

Bifurcation in epigenetics: implications in development, proliferation and diseases

Daniel Jost¹

¹*Laboratoire de Physique, École Normale Supérieure de Lyon, CNRS UMR 5672, Lyon, France*

Cells often exhibit different and stable phenotypes from the same DNA sequence. Robustness and plasticity of such cellular states are controlled by diverse transcriptional and epigenetic mechanisms, among them the modification of biochemical marks on chromatin. Here, we develop a stochastic model that describes the dynamics of epigenetic marks along a given DNA region. Through mathematical analysis, we show the emergence of bistable and persistent epigenetic states from the cooperative recruitment of modifying enzymes. We also find that the dynamical system exhibits a critical point and displays, in presence of asymmetries in recruitment, a bifurcation diagram with hysteresis. These results have deep implications for our understanding of epigenetic regulation. In particular, our study allows to reconcile within the same formalism the robust maintenance of epigenetic identity observed in differentiated cells, the epigenetic plasticity of pluripotent cells during differentiation and the effects of epigenetic misregulation in diseases. Moreover, it suggests a possible mechanism for developmental transitions where the system is shifted close to the critical point to benefit from high susceptibility to developmental cues.

PACS numbers:

Cellular differentiation occurs during the development of multicellular organisms and leads to the formation of many different tissues where gene expression is modulated without modification of the genetic information [1]. These modulations are in part encoded by biochemical tags, called epigenetic marks, that are set down at the chromatin level directly on DNA or on histone tails. These marks are directly or indirectly involved in the local organization and structure of the chromatin fiber, and therefore may modulate the accessibility of DNA to transcription factors or enzymatic complexes, playing a fundamental role in the transcriptional regulation of gene expression.

Each tissue is characterized by a distinct epigenetic pattern [2] that is mainly shaped during cellular differentiation by developmental signals driven mainly by transcription factors. For a differentiated cell, these specific signals disappear and the global epigenetic state of a cell is robustly maintained throughout the cell life and in its daughter cells. This maintenance, despite the fast turnover rate of epigenetic marks [3] or the dilution of epigenetic information during cell division [4], implies the existence of mechanisms to avoid the rapid loss of epigenetic information.

At the gene level, there are active or inactive epigenetic marks that influence the transcriptional activity of the gene [5]. A coherent activity needs that the gene promoter is covered by a majority of active or inactive marks. Recently, many efforts both experimentally [6–8] and theoretically [9–18] have been dedicated to study the mechanisms responsible for the spreading and maintenance of an epigenetic state at the gene level for different biological organisms and contexts. In particular, it has been shown for many examples, like for the locus *MAT* in fission yeast [9, 12] or for the vernalization of the floral repressor *FLC* in *Arabidopsis* [7, 15], that long-range or cooperative in-

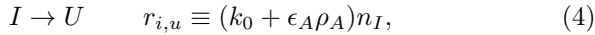
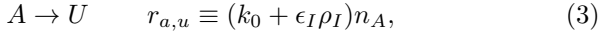
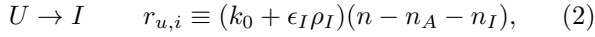
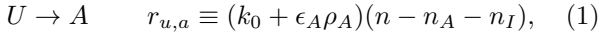
teractions between epigenetic marks are necessary for the emergence of stable coherent states. Of particular interest is also the work of the Sengupta's group showing that differences between activating and silencing rates might be responsible for hysteresis in epigenetic silencing by the SIR system in budding yeast [11, 18].

In this Rapid Communication, we develop a general formalism that describes the dynamics of epigenetic marks and that accounts, at the same time, for the plasticity and for the robustness of the epigenome. Using statistical and non-linear physics methods, we characterize the emergence of coherent epigenetic states from the recruitment of modifying enzymes, and we study the bifurcations occurring in the system when altering the recruitment intensities. Finally, we discuss in details the implications of our findings in concrete biological contexts during development, differentiation or disease. In particular, we propose that the experimentally-observed regulation of the cell cycle during development might have a strong impact on the efficiency of epigenetic switches.

Model. Inspired by the model introduced by Dodd et al for the mating-type switch of the fission yeast (*S. pombe*) [9, 10, 12], we consider a DNA region, consisting of n nucleosomes, that is located between two boundaries that epigenetically isolate the region from neighboring DNA [13, 19]. We assume that the epigenetic state of each nucleosome can fluctuate between three different states: unmarked (U), active (A) and inactive (I). Active marks would be for example associated with acetylation of the lysine 9 of histone 3 (H3K9), and inactive marks with tri-methylation of H3K9 (HP1-type chromatin) or with the methylation of H3K27 (Polycomb-type chromatin) [5]. However, the actual epigenetic marks associated with each state are not important for our purpose and we only need two well separated kinds of marks that can be exchanged by passing through an intermediate

state to each other by passing through an unmarked state ($I \rightleftharpoons U \rightleftharpoons A$).

We assume that the nucleosomal state can be modified by two mechanisms: i) Recruitment of modifying enzymes (for example, histone demethyl- or methyltransferases, and histone deacetyl- or acetyltransferases) by surrounding active or inactive nucleosomes that occurs at a rate $\epsilon_X \rho_X$ where ρ_X is the local density of the modified state X (A or I) that, by considering a spatial mean-field approximation, we identified to the *global* density n_X/n with n_X the corresponding number of nucleosomes. ii) Random transitions between states that occur at a rate k_0 and that represents recruitment-independent enzymatic activity, nucleosome turnover or dilution due to replication. Note that to facilitate the analysis we lumped into k_0 processes (turnover vs replication) with presumably different time-scales and statistical properties. However previous works on epigenetic modeling [9, 10] suggest that the main conclusions of our study should not depend on that approximation. The corresponding system of biochemical reactions is composed of four possible state transitions given below with their respective propensities:



where, for simplicity, we assumed that the rates of random transitions are similar for each reaction and that possible discrepancies between A and I occur at the recruitment level.

Analogy with an Ising model and phase transitions.

From a theoretical perspective, we remark that this model is formally very similar to a zero-dimensional 3-states Ising model where nucleosomal states represent spins (for example $I = -1$, $U = 0$, $A = +1$), recruitment (ϵ_X) corresponds to coupling between spins (J) and random transitions (k_0) are associated with thermal fluctuations ($k_B T$).

