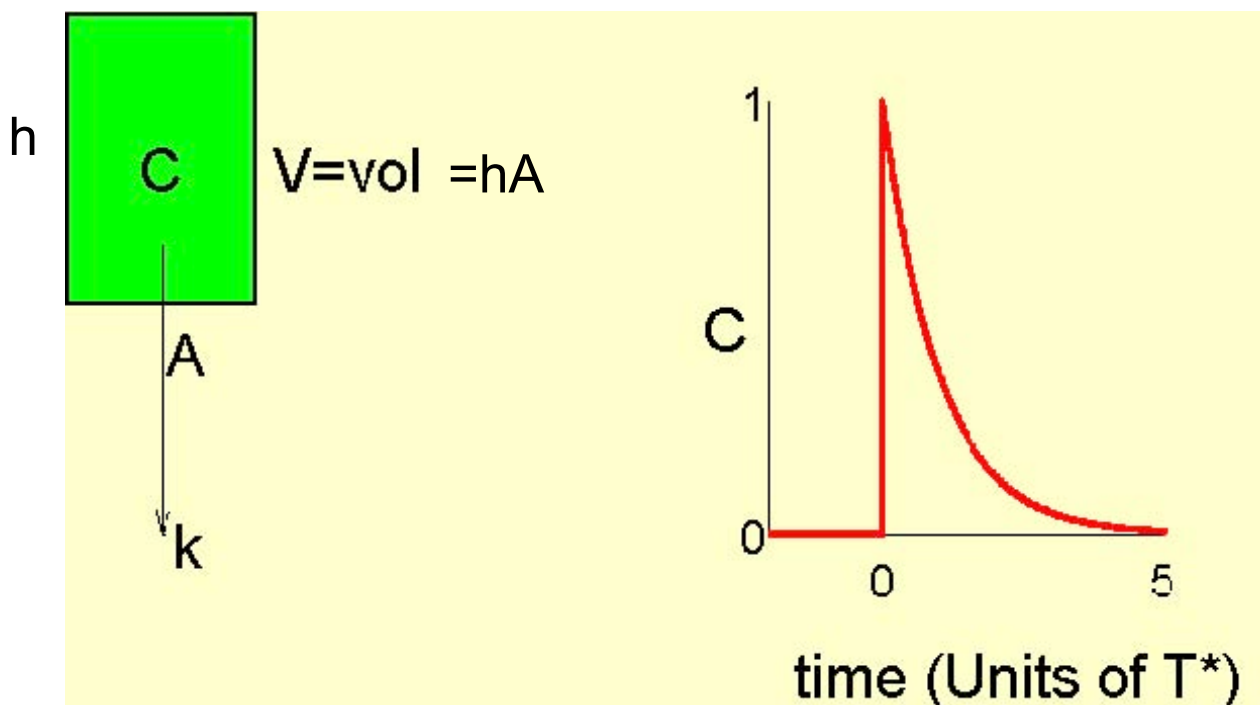


Intro (or review) Linear Systems: One-Box Model or Linear Filter

We start with the deceptively simple problem of computing the concentration of a chemical species (or could be the volume of water) in a box that receives input of mass from the external environment, and it exports mass to a sink (external). The exercise starts us thinking about how to formulate a problem involving an (approximately) linear system of reservoirs, and how to discover the analytical properties of coupled systems of such reservoirs. Linear systems also underlie most filtering and smoothing procedures, interpolation methods, time series analysis, stochastic models (e.g. Markov chains), and many aspects of statistical inference. *We will make the simple box model into a Markov chain at the end of this lecture.*

Mostly, we know these results already, but...a refresher may be in order, and there is always more to learn.

Start with an example. Let us make a simple mathematical model for the concentration C (**units: molecules m^{-3}**) of a gas in a box. The volume of the box is V . Gas inside the box leaks at a rate k proportional to the concentration. The **flux** out of the box is the number of molecules leaving per unit area per unit time, kC , and the total number of molecules leaving the box per unit time is kCA . (Note that k has units of velocity, cm/s, and is often called the "exchange velocity"). We use intensive quantities by design, with total volumes represented explicitly (why do that?).



Equivalent example: Volume of water in a bucket, with a hole in the bottom:

$$\rho \, dh/dt = -k P = -\rho k h; \quad k = \text{conductance (units, } s^{-1}\text{)}. \quad \rho \, dV/dt = \rho A dh/dt$$

This model also represents a linear filter of a time varying input $C(t)$, or even a Markovian system with transition probability k given discrete time steps. We will examine this equivalence shortly.

