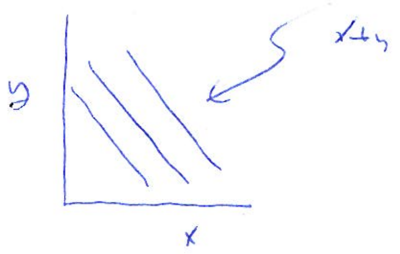
Z

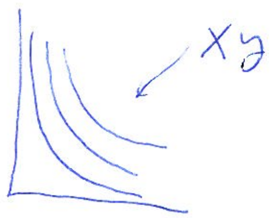
11/2/06

← Nov.

① Computing = throwing away informa'n

$z = f(x,y)$: lots of $x+y$ map to the same z





computing $f(x,y)$ is equivalent to finding which manifold.

← mention the problem of vision.

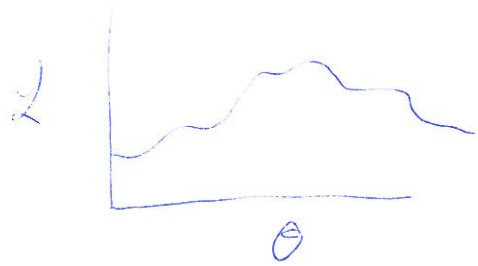
Simple ~~case~~ problem

$$p(r|\theta) = \prod_i \frac{f(\theta-\theta_i)^{r_i} e^{-f(\theta-\theta_i)}}{r_i!}$$

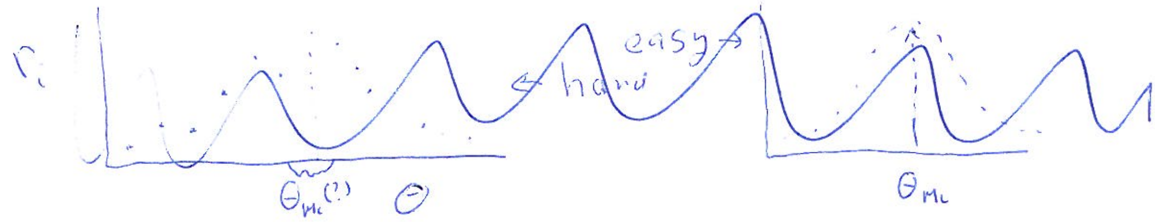
$$\mathcal{L} = \sum_i r_i \log f(\theta-\theta_i) - \sum_i f(\theta-\theta_i) - \log r_i!$$

↓ computer

ML. $\frac{d\mathcal{L}}{d\theta} = 0$



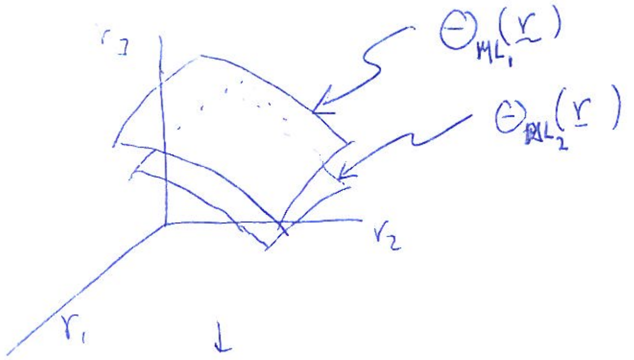
hard



(2)

$$\partial_\Theta \mathcal{L}(\underset{\sim}{r}, \Theta)\Big|_{\Theta_{ML}} = 0$$

