Lecture: Computational Systems Biology
Universität des Saarlandes, SS 2012

05 Structural analysis

Dr. Jürgen Pahle

29.5.2012

Recap

- Standards
  - Systems Biology Markup Language (SBML)
  - Systems Biology Graphical Notation (SBGN)
  - Minimal information required in the annotation of models (MIRIAM)
- Software
  - COPASI, CellDesigner, etc.
- Databases
  - Pathway databases (KEGG, Reactome, ...)
  - Databases providing kinetic parameters (BRENDA, Sabio-RK)
  - Model databases (Biomodels, JWS)

Stoichiometry (reminder)

- **Proportions** of substrate and product molecules in a reaction
- Assignment of stoichiometric coefficients is **not unique**
- Example: $S_1 + S_2 \leftrightarrow 2\,P$

  Stoichiometric coefficients could be
  -1, -1, 2, or

  -1/2, -1/2, 1, or

  1, 1, -2

System equations / balance equations

$$\frac{dS_i}{dt} = \sum_{j=1}^{r} n_{ij} v_j = \boldsymbol{N}\,\boldsymbol{v}$$

Stoichiometric matrix ***N***
rows → chemical species
columns → reactions

Vector of rates ***v***

Direction of reaction arbitrarily assigned
Negative rate is possible (→ net flux in reverse direction)

System representation

System is fully specified by

- Concentration values ***S***
- Rates ***v*** (kinetic functions)
- Parameter set ***p***
- Stoichiometric matrix ***N***
- In case of a steady state, we also have a vector of steady state fluxes ***J***

Information contained in ***N***

- Which combination of fluxes is possible in a steady state?
- Dead ends, unbranched reaction pathways
- Conservation relations / conserved moieties

Information contained in ***N***

- In a steady state (SS)

$$\frac{d\,\boldsymbol{S}}{dt} = \boldsymbol{N}\,\boldsymbol{v} = 0$$

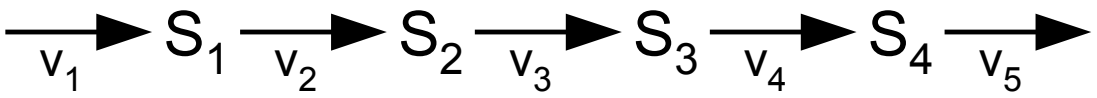
- Linear equation system for the rates in SS. Non-trivial solution when *Rank(**N**) < r*

Kernel matrix ***K***

- Kernel matrix ***K*** fulfilling $\boldsymbol{N}\,\boldsymbol{K} = 0$

  shows the linear dependencies
- Choice of kernel is not unique
- Can be found using the Gauss algorithm
- It contains as columns *r - Rank(**N**)* basis vectors
- Every possible set of steady state fluxes ***J*** is a linear combination of the columns $\boldsymbol{k}_i$ of ***K***

Example



$$\boldsymbol{N} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & -1 \end{pmatrix}$$

