Claude独立攻克理论物理前沿难题，全程无人指导，花费不到两千美元

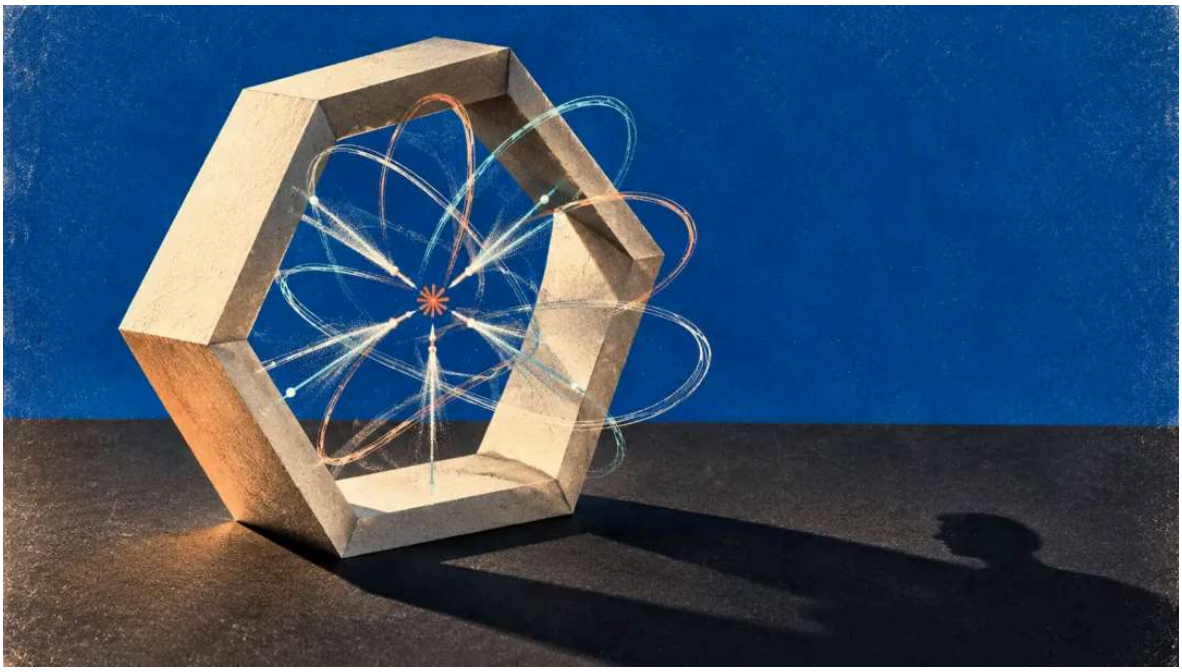
返朴 2026年9月27日 09:26 北京

以下文章来源于DeepTech深科技 ，作者杨梦

DeepTech深科技
DeepTech是一家专注新兴科技的创新赋能机构，致力于推动科学与技术的创新进程。...

返朴
科普中国子品牌，溯源守拙，问学求新，做科学家领航的科普清流。
2204篇原创内容

公众号



撰文 | 杨梦

9 月的第一天，斯坦福大学（Stanford University）物理学家兰斯·狄克逊（Lance Dixon）收到 Anthropic 的消息：Claude 已经完成了平面 N=4 超杨-米尔斯理论（planar N=4 super Yang-Mills theory）中九圈六胶子散射振幅的计算，请他帮忙验证。自 2023 年以来，兰斯与合作者一直致力于解决这一难题。

几天后，中国科学院理论物理研究所的何颂团队也告诉兰斯，他们借助 GPT-6 独立得到了九圈振幅的符号（symbol）。

当地时间 9 月 25 日，Anthropic 邀请前理论物理学家、科学作家马特·冯·希佩尔（Matt von Hippel）撰文，正式公布他们对这一问题的解决过程，兰斯在文末附上了自己的验证经过。

这可能是大模型首次在几乎无人监督的情况下，独立完成一项前沿理论物理计算。整个过程中，Claude 只收到一句话的任务描述和若干条类似于“继续做”的指令，连续运行数日，总花费约一两千美元。



（来源：Anthropic）

算到第九圈，究竟难在哪？

在物理学中，散射振幅（scattering amplitude）用于描述粒子碰撞后以多大概率产生特定结果，是理论预测与对撞机实验相互对照的基础。物理学家通常用微扰展开计算振幅，逐级加入量子修正，每一级就是一个“圈”（loop）。圈数越高，结果越精确，计算复杂度也随之呈指数甚至阶乘式增长。实际研究中，大多数振幅只算到两圈，少数能算到三圈。

平面 N=4 超杨-米尔斯理论是振幅研究者的试验场，其基于杨振宁和罗伯特·米尔斯（Robert Mills）1954 年提出的杨-米尔斯理论，1970 年代由多位物理学家加入超对称性发展而来。该理论不描述真实世界，但粒子种类之间的精细平衡让计算大幅简化，研究者借它打磨新方法，再推广到更贴近现实的理论。

