



**INSTITUTO
FEDERAL**

Brasília

Instituto Federal de Brasília

Campus Estrutural

Licenciatura em Matemática

Arthur Augusto Campos

**INTRODUÇÃO À ANÁLISE DE SÉRIES TEMPORAIS UTILIZANDO A
TRANSFORMADA DE FOURIER DISCRETA EM JANELAS DESLIZANTES**

Brasília

2025

Arthur Augusto Campos

**INTRODUÇÃO À ANÁLISE DE SÉRIES TEMPORAIS UTILIZANDO A
TRANSFORMADA DE FOURIER DISCRETA EM JANELAS DESLIZANTES**

Projeto de Pesquisa apresentado à disciplina Projeto de Conclusão de Curso do curso de Licenciatura em Matemática para obtenção de nota parcial no Instituto Federal de Brasília

Orientador: Doutor Wembesom Mendes Soares
Corientador: Doutor Rodrigo Figueiredo Abdo

Brasília
2025

LISTA DE ILUSTRAÇÕES

Figuras

Figura 1 - Preço de fechamento ajustado das ações da NVIDIA	13
Figura 2 - ECG bruto	13
Figura 3 - ECG discretizado	14
Figura 4 - Série de Fourier quando $n=1$	18
Figura 5 - Série de Fourier quando $n = 5$	18
Figura 6 - Série de Fourier quando $n=15$	19
Figura 7 - Representação visual da DFT	24
Figura 8 - Exemplo visual da SWDFT	28

Quadros

1	Primeiros 10 pontos e seus respectivos coeficientes	26
2	Tabela da proporção de outliers por estado	30

Gráficos

1	Gráfico da função sintética e os pontos amostrados	25
2	Gráfico no domínio da frequência	26

LISTA DE TABELAS

Tabela 1 - Tabela da proporção de NA's por estado	31
Tabela 2 - Tabela de proporção de outliers por estado pós reconstrução . . .	31

LISTA DE ABREVIATURAS E SIGLAS

IFB - Instituto Federal de Brasília

DFT - Transformada de Fourier Discreta

SWDFT - Transformada de Fourier Discreta em Janelas Deslizantes

PPC - Projeto Político Pedagógico de Curso

ECG - Eletrocardiograma

EEG - Eletroencefalografias

CNN - Redes neurais convolucionais

TEA - Transtorno do Espectro Autista

TCC - Trabalho de Conclusão de Curso

SUMÁRIO

Lista de Ilustrações

Figuras

Quadros

Gráficos

Tabelas

1	INTRODUÇÃO E JUSTIFICATIVA	6
1.1	O que o PPC do curso diz	9
1.2	Metodologia	10
2	FERRAMENTAL BÁSICO	12
2.1	Séries Temporais	12
2.2	Séries de Fourier	15
2.2.1	<i>A série de Fourier complexa</i>	19
3	TRANSFORMADA DE FOURIER	21
3.1	Transformada de Fourier Discreta	23
3.2	SWDFT	28
4	APLICAÇÃO SWFFT	29
5	RESULTADOS OBTIDOS	32
	Referências	33

1 INTRODUÇÃO E JUSTIFICATIVA

Este projeto tem como objetivo desenvolver um Trabalho de Conclusão de Curso (TCC) que, por sua vez, será um material introdutório sobre a Transformada Discreta de Fourier em Janelas Deslizantes para a análise de séries temporais. Esse material apresentará a técnica, que atua nos domínios do tempo e da frequência, e destacará sua ampla utilização devido ao menor custo computacional em comparação com métodos como a Fast Fourier Transform, por exemplo. A seguir, apresentaremos uma introdução ao tema e discutiremos a relevância de sua abordagem no curso de Licenciatura em Matemática.

A matriz curricular do curso de Licenciatura em Matemática no Instituto Federal de Educação, Ciência e Tecnologia de Brasília (IFB) é ampla. Entretanto, ainda mais diversos são os tópicos interessantes da Matemática. Nesse sentido, alguns objetivos do curso estão direcionados ao encaminhamento dos alunos para esses temas, a fim de fornecer uma formação completa dos graduandos. Para tal, os componentes curriculares do Núcleo Específico são bons caminhos para despertar interesse por conteúdos extracurriculares.

Entre os diversos tópicos interessantes da Matemática, um que se destaca por sua elegância, versatilidade e relevância histórica são as Séries de Fourier. Para Boyer e Merzbach (2012), durante a Revolução Francesa, as universidades do país não eram referências em Matemática como são hoje. Contudo, após a Revolução, o sistema de educação superior do país foi reformulado, recolocando a França como um dos principais centros matemáticos do mundo. Nesse contexto, entre autores como Lagrange, Laplace e d'Alembert, destacava-se **Jean Baptiste Joseph Fourier (1768-1830)**.

Segundo Boyer e Merzbach (2012) a trajetória do matemático teve início durante a Revolução Francesa, quando ensinou em sua cidade natal, Auxerre. Fourier teve uma atuação política intensa e, por esse motivo, foi preso. Quando livre, tornou-se assistente de Lagrange e Monge na École Polytechnique. Em 1822, foi eleito *secrétaire perpétuel* da Académie des Sciences de Paris, e ainda nesse ano, publicou sua célebre obra *Théorie analytique de la chaleur*.

Pifer e Aurani (2015) afirmam que, em sua obra, quebrando paradigmas da época a respeito da natureza do calor, Fourier reinterpretou o problema de propagação de calor e ampliou as aplicações do cálculo diferencial. A partir da sua interpretação do problema de condução do calor à luz de equações diferenciais parciais, Fourier desenvolveu um método matemático para encontrar soluções para o problema usando a soma de senóides.

Um estudante do curso de Licenciatura em Matemática, após as matérias de Equações Diferenciais Ordinárias, Variáveis Complexas e Cálculo Numérico, deve estar familiarizado com as séries de Taylor, que têm um papel semelhante ao das séries de

Fourier. Entretanto, um dos grandes diferenciais entre essas séries é que as de Fourier não têm tantas exigências técnicas como as de Taylor. Kreyszig (2011), aponta que muitas funções descontínuas e periódicas, que surgem em aplicações, podem ser desenvolvidas como séries de Fourier mas não como expansões das séries de Taylor.

Tome por exemplo a função f de período 2π definida no intervalo $[-\pi, \pi)$

$$f(x) = \begin{cases} 0 & \text{se } -\pi \leq x < 0 \text{ ou } x = \pi, \\ \pi & \text{se } 0 \leq x < \pi. \end{cases}$$

Essa função não tem expansão por séries de Taylor porque, tomando $a = 0$, não existe o limite

$$f'(a) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$$

e, portanto, ela não é derivável no intervalo $[-\pi, \pi)$. Ainda assim, essa função pode ser escrita como a série de Fourier:

$$S[f] = \frac{\pi}{2} + \sum_{k=0}^{+\infty} \left(\frac{2}{2k+1} \text{sen}((2k+1)x) \right)$$

Com o intuito de ampliar a quantidade de funções que podem ser expressas com técnicas derivadas das séries de Fourier, é natural prosseguir com o estudo das transformadas de Fourier. Afinal, como Kreyszig (2011) evidencia, muitos problemas reais envolvem funções que não são periódicas com domínio em \mathbb{R} . Para lidar com tais funções, podemos usar a Transformada de Fourier Contínua. Poderíamos, por exemplo, utilizar a transformada de Fourier contínua para representar uma função que descreve a evolução da posição de uma partícula que se move de forma aleatória ao longo do tempo.

Para tratar de funções com características semelhantes ao exemplo anterior, podemos buscar analisar séries temporais, que tratam de processos que evoluem ao longo do tempo. Ainda sobre o exemplo, é natural o questionamento: “quais séries temporais podem ser representadas por uma Transformada Contínua de Fourier?”. A resposta para essa pergunta depende ao menos da interpretação do intervalo de tempo em que a série temporal está sendo analisada. Para realizar essa transformada, é necessário que a série tenha como domínio um período de tempo que pode ser associado a algum intervalo não degenerado na reta real. Fazendo essa transformada, obtemos uma nova função definida no domínio da frequência.

Segundo Percival e Walden (1993), alguns problemas precisam ser tratados no domínio da frequência, também chamado espectro. Então, é possível pensar na Transformada Contínua de Fourier para descrever séries temporais contínuas, no domínio espectral, a fim de melhorar as análises. Um dos problemas dessa abordagem é o elevado poder computacional requerido, mas um desafio ainda mais frequente é que os

intervalos de tempo nas amostras de dados podem não ser contínuos. Nesses casos, a Transformada de Fourier contínua não pode ser aplicada

Essa limitação no domínio é recorrente quando tratamos de processamento de sinais digitais, que desempenham papel importante em áreas como o processamento de imagens e áudios, por exemplo. É argumentado por Osgood (2007) que mesmo os sinais contínuos devem ser processados por métodos discretos, porque computadores e sistemas de processamento digital podem trabalhar apenas com somas finitas. Para trabalhar com o domínio discreto, utilizamos a Transformada de Fourier Discreta (DFT).

Embora os estudos do tempo e da frequência sejam frequentemente tratados separadamente, a necessidade de uma abordagem que considere ambos os domínios simultaneamente tem sido reconhecida desde os trabalhos pioneiros de Gabor (1946). Cohen (1994) em seu livro, apresenta algumas técnicas para a análise de sinais no domínio do tempo e da frequência. Uma dessas técnicas é apresentada por ele como “The Short Time Fourier Transform”, que também é conhecida como Sliding Window Discrete Fourier Transform (SWDFT). Richardson (2019), menciona em sua tese que o fato de a técnica SWDFT ser uma extensão direta da DFT a torna, dentre as várias técnicas, a mais fácil de compreender.

Portanto, o objetivo do trabalho, desenvolvido a partir deste projeto, é explorar a técnica SWDFT, proporcionando aos licenciandos uma ferramenta versátil para a análise de séries temporais em diversos contextos. Para isso, serão introduzidas noções fundamentais de séries temporais, seguidas pelo estudo das séries de Fourier com exemplos ilustrativos, da Transformada de Fourier e da Transformada de Fourier Discreta, estabelecendo um percurso lógico que conduz ao entendimento da SWDFT na análise de séries temporais.

Para que um graduando do curso de Licenciatura em Matemática do IFB compreenda esses tópicos, é necessário um certo grau de maturidade matemática. Isso ocorre porque o ferramental exige conhecimentos prévios em disciplinas do Núcleo Específico, como Álgebra Linear, Estatística, Variáveis Complexas, entre outras. Por abranger diversos conteúdos, o estudo da técnica promove uma revisão significativa da grade curricular, contribuindo para o primeiro objetivo do Projeto Político Pedagógico do Curso (PPC) de “a) Formar professores para os ensinos fundamental, médio e profissional, com amplo domínio dos conhecimentos específicos em torno dos quais deverá atuar”(BRASIL, 2018, p.15).