This analogy suggests that a good observable for our system will be the magnetization $m = (n_A - n_I)/n$. In biological terms, this magnetization could be interpreted as the relative activity of the DNA region if we consider that, in addition to the favorable effect of active marks, the presence of inactive marks penalizes the activity. In the following, we will consider m as the relevant observable of our system.

As it is well known that zero-dimensional Ising models do exhibit phase transitions between an ordered and a disordered phases [20], we expect our system to display such dramatic changes. To illustrate that we consider the simple mass-action model (equivalent to the mean-field approximation of the Ising model) that captures the mean dynamics of the epigenetic marks in the case of

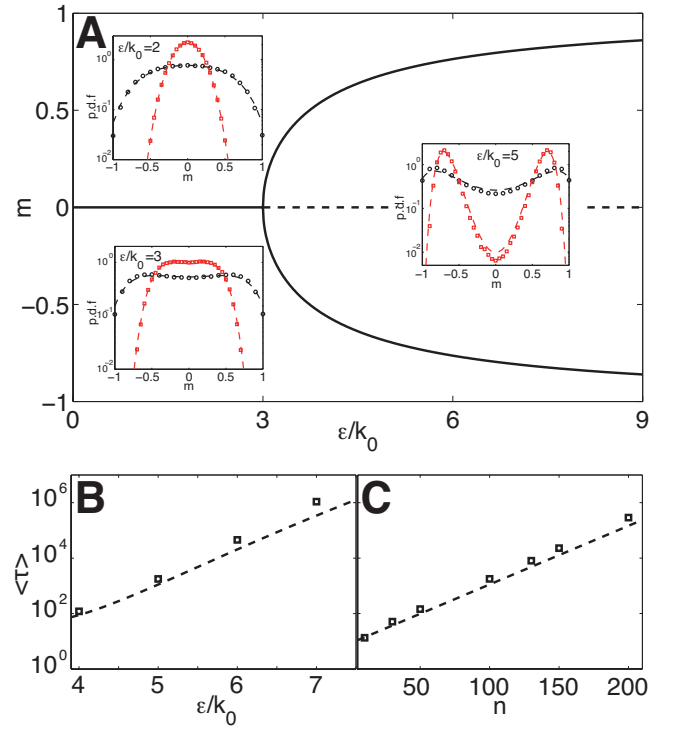


FIG. 1: (Color online) **Symmetric regime.** (A) Bifurcation diagram for m as a function of ϵ/k_0 . Full (dashed) lines represent stable (unstable) fixed points of the dynamical system. (Insets) Probability distribution functions (p.d.f.) of m for $n = 10$ (black circles) or $n = 100$ (red squares) computed from the Fokker-Planck approximation (dashed lines) or from stochastic simulations (squares). (B,C) Mean first passage time ($\langle \tau \rangle$) (in k_0^{-1} -unit) to switch from m_- to m_+ as a function of ϵ (B, for $n = 100$) or of n (C, for $\epsilon/k_0 = 5$), computed from Eq.9 (dashed lines) or from simulations (squares). Standard errors on the estimation of $\langle \tau \rangle$ are smaller than the symbol size.

symmetric recruitment ($\epsilon_A = \epsilon_I \equiv \epsilon$):

$$\frac{d\rho_A}{dt} = (k_0 + \epsilon \rho_A)(1 - \rho_A - \rho_I) - (k_0 + \epsilon \rho_I)\rho_A, \quad (5)$$

$$\frac{d\rho_I}{dt} = (k_0 + \epsilon \rho_I)(1 - \rho_A - \rho_I) - (k_0 + \epsilon \rho_A)\rho_I. \quad (6)$$

At steady-state, the previous dynamical system has at most three relevant fixed points that, in term of the effective magnetization m , are given by $m_0 = 0$ ($\forall \epsilon$) and $m_{\pm} = \pm(k_0/\epsilon)\sqrt{(\epsilon/k_0 + 1)(\epsilon/k_0 - 3)}$ (for $\epsilon > 3k_0$). Fig.1A shows the classical supercritical pitchfork bifurcation occurring at the critical point $\epsilon_c = 3k_0$. For weak recruitment ($\epsilon < \epsilon_c$), m_0 is the only stable fixed point and the activity of the DNA region is not clearly defined. At the critical point, m_0 turns unstable and stable coherent activities (m_{\pm} for active or inactive) are only observed for strong recruitments ($\epsilon > \epsilon_c$).

Distribution and stability of epigenetic states. To go beyond the mean-field approximation, we aim at the full

distribution of probability for m . For simplicity, we assume symmetric coupling in the following. The general situation will be discussed in the next section. The starting point is to write the master-equation related to the set of biochemical reactions given in Eqs.1-4, but for the new variables m and $s = (n_A + n_I)/n$ where s represents the density of marked nucleosomes. Then, assuming that s is always large and that its dynamics is fast compared to the one of m allows to perform a time-scale separation and leads, in the limit of large n , to the Fokker-Planck equation for the probability $P(m)$ (see [21] for details)

$$\frac{\partial P}{\partial t} = -\frac{\partial}{\partial m} ([w_+(m) - w_-(m)]P - (1/2n)\frac{\partial}{\partial m} ([w_+(m) + w_-(m)]P)) \quad (7)$$

with $w_+ = (r_{u,a} + r_{i,u})/n$ and $w_- = (r_{u,i} + r_{a,u})/n$, the propensities to increase (using reactions 1 and 4) or decrease (2 and 3) m by 1. Eq. 8 is a classical Fokker-Planck equation for the heterogeneous diffusion of a particle within an unidimensional potential [23]. At steady-state, the probability distribution is then given by

$$P_\infty(m) = \frac{1}{Z} \frac{\exp \left[2n \int_{-\infty}^m dm' \left(\frac{w_+(m') - w_-(m')}{w_+(m') + w_-(m')} \right) \right]}{w_+(m) + w_-(m)} \quad (8)$$

with Z a normalization factor. Insets of Fig.1A shows the very good agreement between the approximated solutions given by Eq.8 and the distributions computed from exact stochastic simulations of the full system (1-4) using the Gillespie algorithm [24]. While below the critical point, we observe a unimodal distribution centered around m_0 , for high coupling ($\epsilon > \epsilon_c$), the distribution is bimodal and peaked around the coherent epigenetic states m_\pm , the width of each peak being proportional to $[n(\epsilon - \epsilon_c)/k_0]^{-1/2}$ [21]. At the critical point, we observe a nearly flat distribution characteristics of phase transitions.