We can now write the mass balance equation for the box: $dCV/dt = -kAC$ (1.1) or, if we define h as the height of the box and divide by V , $dC/dt = -C/T^*$, (1.2)

$$T^* = h/k \equiv \text{lifetime (time constant)}.$$

Example 1

Let us solve this equation for the case of zero initial concentration, perturbed by an input of concentration C_0 at time zero. The solution (by separation of variables) is

$$C(t) = 0, \quad t < 0 \\ = C_0 \exp(-t/T^*) \quad , \quad t \geq 0$$

We see that the lifetime T^* is the characteristic time for an impulse to decay by factor $1/e$. The concept of lifetime, or time constant, is rigorously defined as this **e-folding time** for a system described by a first order differential equation in t . **T^* entirely defines the temporal behavior of the system.**

Example 2

Consider the response of this system to a constant input P (molecules/s), which starts abruptly at time $t = 0$:

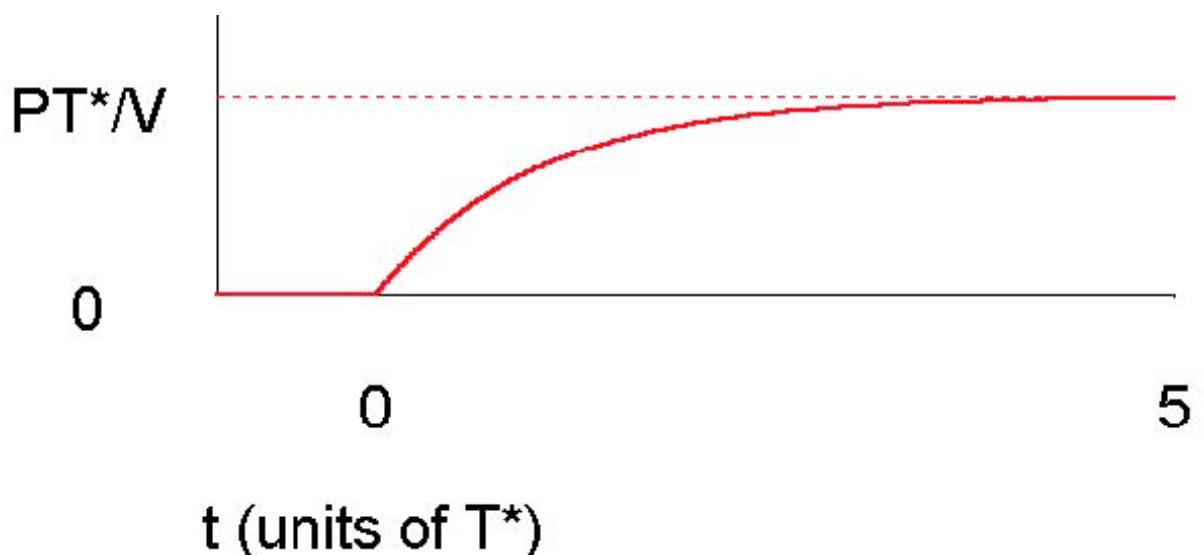
$$dC/dt = -C/T^* + P/V \quad (1.3)$$

The solution to this equation is given by the general solution to the homogeneous equation (1.2) plus a particular solution to (1.3):

$$C(t) = A \exp(-t/T^*) + PT^*/V$$

We get the value for the constant A by substituting for the initial condition $C(0) = 0$ to yield the full solution:

$$C(t) = PT^*/V [1 - \exp(-t/T^*)] \quad , \quad \text{for } t \geq 0 \quad (1.4)$$



The quantity PT^*/V is the **asymptotic concentration** that will persist for all time once the system has adjusted to the input. This stationary solution is given a special name, the **steady-state** of the system.

$t' = t/T^*$ or $t = T^*t'$ <non-dimensional>

$dc/dt' = -c + PT^*/V$ defines invariant problem, with $PT^*/V = I^*$, input rate unit c in unit T^*

Key concepts from this ultra-simple, very familiar example

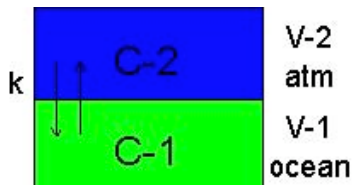
- Set up the problem in terms of a first-order differential equation; attend to units, and the framing.
- The solutions have exponential behavior.
- If the exponent is negative, the solution is unconditionally stable.
- The **time constant** is the characteristic time for a perturbation to decay. It is defined rigorously in terms of time scale derived from parameters in a *first order differential equation* or set of such equations.
- Under some circumstances (forcings, boundary conditions) the system may have a **steady-state** solution towards which the system evolves.

stable: a slight disturbance in a system does not produce too disrupting an effect on that system

2-box model

The [1-box model](#) might represent an atmosphere with an unspecified sink at the bottom (e.g. CO₂ dissolves in the ocean, and is permanently lost). Let's add a second box, to allow for 2-way exchange. This model might represent CO₂ exchange between the atmosphere and the surface ocean.