$$= \Theta_{ML} = \Theta_{ML}(\underset{\sim}{r})$$

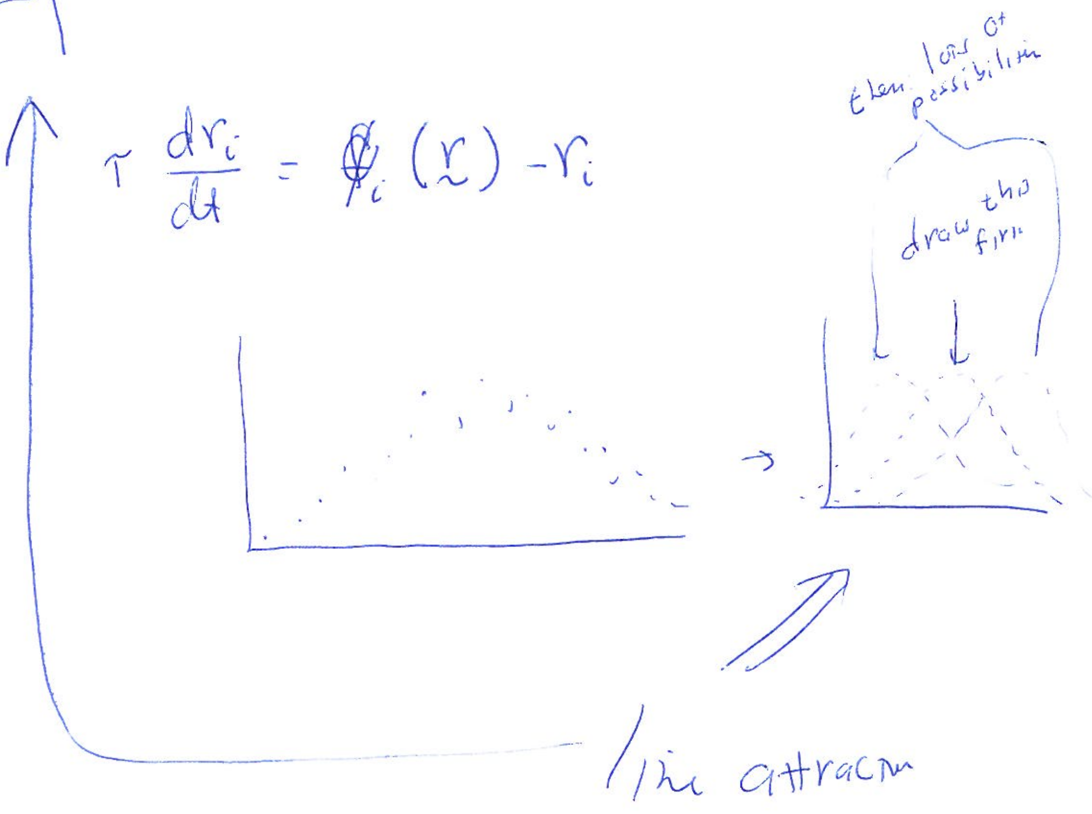


Finding ML estimate = finding which manifold you're on

– mention more complicated manifolds, like object recognition

network dynamics: takes you to a stereotypical point; can use template matching readout

$$\tau \frac{dr_i}{dt} = \phi_i(\underset{\sim}{r}) - r_i$$



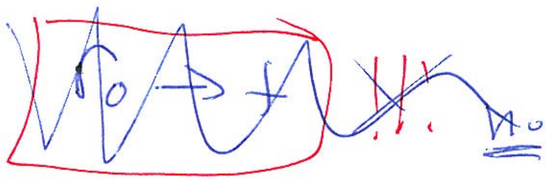
③

How to build a line attractor

$$\tau \dot{r}_i = \phi\left(\sum_j W_{i-j} r_j\right) - r_i$$

Let r_{0i} be a ~~solution~~ fixed point

$$r_{0i} = \phi\left(\sum_j W_{i-j} r_{0j}\right)$$



~~$r_{0(i+k)}$~~ also $\tilde{r}_{0_i} = r_{0_{i+k}}$ is also a solution:

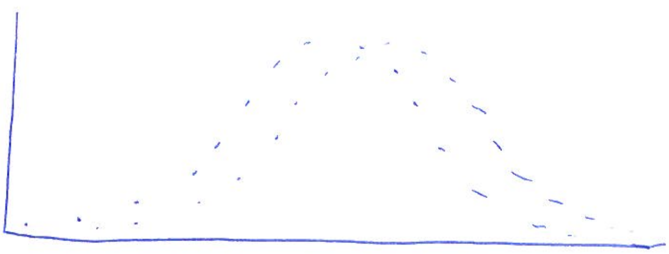
$$\tilde{r}_{0_i} = r_{0_{i+k}} = \cancel{\phi\left(\sum_j W_{i-j} r_{0_{j+k}}\right) = \phi\left(\sum_j W_{i-j} \tilde{r}_{0_j}\right)}$$