In presence of bimodality, we quantify the stability of a coherent epigenetic state by computing the mean first passage time $\langle \tau \rangle$ to switch from m_- to m_+ . Using the Fokker-Planck formalism introduced above, we show that [21]

$$\langle \tau \rangle \approx \frac{18\pi}{(\epsilon - \epsilon_c)\sqrt{3(\epsilon/k_0 + 3)}} \exp[V(0) - V(m_-)] \quad (9)$$

with $V(m) = -\log P_\infty$ the effective potential corresponding to the steady-state "epigenetic landscape". As already pointed out by Micheelsen and al. [12], we find that, similar to transition state theory [25], $\langle \tau \rangle$ mainly depends on the "energy barrier" between the starting state ($m = m_-$) and the transition state ($m = 0$). As the coupling increases, the barrier is higher and the epigenetic state is more stable (see Fig. 1 B). We also remark that $\langle \tau \rangle$ scales exponentially with the size n of the system

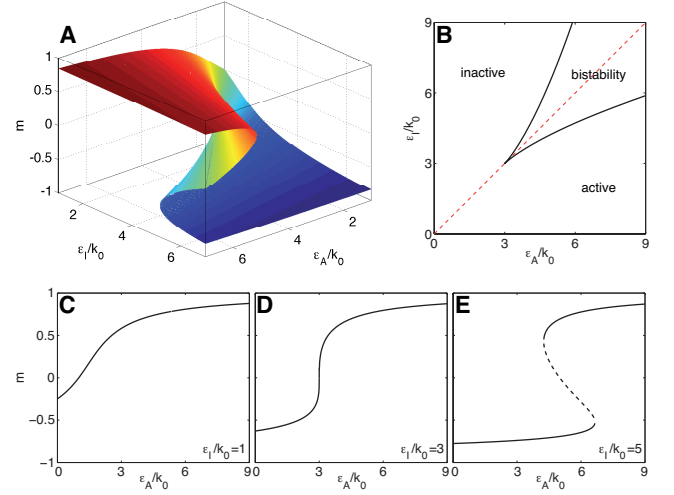


FIG. 2: (Color online) **Asymmetric regime.** (A) Cusp catastrophe surface representing the fixed points of the dynamical system as a function of ϵ_A/k_0 and ϵ_I/k_0 . (B) Stability diagram and boundaries between the mono- and bistable regions. (C,D,E) Bifurcation diagrams for m as a function of ϵ_A/k_0 for fixed values of ϵ_I/k_0 ($=1$, C; $=3$, D; $=5$, E). Legend as in Fig.1A.

(see Fig. 1 C). Such very long relaxation time are typical to system with long-range interactions [20, 26] and means that the larger the system, the more stable the epigenetic state.

Asymmetry and cusp catastrophe. In this section, we consider the general situation where recruitments of enzymes by active or inactive marks are different. In the next, we will focus on the study of the generalization of the dynamical system Eq.(5-6) that captures the main characteristics of asymmetric recruitment.

At steady-state, the system has at most three fixed points and the bifurcation diagram shapes as a cusp catastrophe surface (Fig. 2A), characteristics of dynamical systems with asymmetry [27] and which is also observed for example in the epigenetic SIR system in budding yeast [11, 18] or in insect outbreaks [28]. For every pairs of parameters (ϵ_A, ϵ_I) , the dynamical system is either monostable or bistable (with an unstable fixed point). Using equality conditions on the nullclines and their first derivatives [21], we find an exact parametric expression for the boundary between the mono- and bistable regions (Fig.2B). Depending on the relative asymmetry between ϵ_A and ϵ_I , the single (stable) fixed point of the monostable region corresponds to an active or inactive epigenetic state. Bistability, with the coexistence of an active and of an inactive coherent activity, is observed only for strong recruitments ($\epsilon_A, \epsilon_I > \epsilon_c$) and small asymmetry.

Figures 2 C-E show the bifurcation diagram of the system when increasing ϵ_A/k_0 for fixed values of ϵ_I/k_0 .

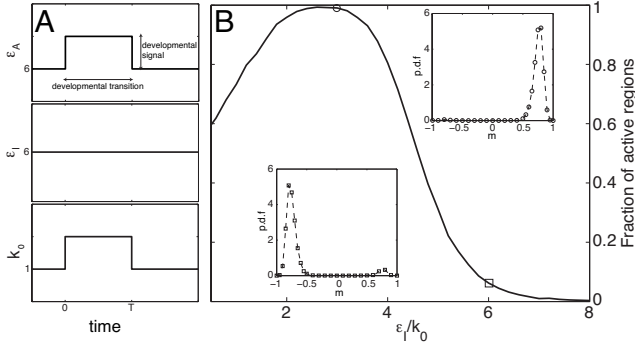


FIG. 3: (Color online) **Epigenetics and criticality.** (A) Proposed strategy for guided epigenetic switching at developmental transitions when temporary asymmetric signal ($\epsilon_A > \epsilon_I = 6$) and modification of k_0 are applied during a finite time period T . (B) Probability to be active ($m > 0$) after a time $t = 100$ starting from a silenced state ($m = m_-$) as a function of the effective recruitment strength ϵ_I/k_0 for $\epsilon_A = 8$ and $T = 5$. (Insets) Corresponding probability distribution function of m at $\epsilon_I/k_0 = 3$ (circles) or 6 (squares).

If $\epsilon_I < \epsilon_c$, the system stays in the monostable region and the unique fixed point goes continuously from an almost incoherent state ($m \sim 0$) to a coherent active state (Fig. 2C). For stronger recruitment ($\epsilon_I > \epsilon_c$), the system crosses the bistable region making a typical hysteresis curve with two saddle-node bifurcations (Fig. 2E). For example, starting from an inactive state, as we increase the recruitment of active marks, the system stays in the inactive state (even for $\epsilon_A > \epsilon_I$) until it switches abruptly to an active state. When crossing the cusp point (Fig. 2D), the system becomes ultra-sensitive and weak asymmetries lead to important changes in the epigenetic state.

