



UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

**DIVISIÓN ACADÉMICA DE CIENCIAS Y TECNOLOGÍAS DE LA
INFORMACIÓN**

**SOLUCIONES FUNDAMENTALES DE LA ECUACIÓN DE ONDA EN
COORDENADAS CILÍNDRICAS UTILIZANDO REDES NEURONALES**

**TESIS PARA OBTENER EL GRADO DE:
MAESTRÍA EN CIENCIAS DE LA COMPUTACIÓN**

PRESENTA:

GABRIEL OMAR DOMINGUEZ RUIZ

BAJO LA DIRECCIÓN DE:

DR. JOSÉ ADÁN HERNÁNDEZ NOLASCO

CUNDUACÁN, TABASCO, A: JULIO 2026



UNIVERSIDAD JUÁREZ AUTÓNOMA DE TABASCO

**DIVISIÓN ACADÉMICA DE CIENCIAS Y TECNOLOGÍAS DE LA
INFORMACIÓN**

**SOLUCIONES FUNDAMENTALES DE LA ECUACIÓN DE ONDA EN
COORDENADAS CILÍNDRICAS UTILIZANDO REDES NEURONALES**

**TESIS PARA OBTENER EL GRADO DE:
MAESTRÍA EN CIENCIAS DE LA COMPUTACIÓN**

PRESENTA:

GABRIEL OMAR DOMINGUEZ RUIZ

BAJO LA DIRECCIÓN DE:

DR. JOSÉ ADÁN HERNÁNDEZ NOLASCO

CUNDUACÁN, TABASCO, A: JULIO 2026

Declaración de Autoría y Originalidad

En la Ciudad de Cunduacán el día nueve del mes de julio del año 2026, el que suscribe **Gabriel Omar Dominguez Ruiz**, alumno del Programa de la **Maestría en Ciencias de la Computación** con número de matrícula **242H21002**, adscrito a la **División Académica de Ciencias y Tecnologías de la Información**, de la Universidad Juárez Autónoma de Tabasco, como autor de la Tesis presentada para la obtención de Grado de Maestro en Ciencias de la Computación y titulada **Soluciones fundamentales de la ecuación de onda en coordenadas cilíndricas utilizando Redes Neuronales**, dirigida por el Dr. José Adán Hernández Nolasco.

DECLARO QUE: La Tesis es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, de acuerdo con el ordenamiento jurídico vigente, en particular, la LEY FEDERAL DEL DERECHO DE AUTOR (Decreto por el que se reforman y adicionan diversas disposiciones de la Ley Federal del Derecho de Autor del 01 de Julio de 2020 regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), en particular, las disposiciones referidas al derecho de cita. Del mismo modo, asumo frente a la Universidad cualquier responsabilidad que pudiera derivarse de la autoría o falta de originalidad o contenido de la Tesis presentada de conformidad con el ordenamiento jurídico vigente.

Cunduacán, Tabasco a 9 de julio de 2026.



Estudiante: Gabriel Omar Dominguez Ruiz



UJAT

UNIVERSIDAD JUÁREZ
AUTÓNOMA DE TABASCO

"ESTUDIO EN LA DUDA, ACCIÓN EN LA FE"



DIVISIÓN ACADÉMICA DE
CIENCIAS Y TECNOLOGÍAS
DE LA INFORMACIÓN



2026
año de
Margarita
Maza

Cunduacán, Tabasco, a 09 de julio de 2026
Oficio No. 1467/2026/DACYTI/D

Asunto: Autorización de impresión de Tesis

C. Gabriel Omar Domínguez Ruiz

Egresado de la Maestría en Ciencias de la Computación

En virtud de que cumple satisfactoriamente los requisitos establecidos en el Reglamento General de Estudios de Posgrado vigente en la Universidad, informo a Usted que se autoriza la impresión del trabajo recepcional **"Soluciones fundamentales de la ecuación de onda en coordenadas cilíndricas utilizando redes neuronales"**, para presentar examen y obtener el Grado de Maestro en Ciencias de la Computación.

Sin otro particular, aprovecho la ocasión para enviarle un afectuoso saludo.

Atentamente

Dra. Laura Beatriz Vidal Turrubiates
Directora

C.c. p. Dr. Eduardo Hernández de la Cruz. – Encargado del despacho de la Coordinación de Posgrado.
Archivo

DRA.*LBVT/EHC

Av. Universidad s/n, Zona de la Cultura, Col. Magisterial,
Villahermosa, Centro, Tabasco, Mex. C.P. 86040.
Tel (993) 358 15 00 e-Mail: rectoria@ujat.mx

Carta de Cesión de Derechos

Villahermosa, Tabasco a 9 de julio de 2026.

Por medio de la presente manifiesto haber colaborado como AUTOR en la producción, creación y/o realización de la obra denominada: **Soluciones fundamentales de la ecuación de onda en coordenadas cilíndricas utilizando Redes Neuronales.**

Con fundamento en el artículo 83 de la Ley Federal del Derecho de Autor y toda vez que, la creación y/o realización de la obra antes mencionada se realizó bajo la comisión de la Universidad Juárez Autónoma de Tabasco; entendemos y aceptamos el alcance del artículo en mención de que tenemos el derecho al reconocimiento como autores de la obra, y a la Universidad Juárez Autónoma de Tabasco mantendrá en un 100 % la titularidad de los derechos patrimoniales por un período de 20 años sobre la obra en la que colaboramos, por lo anterior, cedemos el derecho patrimonial exclusivo en favor de la Universidad.

COLABORADOR



Estudiante: Gabriel Omar Dominguez Ruiz

TESTIGOS



Dr. José Adán Hernández Nolasco



Dr. Mariano Madrigal Domínguez

Dedicatoria

A mis padres, que me enseñaron a no soltar lo empezado y me impulsaron incluso cuando el camino que elegí les resultaba difícil de explicar. Lo que estas páginas tienen de constancia lo aprendí en casa, mucho antes de saber qué era una ecuación diferencial.

A mi esposa, que acompañó esta etapa desde dentro y conoció sus dos caras: los resultados que valía la pena mostrar y las semanas en que no había nada que mostrar. Sostuvo la parte que no aparece en ningún capítulo, y me impulsó a seguir cuando a mí ya se me habían agotado las razones.

A Dios, por proveer en cada etapa lo que hacía falta —el tiempo, la salud, la calma y las personas indicadas en el momento indicado— y por sostenerme en los tramos en que el trabajo no parecía llevar a ninguna parte. Cada meta alcanzada aquí pasó primero por sus manos.

Y a mi fiel compañero felino, que acompañó a su manera: echado sobre el escritorio durante horas interminables, sin entender una sola línea de todo esto, pero sabiendo siempre cuándo su compañía era justo lo que se necesitaba.

Agradecimientos

A Dios, en primer lugar, por proveer en cada etapa lo necesario para llegar hasta aquí.

A mi director de tesis, el Dr. José Adán Hernández Nolasco, por la libertad con que me dejó explorar este problema y por la exigencia con que me hizo sostenerlo.

A la Universidad Juárez Autónoma de Tabasco y a la División Académica de Ciencias y Tecnologías de la Información, por la formación y los medios; y a la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI), antes CONACYT, por la beca de posgrado que hizo posible dedicarme de tiempo completo a esta investigación.

A los integrantes del jurado revisor, el Dr. Pablo Pancardo García, el Dr. Miguel Antonio Wistern Ovando y el Dr. Noel Zacarías Morales, por observaciones que mejoraron de fondo este documento.

A mis padres, a mi esposa y a mi compañero felino, por el sostén de todos los días.

Tabla de contenido

Dedicatoria	I
Agradecimientos	II
Índice de figuras	VII
Índice de tablas	VIII
Resumen	IX
Abstract	X
1. Protocolo de tesis	1
1.1. Introducción	1
1.2. Marco teórico	3
1.2.1. Ecuaciones diferenciales	3
1.2.1.1. Ecuaciones diferenciales ordinarias	3
1.2.1.2. Ecuaciones diferenciales parciales	4
1.2.2. Métodos numéricos	4
1.2.2.1. Método de diferencias finitas	4
1.2.2.2. Método de Euler	4
1.2.3. Inteligencia artificial	5
1.2.3.1. Redes neuronales	5
1.2.4. Ecuación de onda	5
1.2.4.1. Sistema de coordenadas cilíndricas	6

1.2.4.2. Ecuación y función de Bessel	6
1.2.4.3. Solución analítica	7
1.2.5. Estado del arte	7
1.2.5.1. Redes neuronales como aproximadores de soluciones de ecuaciones diferenciales	8
1.2.5.2. PINN aplicadas a la ecuación de onda	9
1.2.5.3. Sesgo espectral y escalabilidad: los límites reconocidos	10
1.2.5.4. Síntesis y vacío de conocimiento	10
1.3. Planteamiento del problema	12
1.4. Justificación	14
1.5. Pregunta de investigación	16
1.6. Hipótesis o supuesto de investigación	16
1.7. Objetivo general	16
1.8. Objetivos específicos	16
1.9. Alcances y limitaciones	17
1.9.1. Alcances	17
1.9.2. Limitaciones	18
1.10. Metodología general	18
1.10.1. Comprensión del problema	19
1.10.2. Comprensión de la información física disponible	19
1.10.3. Preparación de los puntos de colocación	20
1.10.4. Modelado	20
1.10.5. Evaluación	20
1.10.6. Despliegue	20
1.11. Contribución y resultados	20
1.12. Cronograma de actividades	21
2. Redes Neuronales Informadas por la Física para resolver la ecuación de onda en coordenadas cilíndricas (funciones de Bessel de primera especie)	26
2.1. Introducción	28

2.2. Trabajo Relacionado	29
2.2.1. Evaluación Numérica de Funciones de Bessel	29
2.2.2. PINNs para Problemas Oscilatorios y Multiescala	29
2.2.3. Posicionamiento de este Trabajo	30
2.3. Fundamentos Teóricos	30
2.3.1. Redes Neuronales Informadas por la Física (PINNs)	30
2.3.2. Spectral Bias y Fourier Features	30
2.3.3. Caso de Estudio: la Ecuación Diferencial de Bessel	31
2.3.4. Muestreo por Ley de Potencias	32
2.4. Metodología	32
2.4.1. Formulación del Problema	32
2.4.2. Desarrollo Progresivo	32
2.4.2.1. E1: MLP Sencillo con PDE y Condiciones de Frontera.	32
2.4.2.2. E2: Mejoras en el Remuestreo y los Puntos de Colocación.	32
2.4.2.3. E3: Fourier Features con Backbone Compartido.	32
2.4.2.4. E4: Relación de Recurrencia como Restricción Física.	34
2.5. Experimentos y Resultados	34
2.5.1. Configuración Experimental	34
2.5.2. Resultados	35
2.5.2.1. E1: Rendimiento del MLP Sencillo.	35
2.5.2.2. E2: Efecto del Remuestreo y los Puntos de Colocación.	36
2.5.2.3. E3: Fourier Features y Backbone Compartido.	37
2.5.2.4. E4: Impacto de la Relación de Recurrencia.	37
2.5.3. Análisis de Sensibilidad de Hiperparámetros	39
2.6. Discusión	41
2.6.1. Progresión de las Estrategias	41
2.6.2. Alcance del Estudio	42
2.7. Conclusiones y Trabajo Futuro	43

3. Conclusiones y recomendaciones generales	46
3.1. Síntesis general de la investigación	46
3.2. Principales hallazgos	47
3.3. Relación del artículo con los objetivos de la tesis	49
3.4. Comparación con los métodos numéricos tradicionales	50
3.5. Aportes teóricos, metodológicos y prácticos	53
3.6. Limitaciones	54
3.7. Trabajo futuro	55
3.8. Recomendaciones	55
Anexo	58
Bibliografía	59

Índice de figuras

2.1. E1: MLP sencillo (3×20 , activación seno) con pérdida $\mathcal{L}_{\text{PDE}} + \mathcal{L}_{\text{BC}}$.	33
2.2. E2: muestreo uniforme (izquierda) frente al muestreo por ley de potencias $\alpha = 2$ (derecha), que concentra puntos cerca del origen.	33
2.3. E3: arquitectura Multi-Head con Fourier Features ($m = 64$) y <i>backbone</i> compartido (4×256 , SiLU).	34
2.4. E4: la relación de recurrencia acopla las cuatro salidas ($\hat{J}_0 \leftrightarrow \hat{J}_1 \leftrightarrow \hat{J}_2 \leftrightarrow \hat{J}_3$) como restricción en la pérdida.	35
2.5. E1: aproximación de J_0 con el MLP sencillo (similaridad 93.8 %); desviaciones visibles en los máximos.	36
2.6. E1: extensión a J_0 (90.9 %) y J_1 (97.1 %) con redes individuales. J_0 pierde precisión respecto al caso individual.	36
2.7. E2: muestreo por ley de potencias. J_0 (98.3 %) y J_1 (98.1 %) mejoran, pero J_2 (34.4 %) se aparta de la solución analítica.	37
2.8. E3: aproximación simultánea con Fourier Features. Similaridades: $J_0 = 97.4 \%$, $J_1 = 94.6 \%$, $J_2 = 85.4 \%$, $J_3 = 60.0 \%$.	38
2.9. E4, configuración comprimida (800 épocas de Adam + 500 de L-BFGS, $N_0 = 2\,000$): aproximación por orden y convergencia de la pérdida (100 % de similaridad en 197 s en una MacBook Air M4).	39
3.1. Extensión a Neumann y Hankel mediante acoplamiento por Wronskiano	49
3.2. Solución numérica de la ecuación de Bessel y dependencia de la malla	52

Índice de tablas

1.1. Comparación de los trabajos revisados según los ejes pertinentes al problema abordado. Fuente: elaboración propia.	11
1.2. Correspondencia entre las etapas de CRISP-DM y su equivalente en el flujo de trabajo con PINN. Fuente: elaboración propia.	19
1.3. Cronograma de actividades de la investigación. Fuente: elaboración propia.	21
2.1. Comparación costo-precisión entre configuraciones de E4 en una MacBook Air M4 (CPU). <i>Sim.</i> es la similaridad media en J_0-J_3 . Las filas multi-semilla reportan $\text{media} \pm \text{desv.}$ sobre las semillas $\{7, 42, 123\}$	39
2.2. Sensibilidad unidimensional del peso de recurrencia λ_{Rec} (con $\lambda_{\text{BC}} = 50$). La entrada $\lambda_{\text{Rec}} = 0$ elimina la pérdida de recurrencia; se reportan las similaridades por orden porque la media queda dominada por el colapso de J_3	40
2.3. Interacción bidimensional $\lambda_{\text{BC}} \times \lambda_{\text{Rec}}$ en las esquinas de un orden de magnitud alrededor de la configuración base. Todas las entradas son la similaridad media en J_0-J_3 en porcentaje.	40
2.4. Sensibilidad de la escala de Fourier σ y del número de Fourier Features m . Ambos se evalúan de forma independiente con los demás hiperparámetros fijos en sus valores por defecto.	40
3.1. Comparación cualitativa entre los métodos numéricos tradicionales y el enfoque PINN para la ecuación de Bessel. Fuente: elaboración propia.	51

Resumen

Las redes neuronales han impulsado los avances recientes del aprendizaje automático y, más allá de sus usos habituales, hoy se exploran como una vía para resolver ecuaciones diferenciales. En esa frontera destacan las Redes Neuronales Informadas por la Física (PINN), que aprenden la solución de una ecuación *sin datos supervisados*: incorporan las leyes físicas directamente en la función de pérdida y entrenan la red para satisfacer a la vez la ecuación y sus condiciones de frontera. Esta tesis lleva dicho enfoque a la ecuación de onda en coordenadas cilíndricas, cuya componente radial conduce a la ecuación de Bessel, un caso exigente por su comportamiento oscilatorio y de solución analítica conocida. El núcleo del trabajo, publicado como artículo científico, propone cuatro estrategias progresivas y no supervisadas que culminan en una arquitectura Multi-Head con *Fourier Features* y la relación de recurrencia como restricción física; con ella, una sola red aproxima los cuatro órdenes J_0 a J_3 con una similaridad del 99.87 % frente a la solución analítica, y una variante ligera alcanza el 99.7 % en 44 segundos con menos de 100 000 parámetros. La investigación va más allá del artículo: extiende el método a las funciones de Neumann y, mediante el Wronskiano, construye las funciones de Hankel —las ondas viajeras que completan la solución fundamental de la ecuación de onda—. Más que una tabla de valores en puntos discretos, la red entrega un modelo continuo y diferenciable, evaluable en todo el dominio: esa propiedad es la que vuelve competitivo el enfoque frente a los métodos numéricos clásicos cuando la solución debe insertarse en cálculos posteriores basados en gradiente.

Palabras clave: Redes Neuronales Informadas por la Física, Funciones de Bessel, Ecuación de onda, Coordenadas cilíndricas, Aprendizaje no supervisado.

Abstract

Neural networks have driven the recent progress of machine learning and, beyond their usual applications, are now being explored as a means to solve differential equations. At that frontier, Physics-Informed Neural Networks (PINNs) stand out: they learn the solution of an equation *without supervised data*, embedding the physical laws directly into the loss function and training the network to satisfy both the equation and its boundary conditions. This thesis brings that approach to the wave equation in cylindrical coordinates, whose radial component leads to Bessel's equation, a demanding case owing to its oscillatory behavior and known analytical solution. The core of the work, published as a scientific article, proposes four progressive, unsupervised strategies culminating in a Multi-Head architecture with Fourier Features and the recurrence relation as a physical constraint; with it, a single network approximates the four orders J_0 to J_3 with a similarity of 99.87 % to the analytical solution, and a lightweight variant reaches 99.7 % in 44 seconds with fewer than 100 000 parameters. The research goes beyond the article: it extends the method to the Neumann functions and, through the Wronskian, builds the Hankel functions —the traveling waves that complete the fundamental solution of the wave equation—. Rather than a table of values at discrete points, the network yields a continuous, differentiable model that can be evaluated anywhere in the domain: this property is what makes the approach competitive with classical numerical methods when the solution must feed into later gradient-based computations.

Keywords: Physics-Informed Neural Networks, Bessel Functions, Wave Equation, Cylindrical Coordinates, Unsupervised Learning.

Capítulo 1

Protocolo de tesis

1.1. Introducción

El aprendizaje automático (ML, por sus siglas en inglés) es una rama de la inteligencia artificial que permite a los sistemas informáticos aprender de los datos y mejorar su desempeño sin ser programados explícitamente para cada tarea. Con el crecimiento de los volúmenes de datos y el aumento en la capacidad computacional, el aprendizaje automático ha progresado rápidamente, encontrando aplicaciones en diversos campos como la salud, las finanzas, la automatización y el análisis de datos a gran escala (Goodfellow et al., 2016).

La inteligencia artificial (IA), con el aprendizaje automático y en particular las redes neuronales, ha demostrado ser una herramienta poderosa para abordar problemas numéricos complejos, incluidas las ecuaciones diferenciales. Las redes neuronales pueden aproximar soluciones a ecuaciones diferenciales ordinarias y parciales, permitiendo encontrar soluciones en problemas donde los métodos tradicionales son computacionalmente costosos o ineficientes. El uso de redes neuronales profundas y técnicas de aprendizaje automático ha permitido resolver problemas de física computacional y simulación de sistemas dinámicos con mayor velocidad y eficiencia, al tiempo que reduce la necesidad de discretización y la acumulación de errores numéricos (Lagaris et al., 1998).