The net flux (*mass per unit area per unit time*) from the atmosphere to the ocean is given by $-k(C_2 - C_1)$



analogous to the 1-box case, and the net flux from the ocean to the atmosphere is $+k(C_2 - C_1)$. If A is the area of the interface and h_1 and h_2 are the heights (for the moment we ignore a complication due to the solubility of CO₂ in water), then the model equations are:

$$d(C_2 V_2)/dt = -kA(C_2 - C_1) \quad (2.1.1)$$

$$d(C_1 V_1)/dt = +kA(C_2 - C_1) \quad (2.1.2)$$

Divide by A to get

$$h_2 \, d(C_2)/dt = -k(C_2 - C_1) \quad (2.2.1) \quad h_1 \, d(C_1)/dt = +k(C_2 - C_1) \quad (2.2.2)$$

There is a simple way to solve these two coupled equations: add (2.2.1) + (2.2.2) to obtain

$$h_1 \, dC_1/dt + h_2 \, dC_2/dt = d/dt (h_1 C_1 + h_2 C_2) = 0, \quad (2.3)$$

or $h_1 C_1 + h_2 C_2 = M$ (a constant), i.e. mass in the system is conserved (part of the model design).

An important consequence is that, if we place additional mass into the system, such that ΔC_1 and ΔC_2 are concentration changes at $t=0$, the quantity (proportional to the total mass added at $t=0$)

$$\Delta M = h_1 \Delta C_1 + h_2 \Delta C_2$$

will not change with time even though ΔC_1 and ΔC_2 will in general change.

We may solve for the time-variable part of the system by subtracting 2.2.2 from 2.2.1, to obtain

$$-h_1 \, dC_1/dt + h_2 \, dC_2/dt = 2k(C_1 - C_2) \quad (2.4.1)$$

We now want to reformulate (2.4.1) to be a differential equation for a single variable, for which one choice might be $X = C_2 - h_1/h_2 C_1$. This is simple to do: since X and M are linearly independent, any function of C_1 and C_2 (in particular, $C_2 - C_1$), can be so represented.

With a bit of algebra, we obtain

$$\begin{aligned} d/dt (C_2 - h_1/h_2 C_1) &= -k(1/h_1 + 1/h_2) (C_2 - h_1/h_2 C_1) \\ &+ (k/h_2) (h_1 C_1 + h_2 C_2) (1/h_1 - 1/h_2). \end{aligned} \quad (2.4.2)$$

The general solution to this differential equation is: *...constant.....*

$$X(t) = X(0) e^{-k\left(\frac{1}{h_1} + \frac{1}{h_2}\right)t} + \left(1 - e^{-k\left(\frac{1}{h_1} + \frac{1}{h_2}\right)t}\right) \frac{M}{h_2} \left(\frac{h_2 - h_1}{h_2 + h_1}\right) \quad (2.4.3)$$

with X and M as defined above. *We could have used $Y = c_2 h_2 + c_1 h_1$ and $X = c_2 h_2 - c_1 h_1$ and kept symmetry....*

There is a wee surprise here, as may be seen by considering the simple case where $h_1 = h_2$ (note that in this case $X = C_2 - C_1$), which gives $d/dt(C_2 - C_1) = -2k/h (C_2 - C_1)$. *A unit perturbation at $t=0$ decays with time constant $T^* = h/(2k)$, i.e., twice as fast as in the 1 box case*

$$(C_2 - C_1)(t) = (C_2 - C_1)(t=0) \exp(-t[h/(2k)]). \quad (2.4.4)$$

The reason can, of course, be found in the structure of the model. Since mass is being conserved in the exchange between boxes 1 and 2, the **difference** in concentration is reduced by 2 units for every 1 unit exchanged: what is **taken away** from 1 is **added** to 2.

