Emergence of Mutualism

Application of a Differential Model to Endosymbiosis.

Gérard Weisbuch and Guillemette Duchateau

Ecole Normale Supérieure
Laboratoire de Physique statistique
(Associé au CNRS et aux Universités Paris VI et Paris VII)
24 rue Lhomond
75005 Paris
Fax: 33 1 45 87 34 89

Email: weisbuch@physique.ens.fr duchat@physique.ens.fr

Abstract

A population dynamics approach based on a system of differential equations allows us to establish conditions for the emergence of mutualism for cases such as coelenterates-algae symbionts. A central assumption of the model is that a host organism is able to discriminate, via some molecular recognition mechanisms, among different invading organisms and preferentially reject *parasites* rather than *bona fide symbionts*. Large differential rejection rates allow the emergence of mutualism. Different attractors of the population dynamics correspond to the emergence of mutualism, predominance of "selfish" species, or coexistence of many species.

1 Introduction

1.1 The paradox of mutualism

Mutualistic systems are known to occur in nature, e.g., lichen made of algae and fungus, corals-zooxanthellae and *Hydra-Chlorella*. Exchanges among the partners are beneficial for the species involved, which live together in close association. The benefits exchanged can be food, energy, protection-habitat and transport (pollination) {Boucher, James and Keeler, 1982}. According to {Begon, Harper and Trunsend, 1986} most of the world's biomass is composed of mutualists: organisms in forests, meadows and corals are involved in symbiotic associations.

The emergence and stability of mutualism constitutes a paradox in terms of individual selection of the fittest. The paradox is the following: since giving food to the other symbiont should be costly for the donor, we expect the donor to be disadvantaged in terms of fitness with respect to a more selfish species which would give nothing.

The purpose of this paper is to describe a mathematical model which accepts the premises of this assertion but refutes its conclusion. An important part of the argument is that selection occurs not only at the level of individual organisms, but also at the level of their mutualistic associations.

1.2 Previous theoretical studies

The stability of mutualistic ecosystems has been studied by a number of authors using cost-benefits analysis, standard non-linear differential systems or "Artificial Life" numerical simulations.

An early work is that of {Roughgarden, 1975} who did a cost benefit analysis of the exchanges among hosts and guests to describe damselfish-anemones mutualism. {Wilson, 1983} in his study of beetles-phoretic mite association insists on the idea of group rather than individual selection. He predicted that when populations are clustered in groups with a varying proportion of mutualists and selfish types, those group with more mutualists should do better and be selected.

The differential system approach is sumarized in {De Angelis, Post and Travis, 1986}. It is based on either Volterra-Lotka formalism to characterize global stability or a more general form of the *per capita* rate of increase function to monitor local stability. Classical non-linear analysis criteria are used to obtain inequalities among equation coefficients that ensure stability. The absence of stability among a set of host, parasites and true symbionts that exist as individual species leads these authors to suggest that mutualism can only exist when the relationship between the host and the guest involves a one-to-one relationship.

Artificial life simulations of {Ikegami and Kanebo, 1990} do not succeed in establishing mutualism permanently since they postulate the existence of parasitic mutants quite harmful to the host. Their model exhibit transient periods during which mutualistic species predominate, alternating with periods when parasitic species exploit them. Artificial life models have also been used to study stability and breakdown of symbiotic associations {Taylor, Muscatine and Jefferson, 1989}.

A conclusion of many authors is that long term interactions among host and guest should be necessary for the establishment of mutualism. These conclusions also appear in studies of equivalent problems, such as the increase of the stability of hypercycles by compartmentation {Szathmàry and Demeter, 1987} and in the iterated prisonner dilemma model {Lindgren, 1992; Nowak and May, 1992} where cooperation among players can only be established when players play long enough.

The present paper works along these lines. We propose here a differential equation model that takes into account a dynamics of association-dissociation of organisms which is representative of the processes existing among coelenterates-algae symbiosis and we study the range of parameter values for which this aggregation dynamics gives rise to mutualism. The next section presents the mathematical framework and the elementary processes involving individual organisms. Section 3 is a summary of simulation results obtained for a simple model where only binary associations are possible. Section 4 presents the slow manifold analysis of the model, which brings us some insight into the transitions among the different observed dynamical

behaviors. In section 5, we show that the asymptotic behavior of associations with many endosymbionts is equivalent to that of binary association, since the slow manifolds of both models are equivalent.

Section 6 extends the model to those cases when one of the benefits enjoyed by endosymbionts is protection by the host, and when the host eventually digest parasites instead of rejecting them. The conclusions of this study are also discussed in this last section.

1.3 The hyperbolic model

We use a model of population dynamics introduced in {Weisbuch, 1984}. With respect to the Volterra-Lotka approach, this model does not exhibit divergences of populations in the case of positive interactions among organisms and is analytically soluble in the low mutation rate limit.

The following differential system allows us to study population dynamics in the presence of mutations. Each population varies in time due to three terms:

- The first term describes population growth according to a fitness coefficient α and available resources f which are shared by all populations;

- The second term is simply the rate of death.

- The third term is a mutation term which decreases the population because of all possible point mutations or increases it because of mutations from other existing species j one mutation away.

$$\frac{dP_i}{dt} = \frac{\alpha_i.f.P_i}{P_t} - d.P_i + m.(-n.P_i + \sum_j P_j) \;\; ; \;\; \text{every i}$$

$$\text{and} \qquad (1.1)$$

$$P_t = \sum_i P_i$$

where P_i and P_j are the populations of i and j, P_t is the total population, d and m are the death

and mutation rates and n is the number of genes of each organism. A first time scale is given by 1/d, the other one much longer is the time scale related to evolution 1/m. The scale of the populations is proportional to the available resources.

Starting from low population levels, those populations with sufficient fitness grow at an exponential rate. But a total poulation level is then reached when only the fittest populations have a positive growth rate, namely those which fitness is in excess of $\frac{d.P_t}{f}$. The attractor of the system is analytically obtained by a perturbation technique {Morse and Feshback, 1953} when the mutation term is small with respect to the death term. First of all, only the fittest organisms predominate with a population ratio of $\frac{1}{m}$ with respect to other organisms. The fittest (or *dominant*) population P_m is obtained by equating its derivative to 0 and neglecting mutations from other species:

$$P_t = \frac{\alpha_m.f}{d} = P_m \tag{1.2}$$

where α_m is the fitness coefficient of the fittest population (equating P_m to P_t is a guess which is supported by further analysis of higher order terms in m).

The populations of the nearest mutants are obtained by equating to zero their time derivative and taking into account the mutation term which comes from the fittest species:

$$\left(\frac{\alpha_i.f}{P_t} - d\right) . P_i + m . P_m = 0$$

$$P_i = \frac{m . P_m}{d \left(1 - \frac{\alpha_i}{\alpha_m} \right)} \tag{1.3}$$

In the limit of small m, the ratio of the first mutants to the fittest population scales as $\frac{m}{d}$ The same analysis can be carried to next mutants which decay in population by a factor $\frac{m}{d}$ for each further mutation from the fittest.

The perturbation technique described here will allow us to interpret the results obtained later. The main point is the predominance of the fittest populations by a factor $\frac{m}{d}$.

