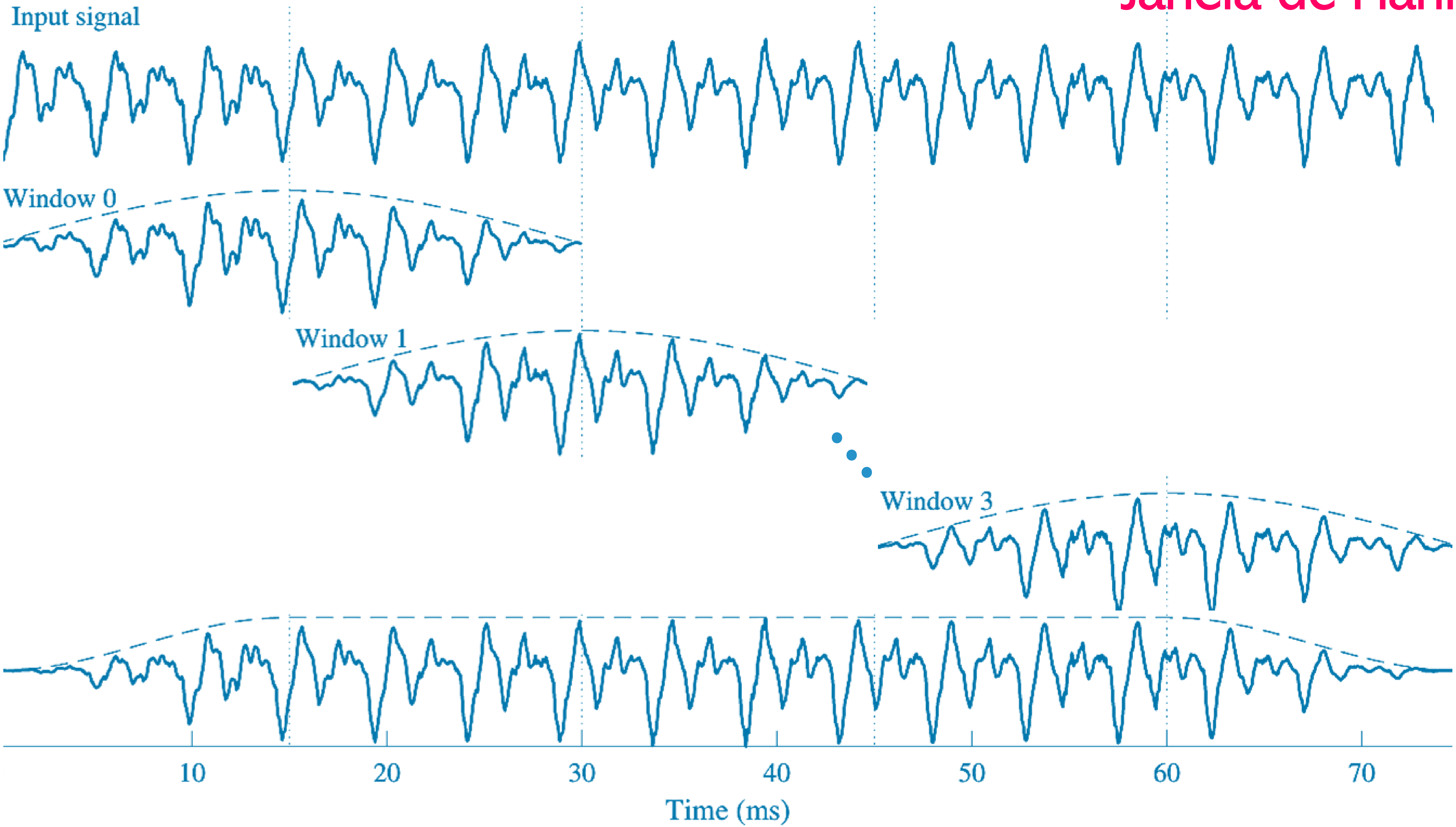
$$w(n) = 0.5 - 0.5cos\left(\frac{2\pi n}{N-1}\right)$$