Não obstante, o estudo dos métodos de análise de séries temporais pode contribuir com a educação no geral. Mao *et al.* (2024) em seu artigo, sugerem que realizar a análise de séries temporais do histórico acadêmico de alunos pode ajudar a obter *insights* e intervir na prática docente, quando necessário. Ainda em sua pesquisa, os autores revelam outras vantagens em análises nessa área como, prever tendências e com-

portamentos, permitindo que educadores, administradores e políticos tomem decisões baseadas em dados.

A análise de séries temporais também encontra uso na reflexão crítica sobre os aspectos financeiros sociais. Tome como exemplo o trabalho de Barbosa e Germano (2023) que sugere, para estudantes de Licenciatura em Matemática, uma modelagem matemática do índice de inadimplentes no Serasa. Esse trabalho evidencia que a análise de séries temporais pode ser uma ferramenta que ajuda a contribuir com o objetivo terceiro específico do PPC do curso de Matemática: “c) Estimular nos futuros professores a prática reflexiva, a fim de que os licenciados vivenciem, enquanto discentes, experiências educativas que contribuam para a sua prática profissional futura”(BRASIL, 2018, p.16).

Barros (2019) contribui com a educação através da análise de séries temporais ao realizar, em sua monografia, uma investigação para conhecer o perfil epidemiológico e a tendência da mortalidade de professores no Estado de Goiás. Pesquisas como essa podem servir como ferramenta de suporte para a construção de políticas públicas para professores.

Gade (2021) aponta que a quantidade de dados que se produz atualmente levou organizações a adotarem a estratégia crítica de tomar decisões baseadas na análise de dados. Assim, se faz necessário a formação de professores que sejam capazes de propiciar certo nível de letramento estatístico a seus alunos. Mesmo que não se tornem analistas de dados, para Engel e Frischmeier (2018), é importante que esses estudantes desenvolvam uma compreensão estatística para participar de debates públicos.

1.1 O que o PPC do curso diz

Este projeto está alinhado com o PPC no que trata dos seguintes objetivos do curso:

d) Criar e estimular condições e oportunidades que promovam a reflexão teórica e prática sobre a Matemática, fazendo uso de novas tecnologias e compreendendo sua formação profissional como processo contínuo, autônomo e permanente.[...] h) Estimular a continuidade dos estudos no nível de pós-graduação. (BRASIL, 2018, p.16)

O Trabalho servirá como apresentação de um recurso utilizado em áreas correlatas à Matemática, estimulando a reflexão teórica e prática sobre as matérias do curso, a partir da técnica proposta. O conhecimento de uma nova técnica revela a importância da formação continuada e amplia a percepção da Matemática para além das salas de aula. Além disso, o trabalho pode servir como material introdutório para tópicos mais

avanzados na área de Matemática aplicada e pura, como a análise harmônica, por exemplo.

Ainda, o Trabalho pode contribuir com o perfil profissional do egresso, uma vez que se espera do licenciando em Matemática as seguintes habilidades e competências:

h) Capacidade de compreender, criticar e utilizar novas ideias e tecnologias para a resolução de problemas. i) Capacidade de aprendizagem continuada, sendo sua prática profissional também fonte de conhecimento [...] k) Habilidade de estabelecer relações entre a Matemática e outras áreas de conhecimento [...] o) Competência potencial para realização de estudos de pós graduação. (BRASIL, 2018, p.18)

Assim sendo, e entendendo a Transformada Discreta de Fourier com Janela Deslizante (SWDFT) como uma técnica relevante para a análise de séries temporais, o objetivo do material proposto é apresentar essa abordagem de forma acessível a estudantes de Licenciatura em Matemática, destacando sua utilidade e aplicações. Ao explorar a SWDFT, busca-se evidenciar a importância da atualização constante sobre novas técnicas matemáticas, especialmente considerando o caráter interdisciplinar da Matemática e suas conexões com diversas áreas do conhecimento.

1.2 Metodologia

Considerando que o trabalho fruto deste projeto tem como finalidade investigar e coletar informações sobre a aplicação da técnica Sliding Window Discrete Fourier Transform (SWDFT) na análise de séries temporais, conforme proposto por Severino (2013), entende-se que a modalidade de pesquisa mais apropriada para esse estudo é a exploratória. Para compor esta pesquisa, está sendo realizado um levantamento bibliográfico e pesquisa documental.

Na pesquisa documental, foi realizada uma análise do PPC do curso de Licenciatura em Matemática a fim de respaldar a construção do TCC, como material introdutório, de acordo com os objetivos do curso apresentados nesse documento e ao perfil esperado do egresso.

O levantamento bibliográfico foi realizado em livros, artigos e dissertações que tratam da técnica principal do trabalho e seus principais temas correlatos. Esse tipo de levantamento “baseia-se no estudo da teoria já publicada, assim é fundamental que o pesquisador se aproprie no domínio da leitura do conhecimento e sistematize todo o material que está sendo analisado”(Sousa *et al.* 2021, p. 66). Seguindo esse propósito, foi elaborado um fichamento dos principais materiais necessários para a compreensão dos principais tópicos de forma isolada e conjunta.

Para selecionar os principais materiais, principalmente foram realizadas buscas nos sites *Google Acadêmico*, *IEEE Xplore*, *Springer Nature Link*, *Science Direct* e *arXiv*.

As buscas no portal de periódicos da CAPES não forneceram materiais relevantes até o momento. Contudo, para as próximas etapas da pesquisa, serão utilizadas buscas nesse e em outros portais de periódicos. Dada a natureza exploratória da pesquisa, a revisão bibliográfica se faz necessária ao longo de todo o processo, pois, como afirma Gil (2002, p.41)

Estas pesquisas têm como objetivo proporcionar maior familiaridade com o problema, com vistas a torná-lo mais explícito ou a constituir hipóteses. Pode-se dizer que estas pesquisas têm como objetivo principal o aprimoramento de ideias ou a descoberta de intuições. (GIL, 2002, p.41)

Com o objetivo de ampliar a quantidade de materiais encontrados, o site ChatGPT foi utilizado para sugerir diferentes buscas com as palavras-chave do levantamento. Além disso, as mesmas buscas foram realizadas também em inglês, a fim de expandir ainda mais o número de fontes consultadas. A construção das representações gráficas foi feita utilizando softwares gratuitos, como o GeoGebra e o R (R CORE TEAM, 2023).

No contexto matemático em que a técnica se insere, os materiais encontrados durante o levantamento bibliográfico se distanciam dos usualmente utilizados em um curso de Licenciatura em Matemática. Isso se deve, em parte, ao fato de muitos desses materiais serem voltados para cursos de engenharia ou estarem em um nível mais avançado, típico da pós-graduação. Assim, para que a ideia por trás das técnicas não ficasse perdida entre resultados da Matemática pura, optou-se por uma abordagem mais prática dos tópicos aqui tratados. Para o trabalho resultante deste projeto, serão aplicadas novas técnicas para o levantamento de materiais, tais como a bibliometria ou apenas a expansão das bases de pesquisa.

Apesar da omissão de alguns resultados em um primeiro momento, para as etapas seguintes da pesquisa, é relevante destacar que

O objetivo dos matemáticos é descobrir e comunicar certas verdades. A Matemática é a linguagem dos matemáticos, e a prova é o método de comunicar uma verdade matemática para outra pessoa que também “fala” essa língua. Uma propriedade notável da linguagem matemática é a sua precisão. Apresentada corretamente, uma prova não contém ambiguidades, isso é, não haverá dúvidas sobre sua validade. (SOLOW, 2013, p. 1, tradução nossa)⁶

Essa preocupação com a clareza na comunicação matemática reforça a necessidade de um material complementar que traduza conceitos avançados, como a técnica

⁶The objective of mathematicians is to discover and to communicate certain truths. Mathematics is the language of mathematicians, and a proof is a method of communicating a mathematical truth to another person who also “speaks” the language. A remarkable property of the language of mathematics is its precision. Properly presented, a proof contains no ambiguity—there will be no doubt about its correctness.

SWDFT, em linguagem adaptada ao contexto da Licenciatura, promovendo a mediação pedagógica entre teoria e prática docente.

Dessa forma, torna-se primordial, para futuros professores, a produção de um material com uma sistematização do percurso conceitual que conecte as séries temporais à técnica SWDFT, de modo acessível aos licenciandos em Matemática e com o devido rigor matemático. Para isso, a revisão contínua dos materiais separados no fichamento será de suma importância. Ainda, serão construídos novos exemplos que ilustrem os passos desenvolvidos, a fim de preservar o foco nas aplicações.

2 FERRAMENTAL BÁSICO

Nas subseções a seguir, estão introduzidos os conceitos de Séries Temporais, Séries de Fourier, Transformada de Fourier Discreta e Contínua e, por fim, há uma breve introdução da técnica SWDFT unindo os tópicos anteriores. Como este é apenas o projeto do trabalho, para fins de celeridade, deu-se preferência a exemplos e ilustrações em vez de expandir teoremas e proposições.

2.1 Séries Temporais

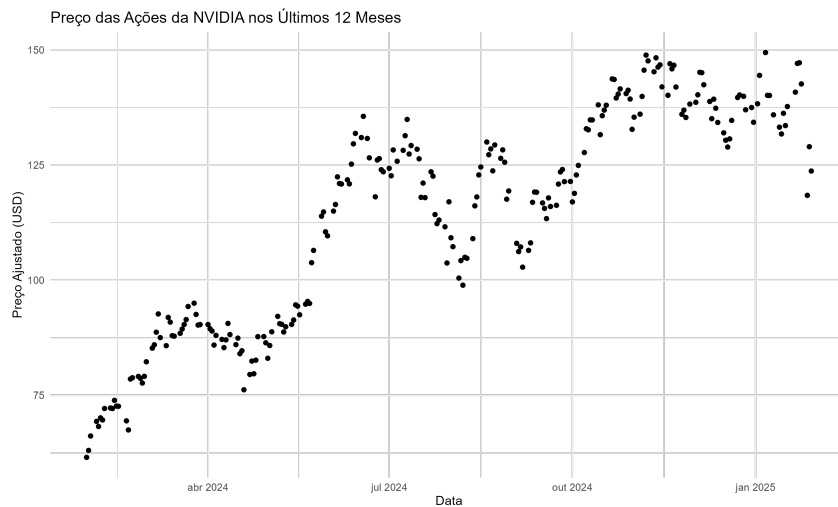
É comum encontrar uma coleção de observações feitas sequencialmente ao longo do tempo, frutos de estudo em áreas diversas, tais como a biologia, meteorologia, geografia, engenharia elétrica e finanças. A essas coleções, atribuímos o termo **séries temporais**. Box *et al.* (2015) define:

Definição 2.1 (Série temporal). ***Série temporal** é uma sequência de observações feitas sequencialmente ao longo do tempo.*

As séries temporais são inicialmente classificadas como discretas ou contínuas. Essa distinção é feita a partir da análise do registro do tempo. Se esse registro foi feito ininterruptamente durante um intervalo de tempo, estamos diante de uma série temporal contínua. Já se for uma coleção de instantes isolados ao longo de um intervalo, a série temporal é dita discreta. Em verdade, uma série temporal discreta pode ser fruto de uma amostragem de intervalos de uma série contínua, por exemplo.