Una gran parte de los avances tecnológicos que vivimos hoy en día se los debemos a la física, ciencia que estudia, entiende y busca darles definición a los fenómenos de la materia, la energía, el espacio y el tiempo, así como a la interacción entre ellos. Es una de las ciencias naturales más

antiguas e importantes y tiene como objetivo general comprender cómo funciona el universo a través de leyes y teorías utilizando el método científico (Tipler y Mosca, 2010).

La física utiliza el lenguaje de las matemáticas para modelar los problemas, para hacer teorías y generalizaciones. Una herramienta muy utilizada en problemas más complejos y elaborados son las ecuaciones diferenciales. Las ecuaciones diferenciales son fundamentales en la modelización matemática de fenómenos naturales y tecnológicos. Estas ecuaciones permiten describir el comportamiento de sistemas dinámicos como el movimiento de los planetas, la propagación del calor, las fluctuaciones en las poblaciones biológicas y el flujo de corriente eléctrica. Su aplicación es clave en la física, la ingeniería, la economía y otras ciencias. Estas ecuaciones diferenciales son complejas de manejar, muchas de ellas no tienen solución analítica y se deben utilizar aproximaciones numéricas para resolverlas; a estos métodos se les llama *métodos numéricos* (Boyce y DiPrima, 2012).

Los métodos numéricos permiten resolver ecuaciones diferenciales en casos donde las soluciones analíticas no son posibles. Sin embargo, estos métodos presentan ciertas limitaciones. Una de las principales es la precisión: las soluciones obtenidas son aproximadas, y el error numérico puede acumularse a lo largo de los cálculos, especialmente cuando se trata de problemas de largo plazo. Además, los métodos numéricos pueden ser inestables, lo que significa que pequeñas perturbaciones en los datos iniciales pueden causar grandes desviaciones en el resultado final. Otro desafío es la eficiencia computacional, ya que algunos métodos requieren grandes cantidades de recursos computacionales para resolver problemas complejos, lo que puede hacer inviable su uso en ciertas aplicaciones. Por último, elegir el método numérico adecuado para un tipo específico de ecuación diferencial es crucial para obtener una solución precisa y eficiente. Cada método tiene sus fortalezas y debilidades dependiendo del tipo de ecuación (ordinaria o parcial), del comportamiento de la solución y del contexto del problema (Burden y Faires, 2011).

En la resolución de la ecuación de onda para diversos sistemas de coordenadas es indispensable encontrar, hallar o aproximar los resultados a una *solución analítica*. Las soluciones analíticas a la ecuación de onda proporcionan una comprensión amplia del comportamiento de las ondas en diferentes medios, que permiten la predicción precisa de fenómenos, lo cual es fundamental en aplicaciones científicas y tecnológicas (Strauss, 2007). Estas ecuaciones, debido a que son muy complejas, no siempre tienen soluciones analíticas conocidas; sin embargo, para el

caso de la ecuación de onda en sistemas de coordenadas cilíndricas sí existe esta solución.

Las redes neuronales profundas se están utilizando cada vez más para resolver problemas basados en ecuaciones diferenciales complejas, especialmente a través de enfoques híbridos que combinan datos experimentales y leyes físicas (J.-X. Wang et al., 2017). El aprendizaje profundo basado en principios físicos (PBDL, *Physics-Based Deep Learning*) combina redes neuronales con el conocimiento físico subyacente para mejorar la precisión y eficiencia en la resolución de problemas científicos y de ingeniería (Raissi et al., 2019).

1.2. Marco teórico

En esta sección se abordan algunos conceptos muy importantes que son base a lo largo de toda esta investigación.

1.2.1. Ecuaciones diferenciales

Una ecuación diferencial es una ecuación matemática que relaciona una función con sus derivadas. Las ecuaciones diferenciales aparecen en una variedad de disciplinas científicas y de ingeniería para describir fenómenos tales como el movimiento, la disipación de calor, el crecimiento de poblaciones y la propagación de ondas, entre otros. Estas ecuaciones pueden ser de diferentes tipos, entre ellas, las ecuaciones diferenciales ordinarias (EDO) y las ecuaciones diferenciales parciales (EDP), dependiendo de la naturaleza de las variables involucradas (Zill, 2018).

1.2.1.1. Ecuaciones diferenciales ordinarias

Las EDO se utilizan para describir sistemas en los que el cambio de una cantidad depende de una sola variable, como el tiempo. Un ejemplo clásico de una ecuación diferencial ordinaria es la ecuación de movimiento de Newton, que describe cómo cambia la velocidad de un objeto en función del tiempo bajo la influencia de fuerzas (Boyce y DiPrima, 2012).

1.2.1.2. Ecuaciones diferenciales parciales

Estas ecuaciones se utilizan para describir fenómenos en los que las cantidades varían con respecto a más de una variable, como el espacio y el tiempo. Ejemplos de ecuaciones diferenciales parciales incluyen la ecuación de calor, la ecuación de ondas y la ecuación de Laplace, que son fundamentales para modelar fenómenos físicos como la difusión de calor, la propagación de ondas y el comportamiento de campos electromagnéticos (Strauss, 2007).

1.2.2. Métodos numéricos

Los métodos numéricos son técnicas matemáticas utilizadas para encontrar soluciones aproximadas a problemas que no pueden resolverse de manera analítica o exacta. Estos métodos se emplean para resolver ecuaciones algebraicas, ecuaciones diferenciales, integrales y otros problemas matemáticos complejos. Algunos de los métodos numéricos más comunes incluyen el método de Euler, el método de Runge-Kutta, el método de Newton-Raphson y el método de diferencias finitas. Estos métodos son esenciales en la física, la ingeniería y la economía, entre otras áreas, donde las soluciones exactas no siempre son posibles o prácticas (Chapra y Canale, 2015).

1.2.2.1. Método de diferencias finitas

El método de diferencias finitas es una técnica numérica utilizada para aproximar soluciones a ecuaciones diferenciales, tanto ordinarias como parciales. Este método se basa en discretizar las derivadas reemplazándolas por diferencias finitas. Las soluciones se calculan en una malla de puntos distribuidos en el dominio de la ecuación, permitiendo aproximar tanto derivadas de primer como de segundo orden. El método es ampliamente utilizado en la simulación de fenómenos físicos y en la ingeniería, como en la propagación de calor, la dinámica de fluidos y la mecánica estructural (G. D. Smith, 1985).

1.2.2.2. Método de Euler

El método de Euler es un método numérico sencillo y directo utilizado para resolver ecuaciones diferenciales ordinarias (EDO). Se basa en aproximar la solución a una EDO en pasos

incrementales, utilizando la pendiente de la función en cada paso para predecir el valor de la función en el siguiente punto. Aunque es fácil de implementar, el método de Euler es generalmente menos preciso que otros métodos más avanzados, como el método de Runge-Kutta, debido a su error de truncamiento. No obstante, es útil para obtener soluciones aproximadas rápidas y para entender los principios básicos de los métodos numéricos para EDO (Burden y Faires, 2011).

1.2.3. Inteligencia artificial

La inteligencia artificial (IA) es el campo de estudio que se centra en el desarrollo de sistemas y algoritmos capaces de realizar tareas que, normalmente, requerirían inteligencia humana. Estas tareas incluyen el reconocimiento de patrones, el aprendizaje automático, la toma de decisiones, la resolución de problemas y la comprensión del lenguaje natural. La IA tiene aplicaciones en una amplia gama de áreas, desde la robótica y los sistemas expertos hasta el procesamiento del lenguaje natural y el aprendizaje profundo (Russell y Norvig, 2016).

1.2.3.1. Redes neuronales

Las redes neuronales son un modelo computacional inspirado en el funcionamiento del cerebro humano, compuesto por capas de neuronas artificiales conectadas entre sí. Estas redes son capaces de aprender patrones complejos a partir de los datos mediante el ajuste de los pesos de las conexiones entre las neuronas. Las redes neuronales se utilizan en tareas como el reconocimiento de imágenes, el procesamiento del lenguaje natural y la toma de decisiones. Existen varios tipos de redes neuronales, como las redes neuronales convolucionales (CNN) y las redes neuronales recurrentes (RNN), que se emplean para diferentes tipos de problemas (Goodfellow et al., 2016).

1.2.4. Ecuación de onda

La ecuación de onda es una ecuación diferencial parcial que describe la propagación de ondas, tales como las ondas de sonido, ondas de luz y ondas en el agua, en un medio determinado. Su forma general en una dimensión es:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$$

donde $u(x, t)$ es la función que describe la amplitud de la onda en un punto x y en el tiempo t , y c es la velocidad de la onda en el medio. Esta ecuación es fundamental en muchas áreas de la física y la ingeniería, ya que modela fenómenos como la propagación de ondas acústicas, electromagnéticas y mecánicas (Strauss, 2007).

1.2.4.1. Sistema de coordenadas cilíndricas

El sistema de coordenadas cilíndricas es un sistema de coordenadas tridimensional en el que un punto en el espacio se describe mediante tres valores: su distancia radial r desde un eje fijo, el ángulo θ medido con respecto a una dirección de referencia en el plano perpendicular al eje, y la altura z sobre el plano. Este sistema es especialmente útil para resolver problemas con simetría cilíndrica, como en el caso de flujos en tuberías o campos electromagnéticos en cables conductores.

$$r = \sqrt{x^2 + y^2}, \quad \theta = \arctan \frac{y}{x}, \quad z = z$$

La ecuación de onda en coordenadas cilíndricas, que describe la propagación de ondas en medios con simetría cilíndrica, es:

$$\frac{\partial^2 u}{\partial t^2} = c^2 \left(\frac{\partial^2 u}{\partial r^2} + \frac{1}{r} \frac{\partial u}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u}{\partial \theta^2} + \frac{\partial^2 u}{\partial z^2} \right)$$

donde $u(r, \theta, z, t)$ es la función de onda que describe la propagación en coordenadas cilíndricas y c es la velocidad de propagación de la onda en el medio. Esta forma de la ecuación es utilizada en problemas con simetría radial, como la propagación de ondas en cilindros o tubos (Kreyszig, 2011).

1.2.4.2. Ecuación y función de Bessel

En el contexto de las coordenadas cilíndricas, la ecuación de onda se reduce a una ecuación de Bessel para la componente radial (Arfken et al., 2013).

$$r^2 \frac{d^2 R(r)}{dr^2} + r \frac{dR(r)}{dr} + (k^2 r^2 - m^2) R(r) = 0$$

donde:

- $R(r)$ es la solución radial,
- k está relacionado con el número de onda,
- m es el orden de la función de Bessel.

Las funciones de Bessel son soluciones a la ecuación de Bessel, y desempeñan un papel fundamental en la resolución de problemas en coordenadas cilíndricas (Arfken et al., 2013). La **función de Bessel** de primera especie está definida como:

$$J_n(x) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(m+n+1)} \left(\frac{x}{2}\right)^{2m+n}$$

donde $J_n(x)$ es la solución regular de la ecuación de Bessel en el origen, y Γ es la función gamma.

1.2.4.3. Solución analítica

En la solución de ecuaciones diferenciales surge el concepto de solución analítica. Una solución analítica de una ecuación diferencial es una solución que se representa en forma de funciones conocidas y explícitas, sin necesidad de aproximaciones numéricas. Estas soluciones permiten comprender por completo el comportamiento de un sistema matemático y permiten tener exactitud en los resultados, lo cual es fundamental en problemas complejos reales que representan problemas físicos del mundo real (Simmons, 1991).

1.2.5. Estado del arte

Tres líneas de trabajo confluyen en el problema de esta tesis: la que estableció que una red neuronal puede aproximar la solución de una ecuación diferencial, la que llevó ese principio a la ecuación de onda y la que documentó por qué el principio se resiente cuando la solución oscila. Conviene leerlas juntas, porque la última explica los límites de las dos primeras. De la revisión —

en cada caso, lo que aporta y lo que deja sin resolver— salen los cinco ejes con que se comparan los trabajos en la tabla 1.1.

1.2.5.1. **Redes neuronales como aproximadores de soluciones de ecuaciones diferenciales**

La idea de fondo es anterior al auge reciente del aprendizaje profundo. Lagaris et al. (1998) mostraron que una red neuronal puede resolver una ecuación diferencial por colocación: se minimiza el residuo de la ecuación en un conjunto de puntos del dominio y las condiciones de frontera se satisfacen por construcción, mediante una transformación algebraica de la salida de la red. La propuesta ya contenía lo esencial del enfoque actual y, sobre todo, no requería ningún dato de la solución. Su alcance, sin embargo, estuvo limitado por la escala de las redes de la época y por el tipo de problemas ensayados: soluciones suaves, de variación lenta, en las que la dificultad que domina este trabajo —la oscilación— no aparece.

El planteamiento se retomó con las *Physics-Informed Neural Networks* (PINN) de Raissi et al. (2019), quienes lo reformularon sobre marcos modernos de diferenciación automática y lo llevaron a ecuaciones diferenciales parciales no lineales, a datos con ruido y a problemas inversos. Su formulación es deliberadamente general y admite dos regímenes: el problema directo, donde solo interviene la ecuación, y el inverso, donde se combinan observaciones con la física para recuperar parámetros desconocidos. Es en el segundo régimen donde la literatura posterior ha concentrado sus esfuerzos, y eso tiene una consecuencia incómoda: buena parte de los resultados que se citan como éxitos de las PINN se apoyan, en algún grado, en datos de la solución. E et al. (2017) argumentan a favor de este tipo de enfoques en problemas de alta dimensionalidad, donde la discretización de un método de malla resulta impracticable, y Thuerey et al. (2022) ofrecen la panorámica más reciente del campo, con especial atención a cómo se articula con los métodos numéricos existentes. Este último es un texto de síntesis, valioso como mapa del área, pero no aporta evidencia experimental sobre el problema concreto que aquí se aborda.

1.2.5.2. PINN aplicadas a la ecuación de onda

Los trabajos que han llevado las PINN a la ecuación de onda provienen en su mayoría de la geofísica, y en todos ellos se repite la misma condición de partida. Karimpouli y Tahmasebi (2020) comparan un proceso gaussiano y una PINN sobre una ecuación de onda sísmica unidimensional, con datos limpios y con ruido: la red resulta más precisa al invertir las velocidades de onda P y S, mientras que el proceso gaussiano aproxima mejor la solución directa. En ambos casos el entrenamiento se apoya en datos de la solución. Moseley et al. (2020) abordan la ecuación de onda acústica en medios heterogéneos —velocidad homogénea, en capas y una región del modelo Marmousi— y logran que la red capture ondas transmitidas y reflejadas y generalice a distintas posiciones de la fuente; es el resultado más sólido de los revisados, si bien el entrenamiento parte de la fuente y de las condiciones iniciales del campo. En Guo et al. (2020) se valida un algoritmo PINN sobre tres ecuaciones (onda, KdV-Burgers y KdV), aunque los propios autores enmarcan el resultado como aprendizaje de la solución «a partir de observaciones limitadas», y Rasht-Behesht et al. (2022) atacan el modelado directo e inverso de la ecuación de onda acústica partiendo de instantáneas tempranas del campo de desplazamiento.

Un caso algo distinto es EikoNet, de J. D. Smith et al. (2021), que resuelve la ecuación eikonal en estructuras de velocidad tridimensionales. El interés del caso está en que sustituye un cálculo que hay que repetir para cada configuración por una red consultable en cualquier punto.

Frente al problema de esta tesis, esa línea deja tres cosas fuera. La más importante es la dependencia de datos: en tres de los cinco casos revisados —Karimpouli y Tahmasebi (2020), Guo et al. (2020) y Rasht-Behesht et al. (2022)— el entrenamiento incorpora datos de la solución, de manera que lo demostrado no es que la física baste, sino que la física ayuda; en los dos restantes la información es puramente física, pero ninguno se ocupa de funciones especiales ni de identidades entre soluciones. Hay además una cuestión de geometría. Todos los casos citados trabajan en coordenadas cartesianas, y al reducir la ecuación de onda a coordenadas cilíndricas aparece el término $\frac{1}{\rho} \frac{\partial u}{\partial \rho}$, singular en el origen, que en esos trabajos no tiene análogo. Y ninguno se ocupa de las funciones especiales que resuelven la componente radial ni de las relaciones que esas funciones satisfacen entre sí.

1.2.5.3. Sesgo espectral y escalabilidad: los límites reconocidos

La tercera línea es la que explica por qué el problema de esta tesis no es una aplicación rutinaria de lo anterior. Rahaman et al. (2019) documentaron que las redes neuronales aprenden primero las componentes de baja frecuencia de una función y solo después, con dificultad, las de alta frecuencia; este sesgo espectral (*spectral bias*) es un obstáculo estructural para aproximar funciones oscilatorias como las de Bessel, y no un defecto de ajuste que se corrija con más épocas de entrenamiento. La propuesta de Tancik et al. (2020) nació para tareas de regresión de imágenes, reconstrucción tomográfica y síntesis de vistas, no para el residuo de una ecuación diferencial, pero es el resultado que más ha cambiado la práctica: proyectar la entrada sobre un conjunto de frecuencias —las *Fourier Features*— mitiga el sesgo y permite representar oscilaciones desde las primeras etapas del entrenamiento. Su traslado al residuo de una ecuación diferencial fue estudiado por S. Wang, Wang y Perdikaris (2021), que lo vinculan al sesgo de autovectores de la red; lo que queda abierto es cómo elegir la escala de las frecuencias cuando la solución oscila con amplitud decreciente, que es precisamente el caso de las funciones de Bessel.

Frente al mismo problema, Moseley et al. (2023) proponen una vía distinta con las FBPINN (*Finite Basis Physics-Informed Neural Networks*): partir el dominio en subdominios y asignar una red a cada uno, de modo que cada red solo deba representar el contenido de frecuencia local. El enfoque escala bien y se valida sobre varios problemas de referencia, pero paga el precio de multiplicar el número de redes y de gestionar la coherencia entre subdominios; ataca la escalabilidad más que la representación. La tesis de Moseley (2022) sitúa ambos trabajos en un programa más amplio sobre aprendizaje automático informado por la física y sus aplicaciones en geofísica y ciencia lunar. Ninguna de estas propuestas combina la codificación de la entrada con las identidades algebraicas que el propio problema proporciona, y es ahí donde arranca esta investigación.

1.2.5.4. Síntesis y vacío de conocimiento

La tabla 1.1 ordena los trabajos revisados según esos cinco ejes.

El enfoque para resolver ecuaciones diferenciales con redes neuronales está consolidado, y también lo están las técnicas para el obstáculo que plantean las soluciones oscilatorias; falta apli-

Tabla 1.1. Comparación de los trabajos revisados según los ejes pertinentes al problema abordado. Fuente: elaboración propia.

