Alberto Scarampi del Cairo 12/2018

Notes on Systems Biology

Imperial College London
Department of Life Sciences

Non Linear Dynamics and Biophysics – Dr. Robert Endres

Modelling in Biology

- **Exponential Growth**

- **Logistic Growth**

- **First-Order Biochemical Reactions**

- **Quasi-steady state approximation**

- **Taylor Expansions**

- **Euler Method**

- **Euler-Maruyama Method**

- **Stochastic master equation**
  If in a cell the number of copies of a certain molecular species is small (e.g. DNA), the concentration cannot be considered a continuum anymore and its discreteness cannot be ignored. As a result, the description of the time evolution of concentration becomes coarser. The system has no memory of the past, but can only depend on present conditions. Therefore, creation and destruction reactions occur with some probability per unit time (proportional to reaction rates). The master equation tries to describe this phenomenon, by formulating probabilistically the reaction kinetics.
  So the master equation describes the evolution of probability distribution.
- **Emergence of deterministic laws**
  It is possible to obtain the moments for the time evolution of the probability distribution without actually solving the master equation. For example, the mean number of molecules n can be calculated with the sue sum rom n=0 to n=+∞ of n times the probability of having n (n*Pn).
- **Poisson distribution**
  Is a discrete probability distribution that describes the probability that a number of events occurs in a fixed period of time, assuming the events occur with known average rate and independently of previous events. This assumption is often fulfilled in biological systems (e.g. basal expression of a protein independent of other regulators or the occurrence of spontaneous mutations in DNA). In Poisson distribution, the variance is equal to the mean, so it only depends on one parameters (on the contrary to a Gaussian

distribution, where you need two parameters; indeed for large averages a Gaussian with same variance and mean is a good approximation for Poisson) If you derive the steady state distribution of the master equation, it's poissonian (rates of production and degradation per molecule are constant and do not depend on number of molecules, not historical).

- **Gillespie algorithm**
  This is a Monte-Carlo simulation that describes the individual stochastic trajectories.

Biophysics

- **Sensing and Signalling**
  Information theory: mutual information between input and output

- **Bacterial chemotaxis:** biased random walk up a chemical gradient.
  Sensing occurs by detecting indirectly temporal gradient. Signal = current – past measurement.
  An attractor-specific chemoreceptor forms a complex with the adaptor protein CheW and the histidine kinase CheA. CheA phosphorylates the response regulator CheY, and this phosphorylated form, CheY-P, stimulates tumbling by interacting with the flagellar motor. When chemoattractant binds receptor, CheA activity is suppressed, the levels of CheY-P decrease, and the bacterium is less likely to tumble. Adaptation results from the methylation of receptor by CheR, which increases CheA activity, promoting CheY phosphorylation. The methylation state of the receptor is balanced by the demethylation enzyme CheB. CheZ acts to dephosphorylate CheY-P. The dynamics of receptor methylation are considered slow (minutes) relative to CheY phosphorylation (milliseconds) .

  - Signal amplification by receptor clustering
  - MWC model (4 microstates for receptors)
  - Two signalling regimes
  - Precise Adaptation (free-energy landscape, dynamics and integral feedback control)
  - The Perfect Monitor
- **Kinetic proofreading**

Dynamics in Gene Expression

- **Models of Gene regulation**
  - **Epigenetics, diauxi, bistability**
    How to get bistability: positive feedback with cooperativity or double negative feedback with cooperativity (toggle switch).
    Diauxie = Greek word coined by Jacques Monod to mean two growth phases.

    Multistability = the capacity to achieve multiple internal states in response to a single set of external inputs (the defining characteristic of a switch).

- **E.g. Lac System** (Ozbudak et al., 2004)
  In Lac system there is both bistability (expression of LacY permease results in even more lactose uptake) and hysteresis (there is memory: LacY persist as long as it is not diluted by cell division).
  The Lac operon has threshold switches. It is induced if [TMG] < 3uM and it is induced if [TMG]>30uM. In between the two values, the system is hysteretic and bistable. If below or above those values, the system is monostable (plot the graph to clearly visualise this). In the bistable region, you can observe a bimodal distribution of cells expressing GFP, as there are stochastic fluctuations that bring some cells to the induced attractor and some other cells to the uninduced attractor. The monostable and bistable states are separated by a saddle-node bifurcation.

  The all-or-none induction of the *E. coli lac* operon has been a paradigmatic example for bistability in gene regulatory networks for many decades. However, so far bistability has been experimentally demonstrated only for induction with gratuitous inducers such as TMG or IPTG, but not for induction with the natural inducer lactose. In fact, based on theoretical analysis of the *lac* circuit architecture, Savageau (2011) argued that bistability is unlikely to occur in the natural lactose utilization system, but that overexpression of the LacI repressor would favor the emergence of bistability. The Bettenbrock group at the Max Plank Institut has recently published a paper in the Biophysical journal where they report to have tested and confirmed this prediction by combining single cell analysis with deterministic computational modelling (Zander et al., 2017).
  In this sense, their experiments can be viewed as opposite to those conducted by Ozbudak et al. (2004), who showed that TMG-induced *lac* operon expression can be driven from a bistable into a monostable regime by successive dilution of LacI repressor. Guided by the computational model, our results support the view that *lac* operon induction in the wild-type strain is graded (monostable) rather than all-or-none (bistable), which is consistent with the Savageau design principle as well as with previous experimental analysis of lactose-induced *lac* operon expression in *E. coli.*

- **Mathematics of bistable gene switch**
  Remember the “existence and uniqueness theorem”: if f(x) is smooth enough, then solutions exist and are unique. One of the consequences of this is that in a phase portrait every trajectory does not intersect other trajectories (because solutions are unique) expect in steady states.

