

UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
FACULDADE DE ENGENHARIA DE SÃO JOÃO DA BOA VISTA

WOLNEY DALLA PRIA NETO

Aplicação de Redes Neurais Informadas por Física na Equação de Burgers

São João da Boa Vista

2025

Wolney Dalla Pria Neto

Aplicação de Redes Neurais Informadas por Física na Equação de Burgers

Trabalho de Graduação apresentado ao Conselho de Curso de Graduação em Engenharia Aeronáutica da Faculdade de Engenharia de São João da Boa Vista, Universidade Estadual Paulista, como parte dos requisitos para obtenção do diploma de Graduação em Engenharia Aeronáutica.

Orientador: Prof^a Dr^a. Leandra Isabel de Abreu

São João da Boa Vista
2025

N469a

Neto, Wolney Dalla Pria

Aplicação de redes neurais informadas por física na equação de Burgers / Wolney Dalla Pria Neto. -- São João da Boa Vista, 2025
54 p. : il.

Trabalho de conclusão de curso (Bacharelado - Engenharia Aeronáutica) - Universidade Estadual Paulista (UNESP), Faculdade de Engenharia, São João da Boa Vista

Orientadora: Leandra Isabel de Abreu

1. Aprendizado do computador. 2. Burgers, Equação de. 3. Turbulência. I. Título.

**UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
FACULDADE DE ENGENHARIA - CÂMPUS DE SÃO JOÃO DA BOA VISTA
GRADUAÇÃO EM ENGENHARIA AERONÁUTICA**

TRABALHO DE CONCLUSÃO DE CURSO

**APLICAÇÃO DE REDES NEURAIS INFORMADAS POR FÍSICA NA EQUAÇÃO DE
BURGERS**

Aluno: Wolney Dalla Pria Neto
Orientador: Prof.^a Dr.^a Leandra Isabel de Abreu

Banca Examinadora:

- Leandra Isabel de Abreu (Orientadora)
- Ivan Aritz Aldaya Garde (Examinador)
- Angelo José Castro Alves Ferreira Filho (Examinador)

Os formulários de avaliação e a ata da defesa, na qual consta a aprovação do trabalho, devidamente assinados pela banca encontram-se no prontuário eletrônico do aluno.

São João da Boa Vista, 08 de dezembro de 2025

AGRADECIMENTOS

Quero dedicar este trabalho, primeiramente, aos meus pais, que sempre me motivaram a seguir meus sonhos e que investiram tempo, esforços e recursos para que eu pudesse alcançar esta etapa tão importante da minha vida. Sem o apoio incondicional deles, dificilmente eu teria conseguido concluir este curso. À minha avó Maria, que sempre foi uma das minhas maiores inspirações na busca pelo conhecimento e um exemplo de resiliência e superação. À minha irmã Amanda, por ser um dos meus pilares emocionais ao longo dessa jornada. E, de forma especial, a todos os membros da minha família com quem convivi durante esses anos e que, à sua maneira, contribuíram para que eu chegasse até aqui.

Registro também minha profunda gratidão a todos os professores deste campus, pelos quais tive o privilégio de ser ensinado. Sem eles — e, principalmente, sem a excepcional didática que demonstraram ao longo do curso — eu não teria desenvolvido a base acadêmica necessária para percorrer este caminho. Dentre esses professores, expresso meu agradecimento especial à minha orientadora, Prof.^a Dr.^a Leandra Isabel Abreu, por me introduzir ao universo da pesquisa científica através da iniciação científica e pelo auxílio imensurável não apenas neste trabalho, mas também nas últimas disciplinas da graduação.

Por fim, dedico este trabalho, com todo o meu carinho, aos amigos e amigas que construí ao longo do curso. Sei que, sem vocês, eu certamente não teria chegado até aqui. Seria egoísmo afirmar que esta conquista é apenas minha; grande parte dela pertence a vocês, que estudaram comigo, riram comigo, enfrentaram dificuldades ao meu lado e que, espero também, possam comemorar comigo a nossa vitória.

*“Se, a princípio, a ideia não é absurda, então não há esperança para ela.”
(Albert Einstein)*

RESUMO

A equação de Burgers desempenha um papel fundamental no estudo de fenômenos envolvendo advecção não linear e difusão viscosa, constituindo-se como um importante protótipo para a análise de métodos numéricos e modelos de aprendizado profundo aplicados à solução de equações diferenciais parciais. Neste trabalho, a equação de Burgers unidimensional foi resolvida por três abordagens distintas: a solução analítica via transformação de Cole–Hopf, a solução numérica obtida pelo método semi-implícito de Crank–Nicolson e a solução aproximada produzida por meio de redes neurais informadas pela física (*Physics-Informed Neural Networks* — PINNs). A solução analítica foi utilizada como referência para a avaliação das demais metodologias. Os resultados mostraram que o método numérico apresenta um bom desempenho global porém, com discrepâncias localizadas na região de gradiente acentuado, decorrentes especialmente do tratamento explícito da parcela advectiva. Em contraste, a PINN apresentou excelente concordância com a solução analítica, mantendo baixos níveis de erro absoluto, uma distribuição estatística concentrada e estabilidade temporal, mesmo em regiões de forte não linearidade. A análise comparativa evidenciou o potencial das PINNs como uma alternativa moderna e eficaz para a solução de equações diferenciais, sobretudo em problemas em que a precisão local desempenha um papel relevante.

Palavras-Chave: equação de Burgers; Crank–Nicolson; *physics-informed neural networks*; métodos numéricos; solução analítica; *machine learning*.

ABSTRACT

The Burgers' equation plays a fundamental role in the study of phenomena involving nonlinear advection and viscous diffusion, serving as an important prototype for evaluating numerical methods and deep learning models applied to partial differential equations. In this work, the one-dimensional Burgers' equation was solved using three distinct approaches: the analytical solution obtained through the Cole–Hopf transformation, the numerical solution computed with a semi-explicit Crank–Nicolson scheme, and the approximate solution generated by Physics-Informed Neural Networks (PINNs). The analytical solution was adopted as the reference for assessing the performance of the other methods. The results showed that the numerical method achieves good global performance, although localized discrepancies arise near the steep gradient region, mainly due to the explicit treatment of the advective term. In contrast, the PINN solution exhibited excellent agreement with the analytical reference, maintaining low absolute errors, highly concentrated statistical distributions, and temporal stability even in regions governed by strong nonlinear behavior. The comparative analysis highlights the potential of PINNs as a modern and effective alternative for solving differential equations, particularly in problems where local accuracy plays a crucial role.

Keywords: Burgers's equation; Crank–Nicolson; physics-informed neural networks; numerical methods; analytical solution; machine learning.

LISTA DE ILUSTRAÇÕES

Figura 1	Superfície tridimensional da solução analítica via transformação de Cole–Hopf, exibindo a dissipação progressiva da inclinação acentuada conforme o tempo avança.	38
Figura 2	Mapa de contorno da solução analítica, evidenciando a suavidade da evolução espaço–temporal e a dissipação progressiva do gradiente acentuado.	39
Figura 3	Superfície tridimensional da solução numérica obtida pelo método de Crank–Nicolson, mostrando boa concordância global com a solução analítica, porém com leve deslocamento na região de maior gradiente.	40
Figura 4	Mapa de contorno da solução numérica obtida pelo método de Crank–Nicolson. A transição espaço–temporal segue o padrão analítico, mas evidencia um pequeno deslocamento do gradiente mais intenso.	41
Figura 5	Superfície tridimensional da solução obtida pela PINN, apresentando excelente acordo com a solução analítica em todo o domínio espaço–temporal.	42
Figura 6	Mapa de contorno da solução via PINN, exibindo níveis de cor extremamente regulares e elevada proximidade com o comportamento analítico.	43
Figura 7	Comparação entre as soluções analítica, numérica e via PINN no instante $t = 0,5$. Observa-se a superior precisão da solução PINN na captura da região de maior gradiente.	44
Figura 8	Erro absoluto entre a solução numérica e a solução analítica. As maiores discrepâncias concentram-se na região de maior variação espacial da solução.	45
Figura 9	Erro absoluto entre a solução via PINN e a solução analítica. As discrepâncias são reduzidas e apresentam distribuição homogênea.	46
Figura 10	Erro absoluto entre a solução via PINN e a solução numérica. As discrepâncias mais intensas coincidem com regiões onde o método numérico se afasta da solução analítica.	47
Figura 11	Distribuição dos erros entre a solução numérica e a solução analítica. A presença de cauda longa reflete a sensibilidade do método numérico à região de maior gradiente.	48
Figura 12	Distribuição dos erros entre a solução via PINN e a solução analítica. A elevada concentração em torno de zero evidencia a precisão da abordagem neural.	48
Figura 13	Distribuição dos erros entre a solução via PINN e a solução numérica. A faixa estreita de dispersão confirma a maior fidelidade da PINN ao comportamento analítico.	49
Figura 14	Estatísticas globais dos erros médio, máximo e RMS para os três pares de comparação. A solução via PINN apresenta desempenho significativamente superior.	50

Figura 15 Evolução temporal do erro RMS. A solução via PINN mantém erro reduzido e estável ao longo do tempo. 51

Figura 16 Histórico da função de perda durante o treinamento da PINN. A tendência global aponta para convergência estável do modelo. 52

LISTA DE ABREVIATURAS E SIGLAS

EDPs	Equações Diferenciais Parciais
LHS	<i>Latin Hypercube Sampling</i>
PINNs	<i>Physics-Informed Neural Networks</i>
MLP	<i>Multilayer Perceptron</i>
RMS	<i>Root Mean Square</i>

LISTA DE SÍMBOLOS

$u(x, t)$	Solução da equação de Burgers unidimensional, dependente do espaço e do tempo.
$u_0(x)$	Condição inicial da equação de Burgers.
$\psi(x, t)$	Função auxiliar introduzida pela transformação de Cole–Hopf.
$F(x, y, t)$	Funcional auxiliar utilizado na formulação integral da solução analítica.
$F_{\min}(x, t)$	Valor mínimo do funcional $F(x, y, t)$, empregado para estabilidade numérica.
x	Coordenada espacial.
t	Tempo.
y	Variável de integração associada à formulação analítica.
ν	Coefficiente de viscosidade cinemática.
Δx	Passo de discretização espacial.
Δt	Passo de discretização temporal.
N_x	Número de pontos da malha espacial.
N_t	Número de passos de tempo.
u_i^n	Aproximação numérica da solução no ponto espacial i e instante n .
\mathcal{L}	Função de perda total da PINN.
$\mathcal{L}_{\text{dados}}$	Termo da função de perda associado às condições iniciais e de contorno.
$\mathcal{L}_{\text{física}}$	Termo da função de perda associado ao residual da equação diferencial.
N_c	Número de pontos de colocação da PINN.
N_d	Número de pontos associados às condições iniciais e de contorno.
ε_{rms}	Erro quadrático médio.
$\varepsilon_{\text{mean}}$	Erro médio absoluto.
ε_{max}	Erro máximo.

SUMÁRIO

1	INTRODUÇÃO	14
2	OBJETIVOS	16
2.1	Objetivos principais	16
2.2	Objetivos específicos	16
3	FUNDAMENTAÇÃO TEÓRICA	17
3.1	A Equação de Burgers	17
3.1.1	Forma conservativa	17
3.1.2	Interpretação física	17
3.1.3	Interpretação matemática	17
3.1.4	Estrutura de choque viscoso	18
3.1.5	Relevância contemporânea	18
3.2	Transformação de Cole–Hopf	19
3.2.1	Linearização da Equação de Burgers	19
3.2.2	Solução integral da equação do calor	19
3.2.3	Determinação da condição inicial transformada	20
3.3	Soluções Analíticas Clássicas	20
3.4	Métodos Numéricos Baseados em Diferenças Finitas	21
3.5	Redes Neurais Artificiais	23
3.6	Perceptron Multicamadas	24
3.7	Physics-Informed Neural Networks	25
4	METODOLOGIA	28
4.1	Ambiente Computacional	29
4.2	Caracterização do Problema e Parâmetros Adotados	29
4.3	Solução Analítica via Transformação de Cole–Hopf	30
4.4	Solução Numérica por Diferenças Finitas (Esquema de Crank–Nicolson)	31
4.5	Solução via Redes Neurais Informadas pela Física	33
4.6	Interpolação das Soluções e Preparação para Comparação	34
4.7	Cálculo das Métricas de Erro	35
5	RESULTADOS E DISCUSSÕES	37
5.1	Visualização das Soluções	37
5.1.1	Solução Analítica	37
5.1.2	Solução Numérica (Crank–Nicolson)	39
5.1.3	Solução via PINNs	41
5.2	Comparação Pontual no Instante $t = 0,5$	43

5.3	Mapas Espaço–Temporais de Erro	44
5.4	Distribuição dos Erros	47
5.5	Estatísticas Globais de Erro	49
5.6	Evolução Temporal do Erro RMS	50
5.7	Convergência do Treinamento da PINN	51
6	CONCLUSÃO	53
	REFERÊNCIAS	55

1 INTRODUÇÃO

A modelagem e a resolução de equações diferenciais parciais (EDPs) constituem pilares fundamentais na descrição matemática de fenômenos físicos presentes em áreas como dinâmica de fluidos, transferência de calor, transporte de massa, acústica e diversos processos de engenharia. Muitas dessas EDPs apresentam termos não lineares que aumentam significativamente a complexidade computacional necessária para sua solução, exigindo métodos numéricos robustos ou técnicas analíticas apropriadas. Entre essas equações, a equação de Burgers ocupa um lugar de destaque por combinar, de forma simples e didática, mecanismos simultâneos de advecção não linear e difusão viscosa (Burgers, 1948; Savović; Ivanović; Min, 2023).

A equação de Burgers é frequentemente interpretada como um modelo simplificado para o estudo de escoamentos viscosos unidimensionais e para a formação de ondas de choque em regimes de baixa viscosidade. Sua estrutura matemática permite analisar fenômenos de interesse em mecânica dos fluidos, gás dinâmico, elasticidade e dispersão longitudinal, conforme apontado por (Benton; Platzman, 1972) e (Kutluay; Bahadır; Özdeş, 1999). Além disso, sua versatilidade faz com que seja amplamente empregada como caso-teste na validação de algoritmos numéricos, como métodos explícitos e implícitos de diferenças finitas, esquemas de alta ordem, métodos espectrais e técnicas analíticas baseadas em transformações, como a de Hopf–Cole (Cole, 1951; El-Danaf; Ramadan, 2007; Hassanien; Salama; Hosham, 2005).

No contexto das soluções analíticas, a transformação de Cole–Hopf desempenha um papel central ao transformar a equação de Burgers em uma equação do calor linear, permitindo a obtenção de soluções fechadas para uma ampla gama de condições iniciais. Benton e Platzman (1972) publicaram um catálogo de 35 soluções analíticas distintas, reforçando a importância dessa equação como ferramenta didática e computacional. Métodos analíticos alternativos, como o método da função *tanh* e o método de decomposição de Adomian, também têm sido empregados com sucesso na obtenção de soluções particulares (El-Danaf; Ramadan, 2007).

Com o avanço da capacidade computacional e o crescimento das técnicas de inteligência artificial, novas abordagens passaram a surgir para a solução de EDPs. Dentre elas, destaca-se o método das *Physics-Informed Neural Networks* (PINNs), introduzido por Raissi, Perdikaris e Karniadakis (Raissi; Perdikaris; Karniadakis, 2019). Diferentemente das redes neurais tradicionais, as PINNs incorporam diretamente as leis físicas — expressas na forma de equações diferenciais — na função de perda do processo de treinamento. Essa característica permite que o modelo aprenda não apenas a estrutura da solução, mas também respeite rigorosamente as restrições impostas pela própria EDP, como condições de contorno, condições iniciais e a consistência física dos termos.

Em particular, sua aplicação à equação de Burgers tem demonstrado excelente desempenho na captura de choques, na representação de gradientes acentuados e na estabilidade da solução em regimes de baixa viscosidade (Savović; Ivanović; Min, 2023).