2 Building-up of the model

2.1 Phylogeny and interactions

We shall now apply a modified version of the above formalism to a system made of five populations, whose interactions and phylogeny are represented on Figure 1. This model differs from the cannonical model of section 1.3 by the presence of two food webs with resources f1 and f2 corresponding repectively to hosts and endosymbionts.

-Populations C and B are unrelated and are not involved in any interaction with other populations. They can be considered as primeval organisms. Their respective fitnesses are γ and β.

- Population A differs from population B by one mutation. Population A can be considered as a *host* for populations D and E. It produces some nutrients that can be used by D or E.

- D is a guest of A and is a *selfish* organism, it is further called an *egoist* . It uses nutrients produced by A, but does not give anything in exchange to A.

- E is a *bona fide symbiont* of A. It uses nutrients produced by A and in exchange provides A with nutrients.

Organisms B and A feed on resource 1 available in quantity f1 and organisms C, D and E on resource 2 available in quantity f2. From the hyperbolic model described in section 1.3, we *a priori* expect to find two dominant species, one for each resource tree.

2.2 The elementary processes

The ecosystem is made of free organisms A, B, C, D and E and couples AE and AD, when A is infested respectively by E or D. To simplify the model, we have first supposed that the host A offers only one site where E or D are able to bind. This simplification limits the set of possible associations to couples AD and AE. The simplified model is to be further called the single site model, SSM. Changes in population sizes are due to elementary processes represented in Figure 2 and listed further.

2.2.a Reproduction

Reproduction of free A, B, C, D and E obey simple rules such as:

$A \rightarrow A + A$ fitness α_F

$B \rightarrow B + B$ fitness β

$C \rightarrow C + C$ fitness γ

$D \rightarrow D + D$ fitness δ_F

$E \rightarrow E + E$ fitness ε_F

The frequency of these processes is described as in equation 1 and involves fitness coefficients which are specific of the concerned species.

In the case of reproduction of one of the symbionts of AD or AE, the other one is liberated. For each one of AE or AD two processes can then occur:

$AE \rightarrow A + AE$ reproduction of A with fitness α

$AE \rightarrow E + AE$ reproduction of E with fitness ε

$AD \rightarrow A + AD$ reproduction of A with fitness α_F

$AD \rightarrow D + AD$ reproduction of D with fitness δ

The frequency of these processes is described as in equation 1 but it involves fitness

coefficients which are different and eventually higher than those of the free species.

The following choice of fitness parameters reflects the fundamental paradox about mutualism. The relations (2.1, 2.2) among fitnesses are:

$$\alpha > \beta > \alpha_F \tag{2.1}$$

The fitness of B (β) is higher than the fitness of A (α_F) because A provides D and E with nutrients. But when E is bound to A, the fitness of A (α) is larger than that of B thanks to the cooperation of E.

$$\varepsilon_F < \delta_F = \gamma < \varepsilon \leq \delta \tag{2.2}$$

The fitness of the E species (ε or ε_F) is inferior to the fitness of D (δ or δ_F) because E produces with some cost a nutrient for A. $\delta_F = \gamma$ since D does not provide nutrient for anyone. Fitnesses of D (δ) and E (ε) when associated with A are larger than fitnesses (δ_F and ε_F) of the free organisms.

2.2.b Death

Death processes for free species A, B, C, D and E occur with frequency d. But in the case of the death of one of the symbionts of AD or AE, the other one is liberated, with frequency d:

$$B \rightarrow .$$

$$AE \rightarrow A$$

$$AE \rightarrow E$$

$$AD \rightarrow A$$

$$AD \rightarrow D$$

2.2.c Mutations

Mutations involving the free species A, B, C, D and E occur with frequency m.n, where n is the number of genes and m the mutation frequency per gene. Only some of these mutations give viable organisms according to the nodes of the phylogenic tree (Figure 1). For AD and AE the mutation processes are:

$$AE \underset{->}{\overset{<-}{}} AD \rightarrow C + A \qquad \text{frequency m}$$

when E or D are mutated. When A is mutated, they are:

$$AE \rightarrow B + E \qquad \text{frequency m}$$

$$AD \rightarrow B + D \qquad \text{frequency m}$$

2.2.d Association and dissociation

Association and dissociation processes involve the symbionts:

$$AE \underset{ks_e}{\overset{kl_e}{<->}} A + E \qquad\qquad AD \underset{ks_d}{\overset{kl_d}{<->}} A + D$$

where kl_e and kl_d are the association (l for linking) rates and ks_e and ks_d are the dissociation (s for separation) rates. Choosing $ks_d > ks_e$ allows E to spend more time inside A than D. This results in a larger time average fitness of E. The basic hypothesis is that the host normally rejects "guests" at a certain rate, but it can somehow appreciate the degree of cooperation of E and reject it less frequently than D. This selective rejection of D might be due to some molecular recognition mechanism, which are known to exist in polypes and sponges, or simply because the increase in the level of nutrient produced by E decreases the rejection rate of E by A.

The only "advantage" of E with respect to D in this model is : $ks_e < ks_d$. We then want to check when, i.e. for what set of parameters, this condition is sufficient to overcome the advantage in fitness of D over E and bring the emergence of mutualism.

3 The Single site model (SSM)

3.1 The differential system

The set of differential equations describing the corresponding population dynamics is:

$$\frac{dA}{dt} = \frac{(\alpha_F.(A+AD) + \alpha.AE).f1}{A+AD+AE+B} - (d+ m.n).A+ d_d.AD + d_e.AE + m.(B+AD)$$
$$+ (ks_e.AE) + (ks_d.AD) - kl_e.(A.E) - kl_d.(A.D) \quad (3.1)$$

$$\frac{dAD}{dt} = - 2.(d + m.n).AD+ (m.AE) - (ks_d.AD) + (kl_d.(A.D)) \quad (3.2)$$

$$\frac{dAE}{dt} = - 2.(d + m.n).AE + (m.AD) - (ks_e.AE) + (kl_e.(A.E)) \quad (3.3)$$

$$\frac{dB}{dt} = \frac{\beta.B.f1}{A+AD+AE+B} - (d+ m.n).B + m.(A+AE+AD) \quad (3.4)$$

$$\frac{dC}{dt} = \frac{\gamma.C.f2}{C+D+E+AD+AE} - (d+ m.n).C + m.(D+AD) \quad (3.5)$$

$$\frac{dD}{dt} = \frac{(\delta_F.D + \delta.AD).f2}{C+D+E+AD+AE} - (d+ m.n).D+ (d.AD) + m.(C+E+AD) + (ks_d.AD) - kl_d.(A.D) \quad (3.6)$$

$$\frac{dE}{dt} = \frac{(\varepsilon_F.E + \varepsilon.AE).f2}{C+D+E+AD+AE} - (d+ m.n).E + d.AE + m.(D+AE) + ks_e.AE - kl_e.(A.E) \quad (3.7)$$

These equations simply sum the contributions of the processes listed in section 2.2 and set the time variation of each population.

3.2 Simulation results

The initial conditions are such that populations are at equilibrium for the populations B, C, A, D and AD, mutant E being absent; for instance:

A=0.01 AE=0 AD=0 B=80000 C=20000 D=60000

At time t=0, E is introduced in the system, with population

E=0.01

For each set of parameters, the initial conditions, determined by numerical simulations of the differential system in the absence of E, are those corresponding to the eventual emergence of mutualism, when a true symbiont E is introduced among an ecosystem at equilibrium containing only primeval species B,C and "egoists" D plus the host species A.

