REACTION NETWORKS

Scales:

- Molecular (diffusion, "motors"; ions, charges, bonds, binding sites, conformation, energetics, single events…)

- Cellular (transport, transporters, chemotaxis; signal transduction, transcription/translation, regulation, metabolism)

- Tissue/organ (blood/lymph flow, muscles, pumps; regulation tissue growth/death, development)

- Whole individual (neurohormonal regulation, pharmacokinetics, pharmacodynamics, genetic variation, cancer, immunology, infection, …)

FOCUS: finite dim. system of coupled differential (and algebraic) equations

- Rate laws (elementary, mass action, approximate, empirical)

- State variables: concentrations, masses, (fraction occupied binding sites, voltages, …)

- Dependent variables (data): concentration, specific activity, flux, flow,… of selected species; linear combinations or ratios

- Independent variables ("design"):
  - time
  - temperature, pH, fixed conc.
  - experimental probes (initial conditions, input fluxes; constant flux input → steady state)

…but will limit the focus:

- no spatial terms/PDE models (may be approximated by compartments coupled with transport)
- no process noise
- no model misspecification (though discrimination among candidate models can be addressed)
- only "noise" (if any) is random measurement error for dependent variables
- large # molecules; well stirred

Today at 2pm Beckman Auditorium
Dan Gillespie
"Stochastic Chemical Kinetics"
http://bnmc.caltech.edu/Events/
(Markov process, chemical master equation, stochastic simulation, …)

EXAMPLES

a) $X_1 \xrightarrow{k_1} X_2$

$$\dot{[X_1]} = -k_1 \cdot [X_1]$$
$$\dot{[X_2]} = k_1 \cdot [X_1]$$

(henceforth drop [.] conc. notation)

b) $X_1 + X_2 \underset{k_2}{\overset{k_1}{\rightleftharpoons}} X_3$

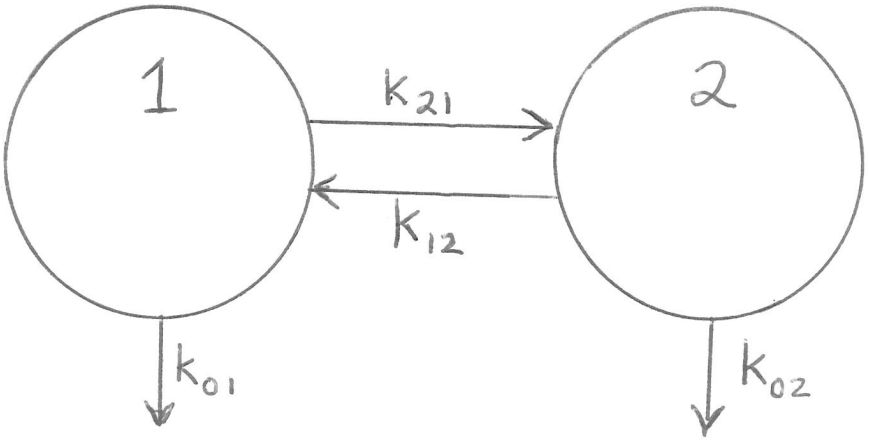
$$\dot{X}_1 = \dot{X}_2 = -k_1 \cdot X_1 \cdot X_2 + k_2 \cdot X_3$$

$$\dot{X}_3 = k_1 \cdot X_1 \cdot X_2 - k_2 \cdot X_3$$

Note: k_1 units [1/time]·[1/conc]
k_2 units [1/time]

$K_D = k_2 / k_1$ = dissociation constant
If $\dot{X}_1 = \dot{X}_2 = 0$, $X_1 \cdot X_2 / X_3 \equiv K_D$

c) compartmental models



q_i = mass of pool i ; $X_i = q_i / V_i$

$$\dot{q}_1 = -(k_{01} + k_{21})q_1 + k_{12}q_2$$

$$\dot{q}_2 = k_{21}q_1 - (k_{02} + k_{12})q_2$$

or $$\dot{\mathbf{q}} = \mathbf{K}\mathbf{q}$$

where $$(K)_{jj} = -\sum_{i \neq j} k_{ij}$$

d) "saturable" binding/kinetics

$$X_1 + R \underset{k_2}{\overset{k_1}{\rightleftharpoons}} X_1R \xrightarrow{k_3} X_2 \ (\ldots + R)$$

- $R_{TOT} \equiv R + X_1R$ is total receptor/enzyme (often fixed)

- $\dot{X}_1 = -k_1 \cdot X_1 \cdot (R_{TOT} - X_1R) + k_2 \cdot X_1R$

- under small parameter assumptions (e.g., if $\varepsilon = R_{TOT}/X_1(0) << 1$) get Michaelis-Menten approximation ("outer solution" for long time scale)

$$X_1 \xrightarrow{Km,\,Vmax} X_2$$

where $\dot{X}_1 = -(Vmax/(Km + X_1)) \cdot X_1$

$Vmax = k_3 \cdot R_{TOT}$

$Km = (k_2 + k_3)/k_1$

e) examples of other approximate rate laws (empirical, lumping, rate-limiting,…)

- Hill equation

$$\dot{X} = - Vmax \cdot (X^n / (K^n_{50} + X^n))$$

- IP_3 = inositol 1,4,5 trisphosphate:

IP_3 induced Ca^{++} flux out of ER (endoplasmic reticulum, a subcellular "compartment")

$$= \alpha \cdot \left(IP_3/(K_1 + IP_3)\right)^3 \cdot [Ca^{++}]_{ER}$$

- S –systems (E. Voit)

$$\dot{X}_1 = \alpha X_3^{1.5} - \beta \cdot X_1^{0.5} \cdot X_3^{-1}$$

…

- various inhibition/stimulation schemes (feedback loops)

A mathematical model of metabolic insulin signaling pathways

Ahmad R. Sedaghat, Arthur Sherman and Michael J. Quon

AJP - Endo 283:1084-1101, 2002. First published Jul 2, 2002; doi:10.1152/ajpendo.00571.2001

