arXiv:2305.05540v1 [math-ph] 7 May 2023

直接泊松神经网络：学习非辛机械系统

Martin Šípka查理大学数学与物理学院数学研究所，Sokolovská 83, 18675 布拉格，捷克共和国 通讯作者：martinsipka@gmail.com

Michal Pavelka查理大学数学与物理学院数学研究所，捷克共和国布拉格 Sokolovská 83，邮编 18675

O˘古尔·埃森盖布泽技术大学数学系，41400 盖布泽，科贾埃利，土耳其

Miroslav Grmela蒙特利尔综合理工学院，C.P.6079 市中心分局，蒙特利尔，H3C 3A7，魁北克，加拿大

2023年5月10日

摘要

本文提出了能够学习辛力学系统（例如质点力学）和非辛力学系统（例如旋转刚体）的神经网络。力学系统具有哈密顿演化，由两个基本组成部分构成：泊松括号和能量泛函。我们将一组哈密顿系统的快照输入神经网络模型，模型随后同时学习这两个基本组成部分。特别是，这些模型能够区分辛系统（具有非退化泊松括号）和非辛系统（具有退化括号）。与早期研究不同，我们的方法除了系统具有哈密顿性质之外，不对动力学作任何额外的先验假设，并且能够给出满足雅可比恒等式的泊松括号。最后，这些模型可以判断一个方程组是否为哈密顿系统。

目录

1 引言 **2**

2 泊松几何中的哈密顿动力学 **4**

2.1 一般表述.4 2.2 $3D$ 哈密顿动力学.7 2.3 $4D$ 哈密顿动力学.8 2.4 向一个 $6D$ 系统作半直积扩张.10

3 学习哈密顿系统 **10**

3.1 刚体.12 3.2 二维粒子.15 3.3 Shivamoggi 方程.15 3.4 三维粒子.18 3.5 重陀螺.18

4 学习非哈密顿量系统 19

5 结论 **21**

1 引言

在物理学和工程学中，未知参数的估计是许多成熟方法和工作流程中的标准步骤。通常，我们从一个模型开始——一组被视为已知的假设和方程；然后根据现有数据，估计所关注系统的演化方程的确切形式。举例来说，我们可以考虑这样一种情况：根据遥远恒星与光的相互作用来估计其质量 [1], ，或者根据小行星的自转推断其转动惯量 [2]。假设的复杂程度各不相同，因此应当恰当地选择参数估计方法。

近年来，动力系统的机器学习技术引起了广泛关注。随着神经网络相关进展的兴起，人们开发了多种方法来捕捉动力系统的行为，各有优缺点。符号方法（例如 [3]）使我们能够从预先定义的可用运算集合中学习方程的精确符号形式。这通常是最高效的方法，往往能使学习得到的系统与目标系统完全匹配，但按定义，可捕捉方程的类别受限于我们考虑作为候选项的代数运算。

或者，也可以直接学习运动方程

$$\dot{\mathbf{x}} = f(\mathbf{x}, \theta) \tag{1}$$

通过学习由权重 θ 参数化的 f。该函数可以由任意函数逼近器表示，在许多情况下由神经网络表示。尽管这种方法非常通用，但它并未将任何已知的物理学原理纳入我们的流程。系统的能量概念并不存在，也没有任何量被隐式守恒，因此该方法可能会给出不符合物理规律的预测。解决这一问题的方法是采用物理信息机器学习的理念，通过约束神经网络模型，使其遵循某些必要的物理定律 [4]。具体而言，可由哈密顿力学描述的力学系统模型会保持若干物理量不变，例如能量或角动量，同时也会保持几何量（例如辛二形式），从而确保系统的自洽性。据我们所知，此前尚无一种神经网络模型能够在不具备任何关于系统的先验知识的情况下，从系统轨迹中学习哈密顿系统，同时又与其底层几何结构相容，

本文的主要目的便是介绍这种模型。此外，我们还提出了若干模型，它们在严格复现底层几何结构的程度上有所不同；模型对系统的学习程度可用于估计该系统是否为哈密顿系统。

动力系统的几何化。动力系统由微分方程描述，尤其是力学系统遵循哈密顿演化方程。这些方程源于几何结构，该结构在坐标变换下保持不变，并在系统运动过程中得以保留。哈密顿系统的几何学可追溯到 Sophus Lie 和亨利·庞加莱 [5, 6]。现代方法将其扩展到无限维系统，并为当今物理学的许多领域奠定了基础 [7, 8, 9, 10]。以几何方式表述力学系统，通常意味着寻找一种括号代数（如辛、泊松、雅可比、莱布尼茨等）和一个生成函数（如哈密顿量、能量或熵）。通常要求括号满足某些代数条件（如雅可比恒等式、莱布尼茨恒等式等）。然而，即便某个系统存在哈密顿表述，也没有通用的算法方法能够通过解析计算求得它。因此，这类分析适合采用机器学习。

除了哈密顿力学之外，还可以纳入耗散 [11, 12, 13] ，或将哈密顿系统的学习扩展到控制问题 [14]。随后，通过选择合适的积分器，这种方法可以确保能量、动量或角动量等具有重要物理意义的量得以守恒。

可逆系统是哈密顿系统的候选者。对于可逆系统，一个起点可以是寻找辛流形和哈密顿函数。可以利用神经网络学习物理系统（包括势场中的粒子、复杂程度各异的摆）的辛特征（如果存在），例如参见 [15, 16]。辛几何只存在于偶数维模型中，并且由于非退化判据的约束，它非常刚性。辛几何的一种推广是泊松几何，其中放宽了非退化要求 [17, 18]。在泊松几何中，存在一个泊松括号（在函数空间上定义一种代数），满足莱布尼茨恒等式和雅可比恒等式。这种推广使得可以在奇数维中进行（哈密顿）分析。泊松几何的退化特征为研究动力学的（完全甚至超）可积性带来了一些优势。

本文旨在更好地学习哈密顿力学。我们构建了能够编码具有泊松几何结构的哈密顿动力学基本构件的神经网络。根据泊松流形的 Darboux-Weinstein 定理 [19]，存在使泊松双向量（决定括号）成为常量的正则坐标。这一表述也具有几何意义，因为它确定了泊松流形的辛叶状结构（更多详情请参见本文正文）。最近，有人提出了泊松神经网络（简称 PNN），通过将系统变换到 Darboux-Weinstein 坐标中来学习哈密顿系统 [20] 。但是，对于许多物理系统而言，泊松括号远非处于正则坐标中，而且 Darboux-Weinstein 泊松双向量的辛部分维数可能事先未知。对于 n–维物理模型，确保底层泊松双向量一致性的雅可比恒等式是一个偏微分方程组。到

确定泊松结构，需要解析地求解该系统，而这通常很困难；或者在学习泊松双向量时强制满足其有效性。

本文的目标。 本文的创新成果是我们称之为直接泊松神经网络（简称 DPNN）的方法，该方法无需假设泊松结构具有任何特定形式，例如 Darboux-Weinstein 坐标。相反，DPNN 直接在数据所提供的坐标系中进行学习。DPNN 具有以下几个优势： **(i)** 我们无需先验了解泊松结构的退化度（换言之，即泊松流形的辛叶状结构的维数）； **(ii)** 更易于学习哈密顿量（能量）； **(iii)** 雅可比恒等式在数据上得到满足，而不仅仅是在只能通过另一个神经网络（ [20]中的可逆神经网络）访问的表示中得到满足。DPNN 通过直接识别作为状态变量函数的泊松双向量和哈密顿量来学习泊松系统。

实际上，我们提供了三种类型的直接泊松神经网络。信息最少的类型直接学习哈密顿函数和泊松双向量，仅假设其具有反对称性，而不假设其满足雅可比恒等式。另一种类型将雅可比恒等式的平方项加入损失函数，从而以软约束的方式强制其成立。几何信息最丰富的类型则通过将泊松双向量构造成三维空间中雅可比恒等式的通解，自动满足雅可比恒等式。尽管几何信息最丰富的版本通常最擅长学习哈密顿系统，但它仅限于三维系统，因为只有在三维情况下才能得到雅可比恒等式的通解。第二种类型通常精确度略低，而信息最少的类型通常精确度最低，不过它仍然能够以相当高的精度学习哈密顿系统。

有趣的是，当我们尝试利用这三种类型的直接泊松神经网络学习非哈密顿模型时，精确度顺序会反转，信息最少的类型反而变得最为精确。因此，直接泊松神经网络各类型的精确度顺序可以表明一个方程组是否为哈密顿系统。

第2节回顾泊松动力学，尤其是辛动力学、刚体力学、Shivamoggi 方程以及重陀螺的演化。第3节介绍直接泊松神经网络，并展示其在学习哈密顿系统中的应用。最后，第4节展示了 DPPNs 在非哈密顿系统（耗散刚体）上的应用。