Vejamos um exemplo de série temporal discreta:

Figura 1 - Preço de fechamento ajustado das ações da NVIDIA

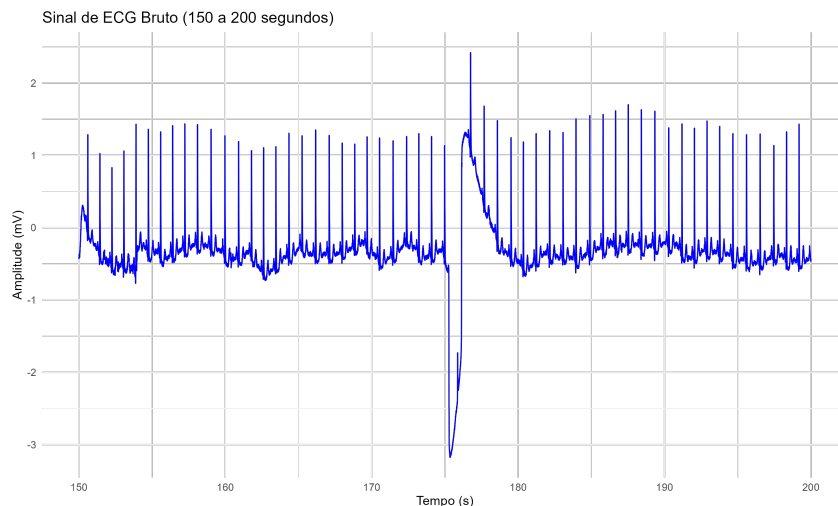


Fonte: Autoria própria. Gráfico desenvolvido em **R**¹ utilizando pacotes do **R**. Dados extraídos do banco **Yahoo Finance**².

A série temporal supracitada representa diariamente os preços em dólar de fechamento das ações da empresa NVIDIA³ do dia 31/01/2024 ao dia 30/01/2025. Perceba que essa é uma série temporal discreta, uma vez que é uma coleção de instantes isolados (dias) ao longo de um intervalo de tempo.

Vejamos um exemplo de série temporal contínua:

Figura 2 - ECG bruto



Fonte: Autoria própria. Gráfico desenvolvido em **R**⁴ utilizando pacotes do **R** e do **Python**. Dados extraídos do banco **MIT-BIH Arrhythmia Database**⁵.

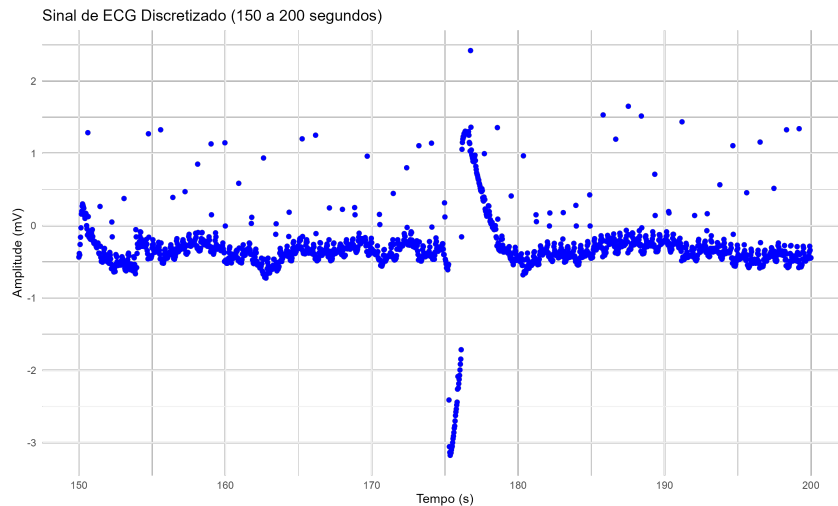
¹Gráfico gerado em **R** (R CORE TEAM, 2023) com o pacote “ggplot2” (WICKHAM, 2016).

²Para acessar no **R** os dados em **Yahoo Finance** (INC., 2024), utilizou-se o pacote “quantmod” (RYAN; ULRICH, 2024).

³A NVIDIA é uma empresa multinacional de tecnologia, conhecida por suas inovações no desenvolvimento de processadores gráficos (GPUs) e unidades de processamento de inteligência artificial.

Perceba que o registro das amplitudes no eletrocardiograma (ECG) foi feito ininterruptamente nesse intervalo de tempo, o que nos leva a classificar essa série temporal como contínua. Perceba ainda que, fazendo uma subamostragem temos o gráfico:

Figura 3 - ECG discretizado



Fonte: Autoria própria. Gráfico desenvolvido em **R**⁴ utilizando pacotes do **R** e do **Python**. Dados extraídos do banco **MIT-BIH Arrhythmia Database**⁵.

Para os fins do trabalho, estaremos interessados apenas nas séries temporais discretas.

Box *et al.* (2015) apontam que um fato intrínseco das séries temporais é que normalmente as observações adjacentes são dependentes. Em relação a esse grau de dependência, quando uma série temporal pode ser prevista com exatidão, ela é chamada de determinística; caso contrário, ela é uma série temporal estocástica. A evolução da posição de um carro, andando em velocidade constante, em uma rodovia ao longo do tempo pode ser considerada uma série temporal determinística. Já os casos ilustrados, como um ECG ou a precificação de um ativo, são exemplos menos previsíveis, e por isso são séries temporais estocásticas. Nesse trabalho, o olhar está voltado para a análise de séries temporais estocásticas.

Tratando de séries temporais estocásticas, Morettin e Tolo (2006) ressaltam que os principais pontos de interesse ao analisar uma série temporal são investigar o que causa as alterações nas séries temporais, prever futuros valores para a série e descrever o comportamento da série através de gráficos, por exemplo, para identificar padrões. Técnicas mais simples, como a identificação de sazonalidade por meio da ob-

⁴Gráfico gerado em **R** (R CORE TEAM, 2023) com o pacote “ggplot2” (WICKHAM, 2016). Para acessar os dados, utilizou-se o pacote “reticulate” (USHEY *et al.* 2024) para rodar a biblioteca “wfdb” (XIE *et al.* 2023) no **Python**.

⁵Os dados foram extraídos do **MIT-BIH Arrhythmia Database** (MOODY; MARK, 2001; GOLDBERGER *et al.* 2000).

servação gráfica, podem ser suficientes para esses objetivos e devem ser a primeira abordagem em qualquer análise. No entanto, este projeto se propõe a explorar uma técnica mais avançada para a análise de séries temporais em que não é suficiente o uso das primeiras táticas.

Muitas vezes, o estudo da série no domínio do tempo não é suficiente para reconhecer padrões, realizar estimativas e cumprir com os objetivos de uma análise. Quando as técnicas do domínio do tempo falham, precisamos fazer um estudo das técnicas do domínio da frequência, também chamado de espectro. A área da análise espectral lida com tais técnicas. Para transformar o domínio de uma função no tempo para o domínio da frequência, utiliza-se a Transformada de Fourier Discreta, permitindo assim uma nova análise.

O estudo das séries temporais não se encerra na análise do domínio do tempo ou da frequência. Gabor (1946), em seu artigo, foi pioneiro na análise de sinais no domínio do tempo e da frequência. Após seu trabalho e a apresentação de seu método, a análise do tempo e da frequência, de forma conjunta, foi amplamente trabalhada gerando outras ferramentas para análise. Uma das ferramentas desenvolvidas foi a Sliding Window Discrete Fourier Transform (SWDFT).

Para melhor compreender a aplicação da técnica SWDFT em séries temporais, é importante conhecer alguns tópicos da Análise Harmônica. Também conhecida como Análise de Fourier, essa trata do estudo das séries e transformadas de Fourier. Os tópicos a seguir tratados nos levarão ao estudo da análise do tempo e frequência de sinais via SWDFT.

2.2 Séries de Fourier

Nesta sessão, estudaremos as séries descritas primeiramente por Jean Baptiste Joseph Fourier, conhecidas como séries de Fourier. Ao descrever uma equação que representasse a propagação do calor, Fourier abriu portas para uma teoria ainda mais ampla, que permite tratar de frequências e descontinuidades. Segundo Cohen (1994), a ideia de que uma função descontínua poderia ser tratada como a soma de funções contínuas se tornou uma das maiores inovações da Matemática e da ciência.

A partir de uma abordagem via equações diferenciais parciais para representar a propagação do calor, Fourier chegou a uma representação de funções periódicas da forma:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{+\infty} \left[a_n \cos\left(\frac{n\pi x}{l}\right) + b_n \sin\left(\frac{n\pi x}{l}\right) \right] \quad (1)$$

em que $a_0, a_1, \dots, a_n, b_1, \dots, b_n$ são chamados coeficientes da série e chamamos ((1)) de **série de Fourier** de $f(x)$.

Os coeficientes dessa série são chamados de **coeficientes de Fourier** de $f(x)$ e são dados pelas *fórmulas de Euler-Fourier*

$$\begin{aligned} a_0 &= \frac{1}{l} \int_{-l}^l f(x) dx \\ a_n &= \frac{1}{l} \int_{-l}^l f(x) \cos\left(\frac{n\pi x}{l}\right) dx \quad n = 1, 2, \dots \\ b_n &= \frac{1}{l} \int_{-l}^l f(x) \operatorname{sen}\left(\frac{n\pi x}{l}\right) dx \quad n = 1, 2, \dots \end{aligned} \quad (2)$$

Assim, sejam $l > 0$ e $f : [-l, l] \rightarrow \mathbb{R}$ uma função integrável. A série de Fourier de f é:

$$S[f] = \frac{a_0}{2} + \sum_{n=1}^{+\infty} \left[a_n \cos\left(\frac{n\pi x}{l}\right) + b_n \operatorname{sen}\left(\frac{n\pi x}{l}\right) \right] \quad (3)$$

Note que as funções $\cos\left(\frac{n\pi x}{l}\right)$ e $\operatorname{sen}\left(\frac{n\pi x}{l}\right)$ são periódicas com período fundamental $T = \frac{2l}{n}$. Em particular, $2l$ é um período comum a todas essas funções. Então se a série convergir para cada $x \in \mathbb{R}$, a função convergirá para uma função periódica de período $2l$.