This ultra-simple model is observed to display some properties that are interesting in a general sense:

- If we decompose the problem into a particular set of linear combinations of the state variables C_1 and C_2 , we obtain two solutions that evolve in very simple ways over time. *We shall show below that these are the eigenvectors of a simple matrix problem, and the time constants are the reciprocals of the associated eigenvalues (i.e. the eigenvalues are the frequencies for relaxation of a perturbation).*
- Since the system conserves mass, one of the time constants is infinite (zero eigenvalue). *This will in general be the case for a closed system, because summing all the dX/dt equations will have to give zero on the RHS (definition of close system, conserved tracer does not decay with time).* right hand side
- The time constant for relaxation of a difference between the compartments is smaller than the time constant for an analogous open system.

Two-box model with loss (2 time scales)

Now let's consider the 2-box model in which we place a radioactive gas with exponential decay time T_d . For simplicity let's assume $h_1 = h_2 = h$. Then the equations become:

$$dC_1/dt = (C_2 - C_1)/T^* - C_1/T_d \quad (2.5.1)$$

$$dC_2/dt = - (C_2 - C_1)/T^* - C_2/T_d \quad (2.5.2)$$

We can solve these equations very much as we did for 2.4, first adding 2.5.1 and 2.5.2,

$$d(C_1 + C_2)/dt = - (C_1 + C_2) / T_d ,$$

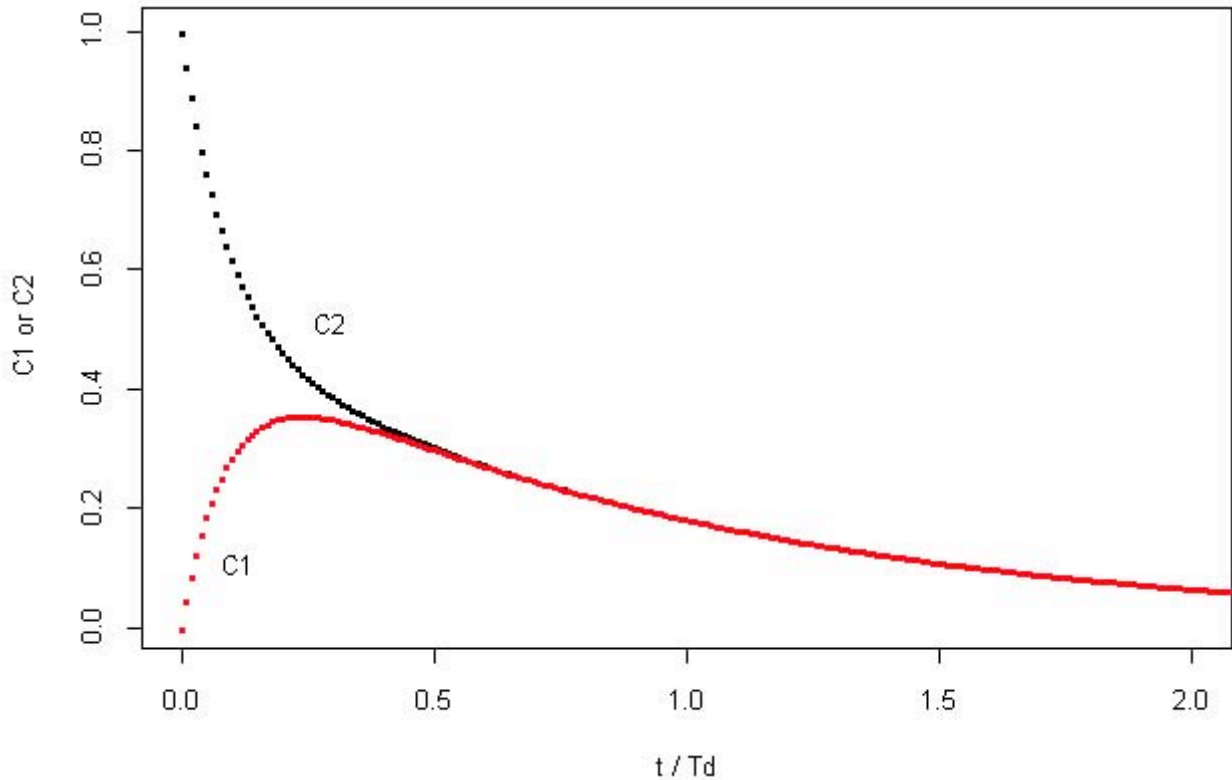
which has solution $(C_1 + C_2) = (C_1 + C_2)_0 \exp(-t/T_d)$.

