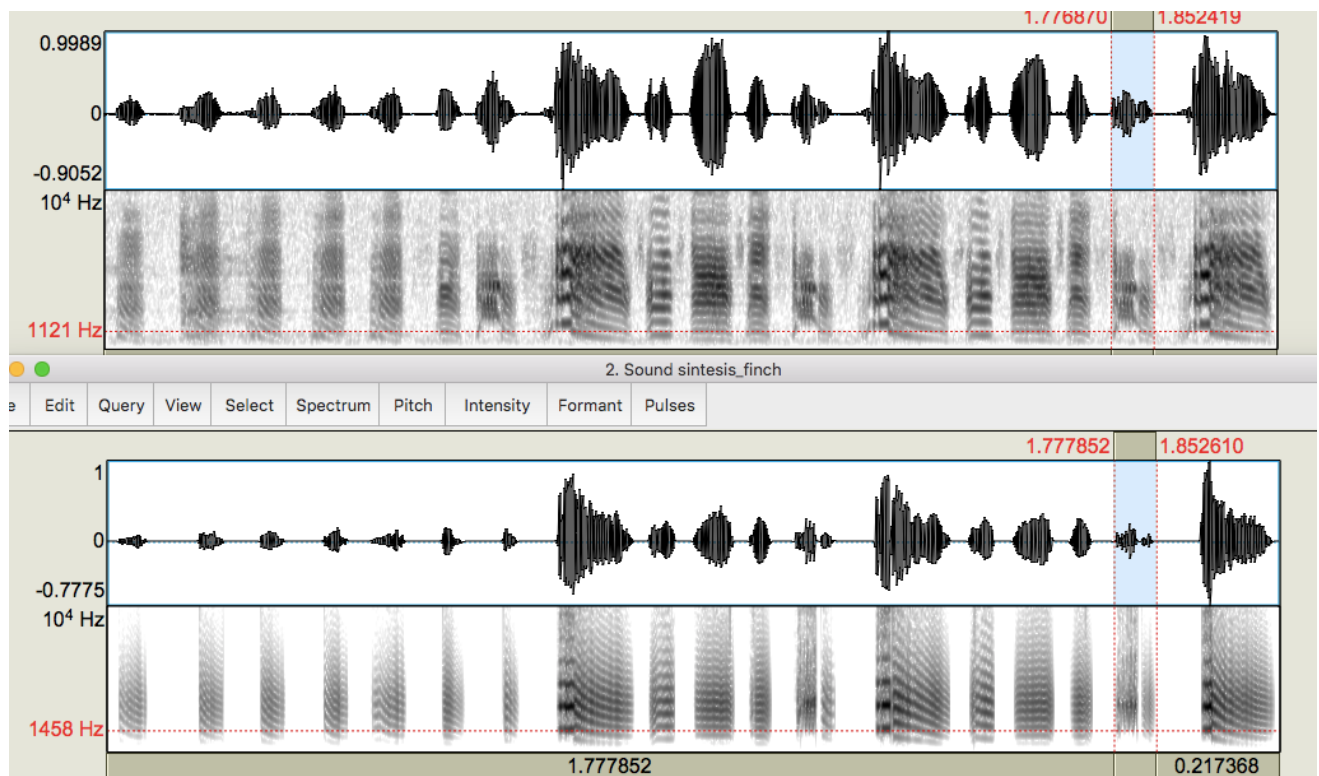


## Clase VII

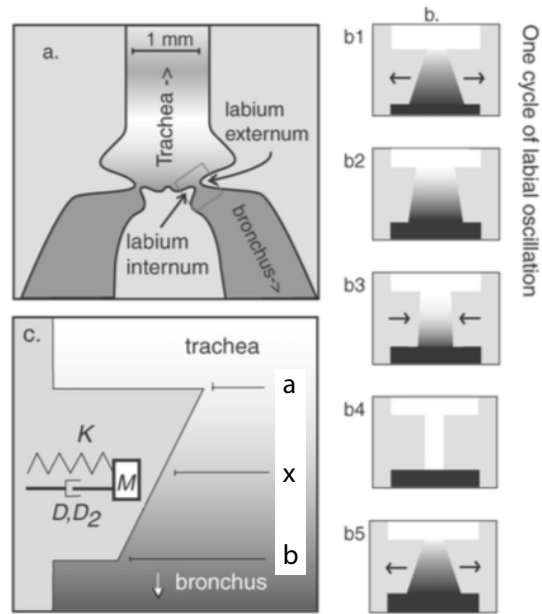
### Descomposición en modos empíricos.

El nacimiento de los sistemas dinámicos fue el estudio de la evolución de la posición de una partícula, por lo que la baja dimensionalidad esta asociada a la noción de partícula puntual. Sin embargo, dinámicas sencillas se encuentran en problemas de significativa complejidad.

Lo que estamos viendo es el sonido y sonograma, para el canto de un ave (diamante mandarín), y la representación generada por un sistema dinámico sencillo.