Trabajo	Información de entrenamiento	Geometría	S.e.	Varias	Ident.
Lagaris et al. (1998)	EDO/EDP y frontera	Cartesiana	X	X	X
Raissi et al. (2019)	EDP y frontera; observaciones en régimen inverso	Cartesiana	X	X	X
Karimpouli & Tahmasebi (2020)	EDP y observaciones	Cartesiana 1D	X	X	X
Moseley et al. (2020)	EDP, fuente y condiciones iniciales	Cartesiana	X	X	X
Guo et al. (2020)	EDP y observaciones limitadas	Cartesiana 1D	X	X	X
Rasht-Behesht et al. (2022)	EDP, instantáneas tempranas y sismogramas	Cartesiana	X	X	X
J. D. Smith et al. (2021)	Ec. eikonal y modelo de velocidad	Cartesiana 3D	X	X	X
Moseley et al. (2023)	EDP y frontera	Cartesiana	✓ ^a	X ^b	X
Este trabajo	EDP, frontera e identidades	Cilíndrica (radial)	✓ ^c	✓ ^d	✓ ^e

S.e.: atiende explícitamente el sesgo espectral. **Varias:** obtiene con un solo modelo varias soluciones linealmente independientes de la misma ecuación (distintos órdenes o especies), no una familia paramétrica de la misma solución. **Ident.:** emplea identidades del propio problema como restricción.

^a Mediante descomposición del dominio y renormalización de la entrada en cada subdominio; no mediante codificación en frecuencias. ^b Emplea una red por subdominio de una misma solución. ^c Mediante *Fourier Features*.

^d Arquitectura *Multi-Head* con *backbone* compartido en el capítulo 2, y red paramétrica en (ρ, n) en la extensión del capítulo 3. ^e Relación de recurrencia entre órdenes y , en la extensión, el Wronskiano.

carlos al caso del que se ocupa este trabajo. La revisión no localizó ningún estudio que reúna las tres condiciones que definen esta investigación: aproximar las funciones de Bessel de primera especie J_0 a J_3 sin emplear valores tabulados ni evaluaciones de la solución analítica durante el entrenamiento; obtener varios órdenes con un modelo único; e imponer la relación de recurrencia entre órdenes contiguos —obtenible algebraicamente sin conocer la solución— como restricción física que acople esos órdenes. Conviene precisar el alcance de esa afirmación, porque por separado varios de esos elementos ya tienen antecedente. Jin et al. (2022) resuelven problemas de autovalores cuánticos con una red cuyo entrenamiento prescinde de datos y que satisface las condiciones de frontera por construcción, e incorporan un mecanismo de barrido con el que una misma red localiza un número arbitrario de soluciones. Zou y Karniadakis (2025) desarrollan la arquitectura *Multi-Head* con cuerpo compartido, precedente directo de la que se emplea en esta tesis. Y Geetanjali y Hiremath (2026) aplican PINN al problema de autovalores de Helmholtz que gobierna los modos de una guía de onda, que es el pariente más próximo en cuanto al tipo de ecuación abordada. Lo que no se encontró es la combinación de esos tres elementos ni, en particular, el uso de las identidades de la ecuación de Bessel sobre la componente radial en coordenadas cilíndricas.

Esa ausencia se explica por la procedencia de los trabajos revisados: dominios en los que hay datos de campo y en los que la geometría es cartesiana, de modo que ni la no supervisión estricta ni las identidades entre soluciones se presentaban como necesidades. Con las funciones de Bessel ambas cosas sí pesan: hay una solución analítica exacta contra la que verificar sin usarla para entrenar, y hay identidades que ninguna de las PINN revisadas explota.

1.3. Planteamiento del problema

La inteligencia artificial ha avanzado muy rápido en los últimos años, impulsada por el aumento en la capacidad de procesamiento, el acceso a grandes cantidades de datos y los avances en algoritmos de aprendizaje profundo. Estos avances han llevado a mejoras en áreas como el procesamiento del lenguaje natural, la visión por computadora, la robótica y la toma de decisiones autónoma. La inteligencia artificial está transformando industrias enteras, desde la salud hasta la manufactura y el transporte, y continúa evolucionando hacia aplicaciones más sofisticadas, como

los modelos de lenguaje grandes y las redes neuronales profundas (Marcus y Davis, 2019).

Por ello deben aprovecharse los avances de la inteligencia artificial en áreas en las que no es común hoy en día, y que tienen potencial para crecer, como el área de la física con la resolución de ecuaciones diferenciales complejas que hoy en día solo se resuelven con métodos numéricos. Estos métodos numéricos consumen muchos recursos computacionales. Chapra y Canale (2015) mencionan que los métodos numéricos requieren una gran cantidad de recursos computacionales, especialmente cuando son problemas de sistemas no lineales o específicamente de ecuaciones diferenciales parciales, lo que puede hacerlos ineficientes e inviables de tratar.

La resolución de la ecuación de onda en un sistema de coordenadas cilíndricas es útil en diversos ámbitos de la ingeniería y la ciencia, especialmente en problemas con simetría cilíndrica. Algunos de ellos son:

- **Ingeniería acústica:** fundamental en el diseño y análisis de la propagación de ondas sonoras en estructuras cilíndricas, como tubos y conductos de sonido. Esto es aplicable en sistemas de escape de motores y en instrumentos musicales (Kinsler et al., 2000).
- **Electromagnetismo:** esencial para analizar la distribución de campos electromagnéticos en cables coaxiales, guías de onda y fibras ópticas (Pozar, 2011).
- **Mecánica estructural:** la propagación de ondas mecánicas, como las vibraciones, en estructuras cilíndricas es clave en ingeniería civil y mecánica (Timoshenko y Goodier, 1970).
- **Dinámica de fluidos:** utilizada para modelar la propagación de ondas de presión en flujos confinados, como en tuberías (Streeter et al., 1998).
- **Física de plasmas:** crucial en dispositivos con geometría cilíndrica como los *tokamaks* utilizados en la investigación de fusión nuclear (Freidberg, 2007).
- **Propagación de calor:** en su forma parabólica (ecuación de calor), esencial para modelar la transferencia de calor en tuberías y reactores (Carslaw y Jaeger, 1959).
- **Geofísica y sismología:** útil para modelar la propagación de ondas sísmicas en pozos y otras estructuras cilíndricas subterráneas (Aki y Richards, 2002).

El problema central que se aborda en esta investigación tiene que ver con la naturaleza misma de la solución que los métodos numéricos tradicionales son capaces de entregar. Las diferencias finitas o el método de Euler resuelven la ecuación de onda en coordenadas cilíndricas sobre una malla y devuelven, en consecuencia, un conjunto discreto de valores atado a esa malla: fuera de los nodos hay que interpolar, ganar resolución obliga a repetir el cálculo sobre una malla más fina (LeVeque, 2007), y las derivadas de la solución solo se recuperan por cocientes de diferencias, con el error de truncamiento que ello arrastra (Burden & Faires, 2011, cap. 4). Lo que estos métodos no producen es un objeto matemático continuo y derivable en todo el dominio, comparable en forma —no solo en valores— con la solución analítica.

Las *Physics-Informed Neural Networks* (PINN) y el *Physics-Based Deep Learning* (PBDL) ofrecen precisamente ese tipo de objeto. Al incorporar la ecuación diferencial en la función de pérdida, la red entrenada es una función continua evaluable en cualquier punto del dominio y derivable de manera exacta mediante diferenciación automática, sin malla que la condicione (Raissi et al., 2019). Esta investigación se centra, por tanto, en dos cuestiones: en qué grado la aproximación que produce una red neuronal concuerda con la solución analítica conocida cuando el entrenamiento se apoya únicamente en la información de la ecuación diferencial, y qué ventajas cualitativas ofrece esa representación frente al resultado de un método numérico.

1.4. Justificación

La resolución de ecuaciones diferenciales parciales (EDP), como la ecuación de onda en coordenadas cilíndricas, es fundamental en numerosos campos de la ciencia e ingeniería, entre ellos la acústica, el electromagnetismo, la mecánica estructural y la dinámica de fluidos. Sin embargo, los métodos numéricos tradicionales, aunque ampliamente utilizados, presentan limitaciones significativas: su alto costo computacional, la acumulación de errores en problemas complejos o de larga duración, y la ineficiencia en sistemas de alta dimensionalidad (Burden y Faires, 2011; Chapra y Canale, 2015; E et al., 2017).

En este contexto, surge la necesidad de explorar enfoques alternativos. Las PINN y el PBDL combinan la capacidad de aproximación de las redes neuronales con principios físicos, y eluden así algunas de las limitaciones de los métodos numéricos clásicos: permiten integrar la ecuación

diferencial en el entrenamiento y prescinden de la discretización del dominio (Raissi et al., 2019). El PBDL, por su parte, aprovecha tanto datos experimentales como conocimiento físico previo (Thuerey et al., 2022). Conviene señalar que el enfoque tiene sus propias dificultades, bien documentadas: existen modos de falla en los que una PINN no converge aunque la formulación de la pérdida sea correcta (Krishnapriyan et al., 2021), y el equilibrio entre los distintos términos de la pérdida puede degradar el entrenamiento si no se atiende (S. Wang, Teng y Perdikaris, 2021). Parte del trabajo de esta tesis consiste precisamente en sortear esas dificultades en el caso concreto que aborda.

La investigación no se propone demostrar la superioridad numérica de las redes neuronales sobre los métodos tradicionales; una comparación así enfrentaría objetos de naturaleza distinta y poco diría del valor de uno u otro enfoque. Lo que se busca es caracterizar las ventajas cualitativas de las redes neuronales como vía alternativa, identificar la arquitectura más adecuada para resolver la ecuación de onda en sistemas con simetría cilíndrica y cuantificar la fidelidad de la aproximación obtenida frente a la solución analítica conocida. Además, el uso de redes neuronales permite explorar una metodología escalable, que podría aplicarse a otros tipos de ecuaciones diferenciales o problemas similares.

Desde una perspectiva social, la optimización en la resolución de ecuaciones diferenciales mediante redes neuronales permite el desarrollo de tecnologías más precisas y eficientes en áreas como la salud, la acústica, la energía y las telecomunicaciones. En el ámbito computacional, las PINN y el PBDL ofrecen una alternativa a los métodos numéricos tradicionales al entregar una representación continua y derivable de la solución en todo el dominio, con lo que desaparece el compromiso entre finura de la malla y costo de resolverla de nuevo. Finalmente, en términos económicos, disponer de un modelo entrenado una sola vez y consultable después en cualquier punto resulta conveniente en escenarios donde la misma ecuación debe evaluarse repetidas veces. Otro aspecto relevante es que esa representación se acerca en forma a una solución analítica, algo que un método numérico no puede ofrecer, pues por definición produce el resultado opuesto: un muestreo discreto de la solución.

1.5. Pregunta de investigación

1. ¿Puede un modelo de red neuronal aproximar la solución fundamental analítica de la ecuación de onda en coordenadas cilíndricas con una fidelidad verificable, apoyándose únicamente en la información de la ecuación diferencial y sin datos de la solución?
2. ¿Cuál es la estructura o arquitectura de red neuronal (PBDL, PINN u otra) que produce la mejor aproximación a esa solución fundamental, y qué ventajas cualitativas ofrece la solución así obtenida frente a la que entregan los métodos numéricos tradicionales aplicados a la misma ecuación?

1.6. Hipótesis o supuesto de investigación

Las redes neuronales, a través de las *Physics-Informed Neural Networks* (PINN) y del *Physics-Based Deep Learning* (PBDL), permiten aproximar con alta fidelidad la solución analítica de la ecuación de onda en sistemas de coordenadas cilíndricas sin recurrir a datos de esa solución, y la representación que producen —una función continua y derivable en todo el dominio— constituye una alternativa a los métodos numéricos tradicionales, cuyo resultado es discreto y dependiente de la malla de discretización.

1.7. Objetivo general

Encontrar una solución fundamental a la ecuación de onda en coordenadas cilíndricas a través del desarrollo de modelos de redes neuronales (PINN o PBDL), que aproxime con alta fidelidad la solución analítica conocida y que constituya una alternativa a los métodos numéricos tradicionales por su carácter continuo, diferenciable y libre de malla.

1.8. Objetivos específicos

1. Desarrollar y evaluar modelos de *Physics-Based Deep Learning* (PBDL) y de *Physics-Informed Neural Networks* (PINN) que combinen conocimiento físico previo con técnicas

de aprendizaje profundo, buscando soluciones relacionadas con la ecuación de onda.

2. Analizar el desempeño de los modelos PINN y PBDL en términos de precisión, velocidad de convergencia y consumo de recursos computacionales, y caracterizar sus ventajas cualitativas frente a los métodos numéricos tradicionales (una solución continua, diferenciable y libre de malla).
3. Obtener una solución a la ecuación de onda en coordenadas cilíndricas que constituya una alternativa a los métodos numéricos tradicionales, con la ventaja de una representación continua y diferenciable del dominio.
4. Obtener, mediante los modelos de red neuronal, una aproximación de alta fidelidad a la solución analítica conocida, y cuantificar esa fidelidad con métricas de error verificables.

1.9. Alcances y limitaciones

1.9.1. Alcances

1. Implementación de modelos PINN y PBDL: se desarrollaron modelos de redes neuronales diseñados específicamente para resolver la ecuación de onda en coordenadas cilíndricas, y se evaluó su desempeño en términos de precisión y de costo de entrenamiento.
2. Comparación cualitativa con los métodos numéricos tradicionales: se contrastó la naturaleza de la solución que produce cada enfoque —continua y derivable frente a discreta y dependiente de la malla— a partir de implementaciones propias de diferencias finitas y Runge-Kutta para la misma ecuación. No se planteó un juicio de superioridad numérica: se comparan cosas que no son del mismo tipo.
3. Propuesta de arquitectura óptima: se identificó la estructura de red más adecuada para resolver este tipo de ecuación diferencial, lo que permite replicar y extender el trabajo a otras ecuaciones con características similares.
4. Aproximación de alta fidelidad: se obtuvo una aproximación de alta fidelidad a la solución de la ecuación de onda en coordenadas cilíndricas.

1.9.2. Limitaciones

1. La investigación se centró en la solución de la ecuación de onda en coordenadas cilíndricas, y no en otros sistemas de coordenadas.
2. El trabajo se concentró en la componente radial de esa solución, es decir, en las funciones de Bessel. El artículo del capítulo 2 cubre las funciones de primera especie J_0 a J_3 ; la extensión a las funciones de segunda especie y a las de Hankel se documenta en el capítulo 3.

1.10. Metodología general

El proceso de investigación se organizó con la metodología CRISP-DM (*Cross-Industry Standard Process for Data Mining*), que es la registrada en el protocolo. Conviene precisar en qué sentido se empleó, porque CRISP-DM nació como estándar para proyectos de minería de datos y este trabajo no parte de un *dataset*: en una PINN entrenada de forma no supervisada no hay observaciones de la solución que la red deba ajustar. De CRISP-DM se retomó su función de **marco de gestión del proceso** —el ciclo iterativo de comprender el problema, construir el modelo, evaluarlo y desplegarlo, con retorno a etapas anteriores cuando la evaluación lo exige—, no su vocabulario orientado al dato. Las dos etapas que en un proyecto de minería se refieren al dato tienen aquí un referente distinto, pero igualmente preciso: la información con la que se entrenó el modelo la constituyen los puntos de colocación del dominio y las restricciones físicas que se impusieron sobre ellos. La tabla 1.2 recoge, etapa por etapa, esa traducción.

Dentro de ese marco, el trabajo experimental siguió un **desarrollo progresivo**: cuatro estrategias sucesivas, denominadas E1 a E4, cada una de las cuales incorporó un elemento nuevo sobre la anterior y se evaluó frente a la misma referencia analítica. El propósito de encadenarlas de esa manera es aislar la mejora atribuible a cada elemento, en lugar de presentar una única arquitectura final sin dar cuenta de qué parte de su desempeño proviene de cada decisión de diseño. Este diseño es el que estructura el capítulo 2, donde se describe en detalle.

Tabla 1.2. Correspondencia entre las etapas de CRISP-DM y su equivalente en el flujo de trabajo con PINN. Fuente: elaboración propia.

Etapas CRISP-DM	Equivalente en esta investigación
Comprensión del problema	Fundamentos de la ecuación de onda en coordenadas cilíndricas, de la ecuación de Bessel y de sus soluciones; estudio de las PINN y del PBDL.
Comprensión de los datos	Caracterización de la información disponible, que aquí es física y no empírica: la ecuación diferencial, las condiciones de frontera y las identidades propias del problema (relación de recurrencia, Wronskiano). Incluye el análisis del carácter oscilatorio de la solución y del <i>spectral bias</i> que ese carácter induce en la red.
Preparación de datos	Construcción del conjunto de puntos de colocación sobre el dominio radial: muestreo por ley de potencias (<i>power-law sampling</i>), densificación adaptativa y codificación de la entrada mediante <i>Fourier Features</i> .
Modelado	Diseño, implementación y entrenamiento de las cuatro configuraciones sucesivas (E1 a E4).
Evaluación	Contraste con la solución analítica conocida, análisis de sensibilidad de los hiperparámetros y comparación cualitativa con los métodos numéricos.
Despliegue	Presentación de los resultados y publicación del artículo científico.

1.10.1. Comprensión del problema

Se profundizó en los fundamentos teóricos de la ecuación de onda y en la relevancia de las soluciones fundamentales frente a su solución analítica, y se reforzó la base técnica sobre redes neuronales, en particular las PINN y el PBDL.

1.10.2. Comprensión de la información física disponible

En esta etapa se identificaron y caracterizaron las propiedades físicas y matemáticas que orientaron el entrenamiento de la red: la propia ecuación diferencial, las condiciones de frontera que distinguen una solución de otra y las identidades que el problema satisface por construcción. Se analizó además el comportamiento oscilatorio y de amplitud decreciente de las funciones de Bessel, que es el que determina las dificultades de aproximación que se abordan más adelante.

1.10.3. Preparación de los puntos de colocación

El dominio radial y el conjunto de puntos de colocación sobre el que se evaluó el residuo quedaron fijados en esta etapa, junto con la codificación de la entrada de la red. Es aquí donde se decide cómo entra la física en el modelo, ya que no hay ejemplos de la solución que sustituyan esa decisión.

1.10.4. Modelado

Se implementaron y entrenaron las redes neuronales, fijando los criterios de optimización, ajustando los hiperparámetros y comparando distintas configuraciones para evaluar su rendimiento. Se comenzó por ecuaciones diferenciales menos complejas antes de abordar de lleno la ecuación de onda. En esta etapa se recurrió tanto a PINN como a PBDL, de manera que el resultado no es un modelo único, sino un conjunto de configuraciones contrastadas entre sí.

1.10.5. Evaluación

Se analizó el desempeño de las redes frente a la solución analítica conocida, con la similaridad media sobre los cuatro órdenes como métrica principal, y se contrastó el resultado con el de las implementaciones propias de diferencias finitas y Runge-Kutta, en los términos cualitativos ya señalados.

1.10.6. Despliegue

Los resultados se documentaron en el artículo del capítulo 2, que consigna las arquitecturas, los hiperparámetros de cada estrategia, el esquema de muestreo y las semillas empleadas, de modo que los experimentos puedan repetirse en sus términos exactos.

