
Physics-constrained Plane Wave Decomposition Network: Solving the Helmholtz Equation in Airborne Acoustics

James Hipperson ^{1,2}

j.r.hipperson@edu.salford.ac.uk

Trevor J. Cox ¹

t.j.cox@salford.ac.uk

Jonathan A. Hargreaves ¹

j.a.hargreaves@salford.ac.uk

¹Acoustics Research Centre, University of Salford, ²Funktion-One Research Ltd.

Abstract

Frequency-domain simulations are very common in Acoustics, which requires solving the Helmholtz equation (the frequency-domain equivalent of the wave equation). Like most other partial differential equations (PDEs), conventional solutions use numerical approaches such as Finite Element (FEM) and Boundary Element (BEM) Methods. These are fast for small domains or low frequencies, but they become computationally intractable for large domains or high frequencies, such as simulation over the full audible frequency range (20 Hz - 20 kHz) in room acoustics or sound propagation outdoors. Neural networks that respect known physics are an emerging alternative to numerical methods. However, physics enforcement and interpretability are significant problems. In our method, we enforce physics constraints by the use of a linear plane wave decoder stage, because plane waves are solutions to the Helmholtz equation and form a differentiable and complete orthogonal basis for single frequency soundfields. The plane wave network was compared to a CNN and FNO on the problem of predicting the scattered field from a cylinder in 2D. Mean absolute error (MAE) was similar compared to FNO and grid artefacts/noise are eliminated. FNO had a lower MAE at low frequencies and higher MAE at high frequencies. CNN was able to learn some aspects of the solution, but failed entirely at 4 kHz.

1 Background

1.1 Frequency domain acoustics

Partial Differential Equations (PDEs) are the foundation of modern science and engineering. However, they are difficult to solve analytically except in the simplest cases. Numerical methods are the conventional solution approach, or simplified methods that ignore wave physics such as geometrical acoustics. The PDE for Acoustics is the wave equation, and its time-harmonic equivalent is the Helmholtz Equation. Because solutions to the Helmholtz equation are by definition for single frequencies, plane waves are solutions to the Helmholtz equation in Cartesian coordinates. k-space is the Fourier transform of a single frequency spatial wavefield, where the energy is concentrated on a ring (in two dimensions) or a sphere (in three dimensions) of radius k_0 (the wavenumber). The angular position of a mode on the ring or sphere corresponds to a plane wave propagating in that angular direction. Plane waves therefore form a complete orthogonal basis for wavefields, with k-space providing a convenient and sparse representation - see Williams [1999].

1.2 Properties of k-space

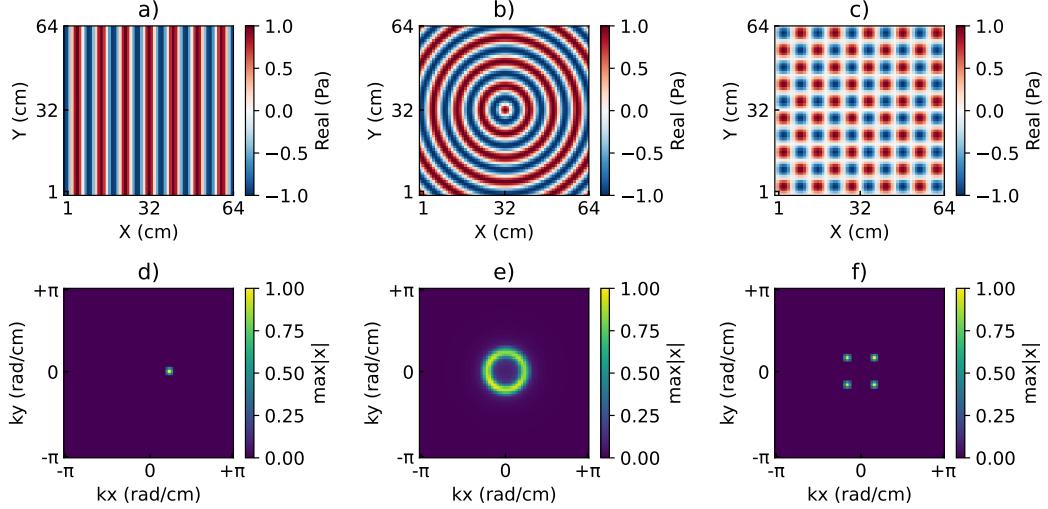


Figure 1: 2D spatial wavefields (top row) and corresponding k-space fields (bottom row)

Figure 1 shows the real part of some simple spatial wavefields at $k_0 \approx 0.69$ and the corresponding Fourier-transformed k-space fields. a) is a plane wave propagating in the $+x$ direction. b) is a point source at 32,32. c) is the 10,10 mode of the 64,64 domain with Dirichlet boundary conditions. On the bottom row are the normalised absolute values of the corresponding k-space fields. This demonstrates a key property of k-space, that the spatial frequency field is sparser than the spatial wavefield. The structure and features remain consistent at all frequencies, simply scaled by k_0 . These properties can be exploited in machine learning as dimensionality reduction and to provide interpretability.