Assim, este trabalho organiza-se da seguinte forma. No Capítulo 2, são apresentados os fundamentos teóricos necessários para o estudo, incluindo a formulação da equação de Burgers, a transformação

de Cole–Hopf, os métodos numéricos baseados em diferenças finitas e os princípios essenciais das redes neurais artificiais e das PINNs. O Capítulo 3 descreve a metodologia adotada, detalhando os procedimentos de implementação da solução analítica, do método de Crank–Nicolson e do modelo PINN, bem como os parâmetros e critérios de comparação utilizados. No Capítulo 4 são apresentados e discutidos os resultados obtidos para os três métodos, incluindo análises quantitativas e qualitativas do desempenho de cada abordagem. Por fim, o Capítulo 5 sintetiza as conclusões do estudo e aponta possíveis extensões para trabalhos futuros.

2 OBJETIVOS

2.1 Objetivos principais

O presente trabalho tem como propósito investigar diferentes abordagens — analítica, numérica e baseada em aprendizado profundo — para a solução da equação de Burgers unidimensional, um modelo amplamente utilizado como referência na validação de métodos para equações diferenciais parciais. Considerando a crescente relevância das PINNs na solução de problemas governados por leis físicas, este trabalho propõe-se a analisar sua eficácia e comparar seu desempenho com soluções tradicionais consolidadas na literatura.

2.2 Objetivos específicos

Os objetivos específicos deste trabalho consistem em:

- formular e implementar a solução analítica da Equação de Burgers unidimensional por meio da transformação de Cole–Hopf, de modo a obter uma referência confiável para validação dos demais métodos;
- desenvolver uma solução numérica baseada no método de diferenças finitas de Crank–Nicolson, com tratamento implícito do termo difusivo e explícito do termo advectivo;
- implementar uma PINN para a solução da Equação de Burgers, incorporando as condições iniciais, de contorno e o residual da equação diferencial na função de perda;
- analisar a influência dos hiperparâmetros da PINN, como arquitetura da rede, número de pontos de colocação e taxa de aprendizado, na qualidade da solução obtida;
- comparar quantitativamente os resultados fornecidos pelos métodos analítico, numérico e pela PINN por meio de métricas de erro absoluto, erro médio, erro máximo e erro RMS;
- avaliar a capacidade da PINN em capturar regiões de gradiente acentuado, típicas da formação de choques viscosos, em comparação com o método numérico tradicional;
- discutir as vantagens, limitações e potencialidades do uso de PINNs como alternativa aos métodos clássicos para a solução de equações diferenciais parciais não lineares.

3 FUNDAMENTAÇÃO TEÓRICA

3.1 A Equação de Burgers

A equação de Burgers desempenha um papel central no estudo de sistemas dinâmicos não lineares e fenômenos de transporte em meios contínuos. Introduzida originalmente por (Burgers, 1948) no contexto da dinâmica de fluidos turbulentos, sua forma unidimensional viscosa é dada por:

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = \nu \frac{\partial^2 u}{\partial x^2}, \quad (1)$$

onde $u(x, t)$ representa o campo escalar (tipicamente, velocidade) e ν é a viscosidade cinemática.

Para fins de simplicidade notacional e melhor legibilidade ao longo do texto, as derivadas parciais serão representadas, sempre que conveniente, por meio da notação em subscrito. Assim, por exemplo, as expressões u_{xx} e ψ_{xx} denotam, respectivamente, as segundas derivadas parciais de u e ψ em relação à variável espacial.

3.1.1 Forma conservativa

A equação (1) pode ser reescrita na forma conservativa:

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left(\frac{u^2}{2} \right) = \nu \frac{\partial^2 u}{\partial x^2}, \quad (2)$$

forma frequentemente empregada na análise de leis de conservação e em métodos numéricos baseados em volumes finitos (LeVeque, 2002), uma vez que evidencia sua relação estrutural com equações de conservação não lineares. Essa formulação enfatiza que o termo não linear atua como fluxo advectivo, enquanto o termo difusivo promove amortecimento e suavização do campo.

3.1.2 Interpretação física

O termo advectivo uu_x produz auto-transporte da solução, característica típica de escoamentos convectivos. Esse mecanismo gera o acúmulo progressivo de gradientes, podendo resultar na formação de frentes abruptas. Por sua vez, o termo νu_{xx} atua dissipando essas variações rápidas.

O próprio Burgers observou que a equação representa um modelo simplificado, porém altamente instrutivo, para estudar os mecanismos presentes na turbulência e na propagação de ondas viscosas. A competição entre não linearidade e difusão controla a morfologia das soluções.

3.1.3 Interpretação matemática

A análise matemática clássica realizada por (Hopf, 1950) mostrou que a equação de Burgers ocupa uma posição intermediária entre as equações hiperbólicas (onde predominam ondas e choques) e as equações parabólicas (onde predominam efeitos difusivos). Em particular:

- quando $\nu \rightarrow 0$, a equação aproxima-se da equação hiperbólica de conservação

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = 0, \quad (3)$$

que desenvolve descontinuidades finitas em tempo finito;

- quando $\nu > 0$, a difusão impede a formação de choques verdadeiros, produzindo em seu lugar *camadas finas* de transição, cuja espessura é proporcional à viscosidade.

3.1.4 Estrutura de choque viscoso

A análise assintótica apresentada por (Whitham, 1974) mostra que a estrutura estacionária de uma frente do tipo choque na equação de Burgers pode ser descrita por uma solução suave que conecta dois estados constantes U_L (a montante) e U_R (a jusante). Uma forma conveniente dessa solução é dada por:

$$u(x) = a + b \tanh\left(\frac{x - x_0}{2\nu}\right), \quad (4)$$

onde:

- x é a coordenada espacial;
- x_0 é a posição do centro da camada de choque;
- ν é a viscosidade cinemática do meio;
- $a = \frac{U_L + U_R}{2}$ representa o valor médio entre os estados a montante (U_L) e a jusante (U_R); e
- $b = \frac{U_R - U_L}{2}$ representa metade do salto de velocidade entre esses dois estados.

Nessa expressão, a função $\tanh(\cdot)$ descreve a transição suave entre U_L e U_R , enquanto o parâmetro 2ν no denominador do argumento controla a espessura da região de transição: quanto menor a viscosidade ν , mais estreita e íngreme se torna a camada de choque.

3.1.5 Relevância contemporânea

A combinação entre simplicidade formal e riqueza fenomenológica torna a equação de Burgers um modelo de referência na literatura moderna de equações diferenciais parciais. Seu caráter híbrido — simultaneamente advectivo e difusivo — permite investigar, de maneira controlada, a formação de choques viscosos, o comportamento de estruturas de fronteira, e diversas questões relacionadas à estabilidade e à convergência de métodos de solução. Por essas razões, a equação de Burgers é amplamente utilizada como caso-teste na validação de esquemas numéricos, incluindo métodos explícitos e implícitos de diferenças finitas, técnicas espectrais e formulações de ordem elevada. Mais recentemente, sua estrutura bem compreendida tem servido também como base para a comparação entre técnicas tradicionais e abordagens modernas baseadas em aprendizado profundo, como as PINNs, que incorporam diretamente as leis físicas na função de perda durante o treinamento.

3.2 Transformação de Cole–Hopf

A obtenção de soluções analíticas para a equação de Burgers permaneceu como um desafio até a contribuição fundamental de Hopf (Hopf, 1950) e Cole (Cole, 1951), que demonstraram ser possível eliminar a não linearidade do termo $u u_x$ por meio de uma transformação adequada. Tal transformação, hoje conhecida como transformação de Cole–Hopf, estabelece uma correspondência direta entre a equação de Burgers e a equação do calor, permitindo formular soluções fechadas para uma ampla classe de condições iniciais.

3.2.1 Linearização da Equação de Burgers

A transformação proposta por Hopf e Cole consiste em definir uma nova função auxiliar $\psi(x, t)$, tal que:

$$u(x, t) = -2\nu \frac{\partial}{\partial x} \ln \psi(x, t) = -2\nu \frac{\psi_x(x, t)}{\psi(x, t)}. \quad (5)$$

Substituindo (5) em (1) e aplicando derivações sucessivas, os termos não lineares podem ser reorganizados de modo a revelar que a função auxiliar $\psi(x, t)$ satisfaz uma equação linear do tipo difusivo, resultando diretamente na forma clássica da equação do calor:

$$\frac{\partial \psi}{\partial t} = \nu \frac{\partial^2 \psi}{\partial x^2}, \quad (6)$$

Dessa forma, a transformação converte o problema não linear original em um problema linear clássico, cuja solução é amplamente conhecida.

3.2.2 Solução integral da equação do calor

A equação do calor (6) admite solução integral, para $t > 0$, dada por:

$$\psi(x, t) = \frac{1}{\sqrt{4\pi\nu t}} \int_{-\infty}^{\infty} \exp\left[-\frac{(x - \xi)^2}{4\nu t}\right] \psi(\xi, 0) d\xi, \quad (7)$$

onde ξ é a variável de integração espacial. O termo

$$G(x - \xi, t) = \frac{1}{\sqrt{4\pi\nu t}} \exp\left[-\frac{(x - \xi)^2}{4\nu t}\right] \quad (8)$$

corresponde ao núcleo fundamental da equação do calor (também chamado de solução fundamental ou *heat kernel*).

Uma vez obtida a função $\psi(x, t)$ pela integral (7), a solução da equação de Burgers é recuperada usando (5) em que cada termo possui significado físico claro: a viscosidade ν controla a espessura das transições difusivas, enquanto a derivada logarítmica determina a intensidade do gradiente no campo de velocidade.

3.2.3 Determinação da condição inicial transformada

Se a condição inicial da equação de Burgers for dada por:

$$u(x, 0) = u_0(x),$$

então, a condição inicial correspondente para $\psi(x, t)$ é obtida diretamente invertendo a transformação (5). Integrando o termo $u_0(x)$, chega-se a:

$$\psi(x, 0) = \exp \left[-\frac{1}{2\nu} \int^x u_0(\eta) d\eta \right], \quad (9)$$

onde η é a variável de integração. Essa expressão mostra que a transformação de Cole–Hopf não apenas lineariza o problema, mas também fornece um mapeamento explícito entre a condição inicial do problema original e a condição inicial da equação linear associada.

Uma vez determinada $\psi(x, 0)$ conforme (9), a solução do problema evolui da seguinte maneira: a função inicial é propagada no tempo pela integral com núcleo de calor (7), e a solução $u(x, t)$ é restaurada por meio da derivada logarítmica dada por (??). Esse procedimento, amplamente documentado em trabalhos posteriores como (El-Danaf; Ramadan, 2007), representa a via analítica fundamental para a construção de soluções exatas da equação de Burgers viscosa.

3.3 Soluções Analíticas Clássicas

A existência de uma transformação capaz de linearizar a equação de Burgers permitiu o desenvolvimento de um conjunto amplo de soluções analíticas aplicáveis a diferentes condições iniciais e de contorno. Após as formulações pioneiras de Hopf (Hopf, 1950) e Cole (Cole, 1951), Benton e Platzman (Benton; Platzman, 1972) sistematizaram um catálogo abrangente contendo diversas soluções fechadas da equação de Burgers viscosa, as quais se tornaram referência tanto em estudos matemáticos quanto em validações numéricas contemporâneas.

A estrutura fundamental dessas soluções decorre da correspondência direta entre a função auxiliar $\psi(x, t)$, que satisfaz a equação do calor, e a solução original $u(x, t)$ da equação de Burgers. Para cada condição inicial prescrita para $u(x, 0) = u_0(x)$, a transformação de Cole–Hopf conduz à condição inicial associada para $\psi(x, 0)$ por meio de

$$\psi(x, 0) = \exp \left[-\frac{1}{2\nu} \int^x u_0(\eta) d\eta \right],$$

onde η representa a variável de integração e ν é a viscosidade cinemática. Em seguida, a evolução temporal de $\psi(x, t)$ é obtida pela convolução da condição inicial com o núcleo fundamental da equação do calor,

$$\psi(x, t) = \frac{1}{\sqrt{4\pi\nu t}} \int_{-\infty}^{\infty} \exp \left[-\frac{(x - \xi)^2}{4\nu t} \right] \psi(\xi, 0) d\xi,$$

em que ξ é a variável de integração espacial. A solução correspondente de $u(x, t)$ é então reconstruída pela expressão (5) que estabelece a ligação entre a solução difusiva linear e o campo de velocidade

associado à equação de Burgers.

No estudo clássico conduzido por (Benton; Platzman, 1972), diversos perfis iniciais foram examinados, revelando comportamentos característicos, como a formação de frentes suaves do tipo choque, o aparecimento de gradientes intensos em regiões localizadas e a dissipação progressiva de estruturas inicialmente acentuadas. Perfis descontínuos, como os saltos de Heaviside, resultam em soluções que assumem a forma de transições suaves governadas pela função hiperbólica tangente; perfis sinusoidais apresentam amortecimento difusivo temporalmente dependente; e distribuições exponenciais e gaussianas evoluem de modo controlado pelas propriedades do núcleo de calor. Em todos esses casos, a solução reflete o equilíbrio entre auto-transporte não linear e difusão viscosa, característica intrínseca da equação de Burgers.

Como discutido por (Whitham, 1974), a partir da solução geral fornecida pela transformação de Cole–Hopf, pode-se analisar ainda regimes assintóticos relevantes, especialmente no limite em que a viscosidade ν assume valores muito pequenos. Nesse regime de fraca dissipação, a solução aproxima-se daquela de uma equação de conservação hiperbólica, desenvolvendo zonas de transição bastante estreitas, cuja estrutura interna continua, entretanto, plenamente descrita pela solução analítica derivada da equação do calor. Esse fenômeno fornece um arcabouço rigoroso para o estudo de choques viscosos, constituindo uma das contribuições mais relevantes das soluções analíticas à teoria da equação de Burgers.

O conjunto organizado de soluções apresentado por (Benton; Platzman, 1972) tornou-se um instrumento fundamental na validação de métodos numéricos, uma vez que permite comparar resultados computacionais com soluções exatas em condições controladas. Tal comparação é especialmente relevante em estudos envolvendo discretizações de diferenças finitas, técnicas espectrais, esquemas implícitos e, mais recentemente, métodos baseados em aprendizado profundo, cuja avaliação de desempenho requer, sempre que possível, uma referência analítica para aferição da precisão. Assim, as soluções analíticas obtidas pela transformação de Cole–Hopf permanecem como referência indispensável não apenas para a análise teórica, mas também para o desenvolvimento e validação de abordagens computacionais modernas aplicadas à equação de Burgers.

3.4 Métodos Numéricos Baseados em Diferenças Finitas

A solução numérica da equação de Burgers desempenha um papel fundamental na análise de sua dinâmica em regimes nos quais a obtenção de soluções analíticas se torna impraticável. A abordagem por diferenças finitas, amplamente estudada em formulações clássicas e contemporâneas, consiste na discretização do domínio contínuo em malhas espaciais e temporais e na substituição das derivadas que compõem a equação diferencial por aproximações algébricas que preservam, em algum grau, a estrutura de transporte e difusão da equação original. Trabalhos como os de (Kutluay; Bahadır; Özdeş, 1999) e (Hassanien; Salama; Hosham, 2005) dedicam-se a investigar a acurácia dessas discretizações para diversos valores de viscosidade e diferentes perfis iniciais, evidenciando sua relevância para estudos comparativos e validações experimentais.

Considere o domínio espacial $x \in [a, b]$, particionado em N subintervalos de comprimento uniforme $\Delta x = (b - a)/N$, e o domínio temporal discretizado em passos de tamanho Δt . Os pontos

da malha são denotados por $x_i = a + i\Delta x$, para $i = 0, 1, \dots, N$, enquanto $t^n = n\Delta t$ identifica o instante temporal correspondente ao passo n . A solução numérica é representada por $u_i^n \approx u(x_i, t^n)$, aproximando o valor da solução exata nos nós da malha.