  Autocatalytic feedback model of gene expression: the activity of a gene is directly induced by two copies of the protein for which it codes.
  The system has three fixed point if $a<1/2b$
  The system experiences a saddle-node bifurcation at the critical value $a=1/2b$
  The system can act like a biochemical switch, but only if the mRNA and protein degrade slowly enough (degradation rates must satisfy $ab<1/2$). If this

is the case, there are two steady states: one at the origin (gene is silent) and one where x and y are large, meaning that gene is active and sustained by the high level of protein. The stable manifold of the saddle acts like a threshold: it determines whether the gene turns on or off, depending on the initial value of x and y.

When plotting phase portraits, remember that trajectories approach the origin tangent to the slow eigendirection, defined as the direction spanned by the eigenvector with the smallest eigenvalue.
Phase curves (or phase trajectories) are solutions of:

$$dx/dy = f(x,y)/g(x,y)$$

General eigensolution:

$$x(t) = c_1 e^{\lambda 1 t} v_1 + c_2 e^{\lambda 2 t} v_2$$

- **Competence/Excitability**
  Excitability: relatively small, threshold-crossing perturbations trigger large-amplitude excursions in phase space that eventually return the system to its initial state. ComK 'master' transcription factor activates expression of a suite of genes necessary for competence, including the comG operon in *B. subtilis*. ComK also activates its own expression. Upon entry into stationary phase, comK is expressed at a basal level, but is also rapidly degraded by the MecA complex. ComS peptide competitively inhibits ComK degradation by the MecA complex. Expression of ComS thus favours induction of competence by allowing ComK levels to build up sufficiently to enable full ComK activation by positive autoregulation. ComK indirectly represses ComS, generating an anti-correlation between P_{comG} and P_{comS} activities. The regulation of comS is, however, known to be complex, having several transcriptional inputs and so occurs with a delay.

  A fundamental question is whether initiation of competence is stochastic or affected by memory of previous events. Two consecutive competence events can be observed in a single cell lineage, showing that cells retain the potential to re-initiate competence. In fact, re-initiation occurred with a frequency not significantly different from the overall competence frequency, that is repeated competence events are neither favoured nor suppressed. These results are consistent with a stochastic and memory-less model for competence initiation and duration.

  In the Elowitz model, noise can induce escape from the otherwise stable vegetative state and turn the system 'on' via the ComK positive feedback. On a slower timescale, this initiates the ComS- mediated negative feedback. Reduction in ComS levels eventually shuts the system back 'off' through an increase in ComK degradation, returning the cell to its vegetative state. In this way, ComS has a dual role in the system: on the one hand it is necessary to initiate competence, by blocking degradation of ComK and allowing positive autoregulation to take effect; on the other hand, repression of ComS is necessary for exit from competence, because reduction in ComS levels favours ComK degradation by MecA.

- **Turing Pattern**

Advanced Modelling

- **Transcriptional Network Motifs**

  Shen-Orr et al. (2002) highlighted three motifs that are most prevalent in the *E. coli* network: the single input module (SIM), a dense overlapping regulon (DOR), and the feed-forward loop (FFL).The three motifs enable the *E. coli* network to be broken down into its basic building blocks. They then showed that there is a single layer of five DORs connecting most TFs to their targets. The DORs are joined together by global regulators. SIMs and FFLs are positioned at the outputs of the DORs. What is apparent is that the *E. coli* regulatory network has a shallow structure and contains only a few long cascades for flagella biogenesis and several metabolic systems.

  a) FFL: consists of a primary TF that regulates a secondary TF, and together, both regulate a final target.
     A FFL is coherent if the direct effect of the primary TF on the target gene has the same sign as its net indirect effect through the secondary
     Example: Arabinose operon

  b) SIM: comprises a group of genes targeted by a lone TF.
     SIM motifs occur in systems of genes that function stoichiometrically to form a protein assembly or a metabolic pathway. They proposed that a single TF could maintain the proportions of transcribed operons at a steady state, and that differences in activation thresholds of targets could provide a temporal program of gene expression. An example of a SIM motif involves ArgR, an *E. coli* TF that uniquely regulates five target operons encoding genes for arginine biosynthesis.

  c) DOR: a close-knit layer of overlapping interactions between genes and a group of input TFs

- **Network Motifs for Different Responses**
  Signaling pathways can be embedded in networks using positive and negative feedback to generate more complex behaviours — toggle switches and oscillators — which are the basic building blocks of the exotic, dynamic behaviour shown by nonlinear control systems.

  a) Linear
  b) Hyperbolic
  c) Sigmoidal: modification of hyperbolic case (b), where the phosphorylation and dephosphorylation reactions are governed by Michaelis-Menten kinetics

- **Cooperativity and Ultrasensitivity**

- **Imprecise Adaptation (Homeostasis)**

- **Precise Adaptation**
  By supplementing the simple linear response element (case a of network responses above) with a second signaling pathway (through a species X in the drawn image behind) it is possible to create a response mechanism that exhibits perfect adaptation to the signal. Perfect adaptation means that although the signaling pathway exhibits a transient response to changes in signal strength, its steady-state response R_{ss} is independent of S. Such behaviour is typical of chemotactic systems, which respond to an abrupt change in attractants or repellents, but then adapt to a constant level of the signal.

  - Integral Feedback Control: is a basic engineering strategy for ensuring that the output of a system robustly tracks its desired value independent of noise or variations in system parameters.

  - Incoherent Feed-Forward

- **Oscillators**
  They are limit cycles, that is steady states with purely imaginary eigenvalues. Often produced by bifurcations (like Hopf).