The set of parameters listed in the Table 1 were used in the simulations, unless otherwise specified.

Varying parameters and initial conditions three types of attractors are obtained. Figures 3, 4 and 5 display typical time evolutions of the populations obtained by numerical integration of system (equations (3.1) to (3.7)). All attractors are point attractors. The most sensitive parameters, in terms of dynamical behavior, are the difference in fitness ($\delta - \varepsilon$) and in the rejection rate ($ks_d - ks_e$) between species D and E. The resource parameters f1 and f2 simply change the scale of populations. Three time scales, fast, intermediate and slow, are fixed by respectively ks_d, d and m. d/m determines the ratio between dominant and less fit species populations.

For small ks_d, i.e., when D spends in A an amount of time comparable to E, AE does not prevail over AD. The "egoist" D benefits from the nutrients it gets from A, and develops

faster than E. A does not benefit from D and is not able to grow faster than B. Finally the primeval population B remains at the higher level. A small population of host A is mainly infested by "egoists" D, which gives a small but sufficient advantage to D to overcome C (Figure 3).

In Figure 4, emergence of mutualism is observed for large ks_d, i.e. when D is rapidly expelled from A. A gets support from E and is able to overcome B. The symbiotic organism AE becomes predominant over the primeval populations C and B. The primeval organisms populations are only maintained by the mutations.

For some values of the fitness parameters a coexistence region is observed for intermediate values of ks_d (Figure 5). In this region, one observes coexistence of the couples AD and AE. The fitness of A in these couples is comparable to that of B, which coexists with them. A high population of free D is also maintained.

The three regimes can be observed on Figure 6 which shows the equilibrium levels of the populations when ks_d increases. The transition value of ks_d between the three dynamics depends upon the difference in fitness between D and E. The "phase diagram" in the plane ($\delta - \varepsilon$, ks_d), Figure 7, shows the different dynamical regimes. These diagrams are complicated because the regime that is reached depends not only from the parameters, but also from the initial conditions. The discontinuities in populations are due to the non-linearity of the equilibrium equations, obtained by setting to zero the time derivatives. For the same set of parameters several solutions exist , one or two of which are attractors. Dependence from the initial conditions and hysteresis are then observed.

One consequence of this hysteresis is that the conditions for the stability of mutualism when facing the invasion of a new "egoist" is less stringent than the condition for the emergence of mutualism facing the same "egoist" already established. The emergence of mutualism implies a transition from a different attractor and is achieved for larger ks_d (or a smaller $\delta - \varepsilon$) than its failure against an invading "egoist" which implies a dynamics starting from the mutualism attractor. In the second case the transition is achieved for a lower ks_d (or a larger $\delta - \varepsilon$).

4 Slow manifold analysis

The slow manifold analysis allows us to interprete the simulation results and to predict the transitions among the different dynamical regimes. The interaction processes among the different populations have very different time scales: the exchange interaction between free and linked species are very fast with respect to population growth and death terms which are of the same order. Mutations are even slower. One can then suppose that the populations of free and linked organisms follow their ratio at exchange equilibrium given by:

$$(ks_d+2.d).AD = kl_d.A.D \tag{4.1}$$

$$(ks_e+2.d).AE = kl_e.A.E \tag{4.2}$$

After a fast decay towards exchange equilibria, the populations are describing a slow manifold whose equations are obtained by equating to zero combinations of the time derivatives in which the association and disssociation terms are cancelled (fast exchange). Mutation terms are neglected since they are small with respect to proliferation and death rates:

$$\frac{dA}{dt} + \frac{dAD}{dt} + \frac{dAE}{dt} = \frac{(\alpha_F.(A+AD) + \alpha.AE).f1}{A+AD+AE+B} - d.(A+AD+AE) = 0 \tag{4.3}$$

$$\frac{dD}{dt} + \frac{dAD}{dt} = \frac{(\delta_F.D + \delta.AD).f2}{C+D+E+AD+AE} - d.(D+AD) = 0 \tag{4.4}$$

$$\frac{dE}{dt} + \frac{dAE}{dt} = \frac{(\varepsilon_F.E + \varepsilon.AE).f2}{C+D+E+AD+AE} - d.(E+AE) = 0 \tag{4.5}$$

$$\frac{dB}{dt} = \frac{\beta.B.f1}{A+AD+AE+B} - d.B = 0 \tag{4.6}$$

$$\frac{dC}{dt} = \frac{\gamma.C.f2}{C+D+E+AD+AE} - d.C = 0 \qquad (4.7)$$

The set of elementary processes that give rise to these equations is reduced to the bold line processes of Figure 2.

The "species" and their effective fitness

The first three equations represent the time derivatives of populations of "species" A, D and E. Here the term "species", by opposition to organisms, refers to a set of organisms, wether free or bound. The total population of species A for instance is A+AD+AE. The reproduction of these species involves an " effective fitness" which is:

$\frac{(\alpha_F.(A+AD) + \alpha.AE)}{A+AD+AE}$ for species A composed of populations A+AD + AE

$\frac{\delta_F.D + \delta.AD}{D+AD}$ for species D composed of populations D+ AD

and $\frac{\varepsilon_F.E + \varepsilon.AE}{E+AE}$ for species E composed of populations E+AE.

The effective fitnesses are thus average fitnesses, intermediate between those of free and bound organisms (see Figure 8).

Since the processes of association and dissociation are faster than those of death and reproduction, the selection process, which is based on the reproduction/death balance, occurs for "species" rather than for individuals. We then expect that those species with larger effective fitness, or larger fitnesses as far as B and C are concerned, will become predominant. Since the effective fitnesses depend on the ratio of free to bound organisms, they vary with the dissociation constant, which gives rise to the observed transitions. The maximum of effective

fitnesses play a role similar to the minimisation of free energies in thermodynamics of solutions. The phases, nearly pure chemical species or mixtures, that are obtained for a given set of thermodynamics conditions are those which minimize free energy under these conditions. The same is true here: the observed dynamical regimes involve either a strict selection of species as in the case of egoism and mutualism, or coexistence of species, but in both case the highest fitnesses are selected. To carry on this analogy, the mutation rate which increases the populations of subdominant species has a role similar to temperature in physical systems. In ordinary physical systems, phase transitions are obtained by changes of intensive variables such as temperature, or concentrations of some chemical species. In the systems of biologicaly interacting species that we describe here, the most biologicaly plausible factors for possible changes in dynamical regime are the apparition of new mutants with different parameters for fitness, death or association. Resources are extensive variables in this model, the equivalent of volumes in thermodynamics of solutions.

4.1 Dominant species

We have seven equations with seven unknown, and we should, in principle, be able to compute all variables which we expect a priori to be of the same order of magnitude. In fact this is never observed; if one tries to solve directly the system, some variables have to be set to zero to avoid contradictions. This means that only some of the variables have strictly positive values; they correspond to the dominant populations with larger effective fitness (see Figure 6) and they can be directly computed by solving the slow manifold equations without mutation terms. The other variables are in fact smaller by a factor $\frac{m}{d}$; they can be computed from the dominant populations by taking into account the mutation terms that we have previously neglected.