兰斯团队从 2011 年起用自举法（bootstrap）逐圈推进。他们先根据物理原理圈定答案所在的函数空间，用一套特定的“字母表”把每个候选函数编码成序列，即保留函数主要结构的

“符号”，再对其逐条施加已知约束、排除候选，直到只剩唯一解。整个过程类似于填数独。马特在转行成为科学作家之前是其中多篇论文的合作者。

到 2019 年，他们才算到六圈和七圈。如果继续推进，直接自举的计算量会变得相当繁重。但 2021 年，兰斯发现，两个胶子碰撞产生四个胶子的振幅，与两个胶子碰撞产生一个胶子和一个希格斯玻色子的“形状因子”（form factor）数值极为相似，后者计算起来更容易。

团队随后证实，把前者符号中每个字母序列的顺序倒过来，经过变量替换就能得到后者，这种现象被命名为“反极对偶性”（antipodal duality）。2022 年，兰斯与合作者先把形状因子算到八圈，2023 年再借助反极对偶性从中反推出八圈振幅。

从那以后，兰斯计划按照这一方法继续完成第九圈的计算。直到今年 9 月，两条不同程度依赖 AI 的路线几乎同时抵达终点。

面向 AI 发起的公开挑战，被 Claude 接下了

今年 8 月 7 日，马特在个人博客发文，称学者们对 AI 的怀疑往往遵循同一个模式：总认为自己的领域更特殊，直到 AI 真的成功解出难题。但他依然认为，自己的老本行是个例外。

在马特看来，AI 此前在振幅领域的成果都停留在学生水平。他举了两个例子：判断 N=8 超引力在七圈是否发散，以及把 N=4 超杨-米尔斯理论的六粒子振幅算到九圈。

和其他学者关注 AI 能否提出新理论不同，马特更看重 AI 能否用学者可及的资源，绕过所有人都以为绕不过去的计算限制。

Anthropic 的两位物理学家利亚姆·菲茨帕特里克（Liam Fitzpatrick）和悉达多·米什拉-夏尔马（Siddharth Mishra-Sharma）读到了这篇挑战，并在询问 Claude 的把握后选定了九圈振幅。

随后，他们在科研工作台 Claude Science 上调用 Fable 5.1，初始提示只有一句话：计算平面 N=4 超杨-米尔斯理论中九圈的六粒子振幅。此后，两人的干预仅限于让它继续，例如，“我去睡了，几个小时内联系不上，继续做，直到我喊停，每 4 到 6 小时汇报一次进展”。

Claude 最终走通了两条路线。第一条复现了兰斯团队的间接思路，第二条是兰斯认为不可行的直接路线：在六边形函数空间里直接自举九圈振幅的符号。两条路线的结果在全部 107,053 个非零系数上逐一吻合。作为对照，同一套程序还在随机抽取的 1,000 个字母序列上复现了已发表的八圈结果。

9 月 16 日，悉达多在个人网站上以计算机可读文件的形式公布了九圈振幅的符号和完整函数，格式与兰斯团队此前发布六圈、七圈、八圈结果时一致。

成本方面，自举计算用 Python 语言和开源符号计算库 SymPy 完成，约合 100 美元，相当于 96 个中央处理器（CPU）运行一周。加上 Claude 长时间运行的费用，整个项目总计耗费一两千美元。

Anthropic 团队没有提前与相关领域的学者打招呼，算完之后才联系兰斯，马特得知后也感到有些措手不及，他事后建议 Anthropic，如果还想挑战清单上的其他问题，应当事先联系相关科学家。

兰斯团队原本就在用自己训练的小型模型尝试预测更高圈数的结果，他们曾表示有能力验证机器给出的任何候选答案。兰斯从振幅倒推回形状因子核对 Claude 的结论，二者完全吻合。

兰斯还注意到，Claude 用到的全是他和合作者多年发展起来的方法，最终结果也按照他们此前约定的格式呈现。他认为，Claude 对他们 2019 年和 2023 年那两篇论文的理解可能在合著者之外的所有人之上。