$$\cancel{= \phi\left(\sum_j W_{i+k} -\right.}$$

$$= \phi\left(\sum_j W_{i+k-j} r_{0_j}\right)$$

$$= \phi\left(\sum_j W_{i+k-(j+k)} r_{0_{j+k}}\right)$$

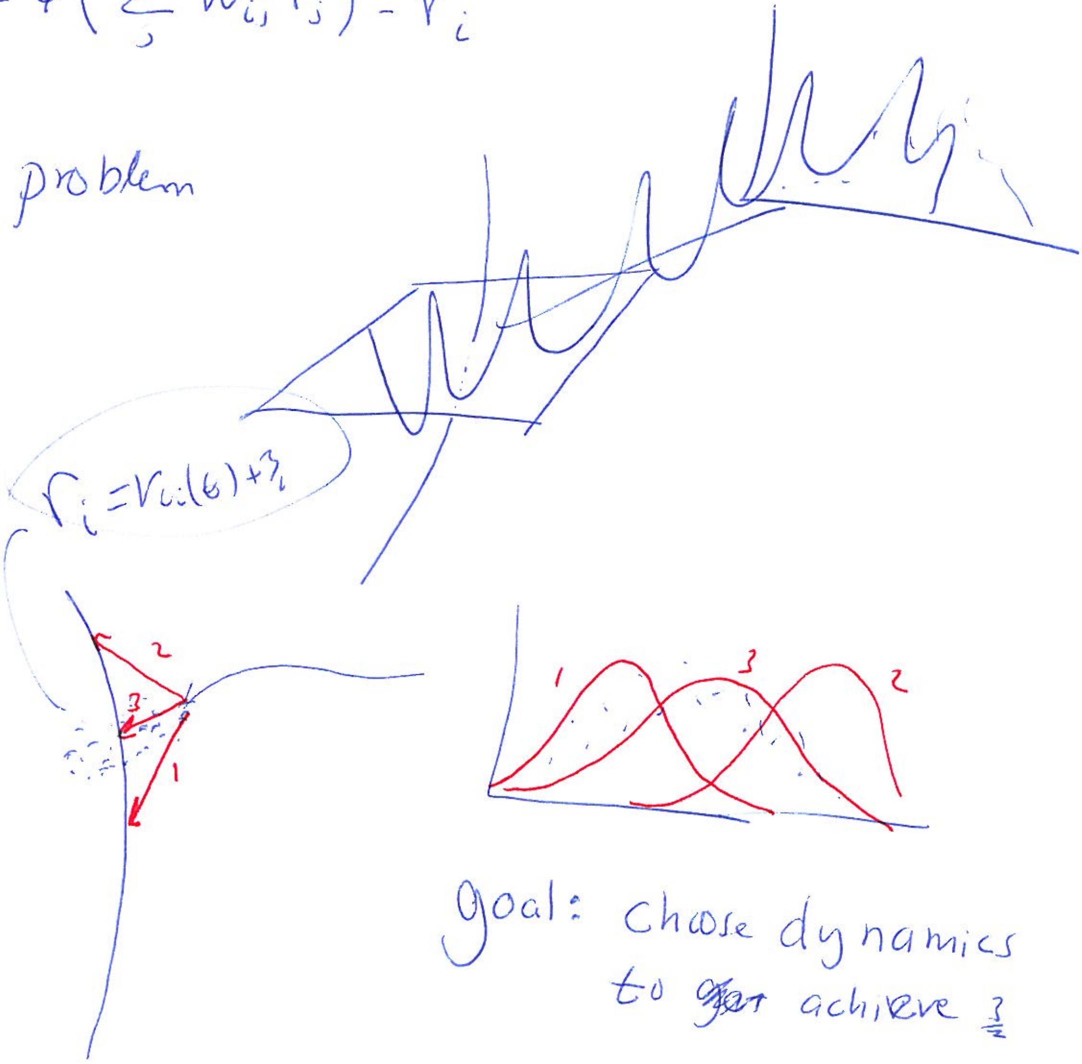
$$= \phi\left(\sum_j W_{i-j} \tilde{r}_{0_j}\right)$$



— Side point about drift?

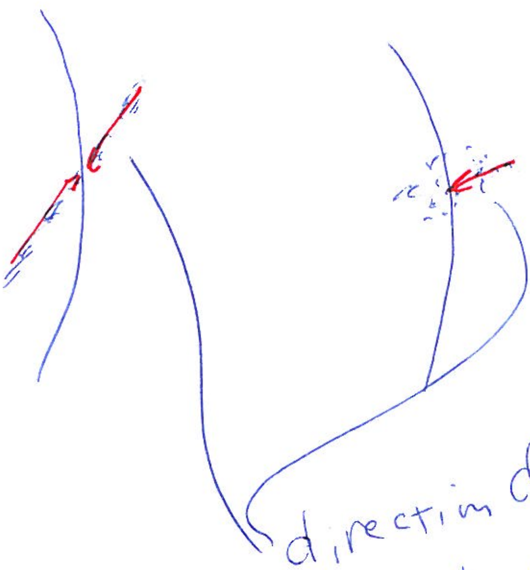
$$\tau \dot{r}_i = \phi\left(\sum_j W_{ij} r_j\right) - r_i$$

The problem



goal: choose dynamics to ~~get~~ achieve ξ

noise is an issue!!



direction depends on Φ + W.
Fix Φ, choose W so it has right dynamics.

