Flux-balance analysis:
Basic concepts

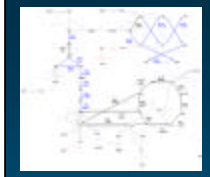
Bernhard Palsson
Lecture #6

September 15, 2003

University of California, San Diego
Department of Bioengineering

Systems Biology Research Group
http:// systemsbiology.ucsd.edu

Outline



1) Closing the flux cone

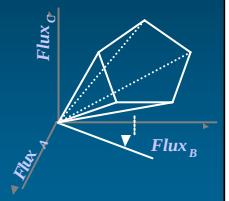
2) Linear programming
   - What is LP?
   - Objective functions
   - Cellular growth optimization

3) Constraints
   - Determining constraints
   - Sensitivity measures

4) Parameter variation

5) Phase plane analysis

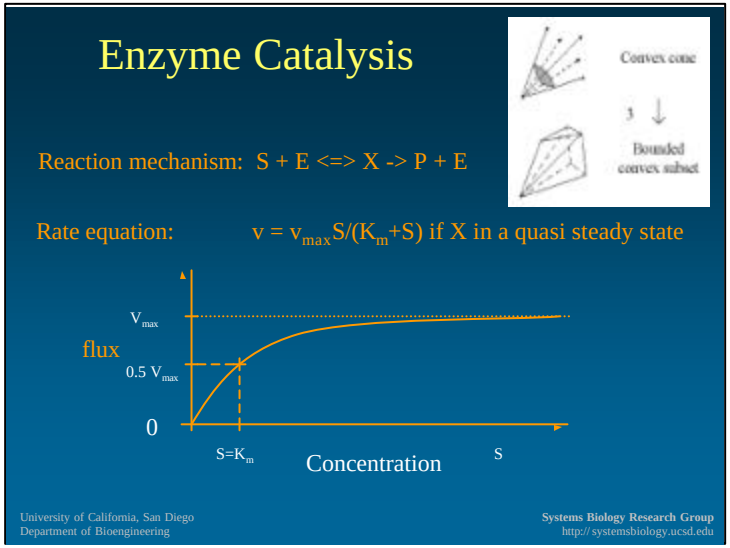
6) Phase planes and extreme pathways



University of California, San Diego
Department of Bioengineering

Systems Biology Research Group
http:// systemsbiology.ucsd.edu

Closing the Flux Cone



Capacity Constraints On Metabolic Fluxes

This slide reviews the basic principle of Michaelis-Menten kinetics. As discussed in the next slide, the flux through any given enzyme will have a finite maximum possible value. To discuss how these maximal fluxes "cap" the solution space, let us briefly review how such a space is generated.

Linear spaces are characterized by a basis set where any linear combination of the basis vectors is found in the space, i.e.;

$$\mathbf{v} = \Sigma_i \, w_i \mathbf{p_i}$$

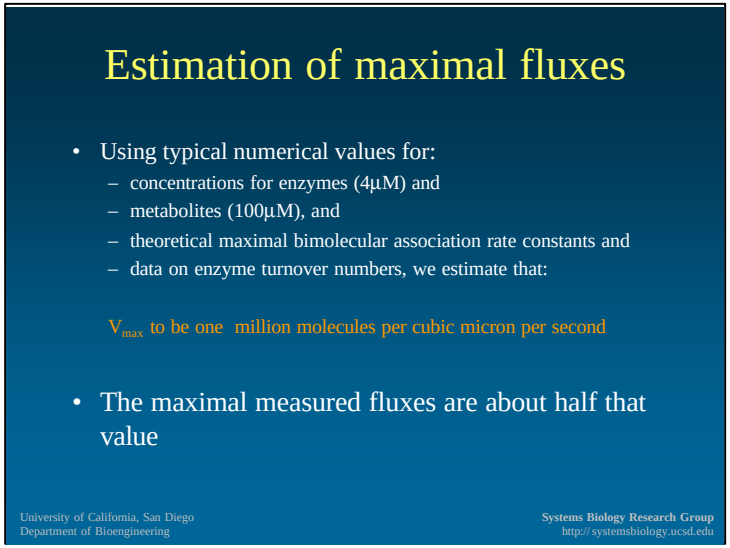
Where $\mathbf{p_i}$ are the conical basis vectors, as introduced in the last lecture. The weights, w_i, used to multiply the basis vectors in the summation are positive.

Since the individual reaction steps (v_i) in a pathway vector are carried out by an enzyme there are limitations placed on the numerical values that w_i can take in a real system:

•Minimum: the reactions are irreversible, thus the weights are positive

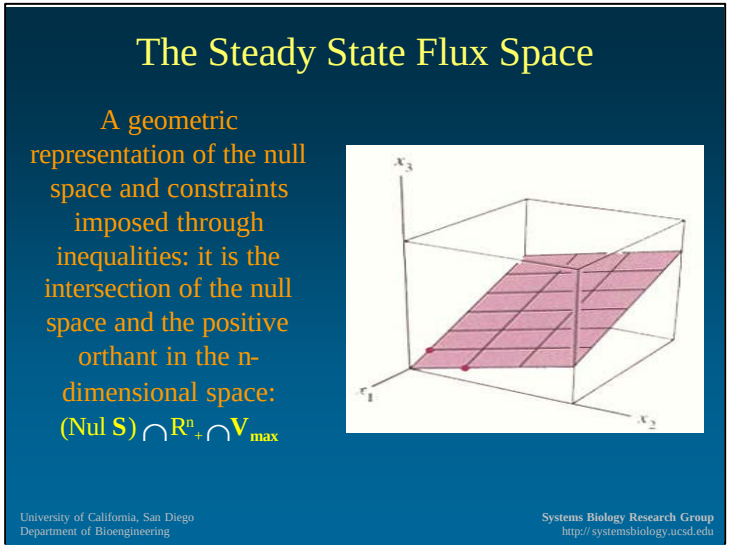
•Maximum: there is maximum flux through an enzymatic reaction, thus there are maximum weights; thus

$$0 < w_i < w_{max}$$



Estimation of Maximal Fluxes

In an earlier lecture, we estimated the typical concentrations for enzymes and metabolites based on the space available in cells, and the number of different metabolites and proteins present inside a cell. This information, along with the theoretically limited biomolecular association rate constant, led to the estimate that the maximum fluxes through a metabolic reaction in a cell is on the order of a million molecules per μ^3/sec. The most rapid metabolic fluxes measured are on the order of half of this numerical value. This limit puts an upper bound on all the fluxes that take place in a cell. Of course, the cell can down-regulate gene expression, and fluxes can be constrained below that maximum.

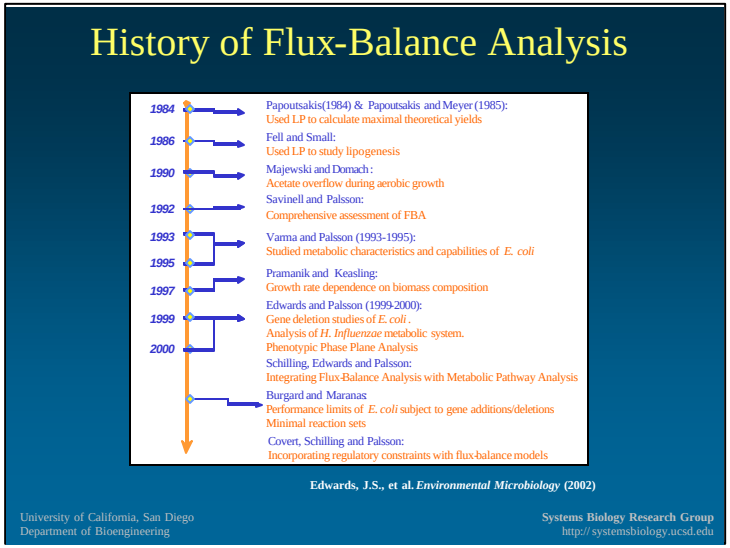


The Confined Solution Space As An Intersection

$$\mathbf{Nul\ S} \cap \mathbf{R}^{\mathbf{n}}_{+} \cap \mathbf{V}_{\mathbf{max}}$$

In linear algebra the term 'null space' is used to describe the space that contains all of the solutions to a system of homogeneous linear equations. The solution space of interest to us is actually the intersection of this null space with the region bounded by the inequalities placed on the weights. This space represents and defines the boundaries and capabilities of a metabolic genotype describing all of the possible flux distributions and routes which can theoretically operate through the system, clearly defining what an organism's metabolic network can and cannot do.

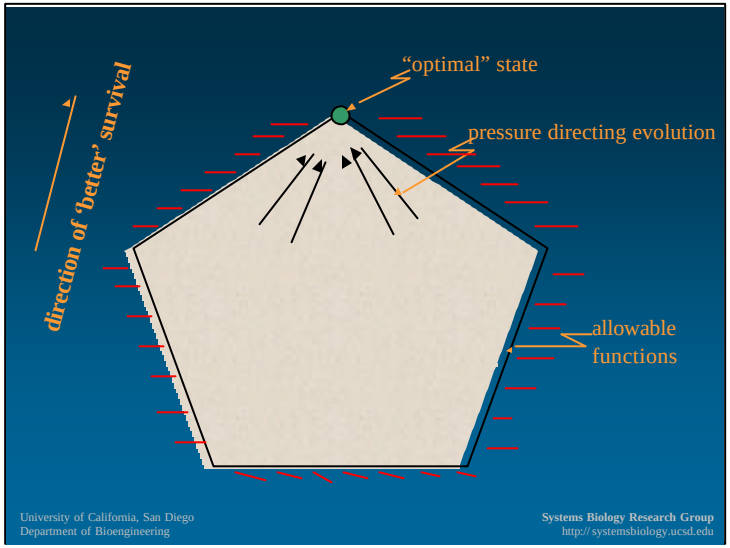
Within the solution space we can find the answers to any and all of our questions which pertain to the structure and production capabilities of an organism.