A equação de Burgers viscosa, definida em (1), pode ser discretizada substituindo-se a derivada temporal por um quociente progressivo, enquanto a derivada espacial de primeira ordem pode ser aproximada por diferenças centradas, de forma a evitar a assimetria numérica associada a esquemas puramente progressivos ou regressivos. A segunda derivada espacial, por sua vez, é comumente aproximada por diferenças centradas de segunda ordem, cuja estrutura simétrica assegura uma representação consistente da difusão. Dessa maneira, um esquema típico assume a forma

$$\frac{u_i^{n+1} - u_i^n}{\Delta t} + u_i^n \frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x} = \nu \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{(\Delta x)^2},$$

em que o termo de advecção é tratado explicitamente e o termo difusivo segue a formulação clássica de segunda ordem. Nesta expressão, o índice i identifica a posição espacial, e o índice n representa o nível temporal. O parâmetro ν é a viscosidade cinemática, enquanto Δx e Δt determinam a resolução da malha.

A estabilidade de esquemas explícitos para a equação de Burgers é fortemente condicionada pelo número de Courant associado ao termo advectivo e pelo número de Fourier associado ao termo difusivo. Assim, o passo temporal Δt deve respeitar simultaneamente as restrições impostas por esses dois mecanismos, de forma que a relação entre transporte não linear e dissipação viscosa seja preservada sem amplificação espúria de perturbações numéricas. Conforme discutido em (Whitham, 1974), a formação de frentes abruptas em regimes de baixa viscosidade exige malhas suficientemente refinadas, de modo que a espessura da camada de transição seja adequadamente resolvida pelo esquema de diferenças finitas. Em situações em que ν assume valores muito pequenos, o esquema explícito tende a se tornar excessivamente rígido, o que motiva o uso de formulações implícitas.

Métodos implícitos, como os estudados em (Hassanien; Salama; Hosham, 2005), substituem as derivadas espaciais calculadas em t^n por aproximações avaliadas em t^{n+1} , resultando em uma equação algébrica a ser resolvida para todos os valores de u_i^{n+1} . A principal vantagem dessa abordagem reside no fato de que a estabilidade tende a ser menos restritiva, permitindo passos temporais maiores. Entretanto, a presença do termo não linear uu_x implica que sistemas algébricos não lineares devam ser resolvidos iterativamente a cada passo, aumentando o custo computacional, mas garantindo maior robustez em regimes dominados por gradientes intensos.

A comparação entre formulações explícitas, implícitas e híbridas permite compreender com clareza a natureza das soluções numéricas da equação de Burgers. Em condições de viscosidade moderada, todos os métodos tendem a reproduzir adequadamente a solução analítica, mas, à medida que ν se aproxima de zero, a fidelidade dos resultados torna-se fortemente dependente da resolução da malha e da forma como o termo advectivo é discretizado. Trabalhos como o de (Kutluay; Bahadır; Özdeş, 1999) evidenciam que discretizações centradas podem apresentar oscilações locais em regiões de choque, enquanto formulações upwind introduzem dissipação numérica adicional que suaviza a solução. O equilíbrio entre precisão e estabilidade, portanto, é um aspecto essencial no estudo computacional da

equação de Burgers.

A relevância desses métodos não se limita ao contexto tradicional das diferenças finitas, mas se estende ao campo contemporâneo das redes neurais informadas por física, uma vez que tais métodos constituem a referência clássica para a validação dos resultados obtidos por técnicas de aprendizado profundo. A comparação direta entre soluções analíticas, soluções numéricas clássicas e soluções obtidas por modelos baseados em inteligência artificial representa um dos caminhos mais consistentes para avaliar as vantagens, limitações e regimes de aplicabilidade de cada abordagem.

3.5 Redes Neurais Artificiais

As redes neurais artificiais constituem um paradigma fundamental na área de aprendizado de máquina e desempenham um papel central em aplicações envolvendo aproximação de funções, regressão contínua e modelagem de sistemas dinâmicos não lineares (Goodfellow; Bengio; Courville, 2016; Haykin, 1999). Sua relevância no contexto da solução de equações diferenciais decorre da capacidade de aproximar, com elevado grau de precisão, funções suaves definidas em domínios de dimensão arbitrária, desde que disponham de uma estrutura adequada de parametrização. Conforme discutido de forma abrangente em (Goodfellow; Bengio; Courville, 2016), tais modelos operam por meio de composições sucessivas de transformações lineares e não lineares, ajustadas por algoritmos de otimização baseados em gradientes, de modo a minimizar uma função de custo que mede o erro entre a saída produzida pela rede e o comportamento esperado.

A formulação mais elementar de um neurônio artificial consiste na aplicação de uma transformação afim seguida de uma função de ativação não linear. Dado um vetor de entrada $\mathbf{x} \in R^d$, uma camada de neurônios produz uma saída $\mathbf{h} \in R^m$ dada por

$$\mathbf{h} = \sigma(\mathbf{W}\mathbf{x} + \mathbf{b}), \quad (10)$$

em que $\mathbf{W} \in R^{m \times d}$ representa a matriz de pesos, $\mathbf{b} \in R^m$ é o vetor de *bias* e $\sigma(\cdot)$ é a função de ativação responsável por introduzir não linearidade ao modelo. A composição sucessiva de várias dessas camadas define uma arquitetura do tipo perceptron multicamadas, cuja profundidade e largura determinam sua capacidade de representação. Redes rasas com um único conjunto de parâmetros são capazes de aproximar funções contínuas em compactos, em conformidade com o teorema da aproximação universal; entretanto, redes profundas apresentam vantagens computacionais e expressivas documentadas de forma detalhada em (Goodfellow; Bengio; Courville, 2016).

A escolha da função de ativação desempenha um papel decisivo no comportamento de uma rede neural. Funções tradicionalmente utilizadas incluem a tangente hiperbólica, a sigmoide logística e variantes modernas, como a unidade linear retificada (*ReLU*), que introduzem características específicas de curvatura e sensibilidade ao gradiente durante o processo de otimização. A função de ativação deve possuir derivadas bem definidas, uma vez que os algoritmos de treinamento baseados em gradiente, como retropropagação, dependem do cálculo da derivada da saída da rede em relação aos parâmetros. Em redes utilizadas para aproximação de soluções de equações diferenciais, esse aspecto assume particular relevância, pois a avaliação de derivadas de ordem superior com respeito às entradas é

frequentemente necessária.

Uma rede neural parametrizada por θ , representada por $u_\theta(x, t)$, pode ser treinada para aproximar funções que dependem simultaneamente das variáveis espaciais e temporais, desde que sejam fornecidos pares de entrada e saída capazes de representar adequadamente o fenômeno modelado. O processo de treinamento consiste na minimização de uma função de custo $J(\theta)$, que, em sua forma mais simples, mede o erro quadrático médio entre a predição da rede e os valores de referência. O ajuste dos parâmetros é realizado, em geral, por métodos baseados em gradientes, como o algoritmo de Adam, que utiliza estimativas adaptativas dos momentos de primeira e segunda ordem do gradiente.

A eficiência computacional de uma rede neural depende tanto da estrutura da arquitetura quanto do processo de otimização. Redes excessivamente profundas podem introduzir dificuldades relacionadas à dissipação ou explosão do gradiente, enquanto redes muito estreitas podem apresentar capacidade insuficiente para representar funções com variações rápidas ou estruturas não lineares complexas. A regularização, por meio de técnicas como penalização L2, abandono (*dropout*) ou normalização de camadas, pode ser empregada para controlar fenômenos de sobreajuste em cenários com dados reais escassos. No contexto da solução de equações diferenciais, porém, a regularização tende a assumir formas alternativas, sendo controlada principalmente pela inserção explícita de termos que expressam a física do problema, como ocorre nas redes neurais informadas por física apresentadas em (Raissi; Perdikaris; Karniadakis, 2019).

A expressividade e a flexibilidade das redes neurais artificiais fazem com que esses modelos sejam ferramentas especialmente relevantes em problemas que exigem a representação de superfícies multidimensionais complexas. A possibilidade de calcular derivadas de ordem arbitrária com respeito às entradas, utilizando técnicas automáticas de diferenciação, estabelece uma conexão natural entre redes neurais e o estudo de equações diferenciais, permitindo que modelos parametrizados desempenhem o papel de aproximadores contínuos das soluções desses problemas. Tal característica constitui a base essencial para a formulação das PINNs, discutidas em capítulos posteriores.

3.6 Perceptron Multicamadas

O perceptron multicamadas (MLP, do inglês *Multilayer Perceptron*) constitui uma das arquiteturas fundamentais no estudo das redes neurais artificiais (Haykin, 1999; Goodfellow; Bengio; Courville, 2016). Trata-se de um modelo de natureza estritamente feedforward, no qual o fluxo de informação progride da camada de entrada para a camada de saída por meio de transformações lineares sucessivas, seguidas de funções de ativação não lineares. A relevância do MLP decorre de sua capacidade de aproximar funções contínuas definidas em subconjuntos compactos de R^n , propriedade formalizada pelos teoremas de aproximação universal e discutida de maneira sistemática em (Goodfellow; Bengio; Courville, 2016).

Em termos formais, um MLP é composto por um conjunto de camadas ocultas conectadas de maneira totalmente densa. Cada camada k é descrita por uma transformação da forma

$$\mathbf{h}^{(k)} = \sigma(\mathbf{W}^{(k)}\mathbf{h}^{(k-1)} + \mathbf{b}^{(k)}),$$

onde $\mathbf{h}^{(k-1)} \in \mathbb{R}^{d_{k-1}}$ representa o vetor de entrada da camada, $\mathbf{W}^{(k)} \in \mathbb{R}^{d_k \times d_{k-1}}$ é a matriz de pesos, $\mathbf{b}^{(k)} \in \mathbb{R}^{d_k}$ constitui o vetor de *bias* e $\sigma(\cdot)$ é uma função de ativação não linear, responsável por conferir expressividade ao modelo. A composição sucessiva dessas transformações produz, portanto, uma função parametrizada pela coleção de pesos e *bias* $\theta = \{\mathbf{W}^{(k)}, \mathbf{b}^{(k)}\}$, o que define a predição final do modelo como $u_\theta(x, t)$.

A estrutura *feedforward* torna o MLP particularmente adequado à modelagem de fenômenos determinísticos que dependem simultaneamente de variáveis contínuas, como espaço e tempo, sendo essa uma característica essencial para aplicações que envolvem a solução de equações diferenciais. O cálculo eficiente de derivadas sucessivas da função aproximada, viabilizado por técnicas de diferenciação automática, permite que o MLP seja empregado como função de teste na formulação fraca ou forte de problemas governados por equações diferenciais parciais, aspecto explorado em métodos modernos, como as redes informadas por física introduzidas em (Raissi; Perdikaris; Karniadakis, 2019).

Do ponto de vista da estabilidade numérica e da capacidade de representação, a profundidade e a largura das camadas do MLP constituem elementos determinantes. Modelos rasos, embora capazes de representar funções contínuas, exigem um número potencialmente elevado de neurônios, o que se traduz em maior custo computacional. Por outro lado, modelos profundos podem distribuir essa representação entre múltiplas camadas, explorando composições sucessivas de funções não lineares, mas enfrentam dificuldades relacionadas à dissipação do gradiente durante o treinamento, conforme discutido em (Goodfellow; Bengio; Courville, 2016). Técnicas como a inicialização apropriada de parâmetros, normalização intermediária e utilização de funções de ativação com comportamento derivável estável, como ReLU ou variantes suaves, mitigam parte desses efeitos.

A dinâmica de treinamento do MLP ocorre por meio da minimização de uma função de custo $J(\theta)$, geralmente expressa como o erro quadrático médio entre as predições e um conjunto de valores de referência. A atualização dos parâmetros é realizada por algoritmos de otimização baseados em gradientes, como o método de Adam, que utiliza estimativas adaptativas dos momentos do gradiente para acelerar a convergência. A robustez desses métodos, associada à flexibilidade estrutural das redes, permite que o MLP se torne uma ferramenta eficaz na aproximação de campos escalares que obedecem a equações diferenciais, desempenhando o papel de aproximador universal de funções contínuas (Hornik; Stinchcombe; White, 1989).

A importância do MLP no contexto deste trabalho reside na sua utilização como elemento central da arquitetura das PINNs, nas quais a rede é treinada não apenas para reproduzir valores observados, mas também para satisfazer a estrutura diferencial da equação de Burgers. Assim, o perceptron multicamadas funciona como o suporte funcional que possibilita a incorporação explícita das leis físicas no processo de aprendizagem, permitindo que a solução aproximada respeite as condições iniciais, de contorno e a forma exata da equação diferencial subjacente. Essa integração entre a estrutura neural e a formulação matemática constitui o alicerce para o capítulo seguinte.

3.7 Physics-Informed Neural Networks

O método das PINNs constitui uma abordagem moderna e unificadora para a solução de equações diferenciais ordinárias e parciais, baseada na incorporação explícita das leis físicas no processo de

treinamento de redes neurais artificiais. Proposto formalmente por Raissi, Perdikaris e Karniadakis (Raissi; Perdikaris; Karniadakis, 2019), esse método utiliza perceptrons multicamadas como aproximadores universais de funções contínuas, nos quais a solução $u(x, t)$ é parametrizada por uma rede $u_\theta(x, t)$ e ajustada de forma a satisfazer simultaneamente as restrições impostas pelos dados e pela estrutura diferencial do problema.

O princípio fundamental das PINNs consiste em impor que a solução aproximada satisfaça a equação diferencial em todos os pontos de um conjunto de colocação distribuído no domínio. Para a equação de Burgers, expressa por (1), a ideia consiste em definir um *residual físico* $R_\theta(x, t)$, obtido ao substituir a rede neural nos operadores diferenciais do problema,

$$R_\theta(x, t) = \frac{\partial u_\theta}{\partial t} + u_\theta \frac{\partial u_\theta}{\partial x} - \nu \frac{\partial^2 u_\theta}{\partial x^2},$$

cuja anulação caracteriza a aderência da solução aproximada à dinâmica governada pela equação. A avaliação eficiente dessas derivadas é realizada por técnicas de diferenciação automática, que permitem computar derivadas de ordem arbitrária com precisão numérica elevada, aspecto que torna as PINNs particularmente adequadas ao tratamento de equações diferenciais não lineares.

A função de custo das PINNs é construída de forma a agregar as condições físicas e os dados disponíveis. Considerando-se um conjunto de pontos iniciais $(x, 0)$ com valores conhecidos $u(x, 0)$, um conjunto de pontos de contorno (x_b, t) , no qual a solução ou seu fluxo é especificado, e um conjunto de pontos internos (x_f, t_f) no domínio, a função de custo assume a forma geral

$$J(\theta) = J_{IC}(\theta) + J_{BC}(\theta) + J_F(\theta),$$

onde J_{IC} , J_{BC} e J_F representam, respectivamente, os termos associados ao erro nas condições iniciais, nas condições de contorno e no resíduo físico. O termo físico é tradicionalmente expresso como

$$J_F(\theta) = \frac{1}{N_f} \sum_{i=1}^{N_f} |R_\theta(x_i, t_i)|^2,$$

enquanto os termos das condições iniciais e de contorno são definidos por erros quadráticos médios entre a solução prevista pela rede e os valores prescritos. A minimização conjunta desses três componentes assegura que a solução aproximada satisfaça simultaneamente as restrições físicas e informacionais do problema.

A estrutura das PINNs promove uma fusão entre técnicas de aprendizado profundo e métodos clássicos de solução de equações diferenciais, permitindo que a rede aprenda representações contínuas da solução em domínios de alta dimensão ou em geometria complexa. A interpretação dessa abordagem como um método de aproximação baseado em colocação e residual reforça sua conexão com formulações fracas e com métodos espectrais, conforme discutido em (Raissi; Perdikaris; Karniadakis, 2019). No caso da equação de Burgers, as PINNs demonstram capacidade de capturar gradientes acentuados e regiões de choque com precisão satisfatória, desde que a rede disponha de profundidade e largura adequadas para representar os comportamentos não lineares associados à solução.