5

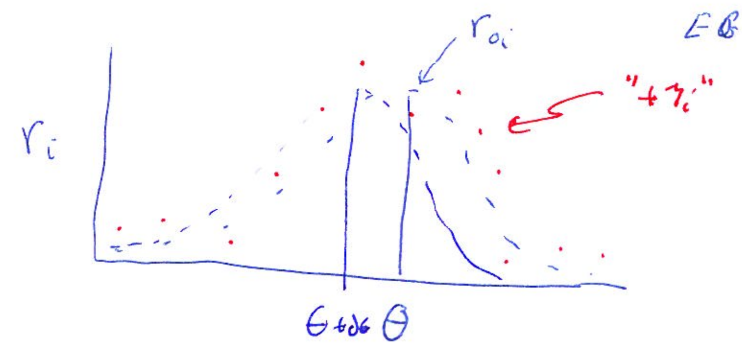
Formalizing the problem

$$\tau \dot{r}_i = \phi\left(\sum_j W_{i-j} r_j\right) - r_i$$

$$r_i(t=0) = r_{0i}(\theta) + \eta_i$$

$\langle \eta_i \rangle = 0 \quad \langle \eta_i \eta_j \rangle = R_{ij}$

E.G. $P(\eta) = \dfrac{e^{-\frac{1}{2}\eta R^{-1}\eta}}{|2\pi R|^{1/2}}$



$$r_i(t=\infty) = r_{0i}(\theta + d\theta)$$

$$d\theta = d\theta(\eta)$$

want: $\langle d\theta \rangle_\eta = 0$

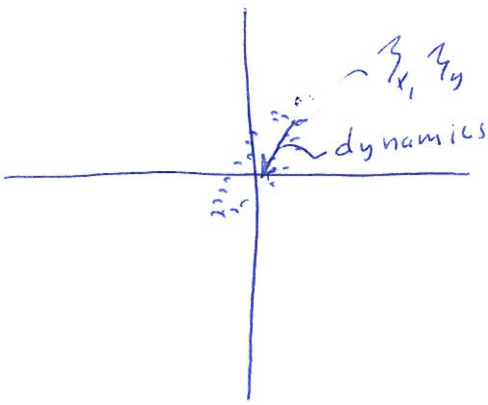
$\langle d\theta^2 \rangle$ = as small as possible.

found by solving ODE

$\boxed{\text{Math problem}}$ remember, this is our goal!

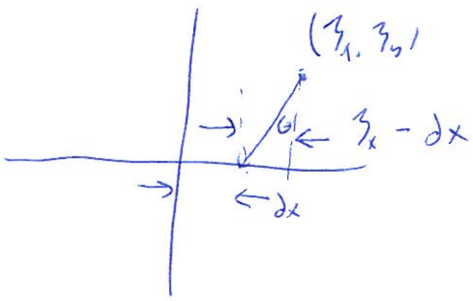
⑥

An easier problem



$p(\eta) \sim \mathcal{N}(0, R)$

$$R = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}$$



$$\tan\theta = \frac{\eta_x - \delta x}{\eta_y} \Rightarrow \delta x = \eta_x - \eta_y \tan\theta$$

$\langle \delta x \rangle = 0$ (good)

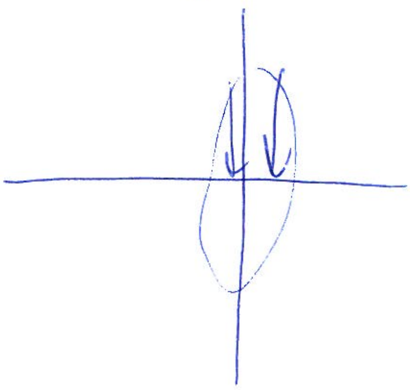
$$\langle \delta x^2 \rangle = \langle \eta_x^2 \rangle - 2\langle \eta_x \eta_y \rangle \tan\theta + \langle \eta_y^2 \rangle \tan^2\theta$$

