Mode Seeking

Input: $P = \{p_1, \dots, p_n\} \subset \mathbb{R}^d$.

Hyp: – the p_i are sampled iid according to some unknown probability distribution μ with (unknown) density f w.r.t. the Lebesgue measure,

– f is regular enough, typically a Morse function: twice differentiable, finitely many critical points, non-degenerate (Hessian matrix is non-singular), all distinct critical values.

Note: the gradient vector field $x \mapsto \nabla f(x)$ is Lipschitz continuous $\Rightarrow$ it can be integrated into a gradient flow $\Phi : \mathbb{R}_+ \times \mathbb{R}^d \to \mathbb{R}^d$ whose trajectories are solutions of the ODE

$$\begin{cases} \gamma_x'(t) = \nabla f \circ \gamma_x(t) \\ \gamma_x(0) = x \end{cases}$$

(Cauchy-Lipschitz Thm.)

Thm (from Morse theory) If f is Morse, then almost every point of $\mathbb{R}^d$ ends up at a maximum of f when following the gradient flow (integration of gradient vector field) of f.

→ **principle:** cluster (almost) $\mathbb{R}^d$ by the ascending regions of the peaks of f.

$\mathrm{Asc}(p) = \{x \in \mathbb{R}^d \mid p \in \mathrm{Im}\, \gamma_x\}$.

↳ in practice: simulation by hill-climbing.