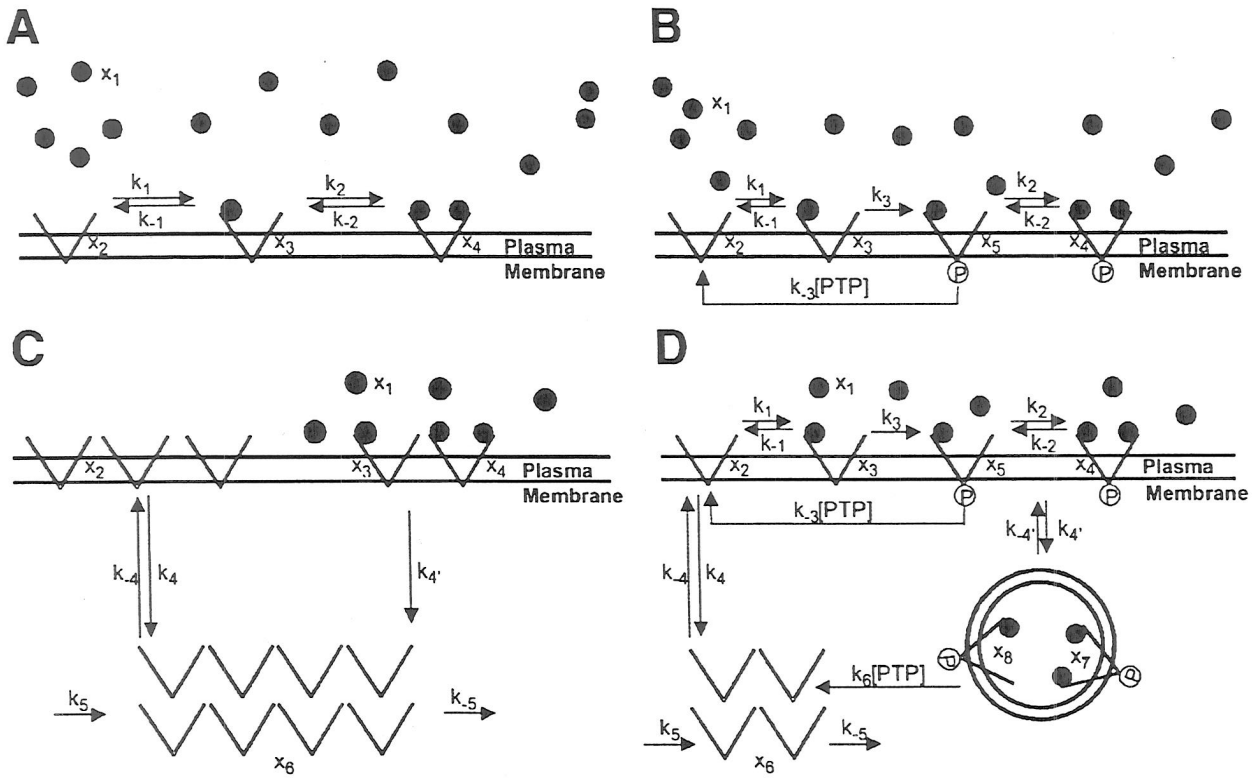


Fig. 1. Schematic of insulin receptor binding and life cycle subsystems. *A*: previously validated model of insulin binding kinetics (57). x_1, Free insulin concentration (system input); x_2, free receptor concentration; x_3, concentration of receptors with 1 molecule of insulin bound; x_4, concentration of receptors with 2 molecules of insulin bound; k_1 and k_{-1}, association and dissociation rate constants, respectively, for the first molecule of insulin to bind the receptor; k_2 and k_{-2}, association and dissociation rate constants, respectively, for the second molecule of insulin to bind the receptor. *B*: receptor binding subsystem extended to include receptor autophosphorylation and dephosphorylation. x_5, Concentration of once-bound phosphorylated receptors; x_4, redefined as concentration of twice-bound phosphorylated receptors; k_3, rate constant for receptor autophosphorylation; k_{-3}, rate constant for receptor dephosphorylation; [PTP], a multiplicative factor modulating k_{-3} that represents the relative activity of protein tyrosine phosphatases (PTPases) in the cell that dephosphorylate the insulin receptor. *C*: previously validated model of insulin receptor recycling (36). x_6, Concentrations of intracellular receptors; k_4, endocytosis rate constant for free receptors; k_{-4}, exocytosis rate constant; $k_{4'}$, endocytosis rate constant for bound receptors; k_5, zero order rate constant for receptor synthesis; k_{-5}, constant for receptor degradation. *D*: extension of insulin receptor binding and recycling subsystems that includes phosphorylated receptors. x_7 and x_8, concentration of twice-bound and once-bound intracellular phosphorylated receptors, respectively; $k_{-4'}$, exocytosis rate for twice-bound and once-bound intracellular phosphorylated receptors; k_6, dephosphorylation rate constant for intracellular receptors that is modulated by the multiplicative factor [PTP].

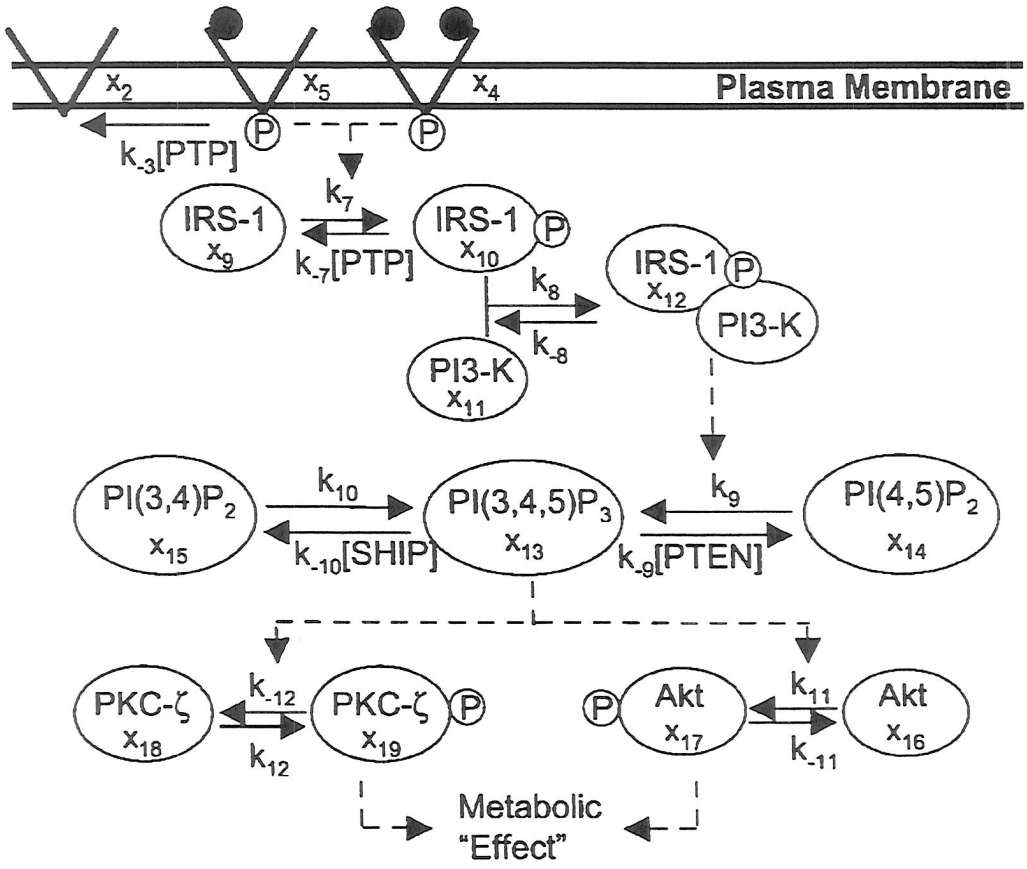


Fig. 2. Schematic of postreceptor signaling subsystem. x_9, Concentration of unphosphorylated insulin receptor substrate (IRS)-1; x_{10}, concentration of tyrosine-phosphorylated IRS-1; x_{11}, concentration of free phosphatidylinositol 3-kinase (PI 3-kinase; PI3-K); x_{12}, concentration of the phosphorylated IRS-1/activated PI 3-kinase complex; x_{13}, x_{14}, and x_{15}, percentages of various phosphoinositol lipids in the cell; x_{16} and x_{17}, percentages of unphosphorylated and phosphorylated Akt in the cell, respectively; x_{18} and x_{19}, percentages of unphosphorylated and phosphorylated protein kinase C (PKC)-ζ in the cell, respectively. The rate constants k_7 to k_{12} and k_{-7} to k_{-12} govern the conversion between state variables as indicated. [PTP] is a multiplicative factor modulating k_{-7} that represents the relative activity of PTPases in the cell that dephosphorylate IRS-1. [PTEN] and [SHIP] are multiplicative factors modulating k_{-9} and k_{-10}, respectively, that represent the relative activity of these lipid phosphatases in the cell. Arrows with solid lines indicate first-order reactions. Arrows with dashed lines indicate reactions where the value of a state variable influences the value of the rate constant.

