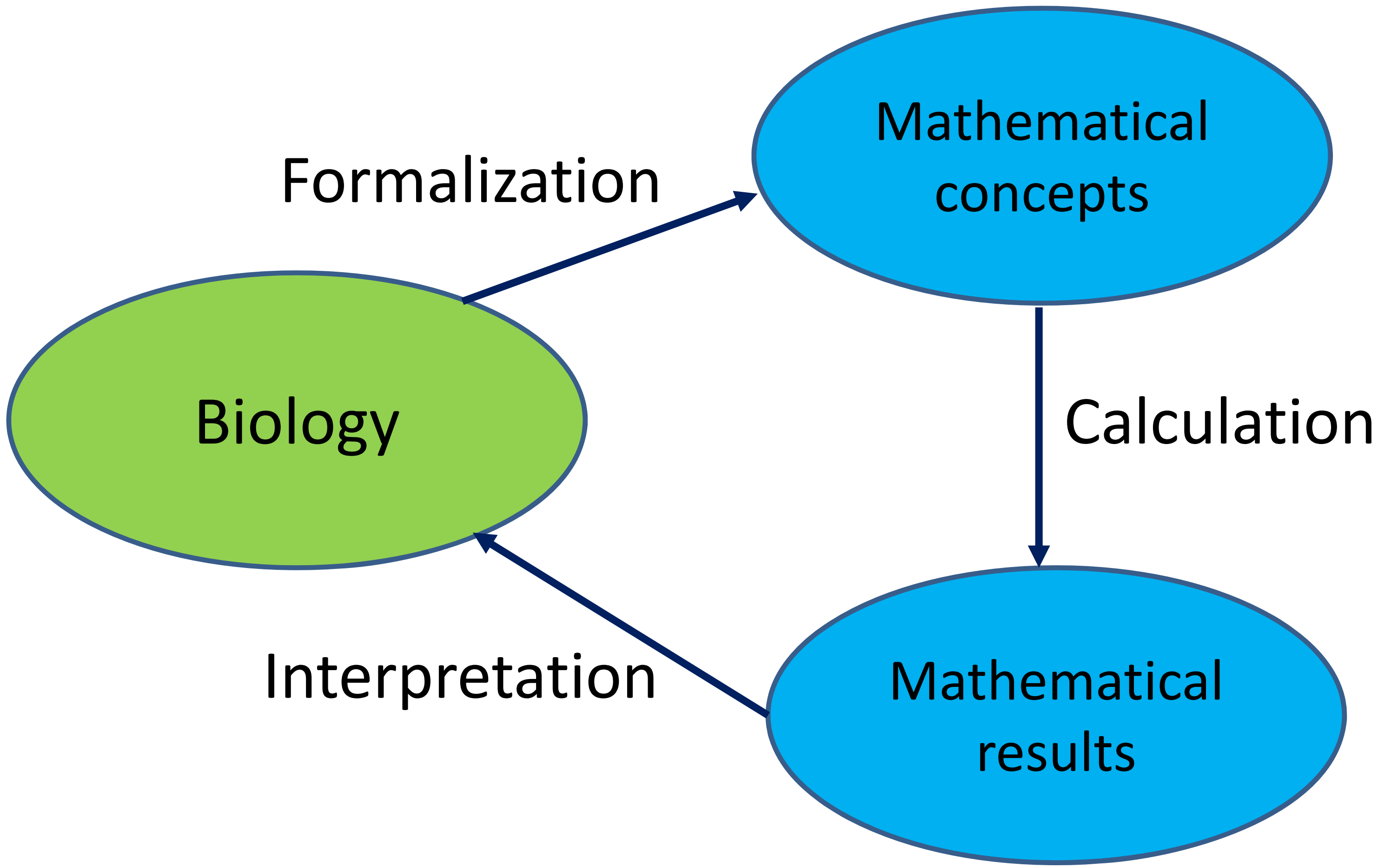
FRIEDRICH-SCHILLER-UNIVERSITÄT JENA

Algebra meets Biology

Stefan Schuster

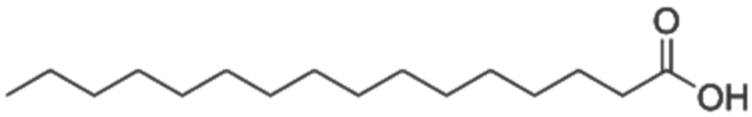
Dept. of Bioinformatics, Jena University Germany


Mathematical concepts
Formalization
Biology
Calculation
Interpretation
Mathematical results

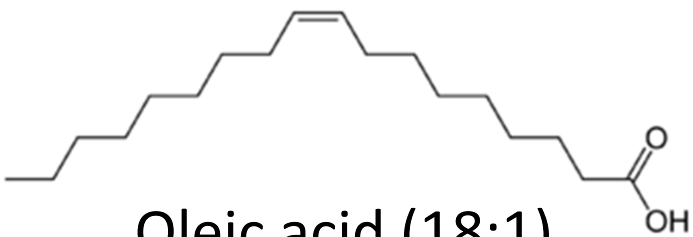
Two topics

- 1st topic: Enumerating fatty acids
- 2nd topic: Calcium oscillations

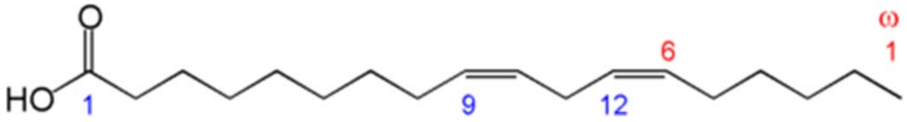
Fatty acids



Examples: Palmitic acid (16:0)



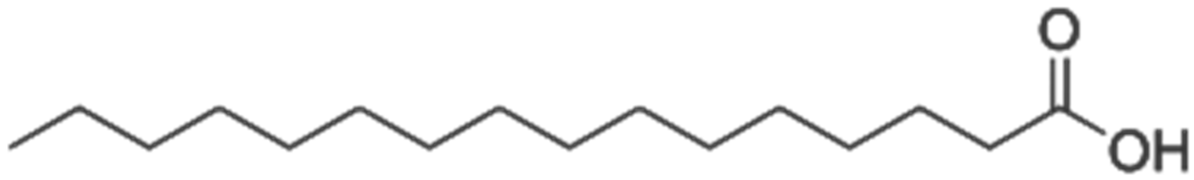
Oleic acid (18:1)



Linoleic acid (18:2)

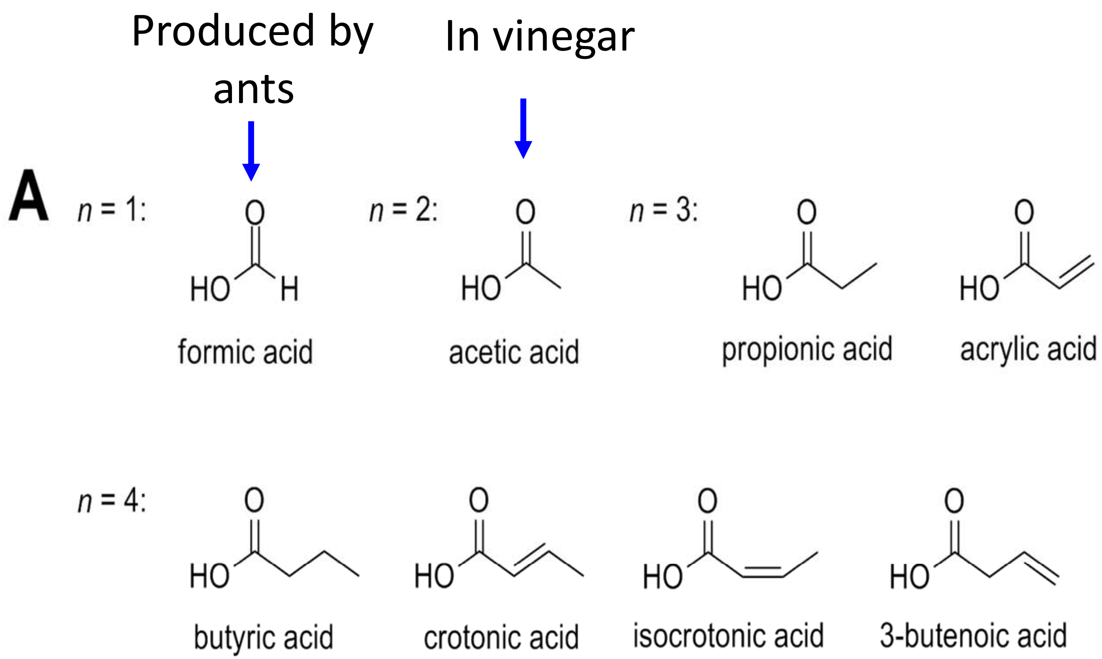
Fatty acids

- Crucial importance for all living beings
- Triglycerides = energy and carbon stores
- Phospholipids in biomembranes
- Signalling substances such as diacylglycerol
- Biomarkers



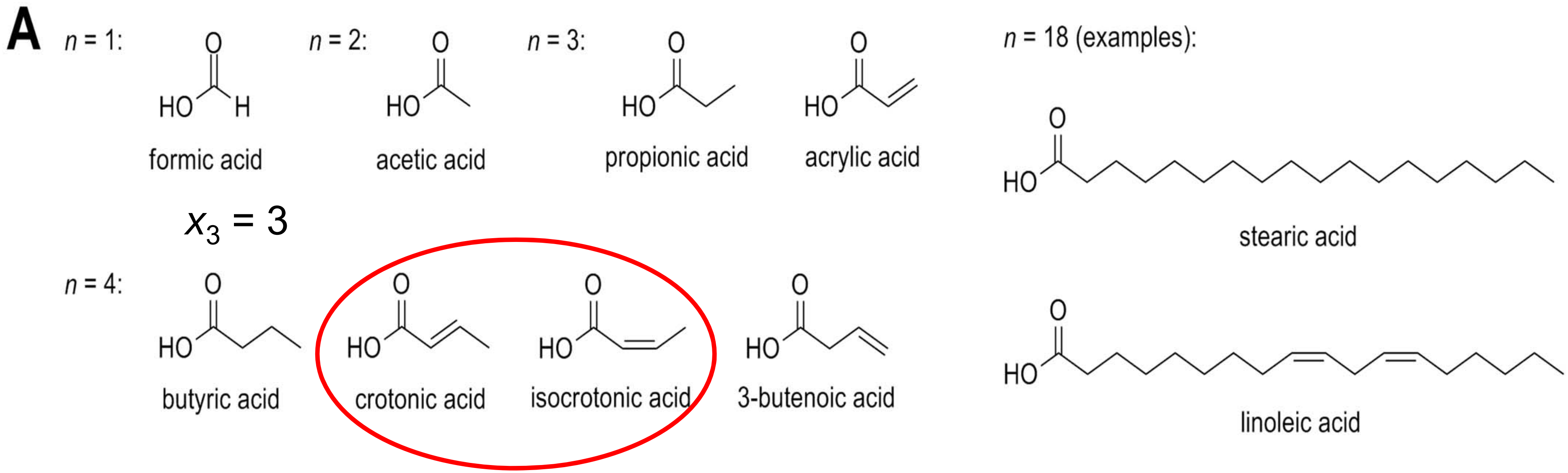
… and short-chain fatty acids (SCFAs)