2 泊松几何中的哈密顿动力学

2.1 一般表述

流形 $\mathcal{M}$ 上的泊松括号（在物理上对应于状态空间，例如物体的位置和动量）是定义在 $\mathcal{M}$ 上光滑函数空间 $\mathcal{F}(\mathcal{M})$ 上的斜对称双线性代数，其定义如下：

$$\{\bullet,\bullet\} : \mathcal{F}(\mathcal{M}) \times \mathcal{F}(\mathcal{M}) \to \mathcal{F}(\mathcal{M}). \tag{1}$$

泊松括号满足莱布尼茨法则

$$\{F, HG\} = \{F, H\}G + H\{F, G\}, \tag{2}$$

以及雅可比恒等式，

$$\{F, \{H, G\} + \{H, \{G, F\} + \{G, \{F, H\} = 0, \tag{3}$$

对于任意函数 F、 H 和 G、 [10, 17, 18]。配备泊松括号的流形称为泊松流形，记作二元组 $(\mathcal{M}, \{\bullet, \bullet\})$。若函数 C与所有其他函数均可交换，即对所有 F都有 $\{F, C\} = 0$，则称其为卡西米尔函数。例如，刚体角动量的大小就是一个卡西米尔函数。

哈密顿动力学。 哈密顿动力学可以看作泊松流形上的演化。对于流形 $\mathcal{M}$上的哈密顿函数（物理上即能量）H，哈密顿向量场和哈密顿方程分别为

$$X_H(F) := \{F, H\}, \qquad \dot{\mathbf{x}} = X_H(\mathbf{x}) = \{\mathbf{x}, H\}, \tag{4}$$

其中，$\mathbf{x} \in \mathcal{M}$ 是流形 $\mathcal{M}$的参数化。泊松括号的代数性质会带来一些物理后果。反对称性意味着能量守恒，

$$\dot{H} = \{H, H\} = 0, \tag{5}$$

而莱布尼茨法则确保动力学不受能量加上常数这一偏置的影响。给定泊松括号，可以据此确定泊松双向量场：

$$L(dF, dH) := \{F, H\}. \tag{6}$$

这种对应关系使我们能够通过一个由流形和泊松双向量组成的元组 $(\mathcal{M},\ L)$ 来定义泊松流形。采用这种记法，雅可比恒等式可改写为 $\mathcal{L}_{\mathbf{X}_H} L = 0$，也就是说，泊松双向量相对于哈密顿向量场的李导数为零 [21]。换言之，雅可比恒等式表达了哈密顿动力学的自洽性：构成动力学的两个基本要素（哈密顿函数和双向量场）在演化过程中均保持不变。

假设在 $\mathcal{M}$ 上采用局部坐标系 $\mathbf{x} = (x^i)$，泊松双向量确定了泊松矩阵 $L = [L^{kl}]$，由此我们可以写出 [22]

$$L = L^{kl} \frac{\partial}{\partial x^k} \wedge \frac{\partial}{\partial x^l}. \tag{7}$$

在这种实现中，泊松括号和哈密顿方程分别写作

$$\{F, H\} = L^{kl} \frac{\partial F}{\partial x^k} \frac{\partial H}{\partial x^l}, \qquad \dot{x}^k = L^{kl} \frac{\partial H}{\partial x^l}, \tag{8}$$

这里，我们假设对重复指标求和。此外，雅可比恒等式（3）可转化为以下偏微分方程组

$$L^{kl}\frac{\partial L^{ij}}{\partial x^k}+L^{ki}\frac{\partial L^{jl}}{\partial x^k}+L^{kj}\frac{\partial L^{li}}{\partial x^k}=0. \tag{9}$$

该方程的左侧称为雅可比子；对于哈密顿系统，它等于零。雅可比恒等式（9）是由偏微分方程组成的微分方程组。在 $3D$ 中，雅可比恒等式（9）是一个偏微分方程，其通解已知 [23]。在 $4D$ 中，雅可比恒等式（9）由四个偏微分方程组成，已有一些部分结果，但据我们所知，对于任意 n，目前还没有通解。接下来的小节将重点讨论 $3D$、$4D$和 $6D$ 情形。

辛流形。如果一个泊松流形不存在非恒定的卡西米尔函数，那么它也是一个辛流形。虽然我们可以将辛流形视为泊松流形的实例，但也可以不借助泊松流形而直接定义辛流形。若一个流形 $\mathcal{M}$ 配备了闭的非退化二形式（称为辛二形式）Ω，则称其为辛流形。当二形式满足以下条件时，称其为非退化：

$$\Omega(X,Y)=0, \qquad \forall X\in\mathfrak{X}(\mathcal{M}) \tag{10}$$

可推出 $Y=0$。当二形式处于德拉姆外微分的核中时，它是闭的，$d\Omega=0$。辛流形 $(\mathcal{M},\Omega)$ 上的哈密顿向量场 X_H 定义为

$$\iota_{X_H}\Omega=dH, \tag{11}$$

其中 ι 是缩并算子（更准确地说，是内导数）[21]。借助辛流形，可以定义泊松括号

$$\{F,H\}:=\Omega(X_F,X_H), \tag{12}$$

其中哈密顿向量场通过公式 (11) 定义。辛二形式 Ω 的闭性保证了雅可比恒等式 (3)。Ω 的非退化条件对括号 (12) 施加了额外限制：卡西米尔函数只能是常数函数。这与泊松流形不同，后者也可能存在非常数的卡西米尔函数。Darboux-Weinstein 坐标通过局部形式，更明确地展现了泊松流形与辛流形之间的关系。

Darboux-Weinstein 坐标。我们从一个配备泊松双向量 L 的 $n=(2m+k)$ 维泊松流形 $\mathcal{M}$ 出发。在泊松流形的每一点附近，Darboux-Weinstein 坐标 $(x^i)=(q^a,p_b,u^\alpha)$（这里 a 从 1 到 m，而 α 从 1 到 k）给出了泊松双向量的一种局部形式

$$L=\frac{\partial}{\partial q^a}\wedge\frac{\partial}{\partial p_a}+\frac{1}{2}\lambda^{\alpha\beta}\frac{\partial}{\partial u^\alpha}\wedge\frac{\partial}{\partial u^\beta} \tag{13}$$

其中系数函数 $\lambda^{\alpha\beta}$ 在原点处等于零。若在局部公式（13）中 $k=0$，则右侧只剩第一项，泊松流形因此成为一个 $2m$ 维辛流形。

例如，牛顿力学就符合这种实现方式。另一方面，若在（13）中 $m=0$，则只剩第二项，它是一个完全退化的泊松双向量。许多泊松流形都属于这种形式，即李代数对偶空间上的李-泊松结构，包括刚体动力学、Vlasov动力学等。一般而言，泊松双向量既有辛部分，也有完全退化部分，例如第2节2.4中的重陀螺动力学。

当泊松双向量非退化时，它会生成一个辛泊松括号，并且与典范泊松双向量对易

$$\mathbf{L}_{can}=\begin{pmatrix}\mathbf{0} & \mathbf{I}\\ -\mathbf{I} & \mathbf{0}\end{pmatrix} \tag{14}$$

从以下意义上说

$$\mathbf{L}\cdot\mathbf{L}_{can}-\mathbf{L}_{can}\cdot\mathbf{L}=\mathbf{0}. \tag{15}$$

这一相容性条件将在第3节中用于衡量 DPNN 在学习辛泊松双向量时的误差。

2.2 $3D$ 哈密顿动力学

在本小节中，我们遵循 [24, 25, 26]，重点讨论三维泊松流形。 $3D$ 中的一项重要观察是，向量空间与反对称矩阵空间之间存在如下同构关系：

$$\mathbf{L}=\begin{pmatrix}0 & -J_z & J_y\\ J_z & 0 & -J_x\\ -J_y & J_x & 0\end{pmatrix}\leftrightarrow\mathbf{J}=(J_x,J_y,J_z). \tag{16}$$

这种同构使我们可以将雅可比恒等式 (9) 写成一个标量方程。

$$\mathbf{J}\cdot(\nabla\times\mathbf{J})=0, \tag{17}$$

参见例如 [25, 27, 28, 29]。雅可比恒等式 (17) 的通解为

$$\mathbf{J}=\frac{1}{\phi}\nabla C \tag{18}$$

对于任意函数 ons ϕ 和 C，其中 C 是一个卡西米尔函数。此时，哈密顿方程具有如下特殊形式

$$\dot{\mathbf{x}}=\mathbf{J}\times\nabla H=\frac{1}{\phi}\nabla C\times\nabla H. \tag{19}$$