Exemplo 1. Para fins de ilustração, vamos calcular a série de Fourier da função:

$$f(x) = \begin{cases} -x & -l \leq x < l, \\ l & l = x. \end{cases}$$

Vamos calcular os coeficientes:

$$\begin{aligned} a_n &= \frac{1}{l} \int_{-l}^l f(x) \cos\left(\frac{n\pi x}{l}\right) dx \\ b_n &= \frac{1}{l} \int_{-l}^l f(x) \operatorname{sen}\left(\frac{n\pi x}{l}\right) dx \\ a_0 &= \frac{1}{l} \int_{-l}^l f(x) dx \end{aligned}$$

Começando por a_n .

$$\begin{aligned} a_n &= \frac{1}{l} \int_{-l}^l f(x) \cos\left(\frac{n\pi x}{l}\right) dx \\ a_n &= \frac{1}{l} \int_{-l}^l -x \cos\left(\frac{n\pi x}{l}\right) dx \\ a_n &= -\frac{1}{l} \int_{-l}^l x \cos\left(\frac{n\pi x}{l}\right) dx \end{aligned}$$

Note que $x \cos\left(\frac{n\pi x}{l}\right)$ é uma função ímpar. Assim, a área sob a curva no intervalo

$[-l, l]$ é nula, pois as regiões abaixo e acima do eixo x possuem áreas de mesma magnitude, mas sinais opostos. Segue que

$$a_n = -\frac{1}{l}(0)$$

$$a_n = 0$$

Façamos o coeficiente b_n .

$$b_n = \frac{1}{l} \left[\int_{-l}^l -x \operatorname{sen} \left(\frac{n\pi x}{l} \right) dx \right]$$

Integrando por partes, segue que

$$b_n = \frac{1}{l} \left[-\frac{l^2}{n\pi} \cos(n\pi) - \frac{l^2}{n\pi} \cos(n\pi) \right]$$

$$b_n = \frac{2l(-1)^n}{n\pi}$$

Por fim, o coeficiente a_0

$$a_0 = \frac{1}{l} \left[\int_{-l}^l -x dx \right]$$

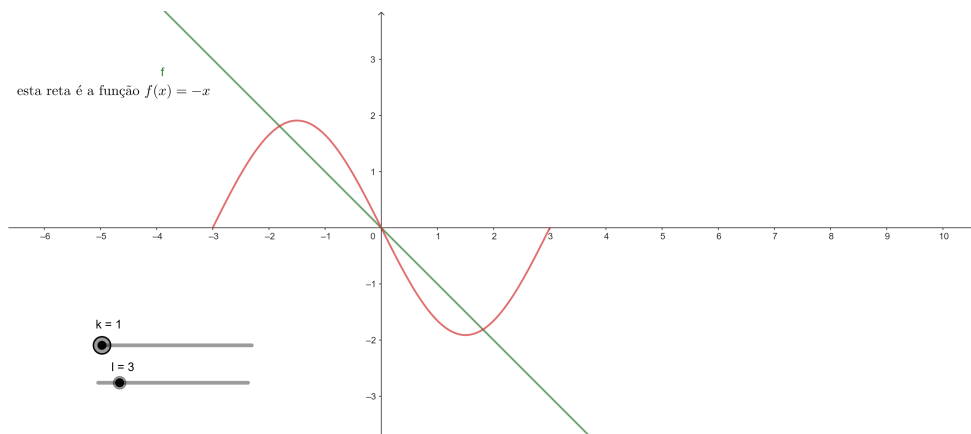
$$a_0 = \frac{1}{l} \left[-\frac{x^2}{2} dx \right]_{-l}^l$$

$$a_0 = 0$$

De ((3)), segue que:

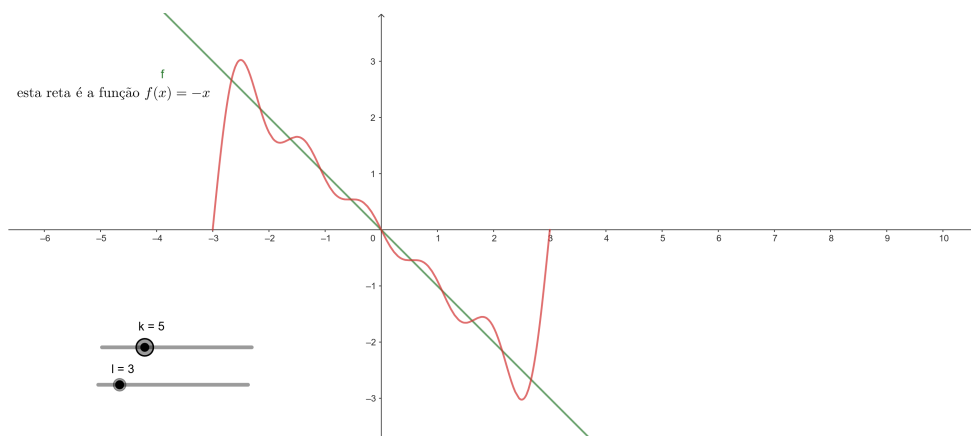
$$S[f] = \frac{-2l}{n\pi} \sum_{n=1}^{+\infty} \left((-1)^n \operatorname{sen} \left(\frac{n\pi x}{l} \right) \right)$$

Perceba no exemplo que $-l \leq x \leq l$, ou seja, a função está restrita a esse intervalo. A seguir estão as representações gráficas da série no geogebra, em que o controle deslizante l representa o tamanho do intervalo, mas foi mantido fixo igual a três para fins de ilustração.

Figura 4 - Série de Fourier quando $n=1$ 

Fonte: Autoria própria. Gráfico desenvolvido no geogebra

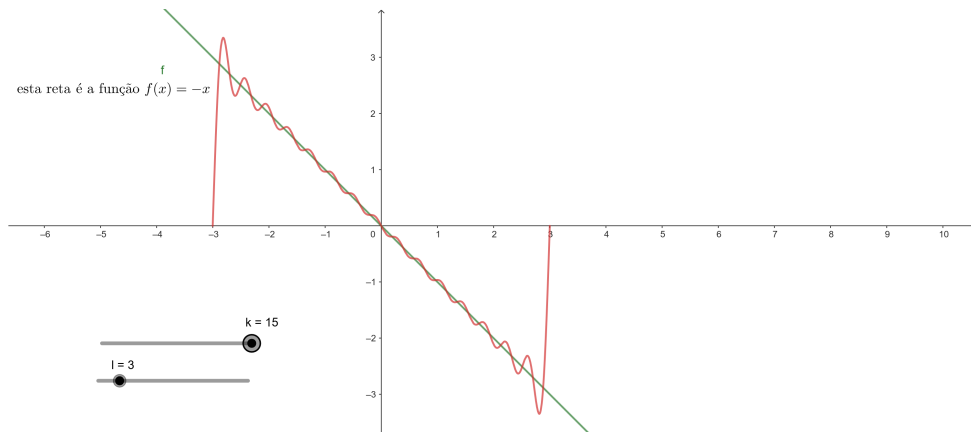
Assim, temos a série de Fourier da função quando $n = 1$. A seguir, veja como a série de Fourier muda para $n = 5$:

Figura 5 - Série de Fourier quando $n = 5$ 

Fonte: Autoria própria. Gráfico desenvolvido no geogebra

Perceba a aproximação das senoides à função, como esperávamos. Por fim, para obter uma nova confirmação visual, façamos $n = 15$

Figura 6 - Série de Fourier quando n=15



Fonte: Autoria própria. Gráfico desenvolvido no geogebra

No exemplo visual construído, a soma está escrita $S[f] = \sum_{n=1}^k \left(\frac{2l(-1)^n}{n\pi} \operatorname{sen} \left(\frac{n\pi x}{l} \right) \right)$ porque o infinito não pode ser representado computacionalmente no software geogebra.

Nesta subseção, representamos funções periódicas como uma soma de senos e cossenos. Essa representação permite representar, em um intervalo, funções com finitas descontinuidades como uma série de senos e cossenos. Embora a representação fornecida em ((3)) seja suficiente para muitos casos, convém uma representação mais sucinta. A seguir, será apresentada uma forma mais breve e usual das séries de Fourier.

2.2.1 A série de Fourier complexa

Segundo Osgood (2007) o trabalho algébrico, para trabalhar séries trigonométricas como a fornecida em ((3)), pode ser simplificado significativamente usando exponenciais complexas para representar o seno e o cosseno. Nesse sentido, vamos utilizar a fórmula de Euler para reescrever a série de Fourier.

Perceba que podemos reescrever a série de Fourier dada em ((3)) utilizando a fórmula de Euler:

$$e^{iy} = \cos(y) + i \operatorname{sen}(y).$$

Fazendo $y = \frac{n\pi x}{l}$, somando e subtraindo os termos $e^{i \frac{n\pi x}{l}}$ e $e^{-i \frac{n\pi x}{l}}$, concluímos que

$$\begin{aligned} \cos \left(\frac{n\pi x}{l} \right) &= \frac{e^{i \frac{n\pi x}{l}} + e^{-i \frac{n\pi x}{l}}}{2} \\ \operatorname{sen} \left(\frac{n\pi x}{l} \right) &= \frac{e^{i \frac{n\pi x}{l}} - e^{-i \frac{n\pi x}{l}}}{2i}. \end{aligned}$$

Reescrevemos então a série de Fourier da seguinte forma:

$$\begin{aligned}
 \frac{a_0}{2} + \sum_{n=1}^{+\infty} \left[a_n \cos\left(\frac{n\pi x}{l}\right) + b_n \sin\left(\frac{n\pi x}{l}\right) \right] \\
 = \frac{a_0}{2} + \sum_{n=1}^{+\infty} \left[\left(\frac{a_n - ib_n}{2} \right) e^{i\frac{n\pi x}{l}} + \left(\frac{a_n + ib_n}{2} \right) e^{-i\frac{n\pi x}{l}} \right] \\
 = \sum_{n=-\infty}^{+\infty} \hat{f}(n) e^{i\frac{n\pi x}{l}}
 \end{aligned}$$

em que a série de $-\infty$ à $+\infty$ deve ser interpretada como o limite

$$\lim_{N \rightarrow +\infty} \sum_{n=-N}^N \hat{f}(n) e^{i\frac{n\pi x}{l}}.$$

Temos então a série de Fourier complexa da função f , representada como:

$$\sum_{n=-\infty}^{+\infty} \hat{f}(n) e^{i\frac{n\pi x}{l}} \quad (4)$$

e os coeficientes $\hat{f}(n)$ serão os chamados coeficientes de Fourier complexos e são dados por:

$$\begin{aligned}
 \hat{f}(0) &= \frac{a_0}{2} \\
 \hat{f}(n) &= \frac{a_n - ib_n}{2} \\
 \hat{f}(-n) &= \frac{a_n + ib_n}{2}.
 \end{aligned} \quad (5)$$

Das equações em ((2)) e ((4)), segue que

$$\hat{f}(n) = \frac{1}{2l} \int_{-l}^l f(x) e^{-i\frac{n\pi x}{l}} dx \quad (6)$$

Essa forma mais concisa é frequentemente encontrada nas abordagens de séries de Fourier.