1.3 Frequency domain neural networks

There are several examples of neural networks for image data that operate in the spatial frequency domain. Wu et al. [2023] utilised hybrid spatial domain/frequency domain convolution blocks to overcome the problem of spectral bias in image super-resolution. In search of improved image restoration performance, Shah et al. [2024] applied complex-valued convolutional layers to learn in the spatial frequency domain and demonstrated improved results over real-valued CNNs operating in the spatial domain only. But complex-valued neural networks need further development in existing machine learning libraries. Much of the theory is well established. Benvenuto and Piazza [1992] and Georgiou and Koutsougeras [1992] presented complex-valued backpropagation algorithms, and Bassey et al. [2021], Abdalla [2023] and Barrachina et al. [2023] are some more recent reviews.

K-space is also well known in medical imaging, particularly Magnetic Resonance Imaging (MRI) and Computed Tomography (CT). Han et al. [2020] demonstrates a CNN operating on complex-valued k-space data, with the loss function computed in the image (spatial) domain. Only a small improvement in accuracy is reported, but inference is significantly faster than conventional linear methods. Du et al. [2021] recognises the importance of truncating spatial frequency data to focus on the centre of the k-space domain, as the majority of the k-space domain does not contribute significantly to the (spatial) image domain. Most recently, Hedayatrasa et al. [2025] demonstrated a PINN (physics-informed neural network) operating in k-space for a vibroacoustics application. Compared to a standard PINN, the k -PINN outperforms in terms of learning and computational efficiency. The most similar network design to that proposed below is Alkhalifah and Huang [2024] which parameterises Gabor basis functions to solve the elastic (geophysics) Helmholtz equation. A related approach is Wang et al. [2020] using plane waves as a periodic activation function, a similar idea to Sitzmann et al. [2020]. A different method of utilising the frequency domain is the Fourier Neural Operator (FNO) (Li et al. [2021]), designed to learn highly complex and non-linear maps between function spaces, such as the solution operator for a PDE. FNO parameterises Fourier modes in a higher-dimensional latent space, exploiting the efficiency of FFTs and convolution theorem.

2 Method

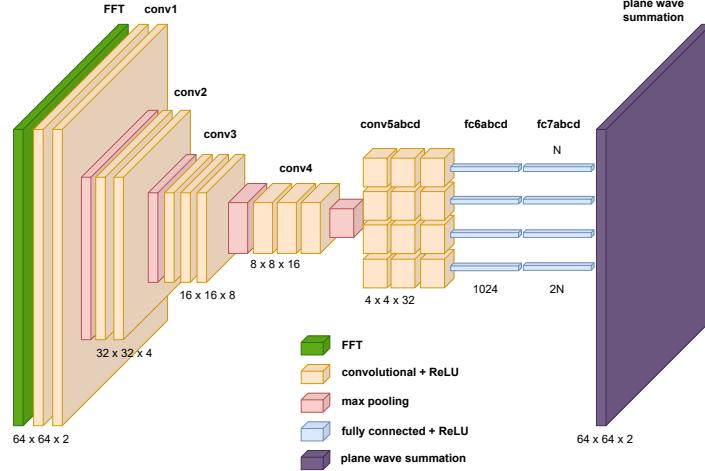


Figure 2: Network architecture

In this work, we integrate frequency domain concepts into a neural network for airborne acoustics.

Fig. 2 shows the network architecture, with single 2D input FFT, a convolutional encoder stage, four branches each with a convolutional stage and a fully connected stage for the plane wave parameters (see Table 1). Input and output are $64 \times 64 \times 2$ grids representing complex-valued pressure. Weights and activations are real, which may not be ideal. Making the network complex end-to-end may provide some additional constraint or guidance. Otherwise, the network has to learn the meaning of the two data channels. The network can learn without the FFT layer, but with diminished performance.