The five last equations (4.3) to (4.7) describe the balance between proliferation and death of species A, D, E, B and C. Not all populations can satisfy these equations at the same time. Equations (4.4) and (4.7), for instance, when combined, give:

$$\gamma = \frac{\delta_F.D + \delta.AD}{D+AD} \tag{4.8}$$

which implies:

$$\delta_F < \gamma < \delta \tag{4.9}$$

in contradiction with expression (2.2), $\delta_F = \gamma$.

In fact, the species with larger effective fitness, D, predominates with a large population of order $\frac{\delta.f2}{d}$ (cf. equation (1.2)). The mutation term has to be re-introduced in equation (4.7) in order to compute C which is smaller than D+AD by a factor of order $\frac{m}{d}$.

The different dynamical regimes which are observed by computer simulations correspond to different branches of the slow manifold where different sets of populations are dominant. There is a limited domain of parameters where a particular attractor solution exists, which gives a necessary condition to reach this attractor. To actually solve the equations, one selects a set of predominant populations, being guided by the simulation results. The self consistency of the choice is checked by solving the equilibrium equations. The following algebraic computations using the slow manifold approximation allows to predict the possible attractors for a given set of parameters. But, unless the attractor is unique, knowing the domains of existence of the attractor does not predicts which possible attractor is reached for every initial condition.

4.2 The coexistence regime

Let us start with the coexistence regime, where all species except C are dominant. The following change of variable is made:

$$x = \frac{AE}{E+AE}\ , \quad y = \frac{AD}{D+AD}\ , \quad z = \frac{E+AE}{D+AD+E+AE}\ , \ K = \frac{(ks_d+2.d)}{(ks_e+2.d)} \tag{4.10}$$

The equations (4.1) and (4.2) being combined in one, the system may be now simplified into:

$$K\frac{y}{1-y} = \frac{x}{1-x} \tag{4.11}$$

A is neglected with respect to AE and AD. This is based on simulation results. A amplitude could also be estimated from equations (4.1) and (4.2), taking into account the fact that $\frac{AD}{D}$ and $\frac{AE}{E}$ are of order 1.

$$\alpha_F y.(1-z) + \alpha.x.z = \beta.\ (x.z + y.(1-z)) \tag{4.12}$$

$$\delta_F.\ (1-y) + \delta.y = \frac{d.(C+D+E+AD+AE)}{f_2} \tag{4.13}$$

$$\varepsilon_F.\ (1-x) + \varepsilon.x = \frac{d.(C+D+E+AD+AE)}{f_2} \tag{4.14}$$

$$A + AD + AE + B = \frac{\beta.f_1}{d} \tag{4.15}$$

Figure 8 is a graphical resolution of equations (4.11), (4.13) and (4.14). This diagram shows the linear increase of the effective fitnesses as a function of the fraction of linked organisms. Equating both left handside of equations (4.13) and (4.14) shows that equality among effective fitnesses, which is the condition for coexistence of species D and E, is obtained *iff* the segments delimited on the horizontal by the fitness lines are in the ratio corresponding to equation (4.11). Two solutions exists for large values of K, but only the one closer to x=1, corresponding to the highest fitness, is stable. The other one is close to y=0. When K decreases,

the two solutions get closer until they collapse for the transition from coexistence to egoism. The algebraic solution is obtained by equating both left handside of equations (4.13) and (4.14) which allows to express y as a function of x. Putting the resulting y in equation (4.11) gives a second degree equation in x. Cancelling the discriminant of this equation gives the lower boundary in ks_d for the existence of the coexistence attractor (the corresponding transition line is the lower dotted line on Figure 6). The boundary only depends upon the ratio of association constants K and upon the fitnesses of species E and D. It is independant of available resources.

An upper boundary in ks_d is obtained by using equation (4.12) to compute the actual magnitudes of the populations; it is reached for rather large values of ks_d when B goes to 0. In between, equations (4.11) to (4.15) allow us to compute the populations of the dominant species and C is computed from the original equation (3.5) with mutations.

4.3 The mutualistic regime

In the other regimes, the set of dominant species is even smaller. We could of course systematically try all possible combinations of species to find out whether they satisfy equations (4.11) to (4.15), but the easiest way is to make use of simulation results to select them.

In the mutualistic regime, the dominant population is AE. It is computed from equation (4.3) by neglecting all other populations on f1 branch :

$$AE = \frac{\alpha . f_1}{d} \tag{4.16}$$

Using this result in equation (4.5) gives a second degree equation for x, the ratio of bound E to total E (as defined above in (4.10)):

$$(\varepsilon - \varepsilon_F) . x^2 + \varepsilon_F . x - \frac{\alpha . f_1}{f_2} = 0 \tag{4.17}$$

The condition x<1 gives the following inequality between the fitnesses and the available foods:

$$\varepsilon.f_2 > \alpha.f_1 \tag{4.18}$$

When both sides are divided by d, they represent the maximum population size of species E and A. Inequality (4.18) is then interpreted as the possibility for E to eventually bind to all available A. This mean that mutualism can be established by compensating the difference in fitness between E and D inside A by a large enough rejection rate for D, provided that the food available to species D and E is in sufficient amount so that E can always saturate A. Otherwise, however large ks_d is with respect to ks_e, the effective fitness of A is not larger than β, and mutualism cannot be established.

The mutualistic regime is observed as long as the effective fitness of species E is larger than that of D. A transition to coexistence occurs when both fitnesses become equal. The effective fitness of E is expressed in terms of x computed from (4.17). x is used to compute y by (4.11), which is then used to compute the effective fitness of D. The transition line of Figure 6 (the upper dotted line), which limits the domain of existence of mutualism, is obtained when the values of x and y are replaced in the equality:

$$\delta_F.(1-y)+ \delta.y = \varepsilon_F.(1-x)+ \varepsilon.x \tag{4.19}$$

The selfish regime

The dominant species in the selfish regime are B and D given by equations (4.4) and (4.6):

$$B = \frac{\beta.f_1}{d} \qquad D = \frac{\delta.f_2}{d} \tag{4.20}$$

The most important mutants AD and E are obtained from equations (4.3) and (4.5)where the

most important mutation terms, those involving B and D, have been reintroduced in the right hand side:

$$AD = \frac{\beta.m.B}{d(\beta-\alpha)} \qquad E = \frac{\delta_F.m.D}{d(\delta_F-\varepsilon_F)} \tag{4.21}$$

A and AE are finally obtained by the kinetics equations (4.1) and (4.2).

The limit of stability of the selfish regime is reached when the effective fitness of A reaches that of B and the same is true for E and D. In the algebraic computation of the stability limit, the fitnesses α and ε_F in the denominator have to be replaced by the effective fitnesses obtained by iteration of the above computation to achieve a reasonable accuracy. Let us note that the stability of the egoism attractor depends upon the mutation rate, which is not the case for the other regimes.

All the quantitative predictions of the slow manifold analysis concerning the values of the populations at equilibrium and the limits of stability of the different attractors have been checked against the numerical integration of the complete model. The agreement is excellent and discrepancies are never greater than the percent. Apart from the transitions between different attractors according to the initial conditions, which can only be obtained by numerical integration, the slow manifold analysis is then a very good predictor of the dynamics.