$$= \sigma_x^2 - 2\rho\sigma_x\sigma_y \tan\theta + \sigma_y^2 \tan^2\theta$$

$$\partial_\theta (\ ) = \partial_\theta \tan\theta \left[-2\sigma_x\sigma_y\rho + 2\sigma_y^2 \tan\theta\right] \Rightarrow \boxed{\tan\theta = \frac{\rho\sigma_x}{\sigma_y}}$$

$$\langle \delta x^2 \rangle = \sigma_x^2 (1 - \rho^2)$$

$\rho = 0$



$T\,\partial\dot{\underline{r}} = \underline{\underline{A}}\cdot\partial\underline{r}$

$\partial\underline{r} = \sum_k a_k \underline{v}_k$

$T\dot{a}_k = -\lambda_k a_k$ $\lambda_0 = 0$

$\partial r(0) = a_0\,\underline{v}_0$ $a_0 = \underline{v}_0^\dagger\cdot\partial\underline{r}(1)$

$= \underline{v}_0^\dagger\cdot[r_{0i}(\theta+\partial\theta) + 7$

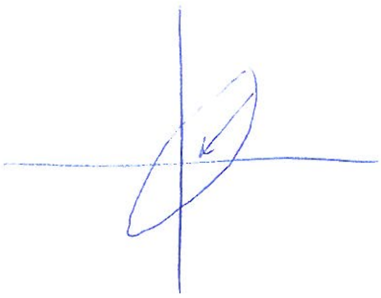
$= v_0^\dagger(r_0(t)) + v_0^\dagger\eta_i + \eta'$

$$\partial\theta = \frac{v_0^\dagger\cdot\eta - \cancel{v_0^\dagger\cdot r_0}}{v_0^\dagger\cdot r_0'}$$

$$\langle\partial\theta^2\rangle = \frac{v_0^\dagger\cdot\langle\eta\eta\rangle\cdot v_0^\dagger}{(v_0^\dagger r_0')^2}$$

min, end up w/ CR bound

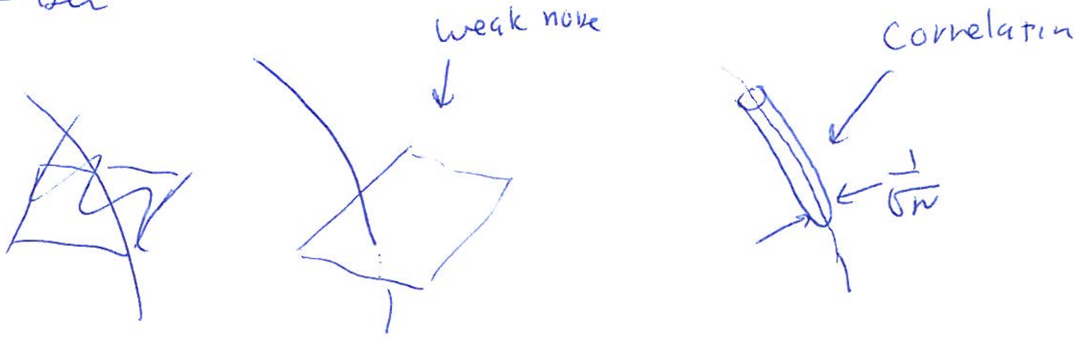
$\rho > 0$



if there's plenty of time, work ~~th~~ real problem

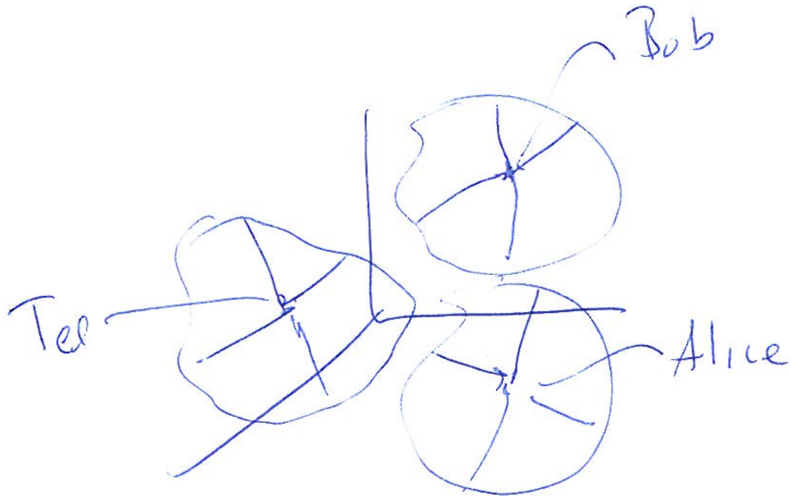
Noise, & especially correlation in noise, has a huge effect on optimal dynamics!!

