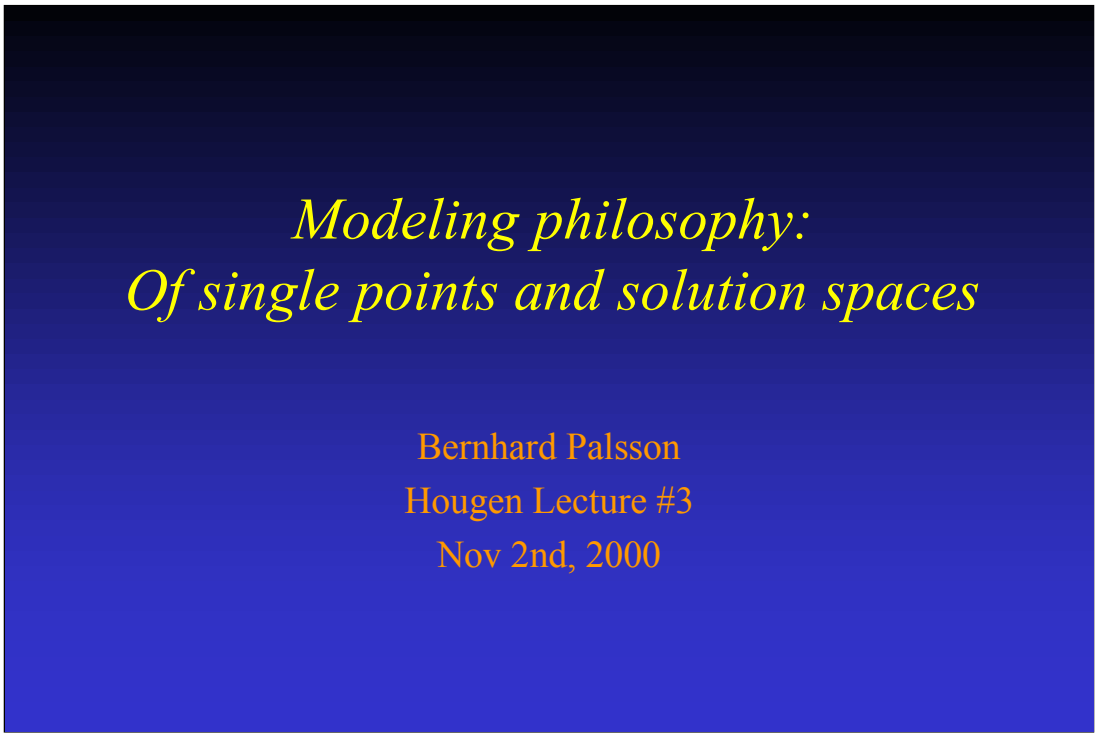
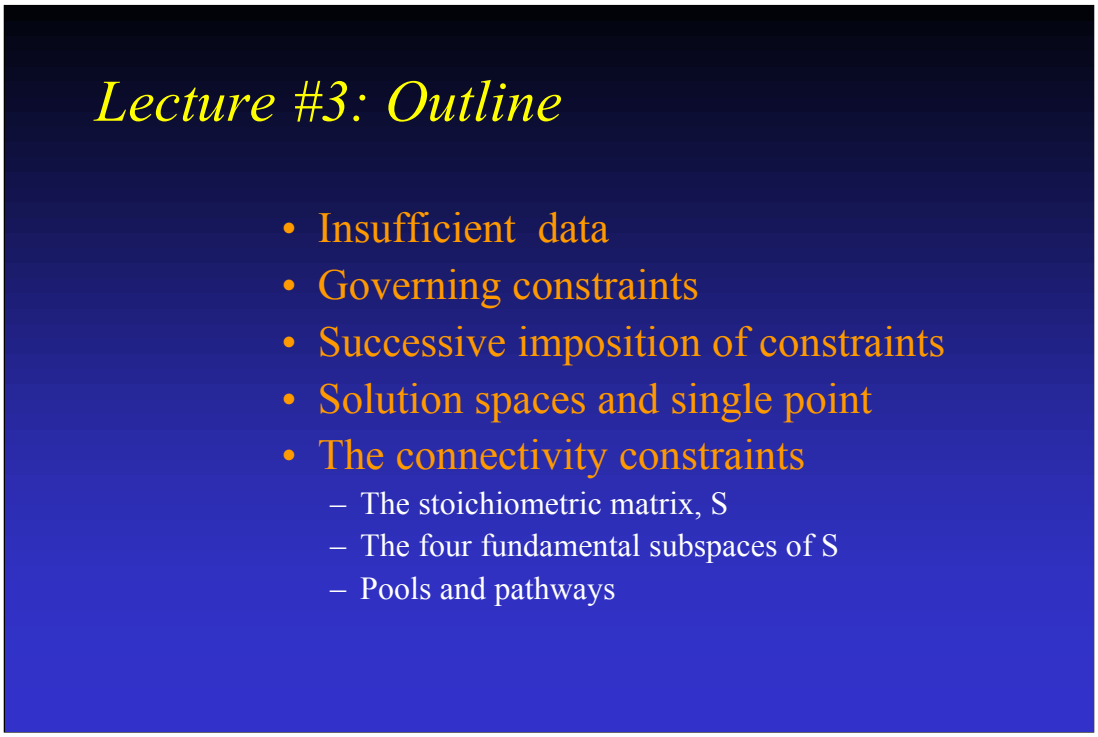


LECTURE #3

The first two lectures discussed the high-throughput technologies and the subsequent determination of cellular component catalogs and reconstruction of biochemical reaction networks. In the third lecture we begin to discuss how one describes the function of such networks in systemic and mathematical terms.



OUTLINE

In spite of the impressive amounts of data that are being generated about cells and their components we do not have all the data that is needed to construct detailed mathematical models of their integrated function. The approaches often used in physicochemical and engineering sciences of stating governing fundamental laws and building detailed mathematical models will thus not work, at least not initially, for the construction of mathematical models of reconstructed biochemical reaction networks. We cannot thus calculate a single solution.

An alternative approach must be developed. The network maps allow us to impose systemic and component constraints on the function of the network as a whole. Thus we can eliminate behaviors but we cannot calculate precise ones. The more governing constraints that we can state the smaller the solution spaces become.

The lecture then ends with a detailed discussion of the consequences of the stoichiometric constraints.

Coping with incomplete constraints: solution spaces vs. single points

--Cannot describe cellular networks in the same detail as we are used to in the P/C sciences

--However, we can subject the networks to known constraints and analyze them given these constraints

3 Problems

•P/C laws may not apply

•Kinetic constraints not known

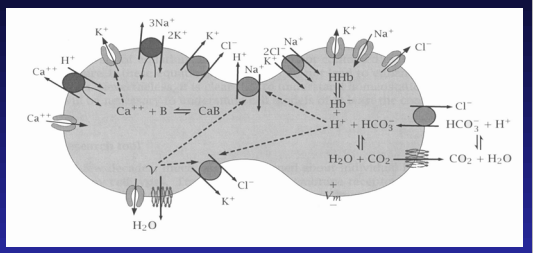
•Even if they were, they

1) Change with time-->evolution

2) Not the same from one individual to the next-->SNPs

Factors Constraining Metabolic Function

- Connectivity:
  - Systemic stoichiometry
- Capacity:
  - Maximum fluxes
- P/C factors:
  - osmotic pressure, electro-neutrality, solvent capacity, molecular diffusion
- Rates:
  - Mass action, Enzyme kinetics, Regulation



CONSTRAINTS

Metabolism is subject to a number of constraints. First, fluxes are balanced in the steady state. For many dynamic metabolic states, the solution does not move far from the steady state. There is an upper limit on the amount of flux that is achievable through every reaction. First there is an upper P/C constraint, a crowding constraint limiting the amount of enzyme present and finally upper limits may be derived from expression and proteomic profiles.

There are a number of physico-chemical constraints that a cell must operate under. These include balancing of osmotic pressure (unless there is a cell wall), maintaining electro-neutrality since charges cannot be separated, the limited solvent capacity of water (i.e. the 30% of cells that is biomass must be divided amongst the thousands of cellular constituents) , and the rate of molecular diffusion limits almost all cellular functions.

Finally, the kinetic parameters that have evolved and the imposed regulatory mechanisms significantly influence the flexibility of the network. These are flexible and the cell can adjust them.

The connectivity and P/C constraints are 'hard' in the sense that the cell cannot manipulate them, the capacity constraints represents fixed upper limit constraints that can be down regulated, while the kinetics may be quite flexible and adjustable by the cell through a evolutionary process.

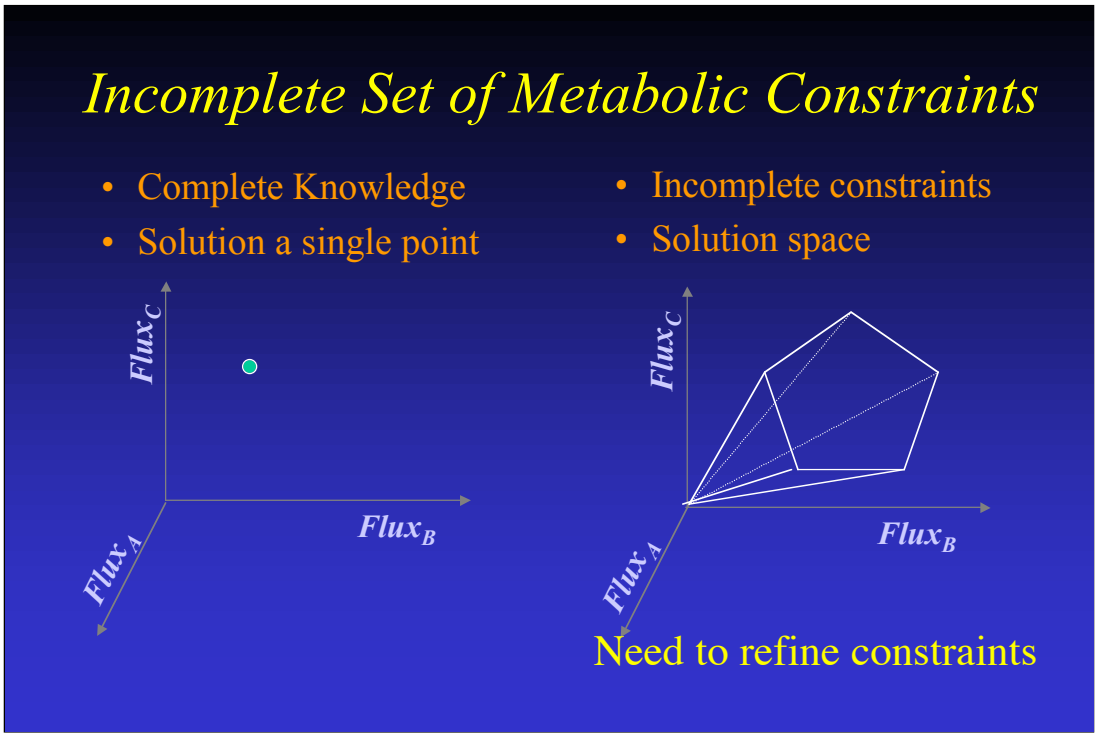
Factors Constraining Metabolic Function

- Connectivity:
  - Systemic stoichiometry
  - Sv=0
- Capacity:
  - Maximum fluxes
  - v_i < maximum value

HARD CONSTRAINTS

In these lectures we will impose two sets of constraints to study possible metabolic functions.