Play role in gut microbiome



First case: Neglecting *cis/trans* isomerism

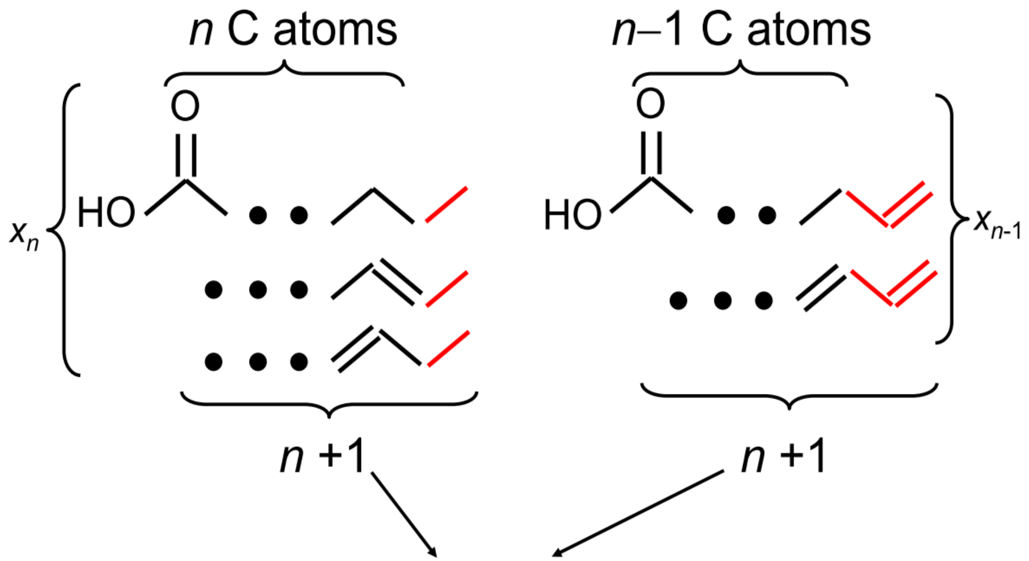
$x_1 = 1$ $x_2 = 1$ $x_3 = 2$



Cis/trans isomers are combined.

We exclude allenic FAs (two neighbouring double bonds) because they are rare.

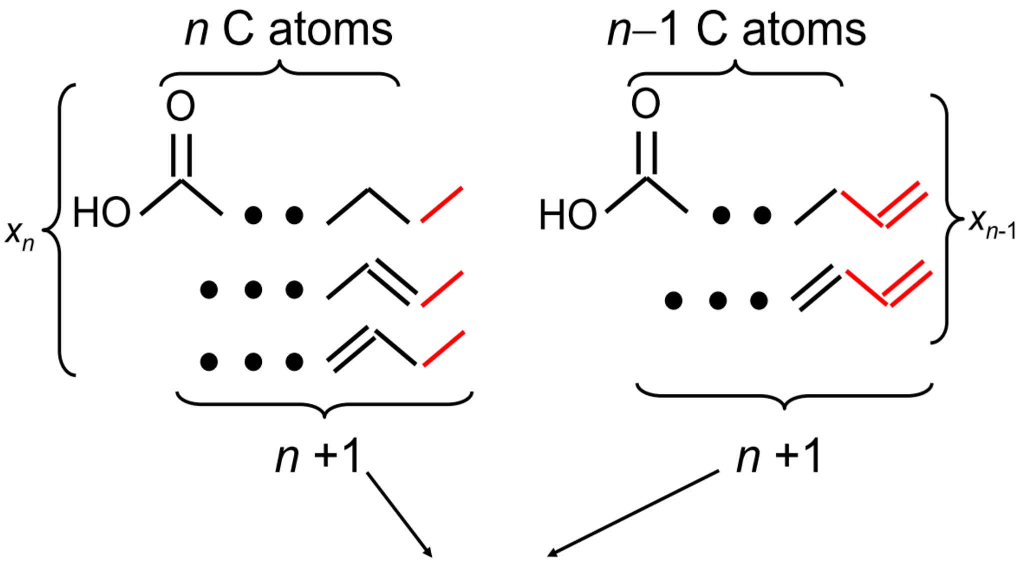
Recursion



$$x_{n+1} = x_n + x_{n-1}\,,\ x_0 = x_1 = 1.$$

(for fatty acids: initial values $x_1 = x_2 = 1$)

Recursion



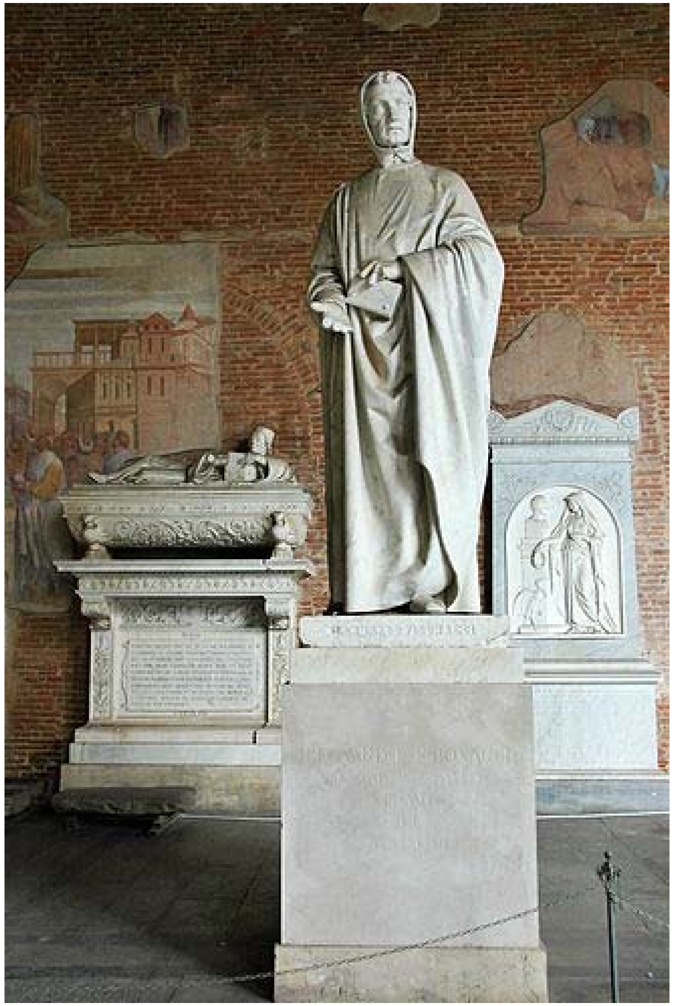
$$x_{n+1} = x_n + x_{n-1}\,,\; x_0 = x_1 = 1.$$

(for fatty acids: initial values $x_1 = x_2 = 1$)

This leads to Fibonacci numbers: 1, 1, 2, 3, 5, 8, 13, 21, …

S. Schuster, M. Fichtner, S. Sasso: Use of Fibonacci numbers in lipidomics - Enumerating various classes of fatty acids. *Sci. Rep.* 7 (2017) 39821

Leonardo Pisano (Fibonacci)



Liber abaci 1202

Pictures: Wikipedia

Sanskrit prosody
and mathematical biology

- Pingala (पिङ्गल, probably ca. 400 BC) in ancient India
- Author of the **Chandaḥśāstra**, the earliest known Sanskrit treatise on prosody.
- First known description of binary numeral system and the (later so-called) **Fibonacci numbers** in systematic enumeration of meters, sequence there called “matrameru”

Picture: Wikipedia

Indian mathematics and Sanskrit prosody

- Short and long syllables S (· ·) and L (– twice as long)
- How many sequences of · and – with exactly *m* beats?
- 0 beat: 1 possibility
- 1 beat: 1 possibility: ·
- 2 beats: 2 possibilities: · · ; –
- 3 beats: 3 " : · · ·; – ·; · –
- 4 beats: 5 " : · · · ·; – · ·; – – · ;...
- Interval between S and S (· ·) → single bond, L (–) → double bond

Interesting properties of Fibonacci (matrameru) numbers

- 1, 1, 2, 3, 5, 8, 13, 21, 34, … → every 3rd n umber is even
- 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144… → every 4th number is divisible by 3.
- Every *k*th number of the sequence is a multiple of F_k (starting with F_1 = 1, F_2 = 2…)

Binet‘s formula

- Explicit formula
- Exponential ansatz: $x_n = a\lambda^n$
- With recursion formula, this leads to Binet‘s formula

$$x_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n$$

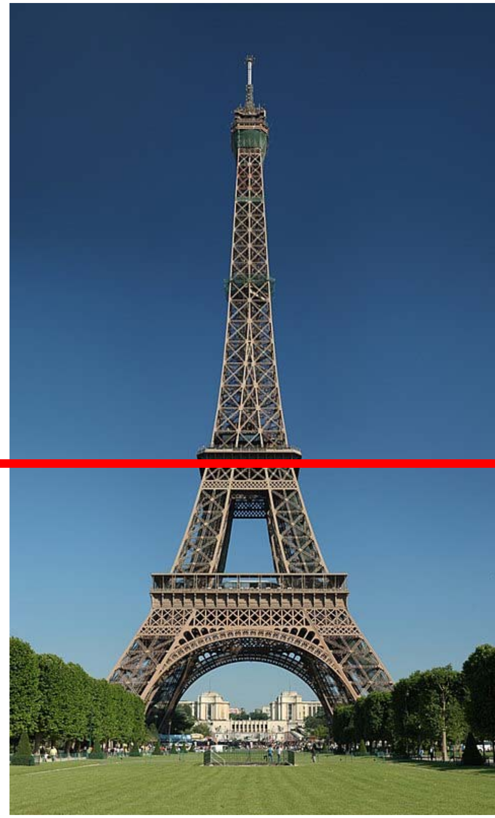
- Can be simplified to

$$x_n = \mathrm{round}\left(\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n\right)$$

- Ratio of Golden Section.
- First discovered by Abraham de Moivre (1667 - 1754) one century before Binet

Golden Section

- (1+SQRT(5))/2 = 1.618…
- Numbers of FAs grow asymptotically exponentially with the basis of 1.618…
- Investing one more carbon, an organism can increase variability of FAs approximately by Golden Ratio



Picture: Wikipedia

Alternative way of calculation

$$\sum_{k=0}^{m} \binom{n-k-1}{k} = n\text{-th Fibonacci number,}$$

- Lucas‘ Formula
- *m* = largest integer <= (n-1)/2
- n–k–1 = number of positions where double bonds can be situated
- Interesting case: limiting *m* by *q* from above. Asymptotic behaviour?