Implications in development, proliferation and diseases. Recent quantitative experiments in mouse [8] have shown that bistable activities take place in an epigenetically-controlled system and that modifications in the epigenome can be stably imprinted by transcriptional factors. In particular, this study suggests that, at least for some chromatin regions, symmetric or weakly asymmetric recruitment may correspond to normal biological situations, and that temporary strong asymmetries in the recruitment of modifying enzymes may lead to long-term modifications of the current epigenetic state.

In vivo, differentiated cells exhibit a robust phenotype within the population and in time. The analysis of the symmetric regime of our formalism (Fig. 1) suggests that such robustness needs strong recruitments ($\epsilon > \epsilon_c$) and large cooperative units ($n \gg 1$) in order to stabilize coherent epigenetic states (active or inactive) and to avoid spurious switches between coherent states. Moreover, Fig. 2 suggests that epigenetic states of differentiated cells are also stable against small fluctuations of the recruitment couplings. Such fluctuations induce asym-

metries in the system that do not impact significantly on the epigenetic state as long as ϵ_A and ϵ_I remains in the bistability region. This property is crucial for the maintenance of a robust phenotype in weakly fluctuating environments.

However, when modifications of the environment are important, it would be beneficial to adapt to the current environment by modifying the epigenetic state, like for example for plants at seasonal transitions [29]. Misadaptation to the current environment may lead to stress signalings that might result in asymmetric recruitment forcing the current epigenetic state towards an adapted state, like in the cold-induced silencing of the *FLC* locus in plant vernalization [7]. Fig. 2 E suggests that such asymmetric signals have to be strong enough to allow an epigenetic switch. Once the switch is performed, the hysteretic shape of the bifurcation insures that the epigenome would remain stably in its new state. This property allows adaptation but only if needed, i.e. only when the organism is strongly misadapted to the current environment.

Many diseases have been related to epigenetic perturbations, from neurologic disorder to cancer [30]. Within our formalism, these perturbations could be interpreted by anomalous values for the recruitment (ϵ/k_0) that shift the steady-state of the system below or close to the critical point, and that make the epigenetic state incoherent or unstable and very sensitive to external noise. For example, cancer is often associated with an increase in the frequency of replication during tumorigenesis [31]. In our model, this means an increase of the random transition rate k_0 due to replication. This may modify the position of the critical point and therefore may lead to epigenetic instability and misregulation of some tumor suppressor proteins for example.

During development, cellular differentiation occurs in successive stages. Cells pass through series of developmental transitions where epigenetic states are modified by developmental cues that presumably force locally the desired state. Previously, we saw that, for differentiated cells, the hysteretic shape of the bifurcation may be valuable for the buffering of environmental fluctuations or for adaptation. Within the context of development, this could represent a hindrance at developmental transitions when the epigenetic state has to efficiently switch in a short time window. Our formalism suggests that a possible strategy to overcome this apparent issue would be to shift temporarily the system at or close to the critical point during developmental transition. Indeed, Fig. 2D shows that at the cusp point, the epigenetic state is very sensitive to weak asymmetries. Going back to the original analogy between our model and an Ising model with phase-transition, this is related to the concept of susceptibility that is maximal at the critical point. In many organisms, experimental evidence suggests that the cell cycle length is regulated during development alternating between short cycles during developmental transi-

tions and longer cycles between the transitions and in the differentiated state [32]. In our model this means that at developmental transitions, k_0 may be increased and therefore that the effective recruitment strength (ϵ/k_0) may be reduced and may become closer to the critical value. Under this hypothesis, we test the ability to switch between two coherent states when applying a weak asymmetric signal (the developmental signal) during a finite time period (the developmental transition), by running Gillespie simulations. As expected, Fig.3 shows that the switching efficiency is optimal when $\epsilon_I/k_0 \sim 3$ during the transition. Compared to the situation where k_0 is not changed during the transition and where only few cells have switched their epigenetic state, going close to the critical point leads to stable switch for almost all the cells. Provided the necessary experimental verifications of the proposed strategy, our results strongly suggest that it could be advantageous, during developmental transitions, of being close to criticality to benefit from the high sensitivity to external stimuli, as already observed in various other biological systems [33].

I thank Cédric Vaillant for fruitful discussions.