1.11. Contribución y resultados

Los productos comprometidos en el protocolo y el estado que alcanzaron al término de la investigación son los siguientes:

- Modelo y arquitectura de red neuronal desarrollados y entrenados, documentados en el artículo del capítulo 2.
- La presente tesis de maestría.
- Un artículo científico, aceptado en el Congreso Mexicano de Inteligencia Artificial (COMIA 2026) y reproducido íntegramente en el capítulo 2.

1.12. Cronograma de actividades

Tabla 1.3. Cronograma de actividades de la investigación. Fuente: elaboración propia.

No.	Actividad	S1	S2	S3	S4
1	Revisión de literatura	✓	✓	✓	✓
2	Redacción del protocolo	✓			
3	CRISP-DM		✓	✓	✓
	3.1 Comprensión del problema		✓		
	3.2 Comprensión de la información física		✓		
	3.3 Preparación de puntos de colocación		✓	✓	
	3.4 Modelado		✓	✓	
	3.5 Evaluación			✓	✓
	3.6 Despliegue				✓
4	Redacción y envío de artículo		✓	✓	
5	Redacción y revisión de la tesis			✓	✓
6	Presentación y defensa de la tesis				✓

Referencias

- Aki, K., & Richards, P. G. (2002). *Quantitative seismology* (2.^a ed.). University Science Books.
- Arfken, G. B., Weber, H. J., & Harris, F. E. (2013). *Mathematical methods for physicists* (7.^a ed.). Elsevier.
- Boyce, W. E., & DiPrima, R. C. (2012). *Ecuaciones diferenciales y problemas con valores en la frontera* (10.^a ed.). John Wiley & Sons.
- Burden, R. L., & Faires, J. D. (2011). *Análisis numérico* (9.^a ed.). Cengage Learning.
- Carslaw, H. S., & Jaeger, J. C. (1959). *Conduction of heat in solids* (2.^a ed.). Oxford University Press.
- Chapra, S. C., & Canale, R. P. (2015). *Métodos numéricos para ingenieros* (7.^a ed.). McGraw-Hill Education.
- E, W., Han, J., & Jentzen, A. (2017). Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *Communications in Mathematics and Statistics*, 5(4), 349-380. <https://doi.org/10.1007/s40304-017-0117-6>
- Freidberg, J. P. (2007). *Plasma physics and fusion energy*. Cambridge University Press.
- Geetanjali & Hiremath, K. R. (2026). Computation of waveguide eigenmodes by physics-informed neural networks. *Machine Learning: Engineering*, 2(1), 015005. <https://doi.org/10.1088/3049-4761/ae510e>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Guo, Y., Cao, X., Liu, B., & Gao, M. (2020). Solving partial differential equations using deep learning and physical constraints. *Applied Sciences*, 10(17), 5917. <https://doi.org/10.3390/app10175917>

- Jin, H., Mattheakis, M., & Protopapas, P. (2022). Physics-informed neural networks for quantum eigenvalue problems. *2022 International Joint Conference on Neural Networks (IJCNN)*, 1-8. <https://doi.org/10.1109/IJCNN55064.2022.9891944>
- Karimpouli, S., & Tahmasebi, P. (2020). Physics informed machine learning: Seismic wave equation. *Geoscience Frontiers*, 11(6), 1993-2001. <https://doi.org/10.1016/j.gsf.2020.07.007>
- Kinsler, L. E., Frey, A. R., Coppers, A. B., & Sanders, J. V. (2000). *Fundamentals of acoustics* (4.^a ed.). John Wiley & Sons.
- Kreyszig, E. (2011). *Advanced engineering mathematics* (10.^a ed.). John Wiley & Sons.
- Krishnapriyan, A., Gholami, A., Zhe, S., Kirby, R., & Mahoney, M. W. (2021). Characterizing possible failure modes in physics-informed neural networks. *Advances in Neural Information Processing Systems*, 34, 26548-26560.
- Lagaris, I. E., Likas, A., & Fotiadis, D. I. (1998). Artificial neural networks for solving ordinary and partial differential equations. *IEEE Transactions on Neural Networks*, 9(5), 987-1000. <https://doi.org/10.1109/72.712178>
- LeVeque, R. J. (2007). *Finite difference methods for ordinary and partial differential equations*. SIAM.
- Marcus, G., & Davis, E. (2019). *Rebooting AI: Building artificial intelligence we can trust*. Pantheon Books.
- Moseley, B. (2022). *Physics-informed machine learning: from concepts to real-world applications* [DPhil thesis]. University of Oxford.
- Moseley, B., Markham, A., & Nissen-Meyer, T. (2020). Solving the wave equation with physics-informed deep learning [Preprint, arXiv:2006.11894]. <https://arxiv.org/abs/2006.11894>
- Moseley, B., Markham, A., & Nissen-Meyer, T. (2023). Finite basis physics-informed neural networks (FBPINNs): A scalable domain decomposition approach for solving differential equations. *Advances in Computational Mathematics*, 49(4), 62. <https://doi.org/10.1007/s10444-023-10065-9>
- Pozar, D. M. (2011). *Microwave engineering* (4.^a ed.). John Wiley & Sons.
- Rahaman, N., Baratin, A., Arpit, D., Draxler, F., Lin, M., Hamprecht, F. A., Bengio, Y., & Courville, A. (2019). On the spectral bias of neural networks. *Proceedings of the 36th International Conference on Machine Learning (ICML)*, 97, 5301-5310.

- Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686-707. <https://doi.org/10.1016/j.jcp.2018.10.045>
- Rasht-Behesht, M., Huber, C., Shukla, K., & Karniadakis, G. E. (2022). Physics-informed neural networks (PINNs) for wave propagation and full waveform inversions [e2021JB023120 2021JB023120]. *Journal of Geophysical Research: Solid Earth*, 127(5), e2021JB023120. <https://doi.org/10.1029/2021JB023120>
- Russell, S., & Norvig, P. (2016). *Artificial intelligence: A modern approach* (3.^a ed.). Pearson Education.
- Simmons, G. F. (1991). *Differential equations with applications and historical notes* (2.^a ed.). McGraw-Hill.
- Smith, G. D. (1985). *Numerical solution of partial differential equations: Finite difference methods* (3.^a ed.). Oxford University Press.
- Smith, J. D., Azizzadenesheli, K., & Ross, Z. E. (2021). EikoNet: Solving the eikonal equation with deep neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 59(12), 10685-10696. <https://doi.org/10.1109/TGRS.2020.3039165>
- Strauss, W. A. (2007). *Partial differential equations: An introduction* (2.^a ed.). John Wiley & Sons.
- Streeter, V. L., Wylie, E. B., & Bedford, K. W. (1998). *Fluid mechanics* (9.^a ed.). McGraw-Hill.
- Tancik, M., Srinivasan, P. P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J. T., & Ng, R. (2020). Fourier features let networks learn high frequency functions in low dimensional domains. *Advances in Neural Information Processing Systems (NeurIPS)*, 33, 7537-7547.
- Thuerey, N., Holl, P., Mueller, M., Schnell, P., Trost, F., & Um, K. (2022). Physics-based deep learning [Preprint, arXiv:2109.05237]. <https://arxiv.org/abs/2109.05237>
- Timoshenko, S. P., & Goodier, J. N. (1970). *Theory of elasticity* (3.^a ed.). McGraw-Hill.
- Tipler, P. A., & Mosca, G. (2010). *Física para la ciencia y la tecnología*. Reverté.
- Wang, J.-X., Wu, J.-L., & Xiao, H. (2017). Physics-informed machine learning approach for reconstructing Reynolds stress modeling discrepancies based on DNS data. *Physical Review Fluids*, 2(3), 034603. <https://doi.org/10.1103/PhysRevFluids.2.034603>

- Wang, S., Teng, Y., & Perdikaris, P. (2021). Understanding and mitigating gradient flow pathologies in physics-informed neural networks. *SIAM Journal on Scientific Computing*, 43(5), A3055-A3081.
- Wang, S., Wang, H., & Perdikaris, P. (2021). On the eigenvector bias of Fourier feature networks: From regression to solving multi-scale PDEs with physics-informed neural networks. *Computer Methods in Applied Mechanics and Engineering*, 384, 113938. <https://doi.org/10.1016/j.cma.2021.113938>
- Zill, D. G. (2018). *Ecuaciones diferenciales con problemas con valores en la frontera* (9.^a ed.). Cengage Learning.
- Zou, Z., & Karniadakis, G. E. (2025). Multi-head physics-informed neural networks for learning functional priors and uncertainty quantification. *Journal of Computational Physics*, 531, 113947. <https://doi.org/10.1016/j.jcp.2025.113947>

Capítulo 2

Redes Neuronales Informadas por la Física para resolver la ecuación de onda en coordenadas cilíndricas (funciones de Bessel de primera especie)

Redes Neuronales Informadas por la Física para resolver la ecuación de onda en coordenadas cilíndricas (funciones de Bessel de primera especie)

Gabriel Omar Dominguez Ruiz, José Adán Hernández-Nolasco

Universidad Juárez Autónoma de Tabasco, División Académica de Ciencias y Tecnologías de la Información

Resumen. Las Redes Neuronales Informadas por la Física (PINN) permiten resolver ecuaciones diferenciales incorporando las leyes físicas directamente en la función de pérdida, sin necesidad de datos supervisados. En este trabajo evaluamos la efectividad de las PINN para aproximar las funciones de Bessel de primera especie J_0 a J_3 , que surgen como soluciones de la ecuación de onda en coordenadas cilíndricas, con aplicaciones en acústica, electromagnetismo y mecánica cuántica. Proponemos cuatro estrategias, progresivas en técnica, complejidad y mejora del desempeño, que parten de un perceptrón multicapa (MLP) sencillo e incorporan de manera incremental: (i) mejoras en el remuestreo y los puntos de colocación, (ii) *Fourier Features* con una arquitectura *Multi-Head* de *backbone* compartido, y (iii) la relación de recurrencia como restricción física adicional. La estrategia final alcanza una similaridad media de $99.87\% \pm 0.06\%$ en J_0-J_3 sobre tres semillas aleatorias respecto a la solución analítica, utilizando únicamente información derivada de la ecuación diferencial; una variante ultraligera logra 99.7% de similaridad en 44 segundos en una MacBook Air M4 con menos de 100 000 parámetros. Estos resultados demuestran que la combinación sistemática de representaciones espectrales, arquitecturas compartidas y restricciones físicas auxiliares constituye una metodología efectiva para aplicar PINN a ecuaciones diferenciales ordinarias oscilatorias de alta complejidad.

Palabras clave: Redes Neuronales Informadas por la Física, aprendizaje automático, funciones de Bessel, ecuación de onda, coordenadas cilíndricas.

2.1. Introducción

Las Redes Neuronales Informadas por la Física (PINNs) (Raissi et al., 2019) integran las ecuaciones diferenciales directamente en la función de pérdida de una red neuronal, lo que permite resolver problemas sin datos etiquetados. Técnicas como las Fourier Features (Tancik et al., 2020) y el muestreo adaptativo (Lu et al., 2021) han mejorado aún más su desempeño.

La ecuación de onda en coordenadas cilíndricas gobierna una amplia gama de fenómenos físicos en la ciencia y la ingeniería: la propagación acústica en tuberías, ductos y resonadores cilíndricos (Morse y Ingard, 1968); los modos electromagnéticos (TE/TM) en guías de onda cilíndricas y fibras ópticas de índice escalonado (Pozar, 2011); la conducción transitoria de calor en cuerpos cilíndricos (Carslaw y Jaeger, 1959); los modos de vibración de membranas circulares; y el confinamiento cuántico en pozos de potencial cilíndricos (Griffiths y Schroeter, 2018). Tras la separación de variables, la componente radial de estas soluciones se reduce a la ecuación diferencial de Bessel, de modo que la capacidad de evaluar J_n con precisión—y, de manera crucial, en una forma *diferenciable y sin malla*—tiene un impacto directo en los flujos de simulación que integran el solucionador dentro de procesos diferenciables más amplios (diseño inverso, identificación de parámetros, optimización de extremo a extremo), donde las evaluaciones clásicas basadas en tablas o en recurrencias introducen error de interpolación e interrumpen el flujo del gradiente.

En este trabajo aplicamos PINNs para aproximar las funciones de Bessel de primera especie J_0 – J_3 , soluciones de la ecuación de onda en coordenadas cilíndricas (Abramowitz y Stegun, 1964; Watson, 1922). Su naturaleza oscilatoria las hace particularmente desafiantes debido al sesgo espectral (*spectral bias*) (Rahaman et al., 2019; Tancik et al., 2020), mientras que su solución analítica conocida permite cuantificar con precisión el efecto de cada decisión de diseño.

Nuestra contribución consiste en cuatro estrategias PINN progresivas, desarrolladas de manera incremental: (1) un MLP básico con la ecuación diferencial; (2) mejoras en el muestreo de los puntos de colocación; (3) Fourier Features con una arquitectura Multi-Head; y (4) la relación de recurrencia como restricción física adicional. Demostramos que su combinación alcanza una alta concordancia con la solución analítica utilizando únicamente información de la ecuación diferencial.

El artículo se organiza como sigue: la Sección 2.2 revisa el trabajo relacionado; la Sección 2.3 presenta los fundamentos teóricos; la Sección 2.4 describe las cuatro estrategias; la Sección 2.5 reporta los resultados; la Sección 2.6 discute los hallazgos; y la Sección 2.7 resume las conclusiones.

2.2. Trabajo Relacionado

2.2.1. Evaluación Numérica de Funciones de Bessel

Los algoritmos clásicos para evaluar $J_n(\rho)$ combinan expansiones en series de potencias cerca del origen, formas asintóticas para ρ grande, y recurrencia hacia atrás (de Miller) para controlar la estabilidad numérica (Abramowitz y Stegun, 1964; Amos, 1986; Press et al., 2007). Estos métodos entregan valores puntuales, pero no una representación continua y diferenciable; al integrarse en procesos basados en gradiente, los datos tabulados deben interpolarse, lo que introduce error adicional y rompe la diferenciable exacta. Los solucionadores de EDP basados en malla, como el método de elementos finitos y el de diferencias finitas (LeVeque, 2007), resuelven la ecuación de onda de forma genérica, pero requieren una discretización fina para resolver los regímenes oscilatorios y producen soluciones discretas y no diferenciables.

2.2.2. PINNs para Problemas Oscilatorios y Multiescala

Desde la introducción de las PINNs (Raissi et al., 2019), diversas líneas de trabajo han abordado las dificultades que plantean las soluciones oscilatorias y multiescala. El *spectral bias* (Rahaman et al., 2019) fue identificado como un obstáculo primordial, y se demostró que las Fourier Features (Tancik et al., 2020), junto con los mapeos de Fourier aleatorios adaptados a la EDP (S. Wang, Wang y Perdikaris, 2021), lo mitigan. Otras estrategias complementarias atacan el panorama de optimización: el balanceo de la pérdida basado en el NTK (S. Wang, Teng y Perdikaris, 2021), los pesos de colocación autoadaptativos (McClenny y Braga-Neto, 2023), las funciones de activación adaptativas (Jagtap et al., 2020) y los esquemas de entrenamiento causal (S. Wang et al., 2022). Los enfoques de descomposición de dominio, como las PINNs de base finita (FB-PINNs) (Moseley et al., 2023), localizan el contenido de frecuencia particionando el dominio. Los

análisis de modos de falla (Krishnapriyan et al., 2021) esclarecen cuándo las PINNs no convergen pese a formulaciones de pérdida formalmente correctas, y revisiones más amplias cubren aplicaciones en mecánica de fluidos (Cai et al., 2021) y bibliotecas de software como DeepXDE (Lu et al., 2021).

2.2.3. Posicionamiento de este Trabajo

Hasta donde sabemos, ningún estudio previo de PINNs aborda las funciones de Bessel de primera especie J_0 – J_3 con una metodología unificada y no supervisada que combine (i) Fourier Features, (ii) una arquitectura Multi-Head de *backbone* compartido, y (iii) la relación algebraica de recurrencia como restricción de acoplamiento entre órdenes derivada de la ecuación que gobierna el problema. La literatura existente de PINNs sobre EDO oscilatorias suele tratar una solución a la vez y no explota las identidades algebraicas entre órdenes. La ablación por etapas que presentamos aísla además la contribución marginal de cada ingrediente, algo que no encontramos en la literatura revisada.

2.3. Fundamentos Teóricos

2.3.1. Redes Neuronales Informadas por la Física (PINNs)

Una PINN aproxima la solución $u(\mathbf{x})$ de $\mathcal{N}[u] = 0$ mediante una red $u_\theta(\mathbf{x})$ entrenada minimizando (Raissi et al., 2019):

$$\mathcal{L}(\theta) = \mathcal{L}_{\text{PDE}}(\theta) + \lambda_{\text{BC}} \mathcal{L}_{\text{BC}}(\theta) \quad (2.1)$$

donde $\mathcal{L}_{\text{PDE}} = \frac{1}{N_c} \sum_{i=1}^{N_c} |\mathcal{N}[u_\theta](\rho_i)|^2$ penaliza el residuo en los puntos de colocación y $\mathcal{L}_{\text{BC}}$ impone las condiciones de frontera. Las derivadas se calculan mediante diferenciación automática.

2.3.2. Spectral Bias y Fourier Features

Las redes neuronales tienden a aprender primero las componentes de baja frecuencia (*spectral bias*) (Rahaman et al., 2019). Para mitigarlo, Tancik et al. (2020) propusieron proyectar la

entrada mediante Fourier Features:

$$\gamma(\rho) = [\sin(2\pi \mathbf{B}\rho), \cos(2\pi \mathbf{B}\rho)] \quad (2.2)$$

donde $\mathbf{B} \in \mathbb{R}^{1 \times m}$ contiene frecuencias muestreadas de $\mathcal{N}(0, \sigma^2)$, lo que permite representar oscilaciones de alta frecuencia desde las primeras épocas.

2.3.3. Caso de Estudio: la Ecuación Diferencial de Bessel

La ecuación de Bessel surge al resolver la ecuación de onda en coordenadas cilíndricas. La componente radial satisface:

$$\rho^2 \frac{d^2 R}{d\rho^2} + \rho \frac{dR}{d\rho} + (\rho^2 - n^2)R = 0 \quad (2.3)$$

donde ρ es la coordenada radial y n es el orden. Las soluciones regulares son las funciones de Bessel de primera especie $J_n(\rho)$ (Abramowitz y Stegun, 1964; Watson, 1922), con serie de Taylor:

$$J_n(\rho) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k! \Gamma(k+n+1)} \left(\frac{\rho}{2}\right)^{2k+n} \quad (2.4)$$

A partir de esta serie se obtienen las condiciones de frontera:

$$J_0(0) = 1, \quad J'_0(0) = 0 \quad (2.5)$$

$$J_1(0) = 0, \quad J'_1(0) = \frac{1}{2} \quad (2.6)$$

$$J_2(0) = 0, \quad J'_2(0) = 0, \quad J''_2(0) = \frac{1}{4} \quad (2.7)$$

$$J_n(0) = 0, \quad J'_n(0) = 0 \quad \text{para } n \geq 2 \quad (2.8)$$

Además, las funciones de Bessel satisfacen la relación de recurrencia:

$$J_{n+1}(\rho) = \frac{2n}{\rho} J_n(\rho) - J_{n-1}(\rho) \quad (2.9)$$

que puede derivarse algebraicamente de (2.3) sin conocer la solución (Abramowitz y Stegun, 1964), y actúa como restricción de acoplamiento entre órdenes en la función de pérdida.