  - Negative Feedback Loops (supercritical Hopf)
  - Activator-Inhibitors (subcritical Hopf)
  - Substrate-Depletion (subcritical Hopf)

- **Bifurcations**
  Points where changes in the qualitative structure of the flow changes as parameters are varied. Bifurcations points are the parameters values where these abrupt changes in dynamics take place (e.g. when a fixed point is destroyed)
  - **Saddle-Node Bifurcation**
    It is the basic mechanism by which fixed points are created or destroyed. In these bifurcations, as parameters are varied, two fixed points move towards each other, collide and mutually annihilate.
  - **Hopf Bifurcation**
    Supercritical: when a stable spiral changes into an unstable spiral surrounded by a small, elliptical limit cycle.
    Subcritical: when an unstable cycle shrinks to zero amplitude and engulfs the origin, thus rendering it unstable. Then solutions that used to remain near the origin are now forced to jump into large amplitude oscillations. Because of these large amplitude jumps, these types of bifurcations are more dramatic and dangerous than super in engineering. However, in biology they enable the creation of activator-inhibitor and substrate-depletion types of oscillations.
  - **Pitchfork Bifurcation**
    reversible transition, symmetry breaking event

Alberto Scarampi del Cairo 12/2018

Emergence – Prof. Mark Isalan

Emergence and Gene Expression

Attractors are stable genetic programs:

- Fixed points (neg feedback, thermostat)
- Limit cycles (neg feedback with delay, e.g. oscillators)
- Bistability (thresholded, e.g. lac operon)

Network topology can define systems attractors (Rene-Thomas Conjecture). However, topology is not everything. The parameters matter. The same network architecture can give rise to two completely different qualitative behaviours (e.g. repressilator).

- **The Repressilator** (Elowitz and Leibler, 2000)
  They built a three-node transcriptional repressor system to build an oscillating network in *E. coli*. An odd number of negative connections in a loop gives negative feedback. A delayed negative feedback leads to oscillations (even though a bit damped). The dynamics of this system depends on factors such as:
  1. Transcription rate
  2. Translation rate
  3. Decay rates of protein and mRNA

  Depending on parameters, two types of solutions are possible.
  1. System converges towards a stable steady state
  2. Steady state becomes unstable, leading to sustained limit-cycle oscillations.

  Oscillations are favoured by:
  1. Strong promoters
  2. Efficient RBS
  3. Tight repression (low leakiness)
  4. Cooperative repression
  5. Similar mRNA and protein degradation rates

  To address requirement 1 and 2, they used hybrid promoters (λ Pl promoters with lac and tet operator sequences)
  To bring protein degradation rate closer to that of mRNA, they added ssRNA –tag on 3' end of mRNA coding for repressor protein (recognised by protease). So now protein has halftime of 4-20 min and mRNA has 2 min.
  They assembled circuit in bacteria and saw that timecourse of GFP florescence had a period of 150 min (longer than cell cycle).
  However, they found that entry into stationary phase causes the repressilator to halt, indicating that the network is coupled to the global regulation of cell growth.

  Parameters matter:
  1. Equal Promoter Strength: oscillations, limit cycle

2. Different Promoter Strength: narrow oscillations (like thermostat), pseudo-oscillator in phase portrait (collapses to a fixed point).

- **The Genetic Toggle Switch** (Gardner, Cantor and Collins, 2000)
  The toggle switch is composed of two repressors and two constitutive promoters. Each promoter is inhibited by the repressor that is transcribed by the opposing promoter.
  The bistability arises from the mutually inhibitory arrangement of the repressor genes (that is equivalent of a positive feedback) and also from the cooperative repression of transcription.
  In the absence of inducers, 2 states are possible: one in which promoter 1 transcribes repressor 2, and the other where promoter 2 transcribes repressor 1.
  Switching is accomplished by transiently introducing an inducer of the currently active repressor. The inducer permits the opposing repressor to be maximally transcribed until it stably represses the originally active promoter.
  Induction by IPTG or aTC alters the dynamic balance between the competing promoters such that the toggle is pushed into a region of monostability.
  The transition from bistability to monostability occurs in a sharp and discontinuous fashion, due to the existence of a saddle-node bifurcation (when one of the stable steady states is annihilated by the unstable state.
  The stochastic nature of gene expression causes variability in the location of the switching threshold and thus blurs the bifurcation point. Near the bifurcation, this blurriness is realised as a bimodal distribution of cells.

  - Stable points attractors

    The toggle with initial state above the separatrix will settle to one stable state, whereas a toggle starting below will settle to the stable steady state below. The position depends on transient fluctuations.

- **Bistable Attractor Selection** (Kashiwagi et al., 2006)
  Can cells choose the ‘right’ attractor via stochastic switching?
  They built a system based on toggle switch and added a fitness pressure to either state 1 or state 2. When cells are cultured in rich medium, they observed o monostable system.
  In order to tease out the attractors, they added nalidixic acid (growth inhibitor) and then they could see bistability (plots here to visualise).
  So turns out that cells can switch state and go to fit state after hours after changing medium. The system is not regulated (no signalling), yet it behaves as if there was signalling. Cells are picking the right attractor. Is this gene expression on demand?
  Under the fluorescence microscope they checked if cells were actively changing the attractors or if just the best fitted cells were surviving.
  They found that the response was not due to proliferation of fit cells but rather by stochastic switching. Individual cells can cross the separatrix and move to one side or another in the dynamical portrait. They described this using a stochastic model.