请注意，通过交换哈密顿函数 H 和卡西米尔函数 C 的角色，可以得到同一系统的另一种哈密顿结构。在这种情况下，泊松向量定义为$\mathbf{J}=-(1/\phi)\nabla H$，哈密顿函数为 C。这是双哈密顿系统的一个例子，体现了可积性 [23, 30, 31, 32]。双哈密顿系统具有两种不同但相容的哈密顿表述。在 $3D$中，两个泊松向量，例如$\mathbf{J}_1$

和$\mathbf{J}_2$ 相容，当且仅当

$$\mathbf{J}_1 \cdot (\nabla \times \mathbf{J}_2) = \mathbf{J}_2 \cdot (\nabla \times \mathbf{J}_1). \tag{20}$$

在第3节中，我们将使用这一相容性条件来衡量 DPNN 学习三维泊松双向量时的误差。

刚体动力学。 我们来考虑一个三维哈密顿系统的例子：自由旋转的刚体。状态变量 $\mathbf{M} \in \mathcal{M}$ 表示在与刚体共同旋转的参考系中的角动量。其泊松结构为

$$\{F, H\}^{(RB)}(\mathbf{M}) = -\mathbf{M} \cdot \frac{\partial F}{\partial \mathbf{M}} \times \frac{\partial H}{\partial \mathbf{M}}, \tag{21}$$

见 [33]。泊松括号(21)是退化的，因为它保持 $\mathbf{M}$ 的模的任意函数不变。哈密顿函数为能量

$$H = \frac{1}{2}\left(\frac{M_x^2}{I_x} + \frac{M_y^2}{I_y} + \frac{M_z^2}{I_z}\right), \tag{22}$$

其中 I_x、 I_y 和 I_z 是刚体的转动惯量。在这种情况下，哈密顿方程

$$\dot{\mathbf{M}} = \mathbf{M} \times \frac{\partial H}{\partial \mathbf{M}} \tag{23}$$

给出了欧拉刚体方程 [34]。

2.3 $4D$ 哈密顿动力学

在四维空间中，我们考虑以下局部坐标 $(u, x, y, z) = (u, \mathbf{x})$。反对称矩阵 L 可表示为一对向量 $\mathbf{U} = (U^1, U^2, U^3)$ 和 $\mathbf{V} = (V^1, V^2, V^3)$，即

$$L = \begin{pmatrix} 0 & -U^1 & -U^2 & -U^3 \\ U^1 & 0 & -V^3 & V^2 \\ U^2 & V^3 & 0 & -V^1 \\ U^3 & -V^2 & V^1 & 0 \end{pmatrix}. \tag{24}$$

在此之后，我有了这个想法 ntification，雅可比恒等式 (9) 结果是由四个方程组成的偏微分方程组 [3 5]

$$\partial_u(\mathbf{U} \cdot \mathbf{V}) = \mathbf{V} \cdot (\partial_u \mathbf{U} - \nabla \times \mathbf{V}), \tag{25a}$$

$$\nabla(\mathbf{U} \cdot \mathbf{V}) = \mathbf{V}(\nabla \cdot \mathbf{U}) - \mathbf{U} \times (\partial_u \mathbf{U} - \nabla \times \mathbf{V}). \tag{25b}$$

注意，当且仅当 $\mathbf{U} \cdot \mathbf{V} = 0$ 时， L 是退化的（其行列式为零）。因此，对于退化的泊松矩阵，雅可比恒等式成立的条件是

$$\nabla \cdot \mathbf{U} = 0, \qquad \partial_u \mathbf{U} - \nabla \times \mathbf{V} = \mathbf{0}. \tag{26}$$

超可积性。 假设给定的动力学系统具有两个与时间无关的第一积分，分别记为 H_1 和 H_2。那么，当向量 $\mathbf{U}$ 和$\mathbf{V}$ 具有如下形式时

$$\mathbf{U} = \nabla H_1 \times \nabla H_2, \qquad \mathbf{V} = \partial_u H_1 \nabla H_2 - \partial_u H_2 \nabla H_1, \tag{27}$$

它们构成一个泊松双向量，因此特别满足雅可比恒等式 (26)。函数 H_1 和 H_2 是卡西米尔函数。对于哈密顿函数 H_3，其哈密顿动力学为

$$\dot{u} = -(\nabla H_1 \times \nabla H_2) \cdot \nabla H_3, \tag{28a}$$

$$\dot{\mathbf{x}} = (\nabla H_1 \times \nabla H_2)\partial_u H_3 + (\nabla H_2 \times \nabla H_3)\partial_u H_1 + (\nabla H_3 \times \nabla H_1)\partial_u H_2. \tag{28b}$$

通过互换函数 H_1、 H_2 和 H_3 的角色（其中两个是卡西米尔函数，一个是哈密顿函数），可以得到另外两种动力学的泊松实现。这是超可积（三哈密顿）表述的情形， [36]。

超可积四维哈密顿系统的两个泊松双向量必须满足相容性条件

$$\mathbf{U}_1 \cdot \mathbf{V}_2 + \mathbf{V}_1 \cdot \mathbf{U}_2 = 0 \tag{29}$$

其中，$\mathbf{U}_{1,2}$ 和 $\mathbf{V}_{1,2}$ 是由公式 (24) 确定的向量。我们将使用这一相容性条件来衡量 DPNN 在学习四维泊松双向量时的误差。

Shivamoggi 方程。四维泊松（非辛）系统的一个例子是 Shivamoggi 方程，它出现在磁流体力学的背景下，

$$\dot{u} = -uy, \qquad \dot{x} = zy, \qquad \dot{y} = zx - u^2, \qquad \dot{z} = xy, \tag{30}$$

见 [37, 38]。该方程组的第一积分为

$$H_1 = x^2 - z^2, \qquad H_2 = z^2 + u^2 - y^2, \qquad H_3 = u(z + x). \tag{31}$$

泊松矩阵 $N^{(i)}$ ($i = 1$、2、3) 对应哈密顿函数 H_1、 H_2 和 H_3 的向量 $\mathbf{U}_i$ 和 $\mathbf{V}_i$ 为

$$\begin{aligned}
&\mathbf{U}_1 = 2u\,(-y, z, y)\,, \qquad \mathbf{V}_1 = 2\left(u^2, y(x+z), u^2 - z(x+z)\right), \\
&\mathbf{U}_2 = 2\,(x+z)\,(0, u, 0)\,, \qquad \mathbf{V}_2 = 2\,(x+z)\,(x, 0, -z)\,, \\
&\mathbf{U}_3 = -4\,(yz, zx, xy)\,, \qquad \mathbf{V}_3 = 4u\,(-x, 0, z)\,,
\end{aligned} \tag{32}$$

分别。注意，这三个泊松矩阵都是退化的，因为对于所有 $i = 1$,2,3，$\mathbf{U}_i \cdot \mathbf{V}_i = 0$ 均成立。运动方程可以写成

$$X = \vartheta N^{(1)} \bar{\nabla} H_1 = \vartheta N^{(2)} \bar{\nabla} H_2 = \vartheta N^{(3)} \bar{\nabla} H_3, \quad \vartheta = -\frac{1}{4(x+z)}$$

¯在三种情况下，均可相差一个共形因子 ϑ 。注意，四维梯度记作 $\nabla H = (\partial_u H, \partial_x H, \partial_y H, \partial_z H)$。

2.4 半直积扩张至 a $6D$ 系统

六维哈密顿系统可以是辛系统，也可以是非辛（退化）系统。前者以三维粒子为代表，后者则以重陀螺动力学为例。由于三维粒子的演化是正则的，因此与二维动力学类似，我们仅回顾重陀螺动力学。

在均匀引力场中受到支撑的旋转刚体称为重陀螺 [39]。刚体的力学状态由质心的位置 $\mathbf{r}$ 和角动量$\mathbf{M}$描述。在这种情况下，泊松括号为

$$\{F, G\}^{(\mathrm{HT})}(\mathbf{r}, \mathbf{M}) = -\mathbf{M} \cdot \left(\frac{\partial F}{\partial \mathbf{M}} \times \frac{\partial G}{\partial \mathbf{M}} \right) - \mathbf{r} \cdot \left(\frac{\partial F}{\partial \mathbf{M}} \times \frac{\partial G}{\partial \mathbf{r}} - \frac{\partial G}{\partial \mathbf{M}} \times \frac{\partial F}{\partial \mathbf{r}} \right), \tag{33}$$

尽管该模型具有偶数维，但它并非辛系统。两个非恒定卡西米尔函数为 $\mathbf{r}^2$ 和 $\mathbf{M} \cdot \mathbf{r}$。在这种情况下，我们将哈密顿函数设为

$$H = \frac{1}{2} \left(\frac{M_x^2}{I_x} + \frac{M_y^2}{I_y} + \frac{M_z^2}{I_z} \right) + Mgl\mathbf{r} \cdot \boldsymbol{\chi}, \tag{34}$$

