doi：10.13413/j.cnki.jdxblxb.2023195

基于物理信息神经网络的光波衍射问题求解

陈旭早，袁利军
（重庆工商大学 数学与统计学院，重庆 400067）

摘要：用物理信息神经网络方法数值求解间断系数光波衍射问题. 结果表明：用光滑函数近似间断系数可大幅度提高物理信息神经网络求解精度；用物理信息神经网络求解散射场比直接求解总场效果更好. 最后通过数值实验验证理论结果的正确性.

关键词：物理信息神经网络；光波衍射；间断系数；光滑函数
中图分类号：O436 **文献标志码**：A **文章编号**：1671-5489(2024)02-0423-08

Solving Light Wave Diffraction Problem Based on Physics-Informed Neural Networks

CHEN Xuzao, YUAN Lijun
(*School of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing* 400067, *China*)

Abstract: We used the physics-informed neural networks method to numerically solve the problem of discontinuous coefficient light wave diffraction. The results show that approximating the discontinuous coefficient with a smooth function can significantly improve the accuracy of the physics-informed neural network solution. Using physics-informed neural networks to solve the scattered field is better than directly solving the total field. Finally, the correctness of the theoretical results is verified through numerical experiments.

Keywords: physics-informed neural network; light wave diffraction; discontinuous coefficient; smooth function

物理信息神经网络(PINN)方法广泛用于数值求解偏微分方程[1]：Mao 等[2]用物理信息神经网络解决了高速空气动力学流动模型的正问题和反问题；Fang 等[3]利用物理信息神经网络解决了频域 Maxwell 方程和超材料设计问题；陆至彬等[4]基于软边界和硬边界两种设定方法构建了神经网络求解传热方程；Lu 等[5]提出了一个通用的深度学习框架 DeepONet，用于学习各种连续非线性算子.

与传统偏微分方程数值求解方法相比，物理信息神经网络方法有如下优点：1）不需进行网格离散化[6]，传统方法如有限差分法和有限元法均需进行网格离散化，从而增加了计算复杂度和计算时间，而物理信息神经网络不需网格离散化过程，且可并行处理，从而提高了求解效率；2）物理信息神经网络方法可处理高维问题[7]，有效避免了“维数灾难”问题；3）物理信息神经网络方法在求解非线性方程中具有优势[8]，由于物理信息神经网络通过引入物理约束条件学习数据中的非线性关系进行求解，因此物理信息神经网络可有效解决非线性方程的求解任务；4）损失函数结构简单易于构造.

收稿日期：2023-05-16.

第一作者简介：陈旭早(1997—)，男，苗族，硕士研究生，从事深度学习的研究，E-mail：407250764@qq.com. **通信作者简介**：袁利军(1982—)，男，汉族，博士，教授，从事深度学习和科学计算的研究，E-mail：ljyuan@ctbu.edu.cn.

基金项目：重庆市自然科学基金(批准号：CSTB2022NSCQ-MSX0610).

光波衍射在实际中应用较多[9-11]，可用一个具有间断系数的 Helmholtz 方程的边值问题描述. 由于方程解在间断点不存在二阶偏导数，因此用物理信息神经网络求解时，损失函数在间断点处无法定义，效果较差. 当介质介电常数变化较大时，物理信息神经网络训练效果极差. Jagtap 等[12]提出了守恒性物理信息神经网络 cPINN，该方法将原始求解区域划分为多个子区域，并在损失函数中强化相邻子区域界面通量的守恒约束，在模拟激波面等光滑性较差甚至间断的情况下可提高精确度；Kharazmi 等[13]对 PINN 方法进行了改进，提出一种基于区域分解的变分物理信息神经网络(hp-VPINNs)方法，该方法采用变分原理描述偏微分方程问题，使用域分解处理复杂的网格结构，从而实现高效求解.

本文针对间断系数边值问题无法训练的问题，引入光滑函数近似间断系数，并将光波衍射问题分为总场和散射场的边值问题分别进行求解. 结果表明：该方法对损失函数的改动较小，并提高了计算效果；物理信息神经网络方法求解散射场的边值问题效果更好.

1 一维光栅衍射问题

1.1 理 论

考虑一个一维多层介质中的光波衍射问题，介质的介电函数 $\varepsilon(x)$ 满足

$$\varepsilon(x)=\begin{cases}\varepsilon_0, & x<D_0,\\ \varepsilon_{n+1}, & D_n<x<D_{n+1}, \quad n=0,1,2,\cdots,m-1,\\ \varepsilon_0, & x>D_m,\end{cases} \tag{1}$$

其中 m 为正整数. 在 $x<D_0$ 一侧给定一个频率为 ω 的入射波，满足

$$u^{(i)}(x)=\mathrm{e}^{\mathrm{i}(k\sqrt{\varepsilon_0}x)}, \tag{2}$$

其中 $k=\omega/c$ 为真空中的波长，c 为真空中的光速，总场 u 满足如下方程[14]：

$$\begin{cases}u''+k^2\varepsilon(x)u=0, & D_0<x<D_m,\\ u'(D_0)=-\mathrm{i}k\sqrt{\varepsilon_0}u(D_0)+2\mathrm{i}k\sqrt{\varepsilon_0}\mathrm{e}^{\mathrm{i}k\sqrt{\varepsilon_0}D_0},\\ u'(D_m)=\mathrm{i}k\sqrt{\varepsilon_0}u(D_m).\end{cases} \tag{3}$$