-
- [1] C.D. Allis, T. Jenuwein, and D. Reinberg (2007) Epigenetics. *Cold Spring Harbor Laboratory Press*.
 - [2] A. Doi, I.-H. Park, B. Wen, P. Murakami, M.J. Aryee, R. Irizarry, B. Herb, C. Ladd-Acosta, J. Rho, S. Loewer, J. Miller, T. Schlaeger, G. Q. Daley, and A.P. Feinberg (2009) Differential methylation of tissue- and cancer-specific CpG island shores distinguishes human induced pluripotent stem cells, embryonic stem cells and fibroblasts. *Nature Genetics*, **41**: 1350-1353.
 - [3] R.B. Deal, J.G. Henikoff, and S. Henikoff (2010) Genome-Wide Kinetics of Nucleosome Turnover Determined by Metabolic Labeling of Histones. *Science*, **328**: 1161-1164.
 - [4] R. Margueron, and D. Reinberg (2010) Chromatin structure and the inheritance of epigenetic information. *Nat. Rev. Genet.*, **11**: 285-296.
 - [5] T. Jenuwein, and C.D. Allis (2001) Translating the histone code. *Science*, **293**: 1074-1080.
 - [6] J.Z. Kelemen, P. Ratna, S. Scherrer, and A. Becksei (2010) Spatial epigenetic control of mono- and bistable gene expression. *PLoS Biol.*, **8**: e1000332.
 - [7] A. Angel, J. Song, C. Dean, and M. Howard (2011) A polycomb-based switch underlying quantitative epigenetic memory. *Nature*, **476**: 105-108.
 - [8] N.A. Hathaway, O. Bell, C. Hodges, E.L. Miller, D.S. Neel, and G.R. Crabtree (2012) Dynamics and Memory of Heterochromatin in Living Cells. *Cell*, **149**: 1447-1460.
 - [9] I.A. Dodd, M.A. Micheelsen, K. Sneppen, and G. Thon (2007) Theoretical Analysis of Epigenetic Cell Memory by Nucleosome Modification. *Cell*, **129**: 813-822.
 - [10] D. David-Rus, S. Mukhopadhyay, J.L. Lebowitz, and A.M. Sengupta (2009) Inheritance of epigenetic chromatin silencing. *J. Theor. Biol.*, **258**: 112-120.
 - [11] S. Mukhopadhyay, V.H. Nagaraj, and A.M. Sengupta (2010) Locus dependence in epigenetic chromatin silencing. *BioSystems*, **102**: 49-54.
 - [12] M.M. Micheelsen, N. Mitarai, K. Sneppen, and I.B. Dodd (2010) Theory for stability and regulation of epigenetic landscapes. *Phys. Biol.*, **7**: 026010.
 - [13] I.B. Dodd, and K. Sneppen (2011) Barriers and silencers; a theoretical toolkit for control and containment of nucleosome-based epigenetic states. *J. Mol. Biol.*, **414**: 624-637.
 - [14] A. Becksei, S. Scherrer, J.Z. Kelemen, and A.E. Murray (2011) Modeling of chromosomal epigenetic silencing processes. *Transcription*, **2**: 173-178.
 - [15] A. Satake, and Y. Iwasa (2012) A stochastic model of chromatin modification: cell population coding winter memory in plants. *J. Theor. Biol.*, **302**: 6-17.
 - [16] K. Sneppen, and I.B. Dodd (2012) A simple histone code opens many paths to epigenetics. *PLoS Comp. Biol.*, **8**: e1002643.
 - [17] H. Binder, L. Steiner, J. Przybilla, T. Rohlf, S. Prohaska, and J. Galle (2013) Transcriptional regulation by histones modifications: toward a theory of chromatin reorganization during stem cell differentiation. *Phys. Biol.*, **10**: 026006.
 - [18] A. Dayarian, and A.M. Sengupta (2013) Titration and hysteresis in epigenetic chromatin silencing. *Phys. Biol.*, **10**, 036005.
 - [19] M. Gaszner, and G. Felsenfeld (2006) Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat. Rev. Genet.*, **7**: 703-713.
 - [20] D. Mukamel, S. Ruffo, and N. Schreiber (2005) Breaking of Ergodicity and Long Relaxation Times in Systems with Long-Range Interactions. *Phys. Rev. Lett.*, **95**: 240604.
 - [21] see Supplementary Materials.
 - [22] N. van Kampen (2001) Stochastic Processes in Physics and Chemistry. *North-Holland*.
 - [23] H. Risken (1989) The Fokker-Planck equation. *Springer-Verlag*.
 - [24] D.T. Gillespie (1977) Exact stochastic simulation of coupled chemical reactions. *J. Phys. Chem.*, **81**: 2340-2361.
 - [25] K.J. Laidler (1969) Theories of Chemical Reaction Rates. *McGraw-Hill*.
 - [26] F. Bouchet, S. Gupta, and D. Mukamel (2010) Thermodynamics and dynamics of systems with long-range interactions. *Physica A*, **389**: 4389-4405.
 - [27] S.H. Strogatz (1994) Nonlinear dynamics and chaos. *Westview Press*.
 - [28] D. Ludwig, D.D. Jones, and C.S. Holling (1978) Qualitative Analysis of Insect Outbreak Systems: The Spruce Budworm and Forest. *J. Anim. Ecol.*, **47**: 315-332.
 - [29] D. Horvath (2009) Common mechanisms regulate flowering and dormancy. *Plant Science*, **177**: 523-531.
 - [30] A. Portela, and M. Esteller (2010) Epigenetic modifications and human diseases. *Nat. Biotechnol.*, **28**: 1057-1068.
 - [31] M. Malumbres, and M. Barbacid (2009) Cell cycle, CDKs and cancer: a changing paradigm. *Nat. Rev. Cancer*, **9**: 153-166.
 - [32] Y. Budirahardja, and P. Gonczy (2009) Coupling the cell cycle to development. *Development*, **136**: 2861-2872.
 - [33] T. Mora, and W. Bialek (2011) Are biological systems poised at criticality? *J. Stat. Phys.*, **144**: 268-302.

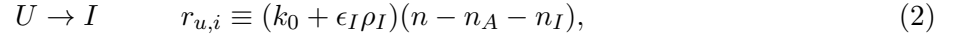
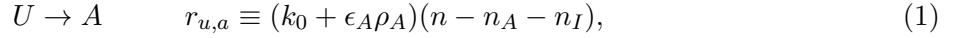
Supporting material

Bifurcation in epigenetics: implications in development, proliferation and disease

Daniel Jost

Laboratoire de Physique, École Normale Supérieure de Lyon, CNRS UMR 5672, Lyon, France

As a reminder, we recall the biochemical reactions corresponding to the epigenetic system of interest, with their respective propensities:



1 Dynamical system

In this section, we detail the study of the mass-action model corresponding to reactions 1-4:

$$\frac{d\rho_A}{dt} = (k_0 + \epsilon_A \rho_A)(1 - \rho_A - \rho_I) - (k_0 + \epsilon_I \rho_I)\rho_A, \quad (5)$$

$$\frac{d\rho_I}{dt} = (k_0 + \epsilon_I \rho_I)(1 - \rho_A - \rho_I) - (k_0 + \epsilon_A \rho_A)\rho_I. \quad (6)$$

In order to introduce the effective magnetization $m = \rho_A - \rho_I$, we change the variables from (ρ_A, ρ_I) to $(m, s \equiv \rho_A + \rho_I)$. This leads to the following dynamical system

$$\begin{aligned} \frac{dm}{dt} &= (r_{u,a} + r_{i,u} - r_{u,i} - r_{a,u})/n \\ &= \left[k_0 + \epsilon_A \left(\frac{s+m}{2} \right) \right] \left[1 - \left(\frac{s+m}{2} \right) \right] - \left[k_0 + \epsilon_I \left(\frac{s-m}{2} \right) \right] \left[1 - \left(\frac{s-m}{2} \right) \right] \end{aligned} \quad (7)$$

$$\begin{aligned} \frac{ds}{dt} &= (r_{u,a} + r_{u,i} - r_{a,u} - r_{i,u})/n \\ &= \left[2k_0 + \epsilon_A \left(\frac{s+m}{2} \right) + \epsilon_I \left(\frac{s-m}{2} \right) \right] (1-s) - k_0 s - \left(\frac{\epsilon_A + \epsilon_I}{4} \right) (s^2 - m^2) \end{aligned} \quad (8)$$