Janela de Hann

$$0 \leq n \leq N-1;\ \ w(0) = w(N-1) = 0$$



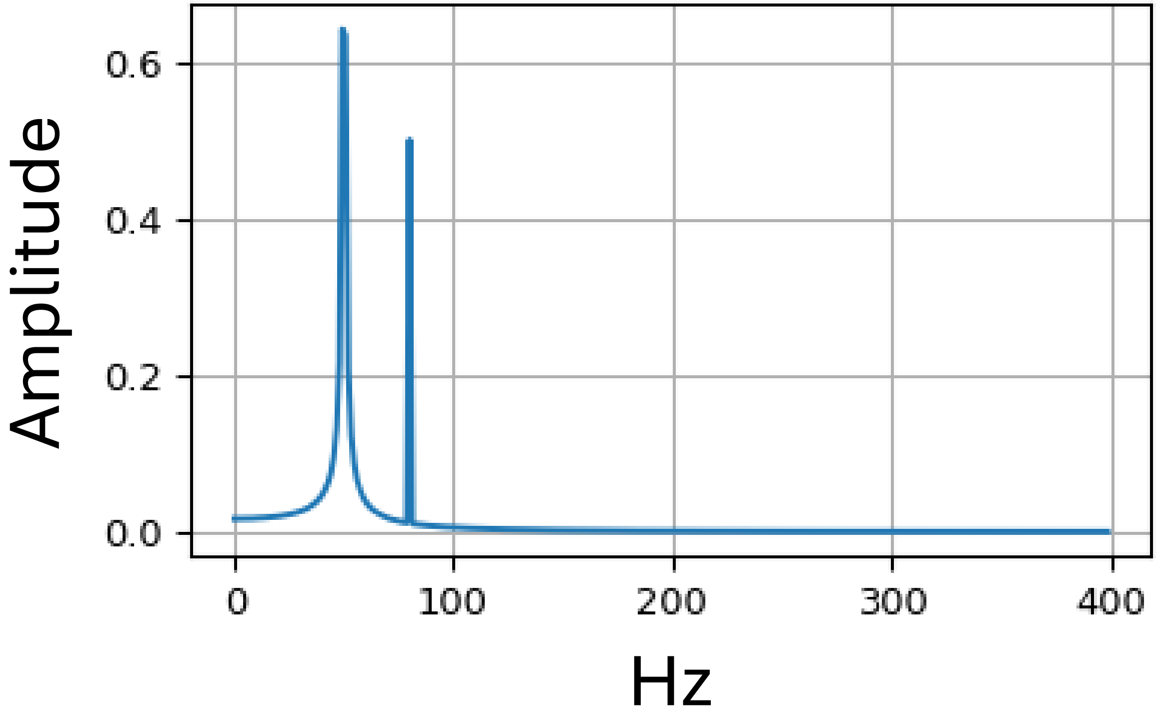
Janela de Hann



FFT de Tempo Curto

Diferença nos resultados gerados:

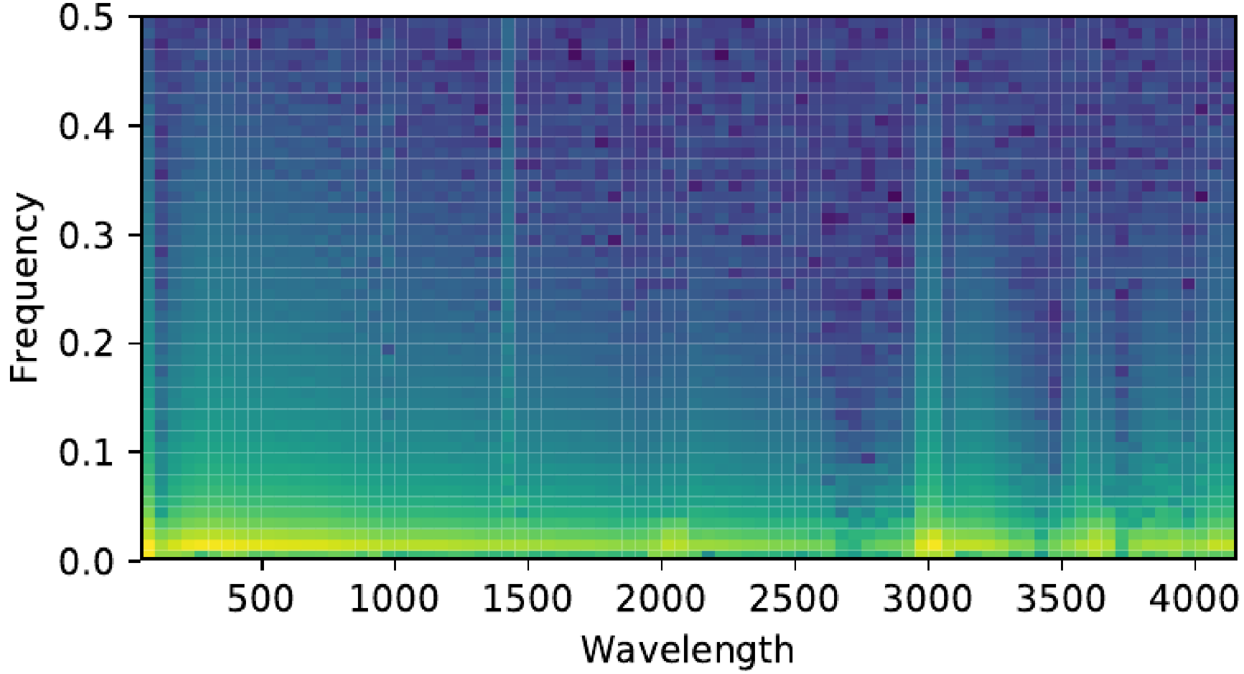
Na FFT normal, temos como resposta um vetor 1D que tem como coeficientes as amplitudes de frequências presentes no sinal original, sem fazer menção a sua variação ao longo do tempo (energia/comp de onda)



FFT de Tempo Curto

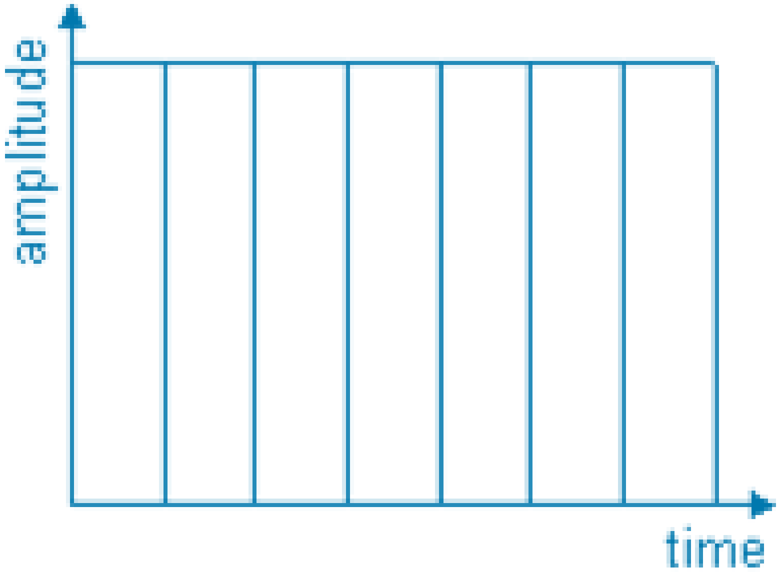
Diferença nos resultados gerados:

Já na STFFT temos como resposta o mesmo vetor 1D, porém, um para cada frame (segmento). Isso gera uma matriz 2D com amplitudes de frequências (frequency bins) versus os frames (variável temporal), fornecendo como elas variam ao longo do tempo ou das variáveis adotadas (energia/comp de onda)