边值问题(3)用总场 u 描述光波衍射问题，也可用散射场 $u^{(s)}=u-u^{(i)}$ 描述光波衍射问题，散射场 $u^{(s)}$ 满足

$$\begin{cases}[u^{(s)}]''+k^2\varepsilon(x)u^{(s)}=k^2(\varepsilon_0-\varepsilon(x))u^{(i)}, & D_0<x<D_m,\\ [u^{(s)}]'(D_0)=-\mathrm{i}k\sqrt{\varepsilon_0}u^{(s)}(D_0),\\ [u^{(s)}]'(D_m)=\mathrm{i}k\sqrt{\varepsilon_0}u^{(s)}(D_m).\end{cases} \tag{4}$$

由于总场 u 和散射场 $u^{(s)}$ 在 $x=D_j(j=0,1,\cdots,m)$ 上不存在二阶导数，因此直接用物理信息神经网络方法求解边值问题(3)或(4)的效果较差，当 $\varepsilon_j(j=0,1,2,\cdots,m)$ 变化较大时，用物理信息神经网络方法求解边值问题(3)或(4)的效果极差. 为提高数值求解效果，将间断系数 $\varepsilon(x)$ 近似为光滑函数 $\varepsilon_\delta(x)$. $\varepsilon_\delta(x)$ 定义为

$$\varepsilon_\delta(x)=\begin{cases}\varepsilon_0, & D_0<x<H_1,\\ \dfrac{1}{2}\left(\dfrac{x-D_n}{\sqrt{(x-D_n)^2+\delta^2}}+1\right)(\varepsilon_{n+1}-\varepsilon_n)+\varepsilon_n, & H_n<x<H_{n+1}, \quad n=1,2,\cdots,m-1,\\ \varepsilon_m, & H_m<x<D_m,\end{cases} \tag{5}$$

其中 δ 为超参数，δ 大小会影响函数的曲率，$H_n=(D_{n+1}+D_n)/2$，$n=1,2,\cdots,m$. 将边值问题(3)和(4)中的 $\varepsilon(x)$ 替换为光滑函数 $\varepsilon_\delta(x)$ 即可得到光滑后的解，分别记为 u_δ 和 $u_\delta^{(s)}$.

1.2 物理信息神经网络求解一维光栅衍射问题

用 $\tilde{u}(x;\theta)$ 表示一个神经网络，其中 x 为神经网络的输入值，θ 为神经网络中所需学习的参数. 由于神经网络不能直接进行复数计算，令神经网络的输出为二维，分别用 $v(x;\theta)$ 和 $w(x;\theta)$ 表示，代表函数的实部和虚部，则有

$$\tilde{u}(x;\theta)=v(x;\theta)+\mathrm{i}w(x;\theta). \tag{6}$$

用 $\tilde{u}(x;\theta)$ 近似问题(3)的解 $u(x)$，设 $\{x_j\}_{j=1}^N$ 为区间 (D_0,D_m) 上 N 个观测点，问题(3)对应的损失函数可写为

$$\begin{aligned}L(\theta)=&\frac{1}{N}\sum_{j=1}^{N}|\tilde{u}''(x_j;\theta)+k^2\varepsilon(x_j)u(x_j;\theta)|^2+\left|\tilde{u}'(D_0)-(-k\sqrt{\varepsilon_0}\tilde{u}(D_0)+2k\sqrt{\varepsilon_0}\,\mathrm{e}^{\mathrm{i}k\sqrt{\varepsilon_0}D_0})\mathrm{i}\right|^2+\\&\left|\tilde{u}'(D_m)-\mathrm{i}k\sqrt{\varepsilon_0}\tilde{u}(D_m)\right|^2=\\&\frac{1}{N}\sum_{j=1}^{N}((\tilde{w}''(x_j;\theta)+k^2\varepsilon(x_j)w(x_j;\theta))^2+(\tilde{v}''(x_j;\theta)+k^2\varepsilon(x_j)v(x_j;\theta))^2)+\\&(\tilde{w}''(D_0)-k\sqrt{\varepsilon_0}\tilde{v}(D_0)+2k\sqrt{\varepsilon_0}\sin(k\sqrt{\varepsilon_0}D_0))^2+\\&(\tilde{v}'(D_0)-(-k\sqrt{\varepsilon_0}\tilde{w}(D_0)+2k\sqrt{\varepsilon_0}\cos(k\sqrt{\varepsilon_0}D_0)))^2+\\&(\tilde{w}'(D_m)+k\sqrt{\varepsilon_0}\tilde{v}(D_m))^2+(\tilde{v}'(D_m)-k\sqrt{\varepsilon_0}\tilde{v}(D_m))^2.\end{aligned}\tag{7}$$

利用优化算法求解最小值问题 $\min\limits_{\theta}L(\theta)$ 即可得到 $u(x)$ 的近似解 $\tilde{u}(x;\theta)$. 同理，对于问题(4)以及问题(3)与(4)光滑后问题的解 $u^{(s)}(x)$，$u_\delta(x)$ 和 $u_\delta^{(s)}(x)$ 分别近似为神经网络 $\tilde{u}^{(s)}(x;\theta)$，$\tilde{u}_\delta(x;\theta)$ 和 $\tilde{u}_\delta^{(s)}(x;\theta)$，用类似方法构造损失函数并求解.