2.3.4. Muestreo por Ley de Potencias

El muestreo por ley de potencias redistribuye los puntos de colocación según:

$$\rho_i = \rho_{\text{máx}} \cdot t_i^\alpha, \quad t_i = \frac{i}{N_c}, \quad i = 0, \dots, N_c \quad (2.10)$$

con $\alpha > 1$, concentrando puntos cerca del origen, donde las funciones de Bessel presentan la mayor variación.

2.4. Metodología

2.4.1. Formulación del Problema

Buscamos aproximar J_0-J_3 en $\rho \in [0, 10]$ utilizando únicamente la ecuación diferencial (2.3), las condiciones de frontera (2.5)–(2.8) y, opcionalmente, la relación de recurrencia (2.9). No se utiliza ningún valor de la solución analítica durante el entrenamiento.

2.4.2. Desarrollo Progresivo

Cada estrategia descrita a continuación surgió de las limitaciones de la anterior.

2.4.2.1. E1: MLP Sencillo con PDE y Condiciones de Frontera.

Un MLP individual por orden (3×20 , activación seno) con 2 000 puntos de colocación uniformes, entrenado durante 5 000 épocas con Adam ($\text{lr} = 0.004$). Véase la Fig. 2.1.

2.4.2.2. E2: Mejoras en el Remuestreo y los Puntos de Colocación.

Se conserva la arquitectura individual, pero con muestreo por ley de potencias ($\alpha = 2$) que concentra puntos cerca del origen. Véase la Fig. 2.2.

2.4.2.3. E3: Fourier Features con Backbone Compartido.

Arquitectura Multi-Head con *backbone* compartido (4×256 , SiLU) y Fourier Features ($m = 64$, $\sigma = 1$). Cuatro cabezales independientes producen $\hat{J}_0-\hat{J}_3$. Entrenamiento en dos fases: Adam

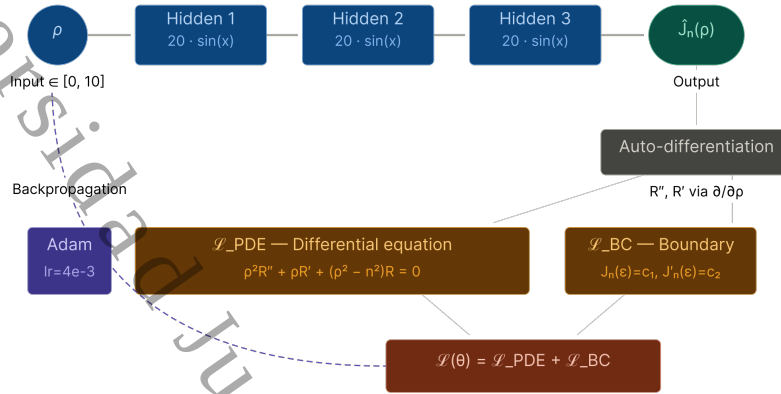


Figura 2.1. E1: MLP sencillo (3×20 , activación seno) con pérdida $\mathcal{L}_{\text{PDE}} + \mathcal{L}_{\text{BC}}$.

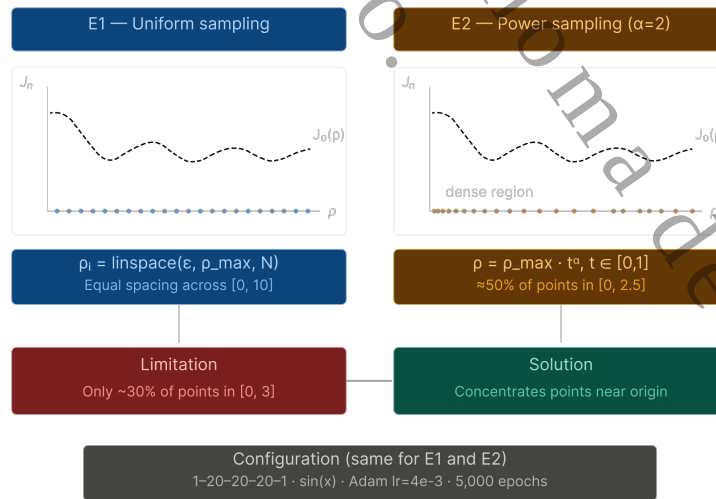


Figura 2.2. E2: muestreo uniforme (izquierda) frente al muestreo por ley de potencias $\alpha = 2$ (derecha), que concentra puntos cerca del origen.

(8 000 épocas) seguido de L-BFGS (5 000 iteraciones). Véase la Fig. 2.3.

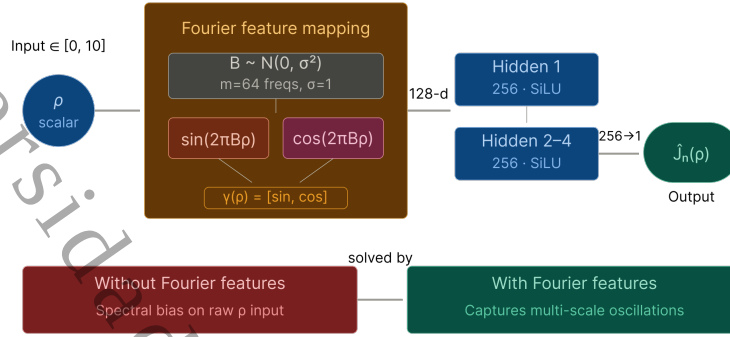


Figura 2.3. E3: arquitectura Multi-Head con Fourier Features ($m = 64$) y *backbone* compartido (4×256 , SiLU).

2.4.2.4. E4: Relación de Recurrencia como Restricción Física.

Conserva la arquitectura de E3 y añade la relación de recurrencia (2.9) a la función de pérdida:

$$\mathcal{L} = \mathcal{L}_{\text{PDE}} + \lambda_{\text{BC}} \mathcal{L}_{\text{BC}} + \lambda_{\text{Rec}} \mathcal{L}_{\text{Rec}} \quad (2.11)$$

donde $\lambda_{\text{BC}} = 50$, $\lambda_{\text{Rec}} = 15$, y:

$$\mathcal{L}_{\text{Rec}} = \frac{1}{N_s} \sum_{i=1}^{N_s} \left| J_2(\rho_i) - \frac{2}{\rho_i} J_1(\rho_i) + J_0(\rho_i) \right|^2 + \left| J_3(\rho_i) - \frac{4}{\rho_i} J_2(\rho_i) + J_1(\rho_i) \right|^2 \quad (2.12)$$

evaluada en $\rho > 0.1$ para evitar singularidades. Véase la Fig. 2.4.

2.5. Experimentos y Resultados

2.5.1. Configuración Experimental

Todos los experimentos se ejecutaron en una MacBook Air M4 con semilla aleatoria 42. Las métricas reportadas son:

- **MSE**: error cuadrático medio $\frac{1}{N} \sum_i (J_n^{\text{PINN}}(\rho_i) - J_n^{\text{exacta}}(\rho_i))^2$.
- **MAE**: error absoluto medio.

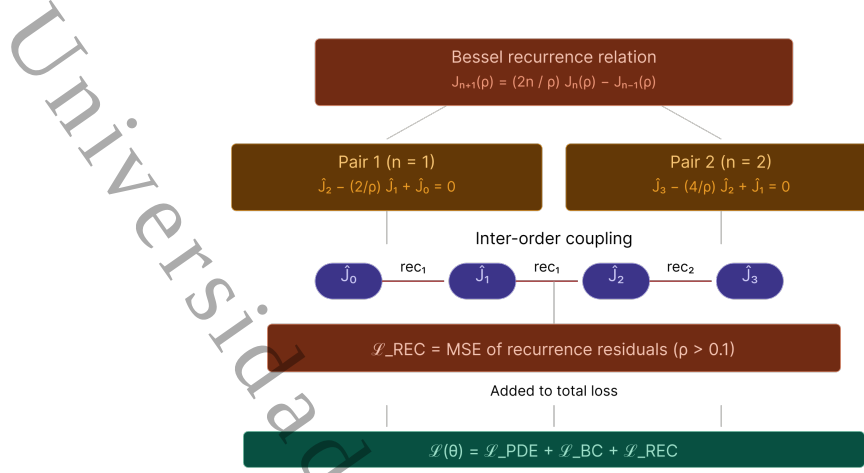


Figura 2.4. E4: la relación de recurrencia acopla las cuatro salidas ($\hat{J}_0 \leftrightarrow \hat{J}_1 \leftrightarrow \hat{J}_2 \leftrightarrow \hat{J}_3$) como restricción en la pérdida.

- **Error máximo:** error absoluto máximo.
- **Similitud:** $(1 - \bar{e}_{\text{rel}}) \times 100\%$, donde $\bar{e}_{\text{rel}}$ es el error relativo promedio en los puntos con $|J_n| > 0.1$.

Todas las métricas se evaluaron sobre 500 puntos equiespaciados en $[0, 10]$. El umbral $|J_n| > 0.1$ en la métrica de similitud evita que el error relativo quede dominado por puntos arbitrariamente cercanos a los ceros de J_n , donde $|J_n| \rightarrow 0$ y $\bar{e}_{\text{rel}}$ queda numéricamente mal definido independientemente de la calidad de la aproximación absoluta (un denominador que se anula infla el error relativo aun cuando el error absoluto es despreciable). De manera crucial, este umbral se aplica *solo* a la métrica de error relativo: el MSE, el MAE y el error absoluto máximo se calculan sobre los 500 puntos de evaluación sin exclusión, de modo que cualquier error de aproximación cerca de los ceros queda capturado por estas métricas complementarias y no puede ser enmascarado por la similitud.

2.5.2. Resultados

2.5.2.1. E1: Rendimiento del MLP Sencillo.

Primero se entrenó un MLP individual únicamente para J_0 . Con la arquitectura mínima (3×20 , activación seno), la PINN capturó el comportamiento oscilatorio general (MSE 4.14×10^{-4} , error

máximo 4.32×10^{-2} , similaridad 93.8 %), con desviaciones visibles en los máximos y los valles (Fig. 2.5). Al extender a dos redes individuales para J_0 y J_1 de forma simultánea se obtuvo 90.9 % y 97.1 % de similaridad, respectivamente; la pérdida de precisión en J_0 muestra que la configuración mínima no escala, lo que motivó las mejoras de muestreo de E2 (Fig. 2.6).

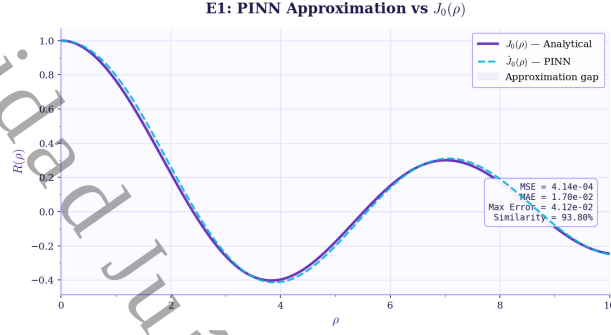


Figura 2.5. E1: aproximación de J_0 con el MLP sencillo (similaridad 93.8 %); desviaciones visibles en los máximos.

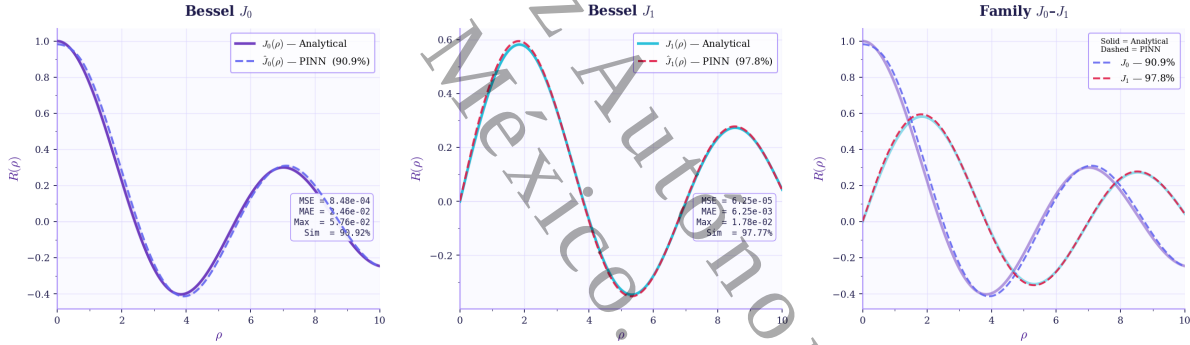


Figura 2.6. E1: extensión a J_0 (90.9 %) y J_1 (97.1 %) con redes individuales. J_0 pierde precisión respecto al caso individual.

2.5.2.2. E2: Efecto del Remuestreo y los Puntos de Colocación.

El muestreo por ley de potencias ($\alpha = 2$) concentra puntos cerca del origen y mejoró notablemente los dos primeros órdenes: J_0 alcanzó 98.3 % de similaridad ($\text{MSE } 3.30 \times 10^{-5}$) y J_1 alcanzó 98.1 % ($\text{MSE } 3.60 \times 10^{-5}$). Sin embargo, J_2 logró solo 34.4 % de similaridad ($\text{MSE } 3.12 \times 10^{-2}$), lo que confirma que el MLP sencillo carece de la capacidad de representación requerida para los órdenes superiores—donde las condiciones de frontera son más débiles y la dinámica inicial más sutil—incluso con un muestreo mejorado (Fig. 2.7).

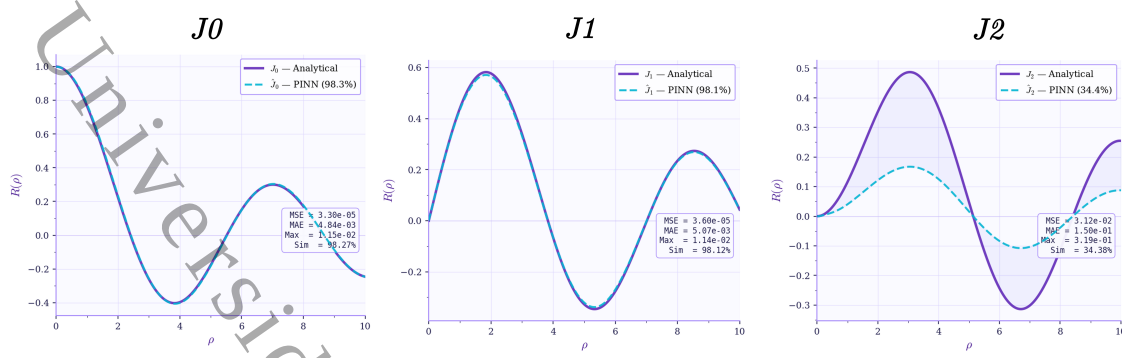


Figura 2.7. E2: muestreo por ley de potencias. J_0 (98.3%) y J_1 (98.1%) mejoran, pero J_2 (34.4%) se aparta de la solución analítica.

2.5.2.3. E3: Fourier Features y Backbone Compartido.

La incorporación de Fourier Features y la arquitectura Multi-Head con *backbone* compartido marcó un salto cualitativo: por primera vez los cuatro órdenes J_0 – J_3 se aproximaron simultáneamente con una sola red, con similaridades entre 60% y 97.4%. Las Fourier Features mitigaron el *spectral bias* al permitir representar oscilaciones de alta frecuencia. Persistieron desviaciones visibles en las regiones de alta variación, en particular para J_2 y J_3 , lo que sugería la necesidad de una restricción adicional (Fig. 2.8).

2.5.2.4. E4: Impacto de la Relación de Recurrencia.

La inclusión de la relación de recurrencia como restricción en la función de pérdida acopla los cuatro órdenes y cierra la brecha de precisión restante. La configuración de presupuesto alto de E4 (10 000 épocas de Adam, 5 000 iteraciones de L-BFGS, $N_c = 6\,000$) alcanza una similaridad media de $99.87\% \pm 0.06\%$ en J_0 – J_3 promediada sobre tres semillas aleatorias. Las curvas de pérdida de los tres componentes—residuo de la ecuación diferencial, condiciones de frontera y recurrencia—convergen de forma estable, lo que confirma que las restricciones no entran en conflicto entre sí. Adicionalmente, exploramos la frontera costo–precisión mediante un ajuste manual, coordinada por coordenada por coordenada, del presupuesto de entrenamiento: una configuración comprimida de 800 épocas de Adam seguidas de 500 iteraciones de L-BFGS con $N_c = 2\,000$ alcanza 100% de similaridad en solo 197 s en una MacBook Air M4 (una sola semilla; Fig. 2.9), mientras que la configuración estándar de 1 000 + 500 alcanza $99.77\% \pm 0.10\%$ en 230 s. Una variante ultraligera (500 épocas de Adam, 300 de L-BFGS, $N_c = 1\,000$, ancho oculto 128, 99 332 parámetros) alcanza

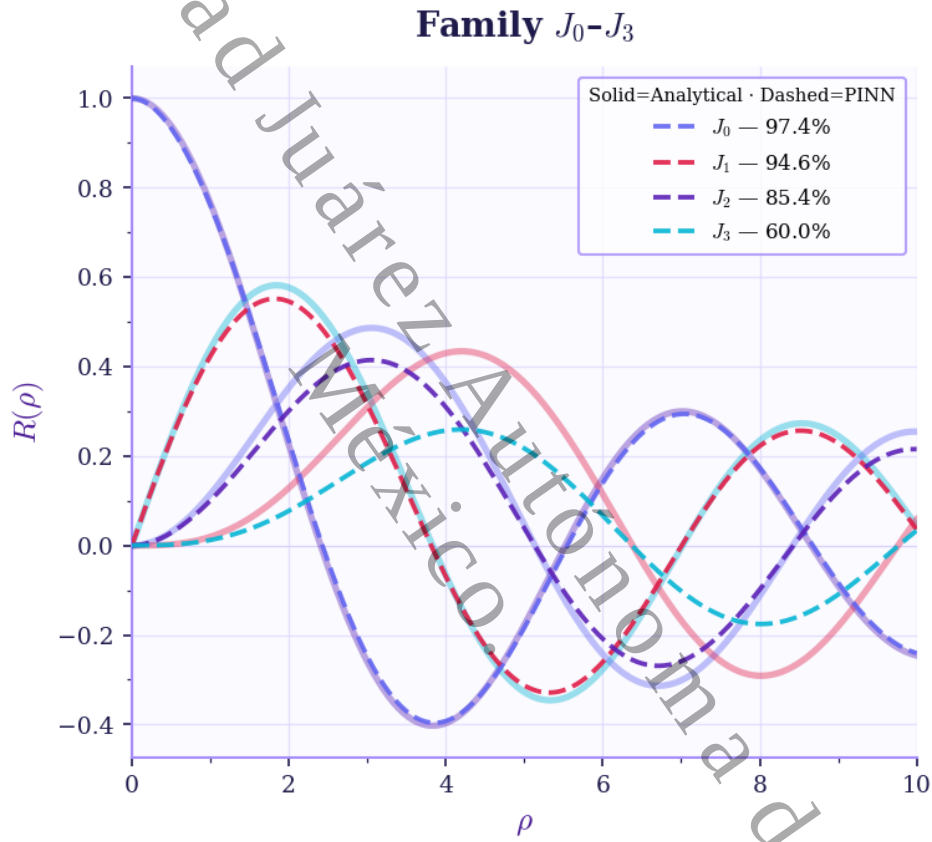


Figura 2.8. E3: aproximación simultánea con Fourier Features. Similaridades: $J_0 = 97.4\%$, $J_1 = 94.6\%$, $J_2 = 85.4\%$, $J_3 = 60.0\%$.