SOME HISTORICAL EVENTS IN THE DEVELOPMENT OF FBA

This slides shows some of the historical events in the development of FBA of under-determined systems. A detailed historical review is found in:

Edwards, et al *Metabolic flux balance analysis* in Metabolic Engineering, Lee and Papoutsakis Editors



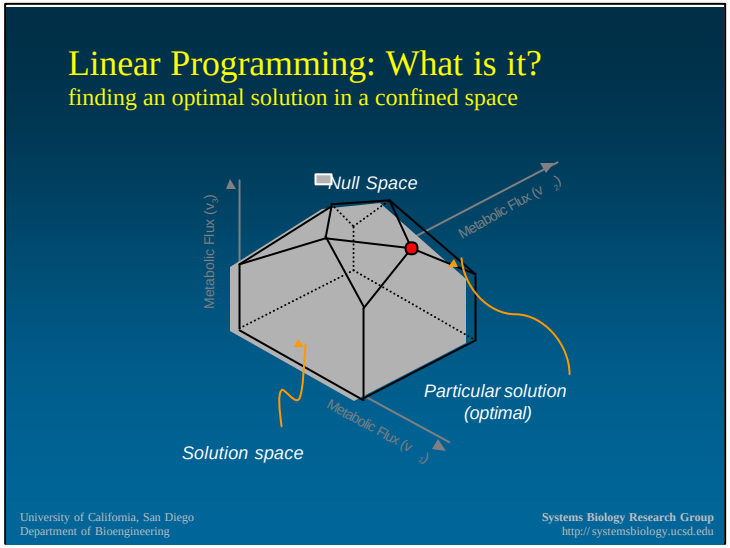
The solution space defined in the previous slides represents all of the *possible* (i.e., allowable) metabolic behavior of the cell. A living cell (or population of cells), however, will tend to exhibit only one phenotype under a given set of conditions. What, then, determines the cell's "choice" of a particular phenotype?

We have assumed that the cell's ability to survive and to grow has led to the evolutionary selection of its optimal growth in a particular set of conditions. Thus, the pressures directing evolution will drive the behavior of the cellular metabolic network toward an optimal edge of the solution space (a concept described quite nicely in Bialy's "Living on the edges," *Nat. Biotechnol.*, 19:111-112). If other selective pressures are present that, for instance, require a certain phenotype for the cell to survive (a phenotype that may not necessarily be "optimal"), then that phenotype will be selected for. Experimental validation of the hypothesis for optimal growth has been published by our research group (Edwards, *et al.*, *Nat. Biotechnol.*, 19:125-130), and will be discussed in detail in Lecture 15.

Linear Programming

University of California, San Diego
Department of Bioengineering

Systems Biology Research Group
http:// systemsbiology.ucsd.edu

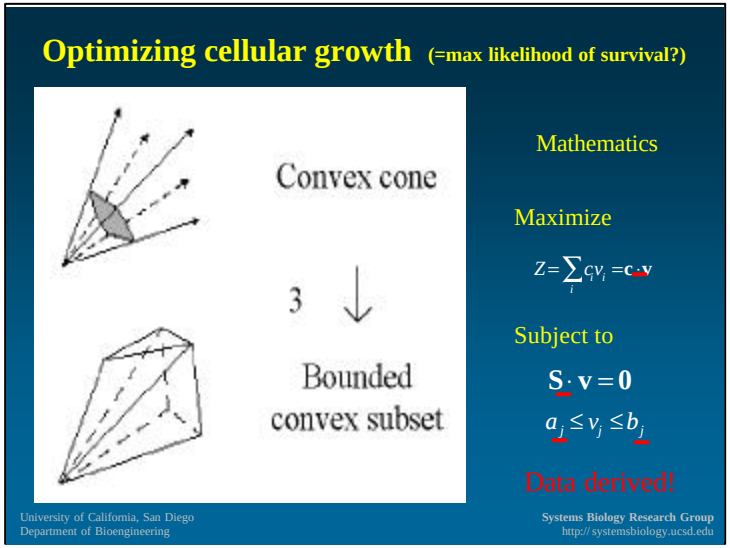


LP: What is it?

This diagram depicts a bounded polytope in 3 dimensions. Imagine that it is the space of possible solutions to a set of linear equalities with constraints, such as the flux balance equations and the capacity constraints. Each point in this space satisfies these conditions. However, the nature of the solutions differs. We can choose a particular solution in this space that is the 'best' in some sense.

This idea underlies LP. We state an objective function that measures what we are interested in. We then try to find the best value for this objective function under the given constraints. The best value normally means the maximum value. Minimization can be performed by simply finding the maximum of the negative of the objective function.

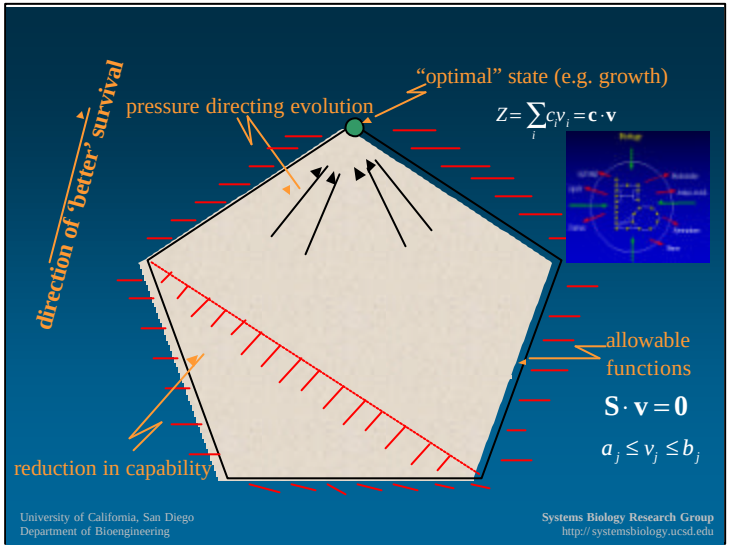
The optimal solution normally lies in a corner of the polytope. Occasionally, the objective function has the same value along a whole edge and all the points on that edge are optimal values. In this rare case the objective function is 'parallel' to the edge of the polytope.



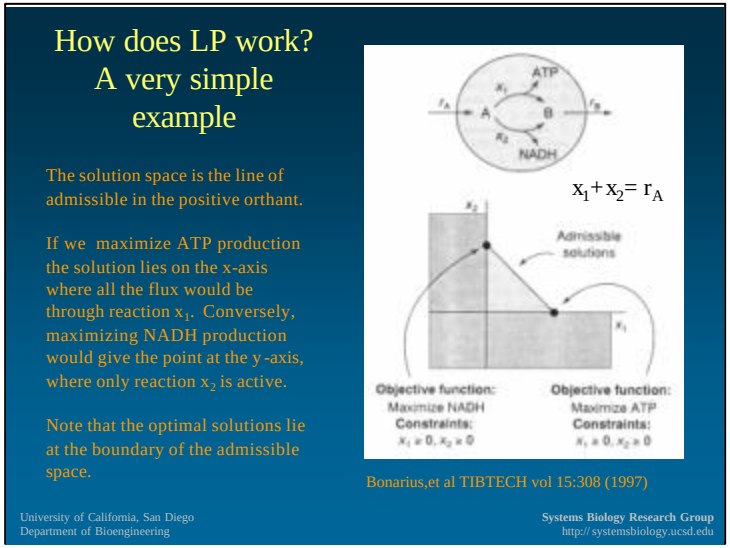
As discussed a few slides ago, the biophysical limitations of each enzyme will limit the maximal flux that each can support. Capacity constraints can thus be imposed on the value of each flux in the metabolic network. These constraints can be used to set the uptake rate for the transport reactions and define the reversibility of each metabolic reaction. Thus, the solution space is defined by the system stoichiometry (**S v**=**0**), bounded by the capacity constraints.

The determination of a particular metabolic flux distribution can be formulated as a linear programming (LP) problem, to be discussed in more detail in a later slide. The solution is then found which maximizes the objective function, *Z*, subject to the stoichiometric and capacity constraints. In the equation for *Z* presented above, the vector **c** is used to select a linear combination of metabolic fluxes to include in the objective function. As stated in the previous slide, we define cellular growth as the objective function; therefore, **c** was defined as the unit vector in the direction of the growth flux, and the growth flux is defined in terms of the biosynthetic requirements (i.e., the proportion of each component of the biomass defined from the literature). This growth flux is thus modeled as a single reaction that converts all biosynthetic precursors into biomass.

It is important to keep in mind that the stoichiometry, the bounds on the solution space, and the biomass composition used in defining the objective function all are data-derived. These data then are used to calculate (via LP) the value of the objective function and the flux map (i.e., **v**) that achieves this objective. It is also worth noting that the flux map may be non-unique (i.e., if the optimal solution lies along an edge, plane, or hyperplane, rather than simply lying at a vertex); thus, several different sets of fluxes may achieve the same optimal objective.



This slide is simply a reiteration of the previous two slides, with the added conceptualization of how a reduction in a cell's metabolic capability will lead to a shrinkage in the solution space and could conceivable lead to a different optimal solution. Such a reduction in capability can result from knocking out a gene(s) that encodes for a particular enzyme(s), or from a lowered expression of a gene(s) due to transcriptional regulation (to be discussed in much greater detail in Lecture 17).

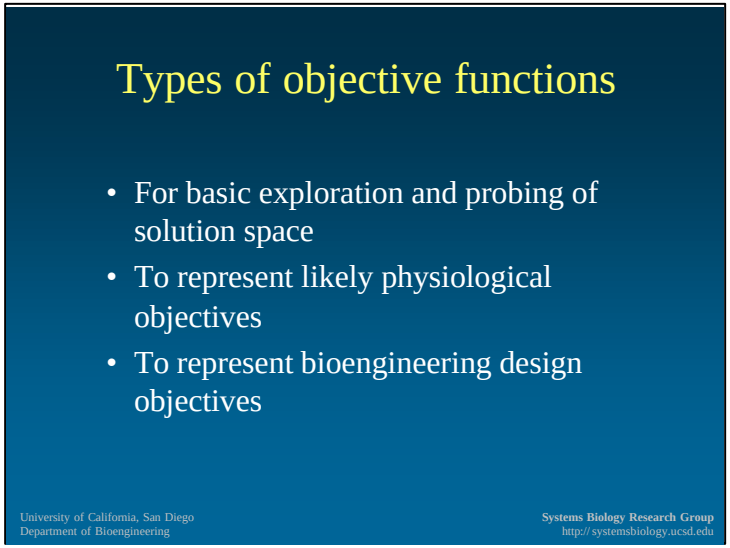


Example

This slide displays a simple and readily understandable example of linear programming. Depicted is a reaction network where a compound *A* is picked up by a cell and is metabolized to *B* via two different routes and then secreted. One route produces high energy phosphate bonds in the form of ATP. The other route produces redox potential in the form of NADH. The only flux balance in this system is that the sum of the two internal fluxes must equal the exchange flux. Once r_A is measured and known, this forms a straight line in the x_1 x_2 plane. Since the fluxes are constrained to be positive, we can only be in the positive quadrant of the plane, and thus the solution plane is the segment of the line shown in the figure.