To achieve fitness-induced attractor selection, the model requires 4 postulates to be satisfied:

1. There must be noise (η_1, η_2)
2. Noise amplitude is independent of activity A (or at least does not vanish with decreasing activity A)
3. Both S(A) and D(A) are increasing functions of activity A (in turn correlates to nutrient condition and growth rates)
4. In inappropriate medium A decreases, whereas by expressing m1 or m2 appropriately A increases

So by increasing fitness by A, there is a positive feedback that increases more and more fitness.

- **Monostable Attractor Selection** (Tsuru et al., 2011)
  Kashiwagi et al. (2006) and others gave experimental evidence of stochastic switching based only on a bistable genetic structure. Bistability allows the cells to stay at a stable state, either fit or unfit, so that the stochastically appearing fit state ca be stabilised without further random switching to the unfit state. Nevertheless, the universality of stochastic switching as an adaptation strategy remains open, as bistability is known to be a special case in native genetic structures.
  Monostable gene expression is much more common and does not rely on complex genetic networks. To achieve a population, shift from a non-adapted stable state to an adapted unstable state, a significant increase in fitness (e.g. growth rate) of the fit cells is necessary (as they new-born fit cells need to proliferate and conquer the population). Otherwise, the random switching will hide the adapted transitions that sometimes occur and would lead to an unchanged population of the stable non-adapted state. Therefore, in contrast to a bistable network, the final adapted state in a monostable system is not determined a priori by the genetic architecture but is determined by the cellular response.
  To answer this question, they deleted the hisC gene from the His operon and placed it under the control of the very orthogonal PtetA promoter in an engineered gene circuit in another chromosomal locus.
  As a result, the native transcriptional regulation of the His operon and only stochastic switching of the engineered hisC monostable circuit can provide cells a chance of survival from histidine starvation.
  They observed stochastic switching-mediated adaptation at both population and microcolony level.
  In conclusion, the stochasticity of the rewired gene itself is insufficient to reach an adapted state, but the coordinated reorganization of global gene expression (quantified by transcriptomics) is essential. These results indicate a certain level of universal discipline during stochastic adaptation, which is represented by cellular plasticity in rewired cells.

  Essentially here there is an implicit attractor from a global fitness relationship. We need to move away from the mechanistic concept of molecule concentrations as only attractors but also global relationships such as fitness can be attractor if we want to understand the output of biological systems.

- **Growth Laws** (Klumpp and Hwa, 2014)

  Bacterial growth provides a model system for studying the coupling between individual genetic circuits and the global state of the cell. For example, gene expression requires ribosomes and RNA polymerases, whose concentrations and availability depends on the growth state of the cell. These observations actually point towards the limitations of the analogies between genetic and electronic circuits and metaphors that describe the host cell as a chassis, onto where the circuit is mounted and orthogonal…
  Changes in gene expression occur in conjunction with adaptation of the physiology of the cell as a whole. Some growth rate-dependent parameters: gene copy number, RNAP and ribosome concentration, mRNA lifetime, dilution by growth, cell volume.

  Passive Effect due to Growth
  Simplest case of growth-rate dependence is a gene, whose product is neutral and in low abundant (passive effect). Constitutively expressed gene can be used to separate growth-rate dependencies from effects of gene regulation:
  1. Constitutive expression in nutrient-modulated growth:
     If a gene product is neutral to fitness and in low abundance, as growth rate increases, protein concentration decreases linearly as growth rate increases (because it gets diluted).
  2. Constitutive expression in translation-modulated growth
     If gene product provides a fitness benefit, as growth rate increases, the protein concentration increases linearly (cellular parameters like transcription rate and gene copy number increase at faster growth).

  Growth Feedback
  If the product of a gene has a (positive or negative) effect on growth, growth provides a feedback mechanism for the expression of that gene.

  1. Genes with toxic product (nutrient-like growth modulation)
     Constitutive expression of a toxic protein in nutrient-limited growth results in positive feedback. An increase in the concentration of the toxic protein leads to a reduction in growth rate, which in turn results in a further increase of the toxin concentration (provided that the toxin's effect modulates growth in similar fashion as nutrient depletion). Bistability arises if positive feedback is sufficiently cooperative and is reflected in bimodal distributions of gene expression levels. The two-subpopulations also exhibit different growth rates. Such feedback may for example be induced by expression of chromosomal toxin-antitoxin systems (the coexistence of slowly and rapidly-growing cells is indeed observed in antibiotic resistance).

  2. Constitutively expressed resistance genes (translation-modulated growth)
     Increased expression of antibiotic resistance gene in growth modulated by translation-targeting antibiotic reduces the antibiotic level in the cell,

thus enhancing growth, which in turn increases the concentration of the gene product.

- **Growth Bistability and Ohm's Law**

Emergence and Pattern Formation

Alan Turing in 1951 proposed that a system of chemical substances (morphogens) reacting together and diffusing can develop from homogenous to a pattern/structure due to an instability of the homogeneous equilibrium (triggered off by stochastic disturbances).
Interestingly, TP can generate self-organised, complex, repetitive patterns of gene expression. For this reason, they are regarded as the driving morphogenetic patterning mechanism in many biological systems.