Figure 9 is a summary of the limits of stability of the different attractors obtained from numerical simulations and algebraic analysis of the slow manifold. The abcissa of this diagram is the ratio of the maximum population on each branch (hosts and parasites):

$$\frac{f1.\alpha}{f2.\varepsilon} \cdot \frac{d^*}{d}$$

The second fraction correspond to the case where the host A offers protection to the

symbiont E, thus decreasing its death coefficient d with respect to free species (see section 6). The ordinate, as in Figure 7, is the dissociation constant ks_d. We are presently concerned with the case when d*= d (the black dots on the figure). Let us note the vertical asymptot of the lower mutualism limit when both maximum populations are equal, as obtained from equation (4.17). The monotonously decreasing upper curve is the upper limit of egoism. When this limit is overcome, transition to mutualism is obtained on the left part of the diagram and to coexistence on the right. The bifurcation point between the two possible regimes after the transition is very close to the left of the intersection with the lower stability of mutualism. In other words, nearly as soon as mutualism is a possible attractor, it is reached when egoism becomes unstable. The condition for this scenario is the possibility for the symbionts to saturate the available hosts. Otherwise, mutualism is impossible and transition to coexistence is observed. The diagram of Figure 6 with several reachable attractors is in fact only obtained for a small parameter region in the vicinity of the equality of maximum populations.

5 The multi-site model (MSM)

Binary associations are not the rule and in most cases the host offers a number of sites, p, to the invading species (Figure 10). This is the case, for instance, for corals with algae inside polyps. The straightforward extension of the above model would be to write a large system of ordinary differential equations which would include all of species such as AD_rE_s, where r and s vary from 0 to p and indicate the number of D and E organisms inside A. In this large system, we can expect that after any reproduction, mutation or death events, the induced fluctuations in the rate of occupancy of the host by the endosymbionts D and E decay very fast while the populations reach their value on the slow manifold. Since we don't care much about these fluctuations, we can enormously simplify the model. A simple approach, which takes into account the different time scales for association and growth, is to suppose that association and dissociation processes are quasi-instantaneous. Apart from the free species, we only have to deal with variables A, AD and AE. AD (resp.AE) is the population of organisms of type D (resp.E)

inside A. A is the population of organisms A which are offering p sites per individual to invading organisms D and E. Among these p sites, AE/A sites are occupied by E organisms and AD/A are occupied by D organisms.

In other words we are neglecting fluctuations of the occupancy rates of D and E inside the different A organisms. This approximation, equivalent to a mean field approach, is based on the fast association kinetics. The differential system is then written:

$$\frac{dA}{dt} = \frac{(\alpha_F.(pA-AE) + \alpha.AE).f1}{p.(A+B)} - (d+m.n).A + m.B \quad (5.1)$$

$$\frac{dAD}{dt} = \frac{\delta.AD.f2}{C+D+E+AD+AE} - (d+m.n).AD + m.AE - ks_d.AD + kl_d.(p.A-AD-AE).D \quad (5.2)$$

$$\frac{dAE}{dt} = \frac{\varepsilon.AE.f2}{C+D+E+AD+AE} - (d+m.n).AE + m.AD - ks_e.AE + kl_e.(p.A-AD-AE).E \quad (5.2)$$

$$\frac{dB}{dt} = \frac{\beta.B.f1}{A+B} - (d+m.n).B + m.A \quad (5.3)$$

$$\frac{dC}{dt} = \frac{\gamma.C.f2}{C+D+E+AD+AE} - (d+m.n).C + m.(D+AD) \quad (5.4)$$

$$\frac{dD}{dt} = \frac{\delta_F.D.f2}{C+D+E+AD+AE} - (d+m.n).D + m.(C+E+n.AD) + ks_d.AD - kl_d.(p.A-AD-AE).D \quad (5.5)$$

$$\frac{dE}{dt} = \frac{\varepsilon_F.E.f2}{C+D+E+AD+AE} - (d+m.n).E + m.(D+n.AE) + ks_e.AE - kl_e.(p.A-AD-AE).E \quad (5.6)$$

The growth term for A is a linear function of the fraction of sites occupied by E. Mutations

in A release infesting organisms, resulting in population changes in AD, AE, D and E. They appear as proportionnal to these populations rather than to A because of the occupancy rates. For instance, when an A organism mutates with frequency m, it frees AE/A organisms of species E. The corresponding source term in the E differential equation is then:

$$m.\ A.\ \frac{AE}{A} = m.AE \tag{5.7}$$

When simulated with p=10 or 100, this model (MSM) gives time evolution of populations which are very similar to those obtained with the single site model (SSM). Populations, are the same in both models for the same set of parameters, provided that the following changes are made:

single site model	multi-site model
f1	$\frac{f1}{p}$
B	p.B
A	p.A - AD -AE

Except for their short time linking dynamics, both SSM and MSM are pretty much the same, once one realizes that free sites of the A species in the MSM play the same role as free A in the SSM.

When the slow manifold analysis is carried out for the multi-site model with the above changes, most equations are identical. The last four equations (4.4) to (4.7) of the simplified dynamical system are the same for both models (single site and multi-site, resp. SSM and MSM). For the MSM the first three equations are written:

$$ks_e.AE = kl_e.\ (p.A\text{-}AD\text{-}AE).E \tag{5.8}$$

$$ks_d.AD = kl_d.\ (p.A\text{-}AD\text{-}AE).D \tag{5.9}$$

$$\frac{(\alpha_F.(pA\text{-}AE) + \alpha.AE).f_1}{p.(A+B)} = d.A \qquad (5.10)$$

The two first equations, once combined, give the same equation, (4.11), in x and y as in the SSM. The third equation (5.10) is equivalent in SSM to equation (4.3) when one notices that the free A population of the SSM corresponds to the number of free sites in A, $\frac{pA\text{-}AD\text{-}AE}{p}$, of the MSM. f1 in the SSM then corresponds to $\frac{f1}{p}$ in the MSM.

Since the dynamical equations on the slow manifold are the same in both models, we expect the attractors and their domains of existence to be the same, which is verified by the computer simulations. The only difference concerns the evolution towards the attractors. At short time scales of order $(ks_d)^{-1}$ some differences are indeed observed, consistent with the approximations made in the MSM on the binding dynamics. Otherwise, at larger time scales the dynamics are identical, and even the transition lines from any given initial conditions are very close. Their relative distance in the (ε, ks_d) plane is never more than 1% of the parameters.

In conclusion of this section, all the dynamical results that were obtained for the simple, but rather unrealistic single site model, remain valid for the evolution of the multi-site model which posess some biological relevance.

6 Discussion

6.1 Selective protection by the host

Mutualistic associations sometimes involve protection in exchange for food, for instance in the case of damselfish and anemon, or coelenterate and algae. This protection against predators might be described in our model by changing the death rate inside the host with respect to the death rate outside. By regrouping the first two terms of the differential equation (1.1):

$$\frac{dP_i}{dt} = (\frac{\alpha.f}{P_t} - d).P_i \ + \ldots .$$

one sees that a change in the death rate is equivalent to the same relative change in the fitness coefficient. We have done series of simulations and algebraic computations on the slow manifold to study the "extreme" case when the only benefit enjoyed by the endosymbionts D and E (or the damsel fish in the case of fishes and protective hosts) is a reduction of their death rate d* with respect to their death rate d when they are free.

$$d=0.01 \quad , \quad d^*= 0.007$$

By contrast with sections 3 and 4, for the present simulations, fitnesses are maintained constant inside and outside the host:

$$\delta_F = \delta = 8 \quad , \quad \varepsilon_F = \varepsilon = 7$$

All other parameters are the same as for previous cases.