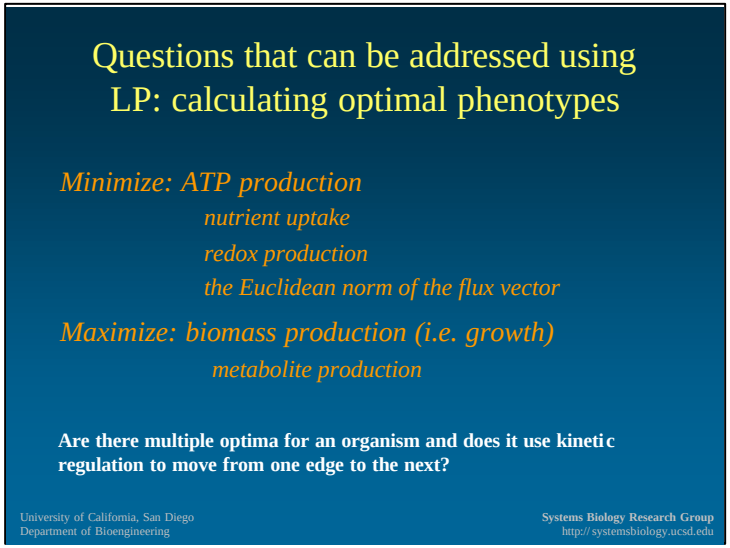
If one maximizes ATP production, it is clear that x_2 should go to zero, and x_1 to the maximum value equal to the uptake rate. That solution lies at the extreme point of the solution space to the right. Conversely, if you tried to maximize the redox production from this metabolite in the form of NADH, the optimal solution is x_2 equal to the uptake rate r_A, and x_1 goes to zero. That optimal solution is in the opposite end of the solution space. This simple example illustrates how optimal solutions in linear spaces are found at the extremities of the allowable solution spaces. We will classify the solutions further below.

What happens if you maximize $x_1 + x_2$?



The Objective Function

Within the solution space defined by the connectivity and capacity constraints, we can search for the best solution using linear optimization. What we search for is determined by the objective function stated. There are several types of objective functions that can be used. First, we can use objective functions to explore the properties of the solution space, and the capabilities of an organism. These objective functions include things like maximizing the ATP from a given substrate, or maximizing the amount of an amino acid produced from a given substrate. These types of objective functions are non-physiological, but can be used to probe the properties of a network. A second class of objective functions would represent objectives that we believe are physiologically relevant. For microbial cells, the belief is that they maximize their growth rate given the constraints under which they operate. In this case, the objective is the balanced exit from the network of all the precursors needed for the synthesis of the cellular mass. The third type of objective function may relate to an intentional engineering objective of a metabolic system. We may wish to maximize a product like Lysine, for instance, and try to figure out what the best flux maps are that lead to the production of Lysine. We can add or delete reactions from the network to determine how those changes affect the yield of the desired product.



Optimal Phenotypes

A number of different objective functions have been used for metabolic analysis. These include:

Minimize ATP production: This objective is stated to determine conditions of optimal metabolic energy efficiency.

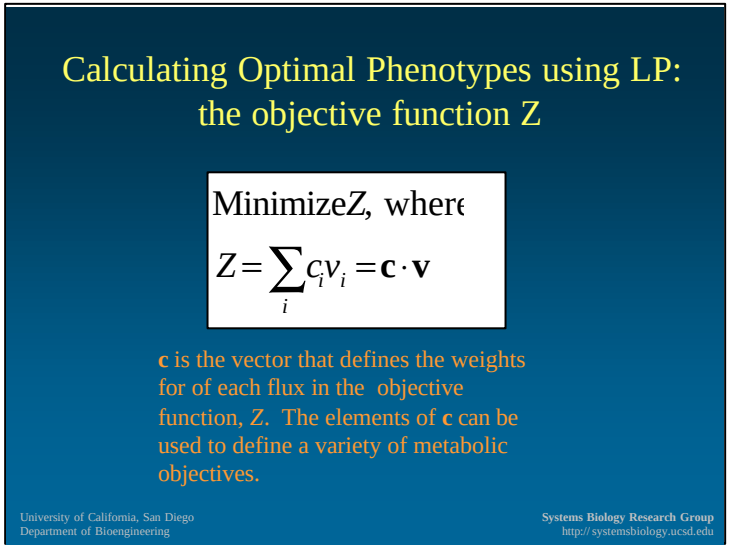
Minimize nutrient uptake: This objective function is used to determine the conditions under which the cell will perform its metabolic functions while consuming the minimum amount of available nutrients.

Minimize redox production: This objective function finds conditions where the cells operate to generate the minimum amount of redox potential.

Minimize the Euclidean norm: This objective has been applied to satisfy the strategy of a cell to minimize the sum of the flux values, or to channel the metabolites as efficiently as possible through the metabolic pathways.

Maximize metabolite production: This objective function has been used to determine the biochemical production capabilities of *Escherichia coli.* In this analysis, the objective function was defined to maximize the production of a chosen metabolite (i.e. lysine or phenylalanine).

Maximize biomass and metabolite production: By weighing these two conflicting objectives appropriately, one can explore the tradeoff between cell growth and forced metabolite production in a producing strain.

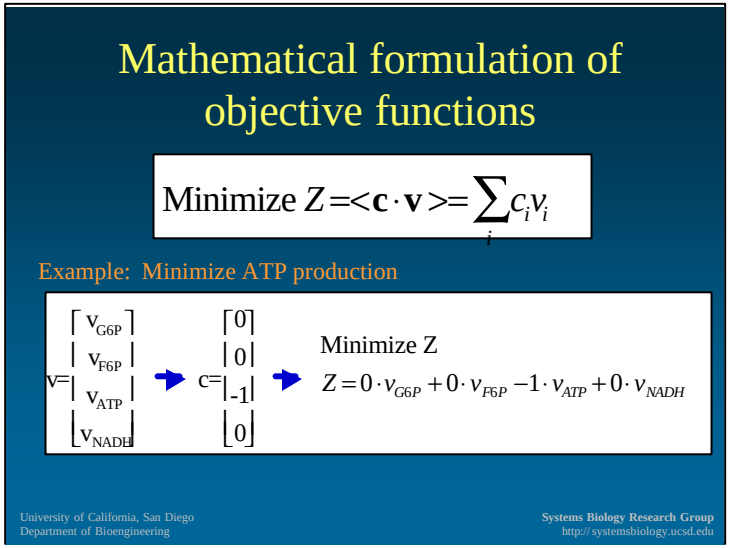


The Objective Function

Numerous questions about metabolic capabilities can be answered using LP. The stoichiometric and capacity constraints define a range of allowable behavior. We can then find the best value within these constraints. Biologically, we have defined the space of all phenotypes (that is particular solutions) that can be derived from a genotype. We can calculate the best phenotype from a particular standpoint. For instance, we can calculate the maximum number of ATP molecules that can be generated from a particular substrate.

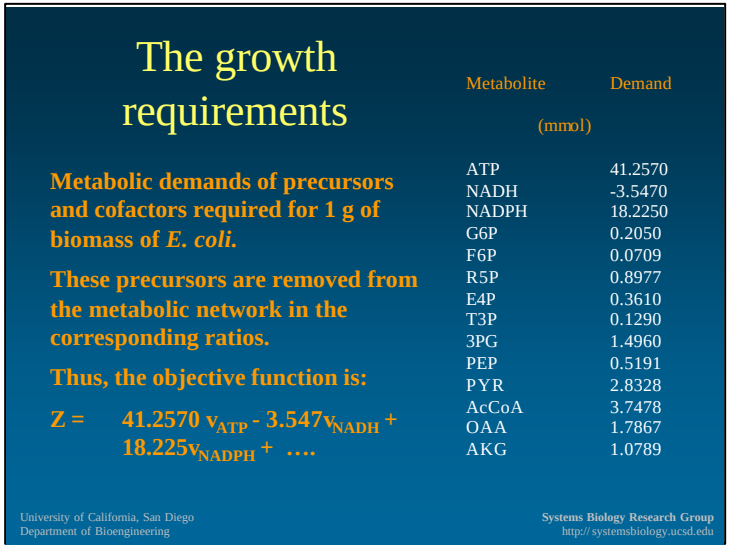
The next slide lists a number of important phenotypic behaviors that can be calculated using LP. The maximum growth function is perhaps the one of greatest interest from an evolutionary standpoint.

This general representation of Z enables the formulation of a number of diverse objectives. These objectives can be design objectives for a strain, exploitation of the metabolic capabilities of a genotype, or physiologically meaningful objective functions, such as maximum cellular growth.



Mathematical Formulation of Objective Functions

This slide illustrates the formation of the objective function using a simple example. In the example there are 4 metabolite fluxes. The objective is to minimize ATP production, therefore the **c** matrix has a zero "weight" on all fluxes except v_{ATP} which has a -1 The coefficient on the ATP flux is negative since it is being minimized.