Do ponto de vista numérico, o treinamento das PINNs envolve desafios relacionados ao equilíbrio entre os diferentes termos da função de custo, à sensibilidade das derivadas na presença de choques e à necessidade de amostragem adequada dos pontos de colocação. Estratégias como reamostragem, ponderação dinâmica dos termos da perda e utilização de otimizadores robustos podem melhorar substancialmente a convergência do modelo. A eficiência computacional das PINNs decorre, em grande medida, da possibilidade de avaliar o residual físico de forma direta, sem a necessidade de discretização explícita da equação diferencial, o que diferencia essa abordagem dos métodos tradicionais baseados em malhas (Raissi; Perdikaris; Karniadakis, 2019).

A versatilidade das PINNs, aliada à sua capacidade de produzir aproximações suaves e diferenciáveis, torna esse método particularmente promissor para problemas em que os dados são escassos ou parcialmente conhecidos. No contexto específico deste trabalho, sua aplicação à equação de Burgers permite a comparação entre métodos puramente numéricos, métodos analíticos clássicos e soluções via aprendizado profundo, estabelecendo uma base consistente para a análise crítica das vantagens e limitações inerentes a cada abordagem.

4 METODOLOGIA

A metodologia adotada neste trabalho foi desenvolvida com o objetivo de comparar, de forma sistemática, três abordagens distintas para a solução da equação de Burgers unidimensional: uma solução analítica obtida por meio da transformação de Cole–Hopf, um método numérico baseado em diferenças finitas e uma abordagem moderna fundamentada em PINNs. A utilização combinada dessas três estratégias permite avaliar, de maneira abrangente, a precisão, a estabilidade e a eficiência computacional de cada método, além de oferecer uma visão integrada que combina técnicas clássicas e contemporâneas de análise de equações diferenciais parciais.

Todo o desenvolvimento computacional foi realizado em ambiente Python, utilizando a plataforma Google Colab, que fornece suporte à diferenciação automática, otimização por métodos de gradiente e aceleração por hardware. O código implementado reflete a estrutura típica das PINNs descritas por Raissi, Perdikaris e Karniadakis (Raissi; Perdikaris; Karniadakis, 2019), adaptada às particularidades do problema estudado.

O desenvolvimento metodológico foi estruturado de forma a garantir a comparabilidade entre os métodos. Para isso, todas as soluções foram obtidas sobre o mesmo domínio espaço–tempo e utilizando os mesmos parâmetros físicos, incluindo a viscosidade do meio, as condições iniciais e as condições de fronteira. Essa padronização assegura que eventuais diferenças entre os resultados sejam atribuídas exclusivamente às características próprias de cada abordagem, e não a discrepâncias no problema formulado.

A primeira estratégia considerada é a solução analítica, viabilizada pela transformação de Cole–Hopf, que converte a equação de Burgers em uma equação do calor e permite a obtenção de expressões exatas para a solução. Essa formulação serve como referência fundamental para todas as comparações realizadas, constituindo um padrão de verdade contra o qual se avaliam os métodos numérico e neural.

Em seguida, emprega-se um método de diferenças finitas para aproximar a solução da equação diferencial no domínio discretizado. A discretização espacial e temporal, bem como o esquema numérico adotado, foram escolhidos de modo a garantir estabilidade e convergência, permitindo a reprodução adequada da dinâmica não linear característica da equação de Burgers.

Por fim, utiliza-se a abordagem das PINNs, na qual uma rede neural profunda é treinada para aproximar a solução da equação por meio da imposição das condições iniciais, das condições de fronteira e da própria dinâmica governada pela equação diferencial. Essa metodologia, baseada na diferenciação automática e na minimização de uma função de perda física, permite obter aproximações contínuas e diferenciáveis sem a necessidade de discretizações explícitas.

Após a obtenção das três soluções — analítica, numérica e via PINN — foi necessário garantir que todas estivessem definidas sobre a mesma malha espaço-temporal, de modo a possibilitar uma comparação direta entre os métodos. Como a solução analítica e a solução obtida pela PINN já foram computadas na mesma malha de referência, nenhum processamento adicional foi necessário para esses dois casos. Entretanto, a solução numérica, originalmente calculada em uma malha mais refinada, precisou ser interpolada para essa malha comum por meio da rotina `griddata`, empregando

interpolação cúbica para preservar a suavidade espacial. A partir das três soluções avaliadas no mesmo conjunto de pontos, foram então calculados o erro absoluto ponto a ponto e estatísticas globais, como o erro médio, o erro máximo e o erro quadrático médio (RMS). Essas métricas permitiram quantificar, de maneira objetiva, a discrepância entre os métodos e avaliar a consistência das soluções ao longo do domínio espaço-temporal.

4.1 Ambiente Computacional

Todas as simulações foram realizadas em ambiente Google Colab, utilizando Python e as seguintes bibliotecas principais:

- NumPy e SciPy para operações numéricas, integração e solução de sistemas lineares;
- Matplotlib para geração de gráficos bidimensionais e tridimensionais;
- TensorFlow/Keras para implementação e treinamento da PINN;
- SciPy.stats.qmc para amostragem por *Latin Hypercube Sampling* (LHS).

O uso de um ambiente em nuvem permite a reprodutibilidade e portabilidade dos resultados, facilitando a execução do código em diferentes máquinas, sem necessidade de configurações específicas de hardware.

4.2 Caracterização do Problema e Parâmetros Adotados

A formulação analisada neste trabalho baseia-se na equação de Burgers unidimensional, definida no domínio espacial

$$x \in [-1, 1],$$

e no intervalo temporal

$$t \in [0, 1].$$

O coeficiente de viscosidade adotado foi

$$\nu = \frac{0.01}{\pi},$$

valor amplamente utilizado em estudos comparativos da literatura por gerar um gradiente acentuado ao longo do tempo, possibilitando avaliar o desempenho de diferentes métodos numéricos e modelos baseados em aprendizado profundo na presença de estruturas semelhantes a choques viscosos.

A condição inicial considerada é dada por

$$u(x, 0) = -\sin(\pi x),$$

expressão tradicional na análise da equação de Burgers por resultar na formação progressiva de gradientes acentuados na região central do domínio (Burgers, 1948; Cole, 1951; Benton; Platzman,

1972; Savović; Ivanović; Min, 2023). As condições de contorno adotadas são homogêneas do tipo Dirichlet:

$$u(-1, t) = 0, \quad u(1, t) = 0.$$

Para a comparação entre os métodos realizada nos capítulos subsequentes, o domínio contínuo foi discretizado de formas distintas, conforme a natureza de cada abordagem. A solução analítica, obtida por meio da transformação de Cole–Hopf, foi avaliada em uma malha de 100×100 pontos no espaço–tempo. Já o método numérico baseado no esquema semi-implícito de Crank–Nicolson empregou uma malha mais refinada, composta por 256 pontos no espaço e 200 passos de tempo, de modo a capturar adequadamente o estreitamento da região de gradiente acentuado.

No caso da PINN, os pontos de treinamento foram selecionados por meio do método LHS. Foram utilizados 50 pontos referentes à condição inicial, 25 pontos em cada contorno espacial e um total de 10 000 pontos de colocalização destinados à imposição do termo residual da equação diferencial.

Esse conjunto de parâmetros assegura coerência entre as metodologias analisadas e reflete fielmente a implementação computacional utilizada neste trabalho.

4.3 Solução Analítica via Transformação de Cole–Hopf

A primeira abordagem considerada neste trabalho consiste na obtenção da solução analítica da equação de Burgers viscosa unidimensional por meio da transformação de Cole–Hopf. Essa técnica, introduzida originalmente por Hopf e desenvolvida por Cole, permite converter a equação não linear (1) em uma equação do calor linear, mediante a introdução da função auxiliar $\psi(x, t)$ definida em (5). A substituição dessa transformação elimina o termo não linear e conduz a (6) permitindo a construção da solução exata a partir da condição inicial do problema.

Para a condição inicial utilizada neste estudo,

$$u(x, 0) = -\sin(\pi x),$$

pode-se introduzir uma função potencial $U(y)$ definida a partir da integral da condição inicial. Essa função é empregada na formulação integral equivalente da solução, usualmente expressa em termos de um funcional $F(x, y, t)$ que combina as contribuições advectivas e difusivas da equação. A solução analítica pode, então, ser escrita na forma de uma razão de integrais, cuja versão utilizada na implementação numérica é dada por

$$u(x, t) \approx \frac{\int_{y_{\min}}^{y_{\max}} \frac{x-y}{t} \exp\left(-\frac{F(x, y, t) - F_{\min}(x, t)}{2\nu}\right) dy}{\int_{y_{\min}}^{y_{\max}} \exp\left(-\frac{F(x, y, t) - F_{\min}(x, t)}{2\nu}\right) dy},$$

em que $F_{\min}(x, t)$ representa o valor mínimo de $F(x, y, t)$ no intervalo considerado. A subtração desse mínimo tem como objetivo evitar instabilidades numéricas devido aos termos exponenciais potencialmente muito pequenos ou muito grandes, proporcionando um melhor condicionamento à integral.

A escolha do intervalo de integração $[y_{\min}, y_{\max}]$ representa um aspecto numérico fundamental dessa abordagem. Embora a formulação contínua da solução envolva integrais definidas no eixo real completo, sua implementação computacional requer a consideração de um intervalo finito. No presente trabalho, adotou-se o intervalo

$$y \in [-1.5, 1.5],$$

que excede o domínio físico do problema, $x \in [-1, 1]$, de forma a garantir que toda a região relevante da integral seja capturada. Essa escolha é justificada pela estrutura da exponencial presente no integrando, a qual contém o termo

$$\exp\left(-\frac{(x-y)^2}{4\nu t}\right),$$

responsável por concentrar a contribuição significativa da integral em uma vizinhança estreita de $y \approx x$. Dado que a viscosidade utilizada neste trabalho é $\nu = 0,01/\pi$, esse termo apresenta um decaimento extremamente rápido, e valores de $|(x-y)| > 1$ já produzem contribuições desprezíveis. Assim, o intervalo estendido $[-1.5, 1.5]$ garante que toda a região de influência esteja incluída, ao mesmo tempo em que evita custos computacionais excessivos associados ao uso de intervalos muito amplos.

Para a avaliação numérica das integrais, empregou-se a regra de Simpson com

$$n_{\text{pontos}} = 1000$$

pontos uniformemente distribuídos em y . Esse valor foi escolhido de forma a assegurar uma boa resolução da região onde a exponencial apresenta maior concentração. A solução analítica foi, então, computada em uma malha retangular composta por

$$n_x = 100 \quad \text{e} \quad n_t = 100,$$

com o intervalo temporal iniciando em

$$t_{\min} = 0.01,$$

a fim de evitar a singularidade da formulação integral em $t = 0$, e estendendo-se até

$$t_{\max} = 1.0.$$

Os valores de $u(x, t)$ obtidos dessa forma constituem a solução analítica amostrada no domínio espaço-tempo, a qual é utilizada como referência para quantificar os erros das soluções numérica e neural. A discretização aplicada neste contexto não introduz erros associados ao método de solução da equação diferencial, representando apenas uma amostragem da solução exata em uma grade compatível com as empregadas nas demais abordagens.

4.4 Solução Numérica por Diferenças Finitas (Esquema de Crank–Nicolson)

A segunda abordagem utilizada neste trabalho corresponde à solução numérica da equação de Burgers viscosa por meio de um esquema de diferenças finitas baseado no método de Crank–Nicolson.

Esse método, amplamente empregado na solução de equações parciais parabólicas, oferece segunda ordem de precisão no espaço e no tempo para a discretização do termo difusivo, além de apresentar estabilidade numérica favorável. De forma geral, o método consiste em uma formulação semi-implícita, na qual o termo difusivo é tratado com uma média entre os instantes t^n e t^{n+1} , enquanto os termos não lineares podem ser tratados de maneira explícita para reduzir o custo computacional.

Considere o domínio discreto formado por uma malha uniforme no espaço e no tempo. No presente trabalho, adotou-se uma discretização com $n_x = 256$ pontos no intervalo espacial $[-1, 1]$, resultando em um espaçamento

$$\Delta x = \frac{x_{\max} - x_{\min}}{n_x - 1},$$

e uma malha temporal com $n_t = 200$ pontos no intervalo $[0, 1]$, com passo

$$\Delta t = \frac{t_{\max}}{n_t - 1}.$$

Denotando por u_i^n a aproximação de $u(x_i, t_n)$, o termo difusivo é tratado pelo esquema de Crank–Nicolson, que medeia os operadores espaciais entre os tempos t_n e t_{n+1} . Introduce-se, então, o parâmetro

$$\alpha = \frac{\nu \Delta t}{2 (\Delta x)^2},$$

o qual aparece nos coeficientes da matriz tridiagonal correspondente ao sistema linear resolvido a cada passo de tempo.

No presente trabalho, a não linearidade advectiva $u u_x$ é tratada explicitamente, utilizando uma discretização central para a derivada espacial, tal como implementado em diversos esquemas semi-implícitos clássicos. Assim, o termo advectivo no vetor do lado direito é aproximado por

$$u_i^n \frac{u_{i+1}^n - u_{i-1}^n}{2\Delta x},$$

enquanto o termo difusivo é incorporado implicitamente via Crank–Nicolson. O sistema linear resultante, de dimensão $n_x \times n_x$, apresenta uma estrutura tridiagonal e é resolvido por um algoritmo eficiente baseado em *banded matrices*.

Em termos computacionais, as matrizes do esquema são construídas a partir dos coeficientes constantes

$$\text{diag principal : } 1 + 2\alpha, \quad \text{diag inferior} = \text{diag superior : } -\alpha,$$

com ajustes apropriados para impor as condições de contorno homogêneas de Dirichlet, isto é,

$$u(-1, t) = u(1, t) = 0.$$

Essas condições são aplicadas impondo-se $u_0^n = u_{n_x-1}^n = 0$ em todos os instantes de tempo, o que coincide com a estrutura do problema físico analisado.

O algoritmo numérico inicia com a aplicação da condição inicial

$$u(x, 0) = -\sin(\pi x),$$

amostrada na malha espacial. Para cada instante t_{n+1} , o termo advectivo é calculado de forma explícita, enquanto o termo difusivo é resolvido implicitamente pela solução de um sistema linear tridiagonal por meio da função `solve_banded`. Esse procedimento garante boa estabilidade numérica mesmo para viscosidades pequenas, como a adotada neste estudo, $\nu = 0,01/\pi$, a qual gera gradientes acentuados ao longo da evolução temporal.

A solução numérica resultante constitui uma aproximação de segunda ordem no espaço e no tempo para o problema considerado, desde que os termos advectivos permaneçam suficientemente regulares. Os resultados obtidos ao longo da malha (x, t) são posteriormente interpolados para uma malha comum com a solução analítica e com a solução obtida pelas redes neurais informadas pela física, possibilitando comparações diretas e o cálculo de medidas estatísticas de erro.

4.5 Solução via Redes Neurais Informadas pela Física

A terceira metodologia utilizada neste trabalho consiste na aplicação das PINNs, uma classe de modelos introduzida por Raissi, Perdikaris e Karniadakis (2019), que incorporam diretamente as equações diferenciais parciais ao processo de treinamento. Diferentemente de arquiteturas tradicionais, nas quais o ajuste ocorre exclusivamente a partir de dados, as PINNs utilizam o residual da EDP como um termo adicional na função de perda, de forma que a rede busca simultaneamente satisfazer a equação física, as condições iniciais e as condições de contorno. Essa estratégia permite explorar as propriedades analíticas do problema e reduzir a dependência de conjuntos extensos de dados observacionais.

No contexto da equação de Burgers viscosa unidimensional a solução é aproximada por uma rede neural que recebe como entrada o par (x, t) e retorna uma estimativa para $u(x, t)$. A rede implementada contempla oito camadas ocultas, cada uma composta por vinte neurônios, e utiliza a função de ativação `tanh` em todas as camadas internas. A camada de saída contém um único neurônio com ativação linear, de modo a permitir a reconstrução de valores reais contínuos para a solução aproximada. Essa arquitetura, relativamente profunda para um problema unidimensional, favorece a capacidade de representar estruturas não lineares e gradientes acentuados presentes na dinâmica da equação de Burgers em regime de baixa viscosidade.