Neste percurso, introduzimos brevemente as séries de Fourier. Uma vez que o ponto deste projeto não é exclusivamente esse tópico, resultados e teoremas a respeito da convergência de tais séries foram omitidos. O objetivo ao abordar esse tópico é ilustrar a forma de representar funções de período $2l$ como uma série de funções trigonométricas.

Embora poderosa por si só, a série de Fourier tem suas limitações, em especial, o fato de que apenas conseguimos decompor funções periódicas em somas de senos e cossenos. Para fugir dessa limitação, é necessário introduzir uma nova ferramenta.

3 TRANSFORMADA DE FOURIER

As séries de Fourier podem auxiliar na interpretação de importantes fenômenos periódicos. Contudo, muitas aplicações que poderiam se beneficiar de uma técnica como a apresentada, envolvem funções que não são periódicas. Para tratar de tais funções, seguindo os passos de Fernández e Cavalcante (2009) motivaremos a definição da transformada de Fourier ampliando o período $[-L, L]$ como um limite da série de Fourier.

$$f(x) = \lim_{L \rightarrow +\infty} f_L(x)$$

Em ((6)) faça $\xi_n = \frac{n}{2L}$. Definindo g como

$$g(\xi) = \int_{-L}^L f(x) e^{-2\pi i x \xi} dx \quad (7)$$

Podemos reescrever a série em (3), da seguinte forma:

$$f_L(x) \approx \sum_{n=-\infty}^{+\infty} e^{2\pi i \xi_n x} g(\xi_n) \frac{1}{2L} \quad (8)$$

Daí, como $\xi_{n+1} - \xi_n = \frac{1}{2L}$, a soma em ((8)) pode ser interpretada como uma soma de Riemann em \mathbb{R} da função $g(\xi_n) e^{2\pi i x \xi}$. Para isso, basta considerar $g(\xi_n) e^{2\pi i x \xi}$ como a altura do retângulo e $\Delta \xi_n$ como a largura da base. Passando o limite quando $L \rightarrow +\infty$, implica diretamente $\Delta \xi_n \rightarrow 0$ e temos as expressões:

$$g(\xi) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i x \xi} dx \quad (9)$$

e

$$\lim_{L \rightarrow \infty} f_L(x) = f(x) \approx \int_{-\infty}^{\infty} g(\xi) e^{-2\pi i x \xi} d\xi. \quad (10)$$

O que nos leva a definição da transformada de fourier contínua de uma função f :

$$\hat{f}(\xi) = \int_{-\infty}^{+\infty} f(x) e^{-2\pi i \xi x} dx \quad (11)$$

Exemplo 2. Para ilustrar a Transformada de Fourier Contínua, seja a função Gaussiana:

$$f(x) = e^{-\pi x^2}$$

Sabendo que $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$, segue que

$$\int_{-\infty}^{+\infty} e^{-\pi x^2} dx = 1.$$

Fazendo a transformada de Fourier contínua, temos

$$\hat{f}(\xi) = \int_{-\infty}^{+\infty} e^{-\pi x^2} e^{-2\pi i \xi x} dx.$$

Utilizando o Teorema de Leibniz para diferenciação sob a integral, podemos derivar com respeito a ξ . Assim, temos

$$\begin{aligned} \frac{d}{d\xi} \hat{f}(\xi) &= \int_{-\infty}^{+\infty} e^{-\pi x^2} \frac{\partial}{\partial \xi} e^{-2\pi i \xi x} dx \\ \frac{d}{d\xi} \hat{f}(\xi) &= \int_{-\infty}^{+\infty} e^{-\pi x^2} (-2\pi i x) e^{-2\pi i \xi x} dx \end{aligned}$$

Agora, integrando por partes, fazemos $dv = -2\pi i x e^{-\pi x^2} dx$ e $u = e^{-2\pi i \xi x}$. Assim, tem-se que $du = (-2\pi i \xi) e^{-2\pi i \xi x}$, e observando que $\frac{d}{dx} e^{-\pi x^2} = -2\pi x e^{-\pi x^2}$, segue que $v = i e^{-\pi x^2}$. Portanto

$$\frac{d}{d\xi} \hat{f}(\xi) = \left(e^{-2\pi i \xi x} (i e^{-\pi x^2}) \right)_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} (i e^{-\pi x^2}) (-2\pi i \xi e^{-2\pi i \xi x}) dx$$

No primeiro termo da diferença, como $e^{-\pi x^2}$ tende a 0 muito rápido quando tem-se $x \rightarrow \pm\infty$, o produto converge para 0. Portanto,

$$\begin{aligned} \frac{d}{d\xi} \hat{f}(\xi) &= - \int_{-\infty}^{+\infty} i e^{-\pi x^2} (-2\pi i \xi e^{-2\pi i \xi x}) dx \\ \frac{d}{d\xi} \hat{f}(\xi) &= - \int_{-\infty}^{+\infty} i^2 e^{-\pi x^2} (-2\pi \xi e^{-2\pi i \xi x}) dx \\ \frac{d}{d\xi} \hat{f}(\xi) &= - \int_{-\infty}^{+\infty} e^{-\pi x^2} (2\pi \xi e^{-2\pi i \xi x}) dx \\ \frac{d}{d\xi} \hat{f}(\xi) &= -2\pi \xi \int_{-\infty}^{+\infty} e^{-\pi x^2} (e^{-2\pi i \xi x}) dx \end{aligned}$$

Lembre-se que $\hat{f}(\xi) = \int_{-\infty}^{+\infty} e^{-\pi x^2} e^{-2\pi i \xi x} dx$. Logo

$$\begin{aligned}\frac{d}{d\xi} \hat{f}(\xi) &= -2\pi i \xi \hat{f}(\xi) \\ \frac{d\hat{f}(\xi)}{\hat{f}(\xi)} &= -2\pi i d\xi\end{aligned}$$

Resolvendo essa equação diferencial de variáveis separadas, obtemos

$$\hat{f}(\xi) = C e^{-\pi \xi^2}. \quad (12)$$

Tomando $\xi = 0$, temos a condição inicial

$$\begin{aligned}\hat{f}(0) &= \int_{-\infty}^{+\infty} e^{-\pi x^2} e^{-2\pi i 0 x} dx \\ \hat{f}(0) &= \int_{-\infty}^{+\infty} e^{-\pi x^2} = 1\end{aligned}$$

e, portanto,

$$1 = \hat{f}(0) = C e^{-\pi 0^2} = C.$$

Assim, de ((12)), segue que $\hat{f}(\xi) = C e^{-\pi \xi^2} = e^{-\pi \xi^2}$.

Assim, calculamos a transformada de Fourier de uma função não periódica no domínio contínuo. Por envolver o conceito de infinito, as técnicas que vimos até aqui, como a série de Fourier e a transformada contínua de Fourier, geralmente são resolvidas analiticamente. O interessante do exemplo acima é que podemos concluir que a transformada de Fourier da função Gaussiana $e^{-\pi x^2}$ é ela mesma.

A ferramenta que acabamos de construir nos permite representar um número ainda maior de funções. Em especial, essa técnica desempenha papel importante na análise espectral uma vez que, segundo Osgood (2007), a transformada de Fourier analisa um sinal em suas componentes de frequência. Apesar do avanço, o ferramental até aqui apresentado abordou apenas funções com domínio contínuo.

Como Smith (1999) argumenta, os computadores digitais trabalham apenas com informações finitas e discretas. Sendo assim, para fazer uma análise computacional de séries temporais, será preciso estender a transformada de Fourier para o domínio discreto.

3.1 Transformada de Fourier Discreta

Conforme nos aproximamos de aplicações reais, a demanda por um método que trate de sinais com domínio discreto aumenta, e é nesse contexto que se destaca a

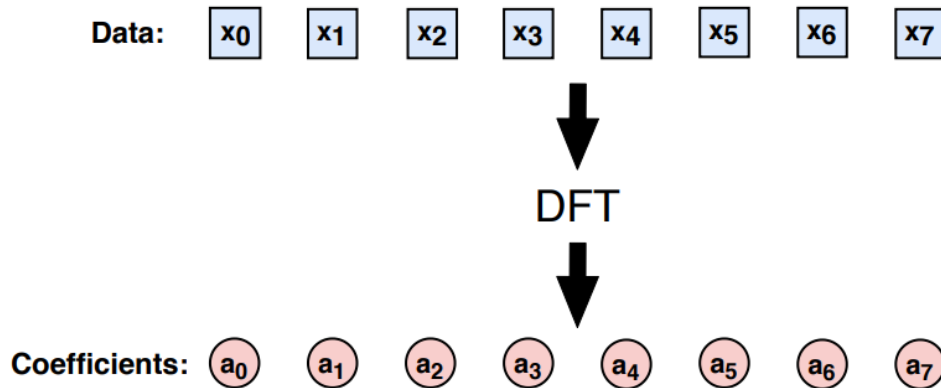
relevância da técnica de Transformada de Fourier Discreta (DFT). Seja um sinal com n valores reais. Segundo Richardson (2019), a DFT associa a cada um desses valores um coeficiente complexo que representa a correlação do sinal com funções seno e cosseno de frequências particulares. Assim, definimos:

Definição 3.1 (Transformada de Fourier Discreta). *Seja $\mathbf{f} = (f(0), f(1), \dots, f(N-1))$ uma n -upla de tamanho N . A Transformada de Fourier Discreta de \mathbf{f} é a n -upla $\mathbf{F} = (F(0), F(1), \dots, F(N-1))$ definida por*

$$F[m] = \sum_{n=0}^{N-1} f[n] e^{-i \frac{2\pi mn}{N}}, \quad m = 0, 1, \dots, N-1. \quad (13)$$

Isso é, dado uma entrada $f(0)$ de um sinal \mathbf{f} , será associado a essa entrada um coeficiente complexo $F(0)$ que indicará a medida de correlação entre o sinal \mathbf{f} e a frequência particular de funções seno e cosseno associadas. Nesse sentido, o módulo do valor complexo F é diretamente proporcional a esse grau de correlação.