Fibonacci (matrameru) numbers in phyllotaxis

- *t = number of turns, n = number of leaves*
- *t*/*n*=1/2 (Opposite distichous leaves), e.g. elm tree
- *t*/*n*=2/3, e.g. beech tree, blueberry
- *t*/*n*=3/5, e.g. oak tree
- *t*/*n*=5/8, e.g. poplar tree, roses
- *t*/*n*=8/13, e.g. plum tree, some willow tree species
- *t*/*n*=Golden section, e.g. agavas, sunflower, *Dracaena*, pine needles on young branches

Picture: Wikipedia

Back to fatty acids
Second case: Considering *cis/trans* isomerism

- Adding the (n+1)-th carbon, there are two cases:

  a) Single bond at position n. Then two possibilities: Adding single or double bond.

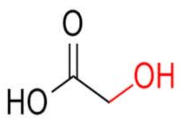
  b) Double bond at position n. Again two possibilities: Adding single bond in *cis* or in *trans* conformation.

- In both cases: $u_{n+1} = 2u_n$.
- Explicit formula: $u_n = 2^{n-2}$ with exception $u_1 = 1$.

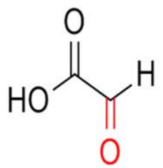
Modified fatty acids

Oxo or hydroxy groups (important in polyketides).
Neither of these can be adjacent to a double bond.
Keto-enol tautomerism: =C-OH → -C=0
Neglecting stereoisomerism at hydroxy groups.

n = 2:

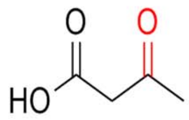


glycolic acid



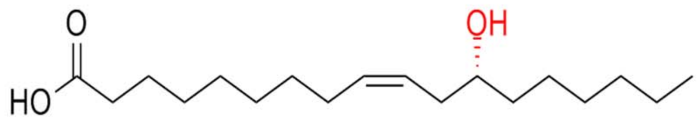
glyoxylic acid

n = 4 (example):



acetoacetic acid

n = 18 (example):



ricinoleic acid

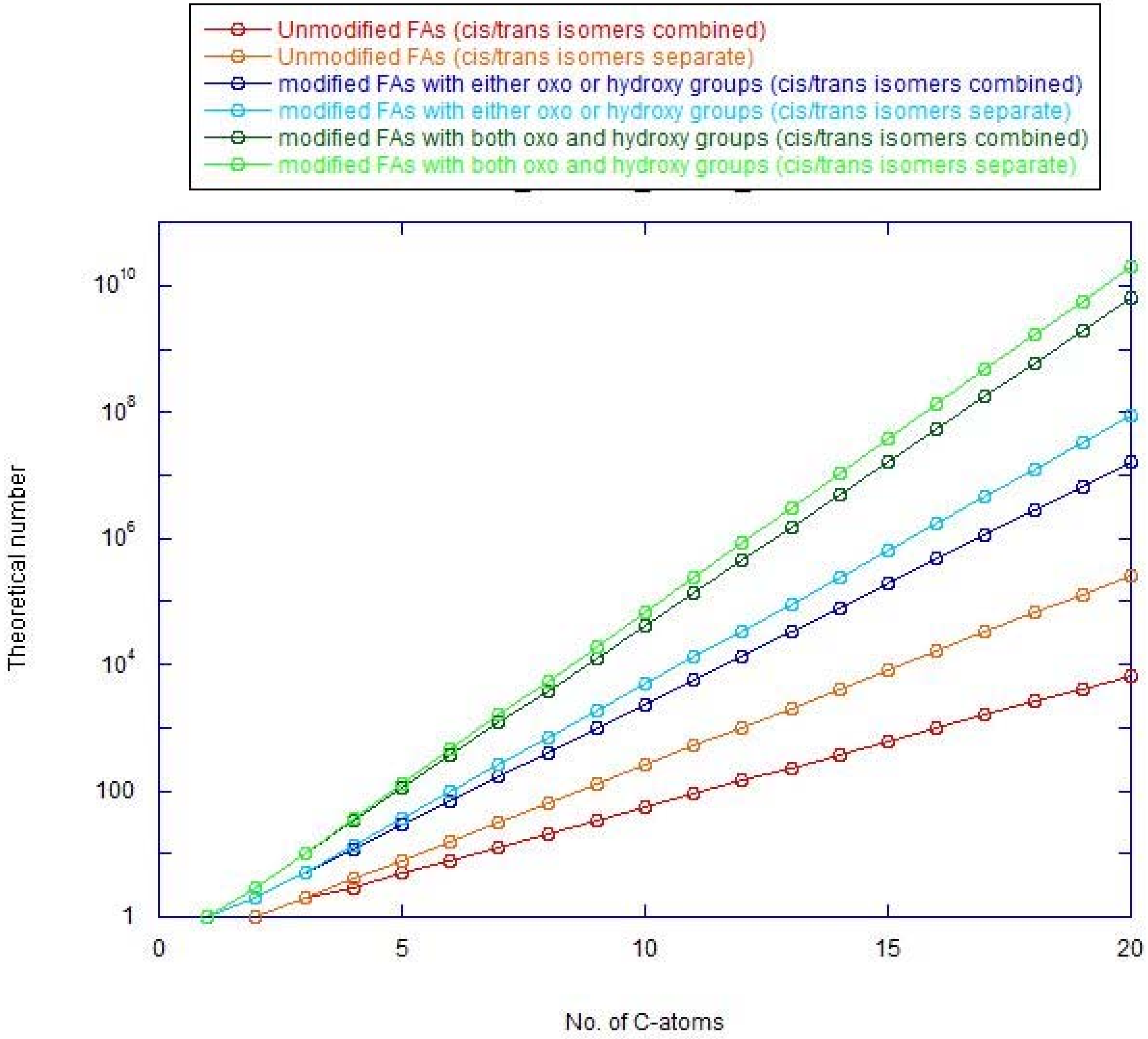
Recursion in the case of functional group(s)

- One functional group:
  - Leads to 2-Fibonacci numbers (Pell numbers)
  - $y_{n+1} = 2y_n + y_{n-1}$
  - 1, 2, 5, 12, 29, 70, …
  - Basis (1 + SQRT(2)), proport. to Silver section
- Two functional groups:
  - 3-Fibonacci numbers
  - $z_{n+1} = 3z_n + z_{n-1}$
  - 1, 3, 10, 33, 109, 360
  - Basis (3 + SQRT(13)), proport. to Bronze section
- All in www.oeis.org

Cis-/trans isomers considered separately, functional groups

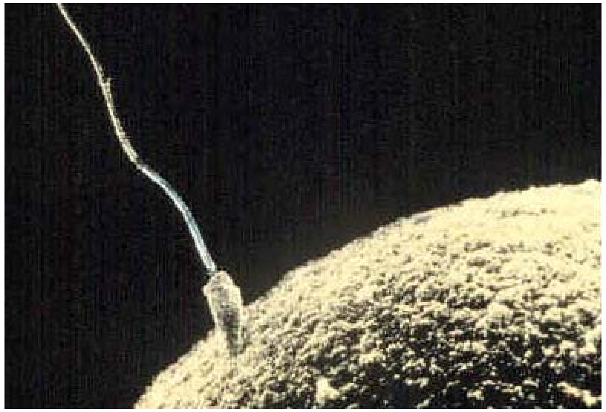
- One functional group: $v_{n+1} = 2v_n + 2v_{n-1}$
- 1, 2, 5, 14, 38… (A052945 in www.oeis.org)

- Two functional groups: $w_{n+1} = 3w_n + 2w_{n-1}$
- 1, 3, 10, 36, 128, … (not yet in www.oeis.org)


Unmodified FAs (cis/trans isomers combined)
Unmodified FAs (cis/trans isomers separate)
modified FAs with either oxo or hydroxy groups (cis/trans isomers combined)
modified FAs with either oxo or hydroxy groups (cis/trans isomers separate)
modified FAs with both oxo and hydroxy groups (cis/trans isomers combined)
modified FAs with both oxo and hydroxy groups (cis/trans isomers separate)
Theoretical number
10^{10}
10^8
10^6
10^4
100
1
0
5
10
15
20
No. of C-atoms

2nd topic: Calcium oscillations

- Oscillations of intracellular calcium ions are important in signal transduction both in excitable and nonexcitable cells (e.g. egg cells)



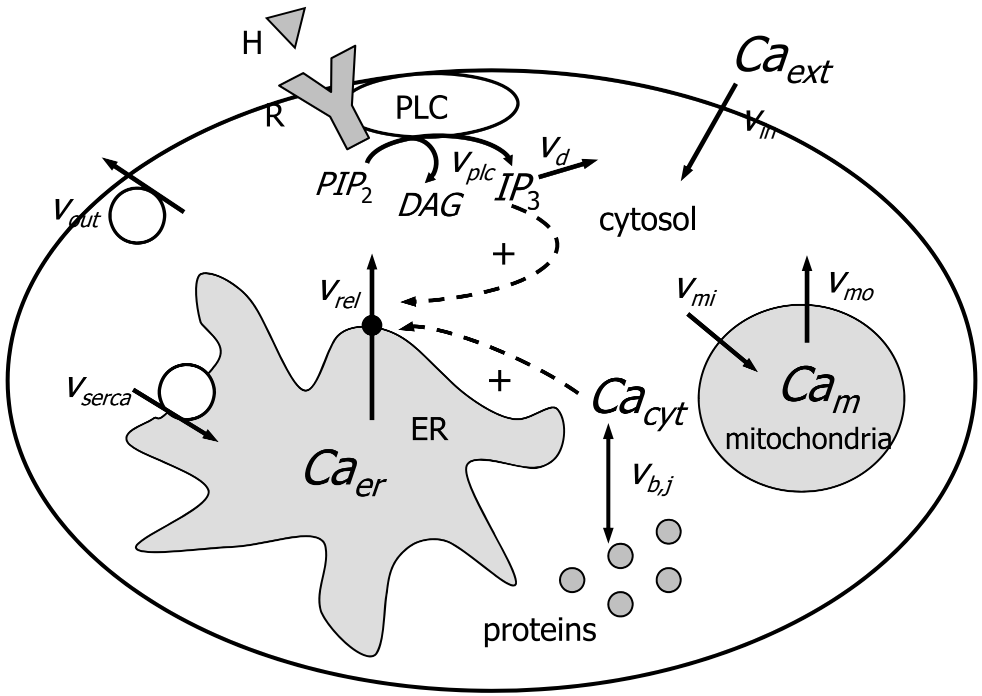
Sperm cell et an egg cell (Wikipedia)