If we subtract the equations we obtain

$$d(C_2 - C_1)/dt = -2 (C_2 - C_1)/T^* - (C_2 - C_1)/T_d ,$$

which has the solution $(C_2 - C_1) = (C_2 - C_1)_0 \exp(-t/T_m)$, where $T_m = T_d T^* / (2 T_d + T^*)$, i.e. once again we have two simple exponentials. The previously conserved mode (with zero decay rate, or $T_d \rightarrow \infty$), now decays with time constant T_d . The other mode relaxes at a faster rate than it did before. Note that $T_m < T_d$ (follows from $(2 T_d + T^*) > T^*$). Therefore the uniform decay rate is the slowest decay rate.

The following shows a graph of the solutions to 2.5 with $T_d = 5 T^*$, where we have put all of the gas initially into box 2.



Note how complex time dependence is obtained from the sum of 2 simple exponentials. Prather [1986] for example discusses the importance of the slowest-decaying mode, in particular, he notes that at long times the entire system relaxes with this time scale, independent of the initial perturbation. He argued that this was true for simple models like this one, and held also for complex models like GCMs with 80,000 degrees of freedom, and even for the real atmosphere.

These ultra-simple examples have shown

- how we obtain eigenvalues (e.g. the decay frequencies $1/T_d$ and $1/T_m$) and eigenvectors (e.g. $(C1 + C2)$ and $(C1 - C2)$) for the simple linear system defined by our 2-box models: the linear combinations of species we have discovered decay like simple exponentials, providing a simple way to express general dynamical behavior in the system; and
- illustrated the fundamental importance of the slowest-decaying mode.

The manual method of solving sets of differential equations for the modes that relax like simple exponentials is evidently laborious; the next section discusses general treatments using matrix algorithms.

Matrix formulation of the 2-box model

Set $h_1 = h_2$, and represent the state of the system as a vector

$$\underline{c} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix}$$

where c_1 and c_2 are the concentrations in boxes 1 and 2 computed from

$$dc_1/dt = k/h (c_2 - c_1)$$

$$h_1 = h_2 = h$$

$$dc_2/dt = k/h (c_1 - c_2)$$

In matrix form these equations become

$$d\underline{c}/dt = \frac{k}{h} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \underline{c}; \text{ note that the determinant } \begin{vmatrix} -1 & 1 \\ 1 & -1 \end{vmatrix} = 0, \text{ i.e. the matrix is singular.}$$

(the inverse does not exist because the first eigenvalue is zero). We already know the two eigenvectors and the associated eigenvalues,

$$\frac{k}{h} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 0 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \frac{k}{h} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = -2 \frac{k}{h} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ for the eigenvalue } 0 \text{ and } \frac{2k}{h} \text{ respectively.}$$

The eigenvalues should be 0 and $-2k/h$

We can readily construct the normalized matrix of eigenvectors \underline{X}_e ,

$$\underline{X}_e = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}; \text{ note that for this case the eigenvectors are orthogonal (this is not true if } h_1 \neq h_2).$$

Eigenvectors are orthogonal, dot product is zero

\underline{X}_e is an orthogonal matrix

It is easy to show that the inverse of an orthogonal matrix with normalized columns, such as \underline{X}_e , is the complex conjugate of the transpose. Since \underline{X}_e is symmetrical in this simple case, $\underline{X}_e^{-1} = \underline{X}_e$.

a matrix \underline{Q} is orthogonal if its transpose is equal to its inverse:

The solution to the differential equation in matrix form is given by

$\underline{c}(t) = \exp(\underline{A}t)\underline{c}(0)$; if we multiply this equation on the left by the unit matrix $\underline{X}\underline{X}^{-1}$, and also insert this matrix in front of $\underline{c}(0)$, we obtain

$\underline{X}\underline{X}^{-1}\underline{c}(t) = \exp(\underline{A}t)\underline{X}\underline{X}^{-1}\underline{c}(0)$, or $\underline{c}'(t) = \underline{X}^{-1}\exp(\underline{A}t)\underline{X}\underline{c}'(0)$, where \underline{c}' is defined as $\underline{X}^{-1}\underline{c}(t)$ and represents the vectors expressed as linear combinations of the eigenvectors, instead of the original basis vectors.