2.1 Physics-constrained output layer (superposition of plane waves)

The summation of plane waves is differentiable, facilitating learning via backpropagation and stochastic gradient descent. There are four learnable parameters for each mode; amplitude A , phase ϕ encoded as $\cos(\phi)$ and $\sin(\phi)$, the propagating direction of the mode encoded as $\cos(\theta)$ and $\sin(\theta)$, and decay terms α_x and α_y . The angular components (ϕ and θ) are mapped to radians using $\text{atan}(y, x)$. Amplitude and Decay use sigmoid activation to limit between 0 and 1. Angular components use tanh activation to limit between -1 and 1. k_0 is provided directly to the plane wave layer as a model input. 64 modes were used in the experiment.

Eq. 1 shows the general function of the plane wave layer for N modes, where α_x and α_y can be adjusted between 0 and 1 to produce propagating or evanescent modes:

$$\sum_{i=1}^N A_i e^{-(\alpha_{xi}X + \alpha_{yi}Y)} e^{j(k_{xi}X + k_{yi}Y + \phi_i)} \quad (1)$$

Where $k_{xi} = k_0 \cos(\theta_i)$ and $k_{yi} = k_0 \sin(\theta_i)$, and X and Y are the grid coordinates. X and Y are independent of the input grid resolution, allowing prediction at arbitrary resolutions.

Table 1: Learnable plane wave parameters

	Amplitude (A)	Phase (ϕ)	Propagating direction (θ)	Decay (α_x, α_y)
Size	N_{modes}	$2 \cdot N_{modes}$	$2 \cdot N_{modes}$	$2 \cdot N_{modes}$
Range	0 to 1	-1 to 1	-1 to 1	0 to 1

Performance was compared between an image-to-image CNN, FNO and the plane wave network (PWN) on the task of predicting the scattered field from a cylinder, given the incident wave and a mask of the scatterer. Using the analytical solution for scattering from an infinite cylinder, the radius of the scatterer and incident angle were varied in the training data, and four frequencies were used (0.5 kHz, 1 kHz, 2 kHz, 4 kHz) to capture a wide range of wave behaviour. The dataset consisted of 8192 samples that were split into 7680 for training, 256 for validation and 256 withheld for testing. Learning rate was reduced by half every 20 epochs. Losses were calculated in the spatial domain (with the scatterer interior excluded) using $0.9 \cdot L_2 + 0.1 \cdot L_1$. For all models, the initial learning rate was 1e-4 and batch size was 16. Weight decay (1e-5) was used, which is likely why the validation loss closely tracks the training loss, but is very slightly lower. Early stopping was implemented by saving the model state with the lowest validation loss. Due to the difficulty in making a fair comparison between architecturally different models, rather than matching parameter or layer count, the model width and modes were adjusted to approximately match training time, indicating a fair comparison in terms of computation required. Further reducing the complexity of the FNO model did not reduce training time, indicating a computational bottleneck. Training was performed on an Apple M3 CPU (without MPS backend).

3 Results

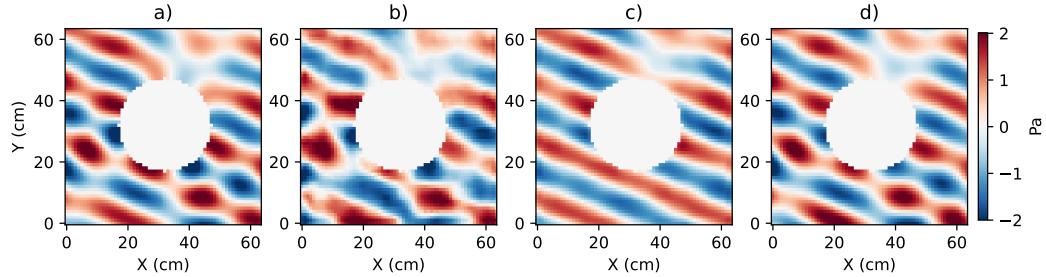


Figure 3: Real part of pressure at 2 kHz, total field (incident + scattered), a) Ground truth b) CNN c) FNO d) PWN (vectorised graphics for close inspection)

Table 2: Test set error

Model	Parameters	Training	0.5-1 kHz			2-4 kHz		
			MAE	SD	RMSE	MAE	SD	RMSE
CNN	1,159,934	0.79 hours	0.229	0.160	0.279	0.908	0.742	1.172
FNO	399,908	1.21 hours	0.194	0.184	0.267	0.498	0.347	0.608
PWN	1,266,920	1.12 hours	0.203	0.212	0.293	0.472	0.388	0.611

All models produced good quality predictions for 0.5 kHz and 1 kHz. 2 and 4 kHz presented a significant challenge due to spectral bias (Fig. 3). The CNN was unable to learn how to predict the scattered field at 4 kHz (see Fig. 7(b), Appendix). The FNO and PWN were able to make an approximation, but both struggled to learn the circularly propagating scattered component. Although PWN had a similar MAE to FNO at 2-4 kHz, a notable aspect is the absence of noise and grid-like artefacts in the PWN predictions as the grid data is not parameterised directly (Appendix).