其中， $-g\boldsymbol{\chi}$是重力加速度向量。哈密顿方程为

$$\dot{\mathbf{M}} = \mathbf{M} \times \frac{\partial H}{\partial \mathbf{M}} + \mathbf{r} \times \frac{\partial H}{\partial \mathbf{r}} \tag{35a}$$

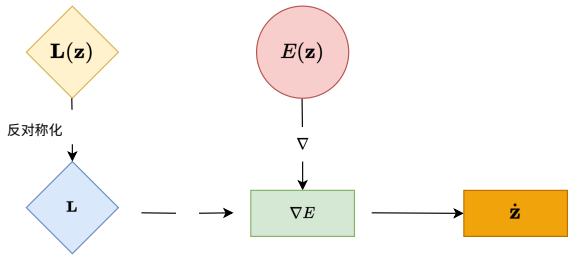
$$\dot{\mathbf{r}} = -\mathbf{r} \times \frac{\partial H}{\partial \mathbf{M}}. \tag{35b}$$

在接下来的章节中，我们将 DPNN 应用于此处回顾的模型，并展示 DPNN 能够从这些模型的模拟轨迹中提取泊松双向量和哈密顿量。

3 学习哈密顿系统

当我们获得哈密顿系统一条轨迹的一组快照时，如何识别其底层力学？换言之，如何从这些快照中学习泊松双向量和能量？机器学习为这类任务提供了一种稳健的方法。此前已有研究表明，机器学习可以重建 GENERIC 模型 [40, 11, 12], ，但泊松双向量通常是已知的且为辛结构。泊松神经网络 [20] 提供了一种学习非辛力学的方法，但这种方法依赖于在 Darboux-Weinstein 坐标中识别辛子动力学的维数，并转换到该坐标系。本文展示了一种稳健的方法，它既无需知道辛子系统的维数，也能在数据给定的坐标系中满足雅可比恒等式。因此，我们将这种方法称为直接泊松神经网络（DPNNs）。

DPNN 通过同时训练一个用于矩阵$\mathbf{L}(\mathbf{x})$的模型和一个用于哈密顿量$H(\mathbf{x})$的模型，直接学习哈密顿力学。编码$\mathbf{L}(\mathbf{x})$的神经网络只学习$\mathbf{L}$ 的上三角部分，因此自动满足反对称性。该网络包含一个采用 Softplus 激活函数的隐藏全连接层。用于学习 $H(\mathbf{x})$的网络结构相同。实际学习是在 PyTorch [41], 标准框架中使用 Adam 优化器 [42]实现的。损失函数包含训练数据与预测轨迹之间偏差的平方；预测轨迹由隐式中点法（IMR）通过数值求解精确方程（针对训练数据）或使用训练后的$\mathbf{L}(\mathbf{x})$和 $H(\mathbf{x})$模型求解哈密顿方程（针对预测数据）得到。尽管这样的模型能使验证轨迹与预测轨迹很好地吻合，但它不一定满足雅可比恒等式。因此，我们还使用另一种模型，将雅可比子（9）的平方加入损失函数，以较软的方式约束雅可比恒等式，见图 1。最后，在三维情形下，由于我们有雅可比恒等式（19）的通解，因此知道泊松双向量的形式。在这种情况下，可以将编码$\mathbf{L}$ 的神经网络简化为学习 $C(\mathbf{x})$的网络，从而自动满足雅可比恒等式，见图 2。总之，我们采用三种方法:



Fig图 1：同时学习能量和泊松双向量的方法方案 SJ（软雅可比）

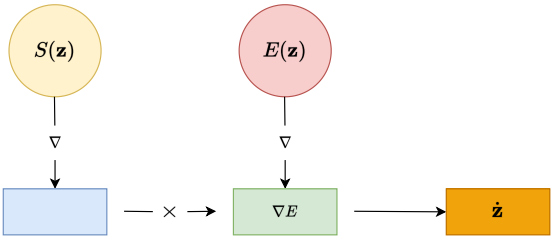


图 2：隐式强制满足雅可比恒等式的学习方法方案 IJ（隐式雅可比）。

- （**WJ**）在不满足$\mathbf{L}(\mathbf{x})$ 和 $H(\mathbf{x})$ 的雅可比恒等式的情况下进行训练。
- （**SJ**）使用软雅可比恒等式训练 $\mathbf{L}(\mathbf{x})$ 和 $H(\mathbf{x})$，其中雅可比子 (9) 的 L_2-范数是损失函数的一部分。
- **(IJ)** 使用基于三维中雅可比恒等式通解（19）且隐式满足雅可比恒等式的方法训练 $C(\mathbf{x})$ 和 $H(\mathbf{x})$。

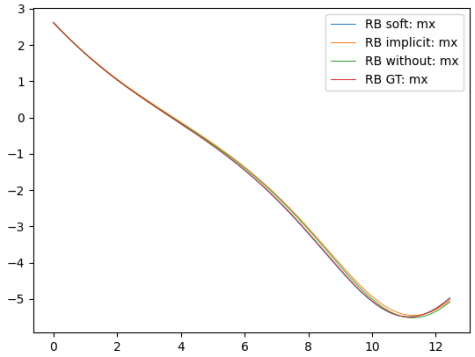
训练本身随后按以下步骤进行:

1. 模拟训练数据和验证数据。对于随机生成的一组初始条件，我们通过 IMR 模拟一组轨迹。然后将这些轨迹拆分为多个步骤，并将这些步骤的集合划分为训练集和验证集。2. 通过反向传播，在训练数据上训练神经网络 WJ、SJ 和 IJ 的参数，使损失函数最小化。随后在验证数据上评估损失函数，以报告误差。3. 随机选择一组新的初始条件，并使用 IMR 生成新的轨迹，从而得到真实值（GT）。4. 使用训练后的模型模拟具有真实值初始条件的轨迹，分别模拟$\mathbf{L}$ 和 H，并将其与 GT 进行比较。

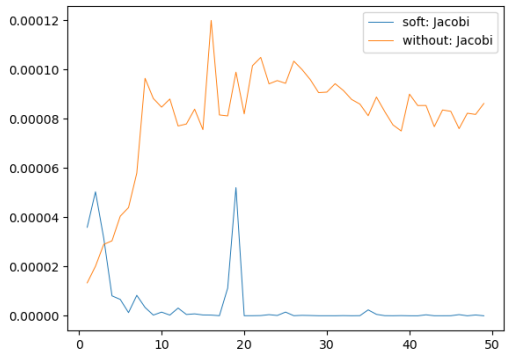
在以下各节中，我们将针对刚体动力学、二维粒子、Shivamoggi 方程、三维粒子和重陀螺动力学，说明这一学习过程。

3.1 刚体

图3a展示了真实值数据集中的一条刚体动力学（23）示例轨迹，以及由DPNN生成的、具有相同初始条件的轨迹。训练采用了200条轨迹，而真实值数据集包含400条轨迹。三种学习方法（WJ、SJ和IJ）的误差见表6。三种方法都能够很好地学习动力学。图3b展示了在验证集上计算得到的雅可比子的范数。对于IJ，雅可比子为零；对于SJ，其值很小；而对于WJ，其值不会趋于零。因此，IJ和SJ满足雅可比恒等式，而WJ不满足。