- For nonexcitable cells found with hepatocytes (liver cells) in 1986

Calcium oscillations

- A change in agonist (hormone) level can lead to a switch from stationary states to oscillatory regimes and, then, to a change in frequency

Scheme of main processes



IP_3 = inositol-trisphosphate

Fluxes of Ca^{2+} across the membrane of the endoplasmic reticulum

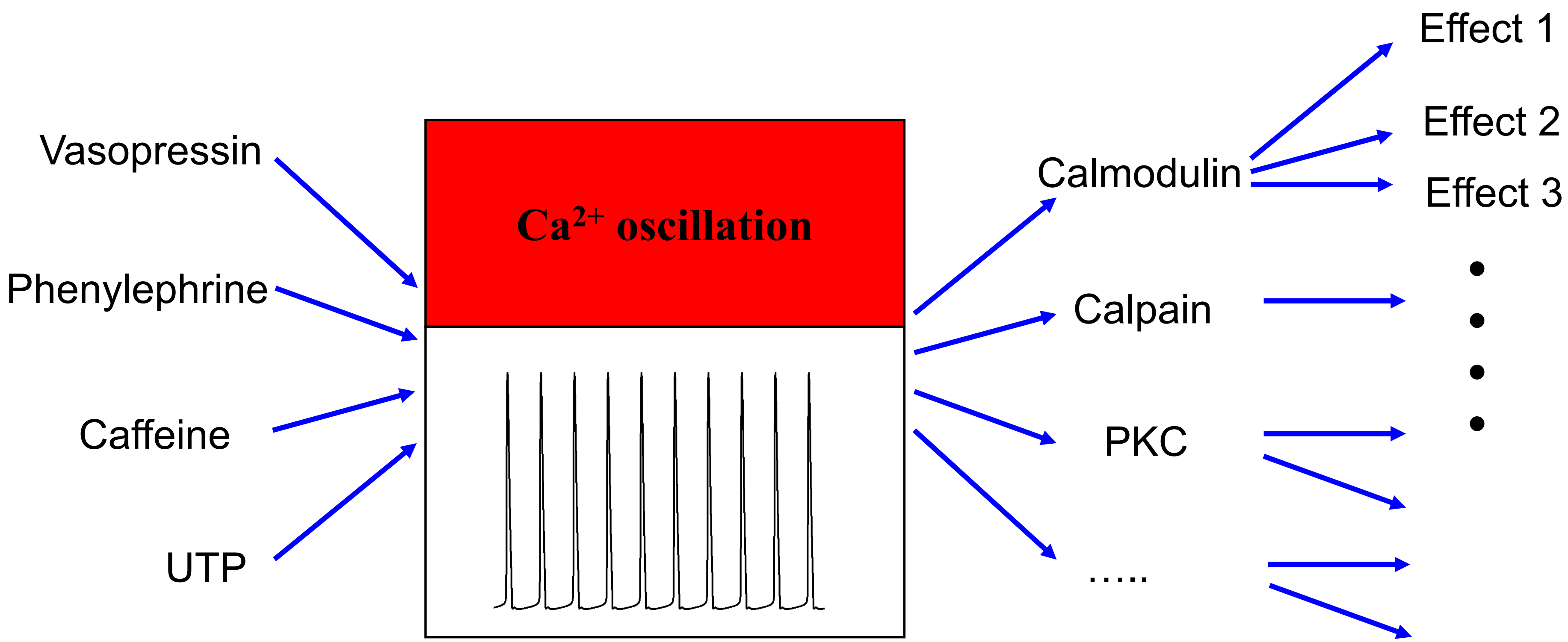
Headlights vs. indicators

Indicators (side repeaters) – signalling function. Oscillating light.

Analogy to intracellular signalling.

Headlights – lighting of street.
Permanent light.

Analogy to metabolism.



Bow-tie structure of signalling

How can one signal transmit several signals?

What is a code?

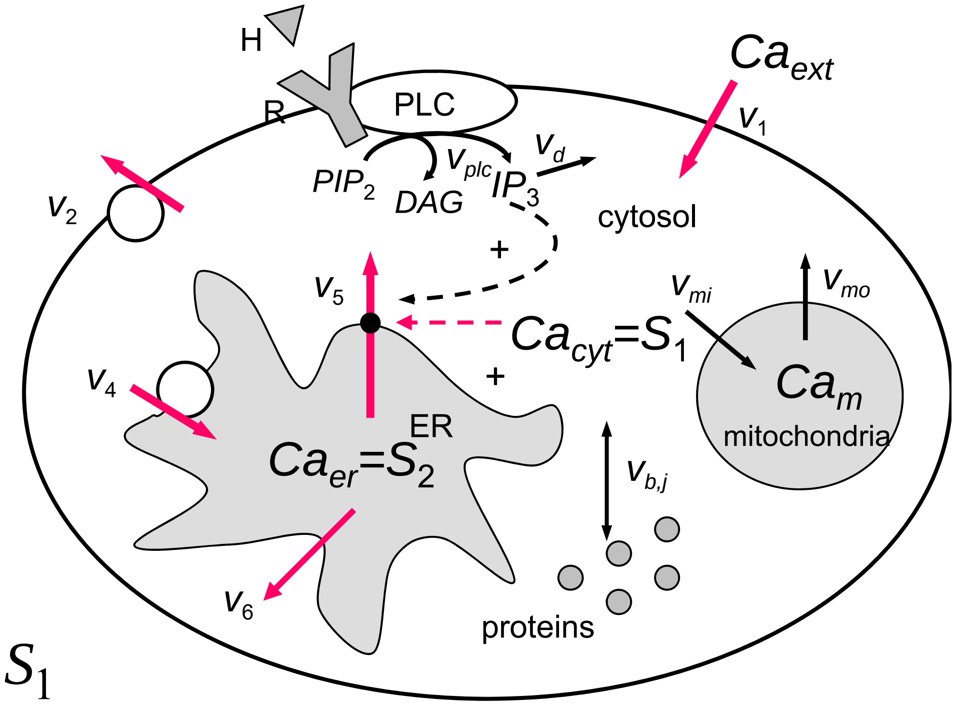
- Mapping in which the rules are not completely determined by physical laws, some bias upon establishment of the code
- Biosemiotics: there are more codes than the genetic code: splicing code, code of calcium oscillations, code of volatiles in plant signalling…

Somogyi-Stucki model

- Is a minimalist model with only 2 independent variables: Ca^{2+} in cytosol (S_1) and Ca^{2+} in endoplasmic reticulum (S_2)
- All rate laws are linear except CICR

R. Somogyi and J.W. Stucki, *J. Biol. Chem*. 266 (1991) 11068

Rate laws of Somogyi-Stucki model



Influx into the cell: $v_1 = const.$

Efflux out of the cell: $v_2 = k_2 S_1$

Pumping of Ca^{2+} into ER: $v_4 = k_4 S_1$

Efflux out of ER through channels (CICR): $v_5 = \dfrac{k_5 S_2 S_1^4}{K^4 + S_1^4}$