A avaliação das derivadas espaciais e temporais necessárias à construção do residual é feita por autodiferenciação, um mecanismo intrínseco ao ambiente TensorFlow. Se $u_\theta(x, t)$ denota a saída da rede para parâmetros θ , então u_t , u_x e u_{xx} são obtidos por meio da aplicação sucessiva de operadores de gradiente sobre essa função. O residual empregado no cálculo da perda é, portanto,

$$F(t, x) = u_t(x, t) + u(x, t) u_x(x, t) - \nu u_{xx}(x, t),$$

consistente com a formulação diferencial da equação de Burgers. Esse residual é avaliado em um conjunto de pontos amostrados no domínio espaço-tempo por meio de um esquema de LHS (McKay; Beckman; Conover, 1979; Raissi; Perdikaris; Karniadakis, 2019). No presente trabalho, adotou-se um total de $N_c = 10\,000$ pontos de colocação, de modo a cobrir uniformemente o domínio $t \in [0, 1]$ e $x \in [-1, 1]$, garantindo diversidade estatística e reduzindo vieses associados à distribuição dos pontos

internos.

A rede também é treinada para satisfazer os dados artificiais associados à condição inicial e às condições de contorno do problema. Para a condição inicial $u(x, 0) = -\sin(\pi x)$, foram considerados 50 pontos amostrados por LHS no intervalo espacial. As condições de contorno homogêneas de Dirichlet, $u(-1, t) = u(1, t) = 0$, foram representadas por 25 pontos em cada fronteira, resultando em um total de $N_d = 100$ pontos de dados físicos. Esses pontos são convertidos para tensores e incorporados ao termo de erro da solução proposta pela rede, de modo que o treinamento busca simultaneamente satisfazer a dinâmica interna da equação e as restrições impostas nas fronteiras do domínio.

A função de perda empregada neste estudo é dada por

$$\mathcal{L}(\theta) = \mathcal{L}_{\text{física}} + \mathcal{L}_{\text{dados}},$$

em que $\mathcal{L}_{\text{física}}$ corresponde à média dos quadrados do residual da equação avaliada nos pontos de colocalização, enquanto $\mathcal{L}_{\text{dados}}$ representa o erro quadrático médio entre os valores previstos pela rede e os valores prescritos pelas condições inicial e de contorno. Importante destacar que, diferentemente de algumas formulações alternativas presentes na literatura, não foram utilizados pesos multiplicadores λ para regularizar os termos da função de perda; ambos foram somados diretamente, preservando a estrutura original proposta por Raissi et al. (2019). Essa escolha simplifica a implementação e evita ajustes adicionais de hiperparâmetros, particularmente em problemas unidimensionais nos quais a escala do residual e dos dados tende a ser comparável.

O treinamento foi conduzido utilizando o otimizador Adam, com taxa de aprendizado igual a 5×10^{-4} e um total de 4000 épocas. Em cada iteração, os gradientes do erro total em relação aos parâmetros da rede foram obtidos por *automatic differentiation*, e as atualizações dos pesos foram realizadas de forma tradicional, segundo a regra adaptativa do Adam. Esse procedimento garante uma convergência estável mesmo em situações em que o residual apresenta regiões altamente não lineares ou gradientes intensos.

Ao final do treinamento, a solução aprendida pela rede foi avaliada em uma malha regular (x, t) composta por 100×100 pontos, compatível com a malha de referência utilizada na solução analítica.

4.6 Interpolação das Soluções e Preparação para Comparação

A etapa final da metodologia consiste na unificação das três soluções obtidas — analítica, numérica e neural — em uma malha comum espaço-temporal, permitindo comparações diretas e o cálculo consistente das métricas de erro. Cada método produz sua solução em uma discretização distinta, seja em razão das características próprias da implementação ou da necessidade de estabilidade numérica. Assim, a solução analítica via Cole–Hopf é obtida em uma malha de 100×100 pontos, enquanto o método de Crank–Nicolson utiliza uma malha significativamente mais refinada no espaço e moderadamente refinada no tempo, com 256 pontos espaciais e 200 pontos temporais. Já a solução obtida pelas redes neurais informadas pela física é reconstruída após o treinamento em uma malha regular de 100×100 pontos, compatível com a utilizada para a solução analítica.

Para possibilitar a comparação quantitativa, torna-se necessário interpolar a solução numérica para uma malha comum (x, t) compartilhada entre os métodos. No presente trabalho, adotou-se uma malha regular composta por 100 pontos no espaço e 100 pontos no tempo, cobrindo o domínio completo $x \in [-1, 1]$ e $t \in [0, 1]$. Essa escolha garante um equilíbrio entre a resolução suficiente para capturar as principais estruturas da solução e o custo computacional reduzido na avaliação das métricas.

A solução numérica obtida pelo método de Crank–Nicolson, que originalmente reside em uma malha mais fina, foi interpolada para a malha comum utilizando o método de interpolação *cubic* fornecido pela função `griddata`. De forma análoga, a solução da rede neural foi avaliada diretamente nessa malha, eliminando a necessidade de interpolação posterior. A solução analítica, construída originalmente na mesma discretização de 100×100 , foi diretamente utilizada como referência, sem transformações adicionais.

Essa padronização das malhas permite que todas as soluções sejam comparadas ponto a ponto ao longo de todo o domínio espaço–tempo. Com isso, é possível calcular diferenças absolutas e relativas entre as abordagens, bem como medidas globais de erro, tais como o erro médio, o erro máximo e a raiz do erro quadrático médio.

4.7 Cálculo das Métricas de Erro

Com as três soluções expressas sobre a mesma malha espaço–temporal, torna-se possível quantificar de forma rigorosa as diferenças entre os métodos analítico, numérico e neural. A avaliação de desempenho foi conduzida por meio de métricas globais de erro definidas sobre o domínio completo (x, t) , permitindo caracterizar não apenas discrepâncias pontuais, mas também a qualidade geral das aproximações em relação à solução de referência.

As comparações foram realizadas entre três pares de soluções: solução numérica versus solução analítica, solução neural via PINNs versus solução analítica, e solução neural versus solução numérica. Para cada um desses pares, foi calculado o erro absoluto ponto a ponto ao longo da malha comum $(X_{\text{common}}, T_{\text{common}})$, resultando em campos bidimensionais de diferenças que refletem a distribuição espaço–temporal das discrepâncias entre as abordagens.

A partir desses campos de erro, foram extraídas três métricas globais. O erro médio absoluto,

$$\varepsilon_{\text{mean}} = \frac{1}{N} \sum_{i=1}^N |u_i^{(1)} - u_i^{(2)}|,$$

fornece uma medida da discrepância média entre as soluções. O erro máximo,

$$\varepsilon_{\text{max}} = \max_{1 \leq i \leq N} |u_i^{(1)} - u_i^{(2)}|,$$

identifica a maior diferença local observada no domínio, sendo particularmente útil para localizar regiões de maior sensibilidade ou gradientes acentuados. Por fim, a raiz do erro quadrático médio

(RMS),

$$\varepsilon_{\text{rms}} = \sqrt{\frac{1}{N} \sum_{i=1}^N \left(u_i^{(1)} - u_i^{(2)} \right)^2},$$

resume a magnitude global do erro de forma mais sensível a discrepâncias maiores, caracterizando a qualidade geral da aproximação com maior penalização de desvios intensos.

No código implementado, essas métricas foram computadas diretamente a partir das matrizes bidimensionais resultantes da interpolação descrita no bloco anterior, empregando funções vetorizadas da biblioteca NumPy. Esse procedimento foi aplicado para os três pares de comparação mencionados, produzindo estatísticas que permitem analisar tanto a acurácia do método numérico quanto a capacidade da rede neural de reproduzir a solução analítica. Os resultados dessas análises são apresentados e discutidos no capítulo seguinte.

5 RESULTADOS E DISCUSSÕES

Este capítulo apresenta e discute os resultados obtidos com os três métodos empregados na solução da equação de Burgers unidimensional: a solução analítica via transformação de Cole–Hopf, a solução numérica por diferenças finitas (Crank–Nicolson) e a solução aproximada por PINN. A análise foi conduzida tanto no domínio completo (x, t) quanto em pontos ou cortes específicos, permitindo uma avaliação abrangente do comportamento das soluções e da precisão relativa entre os métodos.

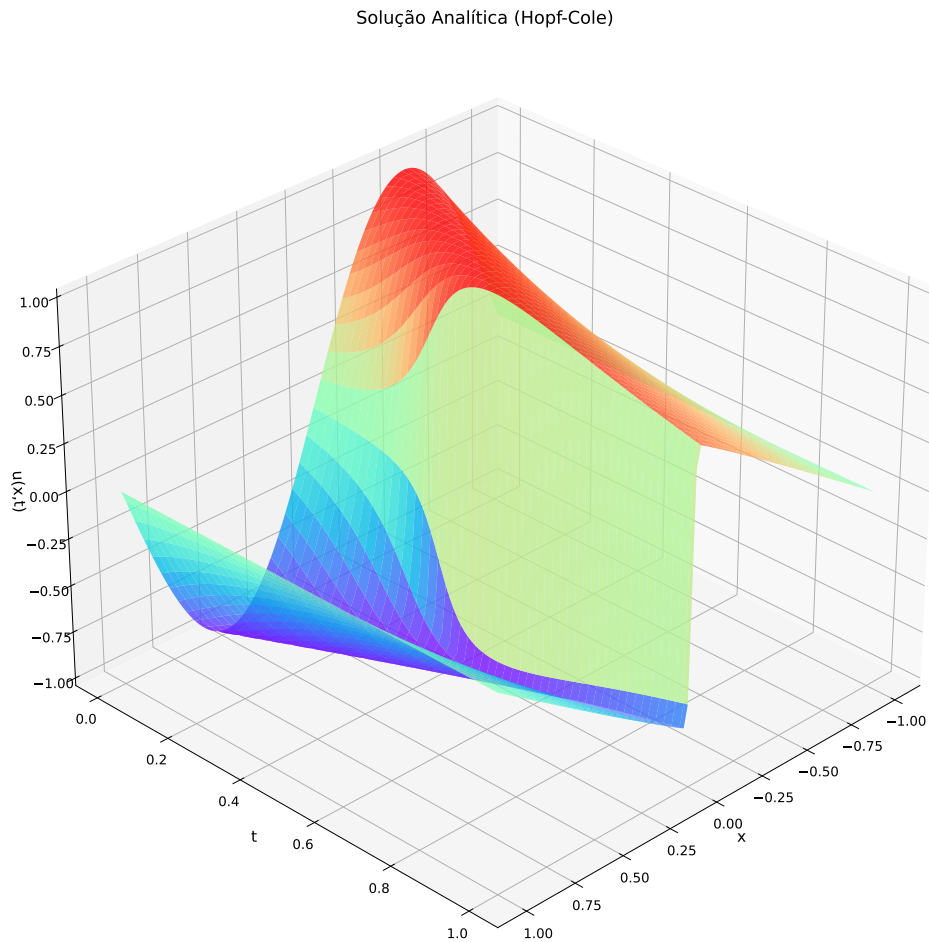
Os resultados foram organizados em cinco partes: (5.1) visualização direta das soluções dos três métodos; (5.2) análise comparativa pontual no instante $t = 0,5$; (5.3) avaliação espaço–temporal dos erros; (5.4) análise estatística da distribuição dos erros (5.5) indicadores globais de desempenho e estabilidade; (5.6) comportamento do erro RMS ao longo do tempo entre os métodos avaliados; e (5.7) histórico da função de perda do treinamento da PINN.

5.1 Visualização das Soluções

5.1.1 Solução Analítica

A Figura 1 apresenta a superfície tridimensional da solução analítica, construída a partir da transformação de Cole–Hopf. Essa solução serve como referência para todas as comparações posteriores. Observa-se que a evolução temporal é caracterizada por uma deformação progressiva da onda inicial, cuja estrutura senoidal, presente no instante inicial $t \approx 0$, suaviza-se gradualmente devido ao efeito difusivo introduzido pela viscosidade $\nu = 0,01/\pi$. À medida que o tempo evolui, o gradiente acentuado que caracteriza a região central do domínio torna-se menos abrupto, evidenciando o balanceamento natural entre os efeitos advectivos e difusivos.

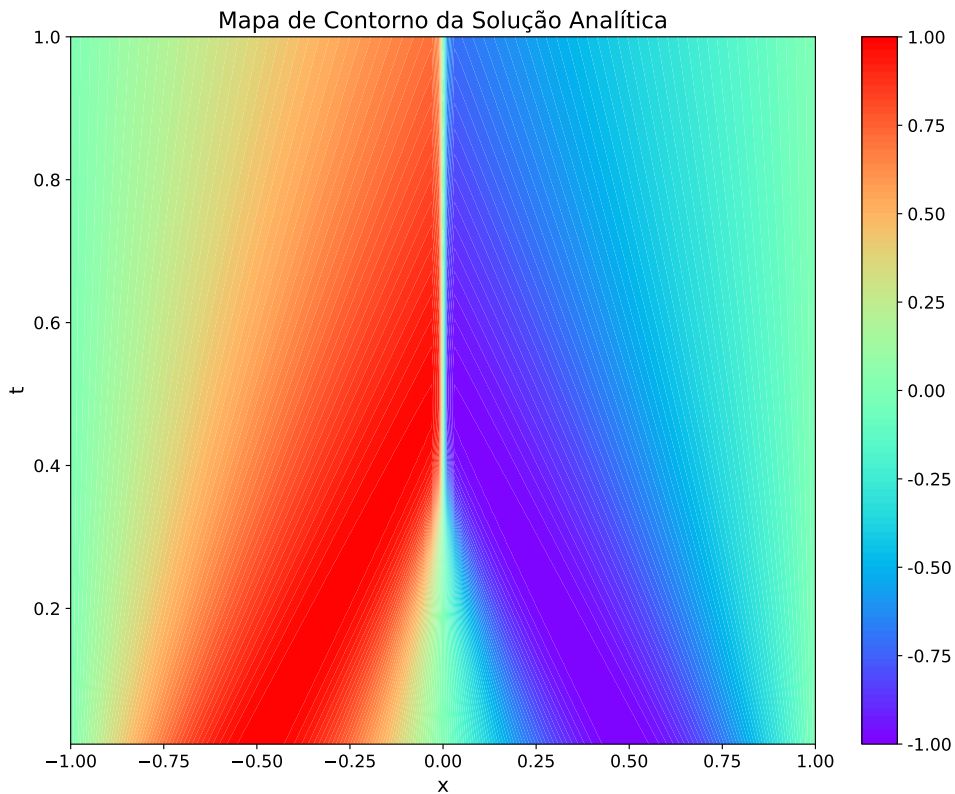
Figura 1 – Superfície tridimensional da solução analítica via transformação de Cole–Hopf, exibindo a dissipação progressiva da inclinação acentuada conforme o tempo avança.



Fonte: Elaborado pelo autor.

O comportamento observado na superfície tridimensional é complementado pelo mapa de contorno apresentado na Figura 2. Nesse plano (x, t) , os níveis de cor mostram claramente a transição suave da solução ao longo do tempo. A região onde ocorre o gradiente mais pronunciado se desloca horizontalmente enquanto é gradualmente dissipado, refletindo o papel da difusão viscosa. A ausência de oscilações espúrias ou padrões irregulares confirma que a solução analítica constitui uma referência estável e suave, típica de equações parabólicas com viscosidade não nula.

Figura 2 – Mapa de contorno da solução analítica, evidenciando a suavidade da evolução espaço-temporal e a dissipação progressiva do gradiente acentuado.