（a）精确轨迹（真实值）与轨迹的比较-
通过对学习得到的模型进行积分所获得的轨迹。三种方法都很好地拟合了这些轨迹。



（b）验证集上的雅可比恒等式误差。IJ 的误差由构造可知为零，SJ 的误差趋于零，而 WJ 的误差则不会。

图 3：刚体：学习得到的模型（WJ、SJ 和 IJ）与真实值的比较。

图4展示了学习泊松双向量（20）时的相容性误差。三种方法都能很好地学习泊松双向量，但 IJ 最精确，其次是 SJ 和 WJ。最后，图5展示了学习轨迹时的误差

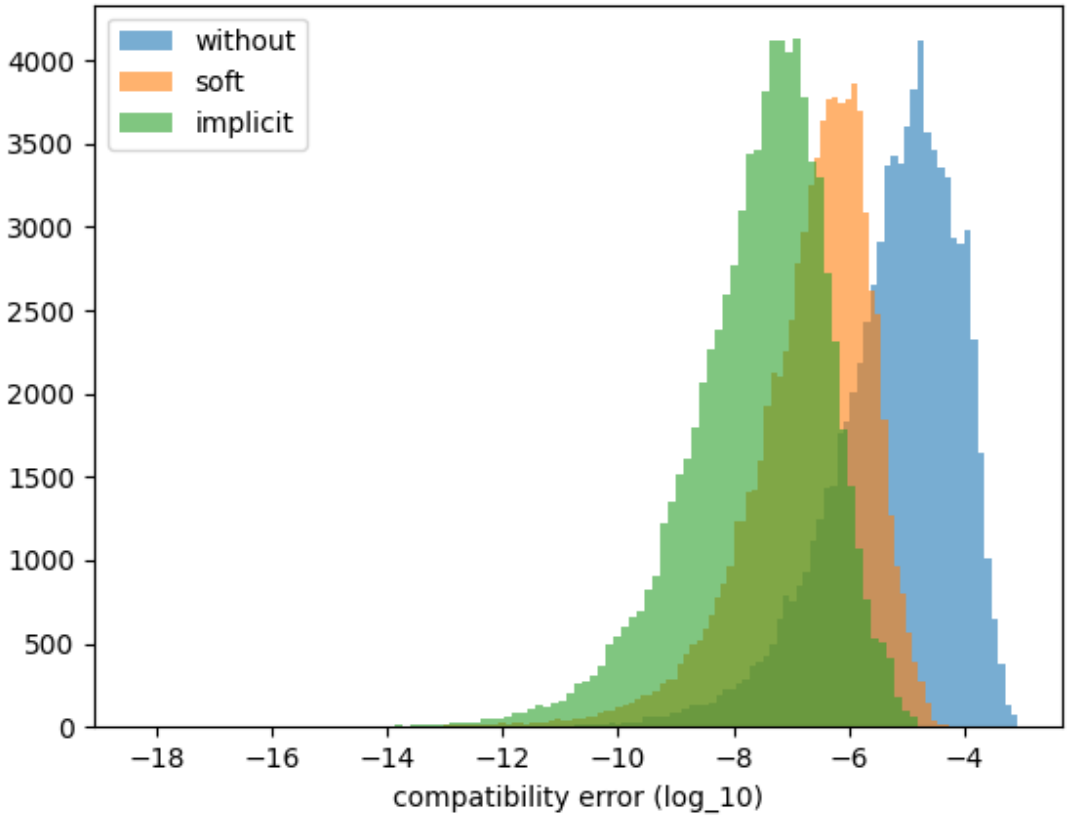


图 4：刚体：RB 的相容性误差以方程（20）的平方的 $\log_{10}$ 表示。误差分布近似服从对数正态分布。IJ 方法的相容性误差最低，其次是 SJ 方法和 WJ 方法。

$\mathbf{M}(t)$。这三种方法都能很好地学习轨迹，但在这种情况下，SJ 方法的表现略好。

表6展示了所有模型和所用学习方法的误差中位数。N/A 表示无需守恒的量。误差 $\Delta\mathbf{M}$ 是$\mathbf{M}^2$ 在所有时间步上的平方偏差中位数。误差 $\Delta\mathbf{r}$,$\Delta\mathbf{M}\cdot\mathbf{r}$、 $\Delta\mathbf{M}^2$, 和 $\Delta\mathbf{r}^2$ 的计算方式类似。在刚体情形下，误差 $\Delta\mathbf{L}$ 的计算方法是：将根据学习得到的$\mathbf{J}$ 计算的相容性条件（20）除以其迹，再乘以系数 1000（生成真实值时使用精确的$\mathbf{J}$ ），然后取其 L^2 范数并求以 10 为底的对数$_{10}$ 。在泊松双向量为辛的 P2D 和 P3D 情形下，误差计算为辛性条件（15）的平方和 Log_{10} 。在 Shivamoggi 方程情形下， $\Delta\mathbf{L}$误差是超可积相容条件（29）的平方 Log_{10} 。 $\Delta\det\mathbf{L}$ 误差是学习得到的 $\det\mathbf{L}$的平方中位数；在 Shivamoggi 和重陀螺情形下，由于行列式理论上应为零，因此这些值取对数。

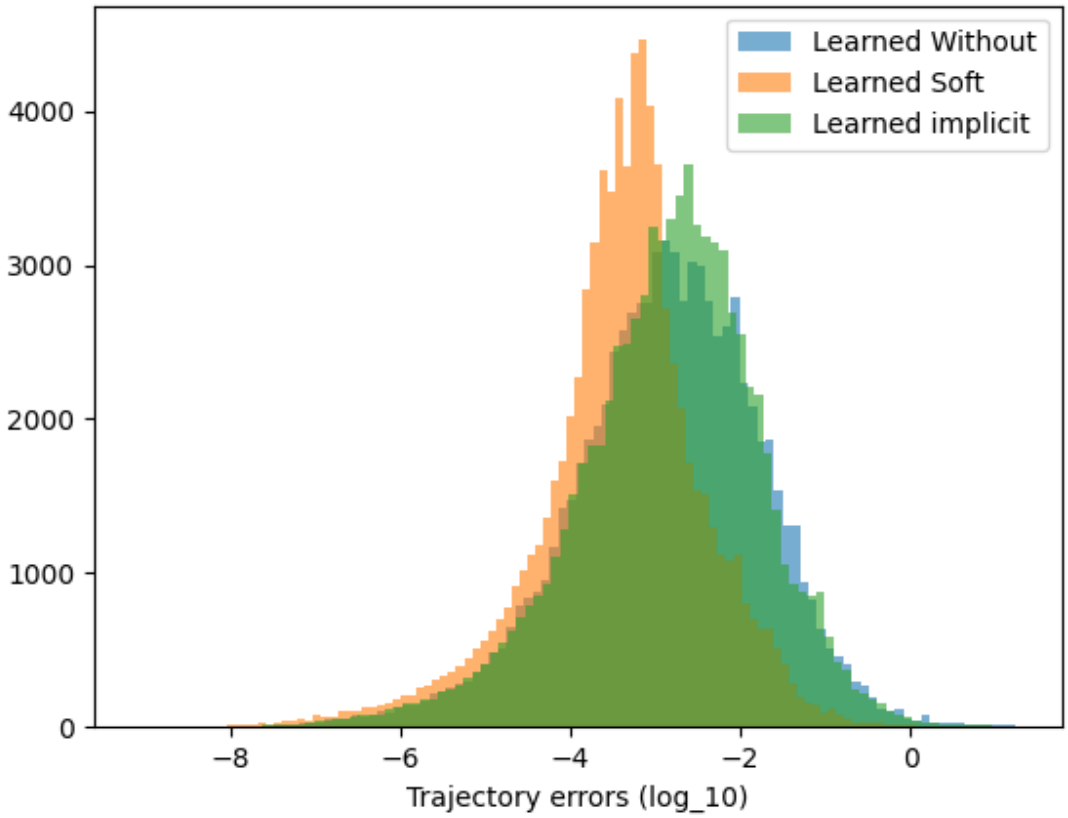


图5：刚体：$\mathbf{M}$ 中误差平方的 $\log_{10}$ 分布。

Mod.	Met.	$\Delta\mathbf{M}$	$\Delta\mathbf{r}$	$\Delta\mathbf{L}$	$\Delta\mathbf{M}^2$	$\Delta(\mathbf{r}\times\mathbf{M})$	$\Delta\mathbf{r}^2$	$\Delta(\mathbf{M}\cdot\mathbf{r})$	$\det\mathbf{L}$
RB	WJ	1.5E-03	N/A	-5.0	5.2E-04	N/A	N/A	N/A	0
	SJ	4.6E-04	N/A	-6.5	3.9E-04	N/A	N/A	N/A	0
	IJ	1.7E-03	N/A	-7.4	4.8E-04	N/A	N/A	N/A	0
P2D	WJ	3.2E-03	3.2E-03	-1.8	N/A	9.7E-03	N/A	N/A	0.83
	SJ	2.9E-03	2.9E-03	-1.8	N/A	8.6E-03	N/A	N/A	0.93
Sh	WJ	N/A	3.4E-04	-2.3	N/A	N/A	N/A	N/A	-2.1
	SJ	N/A	3.0E-04	-3.3	N/A	N/A	N/A	N/A	-3.4
P3D	WJ	1.9E-02	1.9E-02	-1.6	N/A	3.7E-02	N/A	N/A	0.91
	SJ	1.0E-02	9.1E-03	-1.9	N/A	3.5E-02	N/A	N/A	0.70
HT	WJ	3.1E-02	9.9E-03	N/A	N/A	N/A	1.9E-03	2.8E-03	-2.5
	SJ	1.6E-02	4.6E-03	N/A	N/A	N/A	1.3E-03	1.9E-03	-2.6

图6：刚体（RB）、二维粒子（P2D）、Shivamoggiequations（Sh）、三维粒子（P3D）和重陀螺（HT）的学习误差汇总。

3.2 二维粒子

二维势场中运动的粒子构成一个四维辛系统。我们使用 WJ 和 SJ 方法学习模拟得到的轨迹。由于目前没有适用于退化和辛泊松双向量的四维雅可比恒等式通解，因此未使用隐式 IJ 方法。学习结果见表6，WJ 和 SJ 对动力学的学习效果相当。图7展示了一条示例轨迹，图8展示了学习得到的 det(**L**) 分布。行列式中位数（经过归一化，使真实值中的行列式等于 1.0）接近这一数值，表明该系统是辛系统。