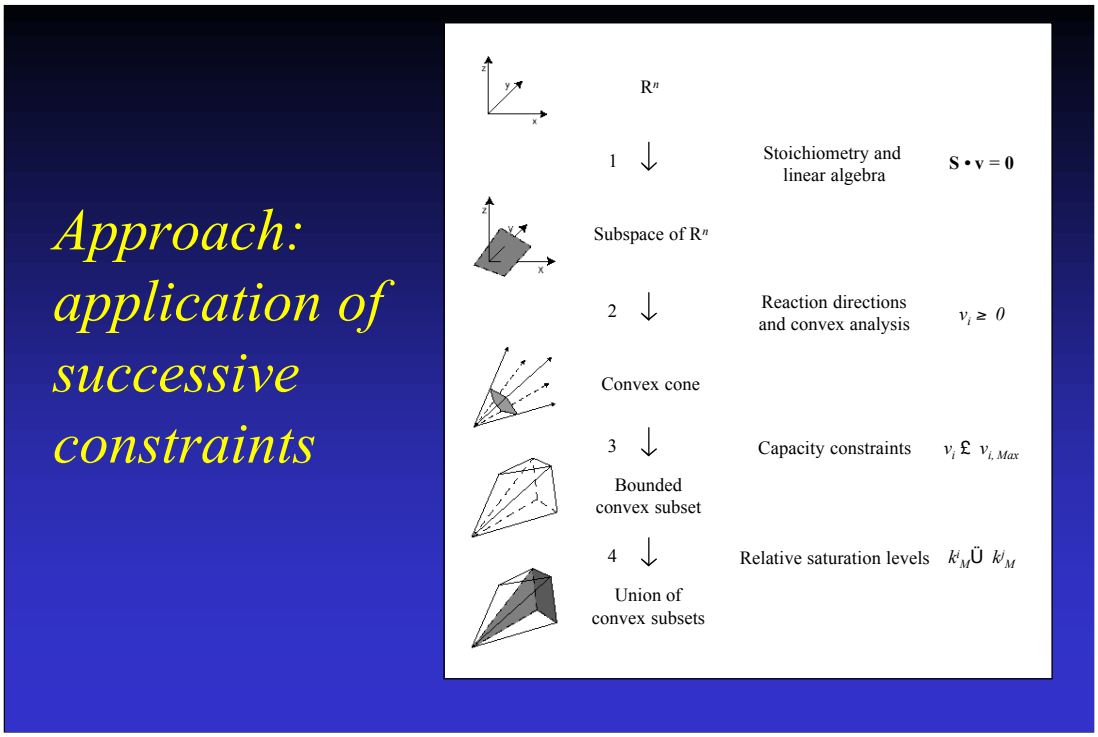
These are the connectivity constraints and the capacity constraints. In Lecture #3 we will cover the consequences of the imposition of the stoichiometric constraints.



CONSEQUENES OF OUR MODELING PHILOSPHY

Normally when we solve a mathematical problem or construct a mathematical model we are looking for 'the solution.' The search for such a solution comes down to a detailed and complete problem statement, and then the use of mathematical or numerical methods to find 'the' solution. It is represented as a single point in the left side of the figure.

However, we are in a situation where we cannot fully define and describe the interior of a cell in all its details. We thus must be content with 'bracketing' the solution. The imposition of governing constraints then eliminates impossible solutions but leaves a range of possible solutions. This range is represented by a solution space that contains all these possible solutions. The more applicable constraints that we find the smaller the solution space.

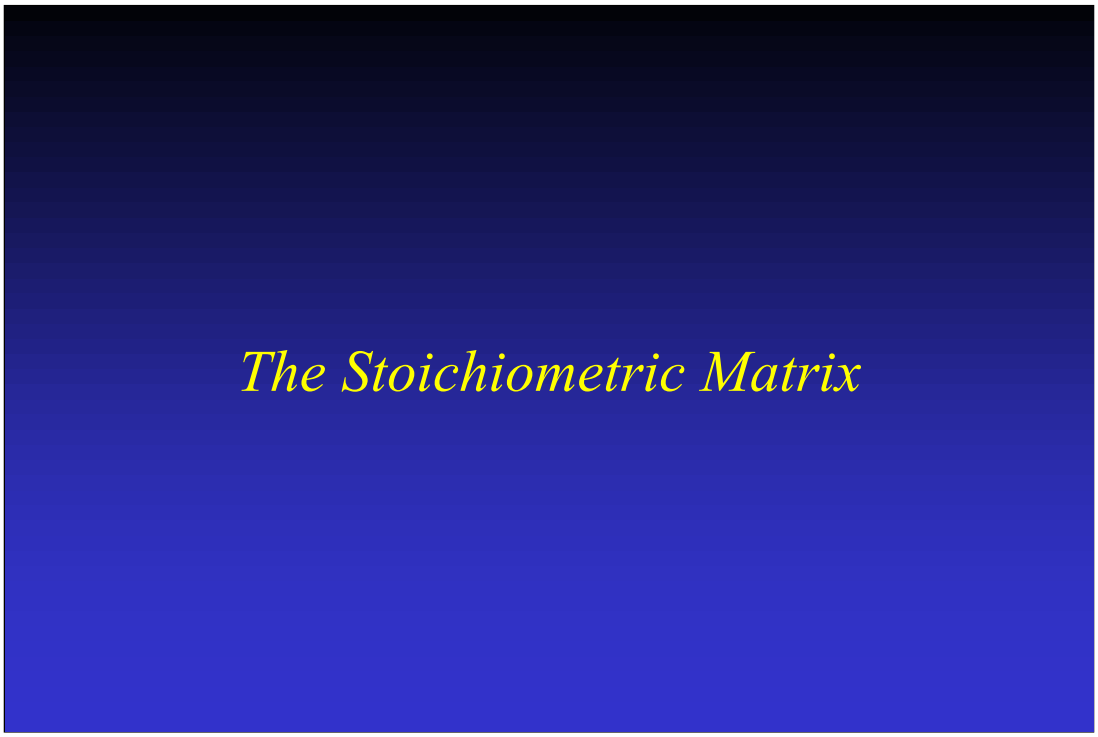


GRAPHICAL ILLUSTRATION OF THE SUCCESSIVE APPLICATION OF GOVERNING CONSTRAINTS

Some years ago it was common to think of each gene/gene product in a cell as an independent element. Genetic engineering came into being and the expectation was that if one would splice a gene into a genome a trait that corresponding to that gene product would be produced. Mathematically, one can represent this as an n-dimensional space (where n is the number of gene products) and any point in this space could be attained.

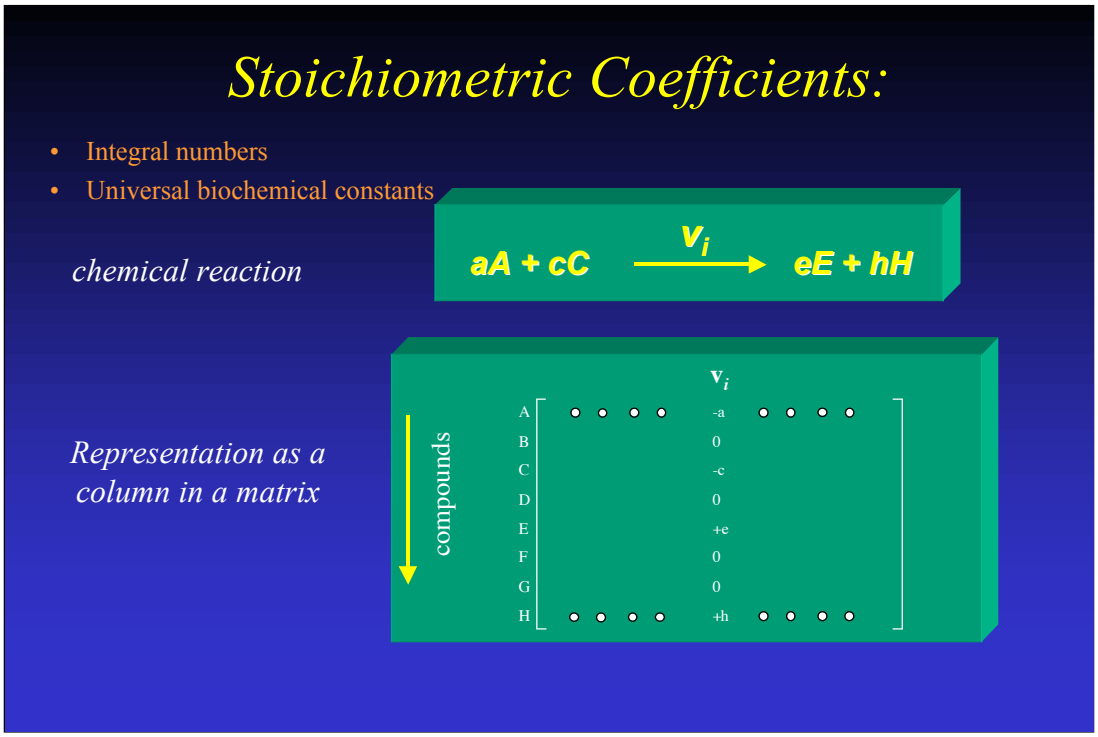
However, every gene product works in the context of many others and is thus constrained in its activity. For instance once can over express an enzyme in a linear pathway and get no increase in flux down the pathway since the flux through all the steps has to be the same. Such connectivity, or stoichiometric constraints reduce the accessible space to a subspace, or a 'hyper-plane' as illustrated. The 'size' of this hyper plane is substantially smaller than the n-dimensional space. Thus these constraints limit the attainable behaviors.

A hyper-plane is infinite in all directions. If we consider all reactions to have positive fluxes (so reversible reactions are represented as two irreversible reactions) the hyper-plane is converted to a semi-finite conical solution space. If we then impose the maximum flux constraints then the solution space is 'capped off' and becomes a 'lock-box' for the solution. This lock box is formed based on hard constraints. Certain kinetic constraints drive the solution to the edge as shown later. These represent adjustable constraints.



