

Recovery of a recessive allele in a Mendelian diploid model

Loren Coquille

joint work with A. Bovier and R. Neukirch
(University of Bonn)

Young Women in Probability and Analysis 2016, Bonn



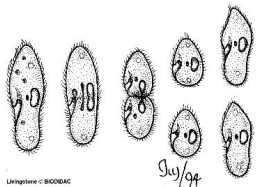
Outline

- 1 Introduction
- 2 Clonal reproduction model
- 3 Mendelian diploid model

- 1 Introduction
- 2 Clonal reproduction model
- 3 Mendelian diploid model

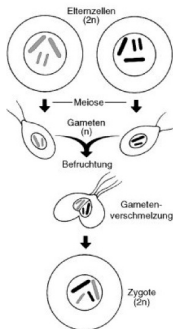
Clonal Reproduction versus Sexual Reproduction

Clonal Reproduction



vs.

Sexual Reproduction



- individual reproduces out of itself
- offspring = copy of the parent
- offspring genom:
 - ▶ 100% of genes of the parent

- individual needs a partner for reproduction
- offspring genom:
 - ▶ 50% genes of mother
 - ▶ 50% genes of father

Outline

- 1 Introduction
- 2 Clonal reproduction model
- 3 Mendelian diploid model

Clonal reproduction model

Bolker, Pacala, Dieckmann, Law, Champagnat, Méléard,...

$\Theta \simeq \mathbb{N}$: Trait space

$n_i(t)$: Number of individuals of trait $i \in \Theta$

Dynamics of the process

$$n(t) = (n_0(t), n_1(t), \dots) \in \mathbb{N}^{\mathbb{N}} \quad :$$

Each individual of trait i

- reproduces clonally with rate $b_i(1 - \mu)$
- reproduces with mutation with rate $b_i\mu$ according to kernel $M(i, j)$
- dies due to age or competition with rate $d_i + \sum_j c_{ij}n_j$

Clonal reproduction model

Bolker, Pacala, Dieckmann, Law, Champagnat, Méléard,...

$\Theta \simeq \mathbb{N}$: Trait space

$n_i(t)$: Number of individuals of trait $i \in \Theta$

Dynamics of the process

$$n(t) = (n_0(t), n_1(t), \dots) \in \mathbb{N}^{\mathbb{N}} \quad :$$

The population n_i

- increases by 1 with rate $b_i(1 - \mu)n_i$
- makes n_j increase by 1 with rate $b_i\mu \cdot M(i, j) \cdot n_i$
- decreases by 1 with rate $(d_i + \sum_j c_{ij}n_j)n_i$

Clonal reproduction model

Bolker, Pacala, Dieckmann, Law, Champagnat, Méléard,...

$\Theta \simeq \mathbb{N}$: Trait space

$n_i(t)$: Number of individuals of trait $i \in \Theta$

Dynamics of the **rescaled** process (with competition c_{ij}/K)

$$n^K(t) = \frac{1}{K}(n_0(t), n_1(t), \dots) \in (\mathbb{N}/K)^{\mathbb{N}} \quad :$$

The population n_i^K

- increases by $1/K$ with rate $b_i(1 - \mu) \cdot Kn_i^K$
- makes n_j^K increase by $1/K$ with rate $b_i\mu M(i, j) \cdot Kn_i^K$
- decreases by $1/K$ with rate $(d_i + \sum_j c_{ij}n_j^K) \cdot Kn_i^K$

Proposition

Let $\Theta = \{1, 2\}$, assume that the initial condition $(n_1^K(0), n_2^K(0))$ converges to a deterministic vector (x_0, y_0) for $K \rightarrow \infty$. Then the process $(n_1^K(t), n_2^K(t))$ converges in law, on bounded time intervals, to the solution of

$$\frac{d}{dt} \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = X(x(t), y(t)) = \begin{pmatrix} (b_1 - d_1 - c_{11}x - c_{12}y)x \\ (b_2 - d_2 - c_{21}x - c_{22}y)y \end{pmatrix}.$$

with initial condition $(x(0), y(0)) = (x_0, y_0)$.