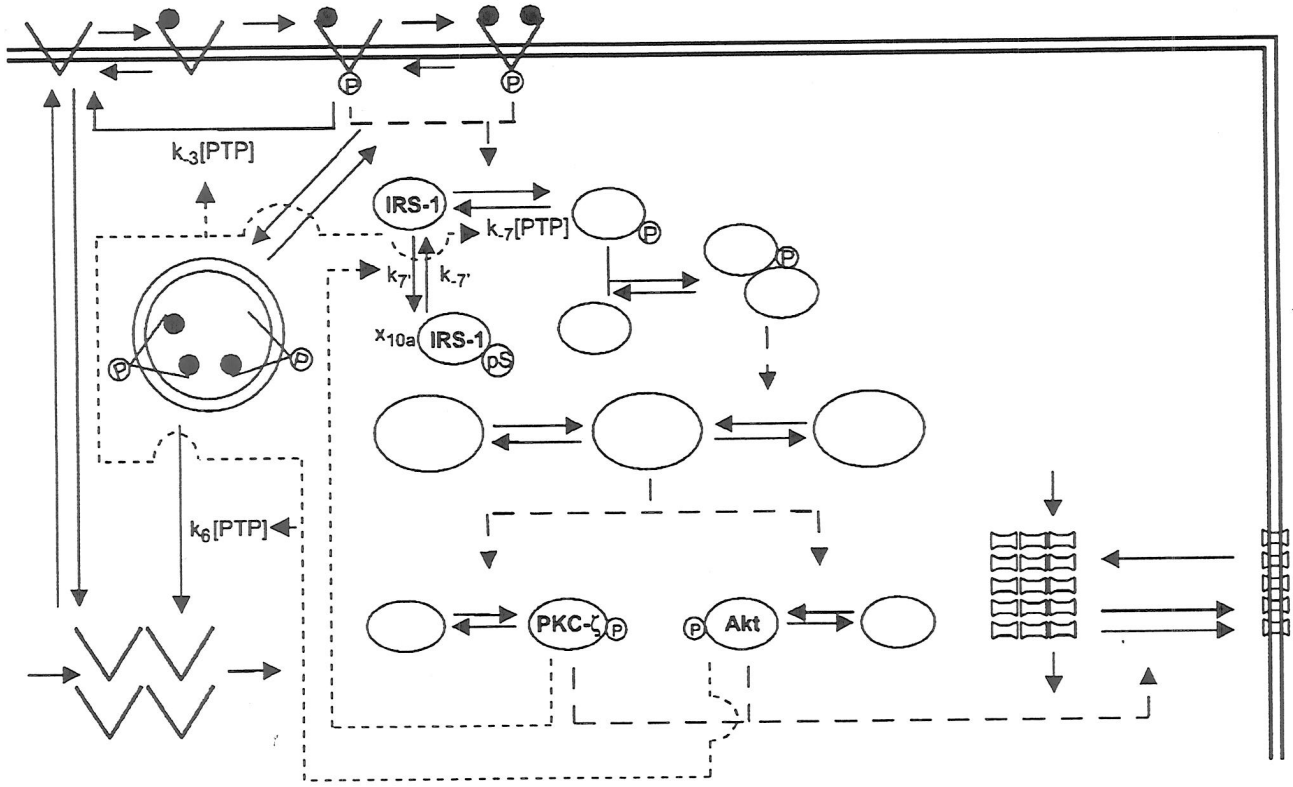


Fig. 5. Complete model of metabolic insulin signaling pathways with feedback. Identical to model shown in Fig. 4 except that new elements comprising positive and negative feedback pathways are indicated by dotted lines. PKC-ζ serine phosphorylates IRS-1 to create a negative feedback pathway, and Akt phosphorylates PTP1B to create a positive feedback pathway.

$$x_1 = \text{insulin input} \tag{6}$$

$$dx_2/dt = k_{-1}x_3 + k_{-3}[\text{PTP}]x_5 - k_1x_1x_2 + k_{-4}x_6 - k_4x_2 \tag{7}$$

$$dx_3/dt = k_1x_1x_2 - k_{-1}x_3 - k_3x_3 \tag{8}$$

$$dx_4/dt = k_2x_1x_5 - k_{-2}x_4 + k_{-4'}x_7 - k_{4'}x_4 \tag{9}$$

$$dx_5/dt = k_3x_3 + k_{-2}x_4 - k_2x_1x_5 - k_{-3}[\text{PTP}]x_5 + k_{-4'}x_8 - k_{4'}x_5 \tag{10}$$

$$dx_6/dt = k_5 - k_{-5}x_6 + k_6[\text{PTP}](x_7 + x_8) + k_4x_2 - k_{-4}x_6 \tag{11}$$

$$dx_7/dt = k_{4'}x_4 - k_{-4'}x_7 - k_6[\text{PTP}]x_7 \tag{12}$$

$$dx_8/dt = k_{4'}x_5 - k_{-4'}x_8 - k_6[\text{PTP}]x_8 \tag{13}$$

Differential equations governing phosphorylation of IRS-1 and subsequent formation of phosphorylated IRS-1/activated PI 3-kinase complex are

$$dx_9/dt = k_{-7}[\text{PTP}]x_{10} - k_7x_9(x_4 + x_5)/(IR_p) \tag{14}$$

$$dx_{10}/dt = k_7x_9(x_4 + x_5)/(\text{IR}_p) + k_{-8}x_{12} - (k_{-7}[\text{PTP}] + k_8x_{11})x_{10} \tag{15}$$

$$dx_{11}/dt = k_{-8}x_{12} - k_8x_{10}x_{11} \tag{16}$$

$$dx_{12}/dt = k_8x_{10}x_{11} - k_{-8}x_{12} \tag{17}$$

Activated PI 3-kinase converts the substrate phosphatidylinositol 4,5-bisphosphate [PI(4,5)P_2] to the product PI(3,4,5)P_3. This is modeled as a linear function so that k_9, the rate constant for generation of PI(3,4,5)P_3, is dependent on x_{12}, the amount of activated PI 3-kinase (see APPENDIX B for detailed derivation). 5′-Lipid phosphatases such as SHIP2 convert PI(3,4,5)P_3 to phosphatidylinositol 3,4-bisphosphate [PI(3,4)P_2] (6), whereas 3′-lipid phosphatases such as PTEN convert PI(3,4,5)P_3 to PI(4,5)P_2 (45). The differential equations describing interconversion between these phosphatidylinositides are

$$dx_{13}/dt = k_9x_{14} + k_{10}x_{15} - (k_{-9}[\text{PTEN}] + k_{-10}[\text{SHIP}])x_{13} \tag{18}$$

$$dx_{14}/dt = k_{-9}[\text{PTEN}]x_{13} - k_9x_{14} \tag{19}$$

$$dx_{15}/dt = k_{-10}[\text{SHIP}]x_{13} - k_{10}x_{15} \tag{20}$$

... etc.