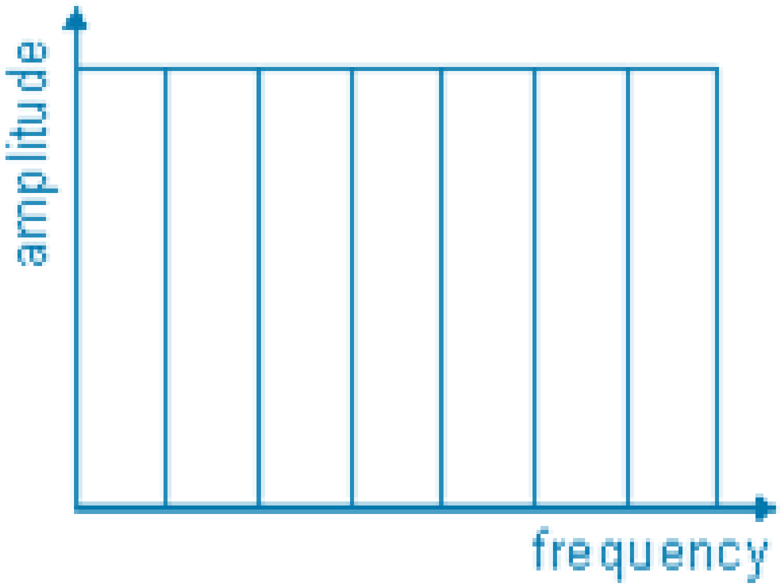


FFT de Tempo Curto

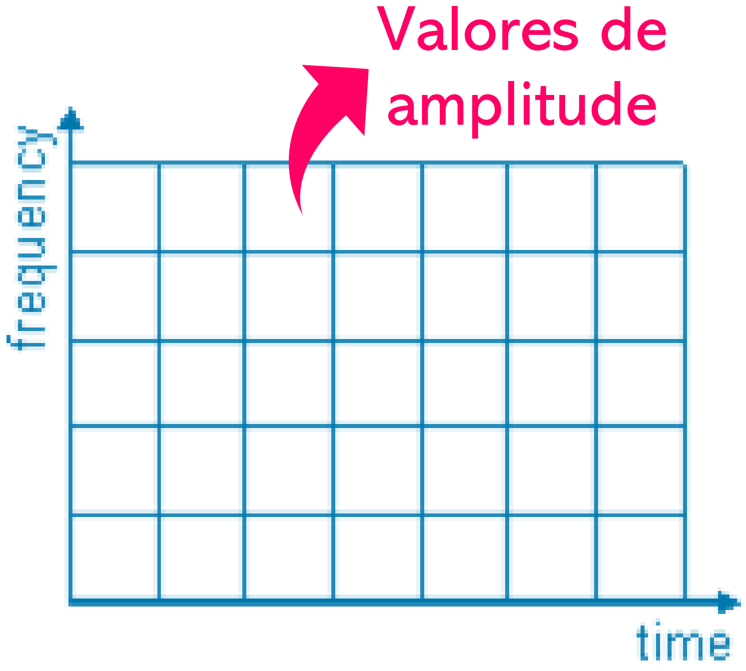
Diferença nos resultados gerados:



Domínio do tempo



Domínio da frequência (FFT)



Short Time FFT

Dimensão das respostas da STFFT

Número de pontos incluídos dentro de cada frame (segmento)

$$\# \, frequency \ bins = \frac{frame \ size}{2} + 1$$

Numero de faixas de frequência englobadas ao longo da STFFT

O numero de variáveis. É como o numero total de "pontos" medidos

$$\# \, frames = \frac{variables \ - frame \ size}{hop \ size} + 1$$

O quanto a janela anda a cada passo

Dimensão das respostas da STFFT

Ex: Um sinal amostrado em 10000 intervalos de tempo (variáveis) cujo o qual aplicamos a STFFT com frame size = 1000 e hop size = 500 gera uma matriz 2D de dimensões:

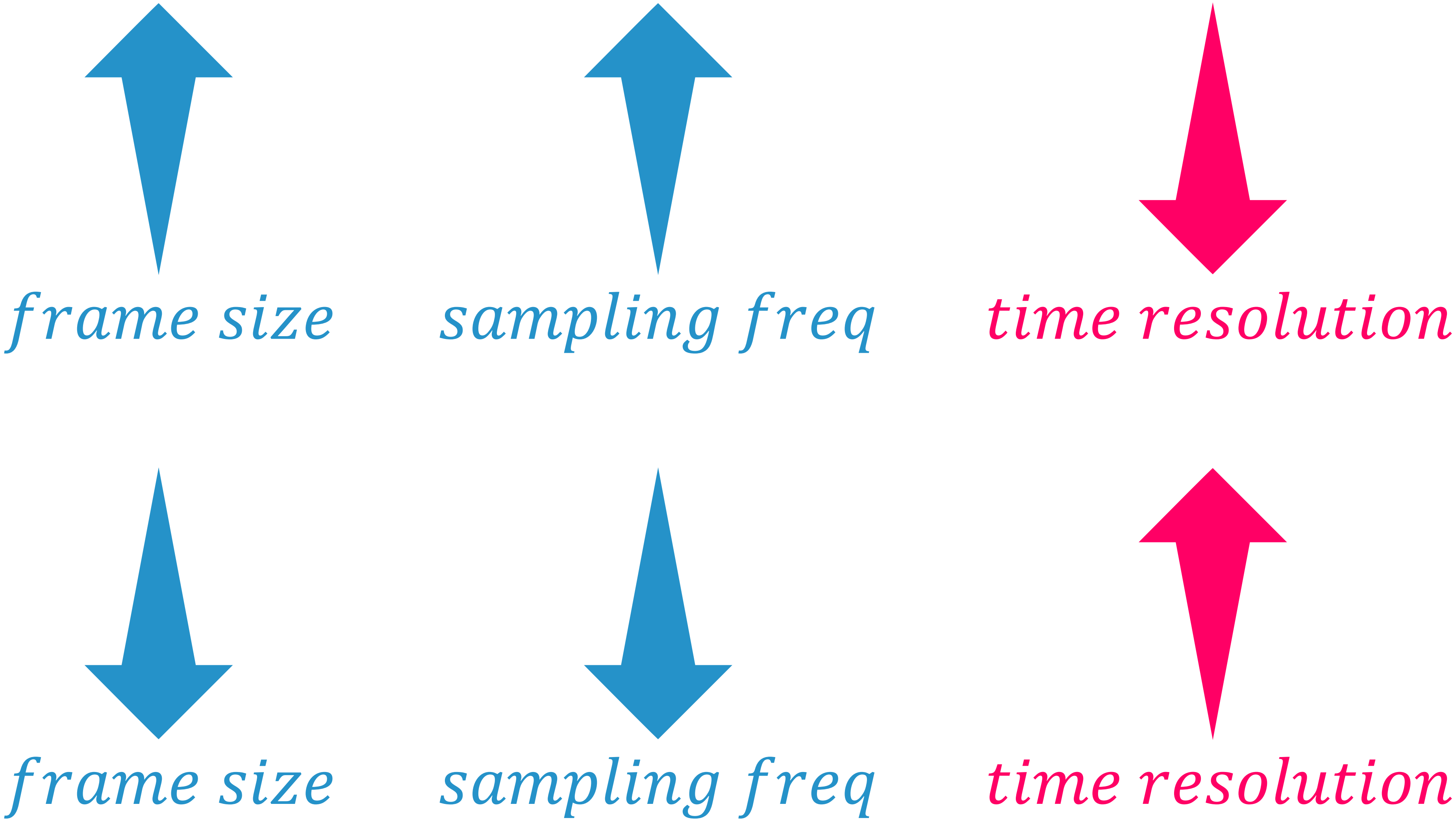
$$\#\, frequency\, bins = \frac{1000}{2} + 1 = 501 \quad \rightarrow \left[0, {}^{f_s}/_{2}\right] \quad f_s = 205.5\ Hz$$