Monomorphic equilibria : $\bar{n}_i = \frac{b_i - d_i}{c_{ii}}$

Invasion fitness : $f_{ij} = b_i - d_i - c_{ij}\bar{n}_j$.

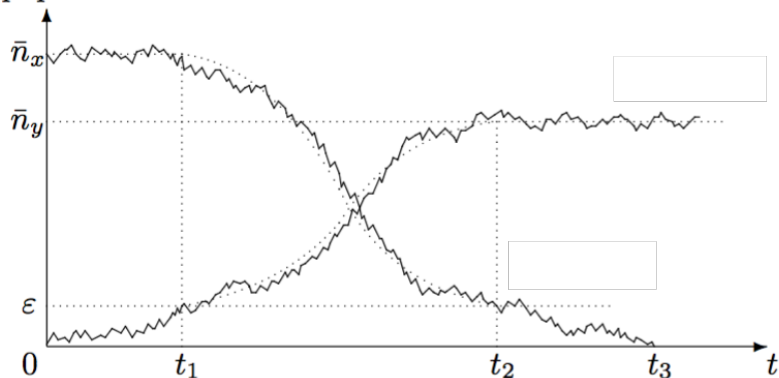
Large population and ONE mutation

Champagnat, 2006

Start with $n_i^K(0) = \bar{n}_x 1_{i=x} + \frac{1}{K} 1_{i=y}$ (time of the first mutation).

If y is fitter than x , i.e. $f_{yx} > 0$ and $f_{xy} < 0$, then with proba $\rightarrow 1$:

population size



$O(\log K)$

Supercritical BP

$O(1)$

LLN

$O(\log K)$

Subcritical BP

Proposition (Trait Substitution Sequence)

Assume

- $\log K \ll \frac{1}{K\mu}$
- $\forall (i, j) \in \Theta^2$, either $(f_{ji} < 0)$ or $(f_{ji} > 0$ and $f_{ij} < 0)$.

Then the rescaled process

$$(n_{t/K\mu}^K)_{t \geq 0} \Rightarrow (\bar{n}_{X_t} 1_{X_t})_{t \geq 0}$$

where the TSS X_t is a Markov chain on Θ with transition kernel

$$P(i, j) = b_i \frac{[f_{ji}]_+}{b_j} M(i, j).$$

Jumps towards higher fitness

Outline

- 1 Introduction
- 2 Clonal reproduction model
- 3 Mendelian diploid model**

Let \mathcal{U} be the allelic trait space, a countable set.

For example $\mathcal{U} = \{a, A\}$.

- An individual i is determined by two alleles out of \mathcal{U} :

- ▶ **genotype:** $(u_1^i, u_2^i) \in \mathcal{U}^2$

- ▶ **phenotype:** $\phi((u_1^i, u_2^i))$, with $\phi : \mathcal{U}^2 \rightarrow \mathbb{R}_+$

- Rescaled population:

let N_t be the total number of individuals at time t ,

$$n_{u_1, u_2}^K(t) = \frac{1}{K} \sum_{i=1}^{N_t} 1_{(u_1^i, u_2^i)}(t)$$

Reproduction rate of (u_1^i, u_2^i) with (u_1^j, u_2^j) :

$$\frac{f_{u_1^i u_2^i} f_{u_1^j u_2^j}}{\text{Number of potential partners of } (u_1^i, u_2^i) \times \text{their mean fertility}}$$

which means, for one individual (u_1^i, u_2^i) :

- $Exp(f_{u_1^i u_2^i})$ -Clock rings,
- choose a partner at random,
- reproduce (proba according to fertility of the partner)