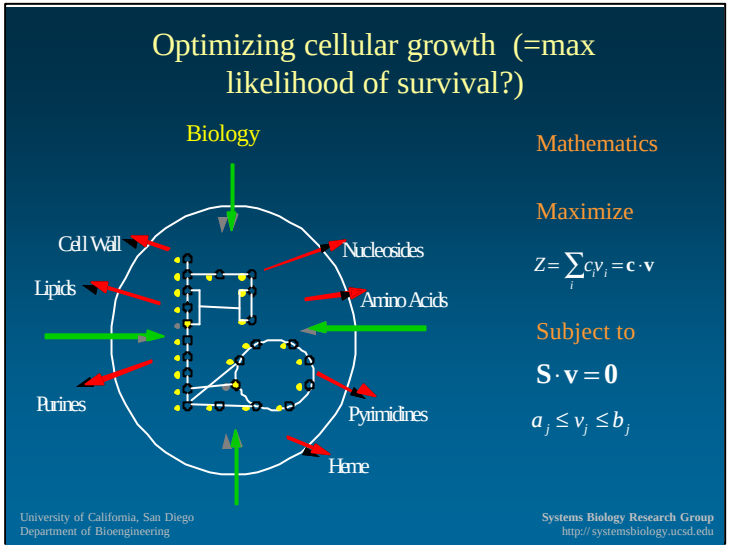
The Growth Function

This table shows the requirements for making one gram of *E. coli*. This means that for the cell to grow, all these components must be provided in these amounts. Thus, a balanced set of metabolic demands makes up the growth objective function:

$$Z = 41.257v_{ATP} - 3.547v_{NADH} + 18.225v_{NADPH} + 0.205v_{G6P} + 0.0709v_{F6P} + 0.8977v_{R5P} + 0.361v_{E4P} + 0.129v_{T3P} + 1.496v_{3PG} + 0.5191v_{PEP} + 2.8328v_{PYR} + 3.7478v_{AcCoA} + 1.7867v_{OAA} + 1.0789v_{AKG}$$

The biomass composition thus serves to define the weight vector **c.**

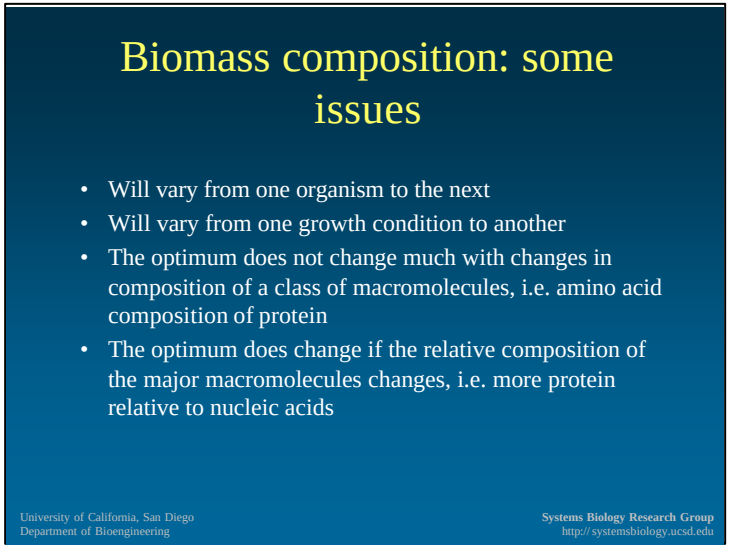
The full growth function for *E. coli* is more complicated than the one given above, since various maintenance functions need to be considered.



The Maximization of Biomass Formation

This slide shows schematically on the left the idea of maximizing biomass formation. There can be one or more inputs (the green arrows) and a balanced (linked) output that corresponds to the biomass composition.

On the right, we show the mathematical formulation of the problem. We wish to maximize the objective function under the stated constraints. These constraints form a closed cone as explained earlier.



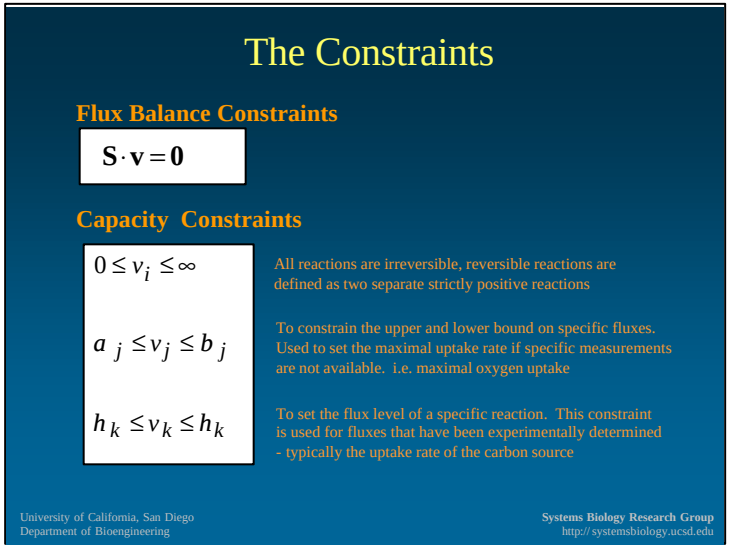
Biomass Composition

The physiologically interesting objective that we wish to study throughout these notes is the maximization of biomass yield. The definition of the solution space has few ambiguities associated with it, but the statement of the objectives has a few uncertainties built into it. First, the biomass composition is variable. It is different from one organism to another. It varies from one growth condition to another, and both of those may potentially be important issues and change the predicted optimum behavior. Legacy databases of biomass composition are needed.

The limited calculations that have been performed show that the optimum solutions do not change significantly with the monomeric composition of the major macromolecules. For instance, if the Valine to Alanine ratio is varied in the protein of a cell, the optimal growth rate does not significantly change. Conversely, if the protein relative to lipid composition in a cell changes, the optimum solution tends to be affected.

As will be shown, one can invert this problem and look at an edge of the solution space and then calculate all the objective functions that are maximized under those conditions. This might give better insight into the objectives that cells are trying to accomplish.

Constraints



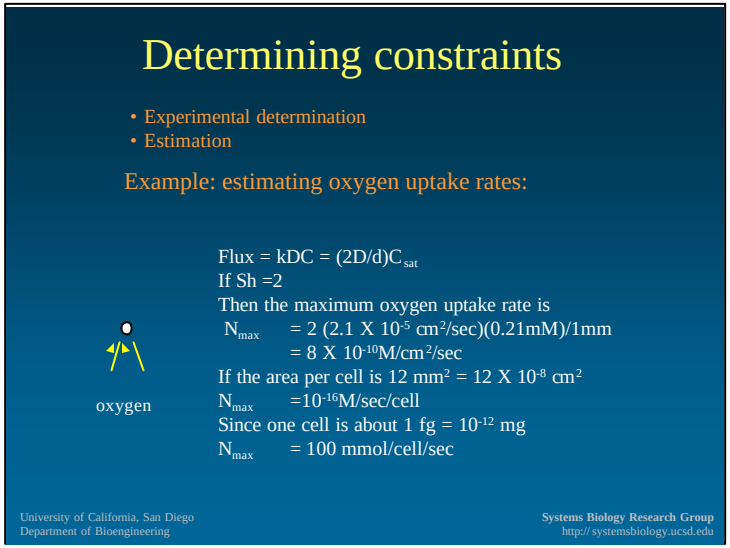
Imposing Constraints

The constraints are stated as the flux balance equations and capacities of the individual fluxes.

Using LP, constraints can be placed on the value of the flux through each of the metabolic reactions.

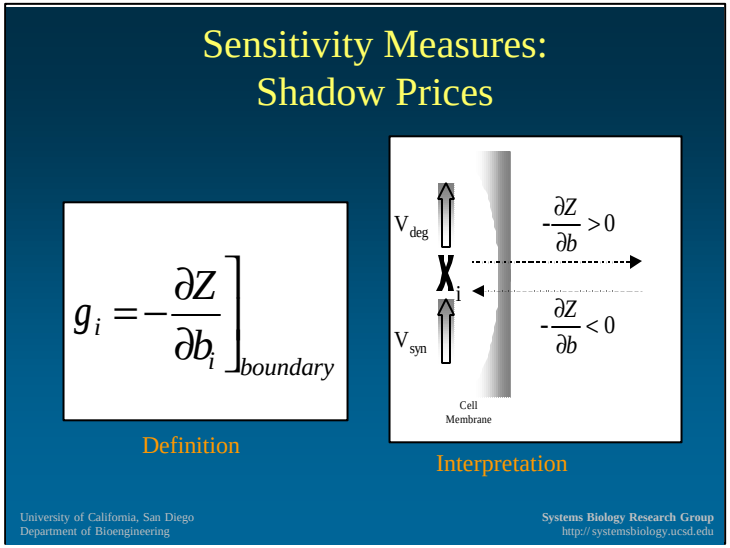
These constraints could be representative of a maximum allowable flux through a given reaction, resulting from a limited amount of an enzyme present. These constraints could also be used to include the knowledge of the minimum flux through a certain metabolic reaction.

The restriction of the flux through certain reactions can be used to model the regulatory events occurring within the cell. Further, any experimental information can be represented by a and b by imposing the error range in the measurements made. The incorporation of regulatory information in the model is also possible. For instance, the catabolite repression of certain enzymes, such as *ppc* and *pps* during growth on glucose can be represented by setting these fluxes to zero or some low level. Thus, the use of such additional constraints provides increased flexibility in analyzing the metabolic network by incorporating additional knowledge about a particular cell.



Stating Constraints

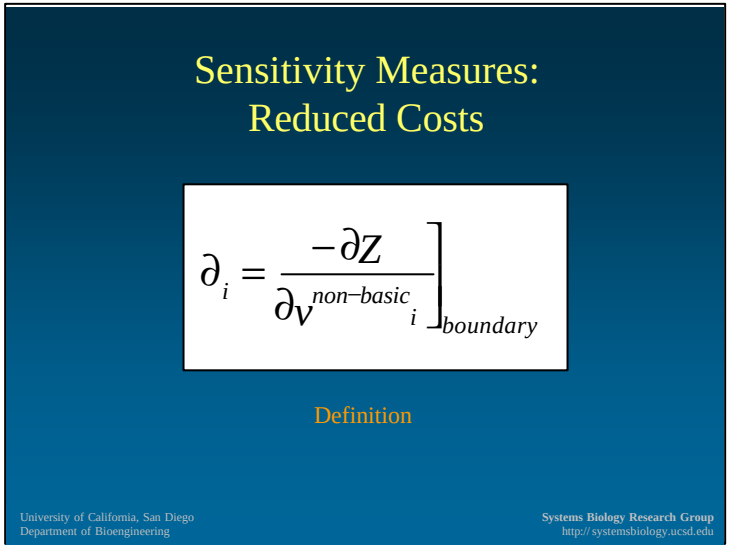
Using the diffusivity equation, the oxygen uptake rate of a cell can be estimated.



The Dual Problem: The Shadow Prices

In designing metabolic engineering strategies, an important question is; to what extent can specific fluxes be altered, and what the ensuing effect will be on the cellular processes of interest, including growth and product formation? These issues can be addressed within the LP formulation by using sensitivity analysis of the optimal solution.