THE MATRIX S

For the rest of this lecture we shall discuss the consequences of the connectivity constraints in metabolism, namely stoichiometry

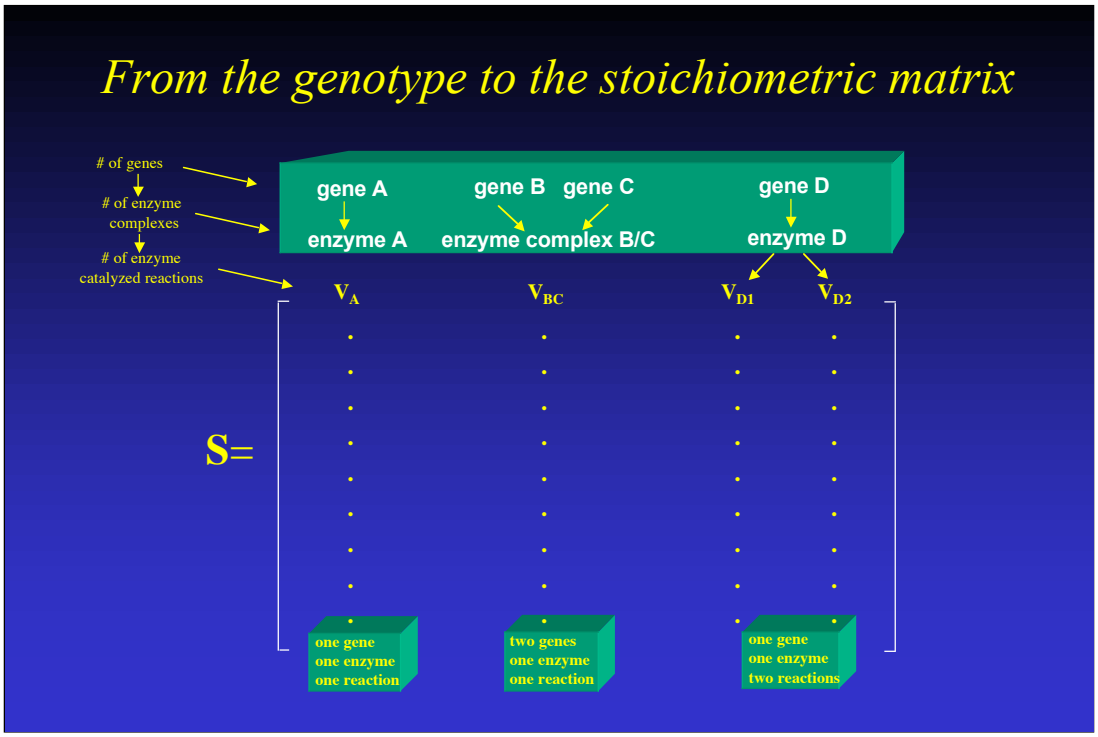


EACH COLUMN IN THE STOICHIOMETRIC MATRIX CORRESPONDS TO A PARTICUALR METABOLIC BIOCHEMICAL REACTION

The stoichiometric coefficients: They are integers (a,c,e,h in the example given) that represent the number of molecules of chemical species (A,C,E,H in the examples) that are transformed in this particular chemical reaction. These coefficients are constants (i.e. are not condition dependent, that is functions of temperature, pressure, pH, etc). Further they are biologically universal, that is the same metabolic reaction proceeds the same way in all cells; for instance hexokinase always catalyzes the reaction:

Glucose + ATP --> Glucose-6-phosphate + ADP

Formation of a column in S: Each metabolite has a row in the stoichiometric matrix, and each reaction has a column. The stoichiometric coefficients are used to form a column, with the stoichiometric coefficient that corresponds to a particular metabolite appearing in the row that it corresponds to. If a metabolite is formed by the reaction the coefficient has a positive sign, if it is consumed by the reaction the stoichiometric coefficient appears with a negative sign. All other rows (corresponding to metabolites that do not participate in the reaction) are zero.



THE NUMBER OF REACTIONS IN A METABOLIC GENOTYPE IS NOT THE SAME AS THE NUMBER OF GENES IN THE GENOTYPE

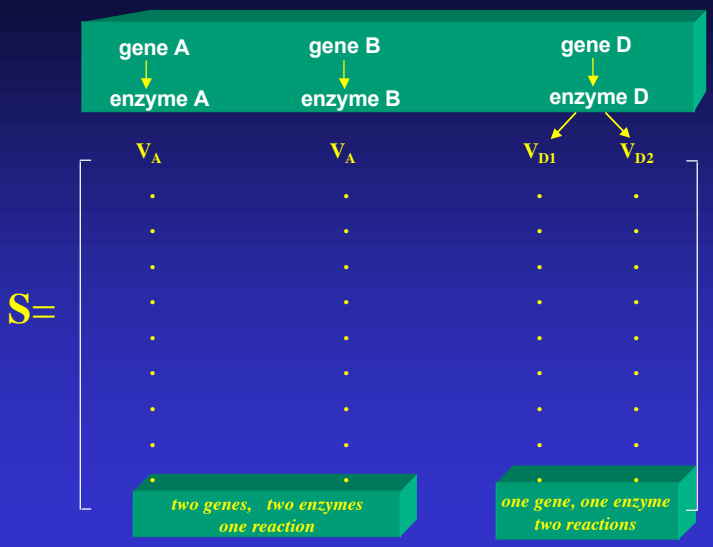
There is not a one-to-one correspondence between the number of genes that are associated with metabolism and the number of chemical transformations that take place. This difference is due to several factors.

First, many enzymes are oligomeric complexes that contain more than one protein chain. These complexes are formed by non-stoichiometric binding, or association of several different protein molecules. Hemoglobin, being a tetramer of two alpha and two beta globins is perhaps the best know example of a protein oligomer.

Second, enzymes can catalyze more than one chemical reaction. This feature is often referred to as substrate promiscuity. These chemical transformations tend to be similar.

These features give rise to a different number of genes from the number of enzymes (or enzyme complexes) and the number of chemical reactions that take place. All of these situations can though be accounted for with the stoichiometric matrix as illustrated.

Redundancy and pleiotrophy in the stoichiometric matrix



Partitioning of the flux vector into internal and external fluxes

- External fluxes are those fluxes that flow across the cellular boundary.
  - These are denoted by b_i. These fluxes are often accessible to measurement or can be estimated based on experimental data. The sign convention adopted for these fluxes is that they are positive if mass is flowing out of the cell.
- Internal fluxes are those that take place with in the cell (within our system boundary).
  - These fluxes are hard to measure, but often we will know their maximum value.

PARTITIONING THE FLUX VECTOR

We draw a systems boundary around the metabolic system that we are interested in. Thus there will be reactions that take place within the system and those that exchange molecules with the surroundings. We partition the flux vector accordingly.

Normally the system boundary is drawn such that the metabolic system being considered is the entire metabolic system in a cell. Then the system boundary effectively becomes the cell membrane. In other cases we may be interested in an organelle, such as the mitochondrion, and we will draw our system boundary around it. In yet other cases we draw system boundaries around certain sectors of metabolism, such as the fueling reactions, or the amino acid synthetic pathways. In such cases the system boundary is a conceptual one and not a physical one.

The concept of a 'system boundary' is frequently used in the physical and engineering sciences, while for life scientists reading these notes may be a new one. It may take some getting used to.

Chemical Reactions vs. Fluxes through them

- The columns of the stoichiometric matrix represent the reactions (n in number)
- The actual reaction rates, or the fluxes that take place through these reaction are denoted by a **v**
- The assignment of a flux through a reaction can be performed by a simple matrix multiplication

$$\mathbf{S} \cdot \mathbf{v} \qquad \begin{vmatrix} S_{11} & \cdots & S_{1n} \\ \vdots & & \vdots \\ S_{m1} & \cdots & S_{mn} \end{vmatrix} \begin{vmatrix} v_1 \\ v_2 \\ \vdots \\ v_{n_i} \\ b_j \\ \vdots \\ b_n \end{vmatrix}$$

METABOLIC REACTIONS AND THE FLUXES THROUGH THEM

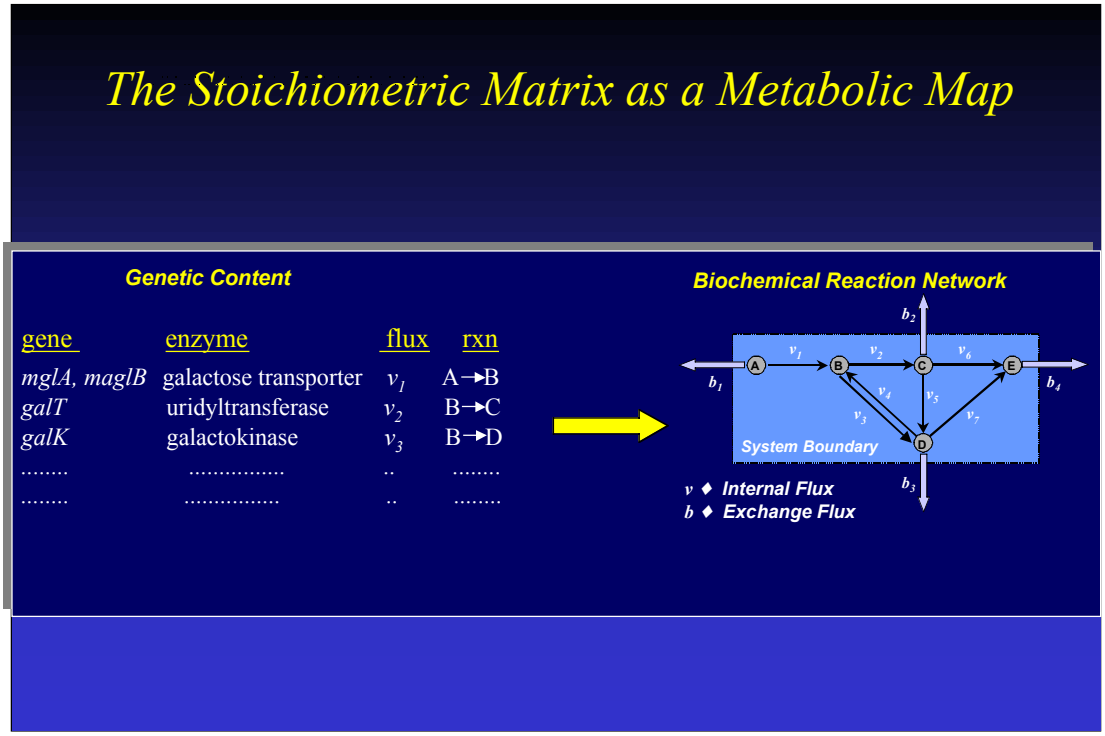
The annotated sequence and biochemical knowledge of the metabolic enzymes lead to the definition of the stoichiometric matrix. Each column in this matrix represents a particular metabolic reaction. However, the flux though a reaction is highly dependent on what the cell is doing. For instance, if an amino acid is available to the cell, it will import it and not synthesize it. Although the cell is capable of carrying out all the reactions that lead to the synthesis of the amino acid they are not used. The flux through them is zero. Later we will see how the cell regulates flux (either by kinetic means or by regulation of gene expression), but for now we introduce the product of the stoichiometric matrix and the flux vector**.** The matrix is a constant, while the flux vector is a variable.