Birth and Death Rate

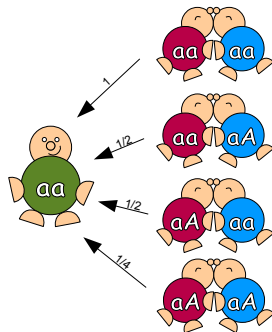
For $\mathcal{U} = \{a, A\}$

- The populations increase by 1 with rate

$$b_{aa} = \frac{(f_{aa}n_{aa} + \frac{1}{2}f_{aA}n_{aA})^2}{f_{aa}n_{aa} + f_{aA}n_{aA} + f_{AA}n_{AA}}$$

$$b_{aA} = 2 \frac{(f_{aa}n_{aa} + \frac{1}{2}f_{aA}n_{aA})(f_{AA}n_{AA} + \frac{1}{2}f_{aA}n_{aA})}{f_{aa}n_{aa} + f_{aA}n_{aA} + f_{AA}n_{AA}}$$

$$b_{AA} = \frac{(f_{AA}n_{AA} + \frac{1}{2}f_{aA}n_{aA})^2}{f_{aa}n_{aa} + f_{aA}n_{aA} + f_{AA}n_{AA}}$$



Birth and Death Rate

For $\mathcal{U} = \{a, A\}$

- The populations increase by 1 with rate

$$b_{aa} = \frac{(f_{aa}n_{aa} + \frac{1}{2}f_{aA}n_{aA})^2}{f_{aa}n_{aa} + f_{aA}n_{aA} + f_{AA}n_{AA}}$$

$$b_{aA} = 2 \frac{(f_{aa}n_{aa} + \frac{1}{2}f_{aA}n_{aA})(f_{AA}n_{AA} + \frac{1}{2}f_{aA}n_{aA})}{f_{aa}n_{aa} + f_{aA}n_{aA} + f_{AA}n_{AA}}$$

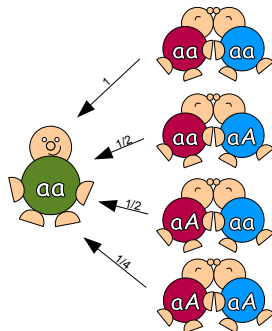
$$b_{AA} = \frac{(f_{AA}n_{AA} + \frac{1}{2}f_{aA}n_{aA})^2}{f_{aa}n_{aa} + f_{aA}n_{aA} + f_{AA}n_{AA}}$$

- The populations decrease by 1 with rate

$$d_{aa} = (D_{aa} + C_{aa,aa}n_{aa} + C_{aa,aA}n_{aA} + C_{aa,AA}n_{AA})n_{aa}$$

$$d_{aA} = (D_{aA} + C_{aA,aa}n_{aa} + C_{aA,aA}n_{aA} + C_{aA,AA}n_{AA})n_{aA}$$

$$d_{AA} = (D_{AA} + C_{AA,AA}n_{AA} + C_{AA,aA}n_{aA} + C_{AA,aa}n_{aa})n_{AA}$$



Proposition

Assume that the initial condition $(n_{aa}^K(0), n_{aA}^K(0), n_{AA}^K(0))$ converges to a deterministic vector (x_0, y_0, z_0) for $K \rightarrow \infty$. Then the process $(n_{aa}^K(t), n_{aA}^K(t), n_{AA}^K(t))$ converges in law to the solution of

$$\frac{d}{dt} \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} = X(x(t), y(t), z(t)) = \begin{pmatrix} b_{aa}(x, y, z) - d_{aa}(x, y, z) \\ b_{aA}(x, y, z) - d_{aA}(x, y, z) \\ b_{AA}(x, y, z) - d_{AA}(x, y, z) \end{pmatrix}.$$

- $b_{AA}(x, y, z) = \frac{(f_{AA}z + \frac{1}{2}f_{aA}y)^2}{f_{aa}x + f_{aA}y + f_{AA}z}$
- $d_{AA}(x, y, z) = (D_{AA} + C_{AA,aa}x + C_{AA,aA}y + C_{AA,AA}z)z$
- and similar expressions for the other terms

ONE mutation : 3-System (aa, aA, AA)

Phenotypic viewpoint

- allele A **dominant**
- allele a **recessive**

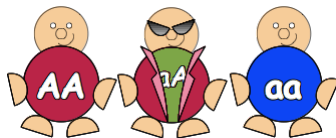


The dominant allele A defines the phenotype $\Rightarrow \phi(aA) = \phi(AA)$

ONE mutation : 3-System (aa, aA, AA)