The 30% decrease in death rate for endosymbionts is comparable to the 30% increase in fitnesses used in sections 3 and 4. We then expect comparable behaviors when the other parameters such as the dissociation rate ks_d and the maximum population ratio $\frac{f1.d^*.\alpha}{f2.d.\varepsilon}$ are the same, which is indeed observed on the phase diagram of Figure 9, where the open dots - corresponding to the model where protection is offered by the host - fall aside the black dots - corresponding to the model where the fitness of the symbionts increases inside the host.

The slow manifold analysis follows the procedures that we have developped in section 4. Only equations (4.4) and (4.5) are changed and become:

$$\frac{dD}{dt} + \frac{dAD}{dt} = \frac{\delta(D+AD).f2}{C+D+E+AD+AE} - (d.D+d^*.AD) = 0 \qquad (6.2)$$

$$\frac{dE}{dt} + \frac{dAE}{dt} = \frac{\varepsilon(E+AE).f2}{C+D+E+AD+AE} - (d.E+d^*.AE) = 0 \qquad (6.3)$$

The equation (6.2) can be rewritten:

$$\frac{dD}{dt} + \frac{dAD}{dt} = \left(\frac{\delta.f2}{C+D+E+AD+AE} - \frac{d.D+d^*.AD}{D+AD}\right) (D+AD) = 0 \qquad (6.4)$$

which shows that the per capita rate of increase is now averaged on the two populations, free and associated, via the death term. The previous analysis (section 4) has shown that instabilities of the different regimes were obtained when average fitnesses became equal. It is now generalized by comparing the averaged per capita rate of increase, which shows that the transitions are obtained by equating average death terms divided by the corresponding fitness. The transitions are then obtained when:

$$\frac{1}{\delta}\,\frac{d.D+d^*.AD}{D+AD} = \frac{1}{\varepsilon}\,\frac{d.E+d^*.AE}{E+AE} \qquad (6.5)$$

The same arguments concerning equalities of the per capita rate of increase allow to predict the transitions in the most general case, when the symbiosis involves simultaneous benefits in fitness and protection.

6.2 Selective digestion by the host.

We have tentatively tried to check wether a selective digestion of the endosymbionts could

result into mutualism. We were driven to this hypothesis by the fact that parasitic algae are sometimes digested by *Hydra* normally involved in association with *Chlorella*. ks_d and ks_e in equations (3.2) and (3.3) now correspond to digestion coefficients. The source terms (ks_d.AD) and (ks_e.AE) are then removed from equations (3.6) and (3.7) since the symbionts die from the digestion process. The positive terms in ks_d are kept in equation (3.1) since the digestion of the symbionts frees A. No mutualism is observed when the digestion is faster or of the same order of magnitude as the death process, which corresponds to the most reasonable hypothesis.

The slow manifold analysis allows to understand the dynamical behavior of the system. Equations (4.4) and (4.5) now become:

$$\frac{dD}{dt} + \frac{dAD}{dt} = \frac{(\delta_F.D + \delta.AD).f2}{C+D+E+AD+AE} - (d+m.n).(D+AD) - ks_d \,.\, AD = 0 \qquad (6.6)$$

$$\frac{dE}{dt} + \frac{dAE}{dt} = \frac{(\varepsilon_F.E + \varepsilon.AE).f2}{C+D+E+AD+AE} - (d+m.n).(E+AE) - ks_e \,.\, AE = 0 \qquad (6.7)$$

The digestion terms add up to the death term and mutation term (d+m.n), which constitutes a strong handicap for species D and E with respect to organisms C. C remains predominant on its branch, except when ks_d is notably smaller than d. Egoism is then the only observed regime, when ks_d is larger than d.

7 Conclusions

Let us summarize the results and discuss the biological significance of the model.

Mutualism does not contradict Darwinian theory of selection of the fittest, provided that one compares species according to their effective fitness, which takes into account the benefits enjoyed by the symbionts while they are associated.

Mutualism can be established when a recognition mechanism allows the host to discriminate among parasites and *bona fide* symbionts. One is then led to look for possible recognition mechanism in biological associations. In the case of rumen/enterobacteria association {Begon, Harper and Trunsend, 1986}, the evident candidate for the recognition function is the immune system. For the rhizobium of legumes associating the legume cells and nitrogen fixing bacteria, lectins able to recognize the polysaccharides on the cells walls of the bacteria also are rather convincing molecular mechanisms {Lis and Sharon, 1986}. Other molecular recognition mechanisms have been documented in very simple organisms (spongae, tunicates, coelenterates) which could be involved in symbiotic associations {Douglas, 1988; Taylor, 1973}. Negative electric charges on the surface of symbiotic chlorellae cell walls allow hydrae to recognize them {McNeil, Hohman and Muscatine, 1981}. Another mechanism, which seems to apply to the *Hydra*/*Chlorella* system, is the detection by the *Hydra* of maltose released by the *Chlorella* (maltose is the benefit enjoyed by the *Hydra* from the *Chlorella*) {Hohman, McNeil and Muscatine, 1982; McAuley and Smith, 1982; Muscatine and McNeil, 1989}.

The differential model that we have built allows to predict the possible emergence of mutualism according to the individual charactistics of the organisms. It applies to most endosymbiotic systems, where association times are smaller than organisms lifetime. It takes into account exchanged benefits such as food and protection. Most of the discussion involved the single site model, but we have shown in section 5 that this simpler model is equivalent to the multiple site model in the time range of interest to us.

The possibity of a coexistence regime with both egoists and true symbionts present with comparable populations was an unexpected result of the model. We are tempted to consider that the possible saturation of the host by the symbionts is the normal case which excludes the possibility of observing coexistence. On the other hand, coexistence could be a possibility in those many systems where we cannot figure out the benefits of the association for each individual organism involved, e.g. lichen where the benefit for the algae is not obvious {Begon, Harper and Trunsend, 1986} or intestinal flora in insects or mammals. Since the coexistence

situation is less favorable to the host than complete mutualism, we might imagine that further mutuations would select host organisms with larger rejection rates of "egoists". The real existence in biological systems of the coexistence regime is still an open question worth investigating.

Although most results can be obtained through direct numerical simulation, the slow manifold analysis brings some insight in the important concepts: effective fitnesses and effective per capita increase, the saturation limit for mutualism, equivalence between single site and multiple site models, and why selective digestion is insufficient to establish mutualism.

Acknowledgments

Computer simulation used GRIND software {De Boer, 1983}. The Laboratoire de Physique Statistique is associated with the CNRS (URA 1306) and Paris VI and Paris VII Universities. This work was started in the Santa Fe Institute which we thank for its hospitality. We thank NATO (CRG 900998) and Fondation Curie external grants program for partial support.

We thank Richard Belew, Bernard Derrida, Nancy Koppel, Luca Peliti, Alan Perelson, Jonathan Roughgarden and Eörs Szathmàry for helpful discussions.