- **Turing RD Model**
  Alan Turing considered a system composed of a ring of cells each in contact with its neighbours, where two morphogens react and diffuse as described by the differential equations:

  Reactions:

$$dX/dt = 5X – 6Y + 1$$

$$dY/dt = 6X – 7Y + 1$$

  Diffusions:

$$\delta X/\delta t = 5X – 6Y + 1 + dx\nabla^2 X$$

$$\delta Y/\delta t = 6X – 7Y + 1 + dy\nabla^2 Y$$

  The system was supposed to be initially in a stable homogeneous condition, but disturbed slightly from this state by some influences unspecified, such as Brownian movement or the effects of neighbouring structures or slight irregularities of form. This had the effect to bring the system from the stable state to an unstable one (Turing instability)

  After the lapse of a certain period of time from the beginning of instability, a pattern of morphogen concentrations appears which can best be described in terms of 'waves'. There are six types of possibility (waves) which may arise:

  1. Uniform, stationary
  2. Uniform, oscillating

3. Stationary with very short wavelengths
4. Oscillatory with very short wavelengths
5. Oscillatory with finite wavelengths
6. Stationary with finite wavelength

Type 6 waves are the most interesting and observed in biological world (now called Turing Patterns!).

- **Gierer-Meinhardt RD Model**

  In 1972, they proposed a theory of biological pattern formation in which concentration maxima of pattern forming substances are generated through local self- enhancement in conjunction with long range inhibition. Local-Activation-Long-range-Inhibition (LALI).

  In their networks, the activator self-activates and activates his inhibitor, resulting in lateral inhibition. This network (like Turing's) is initiated by noise (e.g. does not need maternal info), it is thus self-organising and self-repairing. For instance, if the activated region is removed, the inhibitor-producing region is also removed. After the decay of the remnant inhibitor, a new activation sets in by autocatalysis. The inhibitor production also resumes, ensuring that the newly emerging activated region assumes the correct profile. Gradients, symmetrical arrangements and periodic distributions can be generated in this way. The maxima can have a spot- or a stripe-like shapes.

- **Pillars and Paradoxes of Turing Patterns**

  The four pillars of TP:
  1. Genetic network
  2. Differential diffusion* (even though this is arguable, see Marcon et al., 2016)
  3. Initial noisy expression (stochastic start)
  4. Final patterning

- **Robustness of Turing Patterns**
  Robustness corresponds to the probability of randomly picking a pattern-forming parameter.
  Turing Patterns operate over narrow parameters range. For example, the GM model is very parameter sensitive (e.g. you need very high differential diffusivity) How can evolution discover/pick right parameters?
  Using a mode that employs same promoters, you can use a Hill function with coopertivity (in the lab you can use HGF as X and a truncated version of HGF as Y). Diambra et al. (2014) applied stability analysis on such a model and discovered that increasing the cooperativity in the reaction function (that is making it more non-linear by increasing Hill coefficients) the parameter space increases. So greater steepness increases parameter space and even reduces the requirement for differential diffusion between activator and inhibitor

- **Networks Atlas Approaches**

How to find networks that can give rise to Turing Instabilities?

1. Identify the stable steady states of a given system of differential equations (consider stability of non-spatial system)
2. Study their dispersion relation (dependency of the real part of the largest eigenvalue of the Jacobian matrix of the system on the wavenumber q)
3. Add spatial diffusion: is it possible that a deviation from steady state does not decay into homogeneous steady state but become amplified (that is if eigenvalues become positive)? If yes, you get a Turing Pattern

Using a similar algorithm, Sholes et al. (2018) described two types of Turing instabilities:

1. Turing I: patterns with finite wavelength (for $q \rightarrow +\infty$, $Re(\lambda)<0$)
2. Turing II: patterns with infinite wavelength (for $q \rightarrow +\infty$, $Re(\lambda) +\infty$)

(Marcon et al., 2016) showed using stability analysis that in the presence of cell-autonomous factors (immobile), networks can form a pattern with equally diffusing signals. In particular, using their software RDnet, they described three types of networks:

Type I: only 2 mobile nodes, need differential diffusivity (the classic)
Type II: 2 mobile nodes, one immobile node, allows equal diffusivity
Type III: 2 mobile nodes, one immobile node, unconstrained diffusivity.

The robustness analysis of the networks shows that for unconstrained values of diffusivity, type III networks are more robust to parameters change. However, when diffusion ratios are fixed to experimentally measured values for Lefty and Nodal, type II networks are more robust to biologically relevant parameters. In particular, they say that Nodal and Lefty is a type II network, where Nodal activates Lefty indirectly via the immobile Smad receptor. This means that Nodal and Lefty do not necessarily need to have differential diffusivities to form a pattern but the combination of differential diffusivity and clearance rate constants increase the robustness (of Type II systems).

Importantly, Type III networks form patterns independently of specific diffusion rates; they have never been described before and challenge models of short-range activation and long-range inhibition that dominated most of developmental and theoretical biology for decades.

Modelling Plant Development - Dr. Giovanni Sena

Modelling Auxin Flux in Roots

- **Auxin Distribution in Roots**

Auxin regulates almost every aspect of plant growth and development. Its intracellular concentration is controlled by biosynthesis and degradation. In addition, there is an “auxin pool” that consists of the conjugates with sugars, amino acids, and peptides. Some of the conjugates reversely release auxin, enabling alternative methods to regulate auxin concentrations. Auxin concentrations are also affected by transport. Besides the long distance delivery through the phloem, auxin is transported across the cell by influx and efflux carriers, from the shoot to root with maximum concentration at the root tip. At the root tip, the auxin flow reverses, and shootward auxin transport occurs. An auxin gradient formed this way is indispensable for proper development, maintenance of the meristem, and cell identity. The formation of root hairs is auxin-dependent. Auxin controls not only the initiation of root hairs but also regulates their elongation. In *Arabidopsis thaliana*, auxin accumulates in atrichoblasts and it is supplied to trichoblasts.