## Propuesta para labios siringeos (2001)



Cinemática (asumimos estos dos modos del dibujo de la derecha)

$$b = b_0 + x - \tau \frac{dx}{dt}.$$

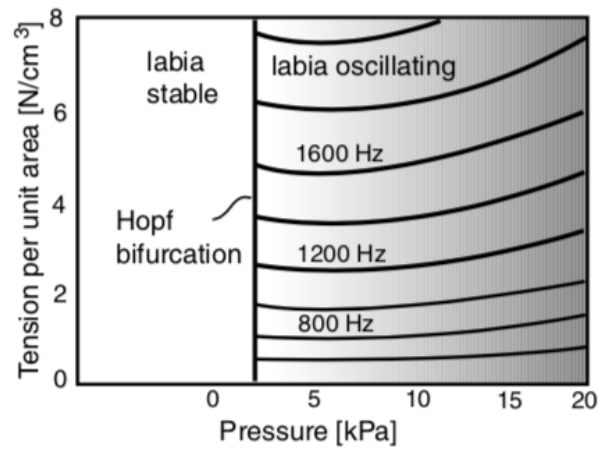
Presión calculada asumiendo esa cinemática

$$P_f = P_b(1 - a/b).$$

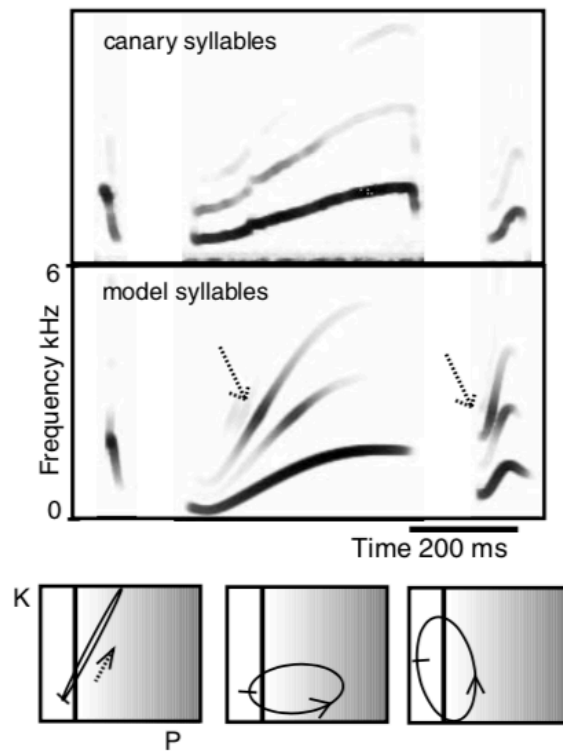
Sistema dinámico para la posición media asumiendo esa presión

$$M\ddot{x} + D\dot{x} + D_2(\dot{x})^3 + Kx = P_b \frac{a_0 - b_0 + 2\tau\dot{x}}{x + b_0 + \tau\dot{x}}.$$

## Diagrama de bifurcaciones

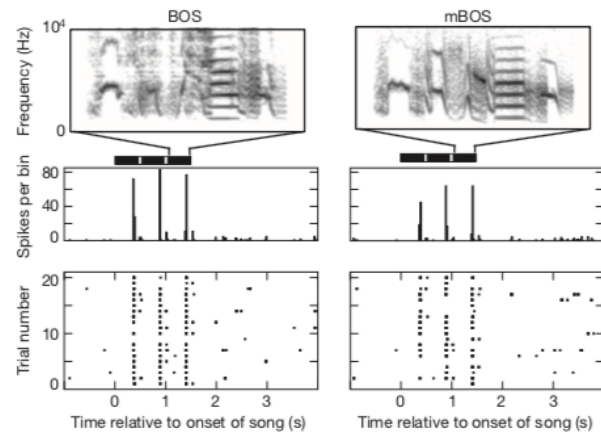


El origen de las silabas según este escenario

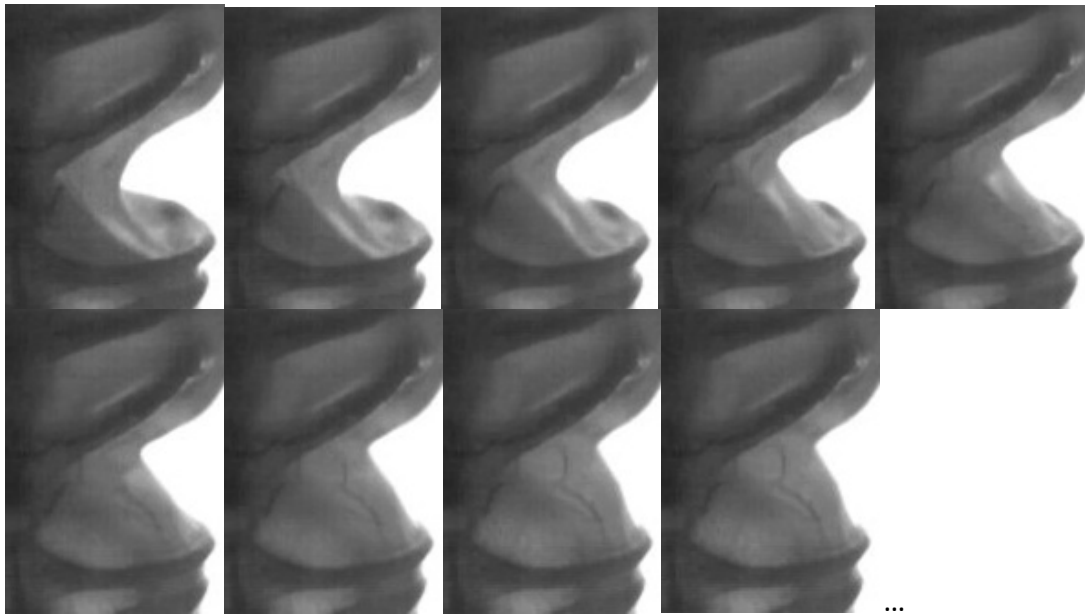


**¿Cuan bueno es este ajuste, basado en estas simplificaciones?**

No solo es parecido: un cerebro de ave responde ante este canto sintético como ante el propio canto.