Leak out of the ER: $v_6 = k_6 S_2$

Relaxation oscillations

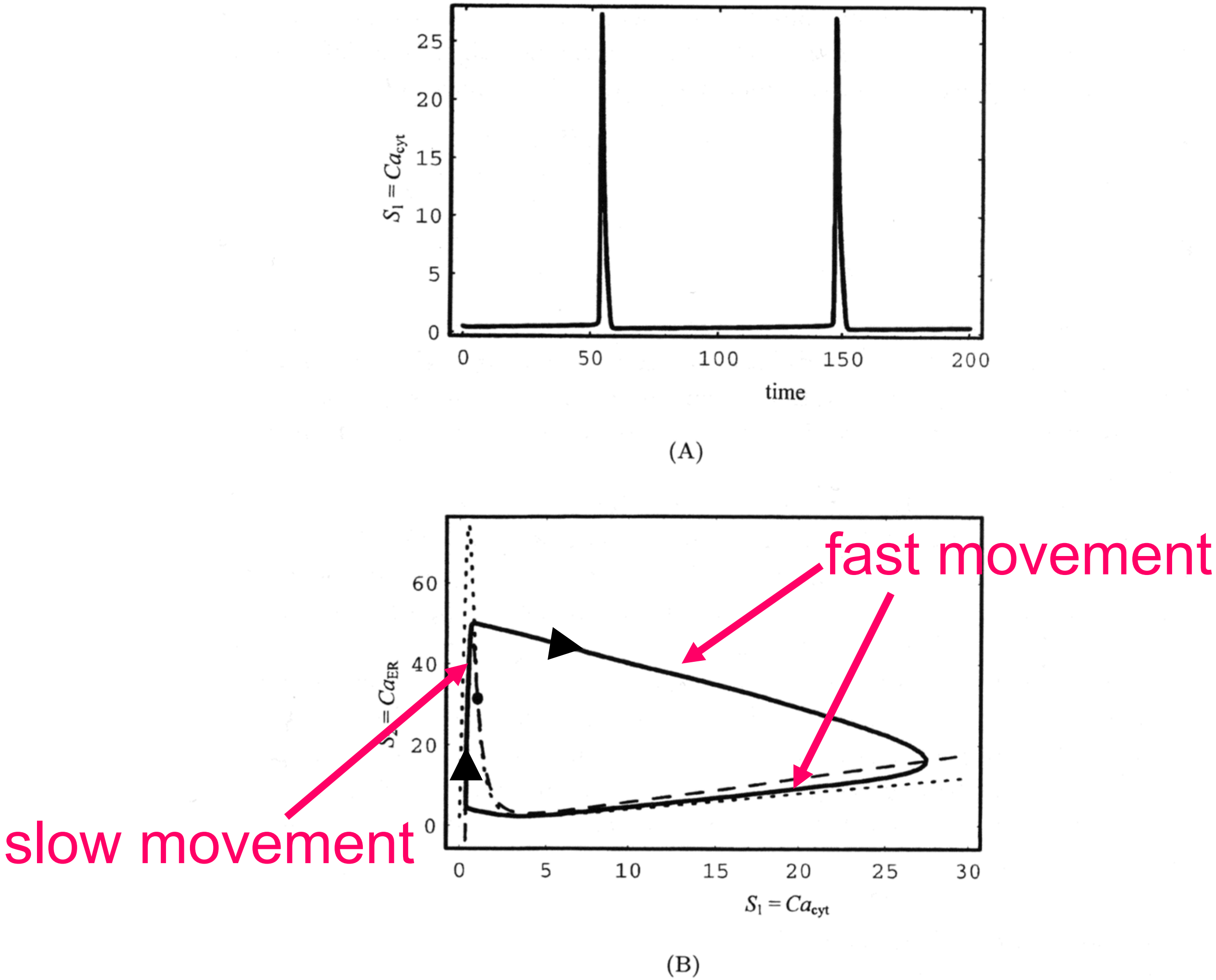
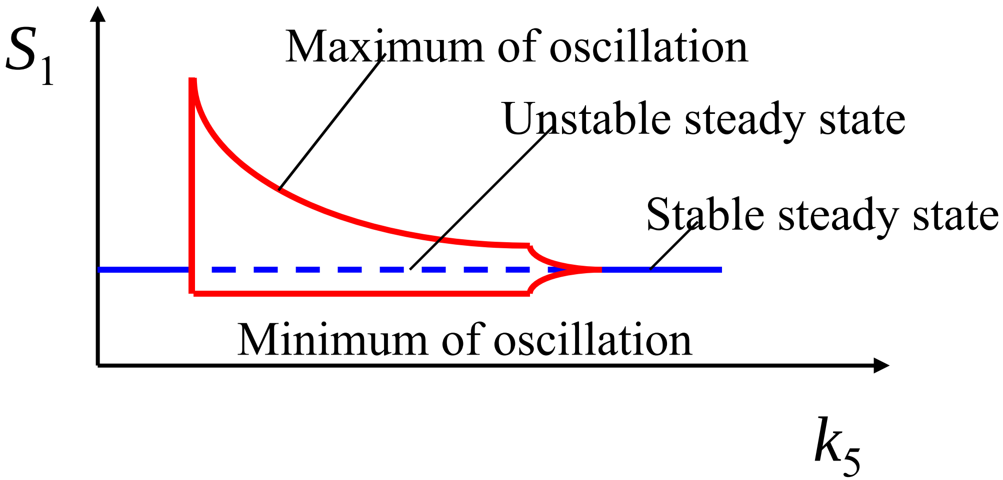


Fig. 2. Time behavior of the Somogyi-Stucki model as obtained by numerical integration using MATHEMATICA. (A) Plot of cytosolic calcium vs. time (in arbitrary units). (B) Plot in the phase plane. The dashed and dotted lines are the nullclines for $\dot{S}_1 = 0$ and $\dot{S}_2 = 0$, respectively. The unique steady state is indicated by ●. Parameter values: $\nu_1 = 1$, $k_2 = 1$, $k_4 = 2$, $k_5 = 5$, $k_6 = 0.01$, $K = 3.1$.

Bifurcation diagram



For small k_5 (i.e. low stimulation of calcium channels by IP_3), $S_1 = Ca_{cyt}$ is at a stable steady state.
Above a critical value of k_5, **oscillations** occur, and above second critical value, again a stable steady state occurs.

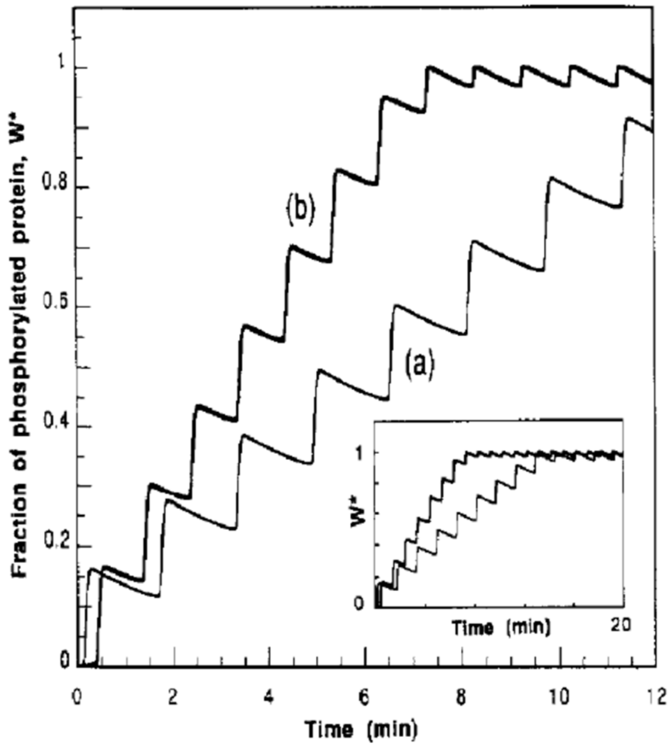
Many other models…

- by A. Goldbeter, G. Dupont, J. Keizer, Y.X. Li, T. Chay etc.
- Reviewed, e.g., in
  - Falcke, M. *Adv. Phys.* (2004) 53, 255-440
  - Schuster, S., M. Marhl and T. Höfer. *Eur. J. Biochem*. (2002) 269, 1333-1355
  - Dupont G, Combettes L, Leybaert L. *Int. Rev. Cytol.* (2007) 261, 193-245.
- Most models are based on calcium-induced calcium release.

„Decoding“ of Ca^{2+} oscillns.

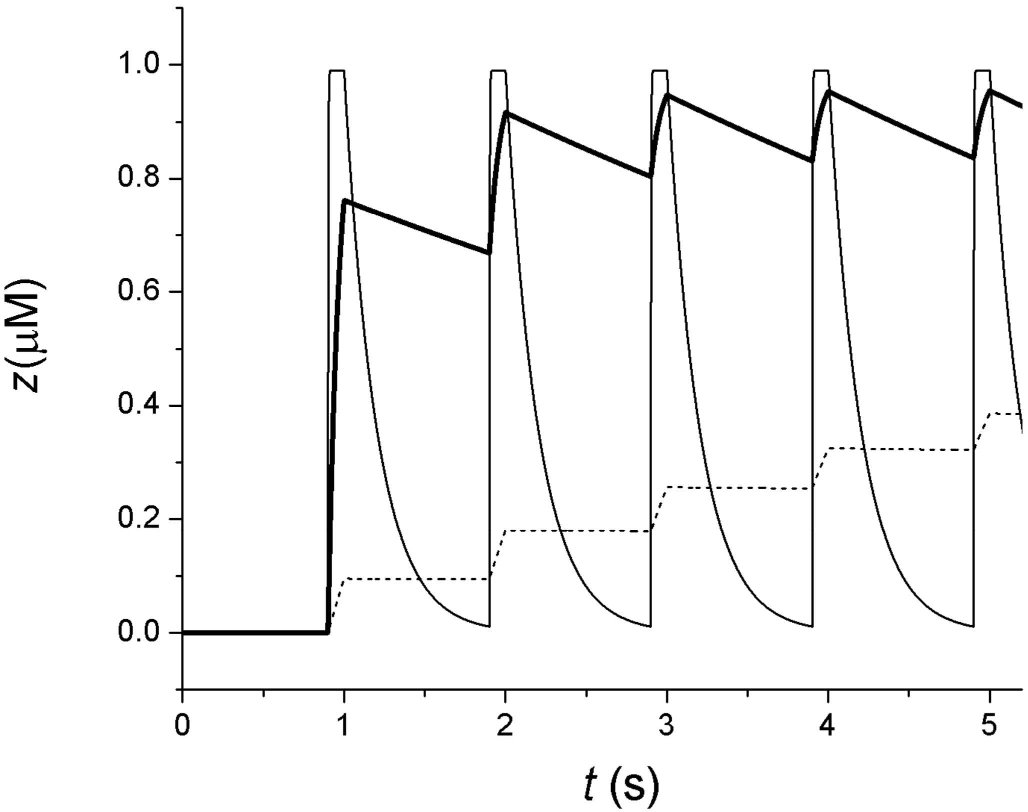
- By calcium-bindung proteins, such as calmodulin
- Two effects.
  - Smoothening / averaging
  - Integration / counting
- Exact behaviour depends on velocities of binding and dissociation

„Decoding“ of Ca^{2+} oscillns.



G. Dupont and A. Goldbeteter,
Biophys. Chem. (1992)

„Decoding“ of Ca^{2+} oscillns.



Thick solid line:
Intermediately fast binding and dissociation

Dotted line:
Slow binding and dissociation

M. Marhl, M. Perc, S. Schuster, *Biophys. Chem*. 120 (2006) 161-167.

What is the point in oscillations?