2　二维光栅衍射问题

2.1　透明边界条件

考虑在 x 方向上无限、在 y 方向周期为 L 的二维圆柱周期结构，如图 1 所示. 介质的介电函数 $\varepsilon(x,y)$ 满足

$$\varepsilon(x,y)=\begin{cases}\varepsilon_0, & (x,y)\notin\Omega_n,\\ \varepsilon_1, & (x,y)\in\Omega_n,\end{cases}\tag{8}$$

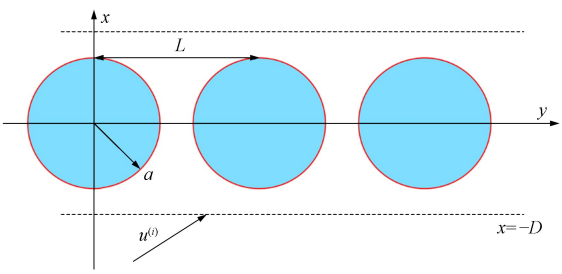


图 1　二维圆柱周期结构

Fig. 1　Two-dimensional cylindrical periodic structure

其中 $\Omega_n=\{(x,y)\,|\,x^2+(y-nL)^2<a^2\}$，$n=0,\pm1,\cdots,\pm m\}$，$a$ 为圆柱半径. 设 D 满足 $D>a$，当 $|x|>D$ 时，介质是均匀的. 在 $x<-D$ 一侧给定如下形式频率为 ω 的入射波

$$u^{(i)}(x,y)=\mathrm{e}^{\mathrm{i}[\alpha(x+D)+\beta y]},\tag{9}$$

其中 (α,β) 为波矢，并且满足 $\alpha^2+\beta^2=k^2\varepsilon_0$.

由于在 y 方向上为周期函数，因此满足拟周期边界条件. 在 x 方向上用透明边界条件对二维光栅衍射问题计算域进行截断. 对任意的整数 m，定义线性算子 T 满足[15]

$$T\mathrm{e}^{\mathrm{i}\beta_m y}=\mathrm{i}\alpha_m\mathrm{e}^{\mathrm{i}\beta_m y}.\tag{10}$$

在透明边界条件下总场 u 满足

$$\begin{cases}\Delta u+k^2\varepsilon(x,y)u=0, & \forall(x,y)\in\Omega_D,\\ u\left(x,\dfrac{L}{2}\right)=\mathrm{e}^{\mathrm{i}\beta L}u\left(x,-\dfrac{L}{2}\right),\quad \dfrac{\partial u}{\partial y}\left(x,\dfrac{L}{2}\right)=\mathrm{e}^{\mathrm{i}\beta L}\dfrac{\partial u}{\partial y}\left(x,-\dfrac{L}{2}\right), & |x|<D,\\ \dfrac{\partial u}{\partial x}(-D,y)=-Tu(-D,y)+2\mathrm{i}\alpha\mathrm{e}^{\mathrm{i}\beta y}, & x=-D,\ |y|<\dfrac{L}{2},\\ \dfrac{\partial u}{\partial x}(D,y)=Tu(D,y), & x=D,\ |y|<\dfrac{L}{2},\end{cases}\tag{11}$$

其中 $\Omega_D=\left\{(x,y)\,\middle|\,|x|<D,\ |y|<\dfrac{L}{2}\right\}$，在数值计算中，线性算子 T 可用矩阵 $\boldsymbol{T}$ 近似，矩阵 $\boldsymbol{T}$ 的计算见文献[15].

与一维情况类似，衍射问题(11)也可通过散射场 $u^{(s)}=u-u^{(i)}$ 求解，$u^{(s)}$ 满足

$$\begin{cases} \Delta u^{(s)} + k^2\varepsilon(x,y)u^{(s)} = k^2(\varepsilon_0 - \varepsilon(x,y))u^{(i)}, & \forall (x,y) \in \Omega_D, \\ u^{(s)}(x,\frac{L}{2}) = \mathrm{e}^{\mathrm{i}\beta L}u^{(s)}\left(x,-\frac{L}{2}\right), \quad \frac{\partial u^{(s)}}{\partial y}\left(x,\frac{L}{2}\right) = \mathrm{e}^{\mathrm{i}\beta L}\frac{\partial u^{(s)}}{\partial y}\left(x,-\frac{L}{2}\right), & |x| < D, \\ \frac{\partial u^{(s)}}{\partial x}(-D,y) = -Tu^{(s)}(-D,y), & x = -D,\ |y| < \frac{L}{2}, \\ \frac{\partial u^{(s)}}{\partial x}(D,y) = Tu^{(s)}(D,y), & x = D,\ |y| < \frac{L}{2}. \end{cases} \tag{12}$$