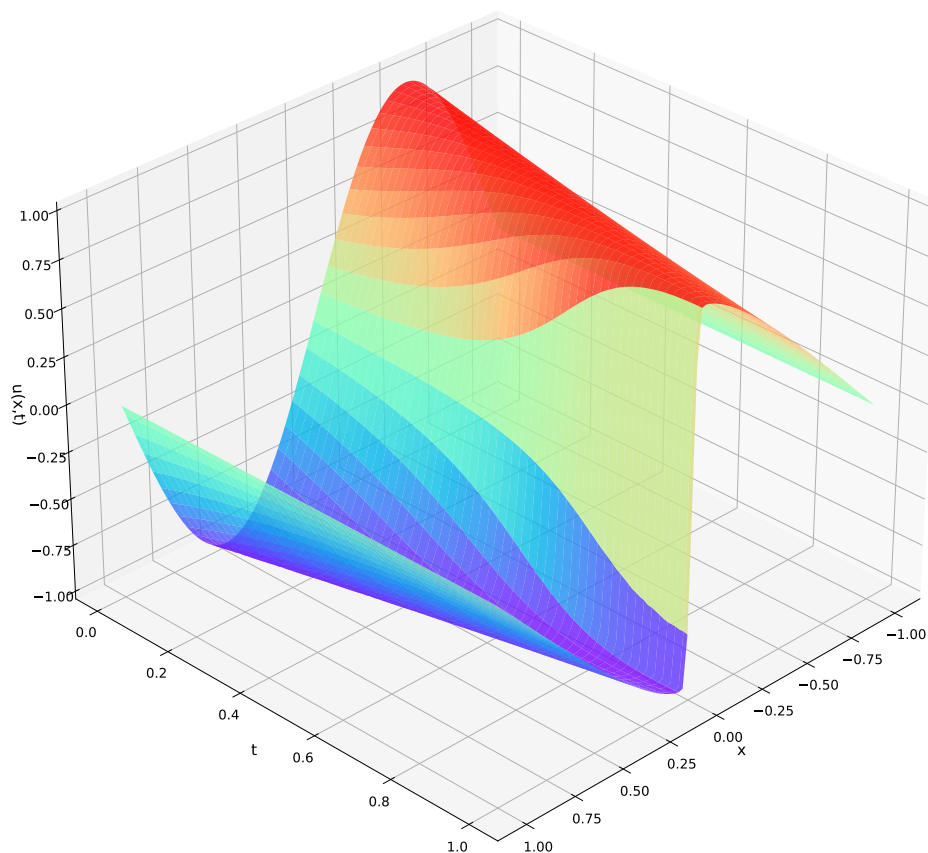
Fonte: Elaborado pelo autor.

5.1.2 Solução Numérica (Crank–Nicolson)

A Figura 3 apresenta a solução obtida numericamente pelo método de Crank–Nicolson, com discretização fina em espaço e tempo. De forma geral, o comportamento da solução acompanha bem o padrão analítico. Contudo, ao observar a superfície com maior atenção, nota-se que a região de maior inclinação, aquela associada à formação do choque viscoso, apresenta uma leve distorção espacial quando comparada com a solução analítica. Essa diferença é esperada, uma vez que o termo advectivo é tratado explicitamente no esquema adotado, o que introduz pequenas defasagens na posição da estrutura onde o gradiente é mais intenso.

Figura 3 – Superfície tridimensional da solução numérica obtida pelo método de Crank–Nicolson, mostrando boa concordância global com a solução analítica, porém com leve deslocamento na região de maior gradiente.

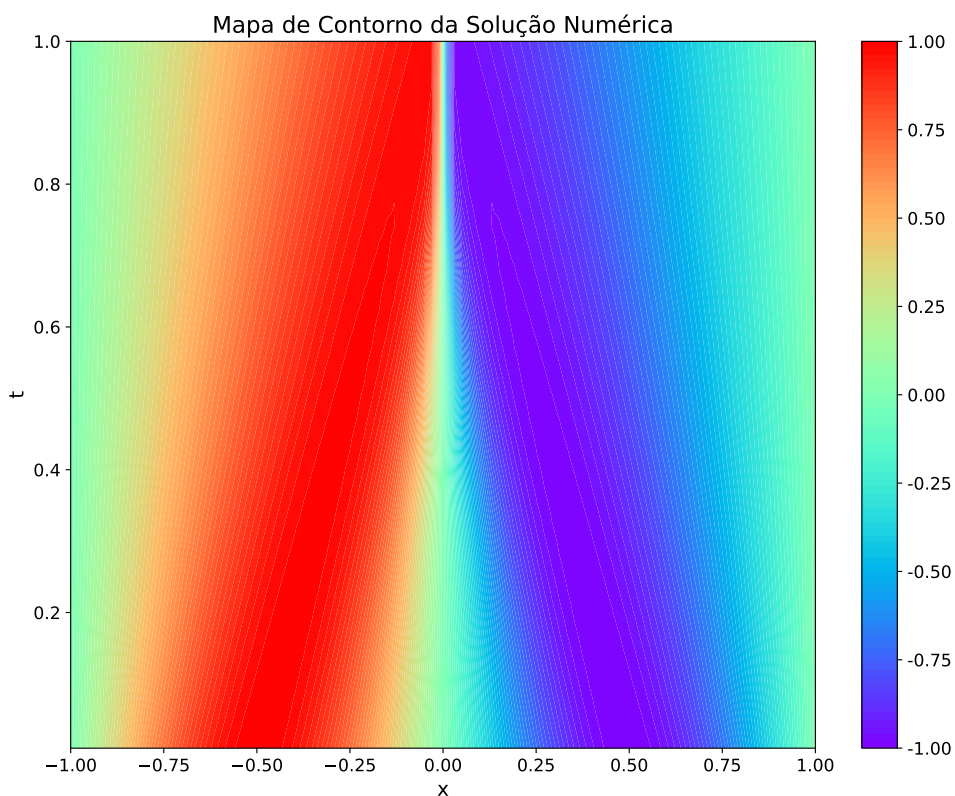
Solução Numérica (FDM)



Fonte: Elaborado pelo autor.

Essas pequenas discrepâncias se tornam mais evidentes ao analisar o mapa de contorno apresentado na Figura 4. O campo de cores revela que, embora o comportamento qualitativo da solução seja reproduzido de modo adequado, o método numérico apresenta uma ligeira tendência de deslocar o ponto de máxima inclinação para a direita, especialmente em tempos intermediários. Além disso, observa-se que as transições entre níveis de cor são um pouco menos suaves do que na solução analítica, refletindo os efeitos cumulativos da discretização temporal e espacial.

Figura 4 – Mapa de contorno da solução numérica obtida pelo método de Crank–Nicolson. A transição espaço–temporal segue o padrão analítico, mas evidencia um pequeno deslocamento do gradiente mais intenso.

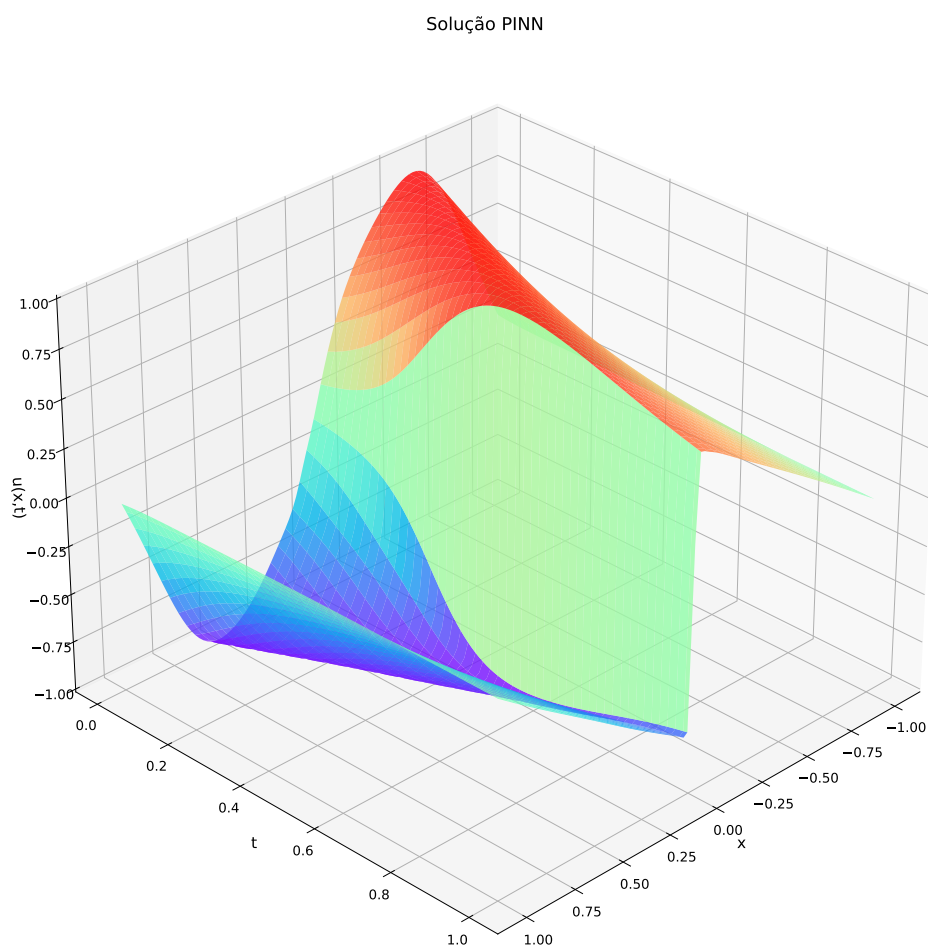


Fonte: Elaborado pelo autor.

5.1.3 Solução via PINNs

A solução obtida pela PINN é mostrada na Figura 5. Uma característica marcante dessa superfície é a elevada fidelidade em relação à solução analítica. A curvatura da solução, a localização da região de maior gradiente e o comportamento difusivo ao longo do tempo são reproduzidos com alta precisão. A suavidade intrínseca das funções aproximadas pela rede neural produz uma superfície que, visualmente, se aproxima mais da solução analítica do que da obtida pelo método de Crank–Nicolson.

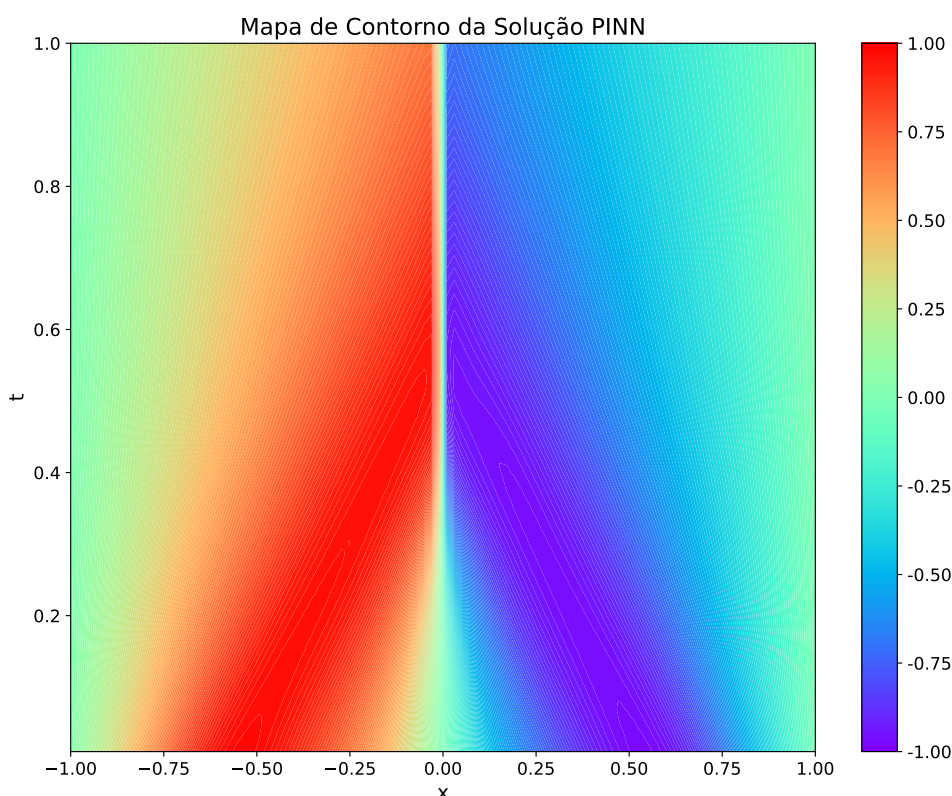
Figura 5 – Superfície tridimensional da solução obtida pela PINN, apresentando excelente acordo com a solução analítica em todo o domínio espaço-temporal.



Fonte: Elaborado pelo autor.

O mapa de contorno correspondente, mostrado na Figura 6, reforça essa observação. As linhas de nível são contínuas e livres de flutuações não físicas, indicando que a rede neural aprendeu adequadamente tanto a solução quanto sua estrutura espaço-temporal. O posicionamento da região de maior inclinação coincide com o da solução analítica, sem o deslocamento observado na solução numérica tradicional.

Figura 6 – Mapa de contorno da solução via PINN, exibindo níveis de cor extremamente regulares e elevada proximidade com o comportamento analítico.



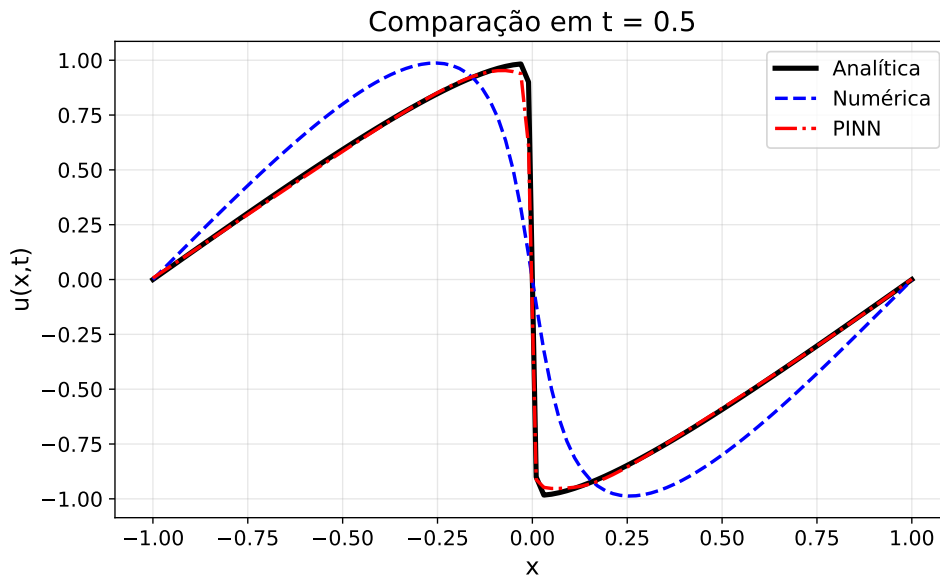
Fonte: Elaborado pelo autor.

5.2 Comparação Pontual no Instante $t = 0,5$

A Figura 7 apresenta as três soluções avaliadas no instante $t = 0,5$, permitindo uma inspeção detalhada da capacidade de cada método em capturar a estrutura da solução na região de maior gradiente. A solução analítica, representada pela curva preta contínua, exibe uma queda abrupta na vizinhança de $x = 0$, característica do choque viscoso em formação. A solução numérica (linha azul tracejada) apresenta uma discrepância perceptível, deslocando ligeiramente a posição do ponto de maior variação e introduzindo uma leve suavização excessiva antes e depois do choque. Esse comportamento é típico de esquemas numéricos que tratam parcialmente o termo advectivo de forma explícita.

Por outro lado, a solução obtida pela PINN (linha vermelha tracejada) praticamente coincide com a solução analítica ao longo de todo o perfil. O choque é reproduzido tanto em posição quanto em intensidade, e a transição suave nas regiões onde o gradiente é menor é capturada com alta precisão. A proximidade entre as curvas analítica e neural indica que a PINN foi capaz de aprender corretamente a estrutura da solução, mesmo na presença de gradientes íngremes, onde redes neurais frequentemente apresentam dificuldades.