r = *5* reactions

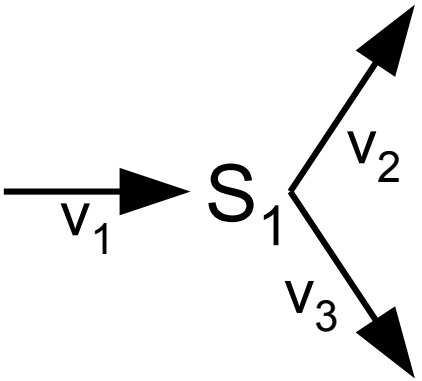
*Rank(**N**) = 4*

Therefore, kernel matrix must contain *1 = 5 - 4* basis vectors

$$k = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 \end{pmatrix}^T$$

In SS flux through all reactions must be equal

Example



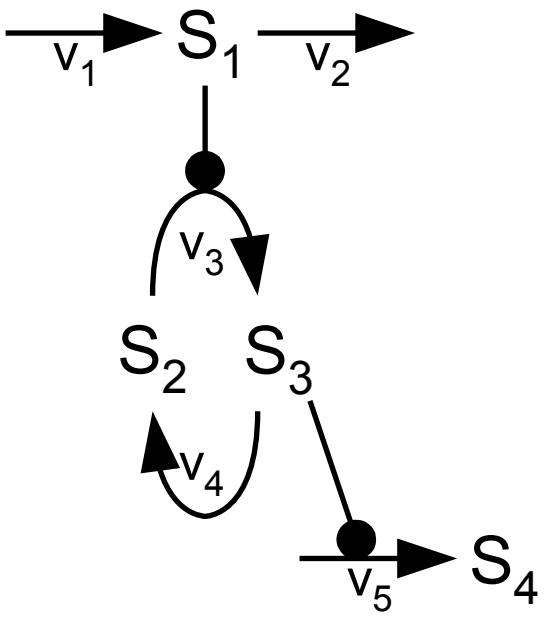
$$\boldsymbol{N} = \begin{pmatrix} 1 & -1 & -1 \end{pmatrix}$$

$$\boldsymbol{K} = \begin{pmatrix} \boldsymbol{k}_1 \boldsymbol{k}_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

r = 3 reactions
*Rank(**N**) = 1*
Therefore, kernel matrix must contain 2 = *3 - 1* basis vectors

For SS fluxes $\boldsymbol{J} = \alpha_1 \cdot \boldsymbol{k}_1 + \alpha \cdot \boldsymbol{k}_2$ *holds*

Example



$$\boldsymbol{N} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

r = 5 reactions

*Rank(**N**) = 3*

Therefore, kernel matrix must contain *2 = 5 - 3* basis vectors

$$\boldsymbol{k}_1 = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \end{pmatrix}^T$$

$$\boldsymbol{k}_2 = \begin{pmatrix} 0 & 0 & 1 & 1 & 0 \end{pmatrix}^T$$

In SS production and degradation of S_1 are balanced ($J_1 = J_2$), fluxes through cycle must be equal ($J_3 = J_4$), and J_5 must be zero

Kernel of ***N***

- **Equilibrium reactions**: zero entries in each basis vector
  → net flow always zero in SS
- **Unbranched path**: same entries for a set of rows in all basis vectors
  → net rate of corresponding reactions is always the same in a SS