4 Conclusions

Parameterising physical basis functions works and appears to be a promising approach for embedding physics constraints into neural networks for wave problems. However, further development is required to be more competitive with state-of-the-art architectures for scientific machine learning.

The most important next step is to investigate and compare generalisability, using scatterers of different shapes, and testing on unseen shapes and frequencies. If this shows potential, the method could be expanded to 3D acoustics tasks. Some combination of operator learning concepts and further embedded prior physics knowledge may be a successful future direction.

Acknowledgments and Disclosure of Funding

This work was supported by EPSRC and Funktion-One Research Ltd. as part of the CDT in Sustainable Sound Futures. (<https://soundfutures.salford.ac.uk/>)

References

- R. Abdalla. Complex-valued neural networks - theory and analysis. *CoRR*, abs/2312.06087, 2023. doi: 10.48550/ARXIV.2312.06087. URL <https://doi.org/10.48550/arXiv.2312.06087>.
- T. Alkhalifah and X. Huang. Physics-informed neural wavefields with Gabor basis functions. *Neural Networks*, 175:106286, July 2024. ISSN 08936080. URL <https://doi.org/10.1016/j.neunet.2024.106286>.
- J. A. Barrachina, C. Ren, G. Vieillard, C. Morisseau, and J.-P. Ovarlez. Theory and implementation of complex-valued neural networks, 2023. URL <http://arxiv.org/abs/2302.08286>.
- J. Bassey, L. Qian, and X. Li. A survey of complex-valued neural networks, 2021. URL <http://arxiv.org/abs/2101.12249>.
- N. Benvenuto and F. Piazza. On the complex backpropagation algorithm. 40(4):967–969, 1992. ISSN 1053587X. doi: 10.1109/78.127967. URL <http://ieeexplore.ieee.org/document/127967/>.
- T. Du, H. Zhang, Y. Li, S. Pickup, M. Rosen, R. Zhou, H. K. Song, and Y. Fan. Adaptive convolutional neural networks for accelerating magnetic resonance imaging via k-space data interpolation. 72:102098, 2021. ISSN 13618415. doi: 10.1016/j.media.2021.102098. URL <https://linkinghub.elsevier.com/retrieve/pii/S1361841521001444>.
- G. Georgiou and C. Koutsougeras. Complex domain backpropagation. 39(5):330–334, 1992. ISSN 10577130. doi: 10.1109/82.142037. URL <http://ieeexplore.ieee.org/document/142037/>.
- Y. Han, L. Sunwoo, and J. C. Ye. k -space deep learning for accelerated MRI. 39(2):377–386, 2020. ISSN 0278-0062, 1558-254X. doi: 10.1109/TMI.2019.2927101. URL <https://ieeexplore.ieee.org/document/8756028/>.
- S. Hedayatrasa, O. Fink, W. V. Paepegem, and M. Kersemans. k-space physics-informed neural network (k-PINN) for compressed spectral mapping and efficient inversion of vibrations in thin composite laminates. 2025.
- Z. Li, N. B. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, and A. Anandkumar. Fourier neural operator for parametric partial differential equations. In *International Conference on Learning Representations*, 2021. URL <https://openreview.net/forum?id=c8P9NQVtmn0>.
- Z. H. Shah, M. Müller, W. Hübner, H. Ortkrass, B. Hammer, T. Huser, and W. Schenck. Image restoration in frequency space using complex-valued CNNs. 7:1353873, 2024. ISSN 2624-8212. doi: 10.3389/frai.2024.1353873. URL <https://www.frontiersin.org/articles/10.3389/frai.2024.1353873/full>.
- V. Sitzmann, J. N. P. Martel, A. W. Bergman, D. B. Lindell, and G. Wetzstein. Implicit neural representations with periodic activation functions, 2020. URL <https://arxiv.org/abs/2006.09661>.
- Z. Wang, T. Cui, and X. Xiang. A neural network with plane wave activation for helmholtz equation, 2020. URL <https://arxiv.org/abs/2012.13870>.
- E. G. Williams. *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. Academic Press, San Diego, CA, 1999. ISBN 978-0-12-753960-7.
- Z. Wu, W. Liu, J. Li, C. Xu, and D. Huang. SFHN: Spatial-frequency domain hybrid network for image super-resolution. 33(11):6459–6473, 2023. ISSN 1051-8215, 1558-2205. doi: 10.1109/TCSVT.2023.3271131. URL <http://ieeexplore.ieee.org/document/10109830/>.