Phenotypic viewpoint

- allele A **dominant**
- allele a **recessive**



The dominant allele A defines the phenotype $\Rightarrow \phi(aA) = \phi(AA)$

- $c_{u_1 u_2, v_1 v_2} \equiv c, \quad \forall u_1 u_2, v_1 v_2 \in \{aa, aA, AA\}$
- $f_{AA} = f_{aA} = f_{aa} \equiv f$
- $D_{AA} = D_{aA} \equiv D$ **but** $D_{aa} = D + \Delta$

ONE mutation : 3-System (aa, aA, AA)

Phenotypic viewpoint

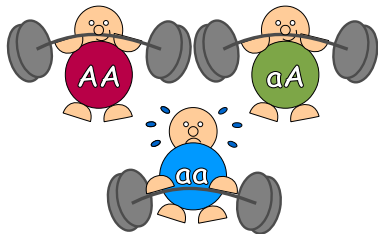
- allele A **dominant**
- allele a **recessive**



The dominant allele A defines the phenotype $\Rightarrow \phi(aA) = \phi(AA)$

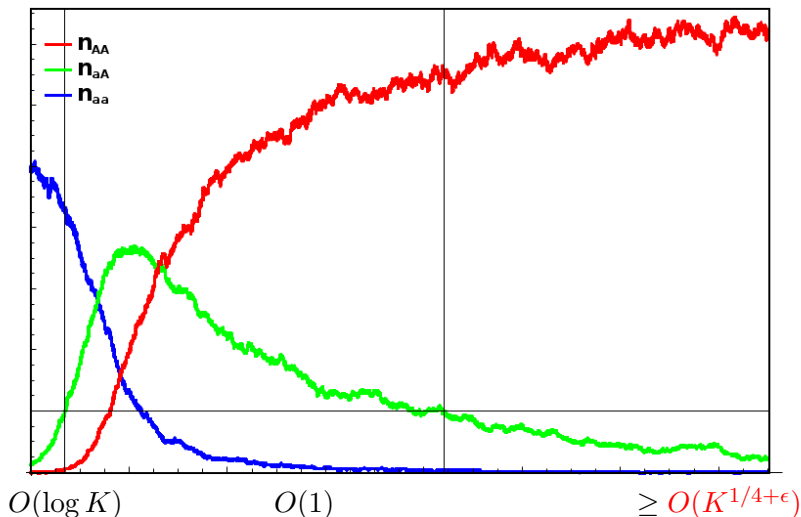
- $c_{u_1 u_2, v_1 v_2} \equiv c, \quad \forall u_1 u_2, v_1 v_2 \in \{aa, aA, AA\}$
- $f_{AA} = f_{aA} = f_{aa} \equiv f$
- $D_{AA} = D_{aA} \equiv D$ **but** $D_{aa} = D + \Delta$

\Rightarrow **type aA is as fit as AA and both are fitter than type aa**



Work of Bovier, Neukirch (2015)

Start with $n_i^K(t) = \bar{n}_{aa}1_{i=aa} + \frac{1}{K}1_{i=aA}$ then :



The system converges to $(0, 0, \bar{n}_{AA})$ as $t \rightarrow \infty$ but **slowly** !

Suppose a new dominant mutant allele B
appears before aA dies out.

Suppose that phenotypes a and B cannot reproduce.

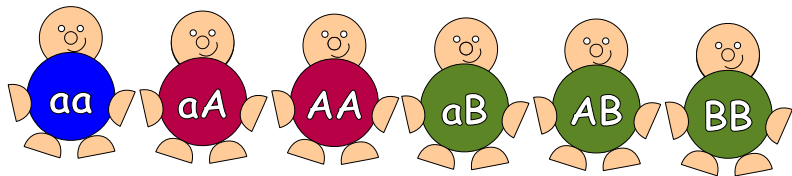
Can the aa -population recover and coexist with the
mutant population ?