(Ir-)reversibility

Some reactions can be considered as being (practically) irreversible

→ no change in stoichiometric matrix ***N***

But, kernel matrix ***K*** will be restricted by this

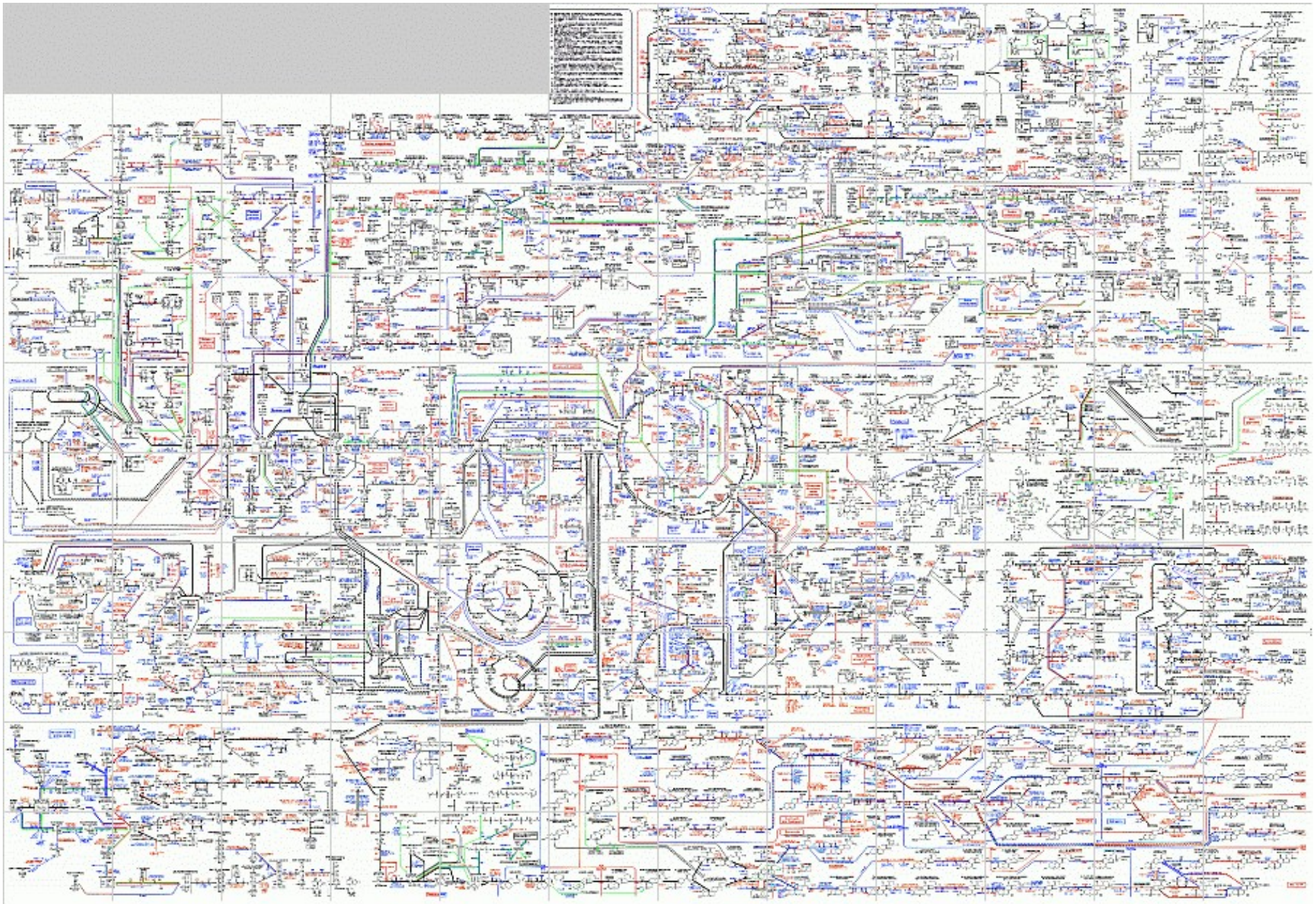
→ Some values may not become negative (or positive)

Pathways

"Set of subsequent reactions that are linked by common metabolites"

Examples: glycolyis, amino acid synthesis pathways, ...

Pathways?



taken from http://www.expasy.ch

Pathways

- Difficult to separate looking at the complete **network** of reactions

- Which direct routes connect external metabolites or go in cycles?

  → "flux modes"

Elementary flux modes

- **Minimal** sets of reactions that allow steady state dynamics
- All steady state fluxes are linear combinations of elementary flux modes
- Calculation only requires structural information (stoichiometry matrix)

Flux modes

- Definition

$$\boldsymbol{M} = \left\{ \boldsymbol{v} \in R^r \middle| \boldsymbol{v} = \lambda \cdot \hat{\boldsymbol{v}}\,,\, \lambda > 0 \right\}$$

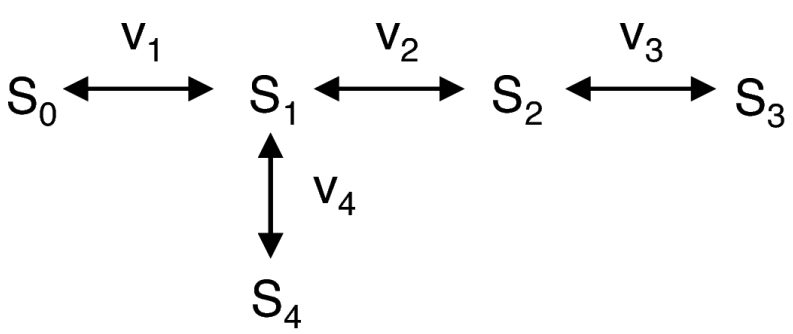
with $\hat{\boldsymbol{v}}$ an *r*-dimensional vector (not the null vector) such that