Model with Feedback

Additional state variables are as follows

x_{10a} = concentration of serine-phosphorylated IRS-1

Additional equations are as follows

$$dx_9/dt = k_{-7}[\text{PTP}]x_{10} - k_7 x_9(x_4 + x_5)/(\text{IR}_p) + k_{-7'}x_{10a} \quad k_7'[\text{PKC}]x_9 \text{ (updated)}$$

$$dx_{10a}/dt = k_7[\text{PKC}]x_9 - k_{-7'}x_{10a}$$

Additional initial conditions are as follows

$$x_{10a}(0) = 0$$

Additional parameters are as follows

$$k_{7'} = \ln(2)/2 \text{ min}^{-1}$$

$$k_{-7'} = k_{7'}\,[(2.5/7.45)(3.70 \times 10^{-13})]/[(6.27 \times 10^{-13}) - (2.5/7.45)(3.70 \times 10^{-13})]$$

✓ $[\text{PTP}] = 1.00\,[1 - 0.25(x_{17}/(100/11)]$ for $x_{17} \leq (400/11)$, otherwise $[\text{PTP}] = 0$ (updated)

$$\text{PI3K} = k_8(3.70 \times 10^{-13})(1 \times 10^{-13})/[k_8(3.70 \times 10^{-13}) + k_{-8}] \text{ (updated)}$$

✓ $[\text{PKC}] = V_{max}x_{19}(t - \tau)^n/[K_d^n + x_{19}(t - \tau)^n]$

$$V_{max} = 20$$

$$K_d = 12$$

$$n = 4$$

$$\tau = 1.5$$

APPENDIX A

Model without Feedback

State variables are as follows

x_1 = Insulin input
x_2 = Concentration of unbound surface insulin receptors
x_3 = Concentration of unphosphorylated once-bound surface receptors
x_4 = Concentration of phosphorylated twice-bound surface receptors
x_5 = Concentration of phosphorylated once-bound surface receptors
x_6 = Concentration of unbound unphosphorylated intracellular receptors
x_7 = Concentration of phosphorylated twice-bound intracellular receptors
x_8 = Concentration of phosphorylated once-bound intracellular receptors
x_9 = Concentration of unphosphorylated IRS-1
x_{10} = Concentration of tyrosine-phosphorylated IRS-1
x_{11} = Concentration of unactivated PI 3-kinase
x_{12} = Concentration of tyrosine-phosphorylated IRS-1/activated PI 3-kinase complex
x_{13} = Percentage of $PI(3,4,5)P_3$ out of the total lipid population
x_{14} = Percentage of $PI(4,5)P_2$ out of the total lipid population
x_{15} = Percentage of $PI(3,4)P_2$ out of the total lipid population
x_{16} = Percentage of unactivated Akt
x_{17} = Percentage of activated Akt
x_{18} = Percentage of unactivated PKC-ζ
x_{19} = Percentage of activated PKC-ζ
x_{20} = Percentage of intracellular GLUT4
x_{21} = Percentage of cell surface GLUT4

Equations are as follows

$$x_1 = \text{insulin input}$$
$$dx_2/dt = k_{-1}x_3 + k_{-3}[PTP]x_5 - k_1x_1x_2 + k_{-4}x_6 - k_4x_2$$
$$dx_3/dt = k_1x_1x_2 - k_{-1}x_3 - k_3x_3$$
$$dx_4/dt = k_2x_1x_5 - k_{-2}x_4 + k_{-4'}x_7 - k_{4'}x_4$$
$$dx_5/dt = k_3x_3 + k_{-2}x_4 - k_2x_1x_5 - k_{-3}[PTP]x_5 + k_{-4'}x_8 - k_{4'}x_5$$
$$dx_6/dt = k_5 - k_{-5}x_6 + k_6[PTP](x_7 + x_8) + k_4x_2 - k_{-4}x_6$$
$$dx_7/dt = k_{4'}x_4 - k_{-4'}x_7 - k_6[PTP]x_7$$
$$dx_8/dt = k_{4'}x_5 - k_{-4'}x_8 - k_6[PTP]x_8$$
$$dx_9/dt = k_{-7}[PTP]x_{10} - k_7x_9(x_4 + x_5)/(IR_p)$$
$$dx_{10}/dt = k_7x_9(x_4 + x_5)/(IR_p) + k_{-8}x_{12} - (k_{-7}[PTP] + k_8x_{11})x_{10}$$
$$dx_{11}/dt = k_{-8}x_{12} - k_8x_{10}x_{11}$$
$$dx_{12}/dt = k_8x_{10}x_{11} - k_{-8}x_{12}$$
$$dx_{13}/dt = k_9x_{14} + k_{10}x_{15} - (k_{-9}[PTEN] + k_{-10}[SHIP])x_{13}$$
$$dx_{14}/dt = k_{-9}[PTEN]x_{13} - k_9x_{14}$$
$$dx_{15}/dt = k_{-10}[SHIP]x_{13} - k_{10}x_{15}$$
$$dx_{16}/dt = k_{-11}x_{17} - k_{11}x_{16}$$
$$dx_{17}/dt = k_{11}x_{16} - k_{-11}x_{17}$$
$$dx_{18}/dt = k_{-12}x_{19} - k_{12}x_{18}$$
$$dx_{19}/dt = k_{12}x_{18} - k_{-12}x_{19}$$
$$dx_{20}/dt = k_{-13}x_{21} - (k_{13} + k_{13'})x_{20} + k_{14} - k_{-14}x_{20}$$
$$dx_{21}/dt = (k_{13} + k_{13'})x_{20} - k_{-13}x_{21}$$