This system has at most three positive fixed points. Their general analytical expressions are quite cumbersome and not informative so we do not give their exact forms here. For the symmetric regime where $\epsilon_A = \epsilon_I \equiv \epsilon$, we find

$$m_0 = 0 \quad s_0 = 2/3, \quad (9)$$

$$m_+ = (k_0/\epsilon) \sqrt{(\epsilon/k_0 + 1)(\epsilon/k_0 - 3)} \quad s_+ = (\epsilon - 1)/\epsilon, \quad (10)$$

$$m_- = -(k_0/\epsilon) \sqrt{(\epsilon/k_0 + 1)(\epsilon/k_0 - 3)} \quad s_- = (\epsilon - 1)/\epsilon, \quad (11)$$

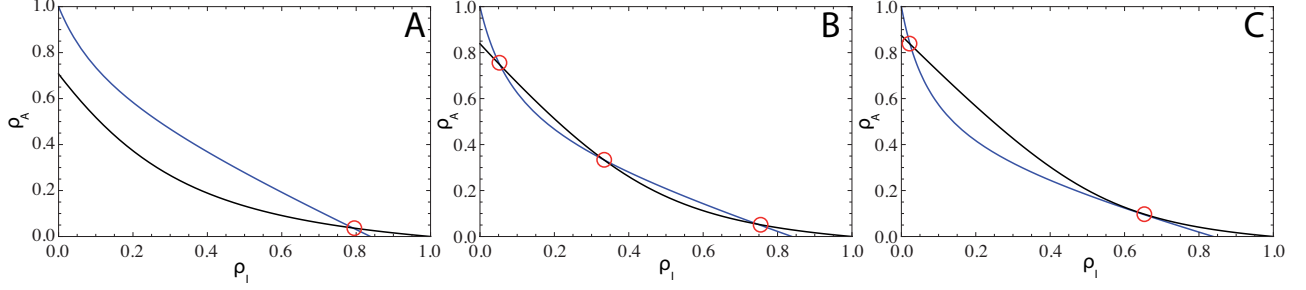


Fig. S 1: Nullclines $\rho_{A,1}$ (black, Eq.12) and $\rho_{A,2}$ (blue, Eq.13) of the dynamical system for $\epsilon_I = 5k_0$ and $\epsilon_A = 2k_0$ (A), $5k_0$ (B) and $6.8k_0$ (C).

where (m_{\pm}, s_{\pm}) exist only for $\epsilon \geq \epsilon_c \equiv 3k_0$.

Back to the general situation, numerical analysis of the fixed points and of their stability (by linear stability analysis) demonstrates that the stability diagram is made of two regions: one region with one stable fixed point and another with two stable and one unstable fixed points (Fig. 2 of the main text). In order to determine the boundaries between these two regions, we use canonical analysis techniques of non-linear dynamical systems [1].

For simplicity, we perform the computation for the system of original variables (ρ_A, ρ_I) . By definition, the fixed points are the intersections of the nullclines of Eq.(5-6), defined as the two lines in the (ρ_A, ρ_I) -plan given by $d\rho_A/dt = 0$ and $d\rho_I/dt = 0$. In the monostability region, the two curves intersect at only one point (Fig.S1A), while in the bistability region, we find three points of intersection (Fig.S1B). At the boundaries, they are two points of intersection and at one of them the nullclines are tangent (Fig.S1C). We use these properties to determine parametric expressions for the boundaries. We first derive the expression of the two nullclines. Setting $d\rho_A/dt = 0$ and $d\rho_I/dt = 0$ gives

$$\rho_{A,1} = \frac{-2k_0 + \epsilon_A - (\epsilon_A + \epsilon_I)\rho_I + \sqrt{4\epsilon_A k_0(1 - \rho_I) + (2k_0 - \epsilon_A + (\epsilon_A + \epsilon_I)\rho_I)^2}}{2\epsilon_A} \quad (12)$$

$$\rho_{A,2} = \frac{k_0 - 2k_0\rho_I + \epsilon_I\rho_I(1 - \rho_I)}{k_0 + (\epsilon_A + \epsilon_I)\rho_I} \quad (13)$$

Then, from the condition $\rho_{A,1} = \rho_{A,2}$ (fixed point), we derive an expression for ϵ_I as a function of ϵ_A and ρ_I . Finally, injecting this relation in the equation related to the tangency condition $(\partial\rho_{A,1}/\partial\rho_I = \partial\rho_{A,2}/\partial\rho_I)$, we find parametric expressions for the boundaries (for $0 \leq \rho_I \leq 1$)

$$\begin{aligned} \epsilon_I/k_0 = & \frac{1}{8(\rho_i - 1)^4 \rho_i^2} \left(\sqrt{2} \left[(\rho_i - 1)^4 \rho_i^2 (3\rho_i - 4) \left(\rho_i \left(\rho_i \left(9\rho_i^2 - 30\rho_i + 55 \right) - 60 \right) + 30 \right) \right. \right. \\ & \left. \left. - 10\sqrt{(4 - 3\rho_i) \rho_i^3} + 21\sqrt{(4 - 3\rho_i) \rho_i^5} - 24\sqrt{(4 - 3\rho_i) \rho_i^7} + 9\sqrt{(4 - 3\rho_i) \rho_i^9} - 8 \right] \right]^{1/2} \\ & - \rho_i \left(2\rho_i (3\rho_i - 1) (\rho_i - 1)^3 + \sqrt{2} \left[-(\rho_i - 1)^4 \rho_i \left(\rho_i \left(\rho_i \left(9\rho_i^2 - 30\rho_i + 55 \right) - 60 \right) \right. \right. \right. \right. \\ & \left. \left. \left. + 30 \right) - 10\sqrt{(4 - 3\rho_i) \rho_i^3} + 21\sqrt{(4 - 3\rho_i) \rho_i^5} - 24\sqrt{(4 - 3\rho_i) \rho_i^7} + 9\sqrt{(4 - 3\rho_i) \rho_i^9} - 8 \right] \right]^{1/2} \right) \\ & - 2\sqrt{(4 - 3\rho_i) \rho_i^3} + 12\sqrt{(4 - 3\rho_i) \rho_i^5} - 24\sqrt{(4 - 3\rho_i) \rho_i^7} + 20\sqrt{(4 - 3\rho_i) \rho_i^9} \\ & \left. - 6\sqrt{(4 - 3\rho_i) \rho_i^{11}} \right) \end{aligned} \quad (14)$$