为解决解在间断点上二阶不可导造成物理信息神经网络求解困难的问题，与一维情形类似，引入二维光滑函数 $\varepsilon_\delta(x,y)$近似介电函数 $\varepsilon(x,y)$. $\varepsilon_\delta(x,y)$定义为

$$\varepsilon_\delta(r) = \frac{1}{2}\left(\frac{r-a}{\sqrt{(r-a)^2+\delta^2}} + 1\right)(\varepsilon_0 - \varepsilon_1) + \varepsilon_1, \qquad (x,y) \in \Omega_D, \tag{13}$$

其中 $r=\sqrt{x^2+y^2}$. 将问题(11)和(12)光滑化后的解记为 $u_\delta(x,y)$和 $u_\delta^{(s)}(x,y)$.

2.2 物理信息神经网路求解二维光栅衍射问题

以 $\tilde{u}(x,y;\theta)$表示一个神经网络，其中(x,y)为神经网络的二维输入值，θ 为神经网络中所需学习的参数，与一维类似，令神经网络输出为二维，分别用 $v(x,y;\theta)$和 $w(x,y;\theta)$表示函数的实部和虚部，则有

$$\tilde{u}(x,y;\theta) = v(x,y;\theta) + \mathrm{i}w(x,y;\theta). \tag{14}$$

用 $\tilde{u}^{(s)}(x,y;\theta)$ 近似问题(12)的解 $u^{(s)}(x,y)$，设 $\{(x_j,y_j)\}_{j=1}^{N}$ 为区域 Ω_D 上 N 个观测点，$\left\{\left(x_j,\frac{L}{2}\right)\right\}_{j=1}^{N_{bx}}$和$\left\{\left(x_j,-\frac{L}{2}\right)\right\}_{j=1}^{N_{bx}}$为 x 边界上分别取的N_{bx}个点，$\{(D,y_j)\}_{j=1}^{N_{by}}$和$\{(-D,y_j)\}_{j=1}^{N_{by}}$为 y 边界上分别取的N_{by}个观测点，问题(12)对应的损失函数可写为

$$\begin{aligned} L(\theta) = {} & \frac{1}{N}\sum_{j=1}^{N}\left|\Delta\tilde{u}^{(s)}(x_j,y_j;\theta) + k^2\varepsilon(x_j,y_j)\tilde{u}^{(s)} - k^2(\varepsilon_0 - \varepsilon(x_j,y_j))u^{(i)}(x_j,y_j;\theta)\right|^2 + \\ & \frac{1}{N_{bx}}\sum_{j=1}^{N_{bx}}\left|\tilde{u}^{(s)}\left(x_j,\frac{L}{2}\right) - \mathrm{e}^{\mathrm{i}\beta L}\tilde{u}^{(s)}\left(x_j,-\frac{L}{2}\right)\right|^2 + \frac{1}{N_{bx}}\sum_{j=1}^{N_{bx}}\left|\frac{\partial\tilde{u}^{(s)}}{\partial y}\left(x_j,\frac{L}{2}\right) - \mathrm{e}^{\mathrm{i}\beta L}\frac{\partial\tilde{u}^{(s)}}{\partial y}\left(x_j,-\frac{L}{2}\right)\right|^2 + \\ & \frac{1}{N_{by}}\sum_{j=1}^{N_{by}}\left|\frac{\partial\tilde{u}^{(s)}}{\partial x}(-D,y_j) + T\tilde{u}^{(s)}(-D,y_j)\right|^2 + \frac{1}{N_{by}}\sum_{j=1}^{N_{by}}\left|\frac{\partial\tilde{u}^{(s)}}{\partial x}(D,y_j) - T\tilde{u}^{(s)}(D,y_j)\right|^2, \end{aligned} \tag{15}$$

其中线性算子 T 可用矩阵 $\boldsymbol{T}$ 近似. 利用优化算法求解最小值问题 $\min\limits_{\theta} L(\theta)$即可得到 $u^{(s)}(x,y)$的近似解 $\tilde{u}^{(s)}(x,y;\theta)$. 同理，对于问题(11)以及问题(11)与(12)光滑后问题的解 $u(x,y)$，$u_\delta(x,y)$和 $u_\delta^{(s)}(x,y)$ 分别近似为神经网络 $\tilde{u}(x,y;\theta)$，$\tilde{u}_\delta(x,y;\theta)$和 $\tilde{u}_\delta^{(s)}(x,y;\theta)$. 用类似方法构造损失函数并求解.