¿Pero esto... es lo que pasa?



Como extraer *la información que deseábamos (\*)* a partir de esta película?

**(\*) una jerarquía de modos, y su evolución temporal**

## Singular value decomposition (SVD).

Supongamos que tenemos un conjunto de observaciones  $\{x_1, x_2, \dots, x_m\}$ , cada una de las cuales  $x_k \in \mathbb{C}^n$ . Armemos el arreglo

$$X = [x_1 \ x_2 \ \dots \ x_m] \in \mathbb{C}^{n \times m}.$$

Por ejemplo, los  $x_k$  pueden ser los valores de un campo escalar en un dominio unidimensional evaluado en  $n$  puntos espaciales, o los valores de una imagen en la cual los  $i \times j = n$  valores en cada píxel se acomodan en un arreglo unidimensional de  $n \times 1$ .

La descomposición en valores singulares es un algoritmo que apunta a descomponer el arreglo original  $X$  en uno en el cual **pesen mas los elementos estadísticamente mas descriptivos** del problema.

La descomposición SVD es una descomposición que existe *siempre*, y consiste en escribir la siguiente identidad:

$$X = U \Sigma V^*,$$

donde  $U \in \mathbb{C}^{n \times n}$ ,  $V \in \mathbb{C}^{m \times m}$  son matrices unitarias (i.e.  $UU^* = U^*U = I$ ), con columnas ortogonales, y  $\Sigma \in \mathbb{C}^{n \times m}$  tiene entradas reales no negativas en la diagonal, y ceros en todas las otras entradas. La matriz  $V^* = \text{conjugada de } V^T$ .

Cuando  $n \geq m$ ,  $\Sigma$  tiene a lo sumo  $m$  elementos no nulos, y puede ser escrita como:

$$\Sigma = \begin{pmatrix} \hat{\Sigma} \\ \mathbf{0} \end{pmatrix},$$

y por lo tanto,  $X$  en forma exacta se puede escribir así:

$$X = U \Sigma V^* = (\hat{U} \ \hat{U}^\perp) \begin{pmatrix} \hat{\Sigma} \\ \mathbf{0} \end{pmatrix} V^*$$

$$\begin{array}{c}
 \text{Full SVD} \\
 \left[ \begin{array}{c} \mathbf{X} \end{array} \right] = \underbrace{\left[ \begin{array}{c|c} \hat{\mathbf{U}} & \hat{\mathbf{U}}^\perp \end{array} \right]}_{\mathbf{U}} \underbrace{\left[ \begin{array}{c} \hat{\Sigma} \\ \hline \mathbf{0} \end{array} \right]}_{\Sigma} \left[ \begin{array}{c} \mathbf{V}^* \end{array} \right] \\
 \\
 \text{Economy SVD} \\
 = \left[ \begin{array}{c} \hat{\mathbf{U}} \end{array} \right] \left[ \begin{array}{c} \hat{\Sigma} \end{array} \right] \left[ \begin{array}{c} \mathbf{V}^* \end{array} \right]
 \end{array}$$

Las columnas de  $\mathbf{U}$  se conocen como los vectores singulares por izquierda de  $\mathbf{X}$ , mientras que los vectores columna de  $\mathbf{V}$  son los vectores singulares por derecha.

**¿Por que es interesante esta descomposición?**

**Teorema** (Eckart Young): Una **aproximación optima** de rango  $r$  de  $\mathbf{X}$ , en el sentido de cuadrados mínimos, esta dada por el truncado de rango  $r$  de la descomposición SVD  $\hat{\mathbf{X}}$  de  $\mathbf{X}$ :

o sea, si  $\tilde{\mathbf{U}}, \tilde{\mathbf{V}}$  son las primeras  $r$  columnas de  $\mathbf{U}, \mathbf{V}$ , y  $\tilde{\Sigma}$  es de  $r \times r$  (ordenados, los valores mas grandes), entonces

$$\tilde{\mathbf{X}} = \tilde{\mathbf{U}} \tilde{\Sigma} \tilde{\mathbf{V}}^*$$

Es la mejor aproximación con ese numero de términos.

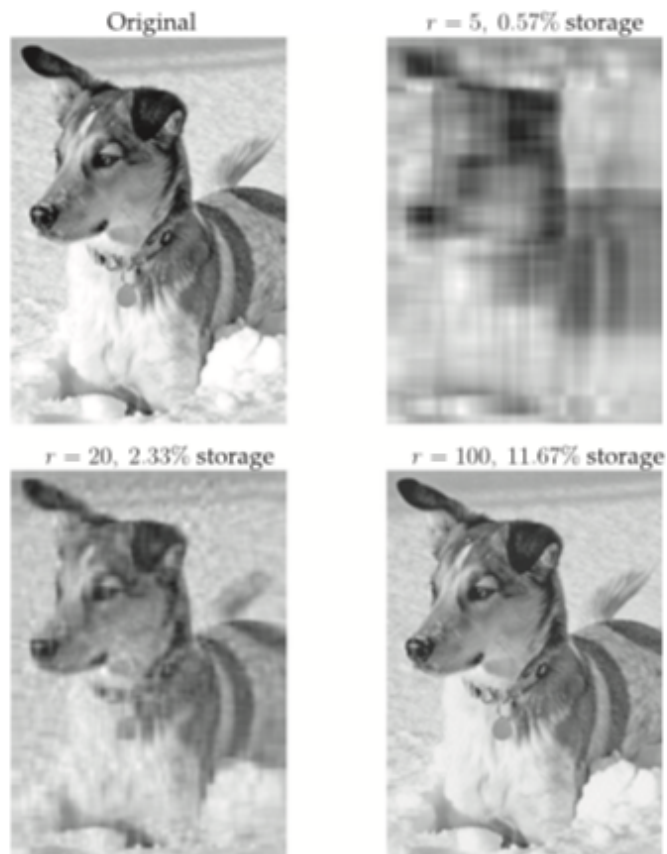
Notemos que como  $\Sigma$  es diagonal, una aproximación de  $r$  términos se podrá escribir:

$$\tilde{\mathbf{X}} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^* + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^* + \cdots + \sigma_r \mathbf{u}_r \mathbf{v}_r^*$$

## Ejemplo: compresión de imágenes.

$$\begin{array}{c}
 \text{Full SVD} \\
 \left[ \begin{array}{c} \mathbf{X} \end{array} \right] = \left[ \begin{array}{c|c|c} \tilde{\mathbf{U}} & \hat{\mathbf{U}}_{\text{rem}} & \hat{\mathbf{U}}^{\perp} \end{array} \right] \left[ \begin{array}{c|c} \tilde{\Sigma} & \hat{\Sigma}_{\text{rem}} \\ \hline \mathbf{0} & \end{array} \right] \left[ \begin{array}{c} \tilde{\mathbf{V}}^* \\ \mathbf{V}_{\text{rem}} \end{array} \right] \\
 \underbrace{\hspace{1.5cm}}_{\mathbf{U}} \\
 \text{Truncated SVD} \\
 \approx \left[ \begin{array}{c} \tilde{\mathbf{U}} \end{array} \right] \left[ \begin{array}{c} \tilde{\Sigma} \end{array} \right] \left[ \begin{array}{c} \tilde{\mathbf{V}}^* \end{array} \right]
 \end{array}$$

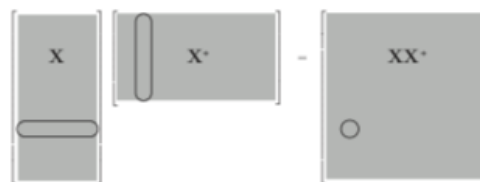
Donde rem denota “remainder”. Entonces, podemos hacer una descomposición, asumiendo que nuestra imagen es un arreglo de  $n \times m$  números (los valores en cada pixel en una escala de grises,  $n$  parametrizando posición vertical y  $m$ , horizontal). Entonces, descomponemos, y reconstruimos la imagen truncando al tomar distinto numero de elementos no nulos para  $\tilde{\Sigma}$ .



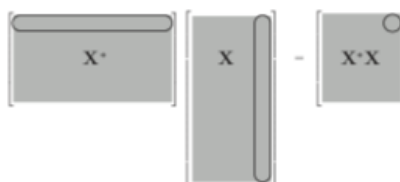
**Figure 1.3** Image compression of Mordecai the snow dog, truncating the SVD at various ranks  $r$ . Original image resolution is  $2000 \times 1500$ .

## Una interpretación en términos de correlaciones

Sean  $XX^*$ ,  $X^*X$ :



**Figure 1.6** Correlation matrix  $XX^*$  is formed by taking the inner product of rows of  $X$ .





Notamos que podemos escribir estas matrices de correlación de la siguiente manera:

$$XX^* = U \begin{pmatrix} \hat{\Sigma} \\ 0 \end{pmatrix} V^* \quad V(\hat{\Sigma} \ 0)U^* = U \begin{bmatrix} \hat{\Sigma}^2 & 0 \\ 0 & 0 \end{bmatrix} U^*,$$

$$X^*X = V(\hat{\Sigma} \ 0)U^*U \begin{pmatrix} \hat{\Sigma} \\ 0 \end{pmatrix} V^* = V\hat{\Sigma}^2V^*.$$

Y como  $V, U$  son unitarias,

$$XX^* U = U \begin{bmatrix} \hat{\Sigma}^2 & 0 \\ 0 & 0 \end{bmatrix},$$

$$X^*X V = V\hat{\Sigma}^2$$

Así, las columnas de  $U$  son auto vectores de la matriz de correlación  $XX^*$ , y  $V$  esta constituido por columnas que son auto vectores de  $X^*X$ .

**Una cosa re interesante (Sirovich):** Es impráctico muchas veces, proponerse realizar un análisis de  $XX^*$  (*grande!!*), por la alta dimensión de  $n$ . Pero podemos intentar calcular la descomposición de  $X^*X$  (*chica!!!*).

De ahí calculamos  $V$ , y  $\hat{\Sigma}^2$  (de la ecuación de auto valores y auto vectores). Tomamos  $r$  valores no nulos,  $\tilde{\Sigma}$ , y de ahí podemos calcular  $\tilde{U}$ :

Como

$$\tilde{X} = \tilde{U}\tilde{\Sigma}\tilde{V}^*,$$

$$\tilde{U} = \tilde{X}\tilde{V}\tilde{\Sigma}^{-1}.$$

## La descomposición de una ecuación a derivadas parciales en modos empíricos.

La proper orthogonal decomposition, o descomposición en modos empíricos, es la aplicación de SVD al problema de estudiar ecuaciones diferenciales a derivadas parciales.