Figura 7 - Representação visual da DFT



Fonte: Richardson (2019)

Exemplo 3. *Seja a função*

$$f(t) = g(t) + h(t) = \text{sen}(2\pi 2t) + 0.5 \text{sen}(2\pi 7t)$$

a soma das funções

- $g(t) = \text{sen}(2\pi 2t)$
- $h(t) = 0.5 \text{sen}(2\pi 7t)$.

Note que a função $g(t)$ tem frequência 2 e a função $h(t)$ tem frequência 7. Naturalmente, podemos tomar 32 amostras da função $f(t)$ igualmente separadas no domínio e registrá-las em um vetor \mathbf{u} .

$$\mathbf{u}(t) = (0.000, 0.873, 0.898, 0.508, 0.646, 1.202, 1.169, 0.285, -0.500, -0.480, -0.245, -0.646, \\ -1.354, -1.340, -0.516, 0.108, 0.000, -0.108, 0.516, 1.340, 1.354, 0.646, 0.245, 0.480, \\ 0.500, -0.285, -1.169, -1.202, -0.646, -0.508, -0.898, -0.873)$$

Vamos visualizar essa função e seus pontos amostrados:

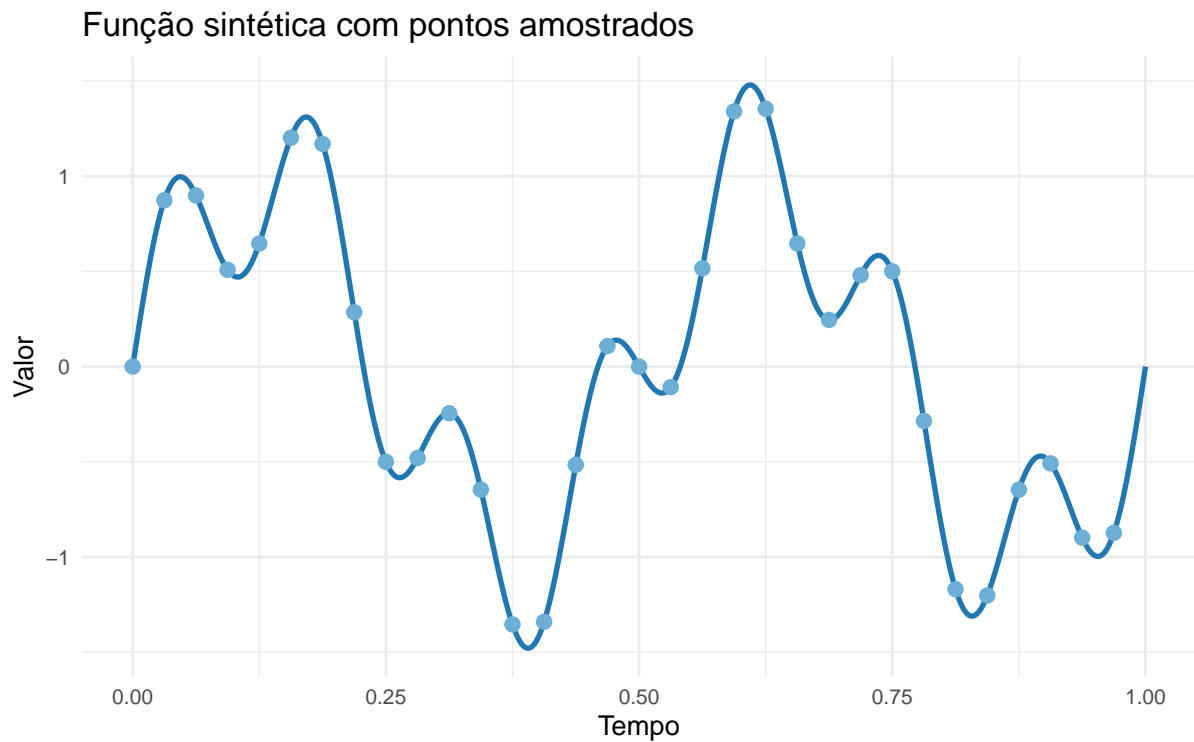


Gráfico 1 — Gráfico da função sintética e os pontos amostrados

Agora, utilizando a fórmula da DFT (13), podemos calcular os coeficientes complexos associados a cada entrada. Vejamos os 10 primeiros, por exemplo:

Quadro 1 — Primeiros 10 pontos e seus respectivos coeficientes

ponto	f_amostrado	X_complexo
0	0.000	0+0i
1	0.873	0+0i
2	0.898	0-16i
3	0.508	0+0i
4	0.646	0+0i
5	1.202	0+0i
6	1.169	0+0i
7	0.285	0-8i
8	-0.500	0+0i
9	-0.480	0+0i

Podemos observar que no intervalo $[0, 9]$ o ponto $t = 2$ e o ponto $t = 7$ são os únicos que em módulo são diferentes de zero. Daí, podemos inferir que as frequências 2 e 7 tem maior impacto na nossa função exemplo, como esperávamos. Veja no gráfico convertido do domínio do tempo para o da frequência:

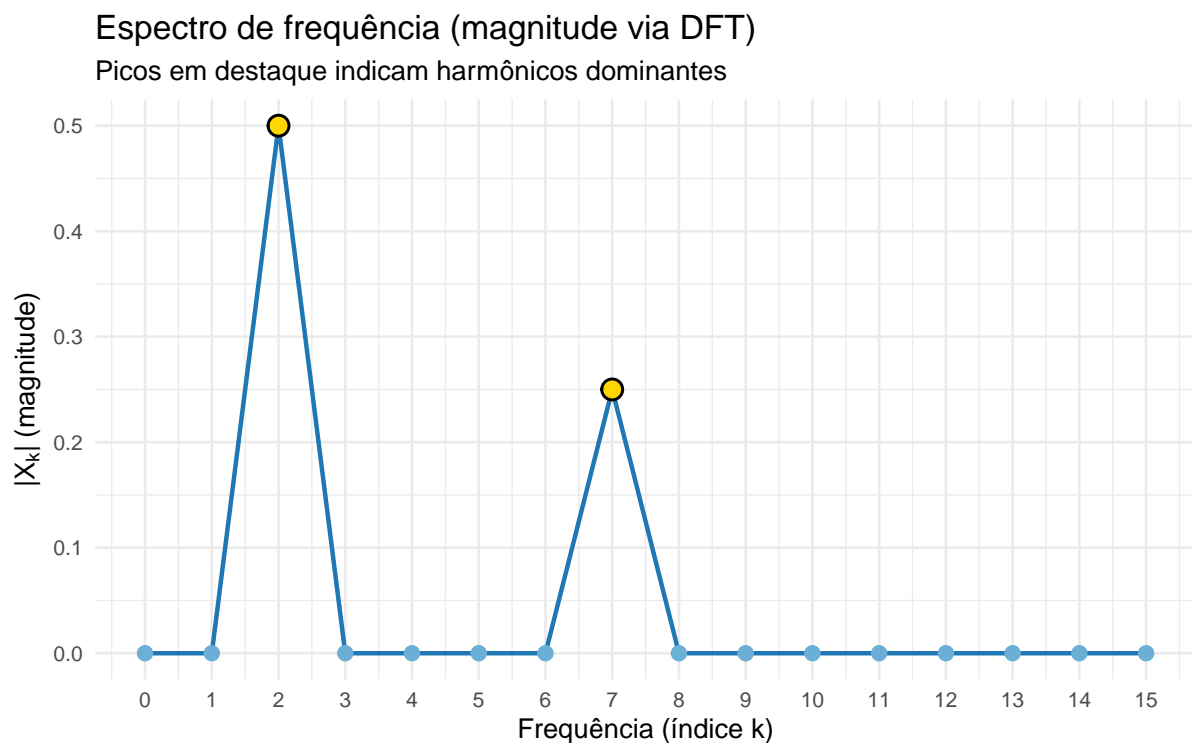


Gráfico 2 — Gráfico no domínio da frequência

Este exemplo foi construído no RStudio com parâmetros controlados e por isso os resultados são tão claros.

Quando estamos interessados em estudar dados reais, a transformada de Fourier discreta confirma sua relevância, afinal, apesar da consideração da natureza contínua dos fenômenos naturais, os aparelhos de registro dessas informações apenas capturam pontos amostrais (isso é, discretizam o fenômeno). Embora funcional, a transformada de Fourier Discreta não é muito eficiente.

(LI) Introduction to Algorithms Mostra que existem algumas notações para descrever tempos assintóticos de execução de algoritmos, entre elas a **notação grande O** (*big O notation* ou *O-notation*). Essa notação descreve um limite superior assintótico de execução, isso é, no pior cenário o tempo que o algoritmo demora para ser completo.

Definição 3.2. *Seja uma função $f(n)$ assintoticamente não negativa, definida em \mathbb{N} , que pode ser estendida a \mathbb{R} . Denotamos como $O(f(n))$ (lê-se "grande-O de f de n " ou apenas "O de f de n ") um conjunto de funções $g(n)$, também assintoticamente não negativas, para as quais existem constantes positivas c e n_0 tais que:*

$$O(f(n)) = \{g(n) : \forall n \geq n_0, \quad 0 \leq g(n) \leq c f(n)\} \quad (14)$$

Isto é, uma função $g(n)$ pertence ao conjunto $O(f(n))$ se existir uma constante positiva c tal que $g(n) \leq c f(n)$ para todo n suficientemente grande.

Para a DFT, perceba que para uma amostra de tamanho N , é preciso fazer N multiplicações complexas, isso é $N \times N$, dizemos então que a DFT tem custo computacional $O(N^2)$.