Jensen‘s inequality:

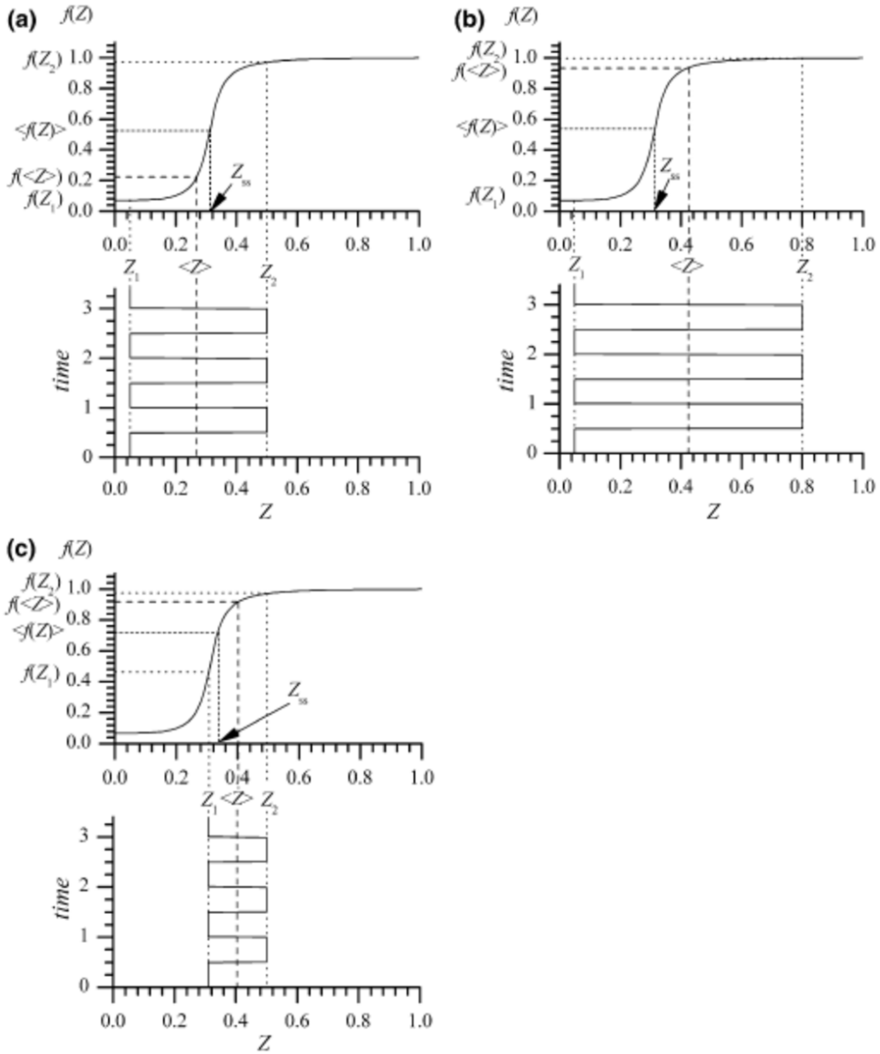
For any convex function $f(x)$:

$\langle f(x)\rangle < f(\langle x\rangle)$

Spikelike oscillations allow signal transmission without increasing average calcium concentration to much. Decoding function must be convex.

Johan Jensen (1859 – 1925)
(Picture: Wikipedia)

Effect of oscillations on protein activation

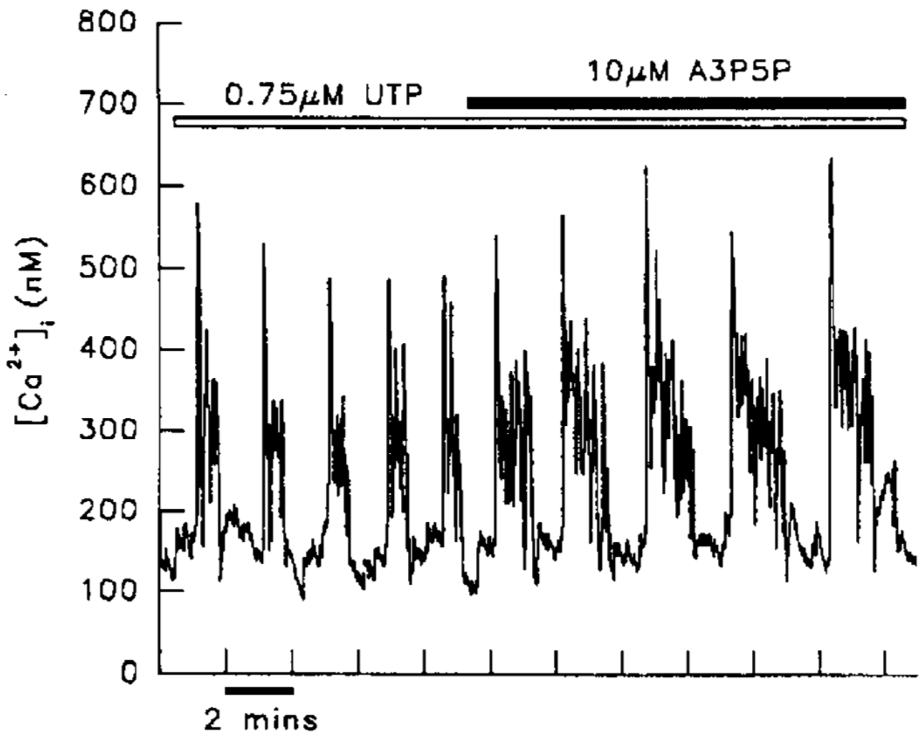


If proteins are activated by convex kinetics (e.g. lower part of Hill kinetics), then average protein activation higher for oscillations than for steady state.

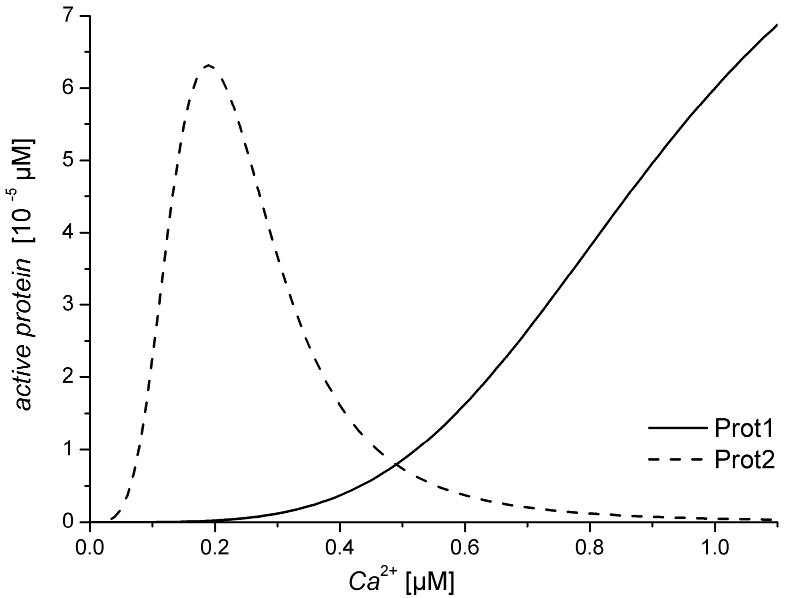
C. Bodenstein, B. Knoke, M. Marhl, M. Perc, S. Schuster: Using Jensen's inequality to explain the role of regular calcium oscillations in protein activation. *Phys. Biol.* 7 (2010): 036009

How can one second messenger transmit more than one signal?

- One possibility: Bursting oscillations



Differential activation of two Ca^{2+} - binding proteins

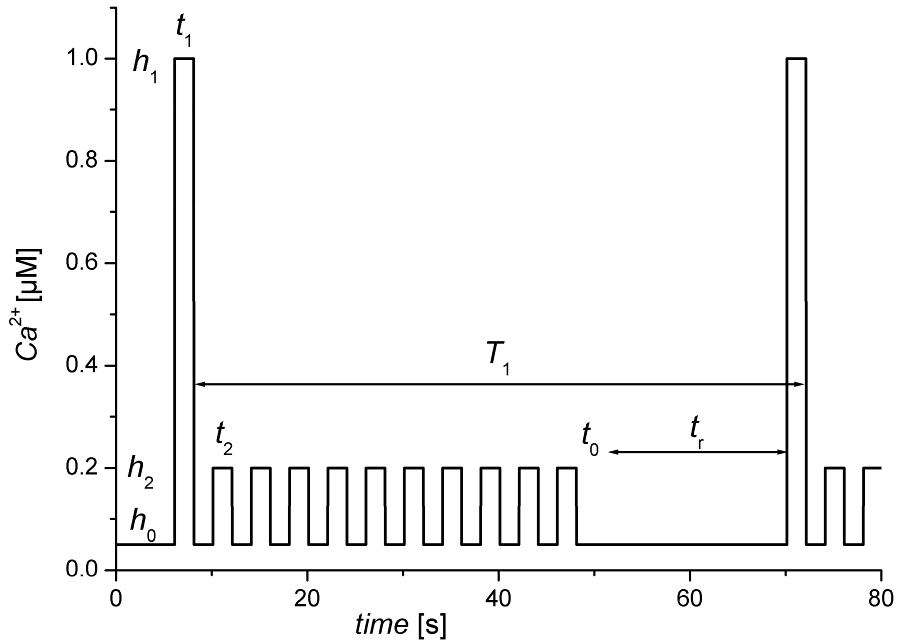


Quasi-equilibrium approximation for binding of Ca^{2+} to proteins

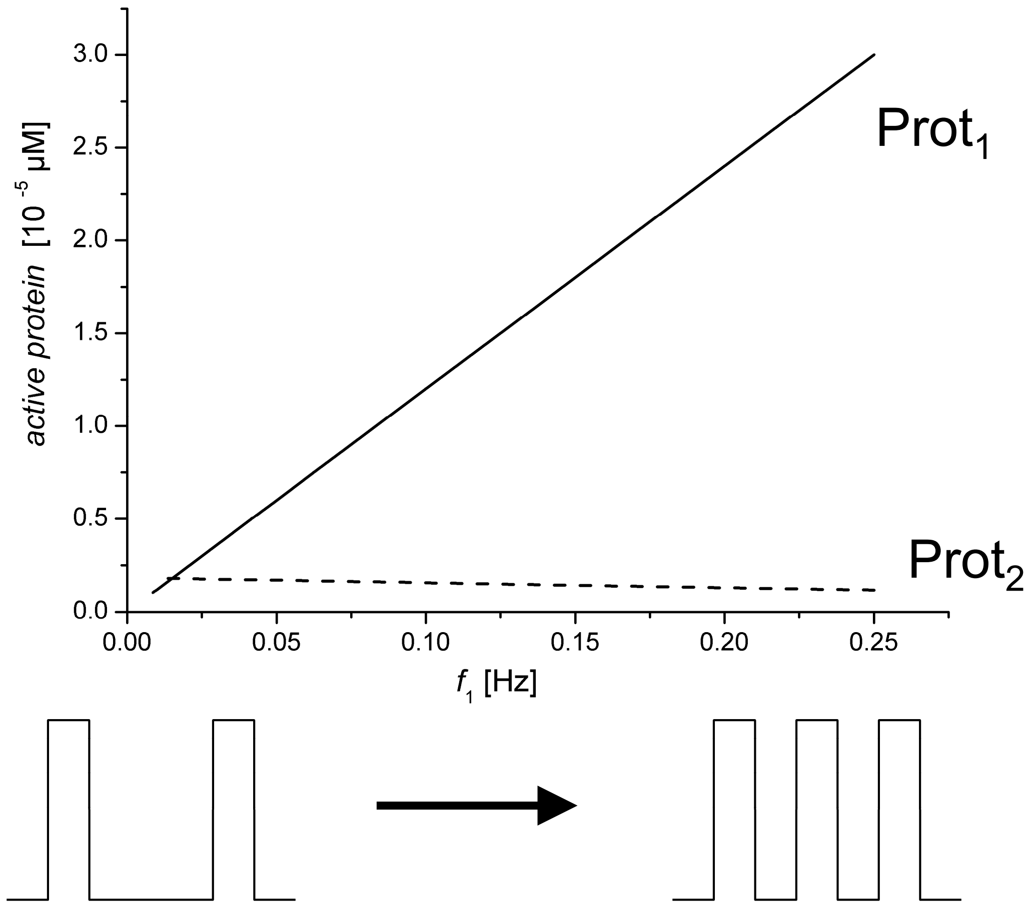
$$Prot_1Ca_4 = \frac{Prot_{1T} * Ca^4}{K_1 + Ca^4}$$

$$Prot_2Ca_4 = \frac{Prot_{2T} * Ca^4}{(K_2 + Ca^4) * \left(1 + \frac{Ca^4}{K_I}\right)}$$