References

Begon, M., J. L. Harper and C. R. Trunsend. 1986. *Ecology*. Blackwell Scientific Publication.

Boucher, D. H., S. James and K. H. Keeler. 1982. Ecology of Mutualism. *Annual Review of Ecology and Systematics* **13**, 315-347.

De Angelis, D. L., W. M. Post and C. C. Travis. 1986. Mutualistic and Competitive Systems. In *Positive Feedback in Natural Systems*. S. A. Levin (Ed.) p 290. Berlin: Springer Verlag.

De Boer, R. J. (1983). GRIND: Great Integrator for Differential Equations. Bioinformatics Group, University of Utrecht, The Netherlands.

Douglas, A. E. 1988. Nutritional interactions as signals in the green hydra symbiosis. In *Cell to cell signals in plant, animal and microbial symbiosis*. S. Scannerini, D. Smith, P. Bonfante-Fasolo and V. Gianinazzi-Pearson (Eds.) p 283-296. Berlin: Springer-verlag.

Hohman, T. C., P. L. McNeil and L. Muscatine. 1982. Phagosome-lysosome fusion inhibited by algal symbionts of *Hydra viridis*. *Journal of cell biology* **94**, 56-63.

Ikegami, T. and K. Kanebo. 1990. Computer Symbiosis - Emergence of Symbiotic Behavior Through Evolution. *Physica D 42* , 235-243.

Lindgren, K. 1992. Evolutionary Phenomena in Simple Dynamics. In *Artificial Life II*. C. G. Langton, C. Taylor, J. D. Farmer and S. Rasmussen (Eds.) p 295-312. Addison-Wesley.

Lis, H. and N. Sharon. 1986. Lectins as molecules and as tools. *Ann. rev. biochem.* **55**, 35-67.

McAuley, P. J. and D. C. Smith. 1982. The green hydra symbiosis. VII Conservation of the host cell habitat by the symbiotic algae. *Proceedings of the Royal Society of London B* **216**, 415-426.

McNeil, P. L., T. C. Hohman and L. Muscatine. 1981. Mechanisms of nutritive endocytosis. II The effect of charged agents on phagocytic recognition by digestive cells. *Journal of Cell Science* **52**, 243-269.

Morse, P. and H. Feshback. 1953. *Methods of Theoretical Physics*. McGrawHill.

Muscatine, L. and P. L. McNeil. 1989. Endosymbiosis in Hydra and the evolution of internal defense systems. *American zoologist* **29**(2), 371-386.

Nowak, M. A. and R. M. May. 1992. Evolutionary games and spatial chaos. *Nature* , 826-829.

Roughgarden, J. 1975. Evolution of Marine Symbiosis - A Simple Cost-Benefit Model. *Ecology* **56**, 1201-1208.

Szathmàry, E. and L. Demeter. 1987. Group selection of early replicators and the origin of life. *Journal of theoretical biology* **128**, 463-486.

Taylor, C. E., L. Muscatine and D. R. Jefferson. 1989. Maintenance and Breakdown of the Hydra-chlorella Symbiosis: a Computer Model. *Proceedings of the Royal Society of London B* **238**, 277-289.

Taylor, D. 1973. The cellular interactions of algal-invertebrate symbiosis. *Advances in marine biology* **11**, 1-56.

Weisbuch, G. 1984. Un modèle de l'évolution des espèces à trois niveaux, basé sur les propriétés globales des réseaux booléens. *Comptes rendus de l'Académie des Sciences de Paris* **298**(III (14)), 375-378.

Wilson, D. S. 1983. The Effect of Population Structure on the Evolution of Mutualism: a Field test involving burying beetles and mites. *American Naturalist* **121**, 851-870.

Parameters	Values
d (frequency of death)	10^{-2}
m (mutation rate)	10^{-5}
n (number of genes)	3
f1 (food shared between A and B)	80
f2 (food shared between C, D and E)	100
Fitness coefficients	
α_F (free organisms A)	8
β (organisms B)	10
γ (organisms C)	8
δ_F (free organisms D)	8
ε_F (free organisms E)	7
α (organisms A in AE)	12
δ (organisms D in AD)	12
ε (organisms E in AE)	between 8 and 12
Association and dissociation constants	
kl_d (organisms D)	1
kl_e (organisms E)	1
ks_d (organisms D)	between 0.1 and 12
ks_e (organisms E)	0.1

Table 1 : Values of the used parameters in the numerical simulations.

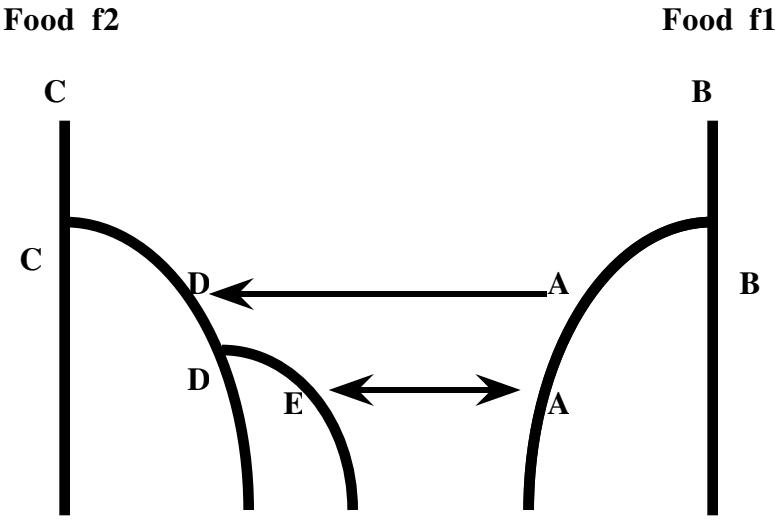

Figure 1 : Phylogeny and interactions of the 5-species system.
The populations of the organisms belonging to the two trees use different resources 1 and 2, available in quantity f1 and f2. Horizontal arrows figure exchanges among organisms.