99.7 % de similitud en 44 s. La Tabla 2.1 resume la comparación costo–precisión, y la Sección 2.5.3 reporta el análisis de sensibilidad de los pesos de la pérdida y de los hiperparámetros arquitectónicos.

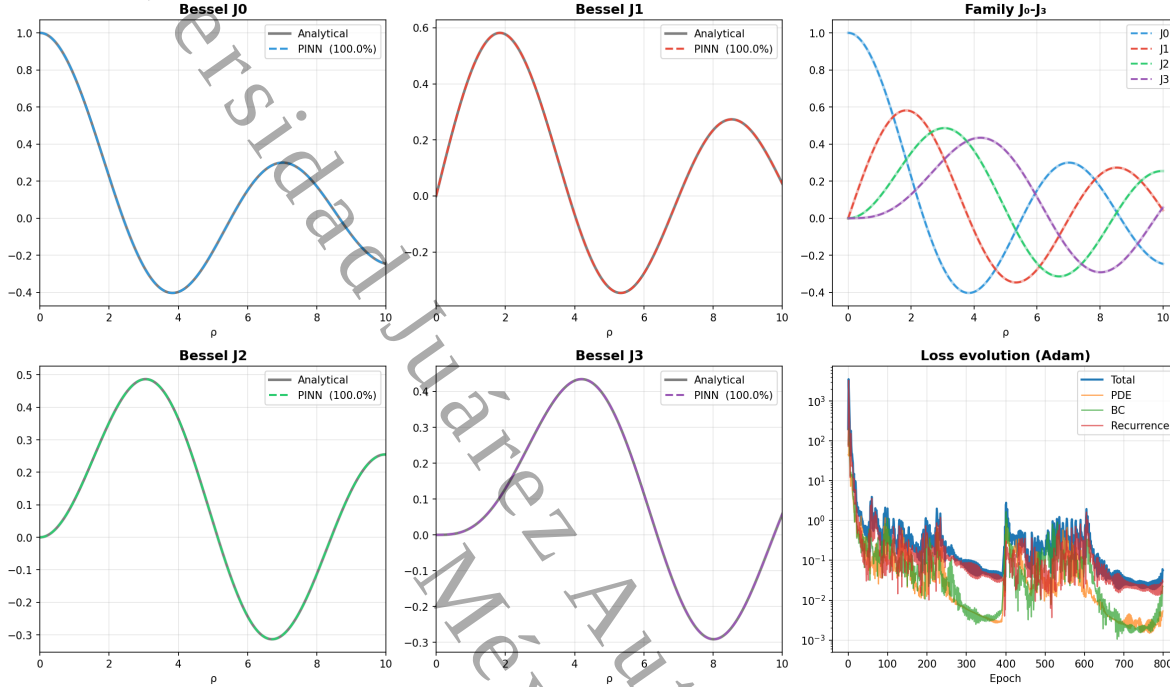


Figura 2.9. E4, configuración comprimida (800 épocas de Adam + 500 de L-BFGS, $N_c = 2000$): aproximación por orden y convergencia de la pérdida (100 % de similitud en 197 s en una MacBook Air M4).

Tabla 2.1. Comparación costo–precisión entre configuraciones de E4 en una MacBook Air M4 (CPU). *Sim.* es la similitud media en J_0 – J_3 . Las filas multi-semilla reportan media±desv. sobre las semillas {7, 42, 123}.

Configuración	Parám.	N_c	Adam+L-BFGS	Tiempo (s)	Sim.
E4 presupuesto alto	362 500	6 000	10k+5k	5 928	99.87 % ± 0.06 %
E4 estándar	362 500	2 000	1k+500	230	99.77 % ± 0.10 %
E4 comprimida	362 500	2 000	800+500	197	100.0 %
E4 ultraligera	99 332	1 000	500+300	44	99.7 %
E4 mínima	99 332	500	200+200	16	98.1 %

2.5.3. Análisis de Sensibilidad de Hiperparámetros

Para abordar el papel de los pesos de la pérdida λ_{BC} y λ_{Rec} y de los hiperparámetros espectrales bajo gradientes potencialmente en competencia, realizamos un barrido sistemático uni- y bidimensional alrededor de la configuración estándar de E4 (1 000 épocas de Adam, 500 de L-

BFGS, $N_c = 2000$, $\sigma = 1$, $m = 64$, ancho oculto 256, semilla 42). Cada celda de las Tablas 2.2–2.4 reporta la similaridad media en J_0 – J_3 sobre 500 puntos de evaluación en $[0, 10]$.

Tabla 2.2. Sensibilidad unidimensional del peso de recurrencia λ_{Rec} (con $\lambda_{\text{BC}} = 50$). La entrada $\lambda_{\text{Rec}} = 0$ elimina la pérdida de recurrencia; se reportan las similaridades por orden porque la media queda dominada por el colapso de J_3 .

λ_{Rec}	0	1	2	5	10	15	25	50
Sim. media (%)	–*	99.7	99.8	99.8	99.7	99.65	99.3	99.6

* $\lambda_{\text{Rec}} = 0$: $J_0 = 99.7\%$, $J_1 = 99.1\%$, $J_2 = 82.4\%$, $J_3 = -0.7\%$.

Tabla 2.3. Interacción bidimensional $\lambda_{\text{BC}} \times \lambda_{\text{Rec}}$ en las esquinas de un orden de magnitud alrededor de la configuración base. Todas las entradas son la similaridad media en J_0 – J_3 en porcentaje.

	$\lambda_{\text{BC}} = 10$	$\lambda_{\text{BC}} = 50$	$\lambda_{\text{BC}} = 100$
$\lambda_{\text{Rec}} = 1$	99.85	99.7	99.6
$\lambda_{\text{Rec}} = 15$	99.5	99.65	99.85
$\lambda_{\text{Rec}} = 50$	98.5	99.6	99.8

Tabla 2.4. Sensibilidad de la escala de Fourier σ y del número de Fourier Features m . Ambos se evalúan de forma independiente con los demás hiperparámetros fijos en sus valores por defecto.

σ	0.5	1 (por defecto)	2	4
Sim. media (%)	100.0	99.65	99.0	28.5
m	16	32	64 (por defecto)	128
Sim. media (%)	99.9	99.9	99.65	100.0

El peso de recurrencia λ_{Rec} admite una amplia meseta de aceptación en $[1, 50]$, con todos los ajustes produciendo una similaridad media superior al 99 % (Tabla 2.2). Eliminar por completo la recurrencia ($\lambda_{\text{Rec}} = 0$) hace que J_3 colapse a una similaridad negativa (-0.7%) y J_2 a 82.4 %, mientras que J_0 y J_1 se mantienen por encima del 99 %; esta asimetría confirma que la recurrencia es un acoplamiento entre órdenes no opcional para los órdenes superiores, más que un regularizador marginal. La interacción bidimensional (Tabla 2.3) muestra que, incluso ante desequilibrios extremos de un orden de magnitud en cualquier dirección, la similaridad media se mantiene por encima del 98.5 %, lo que indica que los componentes de pérdida de frontera y de recurrencia tienen gradientes *cooperativos* en lugar de competitivos.

Entre los hiperparámetros arquitectónicos, la escala de Fourier σ es el más sensible (Tabla 2.4): $\sigma = 0.5$ alcanza una similaridad perfecta, $\sigma = 1$ (usado en todo este trabajo) es estable, y $\sigma \geq 4$ destruye la convergencia. El número de Fourier Features m admite mayor flexibilidad, con

$m \geq 16$ produciendo similaridades superiores al 99.9%. Observamos además que el ancho del *backbone* $\in \{128, 256, 512\}$ produce en todos los casos una similaridad media superior al 99.6% (siendo la variante de ancho 128 entre 4 y 6× más rápida que la de ancho 512), y que el exponente del muestreo por potencias $\alpha \in \{1, 2, 4, 8\}$ varía la similaridad final en menos de 0.5 puntos porcentuales una vez presentes las Fourier Features, lo que indica que la representación espectral absorbe la mayor parte de la carga que antes recaía en el muestreo no uniforme de E2.

2.6. Discusión

2.6.1. Progresión de las Estrategias

Los resultados de las cuatro estrategias ilustran la contribución progresiva de cada técnica incorporada. En primer lugar, la formulación PINN básica (E1) demuestra que es posible aproximar funciones de Bessel usando únicamente la ecuación diferencial y las condiciones de frontera; los resultados iniciales con J_0 fueron satisfactorios, pero la extensión a J_1 y J_2 reveló las limitaciones de una arquitectura mínima con muestreo uniforme. La incorporación del remuestreo y de puntos de colocación adicionales (E2) sugiere que las PINNs base estaban limitadas tanto por la capacidad como por la distribución de los puntos de colocación, siendo el muestreo por ley de potencias particularmente efectivo para las funciones de Bessel, donde la dinámica inicial determina el comportamiento global. La mejora más significativa ocurrió en la transición a las Fourier Features (E3), confirmando que el *spectral bias* es el principal obstáculo para las funciones oscilatorias. La arquitectura Multi-Head con *backbone* compartido consolidó esta representación en los cuatro órdenes, con el beneficio adicional de compartir las características aprendidas. Por último, la relación de recurrencia (E4) aportó el refinamiento final necesario para alcanzar precisión numérica, y su efectividad radica en acoplar los cuatro órdenes, transformando cuatro problemas independientes en un sistema consistente.

Un punto que merece discusión es si la relación de recurrencia (2.9) introduce supervisión. Sostenemos que no, ya que la recurrencia se deriva algebraicamente de la ecuación diferencial (2.3) sin requerir conocimiento de la solución—es una *propiedad de la ecuación*, no de la solución—y no se utilizan valores numéricos precalculados en ningún punto del dominio. Dentro

del marco de las PINNs, la recurrencia es análoga a imponer ecuaciones auxiliares derivadas de la ecuación que gobierna el problema, una práctica estándar en los sistemas de ecuaciones.

No obstante, reconocemos que la recurrencia constituye un *sesgo inductivo* específico de las funciones de Bessel (coordenadas cilíndricas). Una estrategia más general debería derivar restricciones equivalentes a partir de la ecuación diferencial del problema particular. Para extensiones futuras—como la generalización a las funciones de Neumann u otros sistemas de coordenadas—es esencial no depender de la relación de recurrencia.

El análisis de sensibilidad de la Sección 2.5.3 aporta una perspectiva cuantitativa a esta discusión. Eliminar la pérdida de recurrencia manteniendo fijos todos los demás hiperparámetros reduce la similaridad de J_3 de 99.7 % a -0.7 % y la de J_2 de 99.6 % a 82.4 %, mientras que J_0 y J_1 se mantienen por encima del 99 %. Esta marcada asimetría revela que la recurrencia no es un regularizador uniforme, sino un *mecanismo de acoplamiento entre órdenes* que compensa las condiciones de frontera progresivamente más débiles de los órdenes superiores ($J_n(0) = 0$ y $J'_n(0) = 0$ para $n \geq 2$ no aportan información de forma más allá de la segunda derivada). Aumentar λ_{BC} por sí solo o refinar la malla de colocación no recupera este colapso, como se documenta en las Tablas 2.3 y 2.1, lo que indica que el sesgo inductivo introducido por la recurrencia es cualitativamente distinto de—y no sustituible por—los regularizadores genéricos de ajuste a datos o de suavidad.

2.6.2. Alcance del Estudio

El estudio cubre el dominio $[0, 10]$ y las funciones de primera especie J_n para $n = 0, \dots, 3$. Las estrategias E3 y E4 requieren más parámetros y tiempo de cómputo que E1 y E2, de manera consistente con las ganancias de precisión que aportan. La extensión a las funciones de Neumann Y_n y a órdenes superiores se deja como trabajo futuro, dado que las primeras requieren un tratamiento específico de la singularidad logarítmica en el origen.

Evaluamos además la configuración estándar de E4 en dominios radiales extendidos sin reentrenar: en $\rho_{\text{máx}} = 15$ la similaridad media se mantiene en 99.5 %, y en $\rho_{\text{máx}} = 20$ desciende a 97.0 %. La degradación es consistente con el hecho de que el número de ceros de J_n escala aproximadamente de forma lineal con $\rho_{\text{máx}}$, de modo que un presupuesto fijo de Fourier Features

termina por no resolver las oscilaciones de mayor frecuencia. La extensión a dominios sustancialmente mayores es directa en principio—aumentando el número de Fourier Features m o la escala σ —aunque esta última debe permanecer por debajo del umbral de colapso de la convergencia reportado en la Tabla 2.4.

2.7. Conclusiones y Trabajo Futuro

Presentamos cuatro estrategias PINN progresivas para aproximar las funciones de Bessel J_0 – J_3 , desarrolladas de manera incremental a partir de los desafíos observados en cada etapa. Un MLP sencillo con la formulación PINN básica alcanza aproximaciones razonables ($\text{MSE} \sim 10^{-4}$ a 10^{-6}), pero está limitado por el *spectral bias* en los órdenes superiores. El muestreo por ley de potencias concentra el cómputo donde la dinámica es más compleja, las Fourier Features superan el *spectral bias* al representar oscilaciones de alta frecuencia, y la arquitectura Multi-Head con *backbone* compartido consolida estas ganancias en los cuatro órdenes. La relación de recurrencia, derivada exclusivamente de la ecuación diferencial, aporta luego el acoplamiento entre órdenes y cierra la brecha de precisión restante sin introducir supervisión.

Estos resultados sugieren que la combinación de Fourier Features, arquitectura compartida y restricciones físicas auxiliares es una receta efectiva para aplicar PINNs a ecuaciones diferenciales ordinarias oscilatorias. Cuantitativamente, la configuración canónica de E4 alcanza una similaridad media de $99.87\% \pm 0.06\%$ en J_0 – J_3 sobre tres semillas aleatorias, mientras que una variante ultraligera logra 99.7% de similaridad en 44 segundos con menos de 100 000 parámetros en hardware de consumo (MacBook Air M4, solo CPU), lo que demuestra que la metodología propuesta no solo es precisa, sino también computacionalmente accesible.

Trabajo futuro. Planeamos extender esta estrategia a las funciones de Neumann Y_n (funciones de Bessel de segunda especie) utilizando el Wronskiano $W(J_n, Y_n) = 2/(\pi\rho)$ como restricción de acoplamiento y transformaciones logarítmicas para manejar la singularidad en el origen. La combinación de J_n e Y_n permitiría construir las funciones de Hankel $H_n^{(1,2)} = J_n \pm iY_n$, que representan ondas viajeras cilíndricas—completando la solución fundamental de la ecuación de onda. Adicionalmente, buscamos generalizar la metodología a otros sistemas de coordenadas.

Referencias

- Abramowitz, M., & Stegun, I. A. (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Dover Publications.
- Amos, D. E. (1986). Algorithm 644: A portable package for Bessel functions of a complex argument and nonnegative order. *ACM Transactions on Mathematical Software*, 12(3), 265-273.
- Cai, S., Mao, Z., Wang, Z., Yin, M., & Karniadakis, G. E. (2021). Physics-informed neural networks (PINNs) for fluid mechanics: A review. *Acta Mechanica Sinica*, 37(12), 1727-1738.
- Carslaw, H. S., & Jaeger, J. C. (1959). *Conduction of heat in solids* (2.^a ed.). Oxford University Press.
- Griffiths, D. J., & Schroeter, D. F. (2018). *Introduction to quantum mechanics* (3.^a ed.). Cambridge University Press.
- Jagtap, A. D., Kawaguchi, K., & Karniadakis, G. E. (2020). Adaptive activation functions accelerate convergence in deep and physics-informed neural networks. *Journal of Computational Physics*, 404, 109136.
- Krishnapriyan, A., Gholami, A., Zhe, S., Kirby, R., & Mahoney, M. W. (2021). Characterizing possible failure modes in physics-informed neural networks. *Advances in Neural Information Processing Systems*, 34, 26548-26560.
- LeVeque, R. J. (2007). *Finite difference methods for ordinary and partial differential equations*. SIAM.
- Lu, L., Meng, X., Mao, Z., & Karniadakis, G. E. (2021). DeepXDE: A deep learning library for solving differential equations. *SIAM Review*, 63(1), 208-228.
- McClenny, L. D., & Braga-Neto, U. M. (2023). Self-adaptive physics-informed neural networks. *Journal of Computational Physics*, 474, 111722.
- Morse, P. M., & Ingard, K. U. (1968). *Theoretical acoustics*. McGraw-Hill.

- Moseley, B., Markham, A., & Nissen-Meyer, T. (2023). Finite basis physics-informed neural networks (FBPINNs): A scalable domain decomposition approach for solving differential equations. *Advances in Computational Mathematics*, 49(4), 62. <https://doi.org/10.1007/s10444-023-10065-9>
- Pozar, D. M. (2011). *Microwave engineering* (4.^a ed.). John Wiley & Sons.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical recipes: The art of scientific computing* (3.^a ed.). Cambridge University Press.
- Rahaman, N., Baratin, A., Arpit, D., Draxler, F., Lin, M., Hamprecht, F. A., Bengio, Y., & Courville, A. (2019). On the spectral bias of neural networks. *Proceedings of the 36th International Conference on Machine Learning (ICML)*, 97, 5301-5310.
- Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686-707. <https://doi.org/10.1016/j.jcp.2018.10.045>
- Tancik, M., Srinivasan, P. P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J. T., & Ng, R. (2020). Fourier features let networks learn high frequency functions in low dimensional domains. *Advances in Neural Information Processing Systems (NeurIPS)*, 33, 7537-7547.
- Wang, S., Sankaran, S., & Perdikaris, P. (2022). Respecting causality is all you need for training physics-informed neural networks [Preprint, arXiv:2203.07404].
- Wang, S., Teng, Y., & Perdikaris, P. (2021). Understanding and mitigating gradient flow pathologies in physics-informed neural networks. *SIAM Journal on Scientific Computing*, 43(5), A3055-A3081.
- Wang, S., Wang, H., & Perdikaris, P. (2021). On the eigenvector bias of Fourier feature networks: From regression to solving multi-scale PDEs with physics-informed neural networks. *Computer Methods in Applied Mechanics and Engineering*, 384, 113938. <https://doi.org/10.1016/j.cma.2021.113938>
- Watson, G. N. (1922). *A treatise on the theory of Bessel functions*. Cambridge University Press.