A teaser



7'

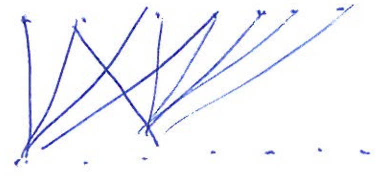
mentim Anni



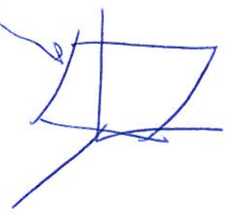
A second way of computing: feed forward networks

$\cancel{x_i = \phi(\sum_j W_{ij} x_j)}$

$$r_i^{(2)} = \phi\left(\sum_j W_{ij} r_j^{(1)} \right)$$



linear manifold

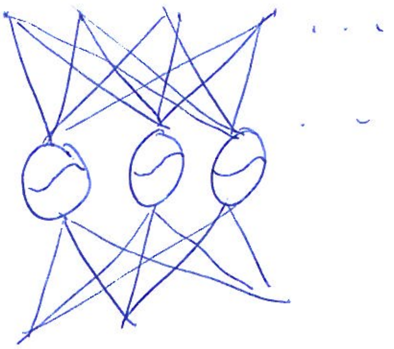


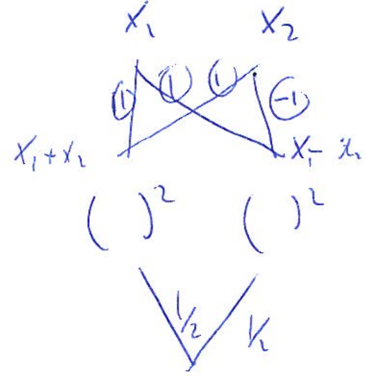
E.G.

$w_1 r_1 + w_2 r_2 = \text{const}$ $\quad w_1 = w_2$, $\quad r_1 + r_2 = \text{const}$

$\quad w_1 = -w_2 \quad r_1 - r_2 = \text{const}$

multiple layer.

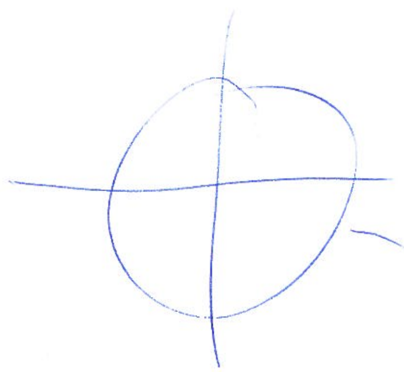




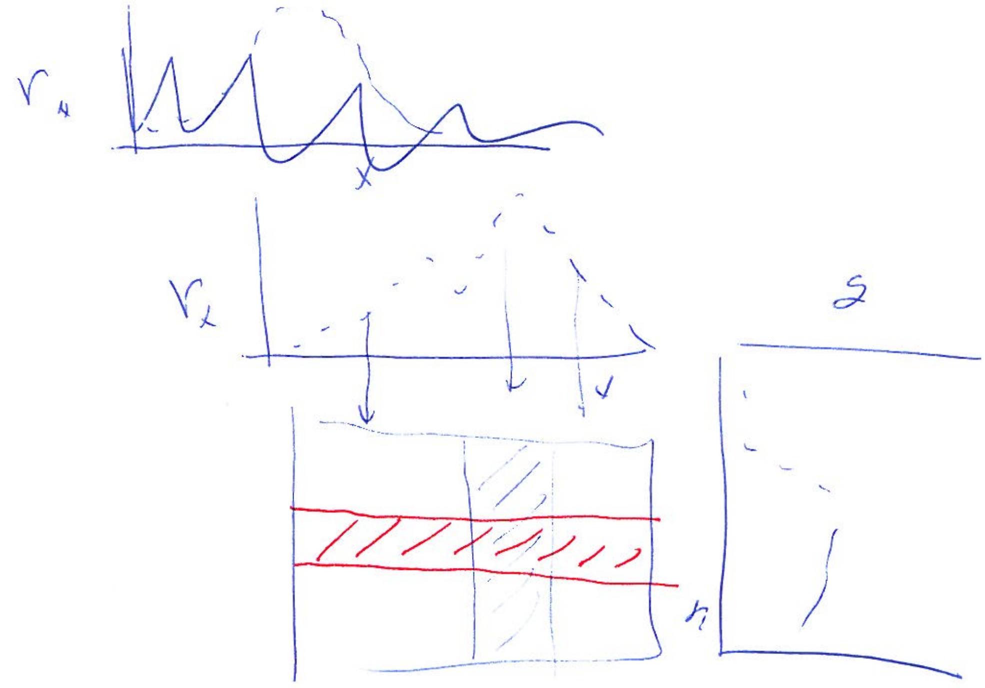
$$\frac{(x_1+x_2)^2 + (x_1 - x_2)^2}{2}$$