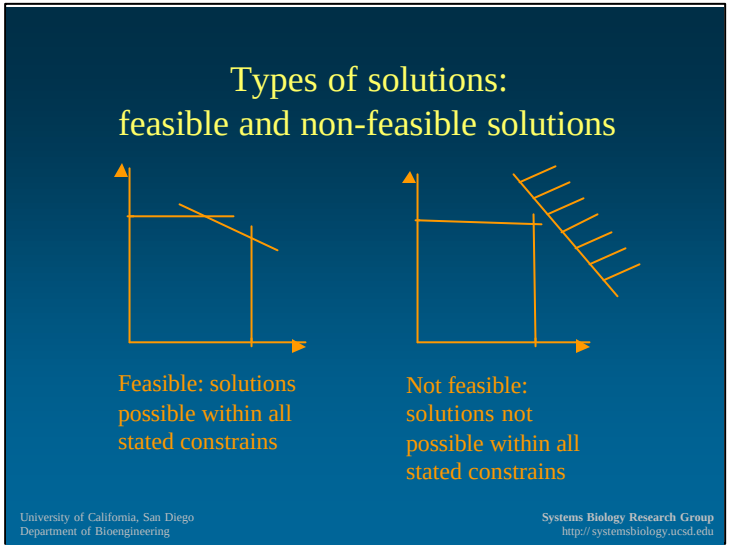
The so-called shadow prices, derived from the dual problem are the derivative of the objective function at the boundary. The shadow prices can be used to determine whether the cell is limited by a particular constraint. This feature has proven to be useful in interpreting optimum solutions, and metabolic decision making.



The Reduced Costs

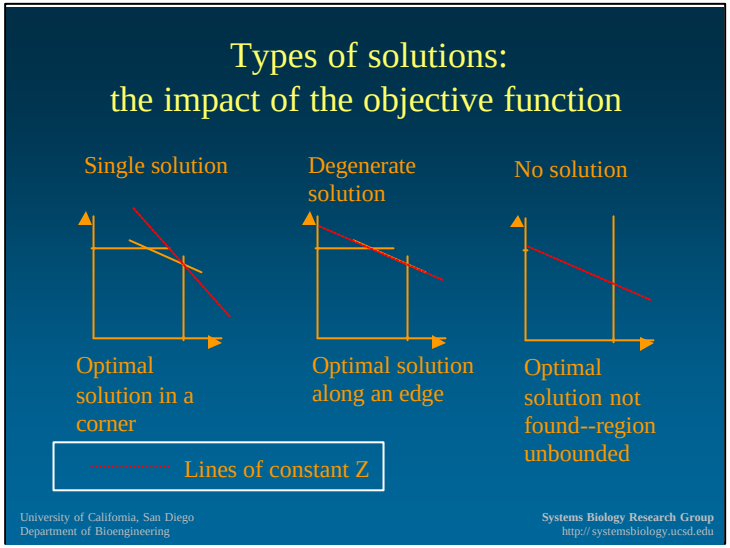
The reduced costs can be defined as the amount by which the objective function will be reduced if the corresponding enzyme is forced to carry a flux (expressed or "turned on").

In the analysis of metabolic systems, several important questions arise that can be addressed with an analysis including the reduced costs. The reduced costs can be used to analyze the presence of alternate equivalent flux distributions, i.e. if the right set of reduced costs are zero. Additionally, the reduced costs are important in examining the effect of gene deletions on the overall function of metabolism.



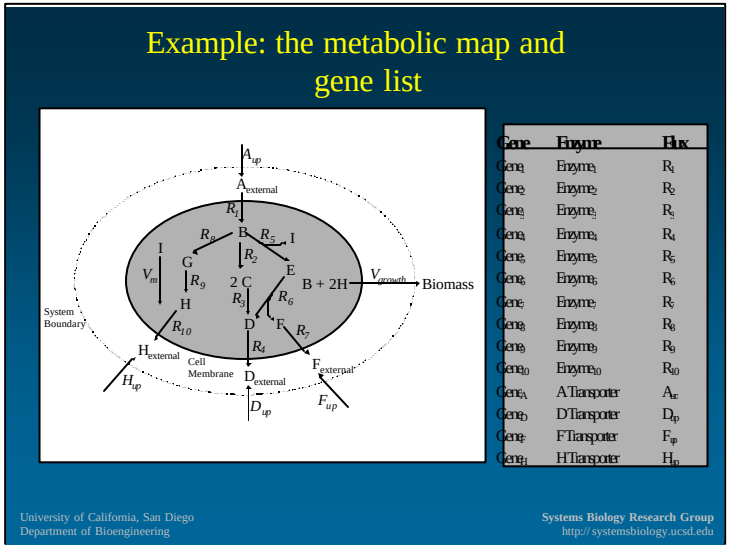
Types of Areas Formed By Constraints

The stated constraints can be consistent and form a feasible set as shown on the left. In some cases, the constraints are such that there are no solutions that can satisfy all the constraints. As seen below, sometimes the feasible set is unbounded and stretches infinitely in one direction.



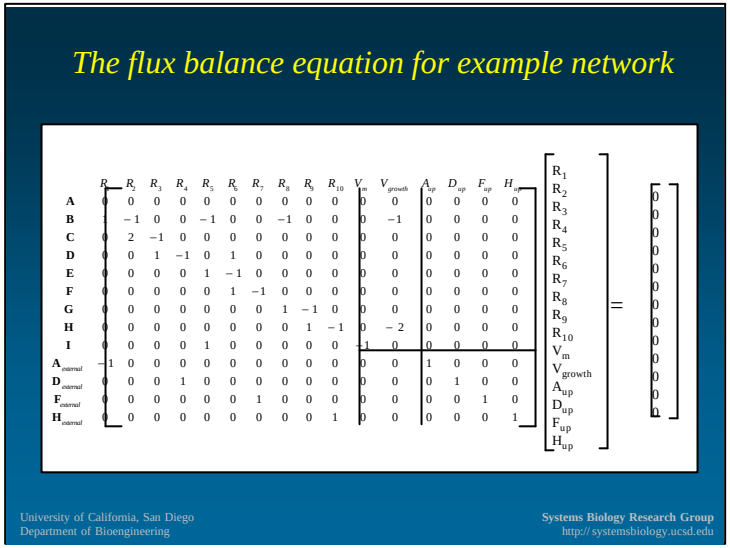
The Types Of Solutions Found

On the left we show the most common situation, namely that the optimal solution lies in a corner of the feasible set. In rare instances, the line formed by the objective function is parallel to a constraint. In this case, the entire edge of the feasible set has the same value of the objective function and all the points along the edge represent an optimal solution. Sometimes the feasible set is unbounded and the objective function increases without limit in the open direction. In this case, no solution is found.

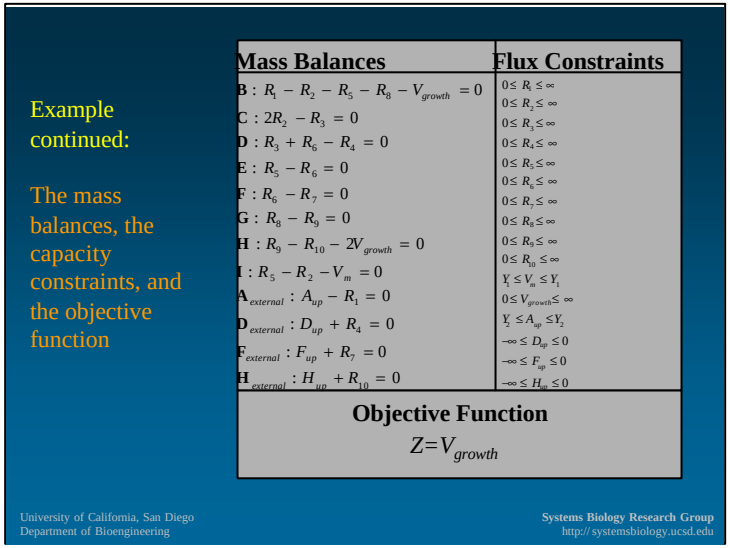


A list of the genes is obtained and the fluxes are indexed. Then the metabolic map is drawn.

In this case the growth function is a simple addition of two biosynthetic precursors.



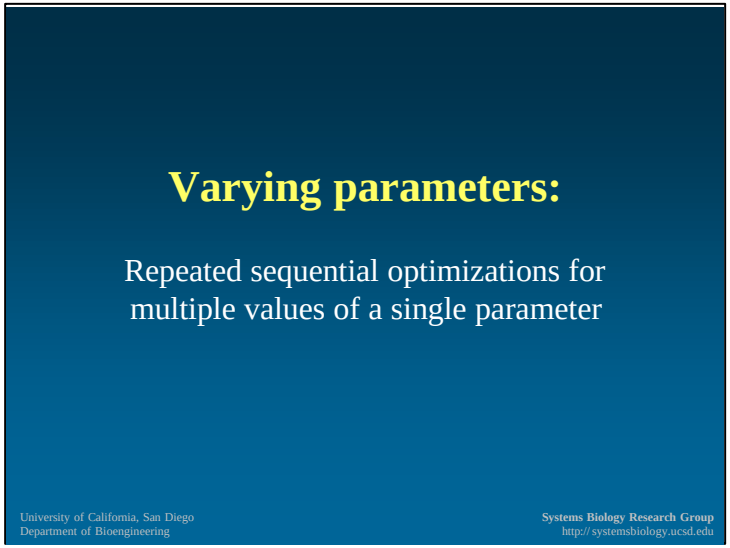
The stoichiometric matrix is formulated and flux balances stated. Note that the matrix is partitioned into external and internal fluxes.



The flux balances in a simplified form, with the capacity constraints. The objective function is the growth function.

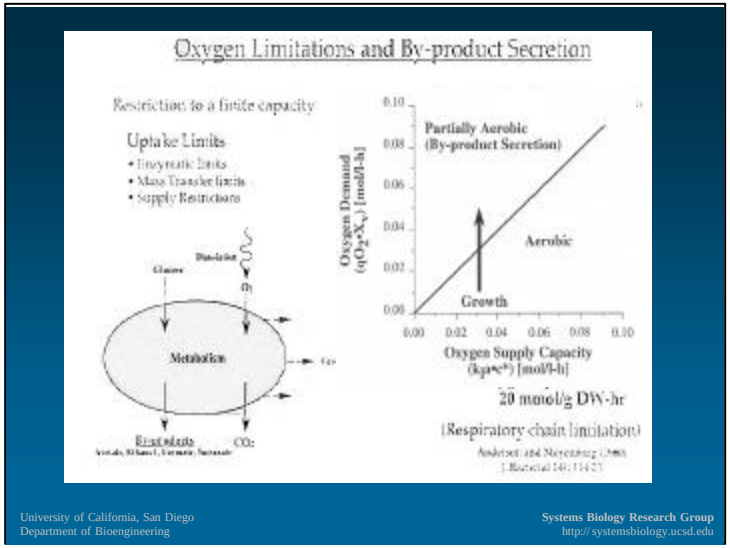
Now these equations can be entered into a software package that allows you to optimize the stated objective.

