

Transformers: From Dynamical Systems to Autoregressive In-Context Learners

Friday 26 September 2025 14:00 (45 minutes)

Transformers have enabled machine learning to reach capabilities that were unimaginable just a few years ago. Despite these advances, a deeper understanding of the key mechanisms behind their success is needed to build the next generation of AI systems. In this talk, we will begin by presenting a dynamical system perspective on Transformers, demonstrating that they can be interpreted as interacting particle flow maps on the space of probability measures, solving an optimization problem over a context-dependent inner objective. We will also discuss the impact of attention map normalization on Transformer behavior in this framework. We will then focus on the causal setting and propose a model to understand the mechanism behind next-token prediction in a simple autoregressive in-context learning task. We will explicitly construct a Transformer that learns to solve this task in-context through a causal kernel descent method, with connections to the Kaczmarz algorithm in Hilbert spaces, and discuss connections with inference-time scaling.

References

- Sander, M. E., & Peyré, G. (2025). Towards understanding the universality of transformers for next-token prediction. International Conference on Learning Representations (ICLR).
- Sander, M. E., Giryes, R., Suzuki, T., Blondel, M., & Peyré, G. (2024). How do transformers perform in-context autoregressive learning? International Conference on Machine Learning (ICML).
- Sander, M. E., Ablin, P., Blondel, M., & Peyré, G. (2022). Sinkformers: Transformers with doubly stochastic attention. International Conference on Artificial Intelligence and Statistics (AISTATS).

Presenter: SANDER, Michaël