The PIN proteins are secondary transporters acting in the efflux of the plant signal molecule auxin from cells. They are asymmetrically localized within cells and their polarity determines the directionality of intercellular auxin flow. *PIN* genes are found exclusively in the genomes of multicellular plants and play an important role in regulating asymmetric auxin distribution in multiple developmental processes
The activity of PIN proteins is regulated at multiple levels, including transcription, protein stability, subcellular localization and transport activity. Different endogenous and environmental signals can modulate PIN activity and thus modulate auxin-distribution-dependent development. A large group of PIN proteins, including the most ancient members known from mosses, localize to the endoplasmic reticulum and they regulate the subcellular compartmentalization of auxin and thus auxin metabolism.

Polar auxin transport (PAT) is sufficient to generate differential intercellular auxin gradients that guide root growth. In vertically orientated roots, root-ward auxin transport is thought to provide equal auxin gradients on all sides of the root. However, in horizontally orientated roots, auxin is redirected and transported more efficiently at the lower side, resulting in an unequal auxin distribution, and thus a steeper auxin gradient at the lower side. This is thought to inhibit cell elongation at the lower side of the root, resulting in downward growth of roots, although underlying mechanism are far from being understood.

PLETHORA (PLT) genes encode auxin-inducible transcription factors expressed in roots, which have been shown to be essential for determining differentiation in a graded manner. PLT protein levels correlate with the auxin response gradient at the root. High levels of PLT activity are required for stem cell niche identity and maintenance, intermediate levels are essential for cell growth and proliferation in the meristem zone (MZ), and low levels are needed for cell expansion in the elongation zone (EZ) and allow further cell differentiation in the differentiation zone (DZ).

- **The source decay model**
  Morphogen production at a localised source and overall decay.

The characteristic length of a morphogen gradient, λ, indicates the distance from the location of maximum concentration, C_0, at which the concentration has fallen to C_0/e (37%) of the maximum value. This can directly be related to the logarithmic slope of the gradient.

This source-decay system which is determined by a linear decay is arguably the simplest gradient generating mechanism that is used in biological development.

$$\delta C(x,t)/\delta t = D \nabla C(x,t) - d \times C(x,t).$$

Boundary conditions: state that at $x = 0$ there is a source of morphogen responsible for a morphogen influx, J, while at $x = L$ (the length of the system) the morphogen cannot leave the system.

$$D\,\delta C(0,t)/\delta x + J = 0$$

$$D\,\delta C(L,t)/\delta x = 0$$

The steady-state distribution of the morphogen concentration over space is (where λ = sqrt(D/d)):

$$C(x) = J\lambda /D(1-e^{-L/\lambda}) * e^{-x/\lambda}$$

when the size of the system is much larger than the characteristic length, this can be approximated as:

$$C = C_0\, e^{-x/\lambda}$$

Due to the very high diffusion coefficient of auxin and its low decay rate, the slope of the established morphogen gradient is extremely shallow. Characteristic length in Arabidopsis is 2.4 cm.

Such a characteristic length is far too large to convey positional information to the root, because concentrations would vary only 4% over the most distal 1 mm of the root tip, where differentiation into stem cell niche, MZ, EZ and DZ take place. This reveals how establishing an auxin gradient through diffusion and decay only is extremely unlikely.

- **The polarised transport model**
  Directed transport of the morphogen into the direction of a ‘dead end’, where a maximum will be formed. Model studies in the early eighties have shown that a unidirectional transport mechanism could underlie the establishment of auxin maxima. These models predicted the existence of polarly localised auxin efflux facilitators, which only much later were experimentally found, i.e. the family of PIN proteins. Unidirectional transport is not only able to generate a maximum, but also a morphogen gradient. The most direct mathematical

way to derive the effects of unidirectional transport is to consider a single cell file containing *n* cells that transport auxin directly into their neighbouring cells in the downward direction (from cell *n*=0 to cell *n*=*N*)
p = transport rate
q = diffusion rate

$$F_{0,1} = (p+q)\, C_0$$
$$F_{1,2} = (p+q)\, C_1$$
$$F_{1,0} = q\, C_1$$
$$F_{1,2} = q\, C_2$$

At equilibrium:

$$F_{0,1} = F_{1,0} \text{ and } F_{1,2} = F_{2,1}$$

In general, (after substituting identities) this results in:

$$C_n = (p+q/q)^n\, C_0$$

After rearrangements and adding characteristic length:

$$C_n = (p+q/q)^n\, C_0 = C_0 e\text{^}n/\ \lambda$$

But in Arabidopsis, measures parameters *p*=19 μm/s; *q*=1 μm/s. It follows that concentrations would drop 20-fold with each cell ($Cn\ /\ _{Cn-1}$ =20), the characteristic length λ being 1/ log ((*p* + *q*) /*q*) =0.33 cell length, or ≃ 5 μm. Thus, within such a cell file, auxin concentrations drop more than 19 orders of magnitude over the first 15 cells from the maximum. This limits the functionality of the auxin gradient to only very few cells close to the maximum, which plots the mathematically predicted auxin gradient using a vascular cell template to correct for cell lengths. The value of C_0 is determined by assuming a total amount of auxin within the vascular bundle equal that used in the source-decay mechanism. This has strong consequences for the positional information over the root, but in an opposite way as was observed for the source-decay mechanism: here the gradient is far too steep, while previously it was too shallow.

- **The reflux loop model** (Grieneisen et al., 2007)
  A combination of a downward and upward flux, linked to each other through a lateral flux, forming an ‘auxin capacitor’.
  Central assumption is that location of PIN proteins alone can describe polar auxin transport.
  A robust auxin gradient associated with the maximum, in combination with separable roles of auxin in cell division and cell expansion, is able to explain the formation, maintenance and growth of meristematic and elongation zones.
  Directional permeability (PINs) and diffusion can fully account for stable auxin maxima and gradients that can instruct morphogenesis.