END OF EXAMPLE



In a previous lecture we looked at optimal mitochondrial flux maps for different substrates and for constraints on several internal fluxes. These are calculated for a discrete set of conditions.

We may however, be interested in the range of numerical values for a particular parameter. Thus, we can calculate a series of optimal solutions for small incremental changes in a parameter in the system. If the increments are small enough, we effectively get a continuous variation in the parameter of interest.

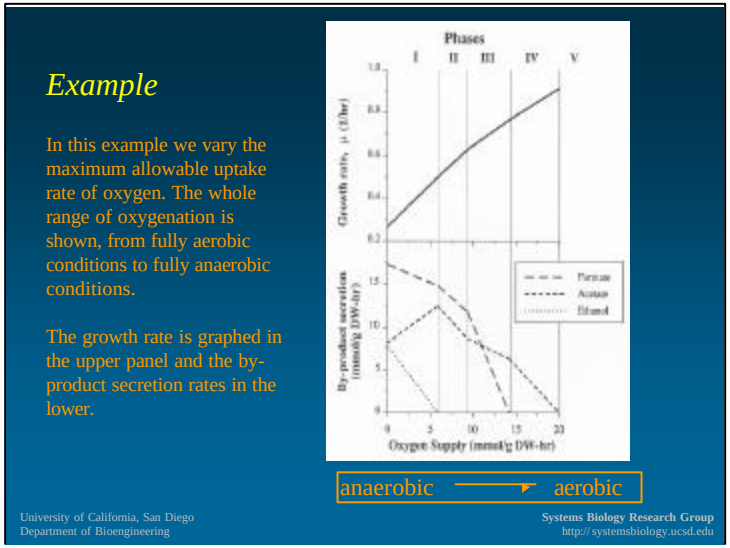


Example: Reducing Oxygen Availability

When cells grow in the laboratory with abundance of substrate they grow to high densities, eventually outstripping the ability for oxygen to be supplied rapidly enough to support fully aerobic growth. As oxygen becomes limiting, the cells must partially oxidize their substrate and secrete metabolic by-products.

The panel on the left illustrates this problem at the cellular level. At the right this problem is illustrated from a bioprocess viewpoint. The arrow indicates growth of a culture supplied with a constant rate of oxygen. As the culture grows, the oxygen demand increases and passes the line indicating the boundary between aerobic and anaerobic growth.

The following slides were prepared with a reduced *E. coli* model in 1993 (Varma, A&EM), but it illustrates how parameter variations can be used to study problems of fundamental physiological relevance, and practical importance.



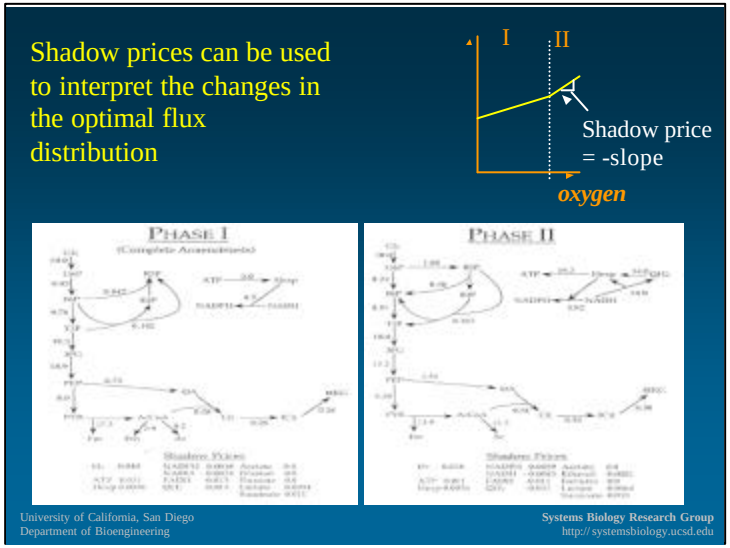
Varying Oxygen Availability

As the dissolution of oxygen cannot keep up with the high volumetric consumption rates at high cell density, the amount available per cell is reduced. Computationally this is represented by lowering the capacity constraints on the oxygen uptake rate.

The results from a series of LP calculations with varying b_{O2} is shown in this slide (glucose uptake rate is fixed at 10mmol Glc/gDW-hr). The optimal growth rate drops as the oxygen uptake rate is reduced, as shown in the upper panel. It does so in piece-wise linear fashion where changes in the slope occur at well defined oxygen uptake rates. This feature naturally divides the range of oxygen uptake rates into distinct phases.

The lower panels shows the secretion rates of metabolic by-products: formate, ethanol and acetate. Each one of these by-products is secreted in a fundamentally different way in each phase. As oxygen is reduced, incomplete oxidation of glucose takes place and metabolic by-products are secreted; acetate is first secreted, then formate, followed by ethanol.

The LP solution in each phase is fundamentally different and the transition from one to another can be interpreted using shadow prices.

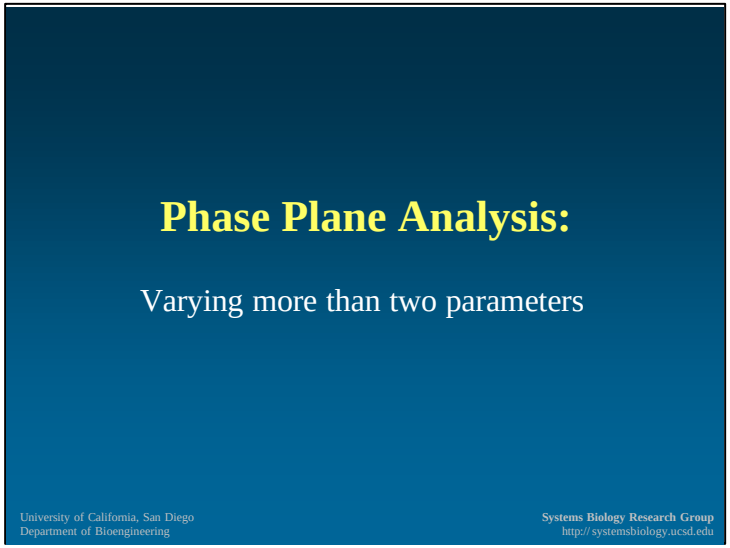


Changes In Shadow Prices At Phase Boundaries

The shadow price changes discontinuously at the boundary from one phase to the next. In fact, the change in the shadow price defines the boundary between the phases. The shadow prices basically tell us how the governing constraints on the objective function change and how the base optimal LP solution changes. This change is reflected in a shift in the flux map.

Phase I shown above is for completely anaerobic growth. The shadow prices for oxygen and ATP are negative, indicating that these are constraining factors, since the objective function would increase as more of these compounds are provided to the cell. Some of the redox carriers have positive shadow prices indicating that the cell has a problem with excess redox potential. This is characteristic of anaerobic metabolism.

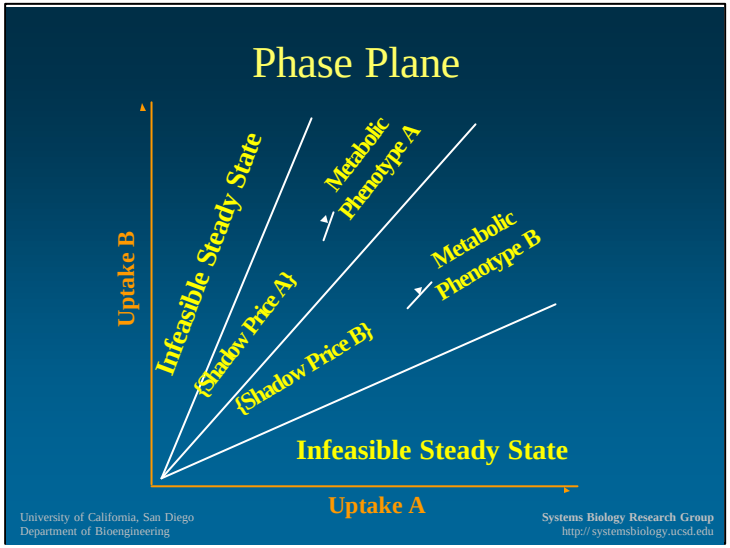
In Phase I, acetate, ethanol, and formate, all have zero shadow prices, indicating that these intermediates are useless to the cell. Thus they are secreted. Notice that in Phase II, ethanol has a negative shadow price. It thus has value to the cell and is not secreted. In fact the defining difference between the optimal flux maps in phase I and II is the secretion of ethanol. The shadow prices are thus key in interpreting the optimal flux maps and changes in the maps as parameters vary.



Phenotype Phase Plane Analysis

A useful way to extend the study of metabolic genotype-phenotype relation is to use two parameters that describe the growth conditions (such as substrate and oxygen uptake rates) as two axes on an x,y-plane. Then the optimal flux-maps can be calculated for all points in this plane. There are a finite number of fundamentally different optimal metabolic flux maps present in such a plane. The demarcations between the different flux maps are determined from the shadow prices of the metabolites. As we have seen, the shadow prices are sensitivity parameters that are calculated in the dual solution to the LP problem, and can be used to interpret shifts from one optimal flux distribution to another . This procedure leads to the definition of distinct regions in the plane in which the optimal use of the pathways is fundamentally different, corresponding to a different phenotypic behavior. We will denote each phase as: $Pn_{x,y.}$ Where P represents phenotype, n is the number of the demarcated region for this phenotype,and x,y the two uptake rates on the axis of the plane.

This phase plane resembles the phase planes used in physical chemistry, which define the different states (i.e., liquid, gas or solid) of a chemical system depending on the external conditions (e.g., temperature, pressure). The plane that we have just described can thus be called the phenotype phase plane (PhPP) for a given genotype. The construction of the phase plane and its main features will now be described, and then conceptually illustrated with a simple example.