3 数值实验

本文的数值实验基于 Python3.8 和 Tensorflow2.11 环境，计算机配置为 AMD Ryzen 7 5800H. 神经网络架构均采用全连接方式，每个隐藏层的神经元数量均相同，激活函数选取双曲正切函数 tanh，优化器选择为 Adam，学习率使用分段学习率.

3.1 一维光栅衍射

3.1.1 十层介质

用物理信息神经网络对一维光栅问题进行求解，取 $\omega=0.55(2\pi c/L)$，$L=1$，$\varepsilon_0=0$，$\varepsilon_1=10$，$\varepsilon_2=6$，$m=10$，$D_n=0.5n$，$n=1,2,\cdots,10$，介电函数 $\varepsilon(x)$为

$$\varepsilon(x) = \begin{cases} \varepsilon_0, & x < D_0, \quad x > D_{10}, \\ \varepsilon_1, & D_n < x < D_{n+1}, \quad n = 0,2,\cdots,8, \\ \varepsilon_2, & D_n < x < D_{n+1}, \quad n = 1,3,\cdots,9. \end{cases} \tag{16}$$

将求解总场且不进行光滑化的解 $\tilde{u}(x;\theta)$记为方法 1，求解散射场且不进行光滑化的解 $\tilde{u}^{(s)}(x;\theta)$记为方法 2，求解总场并进行光滑化的解 $\tilde{u}_\delta(x;\theta)$记为方法 3，求解散射场并进行光滑化的解 $\tilde{u}_\delta^{(s)}(x;\theta)$

记为方法 4. 对 4 种方法进行比较，取训练集观测点数为 $N=400$，选点方式为均匀选点，光滑函数 $\delta=10^{-4}$，神经网络结构为 8 层，每层 40 个神经元，迭代次数为 3×10^5. 由于方法 3 和方法 4 所求结果为散射场，因此需用 $u=\tilde{u}^{(s)}+u^{(i)}$ 计算总场. 在传播过程中没有能量损耗，若反射系数 $|R|^2$ 和透射系数 $|T|^2$ 之和越接近 1，则表示数值计算结果越精确. 各方法的数值结果与总场真解 L_2 误差、数值计算的反射系数 $|R|^2$ 和透射系数 $|T|^2$ 之和、数值计算的反射系数 $|R|^2$ 与真实解 $|R|^2$ 之差的绝对值以及数值计算的反射系数 $|T|^2$ 与真实解 $|T|^2$ 之差的绝对值列于表 1，损失函数以及数值解的实部分别如图 2 和图 3 所示.

表 1　一维问题各方法的数值结果与真解的误差

Table 1　Errors between numerical results of various methods and true solutions for one-dimensional problems

方法	L_2 误差	$\|R\|^2+\|T\|^2$	$\|R\|^2$ 误差	$\|T\|^2$ 误差
1	0.193 37	0.844 86	0.367 53	0.522 663
2	0.134 05	0.876 60	0.214 90	0.338 298
3	0.028 02	0.998 83	0.002 33	0.003 494
4	0.002 69	0.999 76	0.002 27	0.002 294

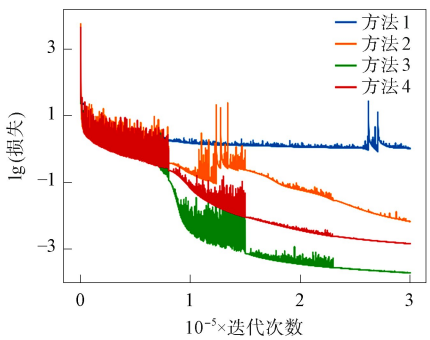


图 2　一维问题各方法的损失函数

Fig. 2　Loss functions of various methods for one-dimensional problems

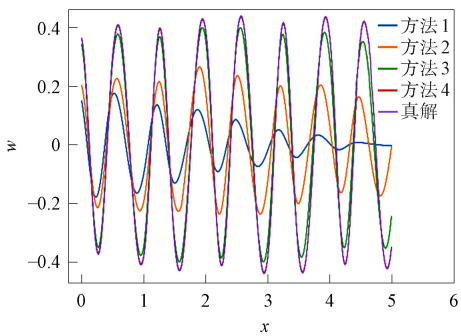


图 3　一维问题各方法解的实部

Fig. 3　Real parts of solutions of various methods for one-dimensional problems

由表 1、图 2 和图 3 可见：方法 2 计算散射场的误差比方法 1 计算总场小，并且方法 2 损失函数曲线下降更快，由于间断系数不可导，因此两种方法的计算结果误差较大；进行光滑化后方法 3 和方法 4 的损失函数曲线下降更快，振荡更少，训练效果更好，并且解的误差更小；方法 4 为所有方法中误差最小且反射系数 $|R|^2$ 和透射系数 $|T|^2$ 之和最接近 1，并且 $|R|^2$ 误差和 $|T|^2$ 误差均最小，因此在一维光栅衍射问题中，用方法 4 求解散射场并进行光滑化的训练方法最有效. 在仅改变神经网络结构、网络层数和每层神经元数量的条件下，利用方法 4 进行数值实验，结果列于表 2.