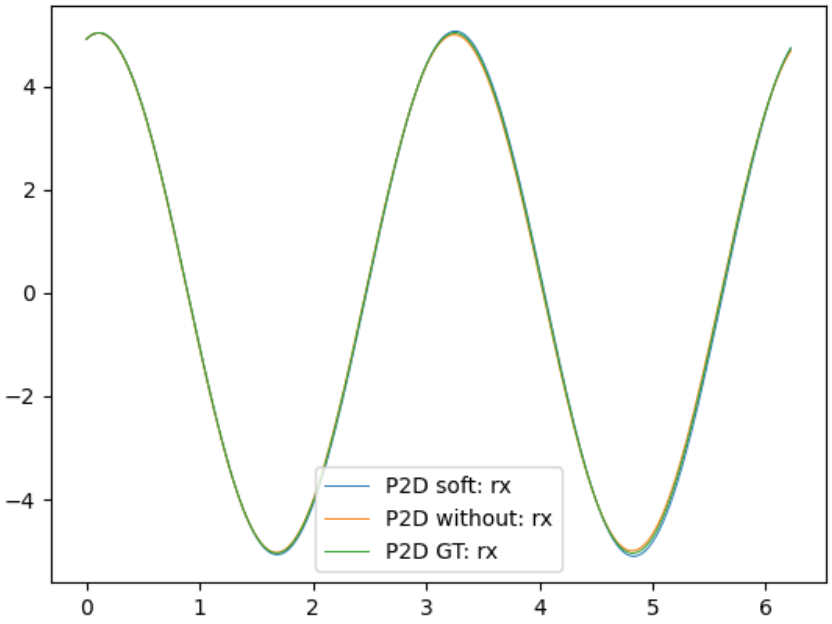


图7：P2D：一条示例轨迹。SJ 和 WJ 均能很好地学习二维粒子的动力学。

3.3 Shivamoggi 方程

Shivamoggi 方程 (30) 表示一个非辛的四维哈密顿系统，因此具有退化泊松双向量。方程在参数 $u \in [-0.5, 0.5]$、$x \in [-0.5, 0, 5]$、$y \in [-0.1, 0.1]$、$z \in [-0.5, 0.5]$ 的范围内求解。必须约束 $\mathbf{r} = (u, x, y, z)$ 的范围，因为例如当 $u = 0$ 时，解会爆炸 [37]。图9显示了示例轨迹上的 u 分量。图10显示了 $\log_{10}$(det(**L**)) 的分布，表明该系统确实是退化的。

与二维粒子（P2D）的辛情形中学习得到的 **L** 的行列式相比（见表6），Shivamoggi 情形下学习得到的行列式相当小（经过与 P2D 情形相同的归一化后）。因此，直接泊松神经网络能够区分辛哈密顿系统与非辛哈密顿系统。

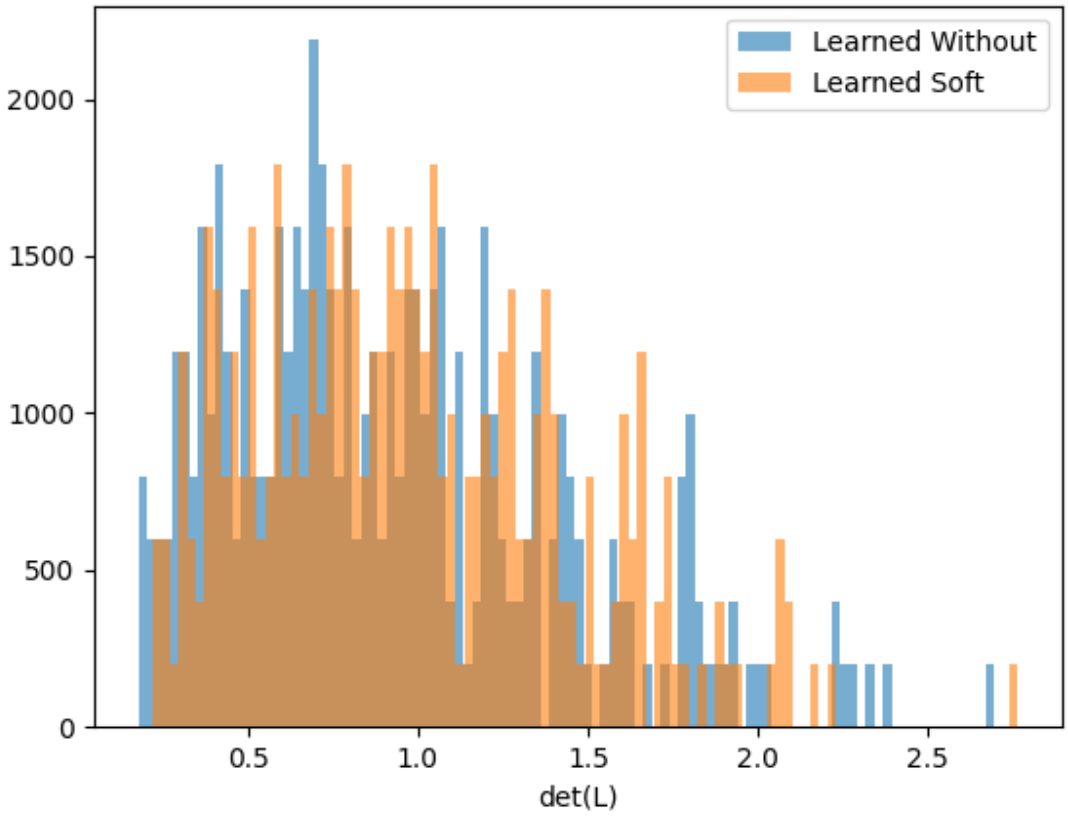


图8：P2D：学习得到的 det(**L**).

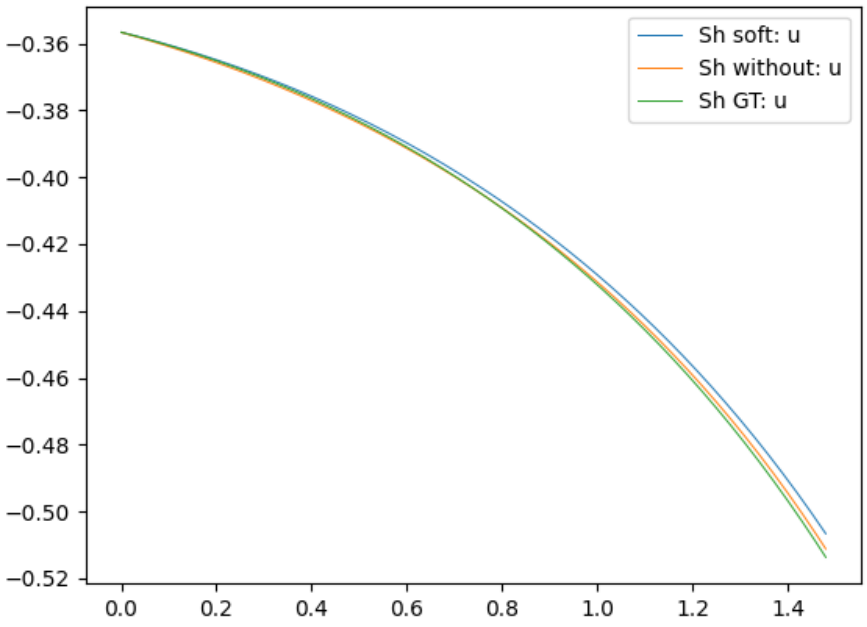


图9：Shivamoggi：一条示例轨迹，分量 $u_{\circ}$

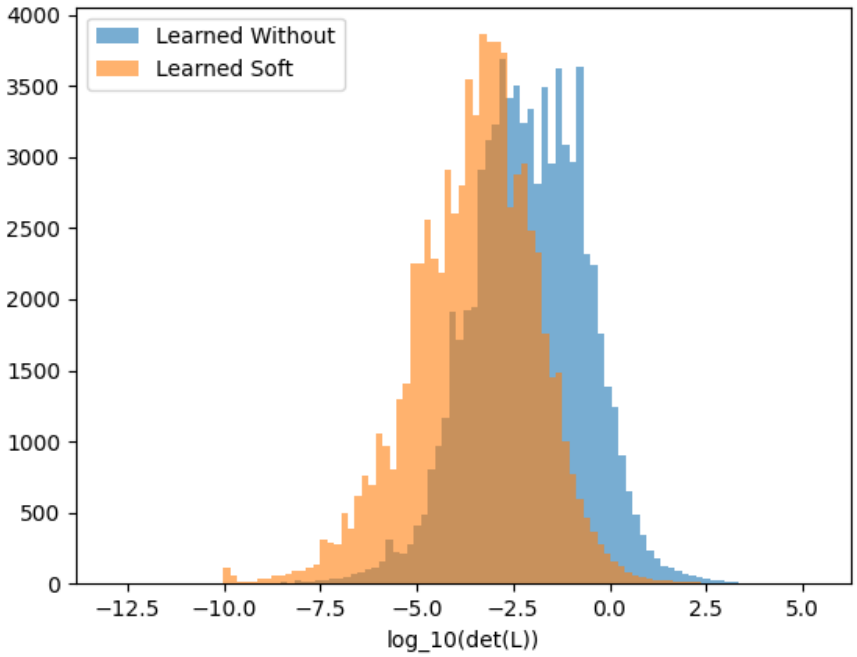


图10：Shivamoggi：学习得到的 $\log_{10}(\det(\mathbf{L}))$.