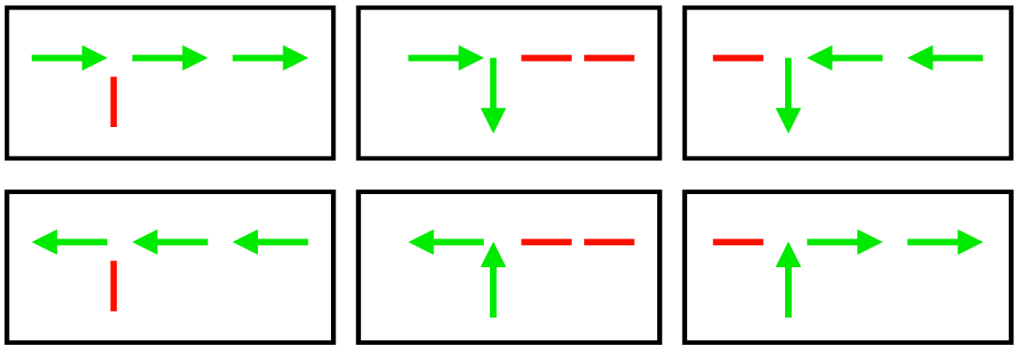
  - it corresponds to a steady state
  - all sign restrictions (irreversible reactions) are fullfilled

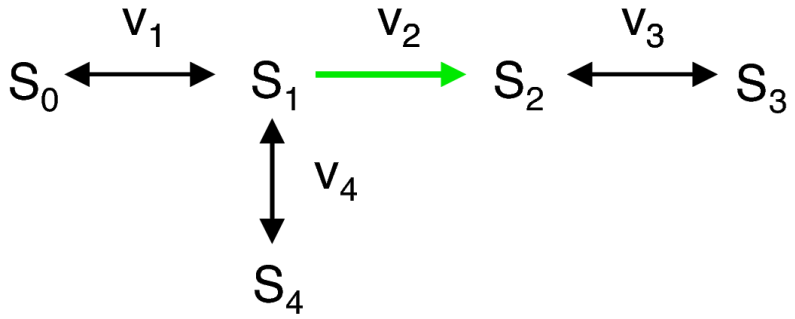
- Can be calculated using COPASI or specialised software, such as Metatool

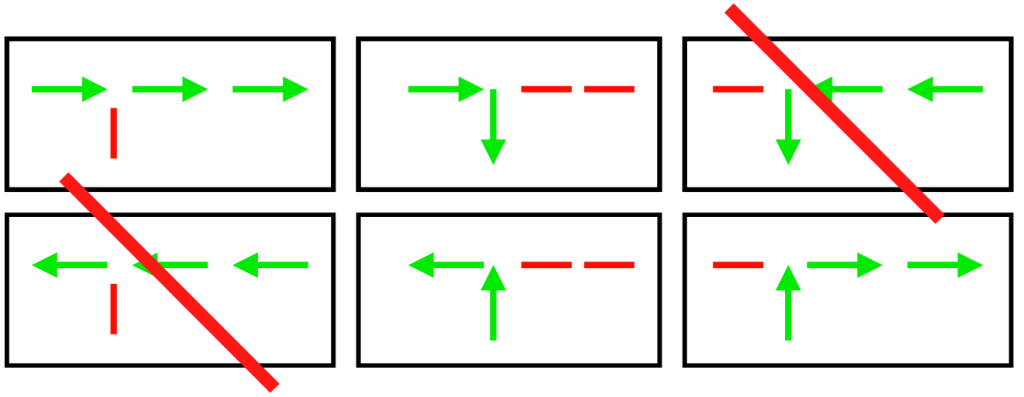
Examples

Elementary Flux Modes









Klipp - Systems Biology
ISBN: 978-3-527-31874-2 fig-02-06

Flux modes

- A flux mode ***M*** comprising ***v*** is called reversible if the set ***M'*** comprising ***-v*** is also a flux mode
- A flux mode is an **Elementary flux mode** (EFM) if it uses a minimal set of reactions and cannot be further decomposed
- An elementary flux mode can be interpreted as a minimal set of enzymes that could operate at steady state (with all irreversible reactions used in the appropriate direction)
- Number of EFMs ≥ number of basis vectors of the null space