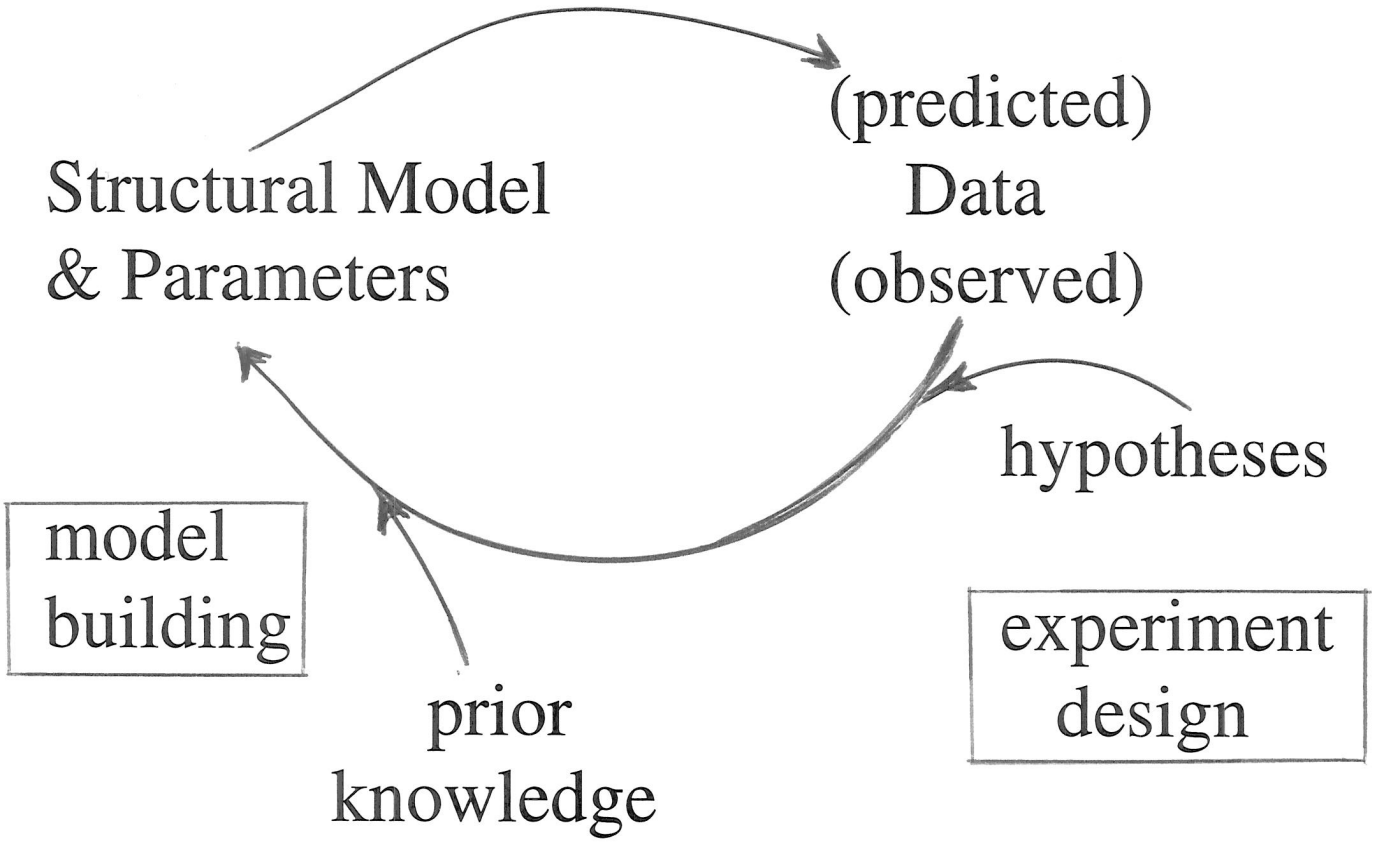
Initial conditions are as follows

$$x_1(0) = 0$$
$$x_2(0) = 9 \times 10^{-13}\ M$$
$$x_3(0) = 0$$
$$x_4(0) = 0$$
$$x_5(0) = 0$$
$$x_6(0) = 1 \times 10^{-13}\ M$$
$$x_7(0) = 0$$
$$x_8(0) = 0$$
$$x_9(0) = 1 \times 10^{-12}\ M$$
$$x_{10}(0) = 0$$
$$x_{11}(0) = 1 \times 10^{-13}\ M$$
$$x_{12}(0) = 0$$
$$x_{13}(0) = 0.31\%$$
$$x_{14}(0) = 99.4\%$$
$$x_{15}(0) = 0.29\%$$
$$x_{16}(0) = 100\%$$
$$x_{17}(0) = 0$$
$$x_{18}(0) = 100\%$$
$$x_{19}(0) = 0$$
$$x_{20}(0) = 96\%$$
$$x_{21}(0) = 4\%$$

Model parameters are as follows

$$k_1 = 6 \times 10^7\ M^{-1}\cdot min^{-1}$$
$$k_{-1} = 0.20\ min^{-1}$$
$$k_2 = k_1$$
$$k_{-2} = 100k_{-1}$$
$$k_3 = 2{,}500\ min^{-1}$$
$$k_{-3} = k_{-1}$$
$$k_4 = k_{-4}/9$$
$$k_{-4} = 0.003\ min^{-1}$$
$$k_{4'} = 2.1 \times 10^{-3}\cdot min^{-1}$$
$$k_{-4'} = 2.1 \times 10^{-4}\cdot min^{-1}$$
$$k_5 = 10k_{-5}\ M\cdot min^{-1} \text{ if } (x_6 + x_7 + x_8) > 1 \times 10^{-13}$$
$$= 60k_{-5}\ M\cdot min^{-1} \text{ if } (x_6 + x_7 + x_8) \le 1 \times 10^{-13}$$
$$k_{-5} = 1.67 \times 10^{-18}\ min^{-1}$$
$$k_6 = 0.461\ min^{-1}$$
$$k_7 = 4.16\ min^{-1}$$
$$k_{-7} = (2.5/7.45)k_7$$
$$k_8 = k_{-8}(5/70.775) \times 10^{12}$$
$$k_{-8} = 10\ min^{-1}$$
$$k_9 = (k_{9(stimulated)} - k_{9(basal)})(x_{12}/PI3K) + k_{9(basal)}$$
$$k_{9(stimulated)} = 1.39\ min^{-1}$$
$$k_{-9} = (94/3.1)k_{9(stimulated)}$$
$$k_{9(basal)} = (0.31/99.4)k_{-9}$$
$$k_{10} = (3.1/2.9)k_{-10}$$
$$k_{-10} = 2.77\ min^{-1}$$
$$k_{11} = (0.1k_{-11})(x_{13} - 0.31)/(3.10 - 0.31)$$
$$k_{-11} = 10 \ln (2)\ min^{-1}$$
$$k_{12} = (0.1k_{-12})(x_{13} - 0.31)/(3.10 - 0.31)$$
$$k_{-12} = 10 \ln (2)\ min^{-1}$$
$$k_{-13} = 0.167\ min^{-1}$$
$$k_{13} = (4/96)k_{-13}$$
$$k_{13'} = [(40/60) - (4/96)]k_{-13}\cdot(\text{Effect})$$
$$k_{14} = 96k_{-14}$$
$$k_{-14} = 0.001155\ min^{-1}$$
$$\text{effect} = (0.2x_{17} + 0.8x_{19})/(AP_{equil})$$
$$IR_p = 8.97 \times 10^{-13}\ M$$
$$[SHIP] = 1.00$$
$$[PTEN] = 1.00$$
$$[PTP] = 1.00$$
$$AP_{equil} = 100/11$$
$$PI3K = 5 \times 10^{-15}\ M$$

"forward problem"
(i.e., simulation)



"inverse problem"

- parameter estimation
- model verification/testing
- model discrimination
- ? parameter "optimization"
- ? reverse engineering

Structural (or *a priori*) Identifiability addresses the question of a unique inverse

Given:

- Specified experiment design
  - design space *S* (e.g., range of times can sample)
  - specified input(s)
  - set of dependent variables ($\mathbf{y}(t)$) to be observed, without error, over *S*

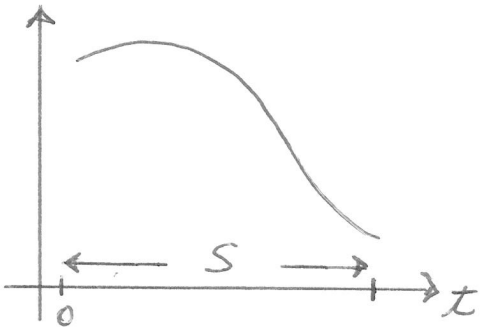
- Structural model $\mathbf{f}$, with unknown parameter vector $\theta \in \Theta \subset R^p$, that relates the design to observed data:

$$\mathbf{y}(t) = \mathbf{f}(t; \theta) \quad (+ \text{ no error})$$

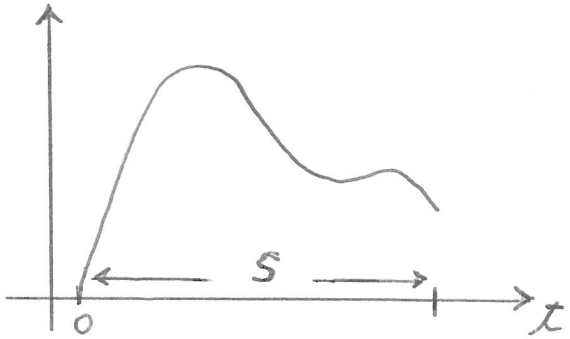
Question: Is there a unique choice of θ for every possible observed data set $\mathbf{y}$?
(i.e., is inverse map $\mathbf{y} \rightarrow \theta$ 1-to-1 or 1-to-many)