The number of genes, enzymes and metabolic reactions for some gram-negative bacteria

	E. coli	*H. influ.*	*H. pylori*	*Yeast*
Total # of Genes	4288	1743	1590	6259
# of metabolic genes	660	400	290	697
# of metabolic enzymes	697	412	272	626
# of metabolic reactions	739	461	381	1212
# of metabolites	442	367	332	569

ACTUAL NUMBERS FOR ACTUAL ORGANISMS

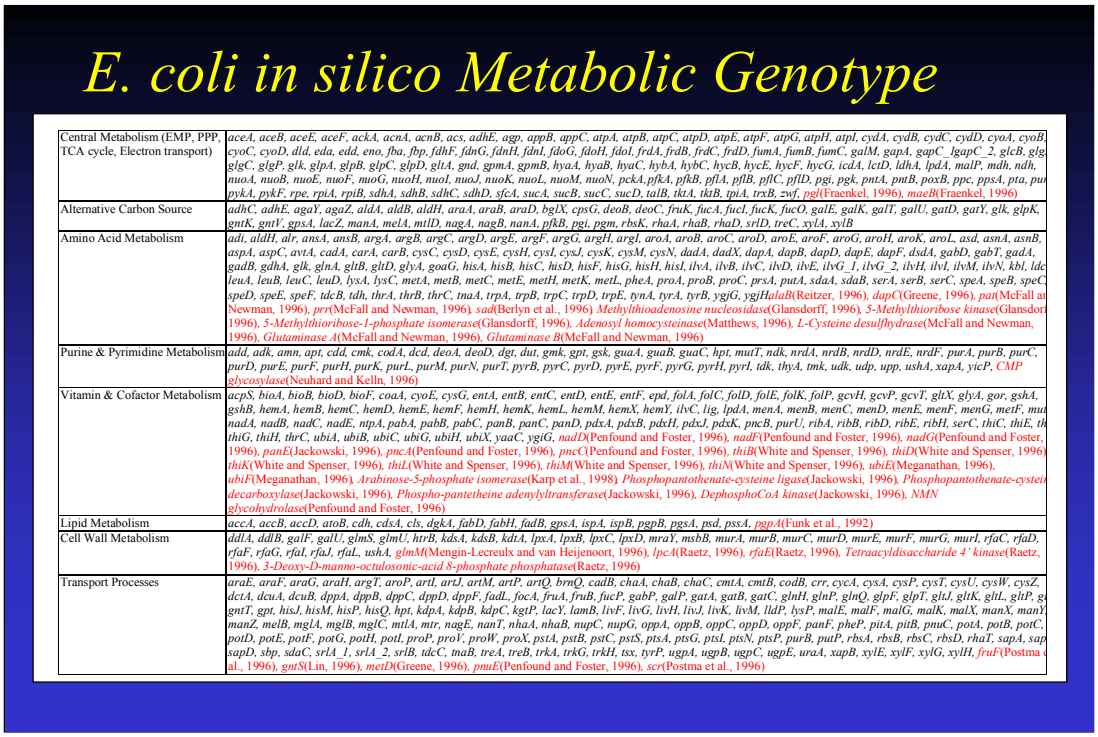
Several in silico genome-scale metabolic maps have been reconstructed. This slide shows actual numbers for three gram-negative bacteria. *E. coli* is a free living organism that can live off of several different individual carbon sources. *E. coli* has thus been called the 'complete' organic chemist as it can synthesize all the chemical structures that it needs for its biomass synthesis. In sharp contrast, *H. influenzae* and *H. pylori* are human pathogens that require several different substrates to grow.



TRANSLATION OF THE STOICHIOMETRIC MATRIX INTO A METABOLIC MAP

Thus we can translate the genomics and biochemistry of a metabolic reaction network directly into the realm of linear algebra in the form of a stoichiometric matrix. Beginning with the gene products of a system we can determine the inter-conversions of metabolites which occur and represent this in the form of a stoichiometric matrix to complete the translation. Within the stoichiometric matrix lies all of the structural information and the architecture of the network. The word structure here is not used to denote the physical structure but the structure of a network.

The stoichiometric matrix is a connectivity matrix that ties all the metabolites, the 'nodes,' in the network together, where the 'edges' or 'connections,' are the metabolic reactions. The stoichiometric matrix is thus a compact mathematical representation of a metabolic map. These maps give us a visual, and easier to understand, representation of the metabolic network in a cell.

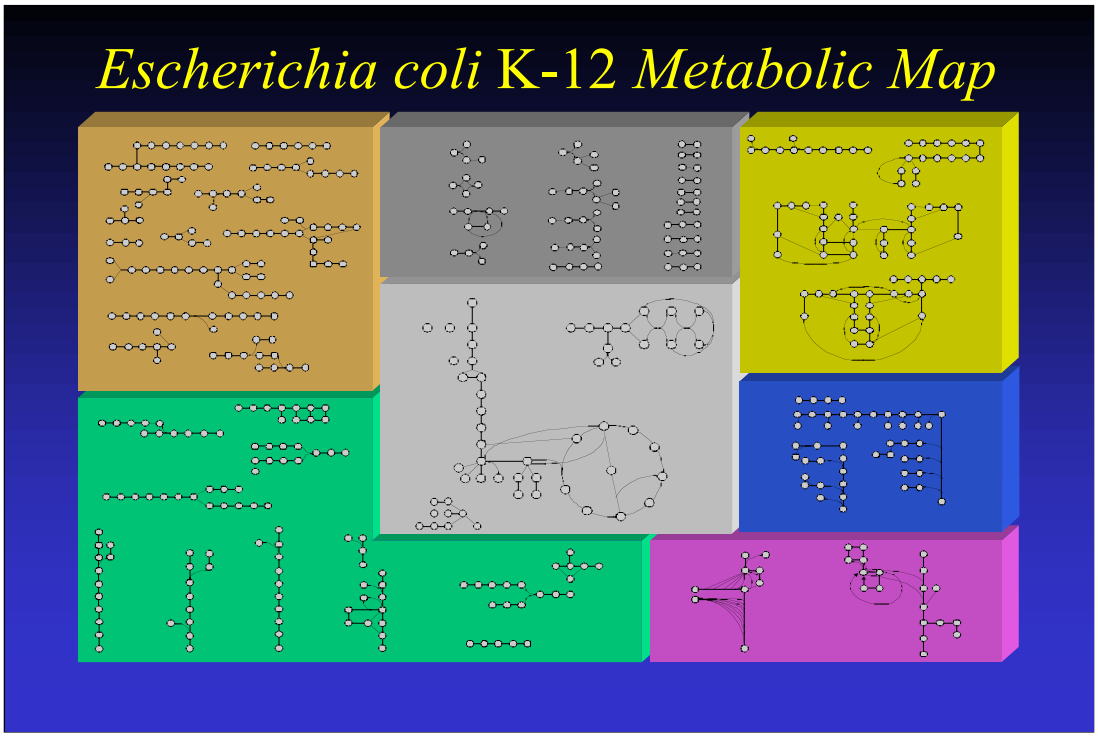


THE E. COLI METABOLIC GENOTYPE

The list of metabolic genes that are found on the E. coli K-12 genome is shown. These genes are categorized into different classes depending on the part of metabolism in which they participate.

The lowest group represents the transporters, or the **b** fluxes.

The enzymes indicated in red represent enzymatic activities that have been found to be in E. coli by biochemical means but the corresponding gene, or ORF, has not been located on the genome.

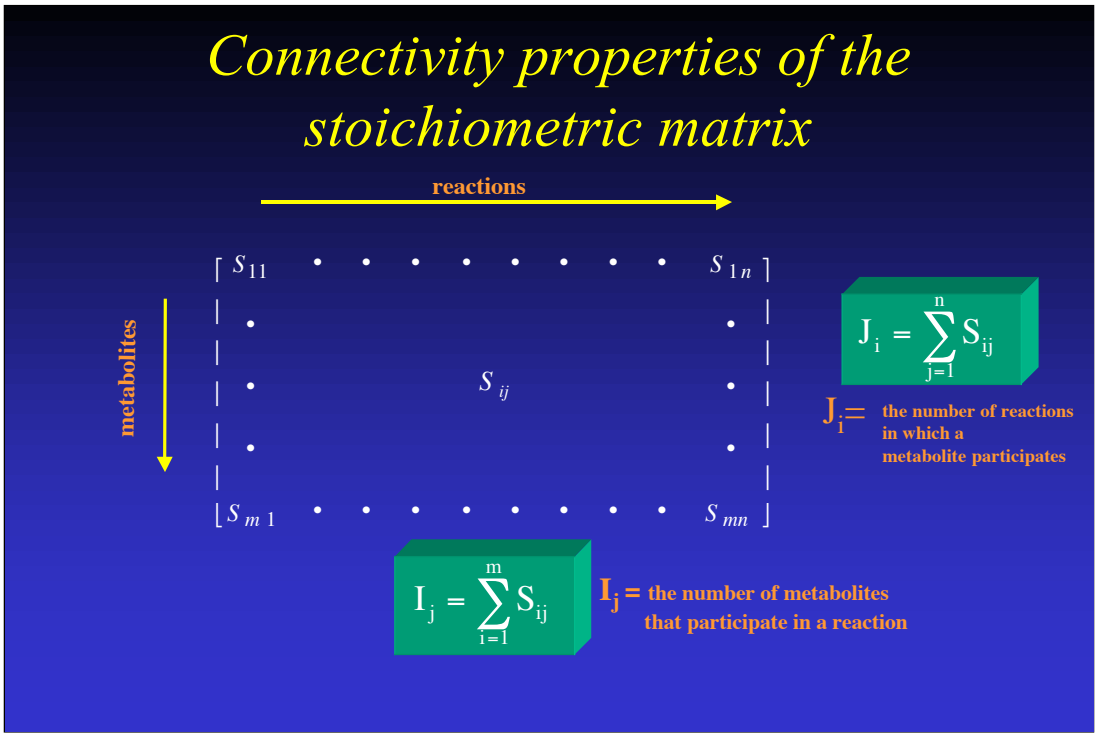


THE METABOLIC MAP REPRESENTATION OF THE *ESCHERICHIA COLI* K-12 METABOLIC GENOTYPE

The metabolic map of the *E. coli* K-12 metabolic genotype divided into metabolic sectors based on a biochemical rationale:

Gray:	Alternative carbon source metabolism
Light gray:	The core metabolic pathways
Orange:	Amino acid biosynthesis
Green:	Vitamin and co-factor metabolism
Yellow:	Nucleotide synthesis
Blue:	Cell wall synthesis
Purple:	Fatty acid synthesis

Not all the 720 reactions are shown. Highly connected metabolites, such as ATP, PEP and pyruvate are likened to dozens of reactions. Showing all of these connections would make this representation visually unattractive. However, these connections should not be overlooked as they play a key role in the stoichiometric characteristics of metabolism.