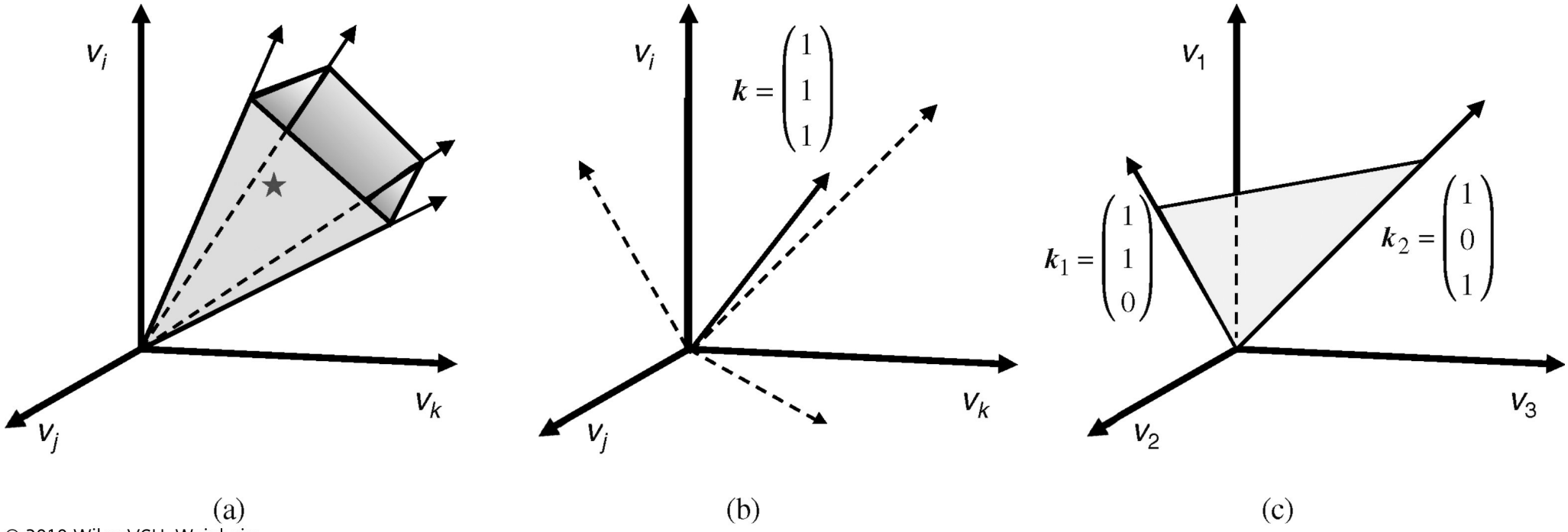
Convex flux cone

Split up all reactions into forward and backward direction → uni-directional representation

SS → flux vector is element of null space of ***N*** spanned by kernel matrix ***K***. Rows of ***K*** can be interpreted as hyperplanes in flux space. The intersection of these hyperplanes forms the null space

Convex cone in flux space formed by basis vectors and considering irreversibility

Convex flux cone



Klipp - Systems Biology
ISBN: 978-3-527-31874-2 fig-02-07

Stoichiometric analyses

Elementary flux mode analysis (and the related concept of **extreme pathways**) can be used:

- to understand the range of metabolic pathways in a network
- to test a set of enzymes for the production of a desired product
- to detect nonredundant pathways
- to analyse enzyme deficiencies
- in an application called **Flux balance analysis (FBA)** to find optimal flux distributions

Conservation relations

- If a substance is neither added to/produced nor removed/degraded from the reaction system its total concentration stays constant.
  Example: enzyme concentration in Michaelis-Menten reaction mechanism

- This also holds if the substance builds complexes with other compounds or for (conserved) subparts of molecules (moieties)

Conservation relations

Consider matrix $\boldsymbol{G}$ with

$$\boldsymbol{G}\,\boldsymbol{N} = 0$$

then

$$\boldsymbol{G}\,\dot{\boldsymbol{S}} = \boldsymbol{G}\,\boldsymbol{N}\,\boldsymbol{v} = 0$$