La observación central es que muchos comportamientos se pueden describir en términos de pocos patrones. Veamos que caminos podemos tomar para representar la dinámica de **sistemas extendidos**.

Supongamos una ecuación diferencial a derivadas parciales:

$$\mathbf{u}_t = N(\mathbf{u}, \mathbf{u}_x, \mathbf{u}_{xx}, \dots, x, t)$$

Con sus condiciones de contorno, supongamos en un dominio acotado  $[-L, L]$ .

Si llevamos a cabo una discretización espacial (como la que uno haría si fuera a resolver numéricamente el problema), tenemos que nuestra variable continua pasa a ser un conjunto de variables discretas:

$$\mathbf{u}(x_k, t) \quad \text{para } k = 1, 2, \dots, n$$

Donde usando que

$$\begin{aligned}\mathbf{u}_x(x_k, t) &= \frac{\mathbf{u}(x_{k+1}, t) - \mathbf{u}(x_{k-1}, t)}{2\Delta x}, \\ \mathbf{u}_{xx}(x_k, t) &= \frac{\mathbf{u}(x_{k+1}, t) + \mathbf{u}(x_{k-1}, t) - 2\mathbf{u}(x_k, t)}{2\Delta x},\end{aligned}$$

Da lugar a un conjunto (¡en principio elevado!). de ecuaciones diferenciales ordinarias:

$$\frac{d\mathbf{u}(x_k, t)}{dt} = N(\mathbf{u}(x_{k+1}, t), \mathbf{u}(x_k, t), \mathbf{u}(x_{k-1}, t), \dots, x_k, t)$$

Ahora, inspirados en la idea de modos lineales, uno puede pensar en otro modo computacional para atacar el problema de las ecuaciones diferenciales a derivadas parciales: el método de separación de variables, que asume que uno puede escribir el campo superponiendo movimientos colectivos:

$$u(x, t) = \sum_{k=1}^n a_k(t) \psi_k(x)$$

Si estas funciones se toman ortogonales, luego se puede proyectar la dinámica de la ecuación diferencial a derivadas parciales a un conjunto de ecuaciones diferenciales ordinarias para los modos, como hacemos habitualmente. Pensemos el caso de un campo escalar  $u(x, t) = \sum_{k=1}^n a_k(t) \psi_k(x)$ :

$$\sum \psi_k(x) \frac{da_k}{dt} = N \left( \sum_{k=1}^n a_k(t) \psi_k(x), \sum_{k=1}^n a_k(t) (\psi_k(x))_x, \sum_{k=1}^n a_k(t) \psi_k(x)_{xx}, \dots, x, t \right)$$

Y usando la ortogonalidad:

$$\langle \psi_k, \psi_j \rangle = \delta_{ij}$$

$$\frac{da_k}{dt} = \langle N \left( \sum a_j \psi_j, \dots \right), \psi_k \rangle$$

**Ahora bien: no hay a priori un criterio único para elegir la base.** Puede ser que uno desee apoyarse en la base del problema linealizado, si va a estar cerca del umbral, o por algún motivo que lo motive. Pero si uno busca disminuir la dimensionalidad del problema, puede intentar esto:

**Medir, tomar “snapshots”, y ¡emplearlos para descubrir los modos óptimos!**

**(acordemosnos de esto:**

*“Una cosa re interesante (Sirovich): Es impráctico muchas veces, proponerse realizar un análisis de  $X^*X$ , por la alta dimensión de  $n$ . Per podemos intentar calcular la descomposición de  $X^*X$ .*

*De ahí calculamos  $V$ , y  $\Sigma^2$  (de la ecuación de auto valores y auto vectores). Tomamos  $r$  valores no nulos,  $\Sigma$ , y de ahí podemos calcular  $\tilde{U}$ :*

Como

$$\tilde{X} = \tilde{U} \tilde{\Sigma} \tilde{V}^*,$$

$$\tilde{U} = \tilde{X} \tilde{V} \tilde{\Sigma}^{-1}. “)$$

Así, tomamos mediciones como:

$$\mathbf{u}_k = [\mathbf{u}(x_1, t_k), \mathbf{u}(x_2, t_k), \dots, \mathbf{u}(x_n, t_k)]^T$$

y armamos:

$$\mathbf{X} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_m \\ | & | & \cdots & | \end{bmatrix}$$

Con cada columna, un conjunto de  $n$  mediciones escalares que constituyen la  $k$ -ésima medición  $\mathbf{u}_k$ . Las  $m$  mediciones pueden ser las mediciones durante una serie temporal de mediciones. Si la resolución espacial es mayor que el número de mediciones, tenemos una matriz alta y flaca. Y ahora estamos listos para hacer entrar a nuestra idea de modos empíricos. Escribimos:

$$\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$$

Con  $\mathbf{U} \in \mathbb{C}^{n \times n}$ ,  $\mathbf{V}^* \in \mathbb{C}^{m \times m}$ ,  $\mathbf{\Sigma} \in \mathbb{C}^{n \times m}$ , donde las primeras dos son unitarias, y  $\mathbf{\Sigma}$  tiene entradas no negativas en la diagonal. Así, las columnas a  $\mathbf{U}$  dan los modos,  $\mathbf{V}$  da la historia temporal, y  $\mathbf{\Sigma}$  los pesos relativos de los modos.