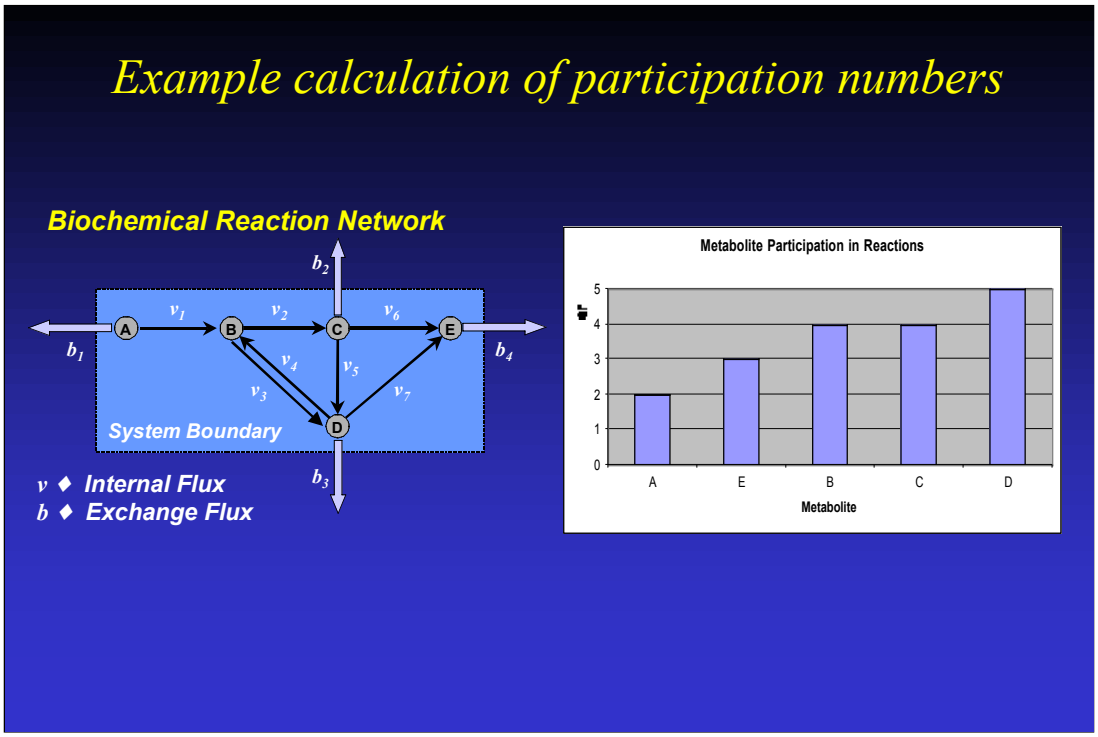


SOME CONNECTIVITY PROPERTIES OF THE STOICHIOMETRIC MATRIX

As illustrated above the stoichiometric matrix is a connectivity matrix that connects all the metabolites in a defined metabolic system. We now introduce some of its connectivity properties:

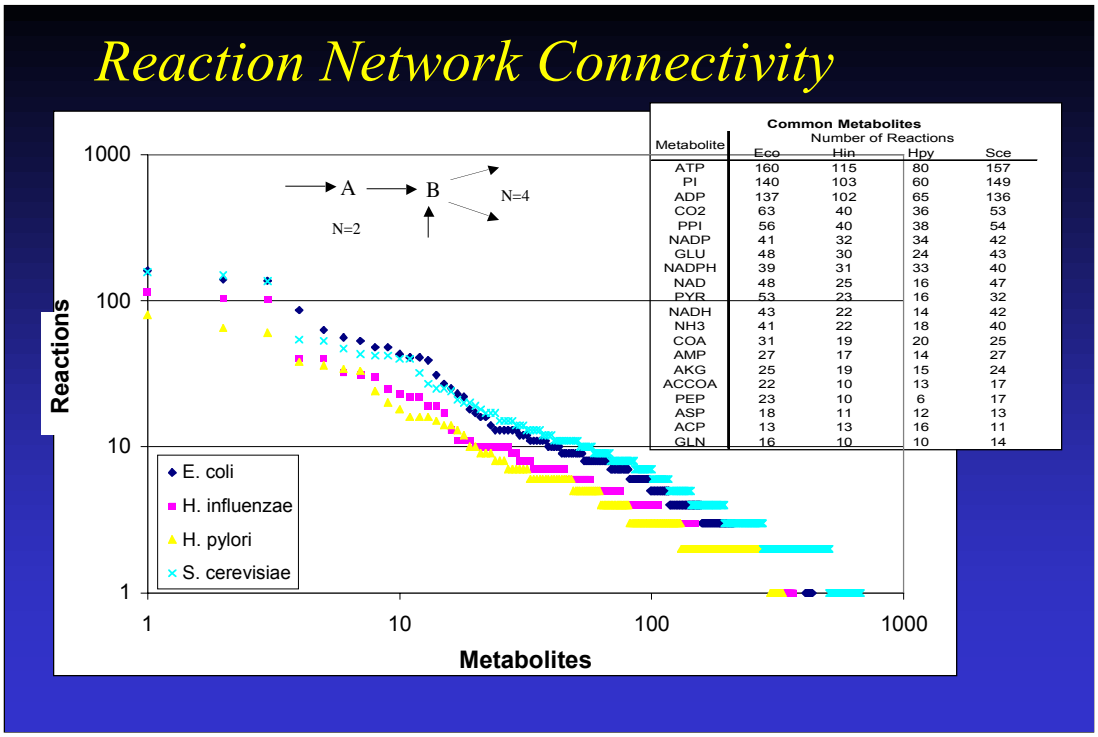
1. The participation number. Metabolites can participate in several metabolic reactions. The number of metabolic reactions that a metabolite participates in can be obtained by simply summing up the number of non- zero elements in the row that corresponds to the metabolite. Note that all internal metabolites must have a participation number of two or more. If not there is a dead end in the network. This feature can be used to curate and diagnose genome annotation, as being either incomplete or erroneous. External metabolites typically will have only a single reaction associated with them, namely membrane transport.

2. The number of molecules participating in a particular metabolic reaction can be obtained by simply summing up the absolute value of all the stoichiometric coefficients that appear in a column. The most frequent number is 4.



CALCULATION OF PARTICIPATION NUMBERS

This slide shows a calculation of the participation number for the simple reaction schema that we have been using. D is the most highly connected metabolite participating in five reactions, while A is the least, participating in the minimum number of two reactions.

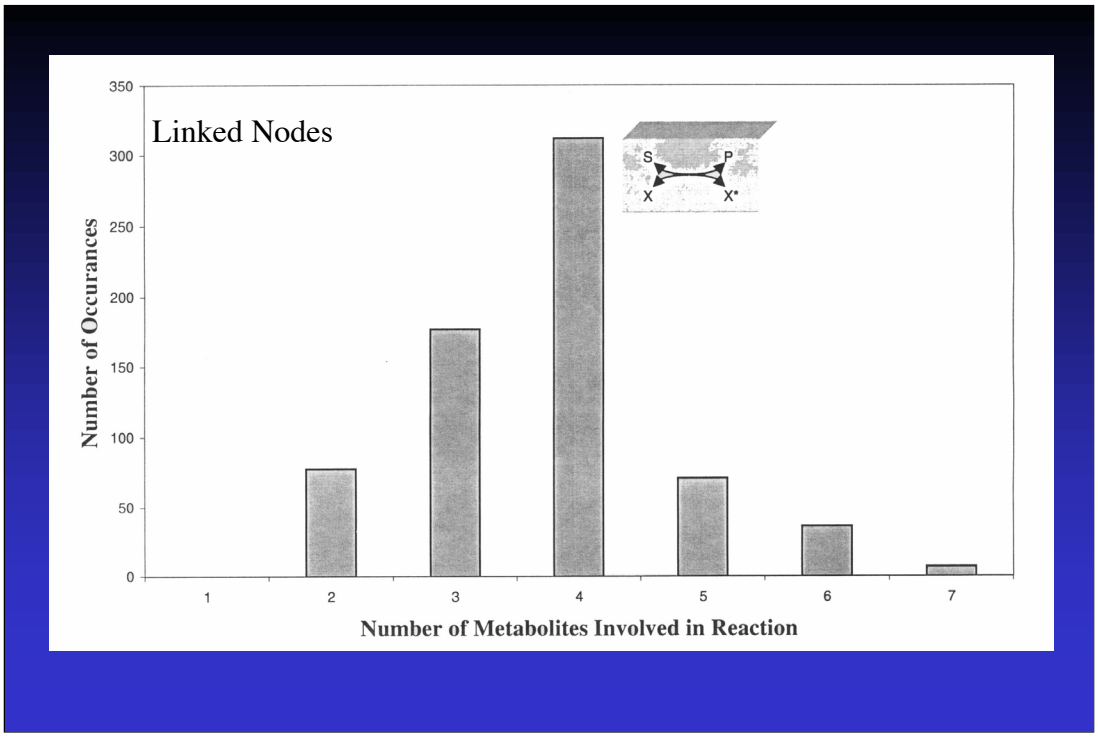


THE PARTICIPATION NUMBERS FOR FOUR METABOLIC MAPS

The 436 metabolites in the *E. coli* K-12 metabolic genotypes all have a participation number associated with them. Here we have calculated them all and rank ordered the metabolites according to the number of reactions that they participate in. This data must be plotted on a log-log scale in order to see the entire range of participation numbers.

ATP is the most highly connected metabolite in *E. coli* K-12. It participates in 161 of the 720 reactions, about one in five reactions. Similarly, ADP and P_i participate in a similar number of reactions. Thus high-energy phosphate metabolism tightly connects the entire metabolic network. Glutamine, the central metabolite of nitrogen metabolism participates in 40+ reactions, and PEP and pyruvate are also highly connected at 25 and 55 respectively. The redox carriers participate in a few dozen reactions. It is therefore not surprising that metabolic regulation must be focused on maintaining the concentrations of these metabolites within a narrow range. Otherwise the entire system would be influenced.

The majority of the metabolites, 198 of the 426, participate in only two reactions, one that forms them and one that degrades them.