Model with a second mutant

Mutation to allele $B \rightarrow \mathcal{U} = \{a, A, B\}$

Model with a second mutant

Mutation to allele $B \rightarrow \mathcal{U} = \{a, A, B\}$

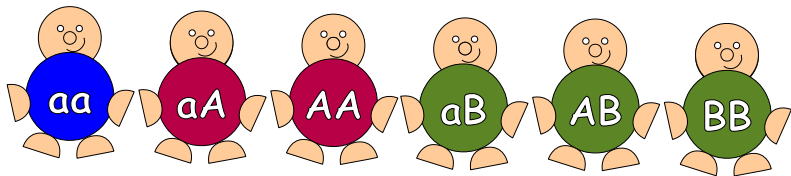


\Rightarrow 6 possible genotypes: aa, aA, AA, aB, AB, BB

\Rightarrow Dominance of alleles $a < A < B$

Model with a second mutant

Mutation to allele $B \rightarrow \mathcal{U} = \{a, A, B\}$



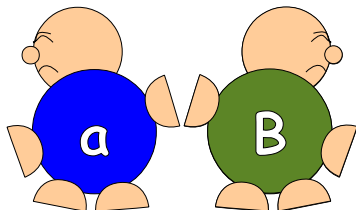
\Rightarrow 6 possible genotypes: aa, aA, AA, aB, AB, BB

\Rightarrow Dominance of alleles $a < A < B$

Differences in Fitness:

- fertility: $f_a = f_A = f_B = f$
- natural death: $D_a = D + \Delta > D_A = D > D_B = D - \Delta$

No recombination between a and B

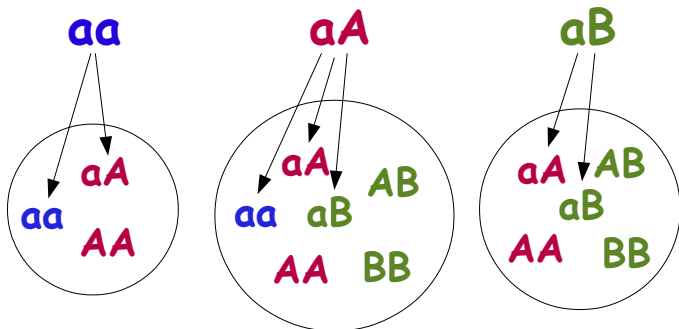


Birth Rates

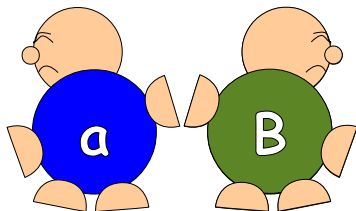
birth-rate of *aa*-individual:

$$b_{aa} = \frac{n_{aa} (n_{aa} + \frac{1}{2}n_{aA})}{\text{Pool}(aa)} + \frac{\frac{1}{2}n_{aB} (\frac{1}{2}n_{aA} + \frac{1}{2}n_{aB})}{\text{Pool}(aB)} + \frac{\frac{1}{2}n_{aA} (n_{aa} + \frac{1}{2}n_{aA} + \frac{1}{2}n_{aB})}{\text{Pool}(aA)}$$

Pools of potential partners:



No competition between a and B



Competition

	aa	aA	AA	aB	AB	BB
aa	<i>c</i>	<i>c</i>	<i>c</i>	0	0	0
aA	<i>c</i>	<i>c</i>	<i>c</i>	<i>c</i>	<i>c</i>	<i>c</i>
AA	<i>c</i>	<i>c</i>	<i>c</i>	<i>c</i>	<i>c</i>	<i>c</i>
aB	0	<i>c</i>	<i>c</i>	<i>c</i>	<i>c</i>	<i>c</i>
AB	0	<i>c</i>	<i>c</i>	<i>c</i>	<i>c</i>	<i>c</i>
BB	0	<i>c</i>	<i>c</i>	<i>c</i>	<i>c</i>	<i>c</i>

Theorem (Bovier, Neukirch, C. 2016/10/07 (yesterday!))