Para melhorar a eficiência do método, alguns algoritmos foram desenvolvidos para aprimorar esse custo computacional. Apesar de se tratar de um conjunto de algoritmos, chamamos neste trabalho de Transformada de Fourier Rápida (Fast Fourier Transform - FFT) a FFT de Cooley-Tukey, descoberta pioneiramente por Gauss em 1805, que não a publicou em vida, **Gauss and the History of the Fast Fourier Transform** mas popularizada por J.W. Cooley e J.W. Tukey **An Algorithm for the Machine Calculation of Complex Fourier Series - Mathematics of Computation Vol. 19, No. 90 (Apr., 1965)**. A FFT, por sua vez tem custo computacional $O(N \log_2(N))$. **aqui devo desenvolver a FFT para explicar o porquê desse custo computacional? Posso apenas justificar que tem um pacote no R que faz isso? Posso apenas indicar bibliografias em que podem ser encontradas essas informações?**

Por via das dúvidas, esse parágrafo é apenas uma coleção de lugares onde encontrar esse tipo de informação: William H. Press - Numerical Recipes página 634 do PDF. Oppenheim - discrete-time signal processing página 655 do PDF. Cohen - time-frequency analysis página 109 do PDF para STFT. LI introduction to algorithms página 910 do PDF

quem sabe aqui posso colocar um gráfico ilustrativo apenas para comparação do

custo computacional $O(N^2)$ versus $O(N\log(N))$). Esse gráfico costuma ser bem representativo.

Agora que temos meios de analisar computacionalmente sinais com domínio discreto, é interessante pensar em como reduzir o poder computacional. Isso é, para grandes quantidades de dados, esse processo requer um custo computacional elevado. Sendo assim, dado um conjunto de amostras grande, seria interessante dividi-lo para calcular a DFT em subconjuntos. Para suprir essa necessidade, podemos utilizar a técnica Sliding Window Discrete Fourier Transform, traduzido para o português, Transformada de Fourier Discreta em Janelas Deslizantes (SWDFT).

3.2 SWDFT

Segundo Richardson (2019), a técnica SWDFT, utilizada como uma transformação no tempo e frequência, é um bom ponto de partida pra transformações - no mesmo domínio - mais complexas, uma vez que é uma extensão da DFT. A SWDFT acrescenta à DFT o recurso de janela deslizante que permite fazer a DFT em P janelas de tamanho n . Assim, Richardson (2019) define:

Definição 3.3 (Transformada de Fourier em Janelas Deslizantes). A SWDFT é dada por

$$a_{k,p} = \frac{1}{\sqrt{n}} \sum_{j=0}^{n-1} x_{p-n+1+j} \omega_n^{-jk};$$

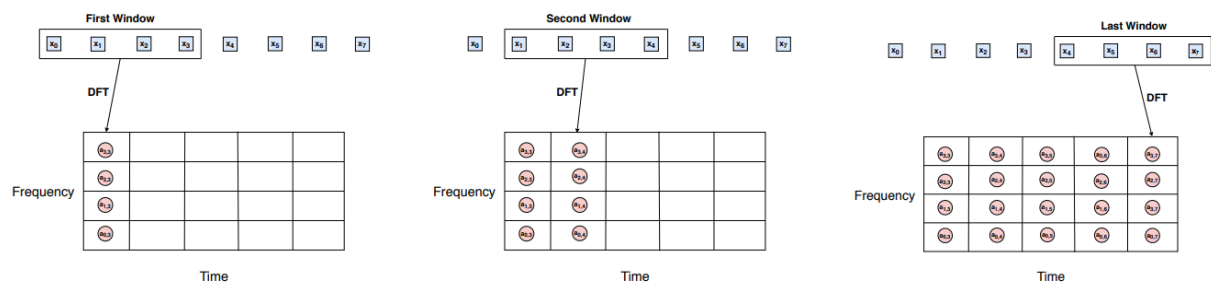
$$k = 0, 1, \dots, n-1;$$

$$p = n-1, n, \dots, N-1.$$
(15)

Em que ((15)) retorna uma matriz $n \times P$ de coeficientes da SWDFT, em que $P = N - n + 1$ é o número de posições da janela. Nesse sentido, a dimensão n representa a frequência e a dimensão P , que corresponde à janela, geralmente representa o tempo.

Segue uma ilustração que representa como a técnica funciona em um conjunto de dados com tamanho $N = 8$ e uma janela deslizante de tamanho $n = 4$:

Figura 8 - Exemplo visual da SWDFT



Fonte: Richardson (2019)

No cenário atual de processamento de sinais e aprendizado de máquina, pesquisadores têm explorado abordagens inovadoras baseadas em SWDFT para diversas aplicações. Sai *et al.* (2024) demonstraram o potencial da SWDFT combinada com redes neurais convolucionais (CNN) para diagnóstico médico, especificamente na detecção de Transtorno do Espectro Autista (TEA) em crianças.

Aplicações biomédicas têm se beneficiado da técnica. Shen *et al.* (2024) integraram SWDFT em tempo real à CNN da GoogleNet para detecção de crises epiléticas em eletroencefalografias (EEG). O tempo de processamento do método foi de 0,02 segundos para cada episódio de EEG de 2 segundos com um atraso médio de 9,85 segundos no início de cada crise. Já Soontornnapar e Ploysuwan (2025), utilizando a SWDFT aliada à magnitude de quaternios de sinais, introduzem uma nova abordagem para a detecção de quedas a fim de garantir intervenções médicas oportunas, especialmente para idosos.

Phungtua-eng e Yamamoto (2024) superaram a limitação dos métodos tradicionais de decomposição sazonal-tendência, que exigem duração fixa da sazonalidade, propondo um método adaptativo integrado à estimativa do período sazonal. Utilizando a SWDFT, reduziram a complexidade computacional para $O(N)$ (vs. $O(N \log N)$ dos métodos atuais), viabilizando a decomposição precisa em fluxos de dados com sazonalidade variável.

Nas últimas subseções foram introduzidas, de maneira informal, as ideias por trás das técnicas do Trabalho aqui proposto. Nas seções a seguir, estão descritos os métodos para sistematizar de maneira rigorosa, mas ainda acessível, essas técnicas, os resultados esperados e o cronograma de execução proposto para o Trabalho.

4 APLICAÇÃO SWFFT

Para melhor ilustrar a técnica de SWFFT foi selecionado, em um banco de dados do Instituto Nacional de Meteorologia (INMET), dados a respeito da radiação solar em quatro estados e o Distrito Federal de diferentes regiões do Brasil, a fim de buscar interpretações dos harmônicos que compõem essa variação. O INMET **achar algo para citar aqui a respeito do INMET** é responsável por prover dados para a Organização Meteorológica Mundial (OMM) a respeito dos dados meteorológicos da América do Sul. No Brasil, a coleta de seus dados é realizada a partir de estações convencionais, em que os dados são coletados manualmente e em estações automáticas, em que os dados são registrados por um datalogger e transmitidos para um supercomputador.

Em seu portal online, o INMET disponibiliza dados das estações automáticas pelo menos desde o ano 2000 a meados de 2025. Esses dados, fornecidos em arquivo .csv, contêm informações hora a hora sobre o vento, a umidade, a temperatura, a

radiação, a precipitação e a pressão. Dentre esses, foi escolhido os dados a respeito da radiação solar em virtude de suas poucas ramificações em registro e seu potencial como energia renovável.

Após ler e separar os dados, utilizando o software RStudio, reunindo mais de um milhão de linhas de informação, foi iniciado o processo de tratamento dos dados a fim de eliminar ruídos, uma vez que esses poderiam interferir na composição de harmônicos. Em momento inicial, buscou-se entender a característica dos dados a serem trabalhados. Separando por estados, foi possível perceber que apesar de apresentarem dados fisicamente inconsistentes, ainda era possível perceber, a partir da mediana calculada a cada hora, que havia um padrão de radiação solar durante os períodos diurnos.

Compreendendo isso, buscou-se identificar o porquê de valores inconsistentes de radiação máxima e mediana, como medianas diferentes de zero na madrugada, por exemplo. Percebeu-se que a quantidade de dados faltantes, como é de se esperar, em períodos noturnos tornava a mediana sensível a dados registrados como erro do sensor nesses momentos. Ainda, os valores de radiação máxima inconsistentes representam um ponto com tal valor. Para fazer uma análise significativa, foi necessário identificar os limites superiores e quantificar os outliers. Nesse sentido, usou-se a fórmula $Q_3 + 1.5IQR = LS$ para identificar os limites superiores teóricos. **aqui podem entrar os boxplots indicando como eram os dados antes do tratamento**

O cálculo do limite superior teórico não registra com grande precisão de fato os limites superiores, uma vez que os dados de radiação podem ser altamente assimétricos. Contudo, esses limites podem servir como um ponto de partida para quantificar os outliers e compreender o quanto esses registros acima do limite representam do total. **tabela de outliers por estado. Mencionar que o gráfico foi criado com RStudio.**

Quadro 2 — Tabela da proporção de outliers por estado

estado	total_observacoes	total_outliers	percentual_outliers
AM	216072	2325	1.0760302
BA	215976	1499	0.6940586
DF	216120	2994	1.3853415
RJ	216120	4222	1.9535443
RS	212808	322	0.1513101
Total	1077096	11362	1.0548735

Aproveitar a deixa do total de observações para falar que os dados disponíveis começam em datas diferentes e para não perder informações, optou-se por trabalhar com os dados assim mesmo

Ainda buscando quantificar dados que podem ser enquadrados como ruídos, fez-se

o estudo do percentual de dados faltantes, expressados pelo RStudio como NA's. Foi possível perceber que durante o período de sol, não havia uma quantidade expressiva de dados faltantes, considerando o volume de dados e o extenso período de coleta desses. Para o período noturno foi registrado um número maior de dados faltantes, como o esperado. **tabela de NA's por estado. Mencionar que o gráfico foi criado com RStudio**

Tabela 1 - Tabela da proporção de NA's por estado

estado	total_observacoes	total_na	percentual_na
AM	216072	94978	43.95664
BA	215976	109979	50.92186
DF	216120	107287	49.64233
RJ	216120	90709	41.97159
RS	212808	101897	47.88213
Total	1077096	504850	46.87140

explicar que isso é antes da limpeza. Falar da quantidade de NA's de noite. Quem sabe modificar a tabela para incluir os períodos noturnos nesses NA's

Compreendendo os principais tipos de ruídos, foi possível elaborar uma estratégia de limpeza e imputação de dados. Para o período noturno, considerando a natureza física do fenômeno de irradiação solar, a todos dados foi atribuído o valor 0. Para o período diurno, primeiramente foram eliminados os dados de irradiação acima de $7200 \frac{kJ}{m^2}h$, uma vez que esse é o limite de captação do sensor, levando a compreender que medições acima desse valor devem ser ao menos imprecisas. Em seguida, foi realizada uma imputação dos dados utilizando o método de bootstrap não paramétrico. Para a imputação, foi utilizada a mediana para períodos solares e o primeiro quartil para períodos de nascer e por do sol.

Tabela 2 - Tabela de proporção de outliers por estado pós reconstrução

estado	total_observacoes	total_outliers	percentual_outliers
AM	216072	80	0.0370247
BA	215976	1463	0.6773901
DF	216120	1447	0.6695354
RJ	216120	775	0.3585971
RS	212808	37	0.0173866
Total	1077096	3802	0.3529862

aqui podemos colocar visualizações do banco de dados limpo, como boxplots por exemplo.