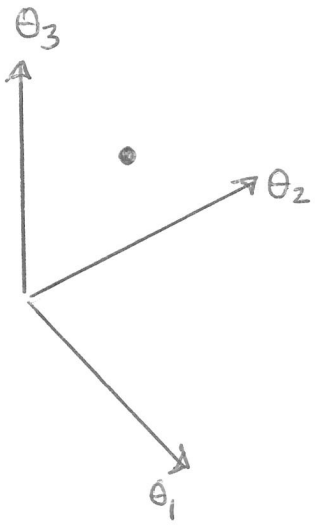
illustration:

$y_1(t) = X_1(t)$

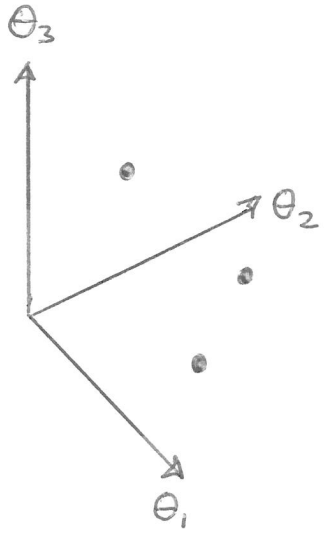


$y_2(t) = X_2(t) + X_3(t)$

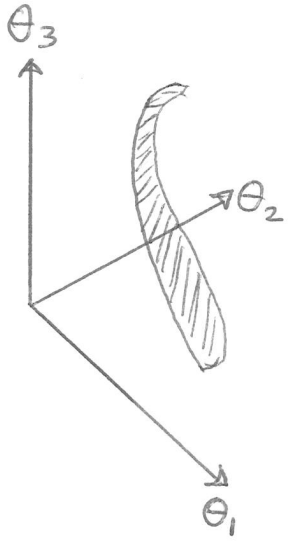




unique:
globally identifiable



separable:
locally identifiable



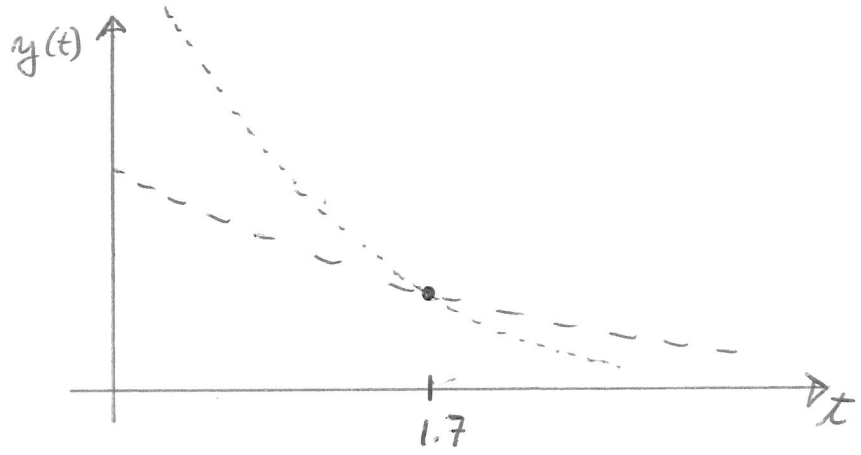
unidentifiable

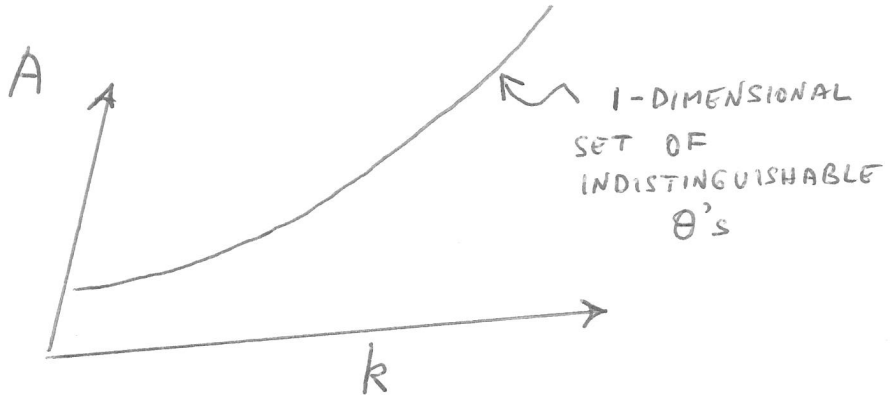
output indistinguishable manifold

trivial example: unidentifiable

$$y(t) = A \exp(-k \cdot t) \qquad \theta = (A, k)$$

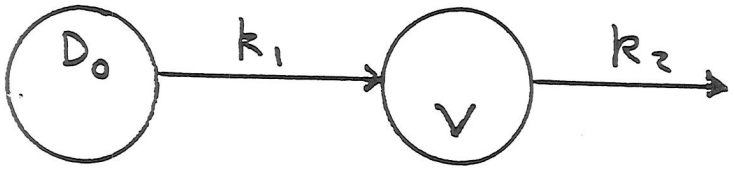
but $S = \{1.7 \text{ only}\}$





Note: *a priori* $k \geq 0$, which implies A is bounded below. If have upper bound for k then A would be "interval identifiable."

Example: local identifiability



y = CONCENTRATION IN CENTRAL POOL

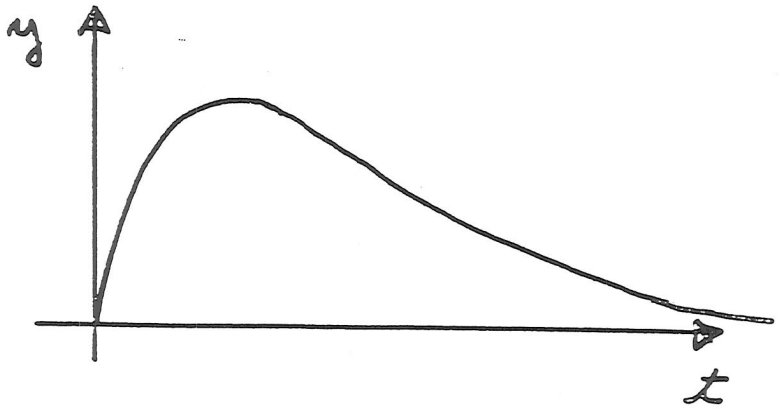
$$= \frac{D_0}{V} \frac{k_1}{k_1 - k_2} \left(e^{-k_2 t} - e^{-k_1 t} \right)$$