The deterministic system started with initial condition

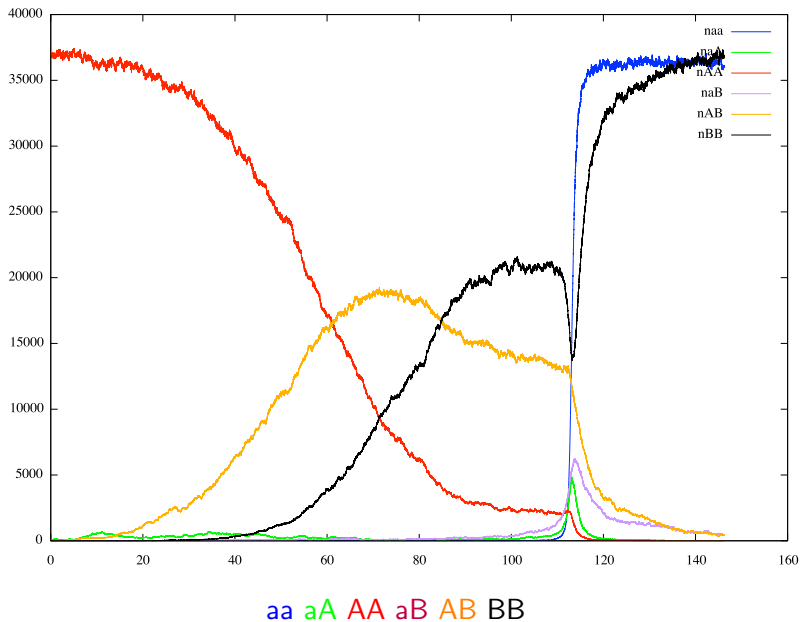
$$n_i(0) = \bar{n}_{AA}1_{i=AA} + \epsilon 1_{i=aA} + \epsilon^2 1_{i=aa} + \epsilon^3 1_{i=AB}$$

and under the following assumptions on the parameters

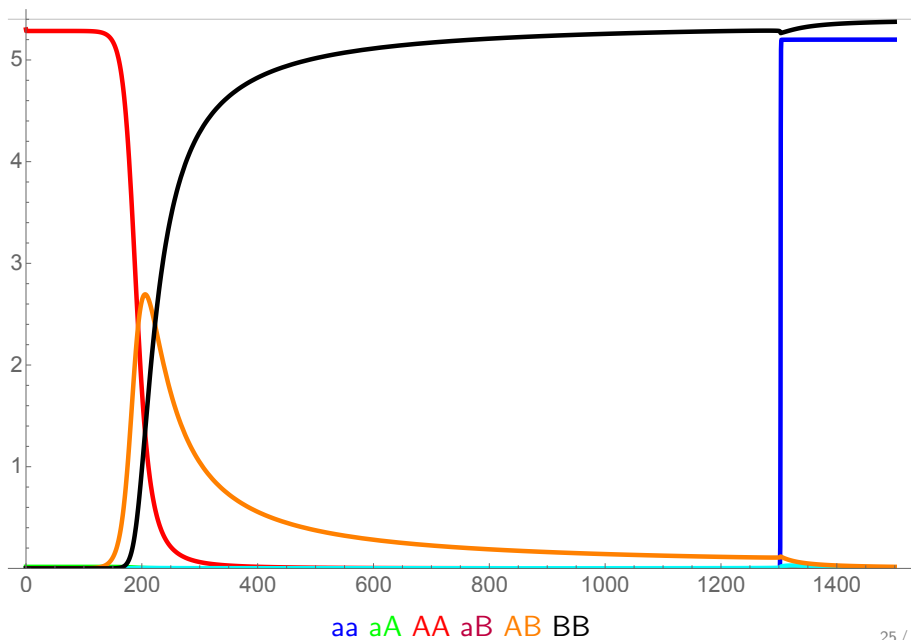
- 1 Δ sufficiently small,
- 2 f sufficiently large,

converges polynomially fast to the fixed point $(\bar{n}_a, 0, 0, 0, 0, \bar{n}_B)$.