Figura 7 – Comparação entre as soluções analítica, numérica e via PINN no instante $t = 0,5$.
 Observa-se a superior precisão da solução PINN na captura da região de maior gradiente.

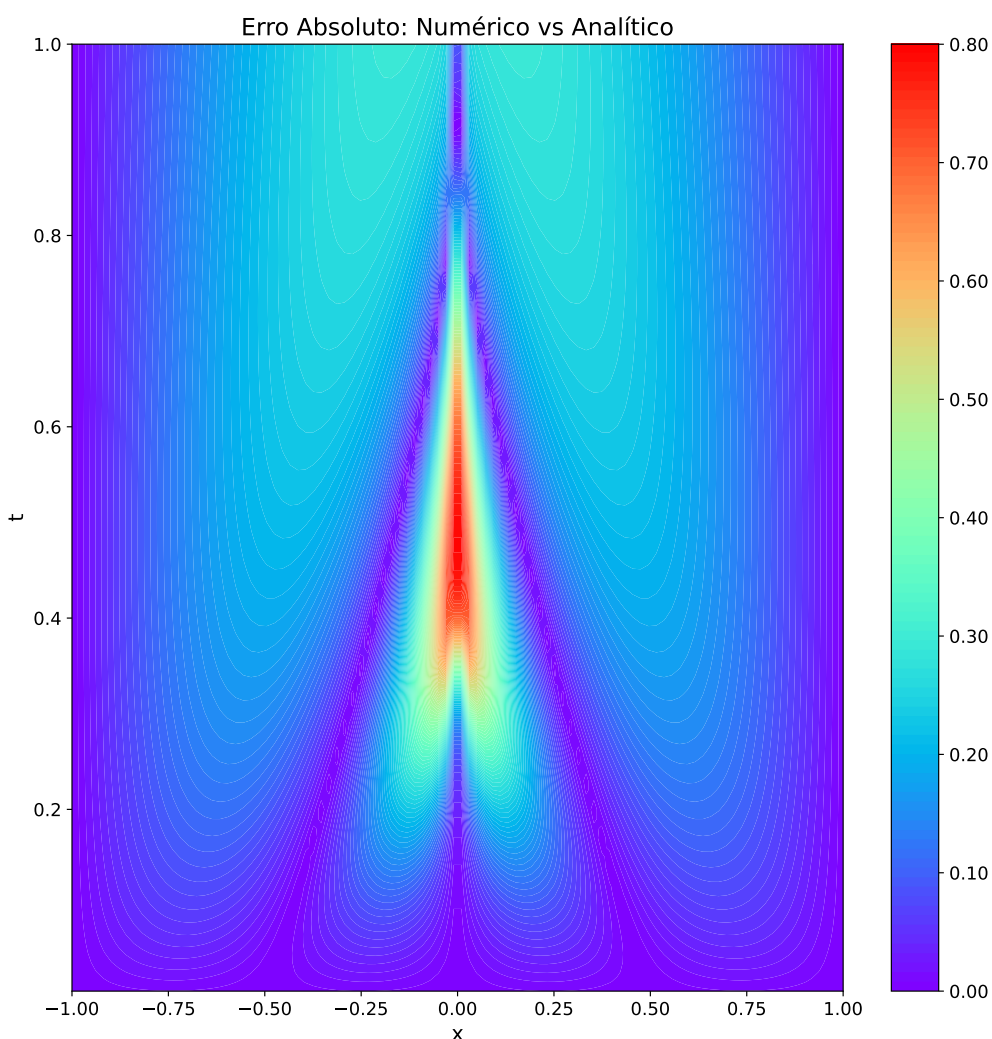


Fonte: Elaborado pelo autor.

5.3 Mapas Espaço–Temporais de Erro

A avaliação comparativa das soluções é aprofundada por meio da análise do erro absoluto ao longo de todo o domínio espaço–temporal (x, t) . Esses mapas permitem identificar regiões em que as discrepâncias entre os métodos se tornam mais pronunciadas e fornecem uma compreensão mais clara da capacidade de cada técnica em reproduzir a solução de referência. A Figura 8 apresenta o erro entre a solução numérica e a solução analítica. Observa-se que as maiores diferenças se concentram na vizinhança da região em que o gradiente da solução é mais acentuado. Esse comportamento decorre do fato de que o método de Crank–Nicolson, na forma implementada, trata o termo advectivo explicitamente, criando uma defasagem temporal que se manifesta justamente quando a não linearidade induz uma variação espacial intensa. À medida que a evolução temporal progride, esse desalinhamento tende a se acumular, produzindo um erro mais significativo, especialmente em tempos intermediários.

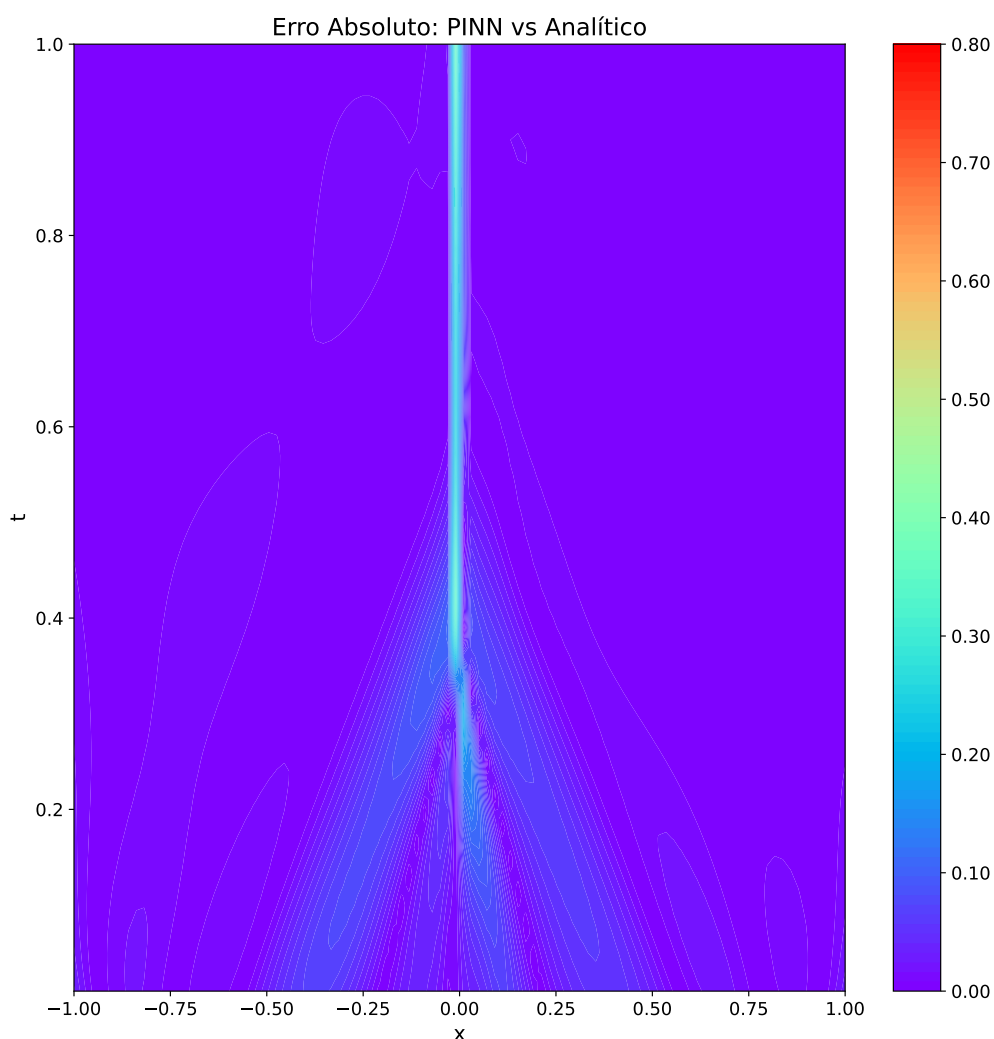
Figura 8 – Erro absoluto entre a solução numérica e a solução analítica. As maiores discrepâncias concentram-se na região de maior variação espacial da solução.



Fonte: Elaborado pelo autor.

A Figura 9 mostra o erro absoluto entre a solução analítica e a solução obtida pela PINN. Nota-se que as discrepâncias são significativamente menores do que aquelas observadas no método numérico, distribuindo-se de forma mais homogênea ao longo do domínio. Embora a região de maior inclinação apresente erros mais elevados, como é natural em soluções que contêm gradientes acentuados, sua magnitude permanece muito inferior à registrada pela solução numérica. Esse comportamento evidencia que a PINN foi capaz de capturar de forma robusta a estrutura global da solução e reconstruir com precisão a posição e a intensidade do choque viscoso.

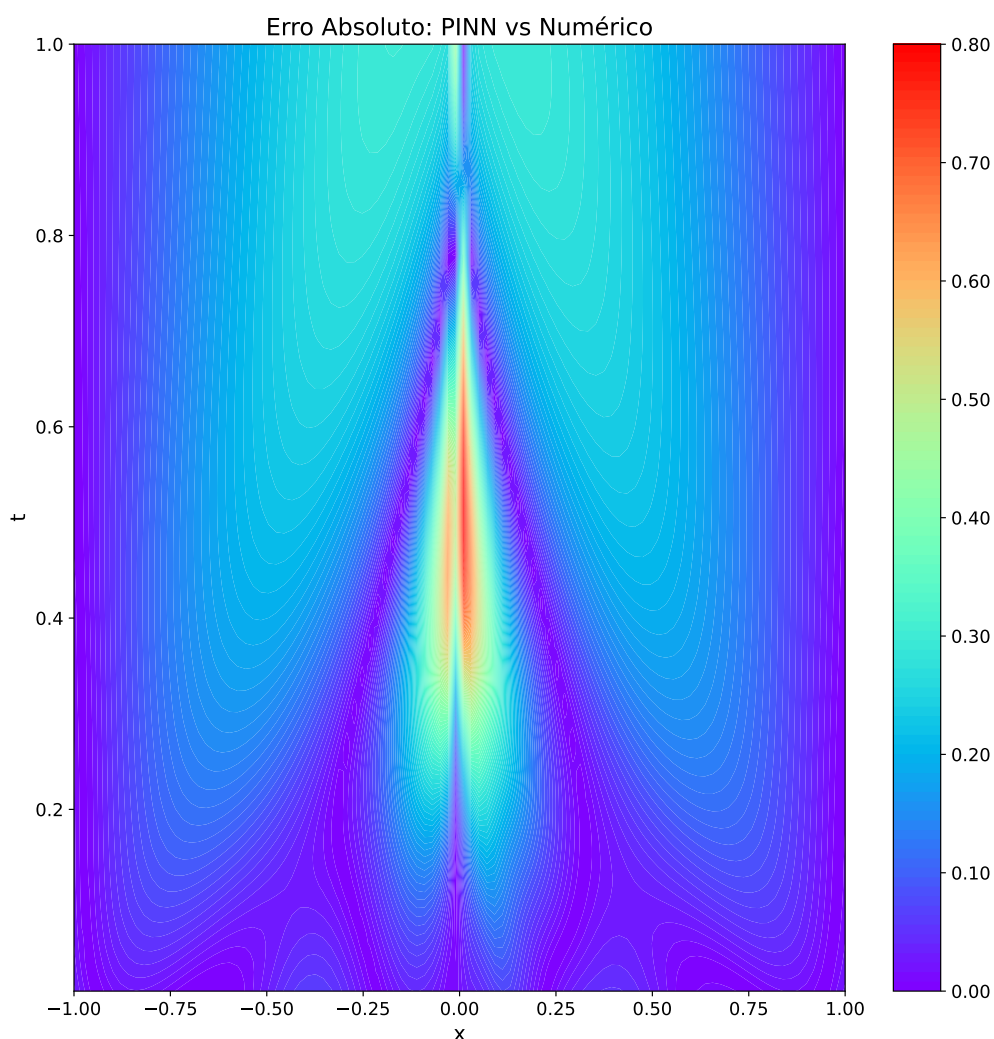
Figura 9 – Erro absoluto entre a solução via PINN e a solução analítica. As discrepâncias são reduzidas e apresentam distribuição homogênea.



Fonte: Elaborado pelo autor.

A Figura 10 apresenta o erro entre a solução via PINN e a solução numérica. Esse mapa revela que as regiões onde o método numérico se afasta mais intensamente da solução analítica correspondem aos locais em que o erro PINN–numérica aumenta. Isso indica que a PINN preserva de forma mais rigorosa as características da solução de referência, enquanto a solução numérica apresenta defasagens decorrentes do tratamento semiexplícito da advecção. Assim, a comparação entre PINN e o método numérico evidencia a capacidade da rede neural de manter maior coerência com a solução analítica, sobretudo em tempos em que a inclinação da solução se torna mais abrupta.

Figura 10 – Erro absoluto entre a solução via PINN e a solução numérica. As discrepâncias mais intensas coincidem com regiões onde o método numérico se afasta da solução analítica.

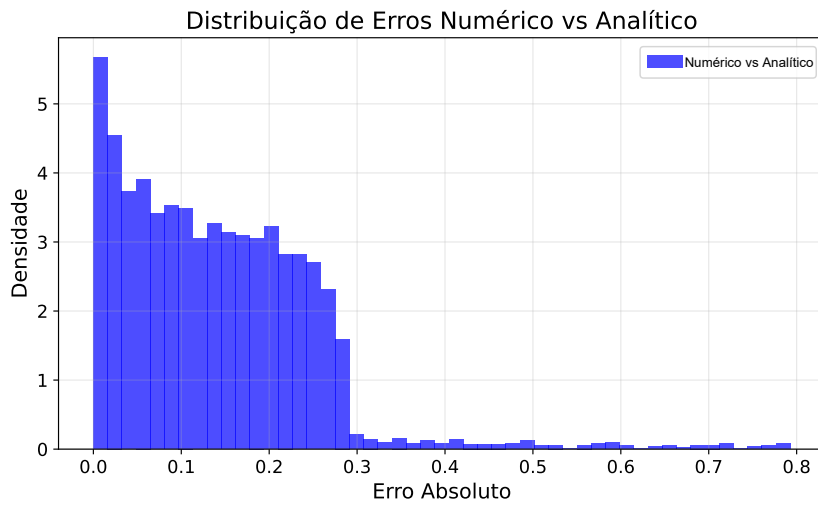


Fonte: Elaborado pelo autor.

5.4 Distribuição dos Erros

A análise estatística das discrepâncias complementa a avaliação espaço-temporal, permitindo caracterizar a frequência e a intensidade com que os erros ocorrem ao longo do domínio. A Figura 11 apresenta a distribuição do erro entre a solução numérica e a solução analítica. Observa-se que a maior parte dos valores concentra-se em uma faixa moderada de erro, mas uma cauda longa está presente devido à dificuldade do método numérico em reproduzir com precisão a região de formação do choque viscoso. Essa cauda está associada à natureza não linear da solução e às limitações inerentes à discretização utilizada.

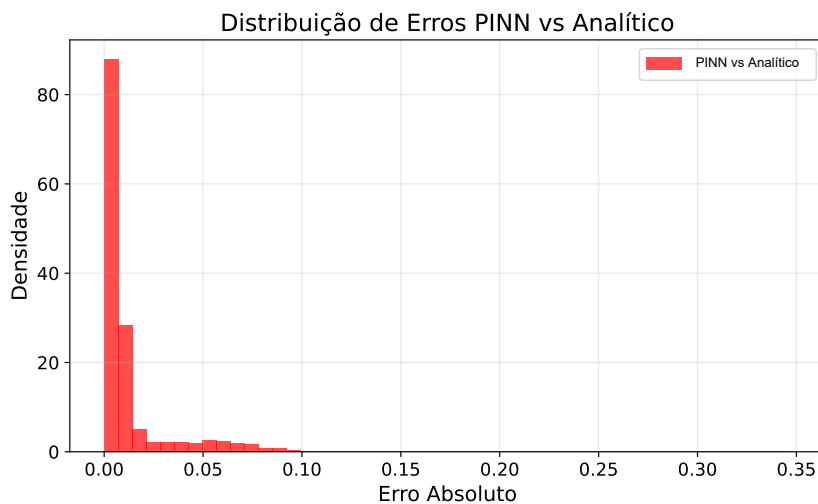
Figura 11 – Distribuição dos erros entre a solução numérica e a solução analítica. A presença de cauda longa reflete a sensibilidade do método numérico à região de maior gradiente.