表 2　一维问题不同网络结构的误差

Table 2　Errors of different network structures in one-dimensional problems

神经网络层数	神经元数量	误差	$\|R\|^2+\|T\|^2$	$\|R\|^2$ 误差	$\|T\|^2$ 误差
6	40	0.006 297	0.989 225	0.009 250	0.020 022
6	60	0.001 461	0.999 043	0.000 667	0.000 021
6	80	0.001 536	0.999 750	0.001 183	0.000 930
8	40	0.001 747	0.999 485	0.000 050	0.000 465
8	60	0.001 494	0.999 822	0.000 385	0.000 210
8	80	0.001 590	0.999 973	0.001 159	0.001 130
10	40	0.001 568	0.999 823	0.001 437	0.001 260
10	60	0.001 568	0.999 439	0.001 441	0.000 880
10	80	0.001 589	0.999 990	0.001 041	0.001 030

由表 2 可见，随着网络层数的增加，模型的误差越来越小，由于神经网络层数从 8 层增加到 10 层误差变化较小，因此本文选取神经网络层数为 8 层.

3.1.2 选取 δ

由于光滑函数中的 δ 是一个超参数，因此需寻找最优的 δ. 在10层介质的基础上进行实验，选择网络结构为8个隐藏层，每层神经元数量为40个，其他条件不变，仅改变 δ，利用方法4进行实验，结果列于表3.

表3 一维问题 δ 对解的影响

Table 3 Effect of δ on solutions for one-dimensional problems

δ	0	10^{-6}	10^{-5}	10^{-4}	10^{-3}	10^{-2}	10^{-1}
误差	0.393 684	0.020 933	0.006 658	0.002 847	0.001 770	0.003 992	0.183 557
$\lvert R\rvert^2+\lvert T\rvert^2$	0.986 398	0.997 242	0.999 587	0.999 765	0.999 831	0.999 104	1.001 930
$\lvert R\rvert^2$ 误差	0.508 560	0.034 410	0.008 010	0.002 050	0.001 250	0.001 848	0.053 310
$\lvert T\rvert^2$ 误差	0.522 165	0.037 167	0.008 421	0.004 397	0.002 937	0.000 950	0.053 113

引入光滑函数后的物理信息神经网络方法与真解的误差可表示为

$$\lvert\tilde{u}_\delta-u\rvert\leqslant\lvert\tilde{u}_\delta-u_\delta\rvert+\lvert u_\delta-u\rvert, \tag{17}$$

其中 u 为原方程的解，u_δ 为光滑后的真解，$\tilde{u}_\delta$ 为光滑后神经网络求得的近似解. 由表3可见，δ 越大函数越光滑，但 δ 太大会导致光滑函数 $\varepsilon_\delta(x)$ 与原介电函数 $\varepsilon(x)$ 差距过大，导致 $\lvert u_\delta-u\rvert$ 增大，而 δ 太小会使光滑函数在间断点处导数过大不够光滑，导致 $\lvert\tilde{u}_\delta-u_\delta\rvert$ 增大，因此为保证 $\lvert u_\delta-u\rvert$ 和 $\lvert\tilde{u}_\delta-u_\delta\rvert$ 都足够小，超参数 δ 应选取适中的值才能使误差最小，在该问题中 $\delta=10^{-3}$ 误差最小.

3.2 二维光栅衍射

用物理信息神经网络对二维光栅问题进行求解，取 $\omega=0.55(2\pi c/L)$，$\varepsilon_0=1$，$\varepsilon_1=10$，$L=1$，$a=0.3L$，$n=2$，介电函数 $\varepsilon(x)$ 为

$$\varepsilon(x,y)=\begin{cases}\varepsilon_0, & (x,y)\notin\Omega_D,\\ \varepsilon_1, & (x,y)\in\Omega_D.\end{cases} \tag{18}$$

二维与一维命名方法相同. 取训练集观测点数为 $N=3\,200$，$N_{bx}=80$，$N_{by}=40$，选点方式为均匀选点，光滑函数 $\delta=10^{-3}$，神经网络结构为8层，每层80个神经元，迭代次数为 2×10^5. 将有限元的解作为参考解，各方法的数值结果与总场参考解 L_2 误差、反射系数 $\lvert R\rvert^2$ 和透射系数 $\lvert T\rvert^2$ 之和、数值计算的反射系数 $\lvert R\rvert^2$ 与参考解的 $\lvert R\rvert^2$ 之差的绝对值以及数值计算的反射系数 $\lvert T\rvert^2$ 与参考解 $\lvert T\rvert^2$ 之差的绝对值列于表4，损失函数以及各方法解与参考解之差的模分别如图4和图5所示.