Capítulo 3

Conclusiones y recomendaciones generales

3.1. Síntesis general de la investigación

Esta investigación se planteó un objetivo de largo alcance: determinar si una red neuronal puede generar una solución fundamental a la ecuación de onda en coordenadas cilíndricas que aproxime con alta fidelidad la solución analítica, y hacerlo con ventajas cualitativas frente a los métodos numéricos tradicionales. Para ello se adoptó el marco de las redes neuronales informadas por la física (PINN), que en lugar de discretizar el dominio incorporan el residuo de la ecuación diferencial directamente en la función de pérdida, entrenando un modelo que satisface de forma simultánea la ecuación y sus condiciones de frontera, sin necesidad de datos supervisados (Raissi et al., 2019). Esta característica es especialmente atractiva en problemas donde la solución analítica es conocida pero costosa de evaluar de forma diferenciable, ya que permite obtener una representación continua, derivable y libre de malla del fenómeno físico.

La ecuación de onda en coordenadas cilíndricas constituye un caso de estudio idóneo para esta pregunta. Al aplicar separación de variables, su componente radial conduce a la ecuación diferencial de Bessel, cuyas soluciones linealmente independientes son las funciones de Bessel de primera especie J_n y las funciones de Neumann o de segunda especie Y_n ; la combinación de ambas da lugar a las funciones de Hankel $H_n^{(1,2)} = J_n \pm i Y_n$, que describen ondas viajeras

cilíndricas (Arfken et al., 2013). Estas funciones combinan un marcado carácter oscilatorio, de amplitud lentamente decreciente, con una solución analítica perfectamente caracterizada, lo que las convierte en un banco de prueba controlado para evaluar con rigor la capacidad de las PINN de aproximar funciones de este tipo y para aislar el efecto de cada decisión de arquitectura y entrenamiento.

El núcleo de resultados de la tesis se desarrolla en el artículo científico del Capítulo 2, deliberadamente acotado a las funciones de Bessel de primera especie J_0-J_3 . Sobre ese alcance bien definido se construyó una metodología progresiva, en la que cada estrategia surgió de las limitaciones observadas en la anterior, hasta alcanzar una aproximación no supervisada de alta precisión y bajo costo computacional. A partir de esa base consolidada, el trabajo experimental de la tesis avanzó más allá del alcance de la publicación, hacia la solución fundamental completa de la componente radial, incorporando las funciones de Neumann y la construcción de las funciones de Hankel, con lo que la investigación cubre el objetivo general en su sentido más amplio.

3.2. Principales hallazgos

El primer hallazgo, y el más directo en relación con la pregunta de investigación, es que una red neuronal sí puede reproducir la solución analítica de la ecuación de onda en coordenadas cilíndricas. Para las funciones de Bessel de primera especie J_0-J_3 , la metodología propuesta alcanza una similaridad media del $99.87\% \pm 0.06\%$ respecto a la solución analítica, promediada sobre tres semillas aleatorias y utilizando exclusivamente información derivada de la ecuación diferencial, sin recurrir en ningún momento a valores precalculados de la solución. Este resultado confirma que el conocimiento físico embebido en la función de pérdida es suficiente, por sí mismo, para guiar a la red hacia la solución correcta, y responde de forma afirmativa al primer interrogante planteado en el protocolo.

El segundo hallazgo se refiere a la arquitectura idónea, y es producto de un desarrollo incremental que constituye en sí mismo una aportación. Una formulación mínima —un perceptrón multicapa sencillo con la ecuación y las condiciones de frontera— basta para capturar el comportamiento oscilatorio de los primeros órdenes, pero pierde precisión al escalar a órdenes superiores, donde las condiciones de frontera son más débiles y la dinámica cerca del origen más sutil.

La concentración de puntos de colocación mediante el muestreo por ley de potencias mitiga parcialmente este problema, pero la mejora decisiva proviene de atacar el *spectral bias*, esto es, la tendencia de las redes neuronales a aprender primero las componentes de baja frecuencia (Rahaman et al., 2019). La incorporación de *Fourier Features* (Tancik et al., 2020) permitió representar las oscilaciones de alta frecuencia desde las primeras épocas de entrenamiento y, combinada con una arquitectura Multi-Head de *backbone* compartido, consolidó la representación de los cuatro órdenes en una sola red, con el beneficio adicional de compartir las características aprendidas entre funciones. El refinamiento final lo aportó la relación de recurrencia de Bessel, que al acoplar los órdenes entre sí transforma cuatro problemas independientes en un sistema consistente y cierra la brecha de precisión restante. Que esta restricción se derive algebraicamente de la propia ecuación, sin requerir conocimiento de la solución, es lo que preserva el carácter no supervisado del método y responde al segundo interrogante sobre la estructura más efectiva de la red.

Un tercer hallazgo, de naturaleza práctica, es que esta precisión no exige grandes recursos de cómputo. Una variante ultraligera del modelo alcanza el 99.7% de similaridad en apenas 44 segundos, con menos de cien mil parámetros y empleando únicamente la unidad central de procesamiento de una computadora portátil de consumo. Más relevante que la cifra de tiempo es la naturaleza del resultado: a diferencia de las evaluaciones tabuladas o basadas en recurrencia de los métodos clásicos, que producen valores discretos e introducen error de interpolación al insertarse en procesos diferenciables, la red entrega una representación continua y derivable en todo el dominio. Esta diferenciabilidad, unida a la ausencia de malla, constituye la ventaja cualitativa del enfoque y facilita su integración en flujos de optimización de extremo a extremo, como el diseño inverso o la identificación de parámetros.

El cuarto hallazgo es el que articula la tesis con su objetivo más ambicioso y la distingue del artículo publicado. Más allá de las funciones de primera especie, el trabajo aproximó las funciones de Neumann Y_0 – Y_3 y, mediante el acoplamiento por el Wronskiano $W(J_n, Y_n) = J_n Y'_n - J'_n Y_n = 2/(\pi \rho)$ —una identidad exacta que vincula ambas familias de soluciones en todo el dominio—, entrenó de manera simultánea las redes de J_n y de Y_n , obteniendo las cuatro familias a la vez: primera especie, segunda especie y las dos funciones de Hankel. La Figura 3.1 resume este resultado. Dado que J_n e Y_n son las dos soluciones linealmente independientes de la ecuación de Bessel, su combinación en las funciones de Hankel completa la solución fundamental de la

componente radial y reproduce las ondas viajeras cilíndricas, que son precisamente la forma en que la energía se propaga en estos sistemas. Con ello, la investigación trasciende el caso particular al que se acotó la publicación y alcanza el objetivo general formulado en el protocolo.

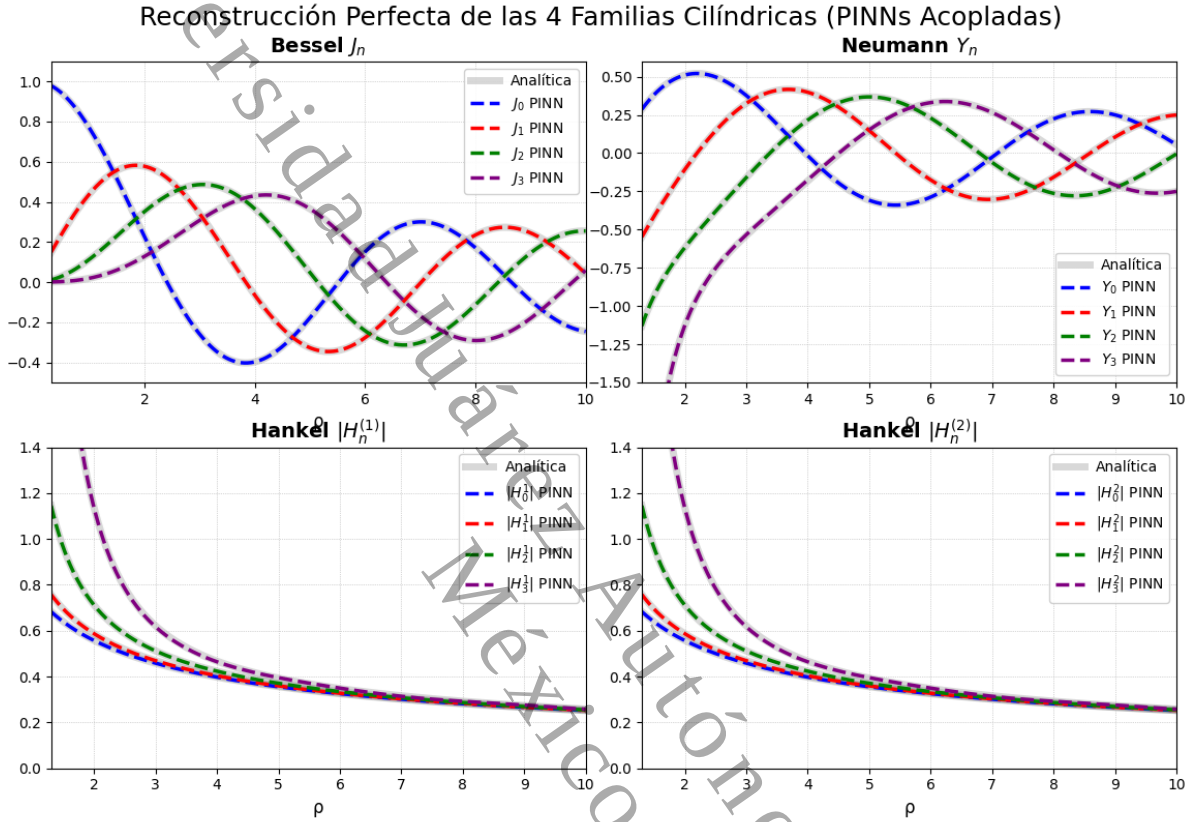


Figura 3.1. Extensión más allá del artículo: las redes de J_n e Y_n , entrenadas de forma simultánea con el acoplamiento por el Wronskiano $W(J_n, Y_n) = 2/(\pi\rho)$, aproximan las cuatro familias de soluciones de la ecuación de Bessel — primera especie J_n , segunda especie o de Neumann Y_n y las funciones de Hankel $H_n^{(1,2)} = J_n \pm i Y_n$, que representan ondas viajeras cilíndricas—, completando la solución fundamental de la componente radial de la ecuación de onda. Fuente: elaboración propia.

3.3. Relación del artículo con los objetivos de la tesis

Los resultados expuestos permiten valorar el grado de cumplimiento de los objetivos planteados en el protocolo. El artículo del Capítulo 2 satisface el objetivo específico de desarrollar y evaluar modelos basados en el conocimiento físico, así como el de analizar su desempeño en términos de precisión, velocidad de convergencia y consumo de recursos; el estudio de ablación progresivo documenta con detalle cómo cada componente contribuye a estos tres ejes. Asimismo,

atiende el objetivo de obtener una aproximación de alta fidelidad a la solución analítica conocida, al alcanzar una concordancia prácticamente total con las funciones de Bessel de referencia y cuantificarla con métricas de error verificables. La extensión a las funciones de Neumann y de Hankel, presentada en este capítulo, completa el panorama y responde al objetivo general en su formulación más amplia, que no se limitaba a una familia de soluciones sino a la ecuación de onda en coordenadas cilíndricas considerada en su conjunto.

En conjunto, la evidencia obtenida respalda la hipótesis de investigación. Las redes neuronales, a través de las PINN, aproximan con alta fidelidad la solución analítica de la ecuación de onda en coordenadas cilíndricas, y lo hacen con ventajas cualitativas frente a los métodos numéricos tradicionales: una representación continua y diferenciable, la ausencia de discretización del dominio y un proceso de aprendizaje que no depende de datos etiquetados, sino del propio contenido físico de la ecuación. La pregunta sobre la arquitectura idónea encuentra también respuesta concreta en la combinación de representaciones espectrales, arquitectura compartida y restricciones físicas derivadas de la ecuación, una combinación reproducible y aplicable a otros problemas oscilatorios de naturaleza similar.

3.4. Comparación con los métodos numéricos tradicionales

La componente radial de la ecuación de onda puede resolverse de forma eficaz con métodos numéricos clásicos. Para situar la contribución de esta tesis y evitar conclusiones apresuradas, se resolvió la misma ecuación de Bessel con dos métodos tradicionales —diferencias finitas, planteadas como un problema de valores en la frontera, e integración de Runge–Kutta, como un problema de valores iniciales—, contrastando sus resultados con la solución analítica (Burden y Faires, 2011; G. D. Smith, 1985).

Para una única resolución, los métodos numéricos son a la vez muy precisos y muy rápidos. La integración de Runge–Kutta alcanzó errores del orden de 10^{-9} en milisegundos, y las diferencias finitas, con una malla suficientemente fina, descendieron hasta el orden de 10^{-6} en tiempos igualmente reducidos. Por tanto, la aportación de las PINN no debe entenderse como una superioridad en exactitud puntual o en velocidad de cálculo, sino en la *naturaleza* de la solución que cada enfoque produce.

La diferencia esencial es que un método numérico entrega una solución *discreta y puntual*, ligada a una malla o a un conjunto de pasos: conoce el valor de J_n únicamente en sus nodos, de modo que obtener el valor en un punto intermedio exige interpolar —lo que introduce error adicional— o volver a resolver el problema. La Figura 3.2 ilustra este carácter discreto y, además, la dependencia del error respecto de la resolución de la malla: para reducir el error de las diferencias finitas de 4.5×10^{-3} (con $N = 50$ nodos) a 2.7×10^{-6} (con $N = 2\,000$) es necesario resolver un sistema sustancialmente mayor. La red, en cambio, una vez entrenada constituye un modelo continuo que se evalúa en cualquier punto del dominio sin recurrir a una malla ni repetir el cálculo. A ello se suma el tratamiento de las derivadas: mientras que en el enfoque numérico deben aproximarse mediante nuevas diferencias finitas sobre la solución discreta —acumulando error—, la PINN las proporciona por diferenciación automática, exactas respecto al modelo aprendido. La Tabla 3.1 resume estas diferencias.

Tabla 3.1. Comparación cualitativa entre los métodos numéricos tradicionales y el enfoque PINN para la ecuación de Bessel. Fuente: elaboración propia.

Aspecto	Métodos numéricos (FDM/RK)	PINN (este trabajo)
Tipo de solución	Valores discretos en nodos o puntos	Función continua en todo el dominio
Evaluación en un punto nuevo	Interpolar o volver a resolver	Evaluación directa del modelo
Malla	Requerida	No requiere malla
Derivadas	Por diferencias finitas (error adicional)	Exactas por diferenciación automática
Mayor resolución	Re-resolver un sistema mayor	El mismo modelo ya entrenado
Varios órdenes J_n	Un cálculo por cada orden	Un único modelo paramétrico
Datos de entrenamiento	No aplica	No supervisado (solo la ecuación)
Exactitud y tiempo en una corrida	Muy altos (10^{-6} – 10^{-9} en ms)	Altos (~99.9 % de similitud), con entrenamiento previo

Para dimensionar esta diferencia con datos concretos, el modelo final del artículo —una sola red de 362 500 parámetros— representa de manera simultánea los cuatro órdenes J_0 – J_3 como funciones continuas en todo el dominio, con una similaridad cercana al 100 % frente a la solución analítica. Una vez entrenado, evaluar los cuatro órdenes en 10 000 puntos toma alrededor de 13 ms

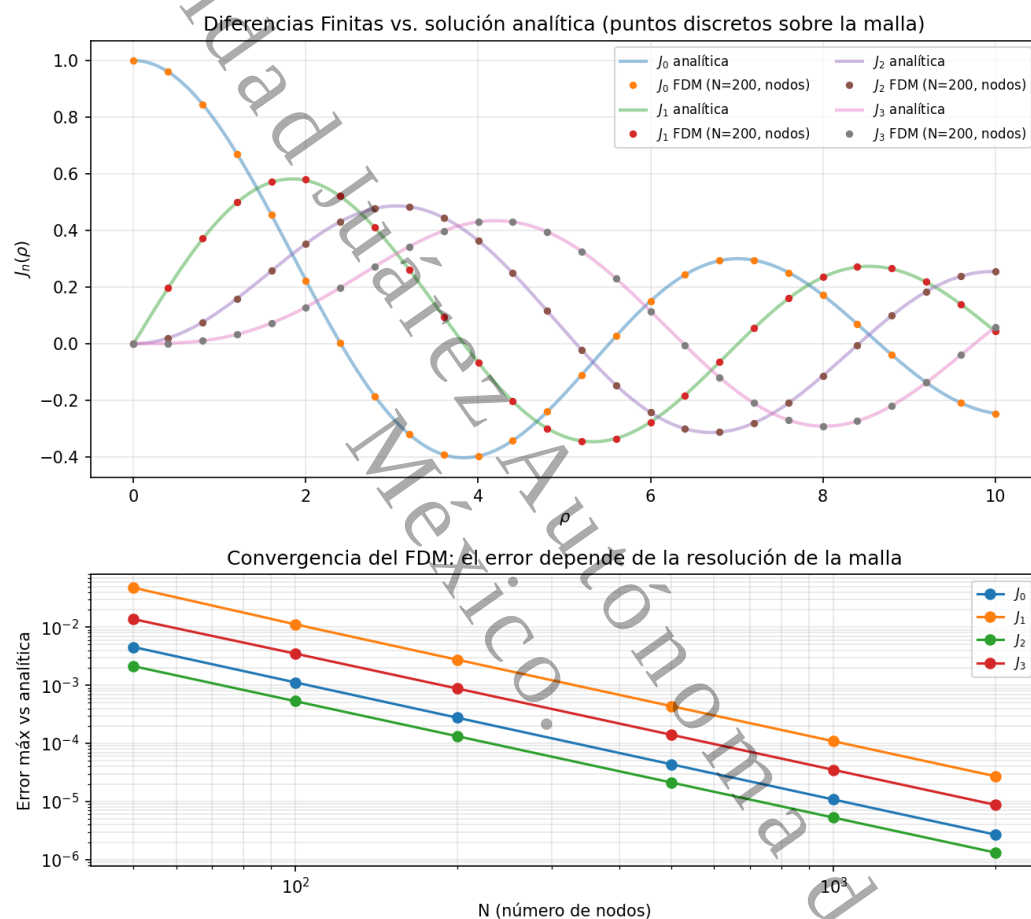


Figura 3.2. Solución de la ecuación de Bessel por diferencias finitas. Arriba: los valores numéricos existen solo en los nodos de la malla (puntos) sobre la solución analítica continua (líneas). Abajo: el error máximo frente a la solución analítica disminuye al refinar la malla, lo que evidencia la dependencia de la precisión respecto de la resolución, a costa de resolver un sistema mayor. Fuente: elaboración propia.

en CPU, y sus derivadas exactas —obtenidas por diferenciación automática— se calculan en una sola pasada, satisfaciendo identidades como $J'_0(\rho) = -J_1(\rho)$ sin aproximación adicional; más aún, una variante ligera de la misma arquitectura alcanza una calidad equivalente entrenándose en 44 segundos con menos de 100 000 parámetros. Reproducir ese mismo conjunto por la vía numérica —cuatro funciones continuas, con sus derivadas, en todo el dominio— exigiría una resolución independiente por cada orden, cada una limitada a su malla, más la interpolación y el nuevo diferenciado necesarios para los valores y las pendientes en los puntos intermedios.