$$\begin{aligned} \epsilon_A/k_0 = & \frac{1}{4(\rho_i - 1)^4 \rho_i} \left(-\rho_i (3\rho_i - 4) (\rho_i - 1)^3 + \sqrt{2} \left[-(\rho_i - 1)^4 \rho_i \left(\rho_i \left(\rho_i (9\rho_i^2 - 30\rho_i + 55) \right. \right. \right. \right. \\ & \left. \left. \left. - 60 \right) + 30 \right) - 10\sqrt{(4 - 3\rho_i) \rho_i^3} + 21\sqrt{(4 - 3\rho_i) \rho_i^5} - 24\sqrt{(4 - 3\rho_i) \rho_i^7} + 9\sqrt{(4 - 3\rho_i) \rho_i^9} - 8 \right]^{1/2} \\ & \left. + 2\sqrt{(4 - 3\rho_i) \rho_i} - 9\sqrt{(4 - 3\rho_i) \rho_i^3} + 15\sqrt{(4 - 3\rho_i) \rho_i^5} - 11\sqrt{(4 - 3\rho_i) \rho_i^7} + 3\sqrt{(4 - 3\rho_i) \rho_i^9} \right) \quad (15) \end{aligned}$$

2 Master-equation and approximations

In this section, we detail the methods and approximations used to derive the probability distribution function (pdf) of m at steady-state and to compute the mean first passage time $\langle \tau \rangle$.

2.1 Derivation of the master-equation for m

From the propensities of the different reactions 1-4, we write the master-equation for the pdf of (m, s)

$$\begin{aligned} \frac{dP(m, s)}{dt} = & r_{u,a}(m - 1/n, s - 1/n)P(m - 1/n, s - 1/n) + r_{u,i}(m + 1/n, s - 1/n)P(m + 1/n, s - 1/n) \\ & + r_{a,u}(m + 1/n, s + 1/n)P(m + 1/n, s + 1/n) + r_{i,u}(m - 1/n, s + 1/n)P(m - 1/n, s + 1/n) \\ & - (r_{u,a} + r_{u,i} + r_{a,u} + r_{i,u})P(m, s) \quad (16) \end{aligned}$$

with

$$r_{u,a}(m, s) = n \left(k_0 + \epsilon_A \frac{s + m}{2} \right) (1 - s) \quad (17)$$

$$r_{u,i}(m, s) = n \left(k_0 + \epsilon_I \frac{s - m}{2} \right) (1 - s) \quad (18)$$

$$r_{a,u}(m, s) = n \left(k_0 + \epsilon_I \frac{s - m}{2} \right) \left(\frac{s + m}{2} \right) \quad (19)$$

$$r_{i,u}(m, s) = n \left(k_0 + \epsilon_A \frac{s + m}{2} \right) \left(\frac{s - m}{2} \right) \quad (20)$$

We make the assumption that s is always closed to 1 and its dynamics is fast compared to the dynamics of m . This means that we neglect the fluctuations of s and that at every time the value of s is given by the solution $s(m)$ of Eq.8=0 for a fixed value of m . For example, in the symmetric regime, we find

$$s(m) = (k_0/\epsilon) \left(\frac{\epsilon/k_0 - 3 + [(\epsilon/k_0 + 3)^2 + 3(m\epsilon/k_0)^2]^{1/2}}{3} \right) \quad (21)$$

Within this limit, the pdf of m is given by $P(m, t) = \sum_s P(m, s, t) \approx P(m, s(m), t)$ and follows the master-equation

$$\frac{dP(m)}{dt} = n [w_+(m - 1/n)P(m - 1/n) + w_-(m + 1/n)P(m + 1/n) - (w_+(m) + w_-(m))P(m)] \quad (22)$$

with

$$w_+(m) \equiv (r_{u,a} + r_{i,u})/n = \left[k_0 + \epsilon_A \left(\frac{s(m) + m}{2} \right) \right] \left[1 - \left(\frac{s(m) + m}{2} \right) \right] \quad (23)$$

$$w_-(m) \equiv (r_{a,u} + r_{u,i})/n = \left[k_0 + \epsilon_I \left(\frac{s(m) - m}{2} \right) \right] \left[1 - \left(\frac{s(m) - m}{2} \right) \right] \quad (24)$$

the propensities to respectively increase (using reactions 1 and 4) or decrease (2 and 3) m by 1. In the limit of a high number of nucleosomes ($n \gg 1$), we derive the Fokker-Planck approximation of Eq.22 by expanding it to second order in $1/n$:

$$\begin{aligned} \frac{\partial P(m)}{\partial t} &\approx n \left[w_+ P - \frac{1}{n} \frac{\partial}{\partial m} (w_+ P) + \frac{1}{2n^2} \frac{\partial^2}{\partial m^2} (w_+ P) + w_- P + \frac{1}{n} \frac{\partial}{\partial m} (w_- P) + \frac{1}{2n^2} \frac{\partial^2}{\partial m^2} (w_- P) \right. \\ &\quad \left. - (w_+ + w_-) P \right] \\ &= - \frac{\partial}{\partial m} \left\{ (w_+ - w_-) P - \frac{1}{2n} \frac{\partial}{\partial m} ([w_+ + w_-] P) \right\} \equiv - \frac{\partial J}{\partial m} \end{aligned} \quad (25)$$

with J defined as the flux of probability.

2.2 Steady-state solution

There is no source of probability, therefore, at steady-state ($\partial P(m)/\partial t = 0$), $J = 0$. This leads to

$$\frac{\partial_m P_\infty}{P_\infty} = 2n \frac{w_+ - w_-}{w_+ + w_-} - \frac{\partial_m (w_+ + w_-)}{w_+ + w_-} \quad (26)$$

$$P_\infty(m) = \frac{1}{Z} \frac{\exp \left[2n \int_{-\infty}^m dm' \left(\frac{w_+ - w_-}{w_+ + w_-} \right) \right]}{w_+(m) + w_-(m)} \quad (27)$$

The local optima of P_∞ are given by solving Eq.26=0. In the limit of large n , this is approximatively equivalent to solve $w_+ - w_- = 0$ which is exactly the steady-state equation of the mass-action model of the first section (Eq.7).