表4 二维问题各方法的数值结果与真解的误差

Table 4 Errors between numerical results of various methods and true solutions for two-dimensional problems

方法	L_2 误差	$\lvert R\rvert^2+\lvert T\rvert^2$	$\lvert R\rvert^2$ 误差	$\lvert T\rvert^2$ 误差
1	0.415 718	0.721 959	0.461 447	0.183 410
2	0.402 212	0.952 089	0.309 327	0.153 090
3	0.278 166	0.843 763	0.131 403	0.179 314
4	0.009 851	0.991 969	0.000 517	0.008 548

由表4、图4和图5可见：方法1和方法2的损失函数曲线下降较慢，进行光滑后方法3和方法4的损失函数曲线下降更快；方法1,2,3的误差均较大，计算效果较差，其中方法2虽然 $\lvert R\rvert^2+\lvert T\rvert^2$ 接近1，但 $\lvert R\rvert^2$ 和 $\lvert T\rvert^2$ 的误差均较大；方法4为所有方法中误差最小，并且 $\lvert R\rvert^2+\lvert T\rvert^2$ 最接近1，因此在二维光栅衍射问题中，用方法4求解散射场并进行光滑化的训练方法同样最有效.

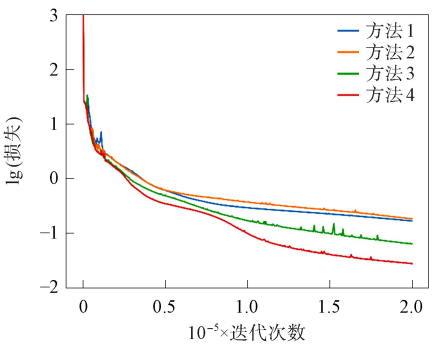


图4 二维问题各方法的损失函数

Fig. 4 Loss functions of various methods for two-dimensional problems

与一维情形类似，在仅改变神经网络结构、网络层数和每层神经元数量的条件下，利用方法4进行数值实验，结果列于表5.

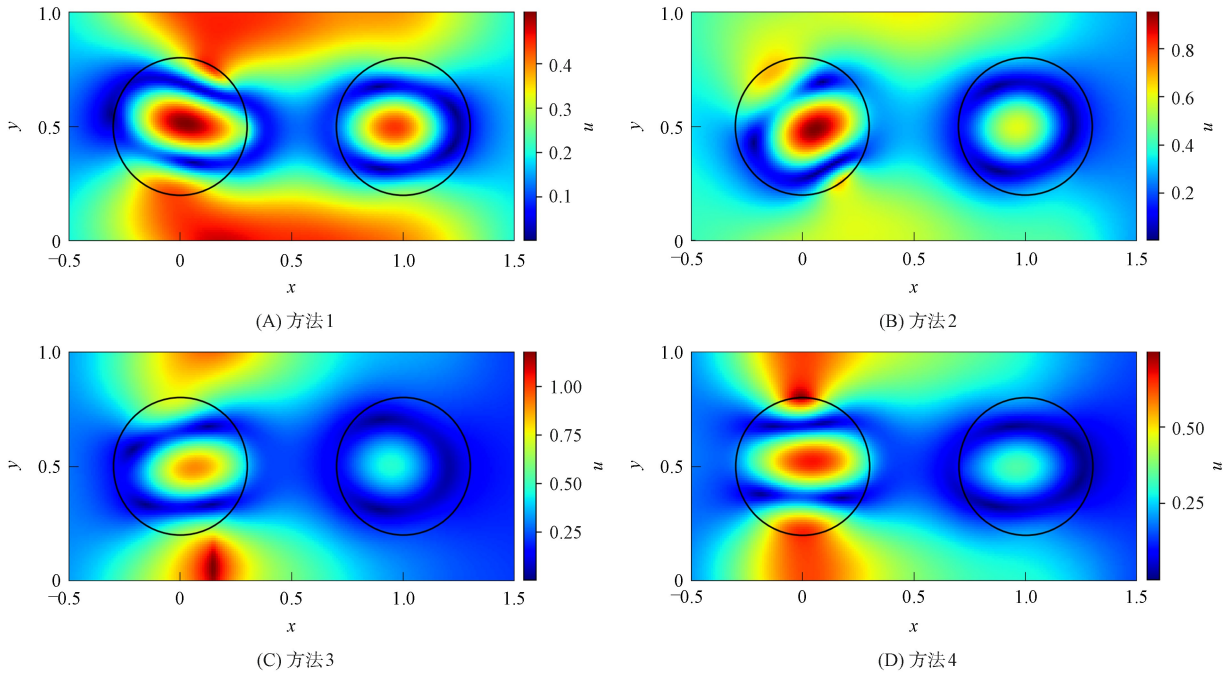


图 5　二维问题各方法解与参考解之差的模

Fig. 5　Modulus of subtraction between solutions of various methods and reference solutions for two dimensional problems

表 5　二维问题不同网络结构的误差

Table 5　Errors of different network structures in two-dimensional problems

神经网络层数	神经元数量	误差	$\|R\|^2+\|T\|^2$	$\|R\|^2$ 误差	$\|T\|^2$ 误差
6	40	0.172 134	0.921 299	0.051 733	0.130 434
6	60	0.041 941	0.974 533	0.006 114	0.031 581
6	80	0.011 568	0.996 116	0.008 625	0.004 791
8	40	0.161 494	0.940 469	0.046 995	0.106 526
8	60	0.014 068	0.989 715	0.001 765	0.008 519
8	80	0.009 851	0.991 969	0.000 517	0.008 548
10	40	0.076 854	0.978 481	0.004 682	0.026 201
10	60	0.013 089	0.990 072	0.005 017	0.004 910
10	80	0.010 504	0.995 267	0.005 186	0.000 452

由表 5 可见，增加神经网络层数和神经元数量即增加网络结构的复杂程度均可降低误差. 与一维问题的结果相比，二维问题需更复杂的网络结构才能达到较好的精度.