3.4 三维粒子

图11展示了从 GT 数据集中选取的三维空间粒子示例轨迹中的动量，以及 DPNN（WJ 和 SJ 两种类型）在相同初始条件下得到的轨迹。训练和验证分别使用了两组轨迹（分别包含 200 条和 400 条轨迹）。表6包含了……的数值

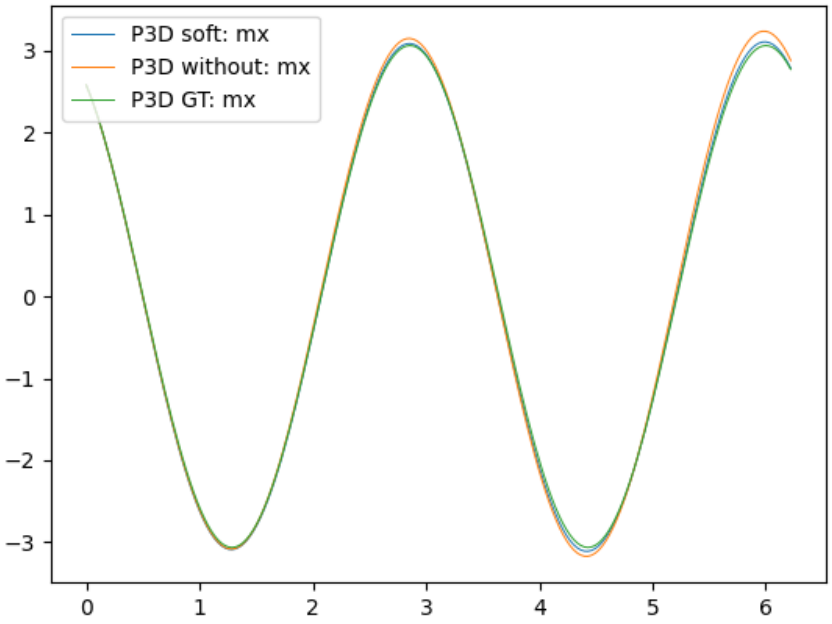
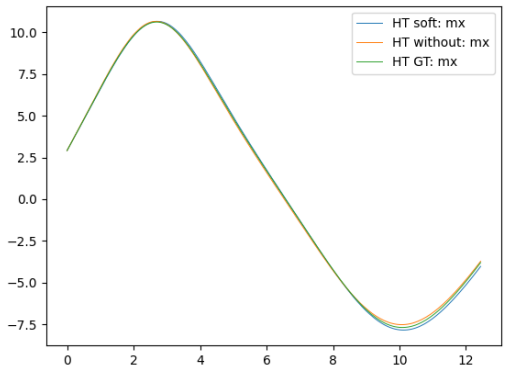


图11：P3D：三维简谐振子情况下，精确轨迹（GT）上的动量$\mathbf{M}(t)$与通过积分学习得到的模型（不含雅可比和含软雅可比）所获得轨迹的比较。

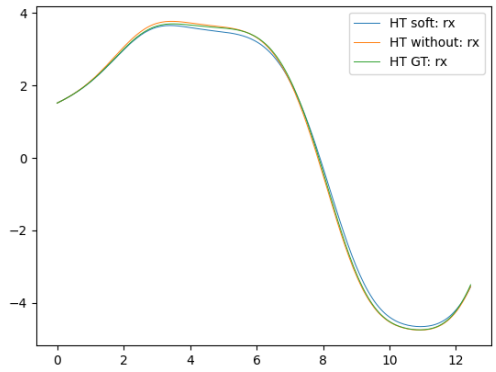
学习误差。WJ 和 SJ 都能很好地学习泊松双向量以及轨迹（因此也能很好地学习能量）。行列式中位数接近 1，这表明该系统是辛系统。

3.5 重陀螺

图12a和图12b展示了 GT 数据集中的重陀螺示例轨迹，以及 DPNN 根据相同初始条件得到的轨迹。训练和验证分别使用两组轨迹（分别包含300条和400条轨迹）进行。学习误差的数值见表6。例如，$\mathbf{L}$ 矩阵接近奇异，表明这是一个非辛系统，但 SJ 的学习效果略优于 WJ。与四维情形类似，DPNN 能够区分辛情形（P3D）和非辛情形（HT）。



（a）示例轨迹中的角动量 $\mathbf{M}(t)$。



（b）同一轨迹中的向量$\mathbf{r}(t)$。

图 12：重陀螺：在重陀螺情况下，对精确轨迹（真实值）以及通过对学习得到的模型（不使用雅可比恒等式和使用软雅可比）进行积分所获得的轨迹进行比较。

4 学习非哈密顿系统

现在，我们尝试将为学习纯哈密顿系统而开发的 WJ、SJ 和 IJ 方法应用于一个非哈密顿系统，具体而言，是一个耗散刚体。描述刚体耗散演化的一种方法称为能量型 Ehrenfest 正则化 [43]，，它在刚体的哈密顿演化中加入耗散项，使角动量的大小保持不变，同时耗散能量。其演化方程为

$$\dot{\mathbf{M}} = \mathbf{M} \times E_{\mathbf{M}} - \frac{\tau}{2} \boldsymbol{\Xi} \cdot E_{\mathbf{M}} \tag{1}$$

其中，τ 是正的耗散参数，$\boldsymbol{\Xi} = \mathbf{L}^T d^2 E \mathbf{L}$ 是一个正定对称矩阵（假设能量为正定），由刚体 $L^{ij} = -\epsilon^{ijk} M_k$ 的泊松双向量和能量$E(\mathbf{M})$构造而成。这些方程满足 $\dot{\mathbf{M}}^2 = 0$，而当 $\dot{E} \leq 0$ 时，其解会收敛到绕刚体主轴（具有最大转动惯量的轴）进行的纯旋转，这是物理上相关的解。通过求解方程（1）生成的轨迹的学习结果如图 13 所示。所有方法（WJ、SJ 和 IJ）都能在一定程度上学习这些轨迹，但 WJ 最为成功，其次是 SJ 和 IJ。由于 SJ，尤其是 IJ，利用了哈密顿系统更深层的性质（雅可比恒等式的软约束和精确有效性），因此它们在非哈密顿系统情况下的鲁棒性较差。

实际上，图 13 可以视为方程 (1) 非哈密顿性的一个标志。在 IJ 学习效果最佳、其次是 SJ 和 WJ 的系统中，系统更可能是哈密顿系统；与之相反，在 WJ 学习效果最佳、其次是 SJ 和 IJ 的非哈密顿系统中，情况则不同。换言之，DPNN 可以根据不同类型的直接泊松神经网络的表现顺序，区分哈密顿系统和非哈密顿系统。

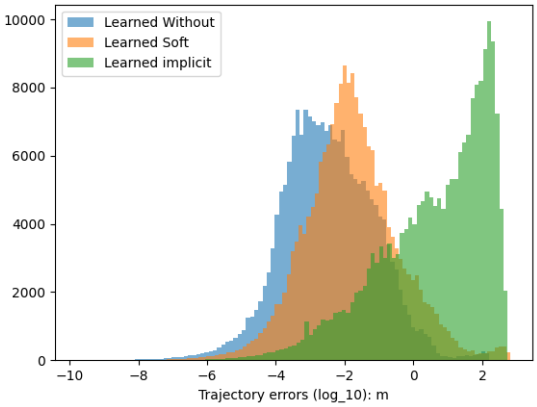


图13：采用假设系统为纯哈密顿系统的方法（WJ、SJ 和 IJ）学习耗散刚体动力学（1）时角动量 M 误差的分布。WJ 方法最为稳健，也能够较好地学习耗散系统（尽管效果不如纯哈密顿系统的情况）。SJ 方法还以软约束的方式施加雅可比恒等式，因此学习耗散系统的能力较弱。IJ 方法在纯哈密顿系统中的表现最佳（见表6），但在耗散情形下的学习能力最差。