Now look at the power series that defines $\exp(\underline{A}t)$, noting first that

$$\underline{X}^{-1}\underline{A}\underline{X} = \underline{\Lambda}, \text{ where } \underline{\Lambda} \text{ is the diagonal matrix of eigenvalues, } \underline{\Lambda} = \begin{bmatrix} 0 & 0 \\ 0 & -\frac{2k}{h} \end{bmatrix}, \text{ and since}$$

$\underline{\Lambda}^n$ is a matrix with each element on the diagonal raised to the n power, the power series

$$\text{defining the exponential } \underline{X}^{-1}\exp(\underline{A}t)\underline{X} = \exp(\underline{\Lambda}t) = \begin{bmatrix} \exp(0) & 0 \\ 0 & \exp(-\frac{2k}{h}t) \end{bmatrix}.$$

We have thus converted the rather awkward infinite sum of matrices to simple diagonal matrix using the eigenvectors and eigenvalues. There is a physical significance to this result--if we start with a perturbation to the system that corresponds to an eigenvector in terms of its distribution between the boxes, then the perturbation decays with time as a simple exponential. This is the reason we designate eigenvectors as the "normal modes" of the problem.

Now let us consider the problem we solved analytically, in which an initial perturbation c_0 is placed into box 2, i.e. $c_2=c_0$ and $c_1=0$ at $t=0$. Then

$$\underline{c}(0) = \begin{bmatrix} 0 \\ c_0 \end{bmatrix} \text{ and applying } \underline{X}^{-1} \text{ we obtain } \underline{c}'(0) = \begin{bmatrix} c_0 \\ -c_0 \end{bmatrix} \frac{1}{\sqrt{2}}, \text{ and } \underline{c}'(t) = \begin{bmatrix} c_0 \\ -c_0 e^{-(2k/h)t} \end{bmatrix} \frac{1}{\sqrt{2}}. \text{ To obtain } \underline{c}(t) \text{ we multiply } \underline{X} \underline{c}'(t) = \frac{c_0}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ -e^{-(2k/h)t} \end{bmatrix} = \frac{c_0}{2} \begin{bmatrix} 1 - e^{-(2k/h)t} \\ 1 + e^{-(2k/h)t} \end{bmatrix}.$$

We can see that this result is the same as obtained by solving the differential equations stepwise, and also how the result follows from a strategy for solution of the equations where we express the initial conditions (for $t=0$) as a linear combination of eigenvectors, each of which then decays with a simple exponential. Thus we obtain $\underline{c}(0) = \sum d_k \underline{v}_k$, $d_1=1$ and $d_2=-1$, where

$$\underline{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \underline{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ and we see that indeed } \underline{c}(t) \text{ is given by } \underline{c}(t) = e^{\underline{\Lambda}t} \underline{c}(0) = \sum_k d_k e^{\Lambda_k t} \underline{v}_k.$$

The eigenvectors are defined so that each one evolves in time with a single, well-defined time constant. In contrast, the individual components have multi-exponential decay rates.

##Two box mass transfer model as a Markov Chain

First, plot the two box model analytical solution

```
k.h=.01                      ##k/h = 0.01 s-1 (unit-1)
tt=0:1000                    #time vector
c1.c2=100*exp(-2*tt*k.h)    #solution (c1 - c2)
c1=(100 + c1.c2)/2
c2=100-c1
plot(tt,c1,type="l",col="red",ylim=c(0,100),lwd=3)
lines(tt,c2,type="l",col="blue",lwd=3,lty=2)

## Markov chain
n1=100;n2=0                  # number of particles, initial (arbitrary)
N1=n1;N2=n2                  # Nj accumulates solutions
for(i in 1:1000){
  if(n1==0)dn1=0 else dn1=sum(runif(n1)<k.h)
  if(n2==0)dn2=0 else dn2=sum(runif(n2)<k.h)
  ##runif: uniform distribution random numbers (Markov)
  n1= n1-dn1+dn2            ## new value in box 1
  n2= n2-dn2+dn1            ## in box 2 NB particle conserv
  N1=c(N1,n1)
  N2=c(N2,n2)
}
##add points to the plot, for the Markov chain problem
points(tt,N1,col="red",pch=16,cex=.6)
points(tt,N2,col="blue",pch=1,cex=.6)
legend("topright",legend=c("Box 1","Box 2"),col=c("red","blue"),
      pch=c(16,1),lty=c(1,2),text.col=c("red","blue"))
## save the figure
dev.copy(png,"Fig.Markov.2box.png");dev.off()
```

QUESTIONS: If we think about the population in each box as a data point, subject to random "shocks" at each time step, estimate the variance of the "shocks" or forcing. What is the variance around the steady state (above)? *Why are they different ???* NB--here we use the language of Time Series Analysis, rather than Markov chain, or filtering, or linear systems, ...

