**Normalizing flows**

Simple distribution complex distribution (through invertible and differentiable functions)

Flow based models

The network is invertible if is contractible.

Lipchitz condition must be followed.

**Continuous Normalizing Flows (CNFs)**

Describe the evolution of hidden states

Apply **adjoint sensitivity** to update parameters .

Problem: compute trace of a matrix is expensive

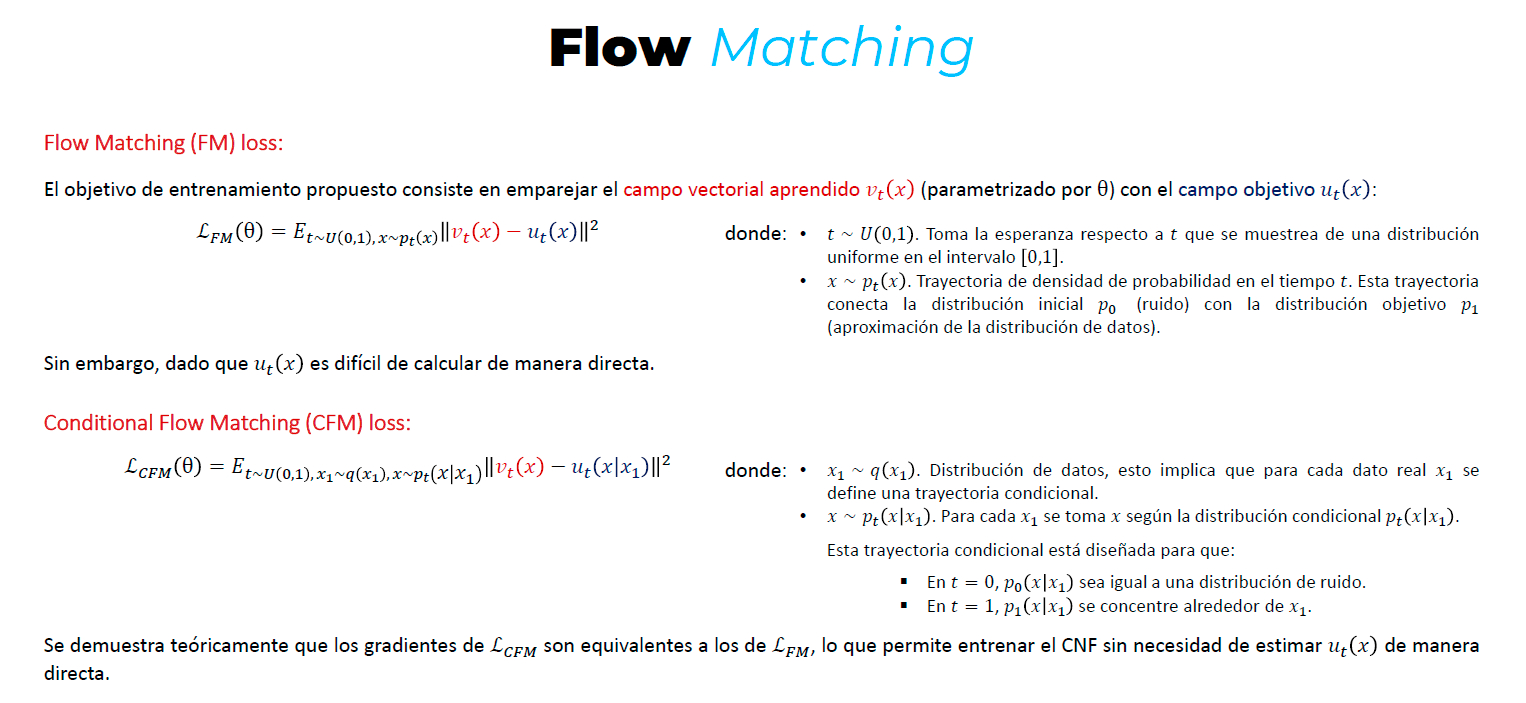
**FFJORD**

Free-form Jacobian of Reversible Dynamics

Solves the trace problem (stochastic estimation)

Flow Matching

Learned vector field



Large Language Models

Byte-Pair Encoding

Words

Splits

Pairs frequencies

Merges

2. Llama models and DeepSeek

Llama 3

Grouped Query Attention (GQA)

Multimodal architecture

Deep Seek

Top k, top p, temperature

Memory optimizations

Latent space

**3. Prompting**

IO, CoT, CoT-SC, ToT

Chain of thoughts (CoT)

Tree of thoughts (ToT)

Active prompt:

1. Uncertainty estimation (fix the prompt?)
2. Selection

**4. LLM fine-tuning**

LoRA: add layers to detect specific changes to modify the original weights!

QLoRA: quantization (16 bits to 4 bits)

**5. Retrieval augmented generation**

RAG