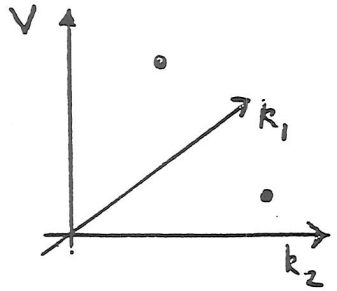
MODEL : ABOVE

CHOICE OF IND. VARIABLE : $0 \leq t < \infty$

ONLY CENTRAL POOL

$$\underline{\theta} = (k_1, k_2, V)$$



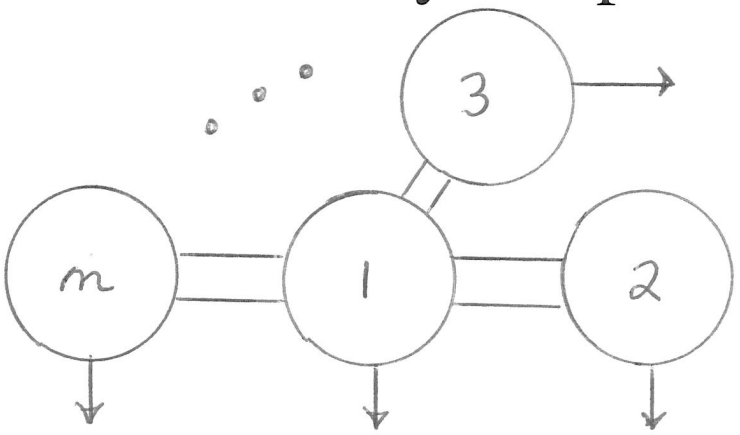


: 2 CHOICES OF $\underline{\theta}$

$$\theta = (k_1, k_2, V)$$

$$\theta' = \left(k_2, k_1, \frac{V k_2}{k_1}\right)$$

Example: mammillary comp. model



Input: known bolus dose into pool 1
Sample concentration of pool 1

$$y(t) = q_1(t)/V_1 = \sum_{i=1}^{n} A_i \exp(\lambda_i t)$$

(*a priori*: if k_{jj} distinct, j=2,…n, then

$$A_i > 0 \ , \lambda_i < \lambda_{i+1} \leqslant 0)$$

2n "macroparameters" $\{A_i, \lambda_i\}$ <u>identifiable</u>

3n-1 "microparameters" $\{k_{ij}, V_1\}$ <u>unident.</u>

2n <u>identifiable</u> parameter combinations:
$\{k_{11},\ldots,k_{nn}, k_{12}\cdot k_{21},\ldots k_{1n}\cdot k_{n1}, V_1\}$ which implies here that micros are interval ident.

Systematic Approaches to Identifiability

a) dimensionality arguments

b) linear systems – convert problem to the solvability of a finite system of nonlinear algebraic (polynomial) equations. (Computer Algebra)

- Laplace transforms
- finite number of $y(0^+)$, $y'(0^+)$, $y''(0^+)$, … (Markov matrices)
- transfer function topological method
- similarity transformation method (count all similarity transformations rendering it output indistinguishable)

c) nonlinear systems – system of equations may be infinite, greater dependence on initial conditions and type of input, etc. See Audoly, Chapman, Chappell, Saccomani references

d) numerical approach:

- pick nominal value θ^*

- generate $\mathbf{y_i} = \mathbf{f}(t_i, \theta^*)$, i=1,…,N
  with t_i sampling *S*

- idea is to evaluate in nbd of θ^*

$$Q(\theta) = \Sigma_i \, (\mathbf{y_i} - \mathbf{f}(t_i, \theta))' W_i (\mathbf{y_i} - \mathbf{f}(t_i, \theta))$$

(e.g., W_i may be a diag., positive matrix).
If system is locally identifiable, then $Q(\theta)$ will have a unique minimum (zero) at θ^*

- Compute PxP "information matrix"

$$M(\theta^*) = \Sigma_i \, M(t_i; \theta) \Big|_{\theta = \theta^*}$$

where $M(t_i; \theta) = (\partial \mathbf{y_i}/\partial\theta)' \, W_i \, (\partial \mathbf{y_i}/\partial\theta)$

- M full rank → locally identifiable

Note need in previous method for computing model sensitivities ($\partial y/\partial \theta$). Also used for real data (i.e., in presence of measurement noise) for

- numerical search for least squares estimate $\hat{\theta}$
- asymptotic covariance matrix for estimate $\hat{\theta}$

($COV(\hat{\theta})$ proportional to M^{-1})

Methods for computing sensitivities:

- one-sided or two-sided differences
- secant approximations (BMDPAR)
- augment state differential equations with sensitivity equations ($\partial \dot{y}/\partial \theta$).

Options if Model is Unidentifiable

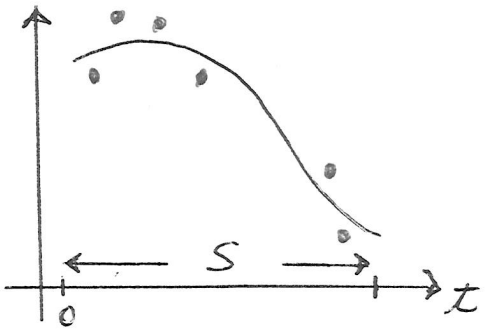
- search for improved design (augment sampled outputs and/or experimental probes)

- evaluate identifiable parameter combinations (e.g., W3MAMCAT www.biocyb.cs.ucla.edu/biocybmodeling.html)
  - may find useful subset of parameters that are identifiable
  - may provide interval identifiability

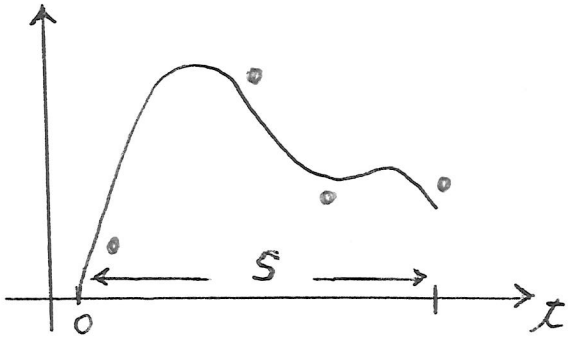
- augment with external information (e.g., Bayesian priors)

illustration:

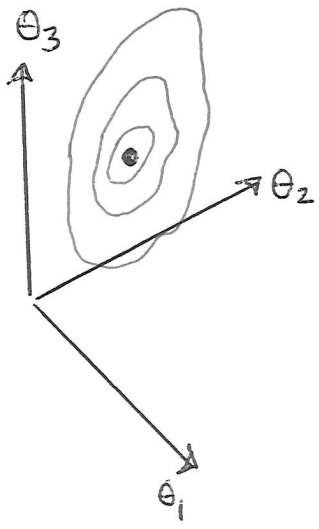
$y_1(t) = X_1(t)$

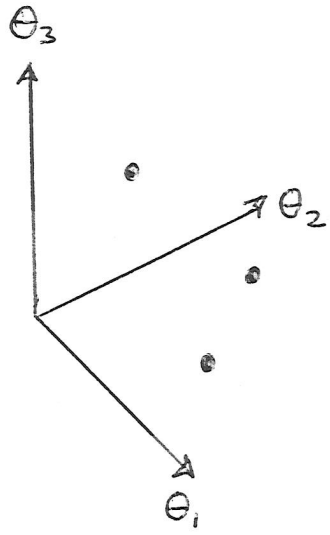


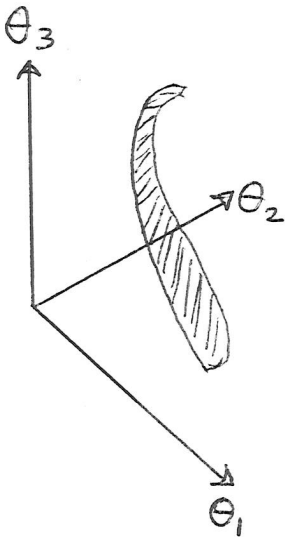
$y_2(t) = X_2(t) + X_3(t)$



FINITE, NOISY DATA







unique:
globally identifiable

separable:
locally identifiable

unidentifiable

output indistinguishable manifold

Even if model is identifiable, for real data must be concerned about "numerical identifiability" i.e., how presence of data noise affects precision of estimates:

$$\mathbf{y}(t) = \mathbf{f}(t; \theta) \quad + \text{ error}$$

- confidence regions for θ (or Bayesian posterior credibility regions) depend critically on design and model of measurement error (e.g., additive vs multiplicative error; error distribution; variance as a function of signal)

- may find subsets of parameters or functions of parameters estimated with adequate precision

- optimal experiment design may improve estimation precision (e.g., choose design $\{t_1,\ldots,t_N\}$ that minimizes $\det(\mathrm{COV}(\hat{\theta}))$.