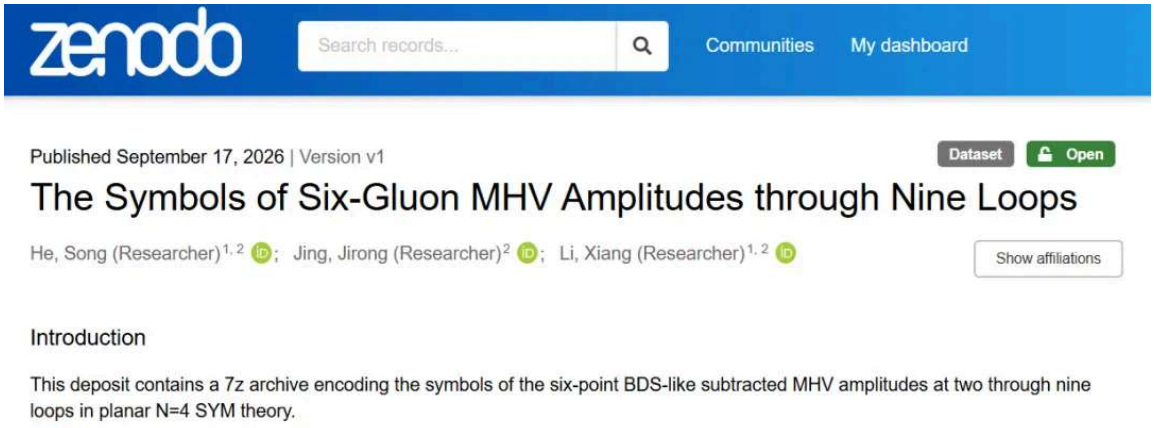
中国声音

在中国，9 月中旬，中科院理论物理研究所的何颂研究员与团队在开放科研数据平台 Zenodo 上发布数据集，公开了从二圈到九圈振幅的符号和完整函数。

何颂长期从事散射振幅研究，他的团队多年来一直用符号自举方法研究 N=4 超杨-米尔斯理论，2025 年 11 月，何颂与合作者发表了七粒子振幅五圈符号的结果。

在最新研究中，何颂团队借助 GPT-6 计算了部分约束条件，整体计算框架仍由人类研究者亲自搭建。马特据此认为，九圈并不是遥不可及的目标。

后续，兰斯、何颂及各自的合作者将继续撰写论文，对结果进行系统的解释和分析，AI 的戏份暂时告一段落。



（来源：Zenodo）

划定 AI 在物理学中的能力边界

马特发起挑战时，期待看到 AI 找到某种出人意料的新思路，但从结果来看，Claude 依然沿用前人的成熟方法，只是多投入了一些算力。

这次计算最大的难点其实是可靠性。整套计算流程非常脆弱，任何一步出错都会导致全盘失败，而且很难定位问题出在哪。兰斯还表示，构造过程中有大量细节从未被写进论文，Claude 缺乏参照，必须从零开始编写全部代码。

马特写道，如果还有人觉得 AI 错误百出、不堪大用，这件事应该会改变他们的看法。

AI 的发展速度也值得关注。今年 3 月，Anthropic 发布的物理研究案例中，AI 还像一名需要手把手指导、频繁犯错的学生；半年后，它就有能力独立完成一项通常由顶尖专家承担的前沿计算。马特提醒外界，预测 AI 的未来时，不应假定眼下就是能力和成本的极限。

在兰斯看来，大模型可以完整执行复杂流程并调度算力，是一项了不起的成就；何时它能先于人类提出新的物理原理和洞见，才真正值得物理学家反思。

马特也认可这一点，他写道，如果一个问题你自己解决不了，但你相信另一个领域的专家或者更擅长编程的人能解决，那 AI 多半也能解决。但反过来，量子引力等问题取决于理论取舍而非技术能力，依赖实验的领域也另当别论。

挑战清单上列举的问题还停留在“玩具模型”的层面，真实世界的振幅计算，与对撞机物理和引力波物理密切相关，参与者更多，竞争也更激烈，现成的突破口可能更少。

马特建议这些领域的研究者尽快行动起来，检验现有 AI 科研框架能否一次性完成他们手头的计算，同时事先想好核对方法。他预估，AI 或许可用合理预算再往前推进一圈。

参考内容

[1] https://www.anthropic.com/research/yes-claude-can-do-nine-loops
[2] https://4gravitons.com/2026/09/25/it-got-to-my-field/
[3] https://4gravitons.com/2026/08/07/it-only-counts-when-ai-gets-to-my-field/
[4] https://smsharma.io/cosmic-nine-loops/
[5] https://doi.org/10.5281/zenodo.22800071

注：封面/首图由AI辅助生成

本文经授权转载自微信公众号“DeepTech深科技”。

相关阅读

1. AI的物理学：用统计物理理解智能
2. “最大的障碍来自物理学界”，MIT物理学家反思AI与物理的结合
3. AI做出四大顶刊级研究，数学家们为什么不安？
4. 菲尔兹奖得主迪米尼-科潘：我不用AI做创造性工作

近期推荐

1. 一个神话版的鸡兔同笼问题，竟蕴含着深刻的前沿物理
2. AI时代，什么才是真正属于人类的数学？
3. 为什么越来越多科研基金，开始用抽签决定谁拿经费？
4. 专访张寿武、袁新意：一篇被拒稿的论文，如何长成一套受欢迎的奠基性理论
5. 王虹、邓煜和北大数学：顶尖人才的成长需要怎样的土壤？

特 别 提 示

1.『返朴』提供按月检索功能。关注公众号，回复四位数组成的年份+月份，如“1903”，可获取2019年3月的内容目录，以此类推。

2. 投稿、转载、应聘、商务合作等，皆可联系：fanpusci@163.com



长按下方图片关注「返朴」，查看更多历史文章



微信实行乱序推送，常点“在看”，可防失联