5 结论

本文提出了一种从数据中学习哈密顿系统的机器学习方法。直接泊松神经网络（DPNN）直接学习力学系统的泊松双向量和哈密顿量，而不对系统结构作进一步假设。具体而言，DPNN 可通过测量所学习的泊松双向量的行列式，区分辛系统与非辛系统。

DPNN 分为三种形式：（i）不满足雅可比恒等式（WJ）；（ii）以软约束方式施加雅可比恒等式（SJ）；以及（iii）隐式满足雅可比恒等式（IJ）。尽管所有方法都能够学习动力学，但只有 SJ 和 IJ 也满足雅可比恒等式。典型情况下，IJ 对哈密顿模型的学习最为精确（见表6），其次是 SJ 和 WJ。

将三种形式的 DPNN 用于学习非哈密顿系统时，预期精度排序会反转：WJ 最为精确，其次是 SJ 和 IJ。精度排序的这种反转可作为区分哈密顿系统与非哈密顿系统的指标。

未来，　我们希望将 DPNN 扩展到由梯度动力学规定耗散的系统。

致谢

我们感谢 E. Cueto、F. Chinesta 和 B. Moya 就雅可比恒等式在物理系统学习中的作用进行的启发性讨论。MP 和 MŠ 得到了捷克资助署资助，资助项目编号为 23-05736S。

参考文献

[1] A. Einstein。引力场中光线偏折所产生的恒星透镜效应。《科学》，84(2188):506–507，1936.[2] S. F. Dermott、C. D. Murray、J. F. James、T. Gold 和 J. R. Donnison。小行星自转速率的分布与演化 [及讨论]。《伦敦皇家学会哲学汇刊：*A*辑，数学与物理科学》，313(1524):157–164，1984.[3] D. M. DiPietro 和 B. Zhu。哈密顿动力系统的辛积分符号回归。2022年9月。[4] G. E. Karniadakis、I. G. Kevrekidis、L. Lu、P. Perdikaris、S. Wang 和 L. Yang。物理信息机器学习。《自然综述：物理学》，3(6):422–440，2021.[5] S. Lie。《连续群讲义》（德文）。莱比锡：B. G. Teubner，1893.[6] H. Poincaré。《论力学方程的一种新形式》。《法国科学院学报》，132:369–371，1901.[7] R. Abraham 和 J. E. Marsden。《力学基础》。Benjamin/Cummings Publishing Co., Inc.，高级图书项目，马萨诸塞州雷丁，1978。

[8] V. I. Arnol’d. 经典力学的数学方法，数学研究生教材第60卷。Springer-Verlag，纽约，第二版，1989年。由 K. Vogtmann 和 A. Weinstein 从俄文译出。[9] P. Libermann 和 C. Marle. 辛几何与分析力学，数学及其应用第35卷。D. Reidel Publishing Co.，多德雷赫特，1987年。由 Bertram Eugene Schwarzbach 从法文译出。[10] J. E. Marsden 和 T. S. Ratiu. 力学与对称性导论，应用数学教材第17卷（第二版）。Springer-Verlag，纽约，1999年。[11] D. González、F. Chinesta 和 E. Cueto. 热力学一致的数据驱动计算力学。连续介质力学与热力学，2018年。[12] B. Moya、D. Gonzalez、I. Alfaro、F. Chinesta 和 E. Cueto. 利用数据学习晃荡动力学。计算力学，64(2):511–523，2019年8月。[13] M. Šípka 和 M. Pavelka. 学习GENERIC演化。2021年9月。[14] Y. D. Zhong、B. Dey 和 A. Chakraborty. 辛 ODE-Net：通过控制学习哈密顿动力学。2019年9月。[15] E. Dierkes 和 K. Flaßkamp. 在神经网络中考虑系统对称性来学习哈密顿系统。*IFAC-PapersOnLine*，54(19):210–216，2021年。[16] S. Greydanus、M. Dzamba 和 J. Yosinski. 哈密顿神经网络。载于 H Wallach、H Larochelle、A Beygelzimer、F d Alché-Buc、E Fox 和 R Garnett 编，神经信息处理系统进展，第32卷。Curran Associates, Inc.，2019年。[17] I. Vaisman. 泊松流形几何讲义，数学进展第118卷。Birkhäuser Verlag，巴塞尔，1994年。[18] A. Weinstein. 泊松几何。第9卷，第213–238页。1998年。辛几何。[19] J. E. Marsden 和 T. S. Ratiu. 力学与对称性导论：经典力学系统基础论述。应用数学教材。Springer New York，2013年。[20] P. Jin、Z. Zhang、I. G. Kevrekidis 和 G. E. Karniadakis. 通过泊松神经网络学习泊松系统及自治系统的轨迹。*IEEE*神经网络与学习系统汇刊，第1–13页，2022年。[21] M. Fecko. 物理学家的微分几何与李群。剑桥大学出版社，2006年。[22] A. Weinstein. 泊松流形的局部结构。微分几何杂志，18(3):523–557，1983年。[23] O. Esen、A. G. Ghose Choudhury 和 P. Guha. 三维混沌动力系统的双哈密顿结构。国际分岔与混沌应用科学工程杂志，26(13):1650215，11，2016年。[24] O. Esen、P. Guha 和 H. Gümral. $3D$-流，由矢势的旋度生成 & Maurer-Cartan方程。土耳其数学杂志，46(8):3234–3244，2022年。

[25] H. Gümral 和 Y. Nutku. 三个自由度动力系统的泊松结构。数学物理杂志，34(12):5691–5723，1993年。[26] H. Gümral. 三维系统中哈密顿量结构的存在性。动力系统应用进展，5(2):159–171，2010年。[27] B. Hernández-Bermejo. 三维泊松结构雅可比方程的新解。数学物理杂志，42(10):4984–4996，2001年。[28] B. Hernández-Bermejo. 三维雅可比恒等式的一个解可用于确定无穷多个解。物理快报 *A*，287(5-6):371–378，2001年。[29] B. Hernández-Bermejo. 雅可比方程的新解族：刻画、不变量与全局达布分析。数学物理杂志，48(2):022903，11，2007年。[30] O. Esen 和 P. Guha. 关于寻找由矢势旋度生成的三维流的广义哈密顿量描述。国际几何方法与现代物理学杂志，17(3):2050042，16，2020年。[31] H. Gümral. N分量 Kodama 方程的双哈密顿结构。物理学杂志 *A*，25(19):5141–5149，1992年。[32] H. Gümral 和 Y. Nutku. d-Boussinesq 方程与 Benney-Lax 方程的双哈密顿结构。物理学杂志 *A*，27(1):193–200，1994年。[33] V. I. Arnold. 关于无穷维李群的微分几何及其在理想流体动力学中的应用。傅里叶研究所年刊，16(1):319–361，1966年。[34] L. D. Landau 和 E. M. Lifshitz. 力学。Butterworth Heinemann。Butterworth-Heinemann，1976年。[35] O. Esen、A. G. Choudhury、P. Guha 和 H. Gümral. 四维动力系统的超可积情形。非线性动力学与混沌，21(2):175–188，2016年。[36] C. Gonera 和 Y. Nutku. 超可积 Calogero 型系统具有最大数量的泊松结构。物理快报 *A*，285(5-6):301–306，2001年。[37] P. Guha 和 A. G. Choudhury. 磁流体力学中出现的一组常微分方程的第一积分与哈密顿量结构。*AIP* 会议论文集，1582(1):116–123，2014年2月。[38] B. K. Shivamoggi. 在等剪切应变率等离子体流存在时双曲磁中性线附近的电流片形成：一个精确解。物理快报 *A*，258(2):131–134，1999年。[39] D. D. Holm、J. E. Marsden 和 T. S. Ratiu. Euler–Poincaré 方程与半直积及其在连续介质理论中的应用。数学进展，137(1):1–81，1998年。[40] D. González、F. Chinesta 和 E. Cueto. 从数据中学习超弹性模型的修正。材料前沿，6(14)，2019年。[41] A. Paszke、S. Gross、F. Massa、A. Lerer、J. Bradbury、G. Chanan、T. Killeen、Z. Lin、N. Gimelshein、L. Antiga、A. Desmaison、A. Kopf、E. Yang、Z. DeVito、M Raison、A. Tejani、S. Chilamkurthy、B. Steiner、L. Fang、J. Bai 和 S. Chintala. PyTorch：一种命令式风格的高性能深度学习库。载于 H. Wallach,

H. Larochelle、A. Beygelzimer、F. d Alché-Buc、E. Fox 和 R. Garnett（编），第*32*届神经信息处理系统进展，第8024–8035页。Curran Associates, Inc.，2019年。

[42] D. P. Kingma 和 J. Ba。《一种随机优化方法：Adam》，2014年。

[43] M. Pavelka、V. Klika 和 M. Grmela。哈密顿系统的埃伦费斯特正则化。《物理学*D*：非线性现象》，399:193–210，2019年。