$$\#\, frames = \frac{10000 - 1000}{500} + 1 = 19$$

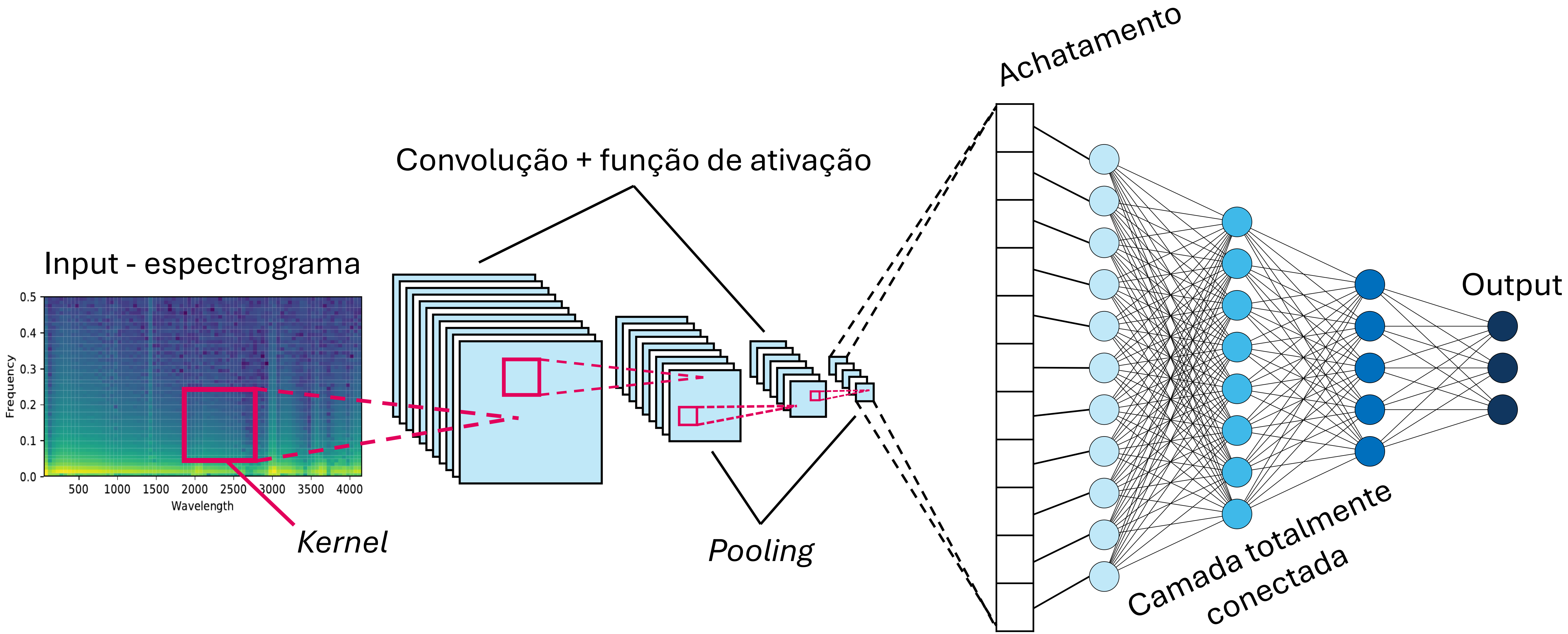
$$[ST]_{501\times19} = \begin{bmatrix} ST_{1\times1} & \cdots & ST_{1\times19} \\ \vdots & \ddots & \vdots \\ ST_{501\times1} & \cdots & ST_{501\times19} \end{bmatrix}$$

O componente da i-ésima linha carrega as intensidades das frequências (decompostas em 501 intervalos) presentes no respectivo i-ésimo frame