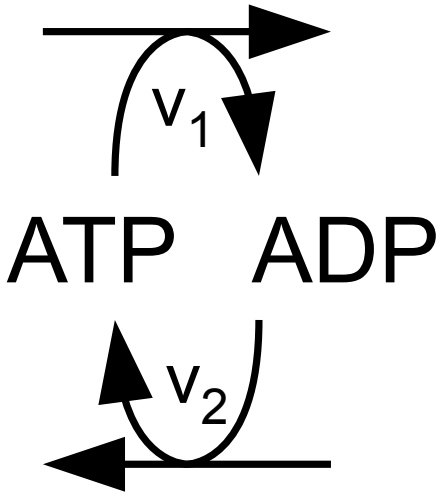
Integrating leads to

$$\boldsymbol{G}\,\boldsymbol{S} = constant$$

Conservation relations

- Number of independent rows of $\boldsymbol{G}$ equals n - $Rank(\boldsymbol{N})$ with n the number of chemical species in the model
- $\boldsymbol{G}^T$ is the kernel matrix of $\boldsymbol{N}^T$
  → Conservation relations correspond to the null space of $\boldsymbol{N}^T$
- It can be found using the Gauss algorithm
- $\boldsymbol{G}$ is not unique, but every linear combination of its rows is again a valid solution

Example



$$\boldsymbol{S} = \left(ATP\ ADP \right)^T$$

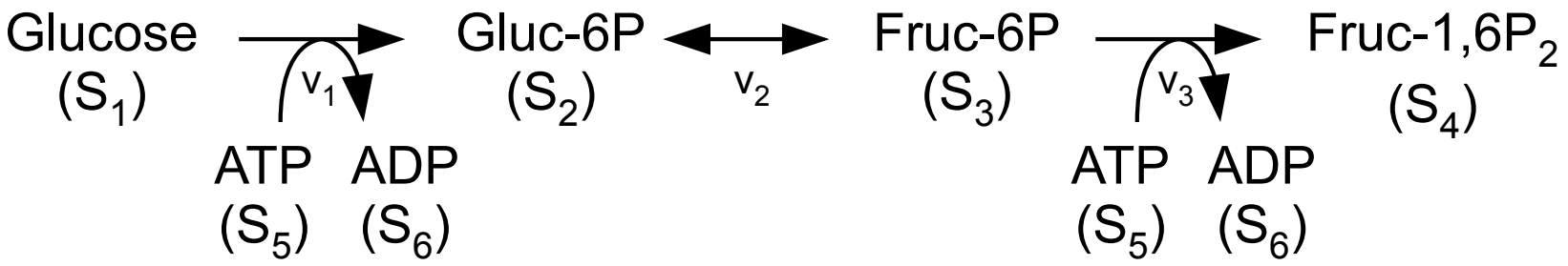
$$\boldsymbol{N} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}$$

$$\boldsymbol{G} = \begin{pmatrix} 1 & 1 \end{pmatrix}$$

ATP + *ADP* = *constant*

The constant is dependent on the initial condition!

Example



$$\boldsymbol{N}^{\boldsymbol{T}}=\begin{pmatrix}-1 & 1 & 0 & 0 & -1 & 1\\ 0 & -1 & 1 & 0 & 0 & 0\\ 0 & 0 & -1 & 1 & -1 & 1\end{pmatrix}$$

$$\boldsymbol{G}=\begin{pmatrix}2 & 1 & 1 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 & 1 & 1\\ 1 & 1 & 1 & 1 & 0 & 0\end{pmatrix}=\begin{pmatrix}g_1\\ g_2\\ g_3\end{pmatrix}$$

g_2: conservation of *ATP* + *ADP*

g_3: conservation of sugars

g_1: with $g_4=-g_1+3\cdot g_2+2\cdot g_3=\begin{pmatrix}0 & 1 & 1 & 2 & 3 & 2\end{pmatrix}$

→ conservation of phosphate groups

Reducing the system

- Idea: Eliminate linearly dependent ODEs (species) and replace them by algebraic equations
- First, reorder

$$\boldsymbol{N} = \begin{pmatrix} \boldsymbol{N}_R \\ \boldsymbol{N}' \end{pmatrix} = \boldsymbol{L}\,\boldsymbol{N}_R = \begin{pmatrix} \boldsymbol{I}_{Rank(\boldsymbol{N})} \\ \boldsymbol{L}' \end{pmatrix} \boldsymbol{N}_R$$