Típicamente nos interesará un truncado con un número  $r$  chico de modos. Esto implica aproximar a  $\mathbf{X}$  por el truncado con los  $r$  primeros modos:

$$\tilde{\mathbf{X}} = \tilde{\mathbf{U}}\tilde{\mathbf{\Sigma}}\tilde{\mathbf{V}}^*,$$

Y entonces tenemos un conjunto de  $r$  modos  $\psi_k$ , que serán las primeras  $r$  columnas de la matriz  $\mathbf{U}$ :

$$\tilde{\mathbf{U}} = \mathbf{\Psi} = \begin{bmatrix} | & | & \cdots & | \\ \psi_1 & \psi_2 & \cdots & \psi_r \\ | & | & \cdots & | \end{bmatrix}$$

Y ahora estamos en condiciones de realizar nuestra proyección:

$$\mathbf{u}(t) \approx \mathbf{\psi}\mathbf{a}(t),$$

Donde  $\mathbf{a} \in R^r$  es el coeficiente dependiente del tiempo, de  $r$  componentes. Reemplazando y empleando la ortogonalidad como antes, nos quedamos con:

$$\frac{d\mathbf{a}}{dt} = \boldsymbol{\psi}^T \mathbf{L} \boldsymbol{\psi} \mathbf{a}(t) + \boldsymbol{\psi}^T \mathbf{N}(\boldsymbol{\psi} \mathbf{a}(t), \boldsymbol{\beta})$$

### Un ejemplo numérico: el oscilador armónico cuántico.

Sea el sistema extendido:

$$iu_t + \frac{1}{2}u_{xx} - \frac{x^2}{2}u = 0$$

$$u(x, t) = \sum_{i=0}^{\infty} a_k (2^k k \sqrt{\pi})^{-\frac{1}{2}} e^{-\frac{x^2}{2}} H_k(x) e^{-i(k+1/2)t} = \sum_{i=0}^{\infty} a_k \psi_k e^{-i(k+1/2)t}$$

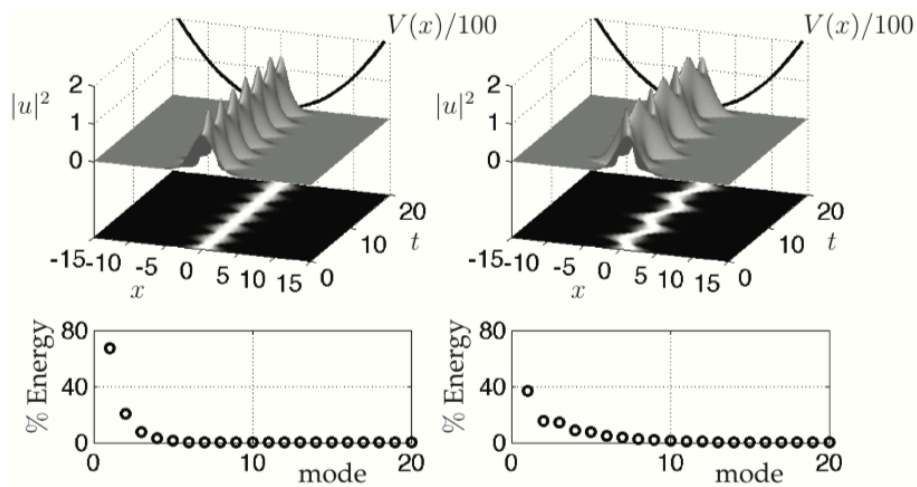
Uno puede generar una solución con condición inicial dada, por ejemplo:

$$u(x, 0) = e^{-0.2(x-x_0)^2}$$

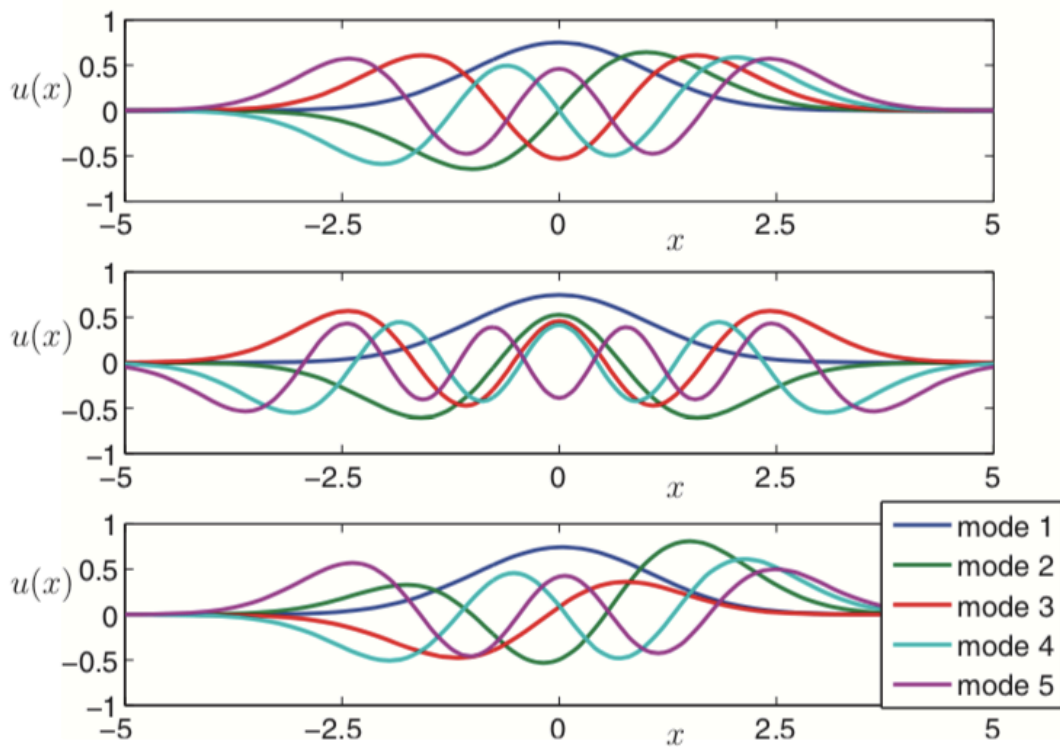
Reemplazando en la expresión los coeficientes:

$$a_k = \langle u(x, 0), \psi_k \rangle.$$

Aquí mostramos simulaciones con  $x_0 = 0$  (izq),  $x_0 = 1$  (der).