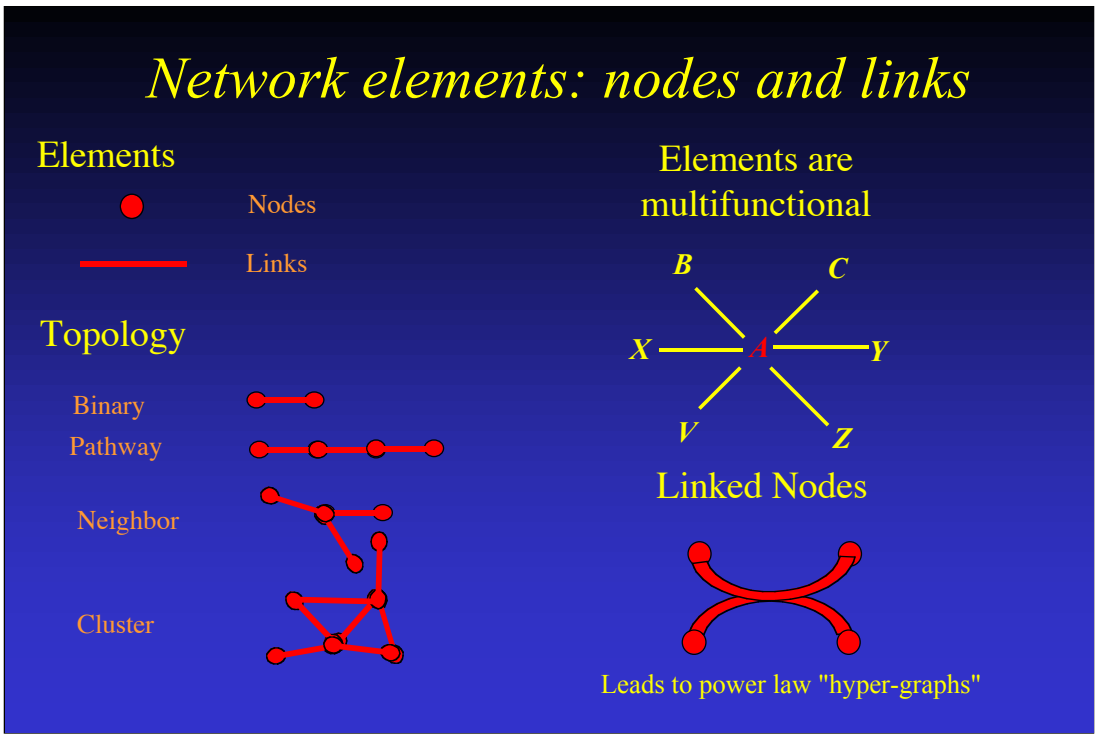
THE NUMBER OF MOLECULES THAT PARTICIPATE IN THE REACTIONS IN THE *ESCHERISCHIA COLI* K-12 METABOLIC GENOTYPE

This histogram shows the number of reactions in E. coli that have 2,3,4, etc molecules participating in the reaction. The most common reaction is of the form:

$$A\text{-}x + B \iff A + B\text{-}x$$

In other words an exchange of a moiety, group, or electrons among molecules. As we saw above, most commonly A-x would be ATP and A would be ADP, and the moiety x is a high energy phosphate group.

We shall see below that this feature has a significant influence on metabolic dynamics. Also this feature makes the map a power-law hyper-graph.



GRAPH THEORY

Much work will be needed to study the structural features of biochemical reaction networks. A few of the issues are illustrated here:

1. There are elements and links in networks. In metabolism, these two correspond to metabolites and the enzymatically catalyzed reactions between them

2. The topological features will be studied.

3. Each element in a network will have many function and potentially many types of links

4. In metabolism, there are linked nodes, i.e. one link will tie together more than two nodes (see previous slide). This changes the nature of the network substantially.

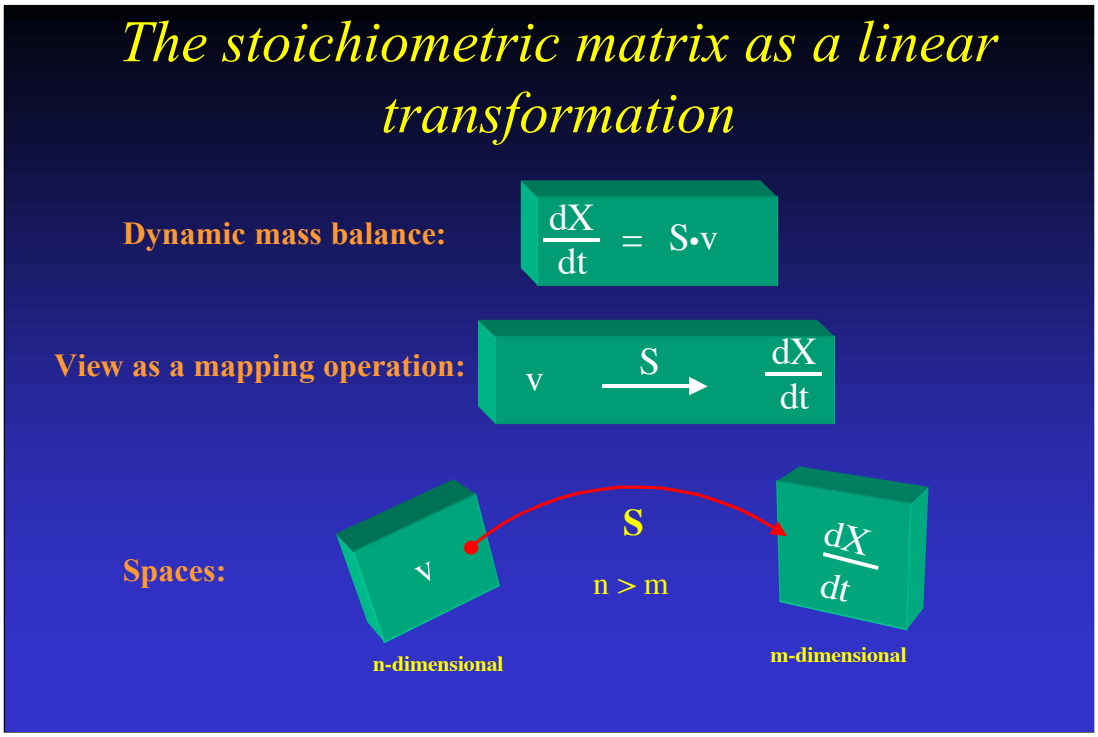
Dynamic Mass Balance: Matrix Form

$$\frac{d}{dt}\begin{bmatrix} X_1 \\ . \\ . \\ . \\ X_m \end{bmatrix} = \begin{bmatrix} S_{11} & . & . & . & . & . & . & S_{1n} \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ S_{m1} & . & . & . & . & . & . & S_{mn} \end{bmatrix} \begin{bmatrix} v_1 \\ . \\ . \\ . \\ v_n \end{bmatrix}$$

THE GENERAL DYNAMIC MASS BALANCE EQUATIONS

This slide shows the details of the general mass balance equations. The time derivatives of the metabolite concentrations (X) is the matrix multiplication of the stoichiometric matrix (S) and the flux vector.

Multiply one row times the vector to see how the summation of fluxes forms the RHS of the differential equation for that metabolite.



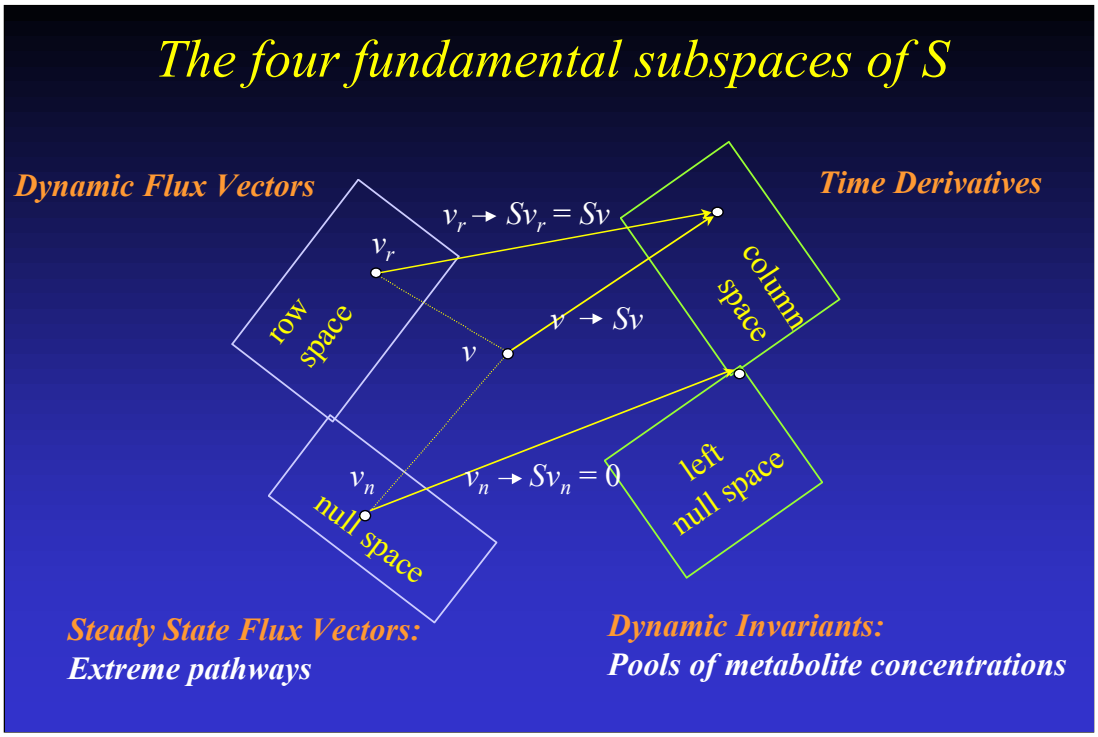
ANY MATRIX MAPS AN ELEMENT FROM ONE VECTOR SPACE INTO ANOTHER; THAT IS IT TRANSFORMS ONE VECTOR INTO ANOTHER

A matrix is a linear transformation;

$$y = A\,x,$$

simply is x mapped into y by the matrix A. The stoichiometric matrix maps the flux vector into the time derivatives. As noted, and as we will discuss in much more detail later, the flux vector is a function of the metabolite concentrations, denoted by the vector X in this slide.

The stoichiometric matrix 'S' thus takes the flux vector at any instant and calculates the time derivative of the concentrations; or how the system will move away from the point that it was located at. The next slide illustrates this point.



A SCHEMATIC DEPICTION OF THE ACTION OF A MATRIX AND THE FOUR SUBSPACES ASSOCIATED WITH IT

Every matrix can be thought of as a mapping operation or a linear transformation. It takes a vector in one space and transforms into a vector in another space. The four fundamental spaces are the row, column, null, and the left null space. These spaces are further described on the next slide.