A escolha de utilizar valores iguais a zero para os dados noturnos, além da justificativa física, é feita para que não haja descontinuidade temporal, pois isso prejudicaria a aplicação da FFT. Além disso, o valor nulo não irá interferir nas frequências observadas.

A aplicação da FFT em janelas deslizantes foi feita em todo o banco de dados. Em coerência com os pressupostos para a aplicação da DFT, utilizou-se janelas de um ano, considerando uma periodicidade anual que se espera em razão do período de translação da Terra ao redor do Sol. Para as janelas, escolheu-se a Hanning por reduzir o vazamento espectral.

5 RESULTADOS OBTIDOS

Para a construção de um bom material, voltado para o curso de Licenciatura em Matemática, que desenvolva de maneira sistemática um percurso para a compreensão da técnica SWDFT, o referencial teórico será revisado e ampliado. O intuito principal do trabalho é desenvolver um TCC que possa facilitar a abordagem da técnica para os graduandos, podendo ser trabalhada em uma Disciplina do Núcleo específico, como Estatística Básica. Ainda, o desenvolvimento dos tópicos da Análise de Fourier pode despertar a curiosidade dos licenciandos, aumentando assim a chance de uma matéria optativa a respeito desse conteúdo.

Em síntese, é esperado que os objetivos específicos e gerais supracitados sejam cumpridos, disponibilizando, enfim, um material a nível de graduação que introduza a técnica SWDFT, apresente exemplos, aplicações e desperte a curiosidade dos licenciandos em áreas da Matemática abordadas em disciplinas de pós-graduação.

REFERÊNCIAS

BARBOSA, Júlia Rodrigues; GERMANO, Renato. Modelagem Matemática da Inadimplência no Serasa: Uma Análise Estatística e de Série Temporal. *In*: RIBEIRO, Felipe Vítório; PEREIRA, Walmir Fernandes (ed.). **Educação, Ensino e Geração de Renda: Estratégias e Propostas em Economia Solidária**. 1. ed. São Paulo: Editora Científica Digital LTDA, 2023. v. 1. cap. 6, p. 118.

BARROS, Célio Ribeiro de. **MORTALIDADE DE PROFESSORES NO ESTADO DE GOIÁS: ANÁLISE DE SÉRIE TEMPORAL, 2008-2017**. 2019. f. 71. Diss. (Mestrado) – Pontifícia Universidade Católica de Goiás, Goiânia.

BOX, George E.P. *et al.* **Time Series Analysis: Forecasting and Control**. 5. ed. Nova Jersey: Wiley, 2015.

BOYER, Carl B.; MERZBACH, Uta C. **História da Matemática**. Tradução: Helena Castro. 3. ed. São Paulo: Blucher, 2012.

BRASIL. **Projeto Pedagógico do Curso Superior de Licenciatura em Matemática**. Brasília, 2018. Acesso em: 30 jan. 2025. Disponível em: <https://www.ifb.edu.br/attachments/article/10493/PPC%20Matem%C3%A1tica%202072018%20revisado%20pelo%20NDE.pdf>.

COHEN, Leon. **Time-Frequency Analysis**. 1. ed. Nova Jersey: Pearson College Div, 1994.

ENGEL, Joachim; FRISCHEMEIER, Daniel. Statistical Literacy and Civic Engagement: Teaching and Learning with Data about Society. *In*: VORTÄGE auf der 52. Tagung für Didaktik der Mathematik - Jahrestagung der Gesellschaft für Didaktik der Mathematik. Paderborn: Gesellschaft für Didaktik der Mathematik, 2018. p. 79–80. Acesso em: 30 jan. 2025. Disponível em: <https://eldorado.tu-dortmund.de/server/api/core/bitstreams/790d7826-c265-4935-ae53-7c1018c916a6/content>.

FERNÁNDEZ, Adán J. Corcho; CAVALCANTE, Marcos Petrúcio de A. **Introdução à Análise Harmônica e Aplicações**. Rio de Janeiro: IMPA, 2009. p. 118. ISBN 978-85-244-0299-9.

GABOR, D. Theory of communication. **Institution of Electrical Engineering**, p. 32, 1946.

GADE, Kishore Reddy. Data-Driven Decision Making in a Complex World. **Journal of Computational Innovation**, v. 1, p. 18, 2021. Acesso em: 30 jan. 2025. Disponível em: <https://researchworkx.com/index.php/jci/article/view/2/2>.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Editora Atlas, 2002.

GOLDBERGER, A. *et al.* PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. **Circulation [Online]**, v. 101, n. 23, e215–e220, 2000. Acesso em: 30 jan. 2025. DOI: 10.1161/01.CIR.101.23.e215. Disponível em: <https://physionet.org/>.

INC., Yahoo. **Yahoo Finance**. [S. l.: s. n.], 2024. Acesso em: 30 jan. 2025. Disponível em: <https://finance.yahoo.com/>.

KREYSZIG, Erwin. **Advanced Engineering Mathematics**. 10. ed. Nova Iorque: John Wiley & Sons, inc., 2011.

MAO, Shengzhong *et al.* Time Series Analysis for Education: Methods, Applications, and Future Directions. **arXiv preprint**, arXiv:2408.13960, 2024. Acesso em: 30 jan. 2025. Disponível em: <https://arxiv.org/abs/2408.13960>.

MOODY, G.B.; MARK, R.G. The impact of the MIT-BIH Arrhythmia Database. **IEEE Engineering in Medicine and Biology Magazine**, v. 20, n. 3, p. 45–50, 2001. DOI: 10.1109/51.932724.

MORETTIN, Pedro A.; TOLOI, Clélia M. C. **Análise de Séries Temporais**. 2. ed. São Paulo: Blucher, 2006.

OSGOOD, Brad. **Lecture Notes for EE 261: The Fourier Transform and its Applications**. [S. l.: s. n.], 2007. p. 428. Notas de aula. Acesso em: 30 jan. 2025. Disponível em: <https://see.stanford.edu/materials/lsoftaee261/book-fall-07.pdf>.

PERCIVAL, Donald B.; WALDEN, Andrew T. **Spectral analysis for physical applications: multipaper and conventional univariate techniques**. 1. ed. Nova Iorque: Cambridge University Press, 1993.

PHUNGTUA-ENG, Thanapol; YAMAMOTO, Yoshitaka. Adaptive Seasonal-Trend Decomposition for Streaming Time Series Data with Transitions and Fluctuations in Seasonality. *In*: BIFET, Albert *et al.* (ed.). **Machine Learning and Knowledge Discovery in Databases. Research Track**. Cham: Springer Nature Switzerland, 2024. p. 426–443. ISBN 978-3-031-70344-7.

PIFER, Anderson; AURANI, Katya Margareth. A teoria analítica do calor de Joseph Fourier: uma análise das bases conceituais e epistemológicas. **Revista Brasileira de Ensino de Física**, Sociedade Brasileira de Física, v. 37, n. 1, 2015. Acesso em: 30 jan. 2025. ISSN 1806-1117. DOI: 10.1590/S1806-11173711681. Disponível em: <https://doi.org/10.1590/S1806-11173711681>.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2023. Acesso em: 30 jan. 2025. Disponível em: <https://www.R-project.org/>.

RICHARDSON, Lee. **The Sliding Window Discrete Fourier Transform**. 2019. f. 169. Tese (Doutorado) – Carnegie Mellon University, Pittsburgh.

RYAN, Jeffrey A.; ULRICH, Joshua M. **quantmod: Quantitative Financial Modelling Framework**. [S. l.], 2024. Acesso em: 30 jan. 2025. Disponível em: <https://CRAN.R-project.org/package=quantmod>.

SAI, Kurma Venkata Keerthana *et al.* Automated ASD detection in children from raw speech using customized STFT-CNN model. **International Journal of Speech Technology**, v. 27, n. 3, p. 701–716, set. 2024. Acesso em: 15 fev. 2025. ISSN 1572-8110. DOI: 10.1007/s10772-024-10131-7. Disponível em: <https://doi.org/10.1007/s10772-024-10131-7>.

SEVERINO, Antônio Joaquim. **Metodologia do trabalho científico**. 2. ed. São Paulo: Cortez editora, 2013.

SHEN, Mingkan *et al.* A real-time epilepsy seizure detection approach based on EEG using short-time Fourier transform and Google-Net convolutional neural network. English. **Heliyon**, v. 10, n. 11, jun. 2024. Acesso em: 15 fev. 2025. ISSN 2405-8440. DOI: 10.1016/j.heliyon.2024.e31827. Disponível em: [https://www.cell.com/heliyon/abstract/S2405-8440\(24\)07858-7](https://www.cell.com/heliyon/abstract/S2405-8440(24)07858-7). Acesso em: 14 fev. 2025.

SMITH, Steven W. **The Scientist and Engineer's Guide to Digital Signal Processing**. 2. ed. California: California Technical Publishing, 1999.

SOLOW, Daniel. **How to Read and Do Proofs: An Introduction to Mathematical Thought Processes**. 6. ed. Ohio: John Wiley & Sons, inc., 2013.

SOONTORNAPAR, Tomorn; PLOYSUWAN, Tuchsanai. A novel approach to enhanced fall detection using STFT and magnitude features with CNN autoencoder. **Neural Computing and Applications**, v. 37, n. 6, p. 4229–4245, fev. 2025. Acesso em: 14 fev. 2025. ISSN 1433-3058. DOI: 10.1007/s00521-024-10845-4. Disponível em: <https://doi.org/10.1007/s00521-024-10845-4>.

SOUSA, Angélica Silva de *et al.* A Pesquisa Bibliográfica: Princípios e Fundamentos. **Cadernos da Fucamp**, p. 64–83, 2021.

USHEY, Kevin *et al.* **reticulate: Interface to 'Python'**. [S. l.], 2024. Acesso em: 30 jan. 2025. Disponível em: <https://CRAN.R-project.org/package=reticulate>.

WICKHAM, Hadley. **ggplot2: Elegant Graphics for Data Analysis**. [S. l.]: Springer-Verlag New York, 2016. Acesso em: 30 jan. 2025. ISBN 978-3-319-24277-4. Disponível em: <https://ggplot2.tidyverse.org>.

XIE, C. *et al.* **Waveform Database Software Package (WFDB) for Python (version 4.1.0)**. [S. l.], 2023. Acesso em: 30 jan. 2025. DOI: 10.13026/9njx-6322. Disponível em: <https://doi.org/10.13026/9njx-6322>.