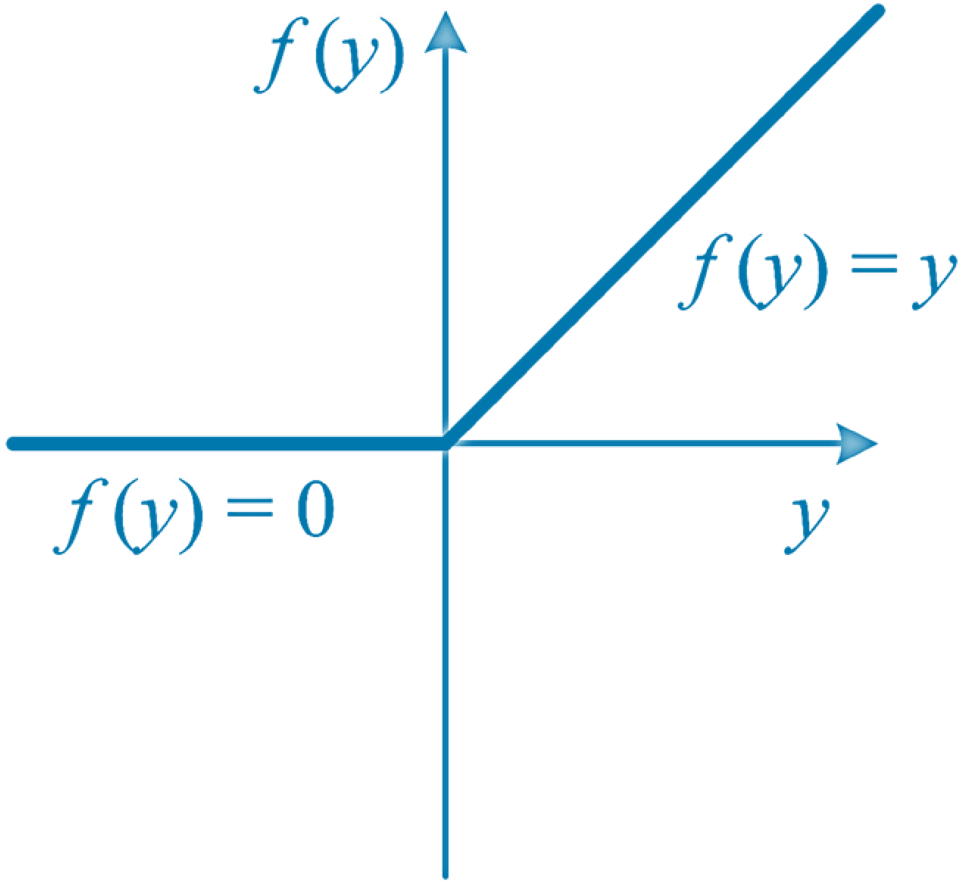
Time / Frequency trade off



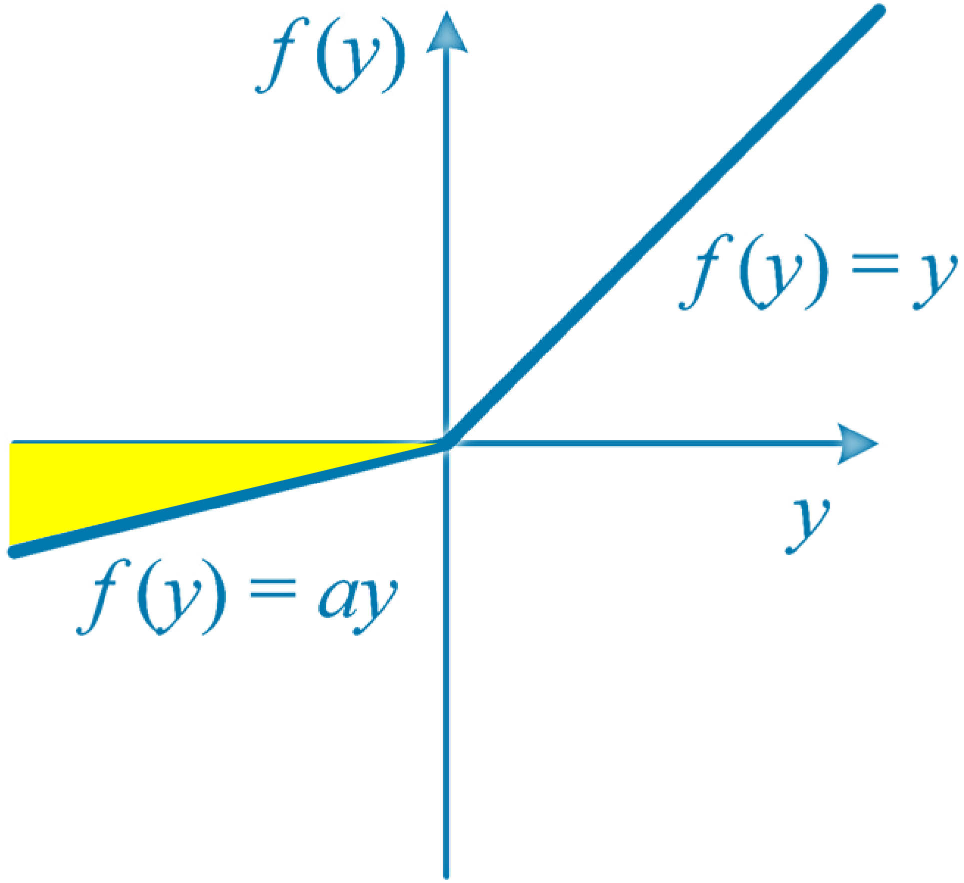
CNN 2D - espectrogramas



Funções de ativação Leaky e Parametric ReLU

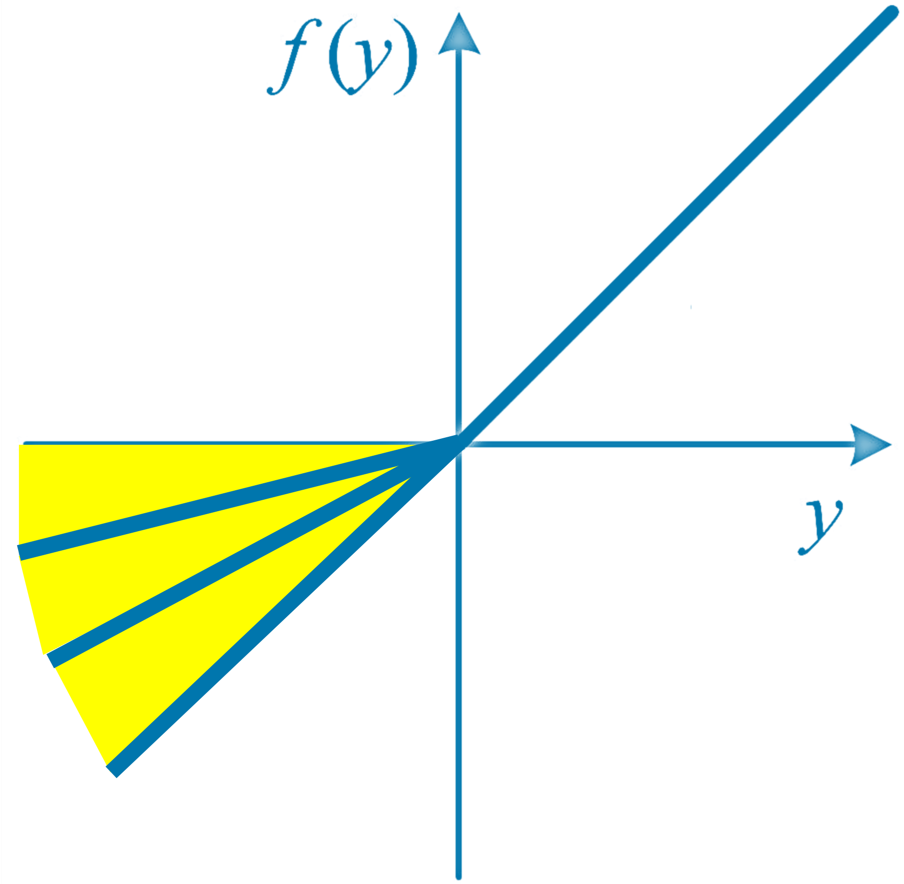


ReLU



Leaky ReLU

Funções de ativação Leaky e Parametric ReLU



Parametric ReLU - PReLU

$$f(y_i) = y_i, \qquad \forall\, y_i \geq 0$$

Introduz um parâmetro aprendível α_i para controlar a inclinação da parte negativa da função. A camada PReLU aprende um α distinto para cada elemento da entrada

$$f(y_i) = \alpha_i y_i, \qquad \forall\, y_i < 0$$

PRÁTICA – GOOGLE COLAB