En síntesis, la comparación no es una competencia por la menor cifra de error —en una sola corrida, los métodos numéricos son rápidos y muy precisos—, sino el reconocimiento de dos objetos distintos: el método numérico resuelve un caso concreto en un conjunto de puntos, en tanto que la PINN aprende una representación continua y diferenciable de la solución, reutilizable en todo el dominio y especialmente valiosa cuando dicha representación debe integrarse en procesos de cómputo diferenciables, como los problemas inversos o la optimización de extremo a extremo, o cuando se extiende a escenarios —altas dimensiones, ausencia de solución analítica— en los que los métodos basados en malla se vuelven costosos.

3.5. Aportes teóricos, metodológicos y prácticos

En el plano metodológico, esta tesis ofrece un procedimiento sistemático y reproducible para construir PINN efectivas frente a problemas oscilatorios. El desarrollo progresivo aísla la contribución de cada técnica —capacidad de la red, distribución de los puntos de colocación, representación espectral mediante *Fourier Features*, arquitectura compartida y restricciones algebraicas de acoplamiento— y convierte ese razonamiento en directrices accionables. No se ofrece un único modelo cerrado, sino una receta que muestra cómo conducir una formulación ingenua hasta una solución con un error por debajo del 0.2 %, justificando cada decisión por la limitación que viene a superar. La reproducción de ese procedimiento es posible en sus términos exactos, pues el artículo consigna las arquitecturas, los hiperparámetros de cada estrategia, el esquema de muestreo y las semillas empleadas.

En el plano teórico, el trabajo demuestra que la estructura matemática del propio problema puede convertirse en señal de aprendizaje sin sacrificar el carácter no supervisado del méto-

do. La relación de recurrencia, lejos de ser un simple regularizador, opera como un mecanismo de acoplamiento entre órdenes que compensa la debilidad creciente de las condiciones de frontera; y el Wronskiano aporta un acoplamiento físico exacto, válido en todo el dominio, capaz de ligar la segunda especie con la primera sin un solo dato de la solución. Este principio — transformar las identidades algebraicas que la ecuación impone en restricciones que potencian el entrenamiento— es reutilizable mucho más allá del caso de las funciones de Bessel.

En la práctica, un único modelo entrega los cuatro órdenes con precisión cercana a la analítica y derivadas exactas: una representación continua, diferenciable y libre de malla de las soluciones de la ecuación de onda en coordenadas cilíndricas, obtenida de forma no supervisada y reproducible en hardware de consumo, lista para integrarse en sistemas de simulación más amplios donde la diferenciabilidad de la solución es un requisito y no un lujo.

Por encima de cada aporte puntual, el trabajo amplía el uso de las PINN como solucionadores diferenciables de funciones especiales: muestra que funciones tan exigentes como las de Bessel —y, mediante el Wronskiano, las de Neumann y Hankel— pueden aprenderse sin supervisión, con un error por debajo del 0.2 % y de forma reutilizable. Si estas familias admiten ese tratamiento, otras funciones especiales del mismo tipo deberían admitirlo también; ahí radica el interés del enfoque para los problemas en los que los métodos basados en malla se encarecen.

3.6. Limitaciones

Los resultados de este trabajo tienen límites bien definidos. El estudio se circunscribe al dominio radial $[0, 10]$ y a los órdenes $n = 0, \dots, 3$; aunque las evaluaciones en dominios extendidos muestran que la degradación de la precisión es gradual y predecible, la escalabilidad a dominios mucho mayores y a órdenes superiores requiere ajustar el presupuesto espectral del modelo y queda como una cuestión abierta. Por otra parte, la relación de recurrencia empleada para las funciones de Bessel, aun siendo eficaz, constituye un sesgo inductivo específico de esta familia de funciones; su efectividad no garantiza por sí sola que la estrategia se traslade sin modificaciones a otros contextos, pues una solución plenamente general debería derivar restricciones equivalentes a partir de la ecuación de cada problema particular. Reconocer estas fronteras no resta valor a los hallazgos, sino que precisa el espacio en el que son válidos y señala con claridad

las direcciones en las que conviene continuar.

3.7. Trabajo futuro

La continuación natural de esta investigación apunta hacia la generalización del enfoque. Una primera línea consiste en extender la metodología a otros sistemas de coordenadas —esféricas, elípticas y, en general, aquellos en los que la separación de variables produce ecuaciones con funciones especiales asociadas—, de modo que las técnicas aquí validadas para el caso cilíndrico se conviertan en un repertorio aplicable a una clase más amplia de problemas de la física matemática. En esta misma dirección, resulta deseable avanzar hacia una formulación capaz de tratar familias de ecuaciones de manera unificada, en la que el modelo no se especialice en una solución concreta sino que aprenda a resolver, a partir de la estructura de la ecuación, distintos regímenes y órdenes con una sola arquitectura. Alcanzar ese grado de generalización —que un mismo marco de aprendizaje resuelva, sin rediseño, ecuaciones diferenciales de distinta naturaleza— es la meta que orienta el trabajo por venir. De forma complementaria, la extensión a dominios radiales mayores y a órdenes superiores, junto con el ajuste sistemático del número y la escala de las *Fourier Features*, permitirá cuantificar los límites de precisión del enfoque y afinar su eficiencia.

3.8. Recomendaciones

A partir de la experiencia acumulada, se derivan algunas recomendaciones para quien aborde problemas semejantes. Ante funciones de carácter oscilatorio, conviene atacar el *spectral bias* antes que cualquier otra cosa: la incorporación de *Fourier Features* resultó ser la decisión de mayor impacto sobre la precisión, por encima de aumentar la profundidad o el ancho de la red, de manera que invertir primero en una representación espectral adecuada suele rendir más que escalar ciegamente el tamaño del modelo. Es igualmente aconsejable examinar la familia de funciones objetivo en busca de identidades algebraicas —relaciones de recurrencia, Wronskianos u otras propiedades deducibles de la ecuación— y emplearlas como restricciones físicas, pues aportan acoplamiento y consistencia sin introducir datos de la solución ni comprometer la naturaleza no

supervisada del aprendizaje. Cuando el problema admita resolver varios órdenes o componentes de forma conjunta, una arquitectura compartida ofrece mayor eficiencia y favorece la reutilización de representaciones entre soluciones afines. Por último, se recomienda reportar siempre el costo computacional —tiempo de entrenamiento y número de parámetros— junto con las métricas de precisión, de modo que la utilidad práctica del modelo pueda valorarse de forma realista frente a las alternativas numéricas tradicionales.

Referencias

- Arfken, G. B., Weber, H. J., & Harris, F. E. (2013). *Mathematical methods for physicists* (7.^a ed.). Elsevier.
- Burden, R. L., & Faires, J. D. (2011). *Análisis numérico* (9.^a ed.). Cengage Learning.
- Rahaman, N., Baratin, A., Arpit, D., Draxler, F., Lin, M., Hamprecht, F. A., Bengio, Y., & Courville, A. (2019). On the spectral bias of neural networks. *Proceedings of the 36th International Conference on Machine Learning (ICML)*, 97, 5301-5310.
- Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686-707. <https://doi.org/10.1016/j.jcp.2018.10.045>
- Smith, G. D. (1985). *Numerical solution of partial differential equations: Finite difference methods* (3.^a ed.). Oxford University Press.
- Tancik, M., Srinivasan, P. P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J. T., & Ng, R. (2020). Fourier features let networks learn high frequency functions in low dimensional domains. *Advances in Neural Information Processing Systems (NeurIPS)*, 33, 7537-7547.

Alojamiento de la Tesis en el Repositorio Institucional	
Título de la tesis:	Soluciones fundamentales de la ecuación de onda en coordenadas cilíndricas utilizando Redes Neuronales
Autor:	Gabriel Omar Dominguez Ruiz
ORCID:	https://orcid.org/0009-0007-6256-2841
Resumen: Las redes neuronales han impulsado los avances recientes del aprendizaje automático y, más allá de sus usos habituales, hoy se exploran como una vía para resolver ecuaciones diferenciales. En esa frontera destacan las Redes Neuronales Informadas por la Física (PINN), que aprenden la solución de una ecuación <i>sin datos supervisados</i>: incorporan las leyes físicas directamente en la función de pérdida y entrenan la red para satisfacer a la vez la ecuación y sus condiciones de frontera. Esta tesis lleva dicho enfoque a la ecuación de onda en coordenadas cilíndricas, cuya componente radial conduce a la ecuación de Bessel, un caso exigente por su comportamiento oscilatorio y de solución analítica conocida. El núcleo del trabajo, publicado como artículo científico, propone cuatro estrategias progresivas y no supervisadas que culminan en una arquitectura Multi-Head con <i>Fourier Features</i> y la relación de recurrencia como restricción física; con ella, una sola red aproxima los cuatro órdenes J_0 a J_3 con una similaridad del 99.87% frente a la solución analítica, y una variante ligera alcanza el 99.7% en 44 segundos con menos de 100 000 parámetros. La investigación va más allá del artículo: extiende el método a las funciones de Neumann y, mediante el Wronskiano, construye las funciones de Hankel —las ondas viajeras que completan la solución fundamental de la ecuación de onda—. Más que una tabla de valores en puntos discretos, la red entrega un modelo continuo y diferenciable, evaluable en todo el dominio: esa propiedad es la que vuelve competitivo el enfoque frente a los métodos numéricos clásicos cuando la solución debe insertarse en cálculos posteriores basados en gradiente.	
Palabras clave: Redes Neuronales Informadas por la Física, Funciones de Bessel, Ecuación de onda, Coordenadas cilíndricas, Aprendizaje no supervisado.	
Referencias citadas:	Se muestran en las páginas siguientes, en la Bibliografía de la tesis (véase la página 59).

Bibliografía

- Abramowitz, M., & Stegun, I. A. (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Dover Publications.
- Aki, K., & Richards, P. G. (2002). *Quantitative seismology* (2.^a ed.). University Science Books.
- Amos, D. E. (1986). Algorithm 644: A portable package for Bessel functions of a complex argument and nonnegative order. *ACM Transactions on Mathematical Software*, 12(3), 265-273.
- Arfken, G. B., Weber, H. J., & Harris, F. E. (2013). *Mathematical methods for physicists* (7.^a ed.). Elsevier.
- Boyce, W. E., & DiPrima, R. C. (2012). *Ecuaciones diferenciales y problemas con valores en la frontera* (10.^a ed.). John Wiley & Sons.
- Burden, R. L., & Faires, J. D. (2011). *Análisis numérico* (9.^a ed.). Cengage Learning.
- Cai, S., Mao, Z., Wang, Z., Yin, M., & Karniadakis, G. E. (2021). Physics-informed neural networks (PINNs) for fluid mechanics: A review. *Acta Mechanica Sinica*, 37(12), 1727-1738.
- Carlsaw, H. S., & Jaeger, J. C. (1959). *Conduction of heat in solids* (2.^a ed.). Oxford University Press.
- Chapra, S. C., & Canale, R. P. (2015). *Métodos numéricos para ingenieros* (7.^a ed.). McGraw-Hill Education.
- E, W., Han, J., & Jentzen, A. (2017). Deep learning-based numerical methods for high-dimensional parabolic partial differential equations and backward stochastic differential equations. *Communications in Mathematics and Statistics*, 5(4), 349-380. <https://doi.org/10.1007/s40304-017-0117-6>
- Freidberg, J. P. (2007). *Plasma physics and fusion energy*. Cambridge University Press.
- Geetanjali & Hiremath, K. R. (2026). Computation of waveguide eigenmodes by physics-informed neural networks. *Machine Learning: Engineering*, 2(1), 015005. <https://doi.org/10.1088/3049-4761/ae510e>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Griffiths, D. J., & Schroeter, D. F. (2018). *Introduction to quantum mechanics* (3.^a ed.). Cambridge University Press.
- Guo, Y., Cao, X., Liu, B., & Gao, M. (2020). Solving partial differential equations using deep learning and physical constraints. *Applied Sciences*, 10(17), 5917. <https://doi.org/10.3390/app10175917>

- Jagtap, A. D., Kawaguchi, K., & Karniadakis, G. E. (2020). Adaptive activation functions accelerate convergence in deep and physics-informed neural networks. *Journal of Computational Physics*, 404, 109136.
- Jin, H., Mattheakis, M., & Protopapas, P. (2022). Physics-informed neural networks for quantum eigenvalue problems. *2022 International Joint Conference on Neural Networks (IJCNN)*, 1-8. <https://doi.org/10.1109/IJCNN55064.2022.9891944>
- Karimpouli, S., & Tahmasebi, P. (2020). Physics informed machine learning: Seismic wave equation. *Geoscience Frontiers*, 11(6), 1993-2001. <https://doi.org/10.1016/j.gsf.2020.07.007>
- Kinsler, L. E., Frey, A. R., Coppens, A. B., & Sanders, J. V. (2000). *Fundamentals of acoustics* (4.^a ed.). John Wiley & Sons.
- Kreyszig, E. (2011). *Advanced engineering mathematics* (10.^a ed.). John Wiley & Sons.
- Krishnapriyan, A., Gholami, A., Zhe, S., Kirby, R., & Mahoney, M. W. (2021). Characterizing possible failure modes in physics-informed neural networks. *Advances in Neural Information Processing Systems*, 34, 26548-26560.
- Lagaris, I. E., Likas, A., & Fotiadis, D. I. (1998). Artificial neural networks for solving ordinary and partial differential equations. *IEEE Transactions on Neural Networks*, 9(5), 987-1000. <https://doi.org/10.1109/72.712178>
- LeVeque, R. J. (2007). *Finite difference methods for ordinary and partial differential equations*. SIAM.
- Lu, L., Meng, X., Mao, Z., & Karniadakis, G. E. (2021). DeepXDE: A deep learning library for solving differential equations. *SIAM Review*, 63(1), 208-228.
- Marcus, G., & Davis, E. (2019). *Rebooting AI: Building artificial intelligence we can trust*. Pantheon Books.
- McClenny, L. D., & Braga-Neto, U. M. (2023). Self-adaptive physics-informed neural networks. *Journal of Computational Physics*, 474, 111722.
- Morse, P. M., & Ingard, K. U. (1968). *Theoretical acoustics*. McGraw-Hill.
- Moseley, B. (2022). *Physics-informed machine learning: from concepts to real-world applications* [DPhil thesis]. University of Oxford.
- Moseley, B., Markham, A., & Nissen-Meyer, T. (2020). Solving the wave equation with physics-informed deep learning [Preprint, arXiv:2006.11894]. <https://arxiv.org/abs/2006.11894>
- Moseley, B., Markham, A., & Nissen-Meyer, T. (2023). Finite basis physics-informed neural networks (FBPINNs): A scalable domain decomposition approach for solving differential equations. *Advances in Computational Mathematics*, 49(4), 62. <https://doi.org/10.1007/s10444-023-10065-9>
- Pozar, D. M. (2011). *Microwave engineering* (4.^a ed.). John Wiley & Sons.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical recipes: The art of scientific computing* (3.^a ed.). Cambridge University Press.

- Rahaman, N., Baratin, A., Arpit, D., Draxler, F., Lin, M., Hamprecht, F. A., Bengio, Y., & Courville, A. (2019). On the spectral bias of neural networks. *Proceedings of the 36th International Conference on Machine Learning (ICML)*, 97, 5301-5310.
- Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378, 686-707. <https://doi.org/10.1016/j.jcp.2018.10.045>
- Rasht-Behesht, M., Huber, C., Shukla, K., & Karniadakis, G. E. (2022). Physics-informed neural networks (PINNs) for wave propagation and full waveform inversions [e2021JB023120 2021JB023120]. *Journal of Geophysical Research: Solid Earth*, 127(5), e2021JB023120. <https://doi.org/10.1029/2021JB023120>
- Russell, S., & Norvig, P. (2016). *Artificial intelligence: A modern approach* (3.^a ed.). Pearson Education.
- Simmons, G. F. (1991). *Differential equations with applications and historical notes* (2.^a ed.). McGraw-Hill.
- Smith, G. D. (1985). *Numerical solution of partial differential equations: Finite difference methods* (3.^a ed.). Oxford University Press.
- Smith, J. D., Azizzadenesheli, K., & Ross, Z. E. (2021). EikoNet: Solving the eikonal equation with deep neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, 59(12), 10685-10696. <https://doi.org/10.1109/TGRS.2020.3039165>
- Strauss, W. A. (2007). *Partial differential equations: An introduction* (2.^a ed.). John Wiley & Sons.
- Streeter, V. L., Wylie, E. B., & Bedford, K. W. (1998). *Fluid mechanics* (9.^a ed.). McGraw-Hill.
- Tancik, M., Srinivasan, P. P., Mildenhall, B., Fridovich-Keil, S., Raghavan, N., Singhal, U., Ramamoorthi, R., Barron, J. T., & Ng, R. (2020). Fourier features let networks learn high frequency functions in low dimensional domains. *Advances in Neural Information Processing Systems (NeurIPS)*, 33, 7537-7547.
- Thurey, N., Holl, P., Mueller, M., Schnell, P., Trost, F., & Um, K. (2022). Physics-based deep learning [Preprint, arXiv:2109.05237]. <https://arxiv.org/abs/2109.05237>
- Timoshenko, S. P., & Goodier, J. N. (1970). *Theory of elasticity* (3.^a ed.). McGraw-Hill.
- Tipler, P. A., & Mosca, G. (2010). *Física para la ciencia y la tecnología*. Reverté.
- Wang, J.-X., Wu, J.-L., & Xiao, H. (2017). Physics-informed machine learning approach for reconstructing Reynolds stress modeling discrepancies based on DNS data. *Physical Review Fluids*, 2(3), 034603. <https://doi.org/10.1103/PhysRevFluids.2.034603>
- Wang, S., Sankaran, S., & Perdikaris, P. (2022). Respecting causality is all you need for training physics-informed neural networks [Preprint, arXiv:2203.07404].
- Wang, S., Teng, Y., & Perdikaris, P. (2021). Understanding and mitigating gradient flow pathologies in physics-informed neural networks. *SIAM Journal on Scientific Computing*, 43(5), A3055-A3081.
- Wang, S., Wang, H., & Perdikaris, P. (2021). On the eigenvector bias of Fourier feature networks: From regression to solving multi-scale PDEs with physics-informed neural networks. *Computer*

- Methods in Applied Mechanics and Engineering*, 384, 113938. <https://doi.org/10.1016/j.cma.2021.113938>
- Watson, G. N. (1922). *A treatise on the theory of Bessel functions*. Cambridge University Press.
- Zill, D. G. (2018). *Ecuaciones diferenciales con problemas con valores en la frontera* (9.^a ed.). Cengage Learning.
- Zou, Z., & Karniadakis, G. E. (2025). Multi-head physics-informed neural networks for learning functional priors and uncertainty quantification. *Journal of Computational Physics*, 531, 113947. <https://doi.org/10.1016/j.jcp.2025.113947>