北京国际数学研究中心
BEIJING INTERNATIONAL CENTER FOR MATHEMATICAL RESEARCH

Titles and Abstracts of 45-Minutes Talks

(Sorted by talk sequence)

Towards an Understanding of the Principles behind Deep Learning

Weinan E (Peking University)

Abstract: The field of deep learning is evolving rapidly, driven by the availability of the vast amount of data and computing resources. Deep learning techniques have also evolved in several different ways, including different formulations such GAN and the diffusion model, different architecture such as CNN and transformers, and different training protocols such as BERT and GPT. This evolution has largely been empirical. Consequently there are a lot mysteries, surprises and “black magics” in this field. Is it possible to decipher some kind of guiding principles behind this? In this talk, we will discuss our thoughts along this line. Specifically, we will discuss how simple mathematical concepts such as symmetry and stability can be used as guiding principles for designing and understanding neural network models.

从 Arikan 访华谈起

Zhiming MA (Chinese Academy of Sciences)

Abstract: 报告人将从 5G Polar 码之父 Arikan 访华谈起，介绍我们在信息编码领域的一些研究成果。同时也回顾近年来我们推动应用数学落地的历程并介绍一些其他应用数学成果。

Non-vanishing of L-values

Ye TIAN (Chinese Academy of Sciences)

Abstract: We discuss non-vanishing of L-values of elliptic curves or modular forms in nice families.

Restriction theory and projection theorems

Hong WANG (NYU Courant)

Abstract: Restriction theory studies functions whose Fourier transforms are supported on some curved manifold in R^n (for example, solutions to the linear Schrodinger equation or to the wave equation). Projection theorems study the Hausdorff dimension of fractal sets under orthogonal projections from R^n to its subspaces. We will survey some recent works in both fields and discuss their interactions.

Compactness of moduli spaces parametrizing Fano varieties

Chenyang XU (Princeton University)

Abstract: (Joint with Blum, Liu, Zhuang) One of the most profound aspects of K-moduli spaces parametrizing K-polystable Fano varieties is their compactness. In this talk, we would like to present a new proof of the compactness. The same as in the original proof by Blum-Halpern-Leistner-Liu-Xu, the new proof still needs the Higher Rank Finite Generation Theorem established by Liu-Xu-Zhuang, but it replaces Halpern-Leistner’s Theta stratification theory by a relative stability theory. As a result, the current proof completely relies on birational geometry arguments without invoking knowledge from stack theory.

Towards AI Virtual Cell Through Dynamical Generative Modeling of Single-cell Omics Data

Peijie ZHOU (Peking University)

Abstract: Reconstructing continuous cellular dynamics from sparse, high-dimensional single-cell omics data remains a fundamental challenge in computatoinal biology. Recently, a paradigm shift has been witnessed by leveraging artificial intelligence—specifically, dynamical generative modeling—to develop an AI virtual cell, a predictive digital twin capable of simulating cellular behavior across time and space. In this talk, we introduce our recent attempts that integrate generative AI models with partial differential equations (PDEs) and optimal transport (OT) theory to infer latent dynamics from scRNA-seq data. For spatial transcriptomics data, we extend this method with stVCR, a generative model that aligns transcriptomic snapshots across biological replicates and temporal stages. To further infer stochastic dynamics from static data, we explore a regularized unbalanced optimal transport (RUOT) formulation and its theoretical connections to the Schrödinger Bridge and diffusion models. Together, these works suggest how generative AI and mathematical tools could work together to unify dynamical modeling, spatial reconstruction, and stochastic inference—transforming fragmented omics data into a predictive virtual cell.

Volume of the moduli of Shtukas and higher derivatives of zeta functions

Zhiwei YUN (Massachusetts Institute of Technology)

Abstract: The volume of a locally symmetric space, properly normalized, is a product of special values of zeta functions. More generally, Hirzebruch's proportionality theorem (extended by Mumford) tells us how to integrate any Chern class polynomial on a locally symmetric space.

We give an analog and extension of these results in the function field case where new phenomena show up. Locally symmetric spaces will be replaced by the moduli space of Drinfeld Shtukas with multiple legs, and special values of zeta functions will be replaced by a linear combination of their derivatives of various order. This is joint work with Tony Feng and Wei Zhang.

球的稳定同伦群和 Kervaire 不变量问题

Zhouli XU (University of California, Los Angeles)

Abstract: 我将介绍球的稳定同伦群理论与计算的近期发展，以及林伟南、王国祯与我在 Kervaire 不变量问题最终情形——126 维上的相关工作。

Stochastic quantization of Yang-Mills

Hao SHEN (University of Wisconsin - Madison)

Abstract: The study of quantum Yang-Mills (as well as quantum field theories in general) at rigorous level has posed various challenges to mathematicians. I will talk about a dynamical approach to the problem. The dynamical formulation brings together many mathematical tools, including analysis, PDE, stochastic calculus, mixing techniques, and ideas from geometry. In particular, I will discuss the construction of gauge covariant Langevin dynamics via the theory of regularity structures. Based on joint work with A.Chandra, I.Chevyrev, and M.Hairer.

Toward a Mathematical Theory of Deep Learning: Lessons from Personal Research

Weijie SU (University of Pennsylvania)

Abstract: A century ago, breakthroughs like relativity and quantum mechanics were enabled by mathematical theories or developed concurrently with their mathematical underpinnings. Today, the development of AI presents a stark contrast: it is predominantly driven by empirical progress, while theories significantly lag behind and offer limited guidance.

In this talk, I will share my perspective on the ongoing efforts to develop mathematical foundations for deep learning, drawing from my personal research experiences and insights. We first argue that an effective mathematical theory for deep learning should fundamentally address three core characteristics simultaneously: the hierarchically structured architectures of networks, the iterative optimization driven by stochastic gradient methods, and the compressive evolution of information extracted from data. As an instantiation of these ideas, we will introduce neurashed, a phenomenological model we have developed; local elasticity, an empirical inductive bias we have identified; and the equi-learning law, a geometric simplicity we observe in deep learning training. Finally, we will examine the current limitations of existing theories and discuss many opportunities awaiting the mathematics community to contribute to advancing the theoretical foundations of deep learning.