  The stable concentration peak is accompanied by large auxin fluxes through the tissues; like a reversed fountain, the central flow down- wards connects to the upward flow in the epidermal tissue through a redistributing root cap. The auxin increase is due to reflux of the upward flow all along the meristem back into the central down- ward flow, which captures auxin within a flux-loop, causing an increase of the concentration maximum, until a steady state is reached. The overall PIN layout specifies the region of the maximum at the junction of the flows, which is positioned centrally just above the cap region.

The density of lateral PINs, both those in the border and epidermal files, are crucial for the reflux and are the most sensitive determinants for the auxin-storage capacity of the root and the magnitude of the maximum

Importantly, auxin equilibrates into the given profile solely as a result of the high auxin flows and PIN localizations. When all cells ubiquitously produce auxin, the same pattern is generated. Even when production is localized outside the auxin-maximum region, the pattern does not change.

To study the dynamics of zonation, they assumed (simplifying) that higher auxin levels promote cell division, whereas cell elongation is correlated with lower auxin levels.

Refining cellular model: To study the dynamics of zonation, they assumed (simplifying) that higher auxin levels promote cell division, whereas cell elongation is correlated with lower auxin levels. They mathematically implemented it

Refining cellular model II: they tried to add the pot model to their refined model. Each pixel in the cell is given a value of sigma. The value of sigma depends on the position: intercellular matrix = 0, cell wall=1. It the pixel is on the interaction between cells the interaction is significant. Then they used a Hamiltonian that is the sum of the interactions, plus the sum of the distance to the target area of the cell saucer. The Hamiltonian evaluates The energy of the system, the further from the target area for a cell the more energy they contribute to the system. The value that calculates the interaction between cells represents the surface tension. The term that evaluates the difference of position from the target position represents the turgor pressure. The target position is located by hand, this scale is rough and artificial. The move of cells and their interactions costs energy the dynamics of the system.

The dynamics proceed stochastically on the basis of a free energy minimization using a dynamic Monte Carlo simulation algorithm. To mimic pseudopod extensions and retractions of the cells, this algorithm randomly selects a lattice site (source site) and attempts to copy its index into a randomly chosen neighbouring site (target site). If this site belongs to a different bio- logical cell (i.e., if it has a different index), the algorithm checks the net energy changes associated with this move. While the index copying occurs in a deterministic manner for the case of energy decrease, it occurs stochastically with the following Boltzmann acceptance function for the case of energy increase:

Hamiltonian energy:

$$\mathcal{H} = \sum_{<nn>} J_{\sigma_{i,j},\sigma_{i\prime,j\prime}} + \sum_{\sigma=1}^{N_{cells}} \lambda \, (A_\sigma - A_{target}(\sigma))^2$$

$$J_{\sigma_{i,j},\sigma_{i\prime,j\prime}} = \begin{cases} 0 \; if \; \sigma_{ij} = \sigma_{i\prime j\prime} \\ a > 0 \;\; if \; \sigma_{ij} \neq \sigma_{i\prime j\prime} \end{cases}$$

σ_{ij} is a pixel on the lattice, part of σ

σ = 0 is apoplast (cell wall)

σ=1...N is symplast (inside cell)

J is surface energy between cells

<nn> = all the possible pixel-pixel borders

$A_\sigma = Area\ cell\ \sigma$

$A_{target}(\sigma) = typical\ area\ of\ cell\ \sigma\ (e.g. elongated\ cell)$

λ= Lagrange operator (>0), specifying the strength of the area constraint (cell sorting is sensitive to lattice discretization, here we used 1)

First term of Hamiltonian gives surface tension (if cells are different it outputs a positive value, that is an energy penalty, consider a =1; second term gives turgor pressure (it is large -and always positive - the larger is the deviation between the actual area of the cell and its "typical" size).

.

Accept transition ($\sigma_{i,j} \rightarrow \sigma_{i\prime,j\prime}$) with statistical mechanics probability:

$$p(\sigma_{i,j} \rightarrow \sigma_{i\prime,j\prime}) = \begin{cases} 1\ if\ \Delta H \leq -Y \\ e^{-\Delta H + Y/kT}\ if\ \Delta H > -Y \end{cases}$$

$\Delta H = H_{after} - H_{before}$

Y = yield of the cell wall resisting deformation

K = Boltzmann constant

T = simulation temperature that determines the magnitude of random biological fluctuations. A higher T causes large fluctuations allowing mesenchymal-like cell behaviours. For extremely high T (melting temperature), the cells tend to disintegrate as the system becomes dominated by random fluctuations.

- **The reflective flow model**

Assumptions

- All cells are equal functionally, in size and in shape
- Auxin flows from the shoot, is degraded in cells, diffuses, is actively transported by PINs
- We only consider PIN1, localised at one cell side

- It has been shown that auxin controls expression of PINs:
    - Normal auxin concentrations activate transcription of PINs genes (positive feedback) .
    - Very high auxin concentrations lead to degradation of PINs (as a negative feedback).
- There are N number of cells. N is the source and 1 is the root tip. The flow goes from N to 1, due to active PIN transport.

Demonstrations

- Under these assumptions there is a formation of an auxin maximum
- Auxin maximum is maintained during early root development
- Demonstrate changes in auxin distribution in root tip cut, root exposure to inhibitors of PINs and treatment of root with exogenous auxin.