5 Appendix

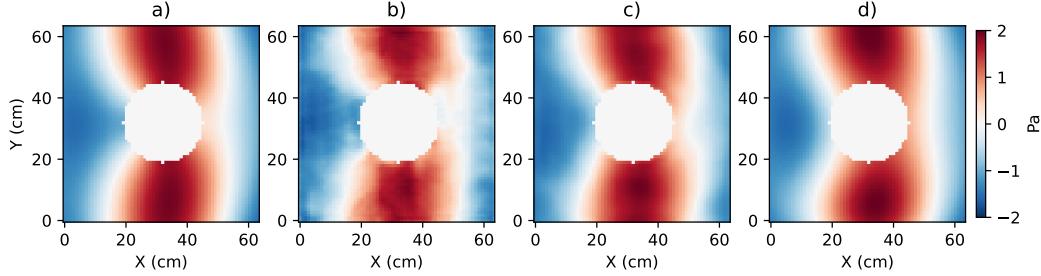


Figure 4: Real part of pressure at 0.5 kHz, total field (incident + scattered),
a) Ground truth b) CNN c) FNO d) PWN

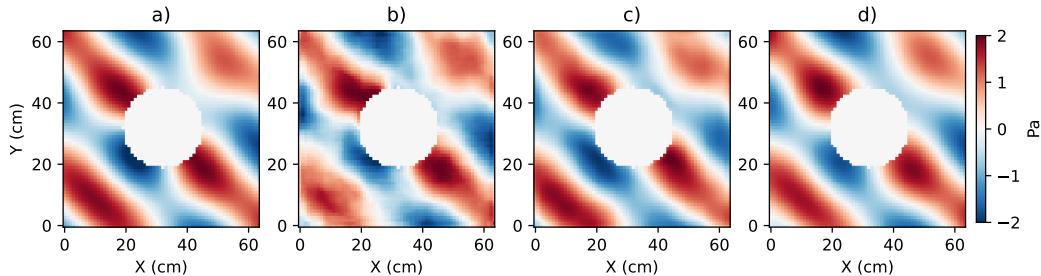


Figure 5: Real part of pressure at 1 kHz, total field (incident + scattered),
a) Ground truth b) CNN c) FNO d) PWN

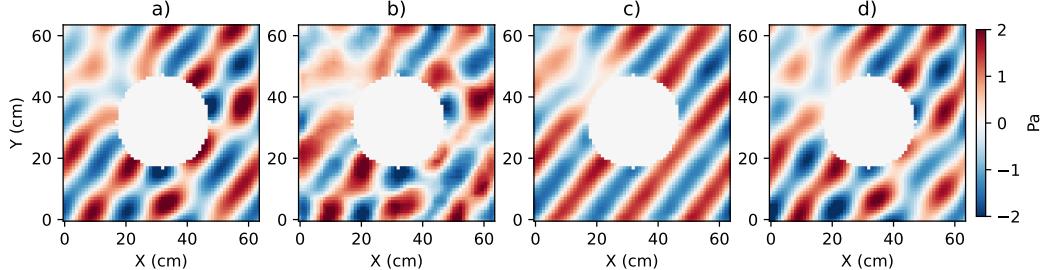


Figure 6: Real part of pressure at 2 kHz, total field (incident + scattered),
a) Ground truth b) CNN c) FNO d) PWN

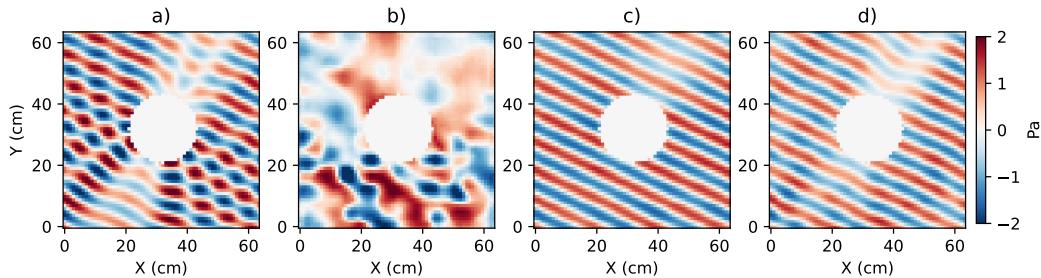


Figure 7: Real part of pressure at 4 kHz, total field (incident + scattered),
a) Ground truth b) CNN c) FNO d) PWN