The four fundamental subspaces of ***S***

- The null space **S**•v=0,
  - contains all the steady state solutions to the flux balance equations
- The column space of **S** (range);
  - contains the time derivatives resulting from the mapping
- The row space of **S**;
  - contains the dynamic flux vectors on which **S** operates
- The left null space of **S**;
  - contains all the dynamic invariants of **S**

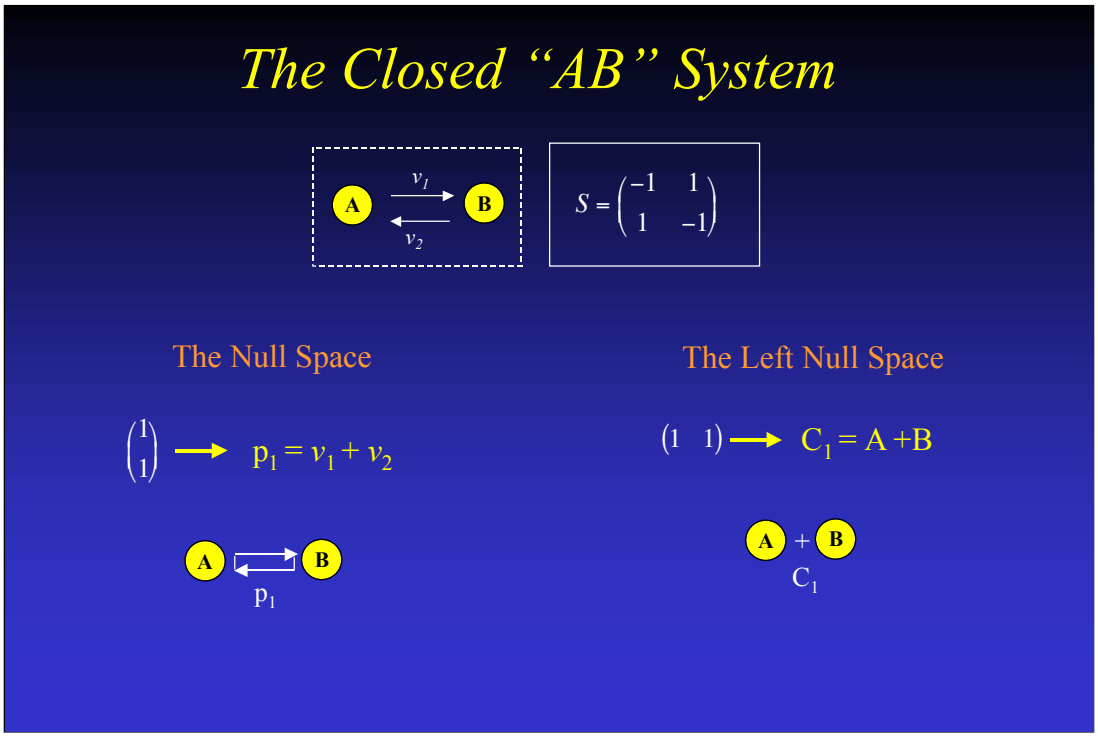
THE FOUR SUBSPACES OF THE STIOCIOMETRIC MATRIX

All the four fundamental subspaces of **S** will be of interest to us. The first spaces that we will study are the right and left null space of **S**, since it contains all the steady state solutions;

$$\mathbf{S}\mathbf{v} = 0$$

And the pooled variables

$$\Sigma_i \, (d\mathbf{X}_i/dt) = 0$$



THE SIMPLE 'AB' EXAMPLE:

Let's consider a reversible reaction. The stoichiometric matrix S is shown and it is rank deficient.

The addition of the two columns gives zero. This can be seen by multiplying the stoichiometric matrix with the column vector $(1,1)^t$. Thus this column vector spans the null space. This vector represents the pathway

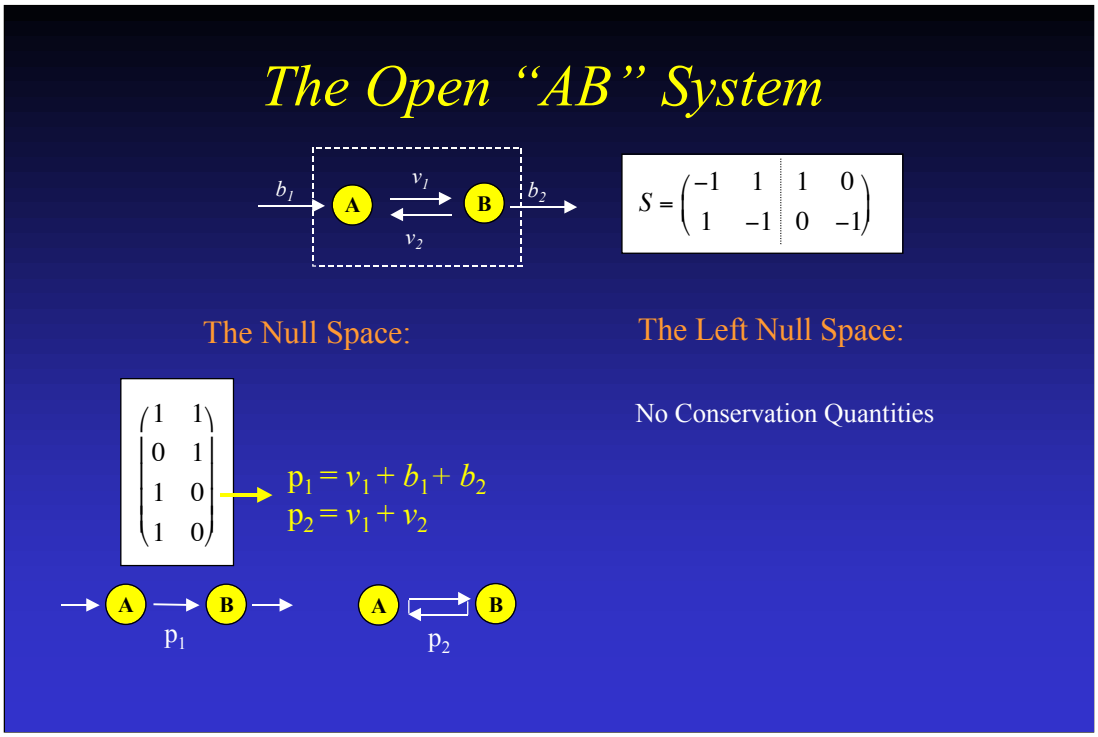
$$v_1+v_2$$

or the reversible back and forth reaction.

The addition of the rows gives a zero. This can be seen by multiplying from the left with the vector (1,1). Thus (1,1) spans the left null space and represents the summation of

$$A+B.$$

It is obvious in this case that this sum is time invariant.



THE OPEN 'AB' EXAMPLE

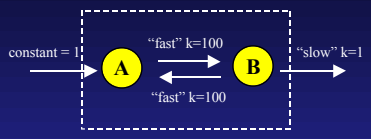
If we now add exchange fluxes the stoichiometric matrix for the closed system is 'appended' with the exchange reactions. The matrix no longer rank deficient. Thus the left null space is of zero dimension and there are no conserved quantities. The sum of A and B will vary with time depending on the exchange fluxes.

The null space is now two dimensional. It is spanned by two pathways. The same pathway as existed for the closed system, corresponding to the reversible reaction, is still there. Later we shall classify this pathway, as Type III.

There is a new pathway vector. It ties the input and the output via a straight pass through the system. Later we shall classify this pathway as Type I.

Any steady state flux distribution in this simple open 'AB' system is a linear combination of these two basis pathways.

Overlaying Order of Magnitude Kinetics



Rate Equations:

$v_1 = 100[A]$

$v_2 = 100[B]$

$b_1 = 1$ (constant)

$b_2 = [B]$

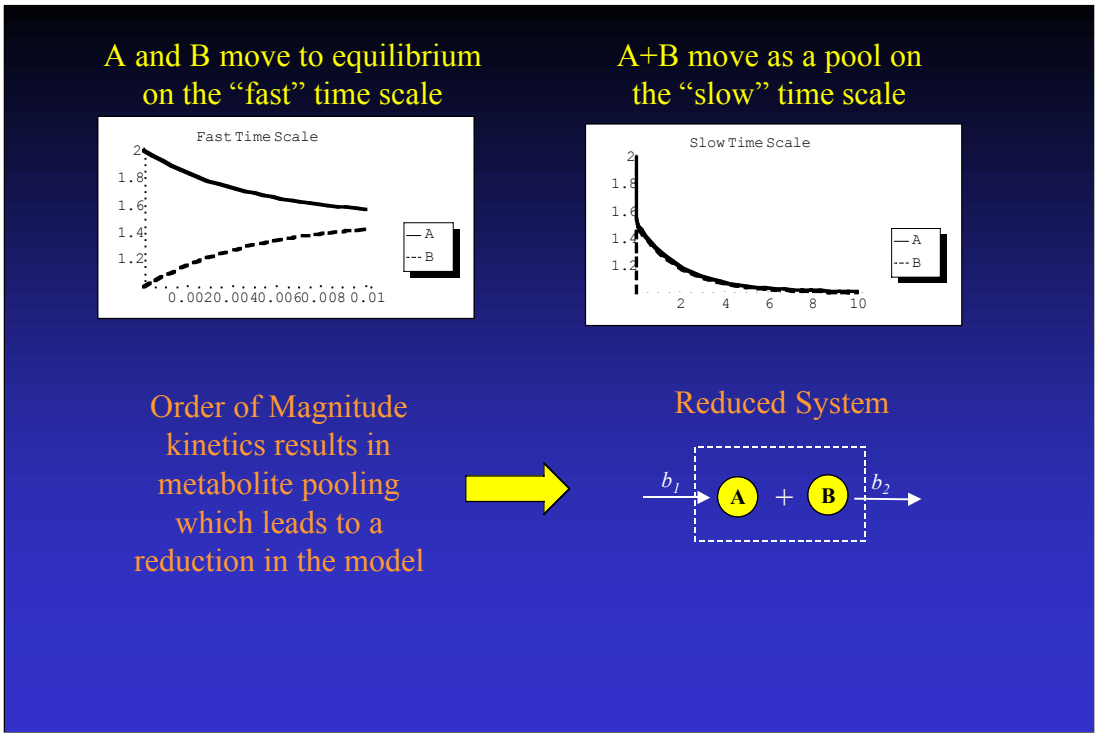
Gradient of the flux vector

$$\frac{\partial v_i}{\partial C_j} = \begin{array}{c} \\ v_1 \\ v_2 \\ b_1 \\ b_2 \end{array} \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} 100 & 0 \\ 0 & 100 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} \end{array}$$

INCLUDING CHEMICAL KINETICS

If one wants to simulate the dynamic states of this system, the kinetics of the reactions need to be known. This means that algebraic expressions for the rate laws must be provided. Here we show a simple mass action type representation of these rate laws assuming that the reaction is first order.

The Jacobian matrix that describes the dynamics is a product of S and the gradient matrix shown at the bottom of the slide.



INTRODUCTION TO TIME SCALE SEPARATION

If the reversible reaction is fast compared to the exchange fluxes, there is very little net exchange with the environment as the reaction equilibrates. Thus the system will behave like a closed system on the fast time scale. Thus a 'pool' of A+B will be formed quickly, and the total inventory in the pool will change slowly as dictated by the exchange fluxes. The state of the system is thus described by only one variable 'A+B'.

This pool formation procedure will be a key element in the massive model reduction challenge that faces us.

Geometric Representation in the Null Space

The steady state solution can be decomposed into weightings on the extreme pathways.

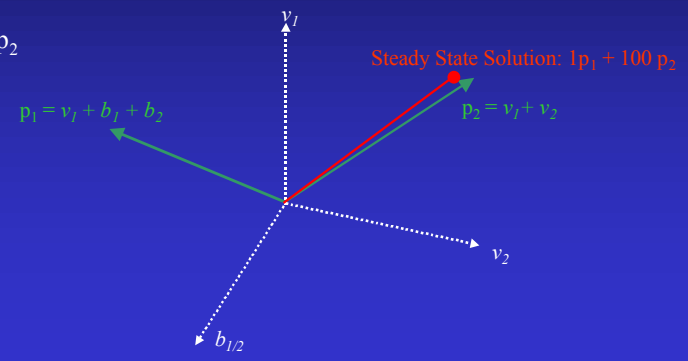
For the sample system above, the steady state solution is:

v_1=101; v_2=100; b_1=1; b_2=1

This can be broken down into weightings on the pathways:

$1p_1 + 100p_2$

The fast kinetics of v_1 and v_2 "push" the steady state solution to the p_2 edge of the cone/plane.