Si tomamos 100 cuadritos, y calculamos los modos, tenemos esto:



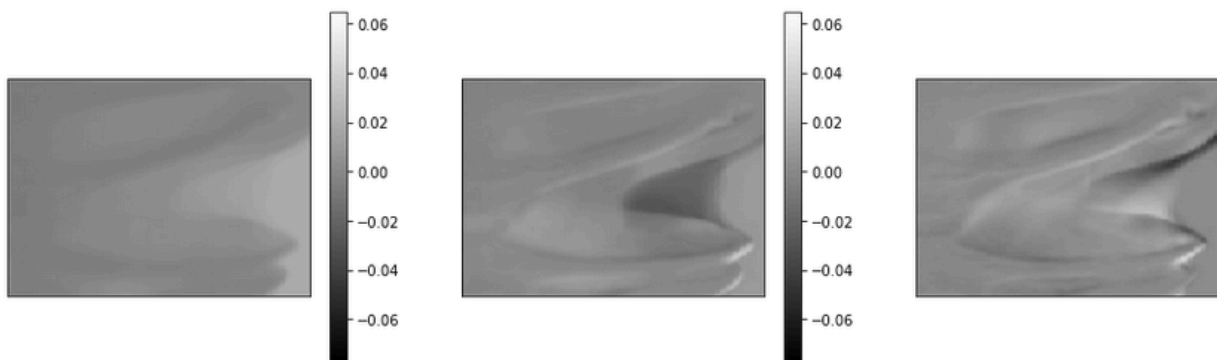
La primera figura muestra los primeros cinco modos de Hermite (arreglados tal que el auto valor es de modulo creciente). En el segundo, lo que sale de hacer SVD, truncando a  $r=5$ , a partir de la primera condición inicial, que es par. Notemos que son solo pares, porque la

evolución será par. Los cinco modos mas importantes que se obtienen de hacer el análisis de los datos que se siguen de la condición inicial con  $x_0 = 1$ , esos si generan los modos asimétricos (comparar el modo rojo del tercer panel, con el verde del primero), y este conjunto si es muy similar a lo que se obtiene con los polinomios de Hermite.

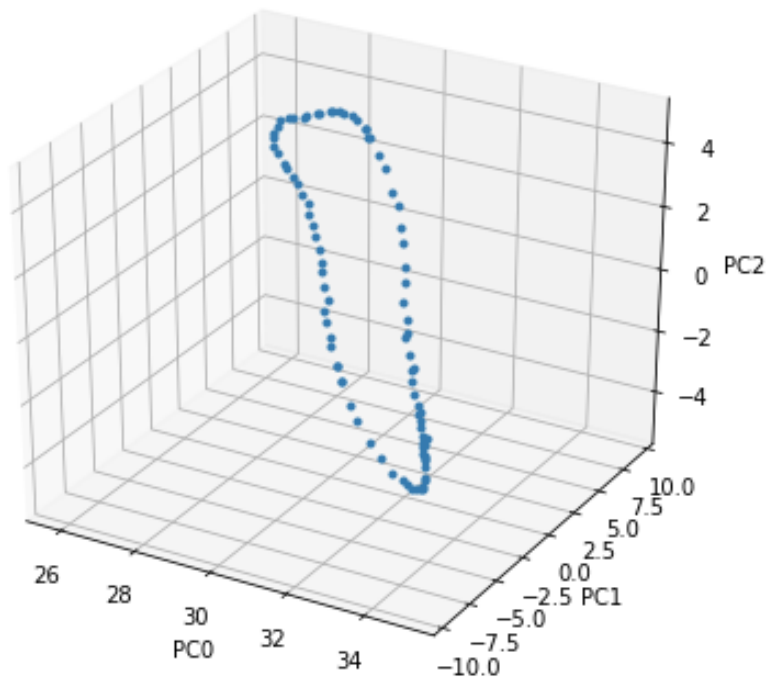
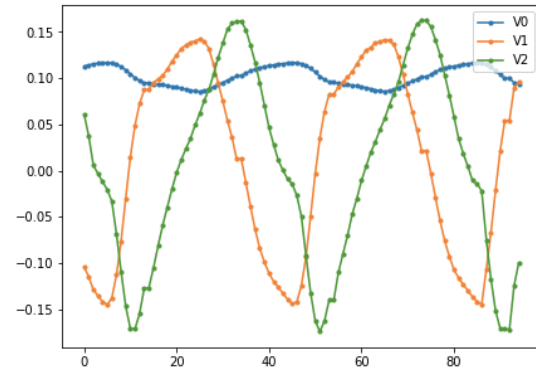
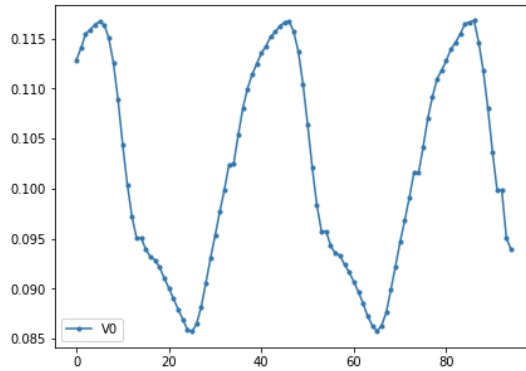
## Practico