Selected Bibiliography (*Reaction Networks & Identifiability*, E. Landaw, IPAM, Mar 15, 2006)

Atkinson A.C. and B. Bogacka (2002). "Compound and other optimum designs for systems of nonlinear differential equations arising in chemical kinetics." Chemometrics and Intelligent Laboratory Systems **61**:17-33.

Atkinson, A. C. and A. N. Donev (1992). Optimum Experimental Designs. Oxford, Clarendon Press (Oxford University Press).

Audoly, S., G. Bellu, et al. (2001). "Global identifiability of nonlinear models of biological systems." IEEE Trans Biomed Eng **48**(1): 55-65.

Audoly, S., L. D'Angio, et al. (1998). "Global identifiability of linear compartmental models--a computer algebra algorithm." IEEE Trans Biomed Eng **45**(1): 36-47.

Barrett, P. H., B. M. Bell, et al. (1998). "SAAM II: Simulation, analysis, and modeling software for tracer and pharmacokinetic studies." Metabolism **47**(4): 484-92.

Bates, D. M. and D. G. Watts (1985). "Multiresponse estimation with special application to linear systems of differential equations." Technometrics **27**: 329-339.

Bates, D. M. and D. G. Watts (1988). Nonlinear Regression Analysis and its Applications. New York, Wiley.

Chapman, M. J., K. R. Godfrey, et al. (2003). "Structural identifiability for a class of non-linear compartmental systems using linear/nonlinear splitting and symbolic computation." Mathematical Biosciences **183**:1-14. (*Erratum* **183**:215)

Chappell, M. J., K. R. Godfrey, et al. (1990). "Global identifiability of the parameters of nonlinear-systems with specified inputs - a comparison of methods." Mathematical Biosciences **102**(1): 41-73.

Chen, C. W., D. Z. Chen, et al. (2002). "An improved differential evolution algorithm in training and encoding prior knowledge into feedforward networks with application in chemistry." Chemometrics and Intelligent Laboratory Systems **64**(1): 27-43.

Cobelli, C. and J. J. DiStefano, 3rd (1980). "Parameter and structural identifiability concepts and ambiguities: a critical review and analysis." Am J Physiol **239**(1): R7-24.

Cobelli, C., A. Ruggeri, et al. (1985). "Optimal design of multioutput sampling schedules--software and applications to endocrine-metabolic and pharmacokinetic models." IEEE Trans Biomed Eng **32**(4): 249-56.

D'Argenio, D. Z. and A. Schumitzky (1979). "A program package for simulation and parameter estimation in pharmacokinetic systems." Comput Programs Biomed **9**(2): 115-34.

Distefano, J. J., 3rd (1982). "Algorithms, software and sequential optimal sampling schedule designs for pharmacokinetic and physiological experiments." Math Comput in Simulation **24**: 531-534.

Distefano, J. J., III, C. Cobelli (1980). "On parameter and structural identifiability: nonunique observability/reconstructibility for identifiable systems, other ambiguities and new definitions." IEEE Transactions on Automatic Control **AC-25**(4): 830-3.

Evans, N. D., M. J. Chappell (2000). "Extensions to a procedure for generating locally identifiable reparameterisations of unidentifiable systems." Math Biosci **168**(2): 137-59.

Fedorov, V. V. (1972). Theory of optimal experiments. New York,, Academic Press.

Kuhn de Chizelle, A. N. and J. J. DiStefano, III (1994). "MAMCAT: an expert system for distinguishing between mammillary and catenary compartmental models." Computers in Biology & Medicine **24**(3): 189-204.

Landaw, E. M. (1982). "Optimal multicompartmental sampling designs for parameter estimation: practical aspects of the identification problem." Math Comput in Simulation **24**: 525-530.
Landaw, E. M. and J. J. DiStefano, 3rd (1984). "Multiexponential, multicompartmental, and noncompartmental modeling. II. Data analysis and statistical considerations." Am J Physiol **246**(5 Pt 2): R665-77.
Le Novere, N. and T. S. Shimizu (2001). "STOCHSIM: modelling of stochastic biomolecular processes." Bioinformatics **17**(6): 575-6.
Marino, A. T., J. J. Distefano, 3rd, et al. (1992). "DIMSUM: an expert system for multiexponential model discrimination." Am J Physiol **262**(4 Pt 1): E546-56.
Mendes, P. (1993). "GEPASI: A software package for modelling the dynamics, steady states and control of biochemical and other systems." Comput. Applic. Biosci. **9**: 563-571.
Michalski R, W. Rode, A. Les (1998). "DerivFit: A program for rate equation parameter fitting using derivatives." Computers and Biomedical Research **31**:71-89
Nash, J.D. and T.K. Quon TK (1996). "Software for modeling kinetic phenomena." The American Statistician **50**:368-378.
Pohjanpalo, H. (1978). "System identifiability based on the power series expansion of the solution." Mathematical Biosciences **41**: 21-33.
Pohjanpalo, H. and B. Wahlstrom (1982). "Software for solving identification and identifiability problems, e.g. in compartmental systems." Math. Comp. in Simulation **24**(6): 490-493.
Saccomani, M. P., S. Audoly, et al. (2003). "Parameter identifiability of nonlinear systems: the role of initial conditions." Automatica **39**(4): 619-632.
Sedaghat, A.R., A. Sherman and M.J. Quon (2002). "A mathematical model of metabolic insulin signaling pathways." Am. J. Physiol. Endocrinol Metab. 283:E1084-E1101.
Tomita, M., K. Hashimoto, et al. (1997). "E-CELL: Software environment for whole cell simulation." Genome Inform Ser Workshop Genome Inform **8**: 147-155.
Toth, J and J-P Rospars (2005). "Dynamic modeling of biochemical reactions with application to signal transduction: principles and tools using *Mathematica*." BioSystems **79**:33-52.
Vicini, P.(2001). "Sparse data and complex models: dealing with nonidentifiability." Resource Facility for Population Kinetics Technical Report TR 2001-1, Seattle, U. Washington. *http://depts.washington.edu/rfpk/rd/index.html.*
Vicini, P., H. T. Su, et al. (2000). "Identifiability and interval identifiability of mammillary and catenary compartmental models with some known rate constants." Mathematical Biosciences **167**(2): 145-161.
Voit, E.O. (2000). Computational Analysis of Biochemical Systems: A Practical Guide for Biochemists and Molecular Biologists. Cambridge UK, Cambridge University Press. (includes PLAS software by A.E.N. Ferreira)
Walter, E. (1982). Identifiability of State Space Models. Berlin, Springer-Verlag.
Walter, E. and Y. Lecourtier (1982). "Global approaches to identifiability testing for linear and nonlinear state-space models." Math. and Computers in Simulation **24**(6): 472-482.
Wang, F. S., T. L. Su, et al. (2001). "Hybrid differential evolution for problems of kinetic parameter estimation and dynamic optimization of an ethanol fermentation process." Industrial & Engineering Chemistry Research **40**(13): 2876-2885.
Zak, D.E., Gonye, G.E., et al. (2003). "Importance of input perturbations and stochastic gene expression in the reverse engineering of genetic regulatory networks: Insights from an identifiability analysis of an *in silico* network. Genome Res. **13**:2396-2405.