Equations

1. Rate of Auxin flux into the cell

$$V_\alpha = \alpha$$

$\alpha = intensity\ of\ auxin\ flux\ from\ the\ shoot$

2. Auxin degradation

$$V_d(a) = K_d a$$

$a = auxin\ concentration\ (normalised\ to\ flux)$
$K_d = degradation\ rate\ constant$

3. Auxin diffusion

$$V_{diff}(a) = Da$$

$D = diffusion\ rate\ constant$

4. Active auxin transport

$$V_a(a, PIN) = K_0 a\ PIN$$

$K_0 = active\ transport\ rate\ constant$
$PIN = concentration\ of\ PIN1\ protein$

5. Auxin-induced PIN synthesis (Hill function), it is zero at [Auxin]=0 and then increases monotonically

$$V_{s,PIN}(a) = k_1 \frac{(\frac{a}{q_1})^{h_1}}{1 + (\frac{a}{q_2})^{h_2}}$$

$k_1 = synthesis\ rate\ constant$
$q_1 = threshold\ for\ auxin - induced\ PIN\ synthesis$
$q_2 = threshold\ for\ saturation\ of\ auxin - induced\ PIN\ synthesis$
$h_{1,2} = Hill\ coefficients$

6. Rate of PIN degradation

$$V_{d,PIN}(a, PIN) = k_2 PIN(1 + (\frac{a}{q_3})^{h_2})$$

$k_2 = basal\ PIN\ degradation\ rate\ constant$
$q_3 = threshold\ for\ auxin - induced\ PIN\ degradation$

Alberto Scarampi del Cairo 12/2018

1-D Model

The one-dimensional (1D) model describes auxin distribution in a linear array of cells located along the central root axis

N = 1 → last cell of columella root cap (root tip)
Then Cells are from 2 to N
Cell = N → cell at the shoot-to-root junction
The model is identical for all the cells, except cell N=1 and Cell N, which have boundary conditions. For cell 1, passive diffusion is necessarily defined only to cell 2 due to the physical boundary conditions - there being no adjacent cell in the other direction. For cell N the boundary condition is that the net effect of active transport and passive diffusion is defined as occurring in only one direction: from the unmodeled shoot toward the root tip, modeled with equation 1.

Auxin from the shoot first enters the Nth cell and then spreads through the linear array of cells by diffusion and active transport.

For cell 1 (root tip):

$$\frac{da_1}{dt} = D(a_2 - a_1) + K_0 a_2 PIN_1 - K_d a_1$$

$$\frac{dPIN_1}{dt} = \frac{k_1(\frac{a_1}{q_1})^{h_1}}{(1 + (\frac{a_1}{q_1})^{h_1}) - k_2 PIN_1(1 + (\frac{a_1}{q_3})^{h_2})}$$

For root cells:

$$\frac{da_i}{dt} = D(a_{i+1} + a_{i-1} - 2a_i) + K_0(a_{i+1}PIN_{i+1} - a_i PIN_i) - K_d a_i$$

$$\frac{dPIN_i}{dt} = \frac{k_1(\frac{a_i}{q_1})^{h_1}}{(1 + (\frac{a_i}{q_2})^{h_1}) - k_2 PIN_i(1 + (\frac{a_i}{q_3})^{h_2})}$$

For cell N (junction):

$$\frac{da_N}{dt} = \alpha + D(a_{N-1} - a_N) - K_0 a_N PIN_N - K_d a_N$$

$$\frac{dPIN_N}{dt} = \frac{k_1(\frac{a_N}{q_1})^{h_1}}{(1 + (\frac{a_N}{q_2})^{h_1}) - k_2 PIN_N(1 + (\frac{a_N}{q_3})^{h_2})}$$

Results
There is a presence of an auxin maximum at the root cap.
There is also a gradient. Mechanism of the gradient:

- In cells with low auxin concentration, the PIN expression in enhanced. This results in rapid auxin accumulation at the tip (N=1).
- High amount of auxin diffuses to cell 2. Maxima is now found in cell2. High concentrations of auxin in cell2, inhibit expression of PINs. This leads to a decrease of active transport to cell 1 and therefore auxin diffuses to cell 3.
- The maxima shift away from the end of the root until the reflected auxin flow becomes balanced.

Advantage over reflux-loop:

- Only one anatomical element needs to be present for the reflected flow mechanism: cells with polarized PIN protein localization, regulated by auxin.
- Sometimes plants that have not developed or where the structure has been disrupted do not have a root tip (with specific PINs and cell shape/size for the tip). Reflected flow model doesn't need that. Reflux-loop does to close the loop of auxin.

- **Comparison and discussion**

  Recently the reflux loop and reflective flow models were combined in a dual-mechanism model that could capture the benefits of both systems. To introduce the reverse fountain PIN2 and PIN3 were included into the reflective flow model and localised in the same manner as the model in Grieneisen et al. (2007); however, the levels of these proteins were controlled by auxin. This dual-mechanism model was able to simulate the realistic regeneration of the root apical meristem upon decapitation (previously seen in the reflective flow model but not the reflux loop model). When a basal rate of auxin synthesis was included, this was able to maintain an auxin maximum at the root apex when the shoot was removed.

  However, maintaining an auxin maximum at the quiescent centre (QC) position is unlikely to be the only requirement for maintaining root growth. Growth is controlled by maintaining the rate of cell division in the meristematic zone (MZ) of the root above QC, cell elongation in the elongation zone (EZ) and above this, cell differentiation in the differentiation zone (DZ). The activity of these domains are controlled by the level of the PLETHORA (PLT) family of genes, which are believed to be controlled mainly through a gradient of auxin.

  Now the question is: does auxin produce a well enough defined gradient to define these zonations within the root?

  It was found that only the reflux loop model was able to form an exponentially increasing auxin gradient that spanned the entire MZ and part of the EZ; the auxin gradient was much too shallow with the source-decay mechanism, or much too steep with the uni-directional transport mechanism.