4. Exploding Gradient.

$$h_t = W^T h_{t-1}$$
$$h_{t-1} = W^T h_{t-2}$$
$$\vdots$$
$$h_1 = W^T h_0$$

Then $h_t = (W^T)^t h_0$

W could be decomposed as
$$W = PDP^{-1}$$
where D is diagonal matrix of eigenvalues.
P is the eigen vector of W

Then we can wright

$$h_t = ((PDP^{-1})^T)^t h_0$$
$$= ((P^{-1})^T D^T P^T)^t h_0$$
$$= ((P^{-1})^T D P^T)^t h_0$$
$$= (P^{-1})^T D P^T (P^{-1})^T D P^T \cdots (P^{-1})^T D P^T h_0$$
as $(P^{-1})^T \cdot P^T = I$
$$= (P^{-1})^T D^t P^T h_0$$

$$\frac{dh_t}{dh_0} = (P^{-1})^T D^t P^T$$

when $t \gg 0$, if a eigenvalues in $D < 1$, the value in $D^t \to 0$
thus vanishing the gradient
if a eigenvalue in $D > 1$, the value in $D^t \to \infty$, thus
exploding the gradient