L is called "link matrix"

- Then

$$\dot{\boldsymbol{S}} = \begin{pmatrix} \dot{\boldsymbol{S}}_{indep} \\ \dot{\boldsymbol{S}}_{dep} \end{pmatrix} = \begin{pmatrix} \boldsymbol{I}_{Rank(\boldsymbol{N})} \\ \boldsymbol{L}' \end{pmatrix} \boldsymbol{N}_R \boldsymbol{v}$$

with

$$\dot{\boldsymbol{S}}_{dep} = \boldsymbol{L}' \cdot \dot{\boldsymbol{S}}_{indep}$$

Reducing the system (cont.)

- Integration leads to

$$\boldsymbol{S}_{dep} = \boldsymbol{L}' \cdot \boldsymbol{S}_{indep} + constant$$

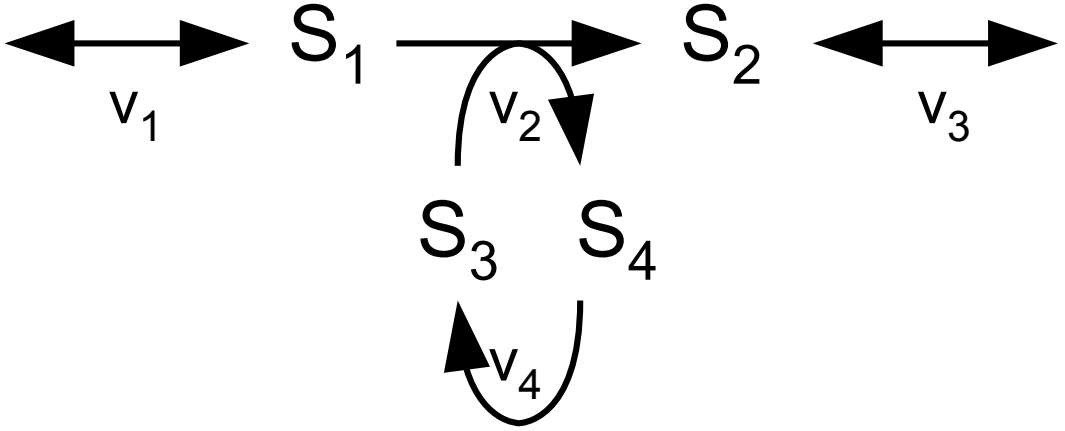
- This holds during the entire time course. Therefore, we can replace the full system by the reduced/simplified system

$$\dot{\boldsymbol{S}}_{indep} = \boldsymbol{N}\,\boldsymbol{v}$$

  supplemented with the set of algebraic equations above

- Reduced system plus algebraic equations to express the dependent species is equivalent to the full system.

Example



$$\boldsymbol{N}=\begin{pmatrix}1 & -1 & 0 & 0\\ 0 & 1 & -1 & 0\\ 0 & -1 & 0 & 1\\ 0 & 1 & 0 & -1\end{pmatrix}$$

$$\boldsymbol{N}_R=\begin{pmatrix}1 & -1 & 0 & 0\\ 0 & 1 & -1 & 0\\ 0 & -1 & 0 & 1\end{pmatrix}$$

$$\boldsymbol{L}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & -1\end{pmatrix}$$

$$\boldsymbol{L}'=\begin{pmatrix}0 & 0 & -1\end{pmatrix}$$

$$\boldsymbol{G}=\begin{pmatrix}0 & 0 & 1 & 1\end{pmatrix}$$

Reduced system:

$$\dot{S}_1=v_1-v_2$$
$$\dot{S}_2=v_2-v_3$$
$$\dot{S}_3=v_4-v_2$$
$$S_4=constant-S_3$$

Conservation relations (summary)

Linear combinations of species whose overall concentration stays constant

Might not be obvious from visual inspection of the network diagram only

Important for reducing the system