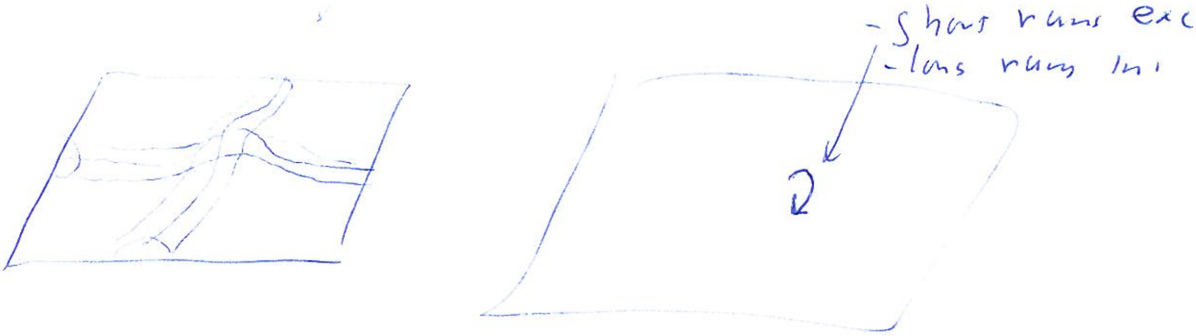
$$= \frac{x_1^2 + 2x_1x_2 + x_2^2 + x_1^2 - 2x_1x_2 + x_2^2}{2}$$