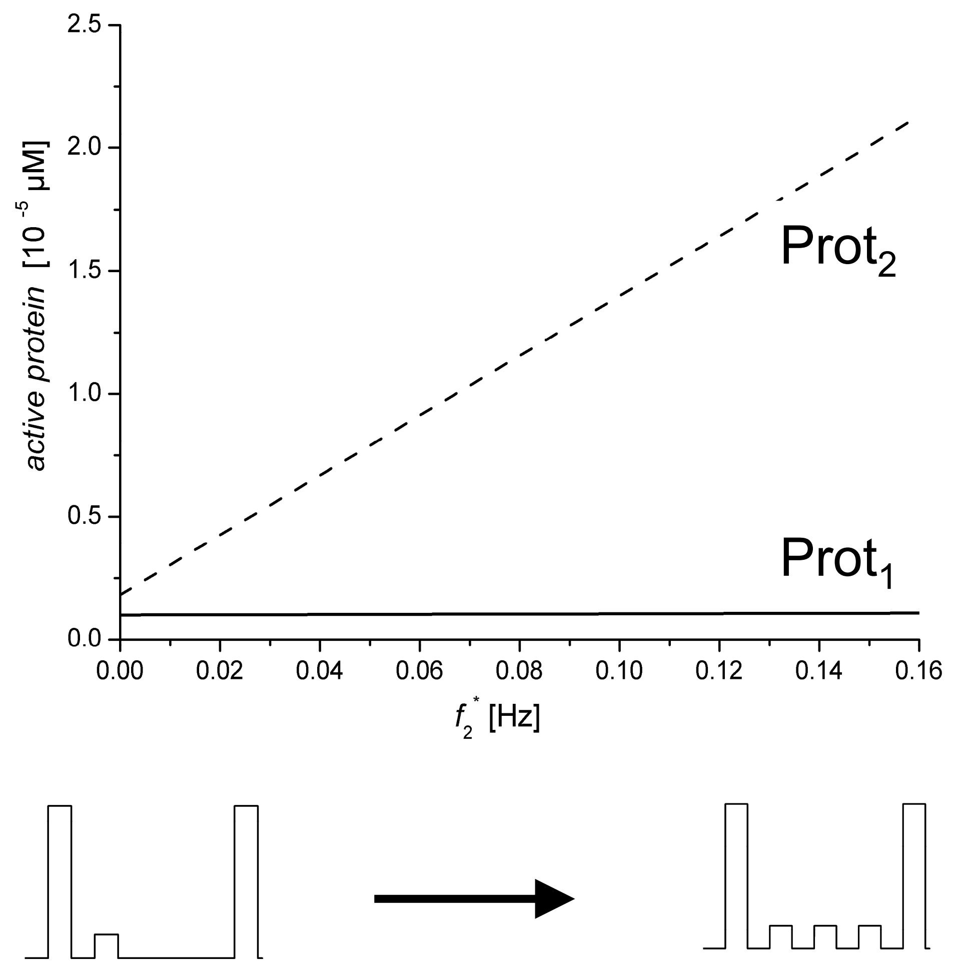
Simplification: rectangular shape of calcium spikes



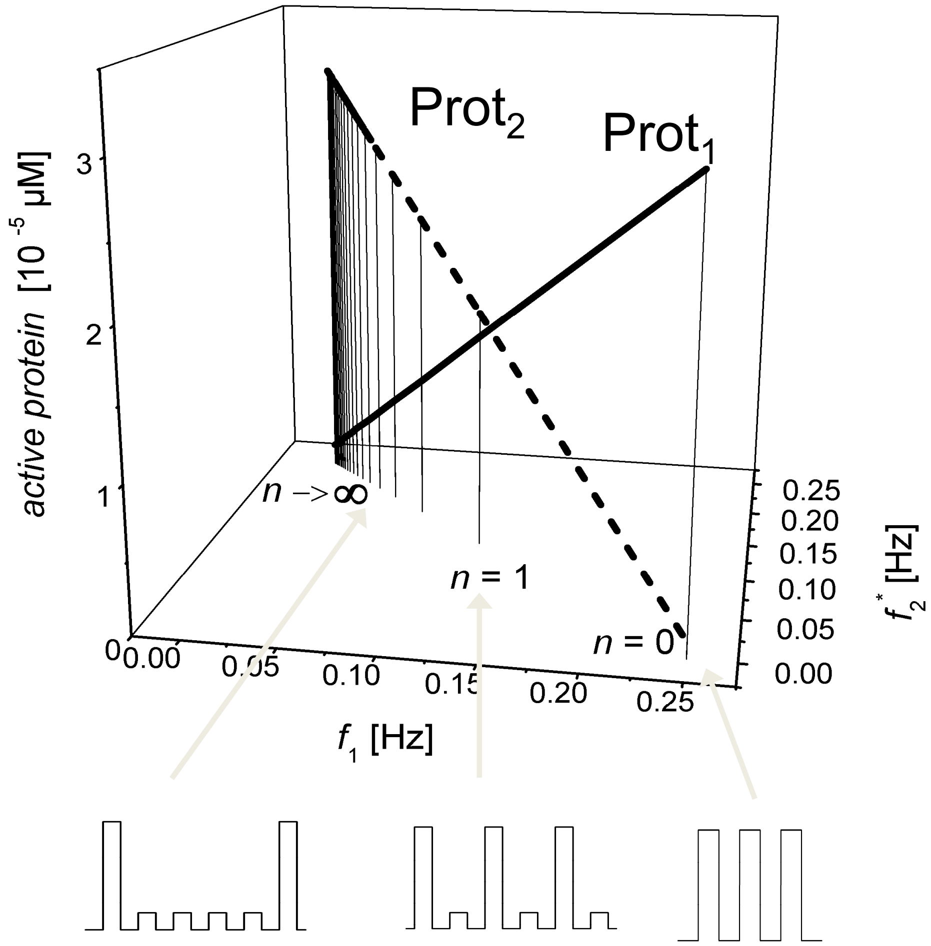
Selective activation of protein 1



Selective activation of protein 2



Simultaneous up- and downregulation



S. Schuster, B. Knoke, M. Marhl: *BioSystems* 81 (2005) 49-63.

See also A.Z. Larsen, L.F. Olsen and U. Kummer, *Biophys Chem.* 107 (2004) 83–99.

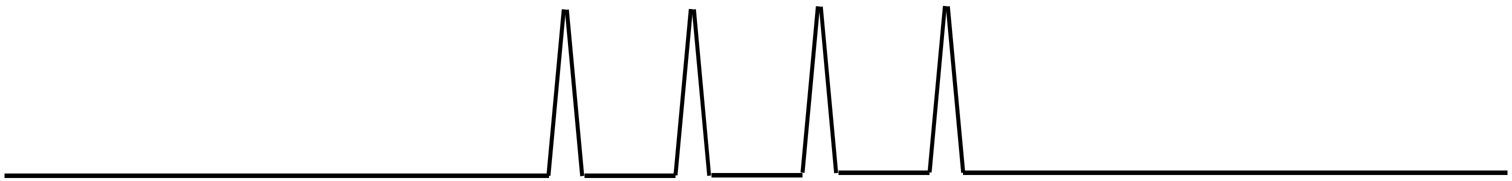
Aperture of stomata in plants

- Certain number (e.g. 5), period and duration of spikes in guard cells optimal for aperture of stomata (openings in leaf surface)

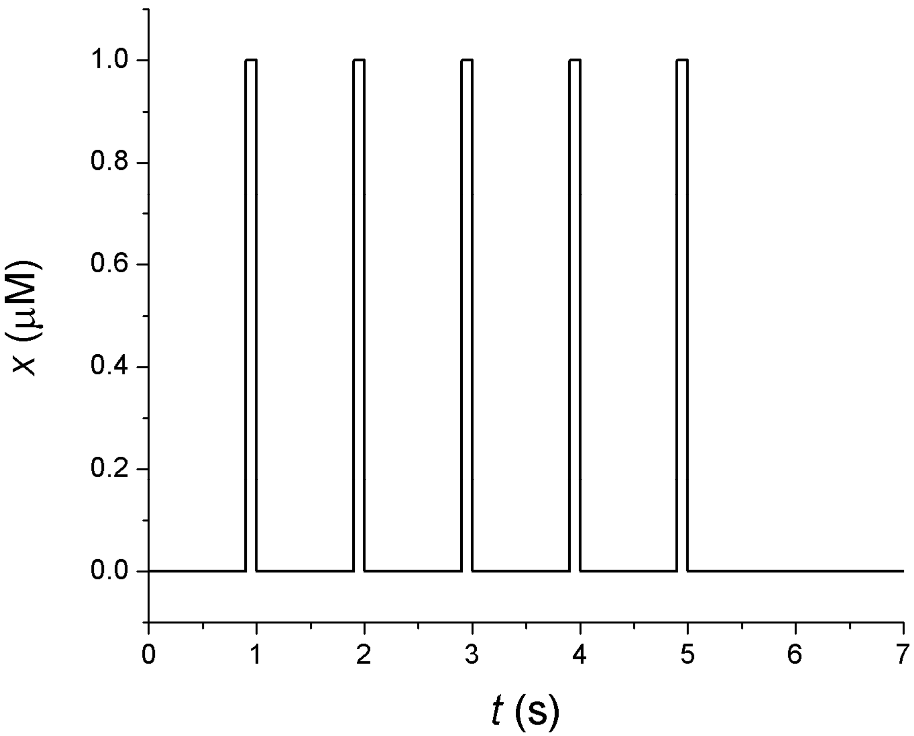
Allen et al.: A defined range of guard cell calcium oscillation parameters encodes stomatal movements. *Nature* 411(2001):1053-1057.