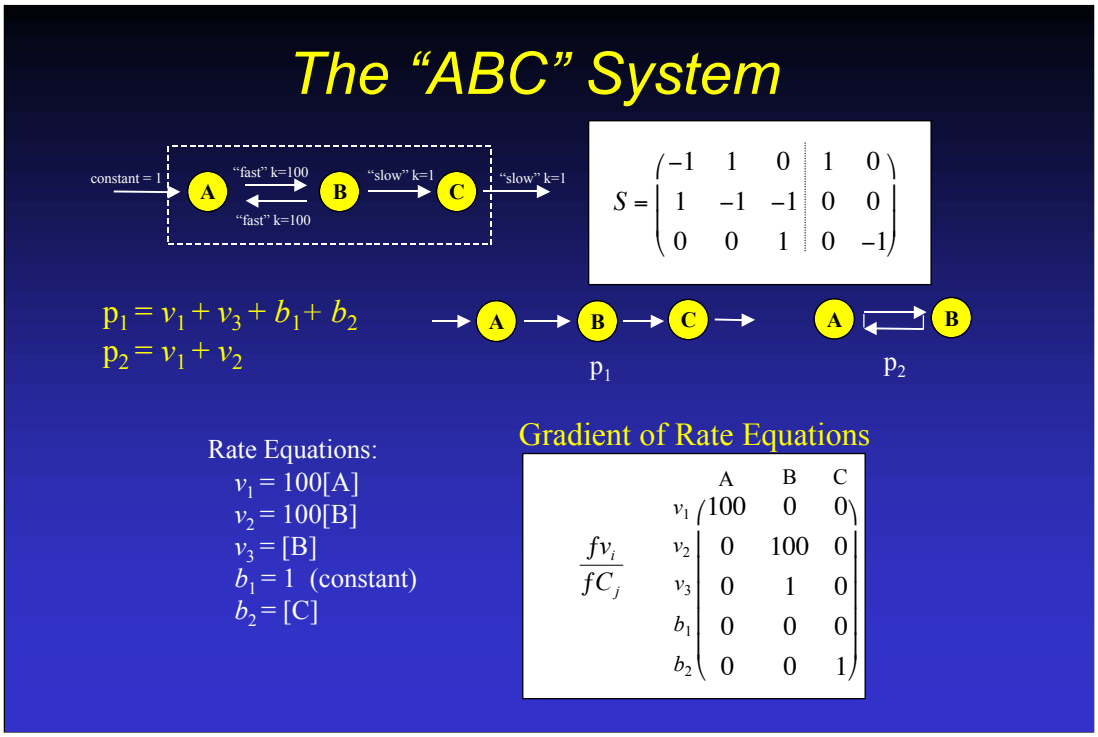


INSIGHTS FROM GEOMETRY

The null space is a cone, as we shall see in much more detail later, and the two pathways are the edges of this cone. If the kinetic parameters are well separated as indicated the steady state flux distribution is

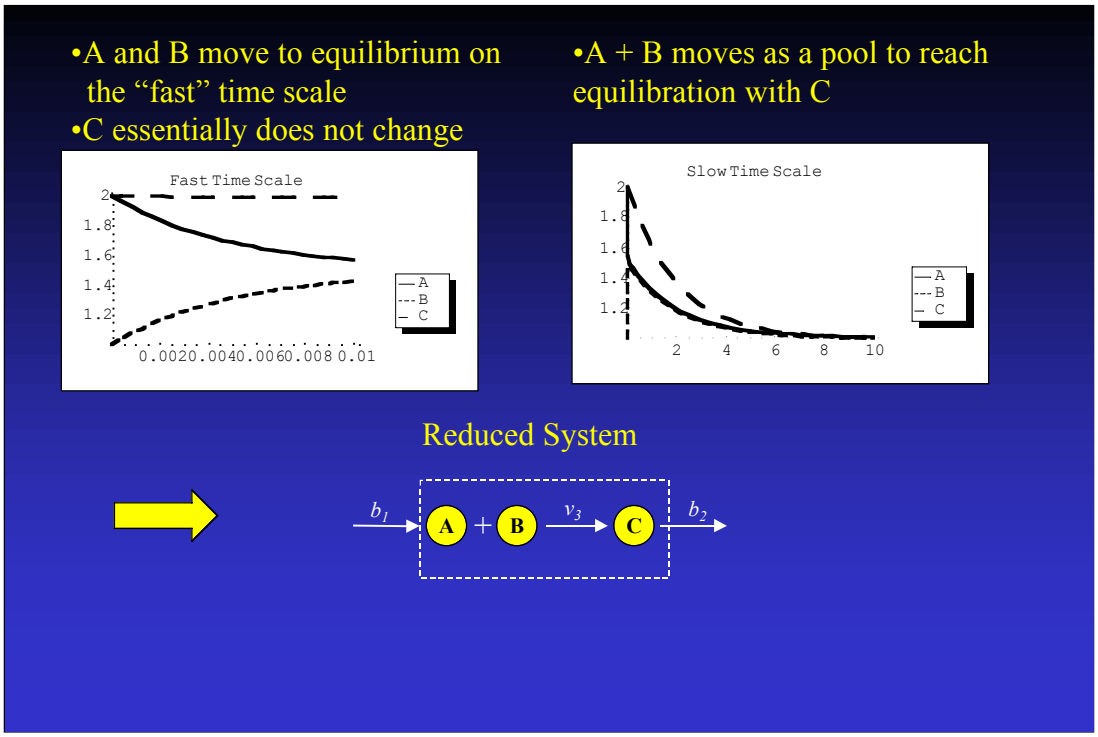
$$v_{ss} = p_1 + 100\ p_2$$

This solution lies 'close to the edge' of the solution space. We shall see this feature emerge as a principle later on



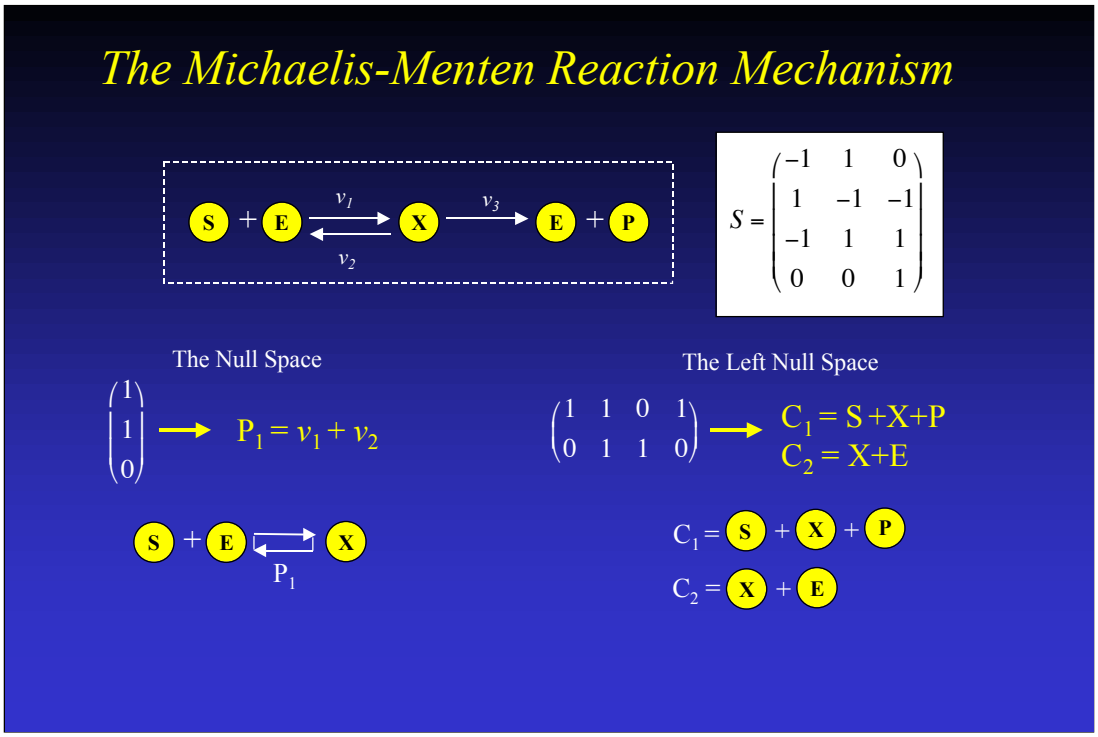
A SLIGHTLY MORE COMPEX EXAMPLE

The next two slides have a slight variation on the previous example. Now we are examining a 3 component system but the analysis is the same. A and B equilibrate on the fast time scale forming a pool (A+B). On the slower time scale the the pool (A+B) is filled via the input reaction and drained via the conversion to C.



These graphs show the pooling effect over the 2 times scales. On the fast time scale it is evident that A and B are equilibrating while C is unchanging (the concentration does not change under the fast "window of observation"). On the slower time scale, you can see that A and B move as a pool to equilibrate with C.

The reduced network that is a result of the pooling is diagramed at the bottom of the slide.



THE MICHAELIS-MENTEN REACTION MECHANISM

The classical MM mechanism can be studied along the lines introduced. For the closed system there is one pathway and two conserved moieties: the total enzyme and the total substrate species. The latter disappears as we open the system on the next slide.

A detailed kinetic analysis of the irreversible MM mechanism is found in:

B. O. Palsson (1987), "On the Dynamics of the Irreversible Michaelis-Menten Reaction Mechanism", *Chem. Eng. Sci*., **42**, 447-458.

Stoichiometric Matrix

- Can now be derived from annotated genomes given knowledge of enzyme stoichiometries
- A mathematically compact description of metabolic maps
- Has characteristic connectivity and graph properties
- Its size for simple prokaryotic cells is (300-450) X (500-750)
- The stoichiometric matrix is 'sparse', i.e. few non-zero elements
- It has well defined associated fundamental sub-spaces
- These subspaces are keys to understanding pool and pathway formation, and thus model reduction and conceptual simplification

SUMMARY OF POINTS MADE ABOUT THE STOICHIOMETRIC MATRIX

This list is a summary of the points made about the matrix. Next we shall look into linear algebra and examine the matrix properties of S.

References

- Gilbert Strang, Linear Algebra and Its Applications, Academic Press, New York, 1981.
- B. O. Palsson, "On the Dynamics of the Irreversible Michaelis-Menten Reaction Mechanism", *Chem. Eng. Sci*., **42**, 447-458 (1987).
- C.H. Schilling, S. Schuster, B.O. Palsson, and R. Heinrich, "Metabolic Pathway Analysis: Basic Concepts and Scientific Applications in the Post-Genomic Era," *Biotechnology Progress*, **15**: 296-303 (1999).
- J.S. Edwards and B.O. Palsson, "The *Escherichia coli* MG1655 *in silico* metabolic genotype; Its definition, characteristics, and capabilities," *Proc. Natl Acad Sci (USA)*, **97**: 5528-5523 (2000).
- B.O. Palsson, "The challenges of *in silico* biology," Nature Biotechnology, **18:** 1147-1150 (2000).