The Phase Plane

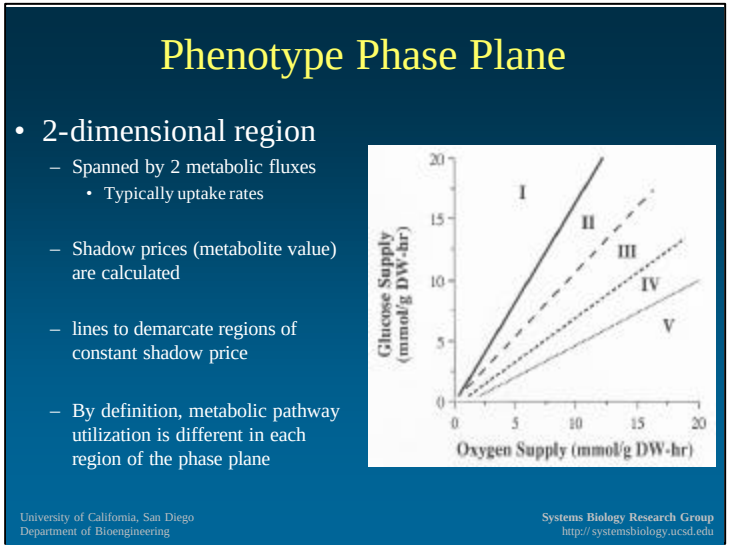
Using the shadow prices, we can define a phase plane.

A phase plane is a two dimensional region that is spanned by 2 metabolic fluxes. These fluxes are typically uptake rates, but this isn't required. And then the shadow prices for all the metabolites are calculated for all the points within this space, and lines are drawn to demarcate regions of constant shadow prices.

The shadow prices are constant within each region and will be different in the other regions.

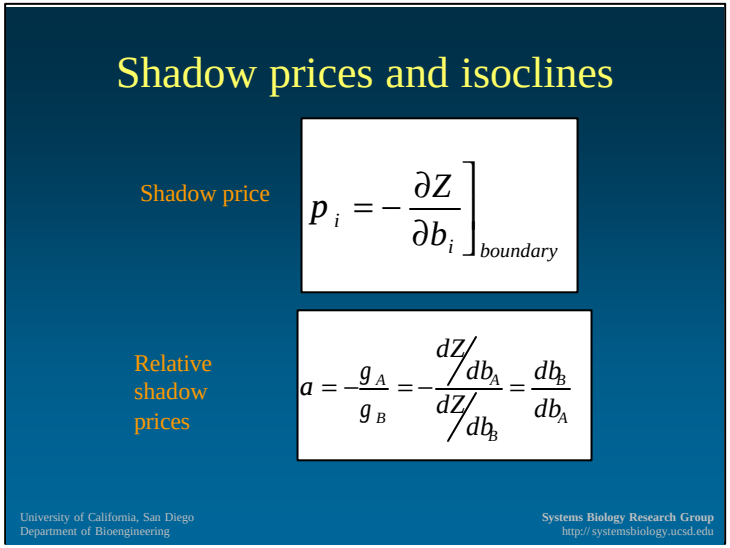
Each region typically refers to a different basis solution, which implies a different utilization of the metabolic pathways or a different metabolic phenotype.

Thus, the utilization of the metabolic pathways will be qualitatively different depending on the region of operation within the phase plane.



Summary of Phenotype Phase Planes

The example on the right indicates 5 distinct phases when comparing glucose supply to oxygen supply. Typically, PhPPs are drawn with a carbon source on the x-axis, and oxygen uptake rates on the y-axis.

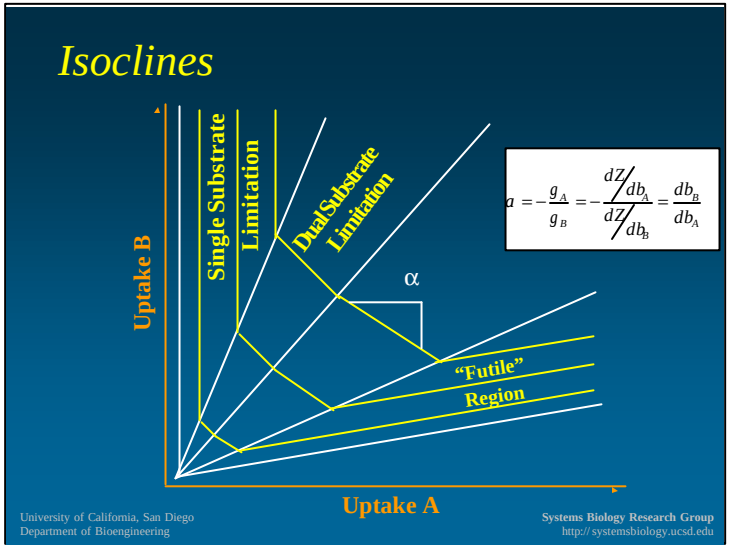


Isoclines

The isoclines represent the combinations of the metabolite uptake rates that will lead to the same value of the objective function. The slope of the isoclines within each region is calculated from the shadow prices; thus, it follows that the slope of the isoclines will be different in each region of the PhPP.

The shadow price is the sensitivity of the objective function (Z) to changes in the availability of metabolites (the **b** vector defines the right hand side of the mass balance constraints). The numerical value on the shadow price can be negative, zero, or positive, depending on the intrinsic value of the metabolite. A ratio of shadow prices between the two external metabolites can be defined.

The negative sign on α is introduced in anticipation of its interpretation. The ratio α is the relative change in the objective function for the two key exchange fluxes. In order for the objective function to remain constant, an increase in one of the exchange fluxes will be accompanied by a decrease in the other and thus we introduce the negative sign on the definition of α. Therefore, the slope of the isoclines (within each region of the PhPP) will be equal to the negative ratio of the x-axis variable shadow price and the y-axis variable shadow price, and this parameter is termed α.



Phase Plane With Objective Function Isoclines

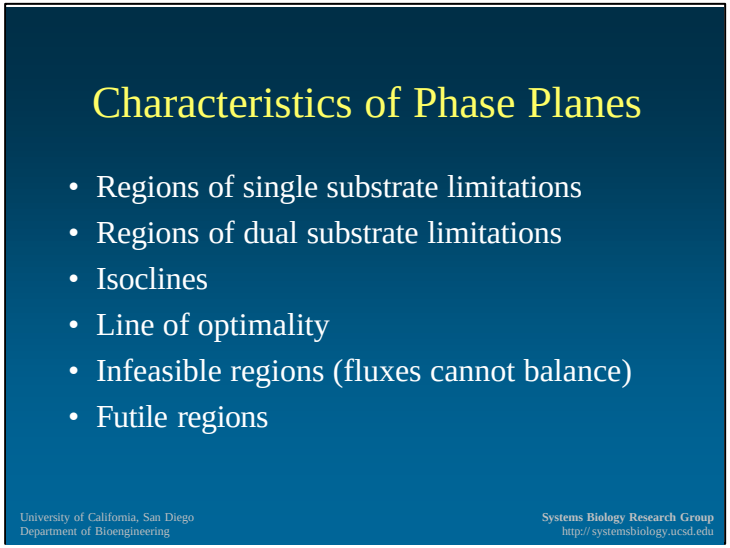
The definition of the shadow prices can be used to determine the slope of the isoclines within each region. Due to the definition of the phase plane, the slope of the isoclines will be constant within each region, however it will be different in the other regions.

We can draw isoclines for the objective function on the phase plane. The Isoclines are defined as the lines that will provide the same value of the objective function as the parameters on the X and Y axes are changed.

For example, as you follow this line, the objective function (here taken as growth rate) will be constant.

The state of the metabolic network can be classified by the value of alpha.

For example, a negative slope as shown here indicates dual substrate limitation. Isoclines can also be horizontal or vertical, and this corresponds to single substrate limitation. These situations occur when the shadow price for one of the substrates goes to zero, and thus has no value to the cell. Finally, the sign can change on one of the shadow prices, this will cause the isoclines to have a positive slope. This indicates a situation where one of the substrates is in excess and is actually inhibitory toward the cell. This defines a "futile" region.



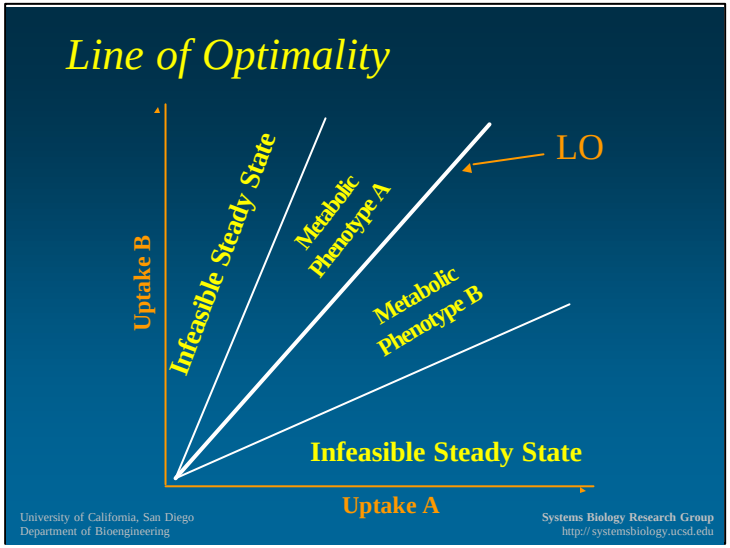
Characteristics of defined phases: The regions of the phase plane can be defined based on the contribution of the two substrates to the overall objective function:

In regions where the α value is negative, there is dual limitation of the substrates. Based on the absolute value of α , the substrate with a greater contribution toward obtaining the objective (here considered to be biomass production) can be identified. If the absolute value of α is greater than unity, the substrate along the x-axis is more valuable toward obtaining the objective, whereas if the absolute value of α is less than unity, the substrate along the y-axis is more valuable to the objective.

The regions where the isoclines are either horizontal or vertical are regions of single substrate limitation, the α value in these regions will be zero or infinite, respectively. These regions arise when the shadow price for one of the substrates goes to zero, and thus has no value to the cell.

Regions in the PhPP can also have a positive α value; these regions are termed "futile" regions. In these regions one of the substrates is inhibitory toward obtaining the objective function, and this substrate will have a positive shadow price. The metabolic operation in this region is wasteful, in that it consumes substrate that it does not need, and is thus unavailable for later utilization.

Finally, due to stoichiometric limitations, there are infeasible steady state regions in the PhPP. If the substrates are taken up at the rates represented by these points, the metabolic network is not able to obey the mass, energy, and redox constraints while generating biomass. The operation of the metabolic network can only transiently operate in such a region.