Finite calcium oscillations

- Previously, in theoretical analyses of calcium oscillations, the idealized situation of infinitely long self-sustained oscillations was considered
- However, in living cells, only a finite number of spikes occur
- Question: Is finiteness relevant for protein activation (decoding of calcium oscillations)?

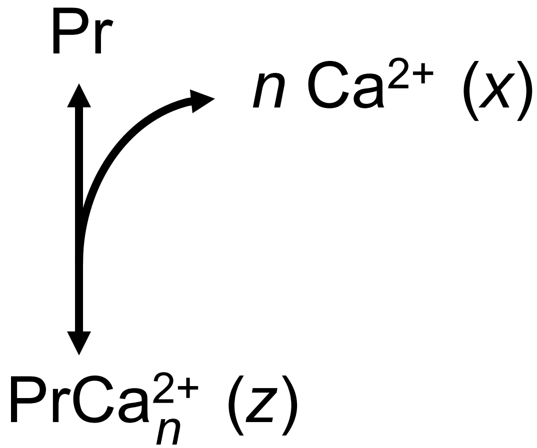


Again simplification by using rectangular shape of calcium spikes



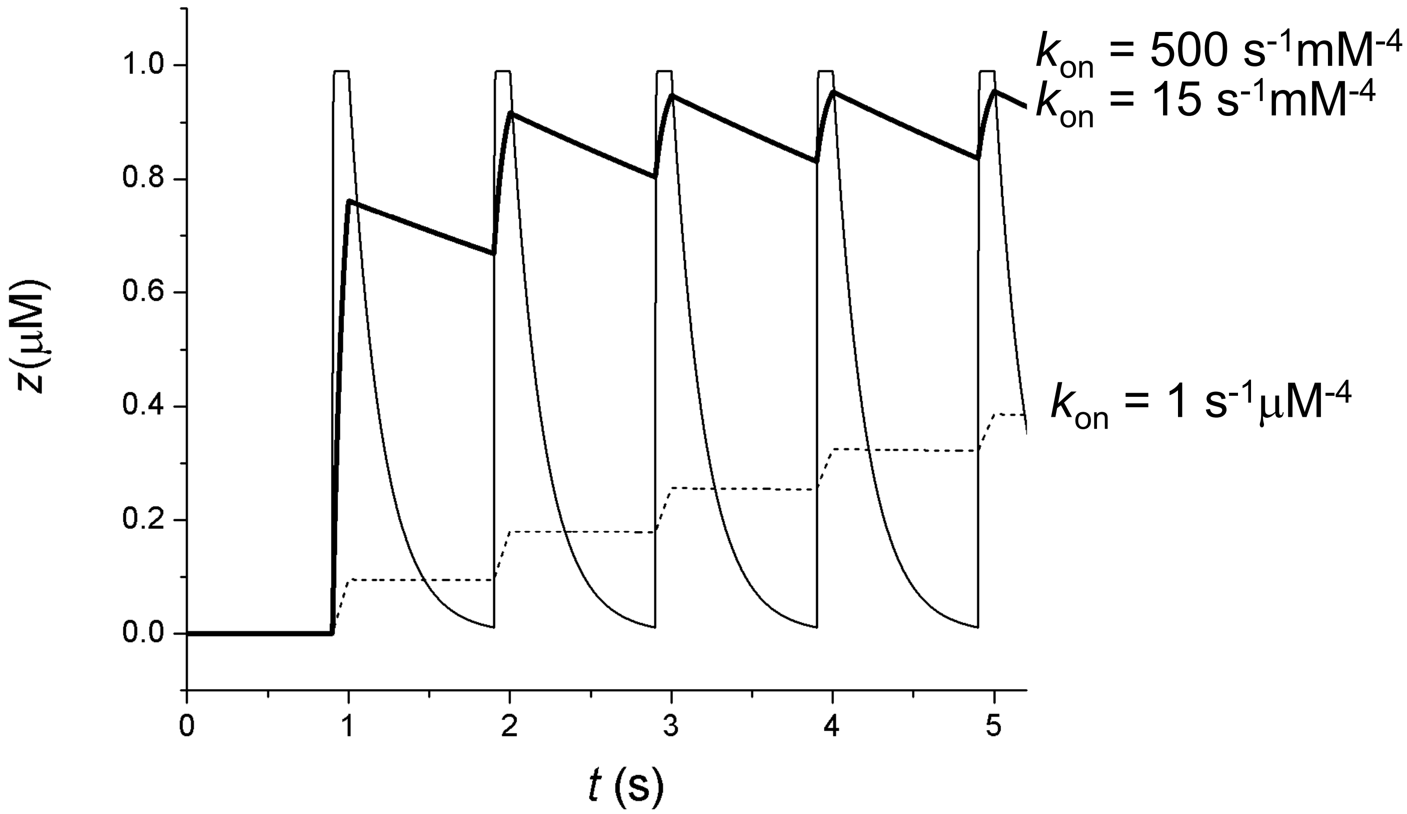
In most calculations, we consider trains with 5 spikes

Binding of calcium to proteins



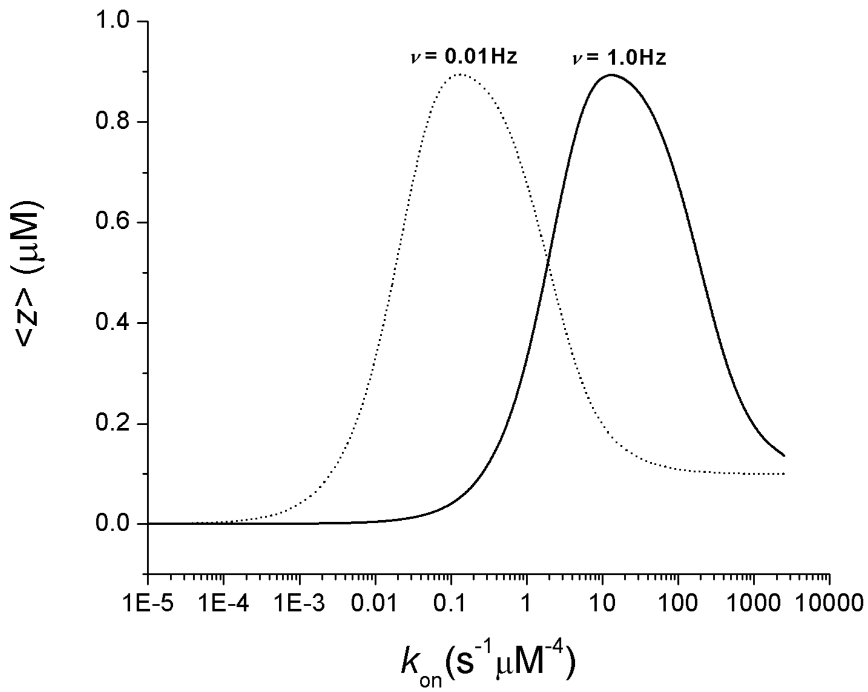
$$\frac{\mathrm{d}z}{\mathrm{d}t} = k_{\mathrm{on}}(z_{\mathrm{tot}} - z)x^n - k_{\mathrm{off}}z$$

Intermediate velocity of binding is best



k_{off}/k_{on} = const. = 0.01 μM^4

Resonant activation of protein

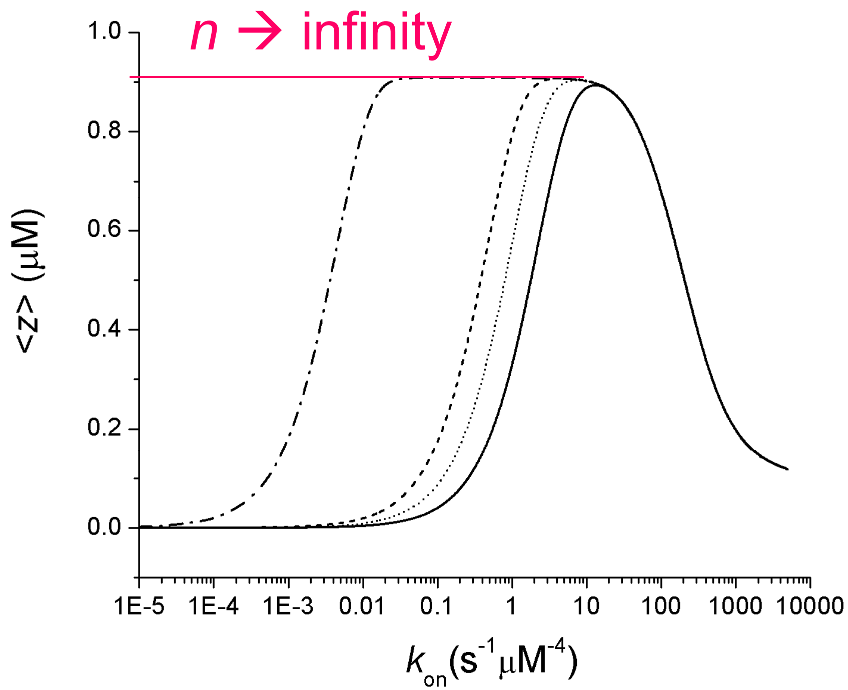


<*z*> = average of activated protein during 5th spike

Proteins with different binding properties can be activated selectively.

M. Marhl, M. Perc, S. Schuster, *Biophys. Chem*. 120 (2006) 161-167.

Finiteness resonance



Gets lost in the limit of infinitely many spikes.
More pronounced when spikes are narrow (not shown).

Discussion

- There are many examples of biological phenomena that can be described algebraically so that result is interpretable and useful.
- Both continuous and discrete mathematics useful for biology.
- Relatively simple models (e.g. Somogyi-Stucki with 2 variables) can describe biological oscillations.

Acknowledgments

- Maximilian Fichtner, Severin Sasso (FSU Jena)
- Christian Bodenstein, Ines Heiland, Beate Knoke (formerly FSU Bioinformatics)
- Marko Marhl, Matjaz Perc, Marko Gosak (U of Maribor, Slovenia)
- Special thanks to Dr. Ina Weiß for literature search.

- FSU Jena, DFG and BMBF for financial support

Picture:
FSU Jena