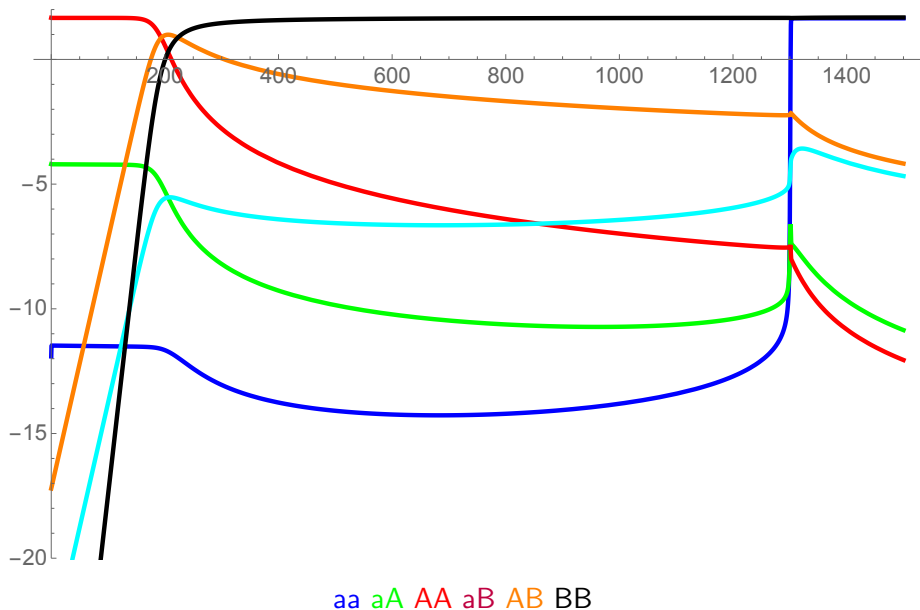
Simulation



Large population : deterministic system

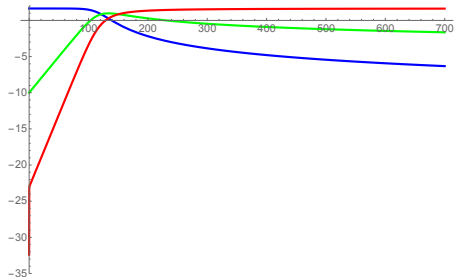
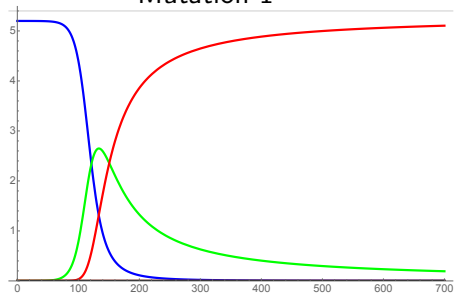


Large population : deterministic system

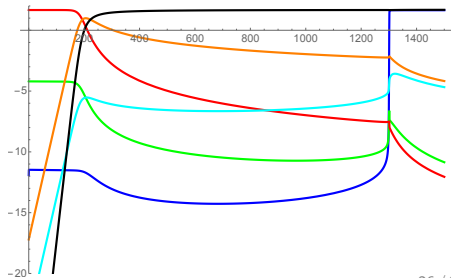
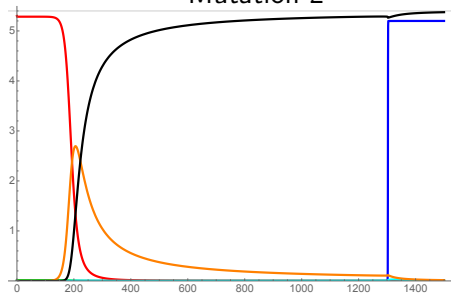


Recap - Dimorphism in two mutations

Mutation 1



Mutation 2



- Phase 1 : Exponential growth of AB at rate Δ inside the 3-system (aa, aA, AA)
- Phase 2 : Effective new 3-system (AA, AB, BB) until $AA \approx aA$
- Phase 3 : Exponential growth of aa at rate $f - D - \Delta \gg \Delta$
- Phase 4 : Convergence to $(\bar{n}_a, 0, 0, 0, 0, \bar{n}_B)$