综上，本文将物理信息神经网络方法用于求解光栅衍射问题. 结果表明，将间断系数近似为一个光滑函数可提升计算效果，并且计算散射场比总场的结果更好. 通过一维和二维光栅衍射的数值实验验证了结论，并在多组对照实验中找到合适的超参数(神经网络层数、每层神经元数量和光滑函数 δ 值)以达到最小的误差.

参　考　文　献

[1] 李野，陈松灿. 基于物理信息的神经网络：最新进展与展望 [J]. 计算机科学，2022，49(4)：254-262. (LI Y，CHEN S C. Physics-Informed Neural Networks：Recent Advances and Prospects [J]. Computer Science，2022，49(4)：254-262.)

[2] MAO Z P，JAGTAP A D，KARNIADAKIS G E. Physics-Informed Neural Networks for High-Speed Flows [J]. Computer Methods in Applied Mechanics and Engineering，2020，360：112789-1-112789-26.

[3] FANG Z W，ZHAN J. Deep Physical Informed Neural Networks for Metamaterial Design [J]. IEEE Access，2019，8：24506-24513.

[4] 陆至彬，瞿景辉，刘桦，等. 基于物理信息神经网络的传热过程物理场代理模型的构建[J]. 化工学报，2021，72(3)：1496-1503.(LU Z B，QU J H，LIU H，et al. Surrogate Modeling for Physical Fields of Heat Transfer Processes Based on Physics-Informed Neural Network [J]. CIESC Journal，2021，72(3)：1496-1503.)

[5] LU L，JIN P Z，PANG G F，et al. Learning Nonlinear Operators via DeepONet Based on the Universal Approximation Theorem of Operators [J]. Nature Machine Intelligence，2021，3(3)：218-229.

[6] RAISSI M，PERDIKARIS P，KARNIADAKIS G E. Physics-Informed Neural Networks：A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations [J]. Journal of Computational Physics，2019，378：686-707.

[7] HAN J Q，JENTZEN A E W N. Solving High-Dimensional Partial Differential Equations Using Deep Learning [J]. Proceedings of the National Academy of Sciences，2018，115(34)：8505-8510.

[8] RAISSI M，KARNIADAKIS G E. Hidden Physics Models：Machine Learning of Nonlinear Partial Differential Equations [J]. Journal of Computational Physics，2018，357：125-141.

[9] 朱磊，邵晓鹏. 散射成像技术的研究进展[J]. 光学学报，2020，40(1)：77-91.(ZHU L，SHAO X P. Research Progress on Scattering Imaging Technology [J]. Acta Optica Sinica，2020，40(1)：77-91.)

[10] 查显昊，胡真. 基于 Dirichlet-to-Neumann 映射计算太赫兹频段旋电光子晶体带隙结构[J]. 吉林大学学报(理学版)，2023，61(2)：400-406.(ZHA X H，HU Z. Computing Bandgap Structures of Gyroelectric Photonic Crystals in Terahertz Band Based on Dirichlet-to-Neumann Map [J]. Journal of Jilin University (Science Edition)，2023，61(2)：400-406.)

[11] 洪亮，高尚，李翔. 基于网络特征的分层剪枝方法[J]. 吉林大学学报(理学版)，2022，60(6)：1407-1415.(HONG L，GAO S，LI X. Layer-Wise Pruning Method Based on Network Characteristics [J]. Journal of Jilin University (Science Edition)，2022，60(6)：1407-1415.)

[12] JAGTAP A D，KHARAZMI E，KARNIADAKIS G E. Conservative Physics-Informed Neural Networks on Discrete Domains for Conservation Laws：Applications to Forward and Inverse Problems [J]. Computer Methods in Applied Mechanics and Engineering，2020，365：113028-1-113028-27.

[13] KHARAZMI E，ZHANG Z Q，KARNIADAKIS G E M. hp-VPINNs：Variational Physics-Informed Neural Networks with Domain Decomposition [J]. Computer Methods in Applied Mechanics and Engineering，2021，374：113547-1-113547-25.

[14] CHEW W C. Waves and Fields in Inhomogenous Media [M]. New York：John Wiley & Sons，1999：22-24.

[15] WANG Z F，BAO G，LI J Q，et al. An Adaptive Finite Element Method for the Diffraction Grating Problem with Transparent Boundary Condition [J]. SIAM Journal on Numerical Analysis，2015，53(3)：1585-1607.

(责任编辑：王 健)