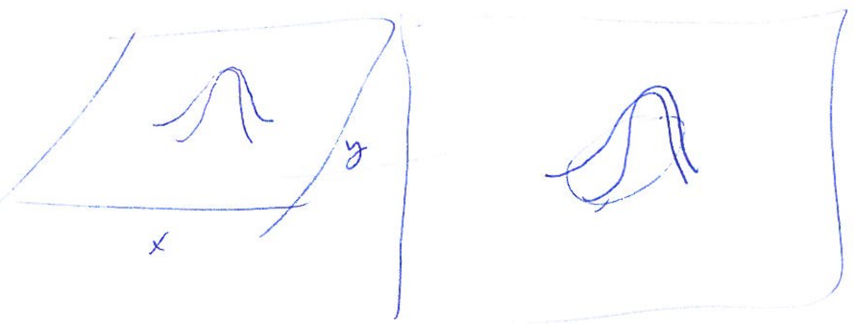
$$= x_1^2 + x_2^2$$

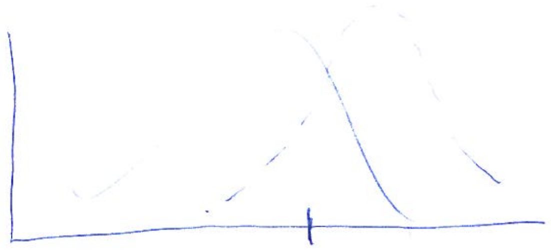


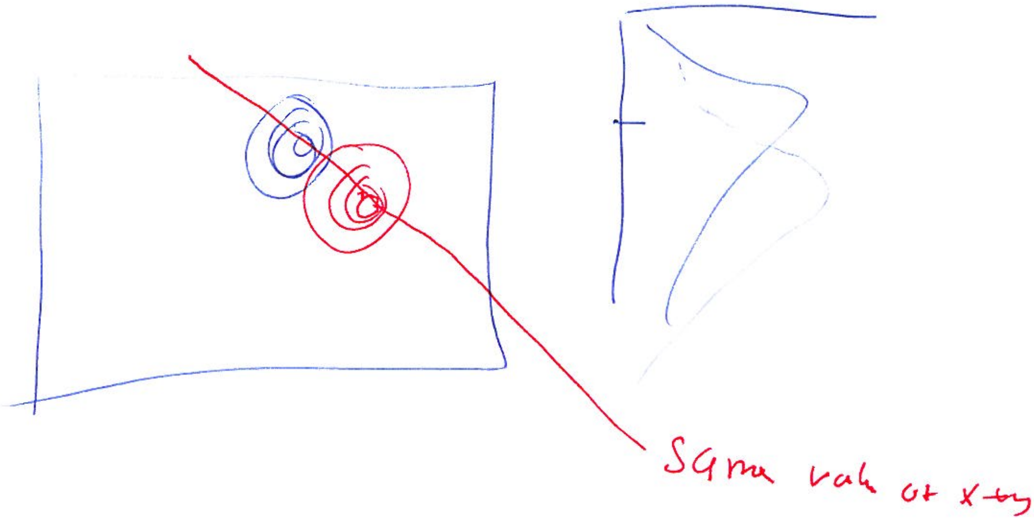
Combine ff & recurrent to compute function

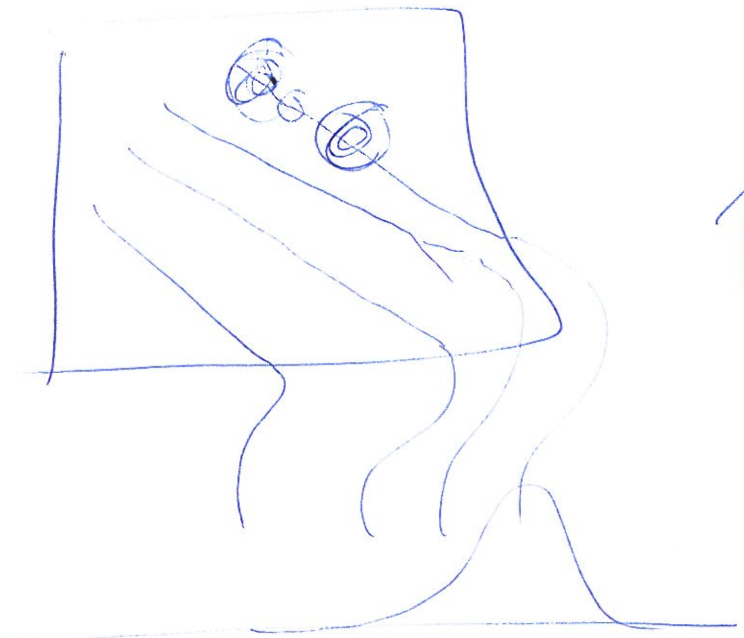


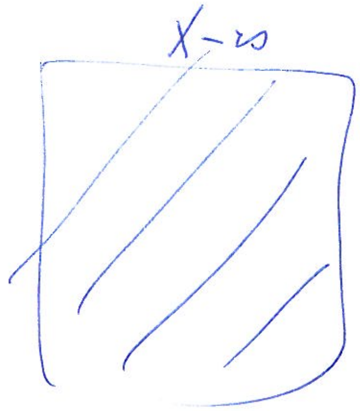


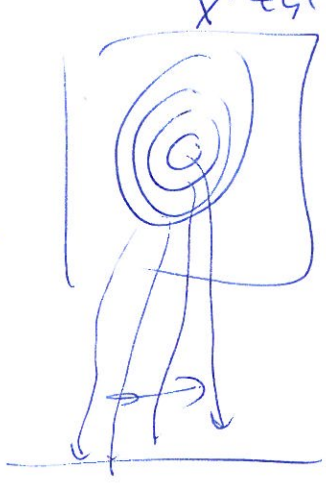





Same vals of x+y




X-ys


x^2+y^2

$$r_i^x = \phi\Big(\sum_j W^x_{i-j} r^x_j + W^{xI}_{ijk} r_{jk}\Big)$$

$$r_i^y = \phi\Big(\sum_j W^y_{i-j} r^y_j + W^{yI}_{ijk} r_{jk}\Big)$$

$$r_i^z = \phi\Big(\qquad \qquad \Big)$$

$$r_{ij} = \phi\Big(W^I_{ijkl} r_{kl} + W^{Ix}_{ijk} r^x_k + W^{Iy}_{ijk} r^y_k + W^{Iz}_{ijk} r^z_k\Big)$$

→ This is the state of the art (almost) in computing functions

if there's time

1. mention that this is 2-D ANN
2. throws away information about prob. dist.
3. PPC