Fonte: Elaborado pelo autor.

A Figura 12 apresenta a distribuição dos erros entre a PINN e a solução analítica. A concentração dos valores em torno de zero é marcante, com dispersão muito inferior àquela observada no método numérico. Esse comportamento indica que a PINN reproduziu a solução analítica de forma precisa em praticamente todo o domínio, mesmo na presença de gradientes acentuados. A ausência de uma cauda longa revela ainda que a rede neural evitou discrepâncias pontuais significativas, refletindo sua capacidade de internalizar a estrutura da equação diferencial.

Figura 12 – Distribuição dos erros entre a solução via PINN e a solução analítica. A elevada concentração em torno de zero evidencia a precisão da abordagem neural.

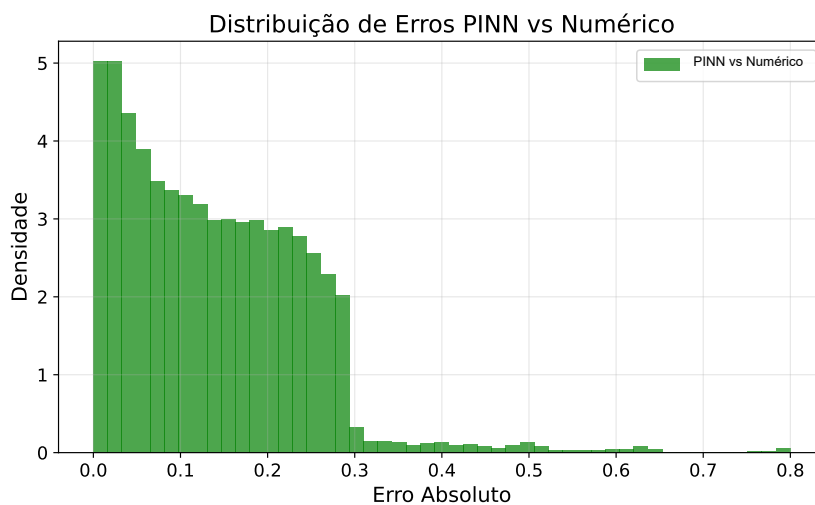


Fonte: Elaborado pelo autor.

A Figura 13 apresenta a distribuição dos erros entre a PINN e a solução numérica. O padrão observado revela que a maior parte dos erros encontra-se em uma faixa estreita, destacando que a PINN preserva mais fielmente o comportamento da solução analítica. As discrepâncias mais pronunciadas, que surgem na comparação numérica, praticamente desaparecem aqui, indicando que a rede neural

reduz significativamente as limitações introduzidas pelo método de diferenças finitas, sobretudo na região de transição rápida da solução.

Figura 13 – Distribuição dos erros entre a solução via PINN e a solução numérica. A faixa estreita de dispersão confirma a maior fidelidade da PINN ao comportamento analítico.



Fonte: Elaborado pelo autor.

5.5 Estatísticas Globais de Erro

A Figura 14 apresenta as estatísticas globais de erro médio, erro máximo e erro RMS para os três pares de comparação. Os valores obtidos corroboram as análises anteriores: o método numérico apresenta os erros mais elevados, com erro máximo significativamente mais alto devido à presença de regiões de gradiente intenso. A solução via PINN, por sua vez, apresenta erros de ordem muito inferior, tanto em termos médios quanto RMS. Esse último indicador é particularmente relevante no contexto da equação de Burgers, pois sintetiza a precisão global da solução em um problema cuja estrutura envolve, simultaneamente, regiões de forte variação e trechos suavemente difusivos.

Figura 14 – Estatísticas globais dos erros médio, máximo e RMS para os três pares de comparação. A solução via PINN apresenta desempenho significativamente superior.

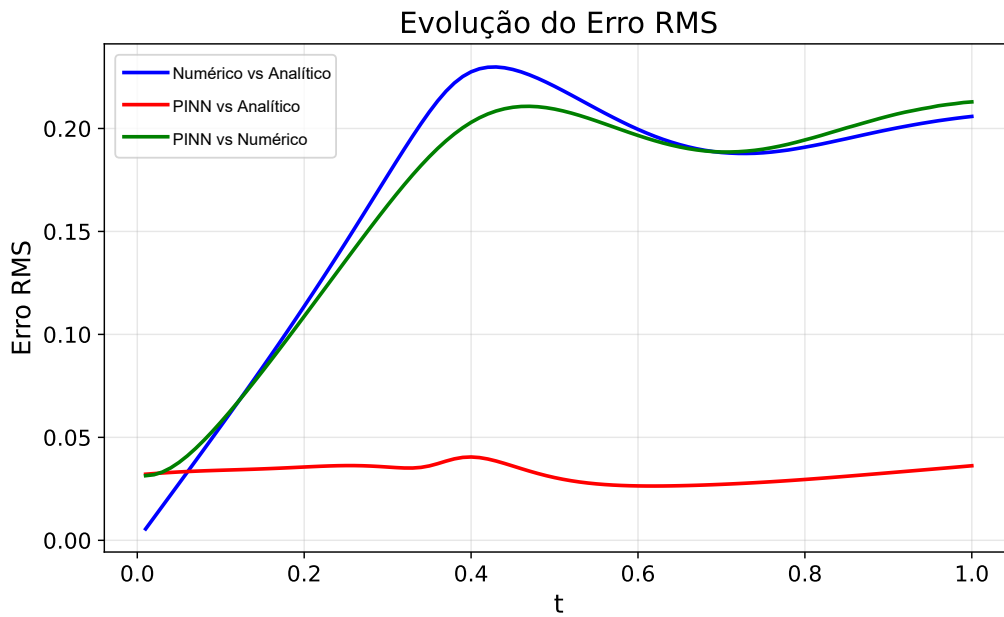
ESTATÍSTICAS DE ERRO	
Numérico vs Analítico:	
RMS:	1.80e-01
Médio:	1.41e-01
Máximo:	7.93e-01
PINN vs Analítico:	
RMS:	3.28e-02
Médio:	1.36e-02
Máximo:	3.47e-01
PINN vs Numérico:	
RMS:	1.75e-01
Médio:	1.39e-01
Máximo:	8.00e-01

Fonte: Elaborado pelo autor.

5.6 Evolução Temporal do Erro RMS

A Figura 15 apresenta a evolução temporal do erro RMS ao longo de todo o intervalo analisado. A curva referente ao método numérico apresenta crescimento progressivo à medida que o tempo avança, refletindo a dificuldade da discretização semiexplícita em capturar com precisão a região de maior variação espacial, particularmente na proximidade do choque viscoso. Por outro lado, a curva associada à PINN mantém-se baixa e praticamente constante durante todo o intervalo, evidenciando que a rede neural preserva a coerência com a solução analítica inclusive em instantes em que o gradiente da solução se torna mais abrupto.

Figura 15 – Evolução temporal do erro RMS. A solução via PINN mantém erro reduzido e estável ao longo do tempo.

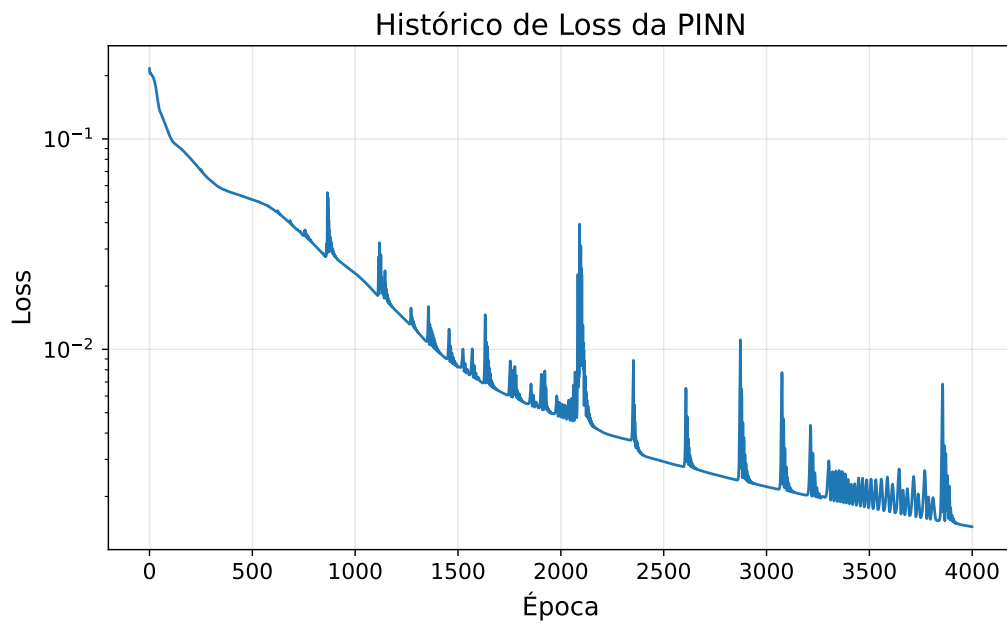


Fonte: Elaborado pelo autor.

5.7 Convergência do Treinamento da PINN

A Figura 16 apresenta o histórico da função de perda ao longo das épocas de treinamento. A curva exibe tendência consistentemente decrescente, acompanhada de oscilações de baixa amplitude, típicas de modelos treinados com o otimizador Adam. Conforme o treinamento avança, a função de perda se estabiliza, indicando que a rede neural foi capaz de minimizar simultaneamente os termos associados às condições iniciais, condições de contorno e ao residual da equação de Burgers. A convergência observada demonstra que a PINN assimilou adequadamente a estrutura física do problema e alcançou um regime de estabilidade numérica e funcional.

Figura 16 – Histórico da função de perda durante o treinamento da PINN. A tendência global aponta para convergência estável do modelo.



Fonte: Elaborado pelo autor.

6 CONCLUSÃO

A equação de Burgers, apesar de sua formulação compacta, constitui um modelo de grande relevância para o estudo de fenômenos envolvendo advecção não linear e difusão viscosa. Neste trabalho, três abordagens distintas foram empregadas para resolver essa equação: a solução analítica obtida pela transformação de Cole–Hopf, a solução numérica baseada no método semiexplícito de Crank–Nicolson e a solução aproximada fornecida por meio de redes neurais informadas pela física (PINNs). A comparação sistemática entre essas abordagens permitiu avaliar, de forma abrangente, suas capacidades, limitações e coerência com o comportamento físico esperado.

A solução analítica desempenhou um papel central como referência para a validação das demais técnicas. Sua estabilidade e precisão permitiram identificar, com clareza, as discrepâncias introduzidas pelos métodos alternativos. Os resultados mostraram que o método numérico de Crank–Nicolson, embora estável e amplamente utilizado em problemas de difusão, apresenta limitações quando aplicado a uma equação com forte não linearidade. Em particular, observou-se que a região de gradiente acentuado — associada à formação do choque viscoso — é mais sensível à discretização adotada, produzindo erros localizados e um leve deslocamento da solução quando comparada à referência analítica.

Por outro lado, a abordagem via PINNs demonstrou um desempenho notável. A rede neural foi capaz de internalizar a estrutura física do problema, respeitando simultaneamente o residual da equação diferencial, bem como as condições de contorno e iniciais. A solução obtida apresentou excelente concordância com a solução analítica, tanto qualitativa quanto quantitativamente. O erro absoluto manteve-se reduzido e bem distribuído ao longo de todo o domínio, enquanto o erro RMS permaneceu praticamente constante durante a evolução temporal. A distribuição estatística dos erros reforçou esse comportamento ao revelar valores concentrados em torno de zero e a ausência de caudas longas, indicando que a PINN capturou com precisão a dinâmica espaço–temporal da equação.

Apesar de seu desempenho superior, a utilização de PINNs envolve custos computacionais elevados, sobretudo durante o processo de treinamento, que pode demandar milhares de épocas. Além disso, o desempenho final pode depender de escolhas específicas, como a arquitetura da rede, a função de ativação, a estratégia de amostragem dos pontos de colocação e os hiperparâmetros da otimização. Embora os resultados obtidos neste trabalho tenham sido consistentes e estáveis, a robustez das PINNs frente a diferentes regimes ou classes de equações continua sendo um tema de intensa pesquisa.

Em síntese, os resultados obtidos demonstram que a abordagem via PINNs apresenta desempenho superior ao método numérico tradicional no problema estudado, reproduzindo com alta fidelidade a estrutura espaço–temporal da solução analítica, incluindo a região de choque viscoso. A aplicação conjunta dos três métodos permitiu caracterizar de forma ampla os fenômenos envolvidos e destacou a importância de integrar técnicas clássicas e modernas no contexto da modelagem matemática e computacional.

Como continuidade natural deste trabalho, sugere-se investigar diferentes arquiteturas de redes neurais, explorar estratégias de ponderação adaptativa dos termos da função de perda, incorporar

técnicas de regularização e estender a análise para versões bidimensionais e tridimensionais da equação de Burgers. Estudos envolvendo valores de viscosidade ainda menores, bem como comparações com métodos numéricos de alta ordem, também representam caminhos promissores para pesquisas futuras.

REFERÊNCIAS

- BENTON, E. R.; PLATZMAN, G. W. A table of solutions of the one-dimensional burgers equation. **Quarterly of Applied Mathematics**, v. 30, n. 2, p. 195–212, 1972.
- BURGERS, J. M. A mathematical model illustrating the theory of turbulence. **Advances in Applied Mechanics**, v. 1, p. 171–199, 1948.
- COLE, J. D. On a quasi-linear parabolic equation occurring in aerodynamics. **Quarterly of Applied Mathematics**, v. 9, p. 225–236, 1951.
- EL-DANAF, T. S.; RAMADAN, M. A. On the analytical and numerical solutions of the one-dimensional nonlinear burgers' equation. **The Open Applied Mathematics Journal**, v. 1, p. 1–8, 2007.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. Cambridge, MA: MIT Press, 2016.
- HASSANIEN, I. A.; SALAMA, A. A.; HOSHAM, H. A. Fourth-order finite difference method for solving burgers' equation. **Applied Mathematics and Computation**, v. 170, p. 781–800, 2005.
- HAYKIN, S. **Neural Networks: A Comprehensive Foundation**. 2. ed. Upper Saddle River: Prentice Hall, 1999.
- HOPF, E. The partial differential equation $u_t + uu_x = \nu u_{xx}$. **Communications on Pure and Applied Mathematics**, v. 3, n. 3, p. 201–230, 1950.
- HORNIK, K.; STINCHCOMBE, M.; WHITE, H. Multilayer feedforward networks are universal approximators. **Neural Networks**, v. 2, n. 5, p. 359–366, 1989.
- KUTLUAY, S.; BAHADIR, A. R.; ÖZDEŞ, A. Numerical solution of the burgers' equation by an explicit finite difference method. **Journal of Computational and Applied Mathematics**, v. 103, p. 251–261, 1999.
- LEVEQUE, R. J. **Finite Volume Methods for Hyperbolic Problems**. [S.l.]: Cambridge University Press, 2002.
- MCKAY, M. D.; BECKMAN, R. J.; CONOVER, W. J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. **Technometrics**, v. 21, n. 2, p. 239–245, 1979.
- RAISSI, M.; PERDIKARIS, P.; KARNIADAKIS, G. E. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. **Journal of Computational Physics**, v. 378, p. 686–707, 2019.
- SAVOVIĆ, S.; IVANOVIĆ, M.; MIN, R. A comparative study of the explicit finite difference method and physics-informed neural networks for solving the burgers' equation. **Axioms**, v. 12, n. 10, p. 982, 2023.
- WHITHAM, G. B. **Linear and Nonlinear Waves**. New York: John Wiley & Sons, 1974.