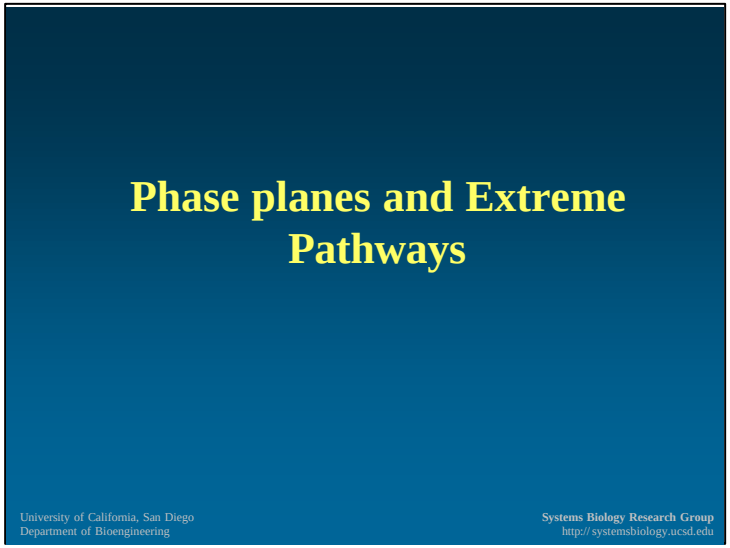


Line of Optimality:

The line of optimality is defined as the line representing the optimal relation between the two metabolic fluxes corresponding to the axis of the PhPP. For aerobic growth, this line is interpreted as the optimal amount of oxygen to be taken up to allow for the complete oxidation of the substrate.

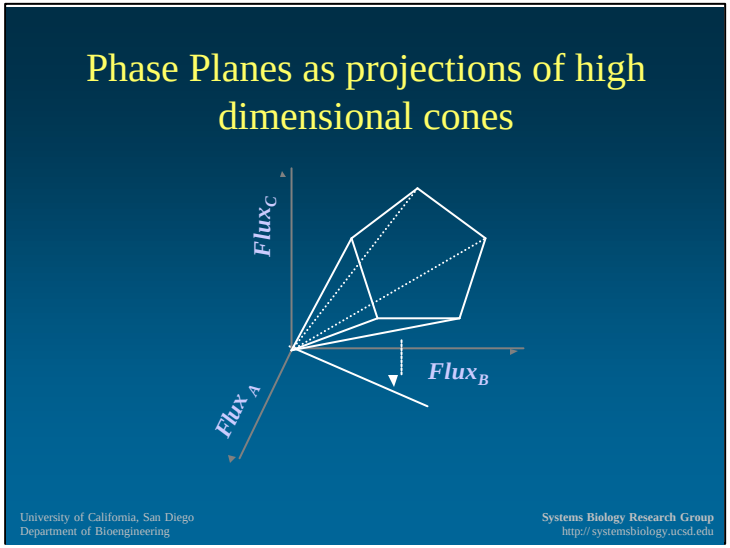
The line of optimality is determined by specifying the uptake rate of the substrate along the x-axis and allowing any value for the flux along the y-axis. LP is then used to calculate the optimal value of the objective as a function of the y-axis flux. Once the objective is determined, the corresponding flux value for the y-axis is used to plot the line of optimality (LO).

The LO defines the optimal utilization of the metabolic pathways without limitations on the availability of the substrates.



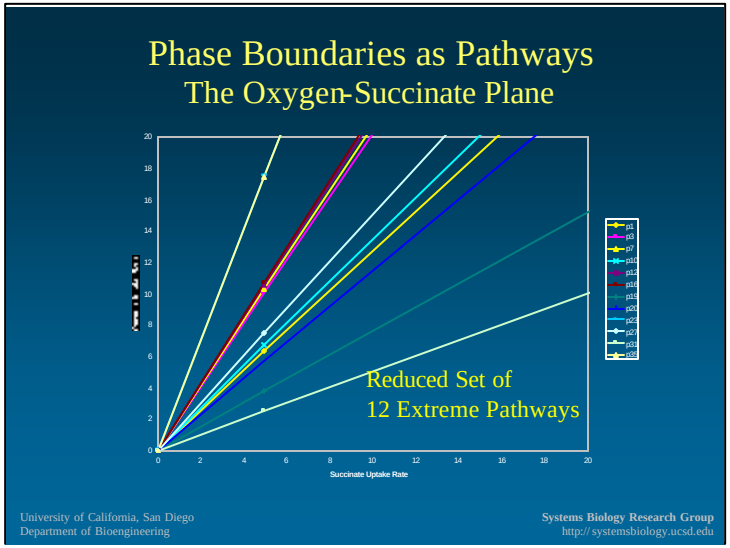
The Relationship Between Phase Planes and Extreme Pathways

In previous lectures we covered the topic of extreme pathways as the generating vectors for cones in high-dimensional spaces. It turns out that there is a close relationship between these extreme pathways and what is shown in the phase plane.



Projections of Extreme Pathways

This slide illustrates the projection of the edge of a cone onto a 2-dimensional space. The 2-dimensional space would be formed by the two uptake fluxes or any other two fluxes of interest, and the vector corresponding to the edge is drawn in that particular 2-dimensional phase plane. If that edge corresponds to an extreme pathway that is physiologically meaningful, and the cell positions itself close to it, then the data will project onto the phase plane very close to that line. This indeed corresponds to the line of optimality shown in the numerous slides before this. The line of optimality is an edge on the cone in a high dimension.



The Oxygen-Succinate Extreme Pathways in the Phase Plane

This slide shows the projection of a number of extreme pathways calculated from the core *E. coli* model with succinate as the carbon source. We see that all these pathways form a straight line in the phase plane. One of these pathways corresponds to the line of optimality, and it in return corresponds to the extreme pathway with the highest biomass yield.

Summary

- Flux balance and capacity constraints form closed polyhedral spaces
- Linear programming can be used to find optimal solutions in this space
- Does require the statement of an objective function and the solution will be the corresponding optimal phenotype
- Optimal solutions lie at the edge of the polyhedra
- Shadow prices and reduced costs are used to characterize the optimal solution
- A series of LP can be solved to represent a continuous variation in a parameter of interest
- All possible combinations of the values of two parameters leads to the definition of a phase plane
- The boundaries in the phase plane are edges on the polyhedral cone projected into the plane. Thus, the boundaries represent systemic pathways

References

- Varma, A. and Palsson, B.O., "Predictions for oxygen supply control to enhance population stability of engineered production strains," *Biotechnol Bioeng*, **43**: 275-285 (1994).
- Varma, A. and Palsson, B.O., "Parametric sensitivity of stoichiometric flux balance models applied to wild-type *Escherichia coli* metabolism," *Biotechnol Bioeng*, **45**: 69-79 (1995).
- Pramanik, J. and Keasling, J.D., "Stoichiometric model of *Escherichia coli* metabolism: Incorporation of growth-rate dependent biomass composition and and mechanistic energy requirements," *Biotechnol Bioeng*, **56**: 398-421 (1997).
- Bonarius, H. P. J., Schmid, G. & Tramper, J. (1997) *Trends in Biotechnology* **15,** 308-314.
- J.S. Edwards, R. Ramakrishna, C.H. Schilling, and B.O. Palsson, "Metabolic Flux Balance Analysis," *Metabolic Engineering* **2**: 13-57, (1999), S.Y. Lee, Papoutsakis, E.T., Eds; Springer-Verlag: New York.
- J.S. Edwards and B.O. Palsson, "Systems Properties of the *Haemophilus influenzae* Rd Metabolic Genotype," *The Journal of Biological Chemistry*, **274**: 17410-17416 (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).

References

- Edwards J.S., Ramakrishna R., and Palsson, B.O., "Characterizing the metabolic phenotype: a phenotypic phase plane analysis," *Biotech. & Bioeng.*, **77**: 27-36 (2002).
- Schilling C.H., Edwards, J.S., Letscher, D.L., and Palsson, B.O., "Combining pathway analysis with flux balance analysis for the comprehensive study of metabolic systems," *Biotech. & Bioeng.*, **71**: 286-306 (2001).
- Edwards, J.S. and Palsson, B.O., "Metabolic flux balance analysis and the in silico analysis of Escherichia coli K-12 gene deletions," *BMC Bioinformatics*, 1(1):1 (2000).
- Burgard, A.P and Maranas, C.D., "Probing the performance limits of the Escherichia coli metabolic network subject to gene additions or deletions," *Biotech. & Bioeng.*, **74(5)**: 364-375 (2001).
- Burgard, A.P, Vaidyaraman, S., and Maranas, C.D., "Minimal reaction sets for Escherichia coli metabolism under different growth requirements and uptake environments," *Biotechnol. Prog.*, 17: 791-797 (2001).
- Covert, M., Schilling, C.H., and Palsson, B.O., "Regulation of gene expression in flux balance models of metabolism," *J. Theor. Biol.*, 213:73-78 (2001).
- Edwards, J.S., Covert, M., and Palsson B.O., "Metabolic modelling of microbes: the flux balance approach," Environmental Microbiology, 4(3):133-140 (2002).

References

- Papoutsakis, E.T., "Equations and calculations for fermentations of butyric acid bacteria," *Biotechnol Bioeng*, **26**: 174-187 (1984).
- Papoutsakis, E. and Meyer, C., "Equations and calculations of product yields and preferred pathways for butanediol and mixed-acid fermentations," *Biotechnol Bioeng*, **27**: 50-66 (1985).
- Fell, D.A. and Small, J.A., "Fat synthesis in adipose tissue. An examination of stoichiometric constraints," *J. Biochem*, **238**: 781-786 (1986).
- Majewski, R.A., and Domach, M.M., "Simple constrained optimization view of acetate overflow in *E. coli*," *Biotechnol Bioeng*, **35**: 732-738 (1990).
- Savinell, J.M., and Palsson, B.O., "Network analysis of intermediary metabolism using linear optimization. II. Interpretation of hybridoma cell metabolism," *J Theor Biol*, **154**: 455-473 (1992).
- Varma, A. and Palsson, B.O., "Metabolic capabilities of *Escherichia coli*. I. Synthesis of biosynthetic precursors and cofactors," *J Theor Biol*, **165**:477-502 (1993).
- Varma, A. and Palsson, B.O., "Stoichiometric flux balance models quantitatively predict growth and metabolic by -product secretion in wild-type *Escherichia coli* W3110," *Appl Environ Microbiol*, **60**: 3724-3731 (1994).