¿Cuan difícil es calcular los modos empíricos?

```
>>> import numpy as np
>>> X = np.random.rand(5, 3) % create random data matrix
>>> U, S, V = np.linalg.svd(X, full_matrices=True) % full SVD
>>> Uhat, Shat, Vhat = np.linalg.svd(X, full_matrices=False)
    % economy SVD
```

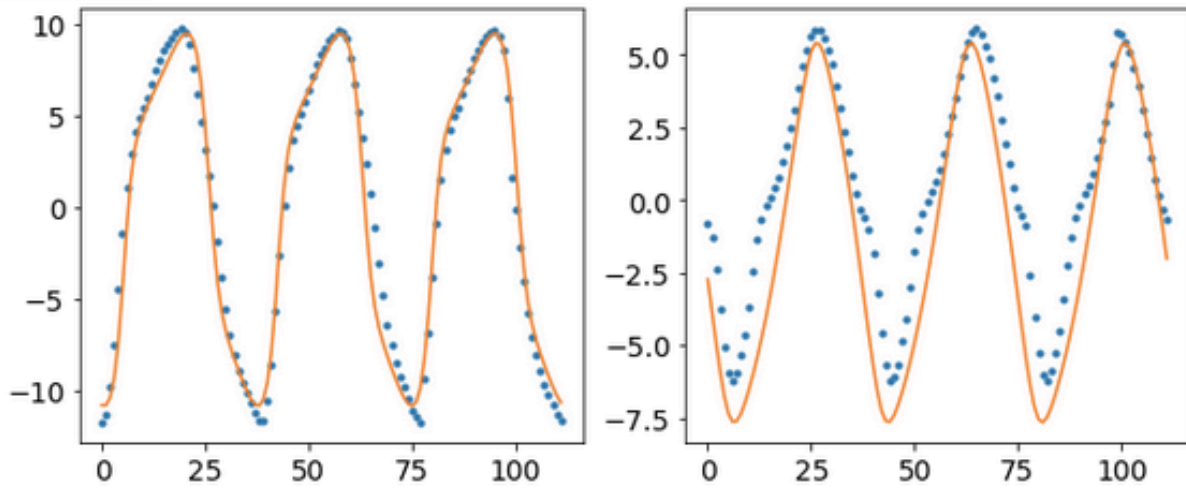
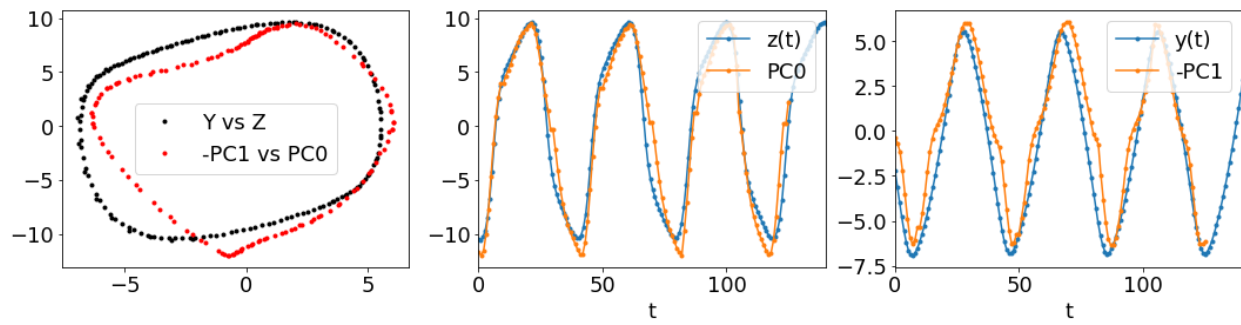


<matplotlib.legend.Legend at 0x7f2100fb9908>



$$\begin{cases} \dot{y} = az \\ \varepsilon \dot{z} = (z_o + \alpha y - z)(-z_o + \alpha y - z)(m y + b - z) - \varepsilon cz \end{cases}$$





name	value	initial value	min	max	vary
A	976.835603	906.5554992211789	0.00000000	1000.00000	True
B	1.79037878	1.4586126325004711	-4.00000000	4.00000000	True
C	523.547412	529.0970608690453	0.00000000	1000.00000	True
m	2.64295021	2.809144517200113	0.50000000	3.00000000	True
eps	0.06802982	0.0628527515210699	1.0000e-04	0.10000000	True
alpha	0.98095800	0.8264166097250599	0.50000000	2.00000000	True
z0	12.5217663	12.550002008816623	8.00000000	16.0000000	True
gamma	1.0334e-04	9.4e-05	1.0000e-06	1.0000e-03	True

### **Conclusión:**

1. Existe un modo algorítmico, a partir de los datos, de proyectar una dinámica espacio-temporal a un sistema dinámico: la proyección en auto estados empíricos (SVD, PCA, ...)
2. Este modo es optimo desde el punto de vista lineal
3. ¿Como escribir las ecuaciones para esta dinámica proyectada?
4. ¿Hay algo mejor para hacer?