In the symmetric regime, we find analytical expression for the optima

$$m_0^* = m_0 = 0 \quad (28)$$

$$m_\pm^* = \pm (k_0/\epsilon) \sqrt{(\epsilon/k_0 + 1)(\epsilon/k_0 - 3) + \frac{(\epsilon/k_0)^2}{3n^2} \left[2 + 3n - 2n^2 + 2(1+n)\sqrt{1+n+n^2} \right]} \quad (29)$$

$$\approx m_\pm \left(1 + \left(\frac{1}{n} \right) \left(\frac{(\epsilon/k_0)^2}{(\epsilon/k_0 + 1)(\epsilon/k_0 - 3)} \right) + o\left(\frac{1}{n^2} \right) \right) \quad (30)$$

Where m_\pm^* exist only for

$$\epsilon \geq \epsilon_c^* = \frac{3k_0 n}{\sqrt{2n^2 + 3n + 2 + 2(1+n)\sqrt{1+n+n^2}}} \quad (31)$$

$$\approx \epsilon_c \left(1 - \frac{3}{2n} + o\left(\frac{1}{n^2} \right) \right) \quad (32)$$

We remark that the local optima of P_∞ are approximatively given by the fixed point of the dynamical system up to $1/n$ - correction.

In the bistable regime, the widths of the peaks around m_\pm^* are determined by modeling locally the probability by a gaussian distribution $P_\infty(m) \approx \exp[-(m - m_\pm^*)^2/(2\sigma^2)]/z$ centered around m_\pm^* with

$$\sigma = \left[\frac{\partial^2 V}{\partial m^2} \right]_{m_\pm^*}^{-1/2} \quad (33)$$

$$\approx \sqrt{\frac{2k_0}{n(\epsilon - \epsilon_c)}} + o\left(\frac{1}{n^{3/2}} \right) \quad (34)$$

where $V = -\log P_\infty$ is the effective landscape.

2.3 Mean first-passage time

By definition, the mean first-passage time to switch from m_-^* to m_+^* is given by $\langle \tau \rangle = \int_0^\infty dT p(T) \times T$ with $p(T) = -\frac{d}{dT} \left[\int_{-\infty}^{m_+^*} P(m, T) dm \right]$ the probability distribution function of first-passage time and $P(m, T)$ the solution of the Fokker-Planck equation (Eq.25) for the initial condition $P(m, 0) = \delta(m - m_-^*)$. By inverting integration over m and over T and by making integration by parts, it is easy to show that [2]

$$\langle \tau \rangle = \int_{-\infty}^{m_+^*} f(m) dm \quad (35)$$

with $f(m)$ solution of the ordinary differential equation

$$-\frac{d}{dm} \left\{ (w_+ - w_-)f - \frac{1}{2n} \frac{d}{dm} ([w_+ + w_-]f) \right\} = -\delta(m - m_-^*) \quad (36)$$

$$f(m_+^*) = f(-\infty) = 0 \quad (37)$$

Note that the left-hand side of Eq.36 is the same Fokker-Planck operator than in Eq.25. Since P_∞ is the homogeneous solution of Eq.36, we can express $f(m)$ as

$$f(m) = 2nP_\infty(m) \left[\int_{m_-^*}^{m_+^*} \frac{dm'}{[w_+(m') + w_-(m')]P_\infty(m')} - \Theta(m - m_-^*) \int_{m_-^*}^m \frac{dm'}{[w_+(m') + w_-(m')]P_\infty(m')} \right] \quad (38)$$

with $\Theta(x)$ the Heaviside function. Injecting this expression in Eq.35 and inverting integration over m and over m' , leads to

$$\langle \tau \rangle = 2n \int_{m_-^*}^{m_+^*} \frac{dm'}{[w_+(m') + w_-(m')]P_\infty(m')} \int_{-\infty}^{m'} dm \frac{P_\infty(m)}{P_\infty(m')} \quad (39)$$

Remarking that $P_\infty(m)$ is maximal around $m = m_-^*$ and that $1/([w_+(m') + w_-(m')]P_\infty(m'))$ is maximal around $m' = 0$, we approximate these functions by gaussians around their respective maximum. This leads to

$$\langle \tau \rangle = \frac{4\pi n \sigma_0 \sigma_-}{[w_+(0) + w_-(0)]} \exp[V(0) - V(m_-^*)] \quad (40)$$

$$\approx \frac{18\pi}{\sqrt{\epsilon/k_0 + 3}(\epsilon - \epsilon_c)\sqrt{3}} \exp[V(0) - V(m_-^*)] \quad (41)$$

with $\sigma_0 = \left[-\frac{\partial^2 V}{\partial m^2} \right]_{m_0}^{-1/2}$, $\sigma_- = \left[\frac{\partial^2 V}{\partial m^2} \right]_{m_-^*}^{-1/2}$ and

$$\begin{aligned} V(0) - V(m_-^*) &\approx \frac{n}{(\epsilon/k_0)(2 + \epsilon/k_0)} (-6 + \log[8] + \epsilon/k_0(-1 + \epsilon/k_0 + \log[9]) - 2\epsilon/k_0 \log[\epsilon/k_0]) \\ &\quad + 3 \log \left[\frac{3 + \epsilon/k_0}{\epsilon/k_0(\epsilon/k_0 + 1)} \right] + \epsilon/k_0(4 + \epsilon/k_0) \log \left[\frac{(2(3 + \epsilon/k_0))}{3(1 + \epsilon/k_0)} \right] \\ &\quad - \log \left[\frac{9(1 + \epsilon/k_0)}{2\epsilon/k_0(3 + \epsilon/k_0)} \right] + o\left(\frac{1}{n}\right) \end{aligned} \quad (42)$$

References

- [1] S.H. Strogatz (1994) Nonlinear dynamics and chaos. *Westview Press*.
- [2] H. Risken (1989) The Fokker-Planck equation. *Springer-Verlag*.