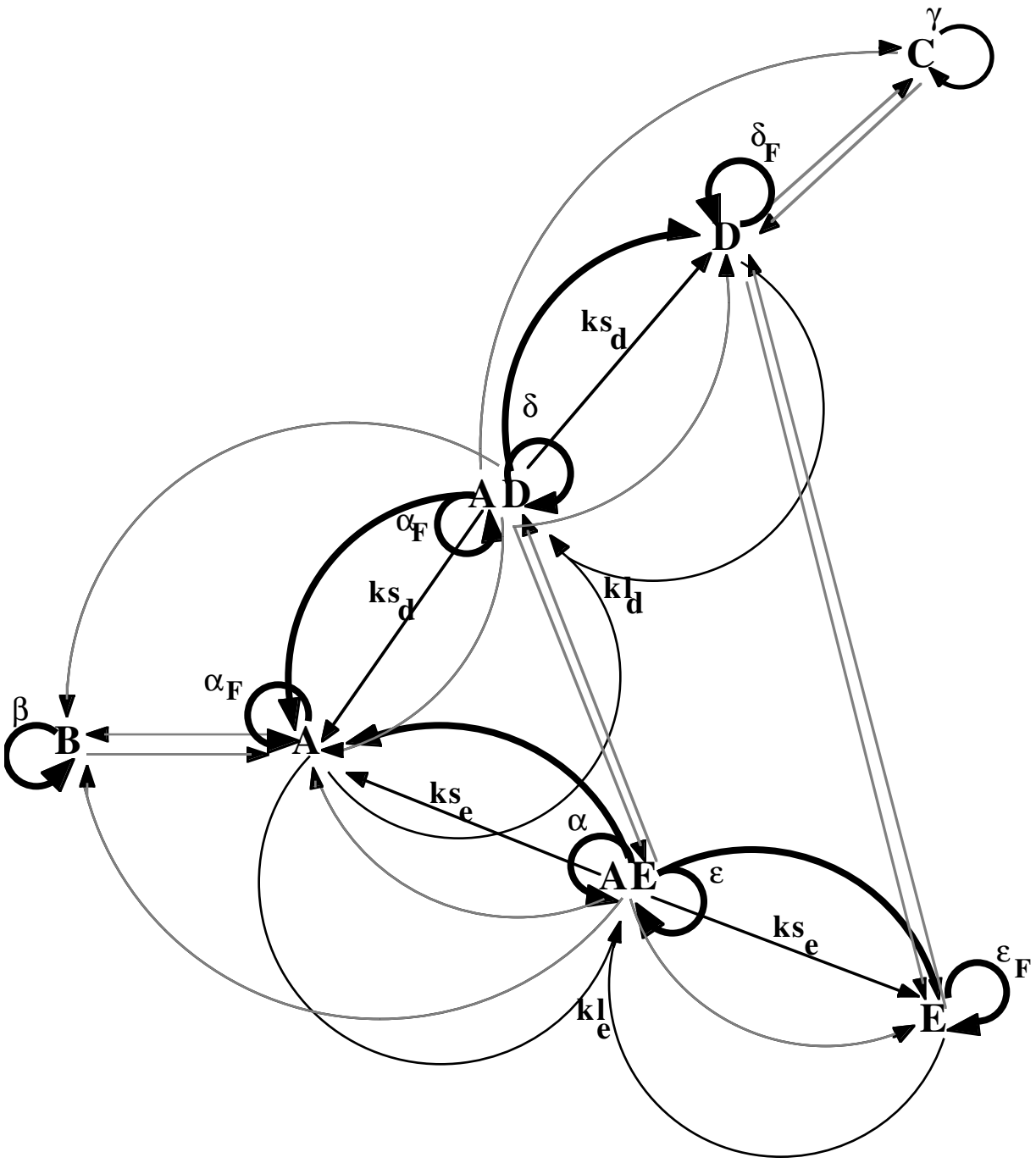

Figure 2: Set of the elementary processes that modify the populations.
Thin continuous lines represent fast association-dissociation processes between hosts and guests organisms with kinetic constants kl_e, ks_e, kl_d and ks_d. Bold lines represent reproduction and death processes (with rate d). The greek letters represent the fitness coefficients associated with the reproduction processes of the organisms, which are either free or associated with another organism belonging to a different species. Dotted lines represent mutation processes with rate m.

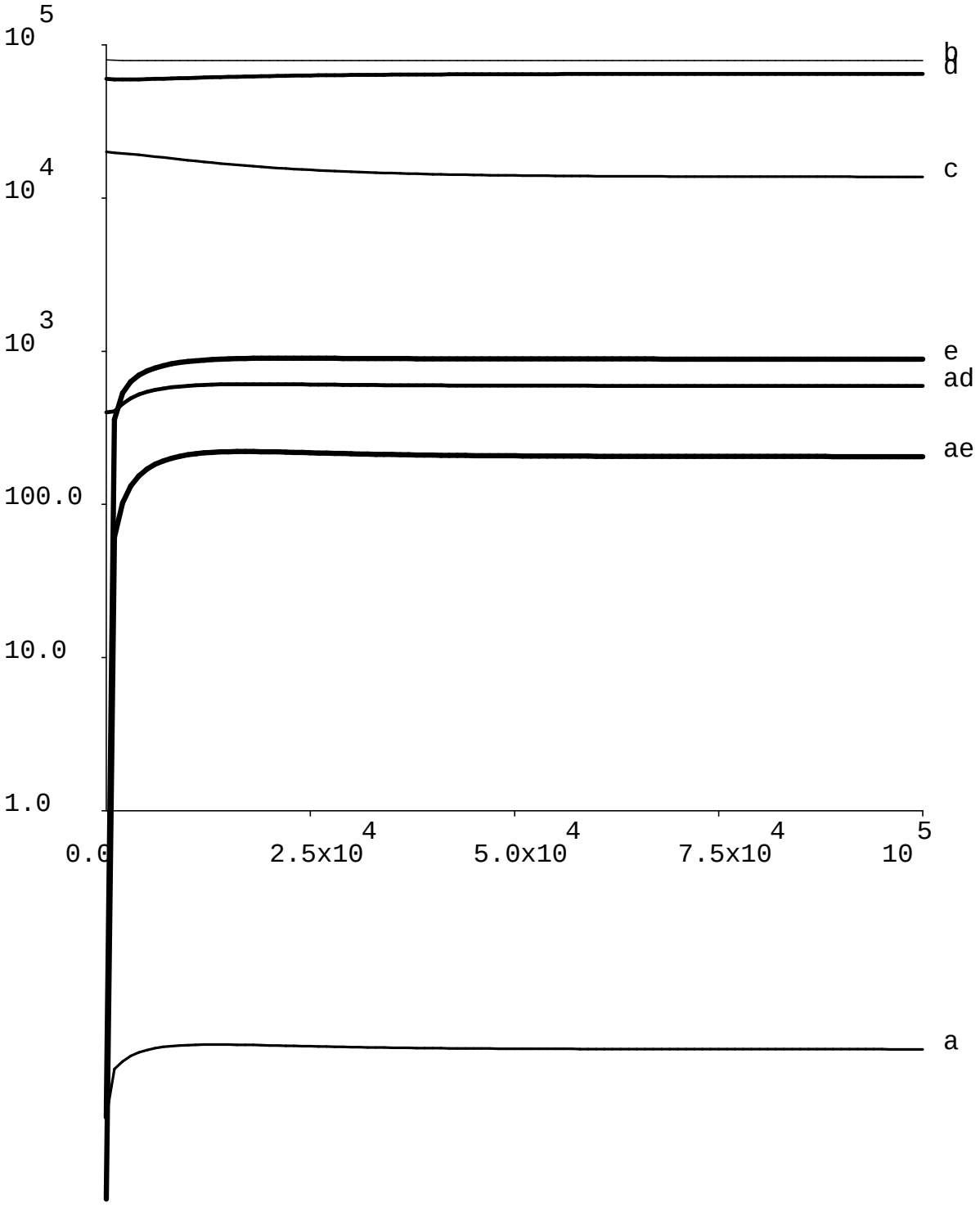

Figure 3 : Prevalence of the primeval species ("egoism attractor") (ks_d=3).
Log plot of the populations versus time obtained by computer simulations using GRIND software. Parameters and initial conditions are the same as in table 1, the rejection rate of D is set at a low value ks_d=3 and ε=10. Mutations from D feed species E. But the insufficient difference in rejection rates between D and E prevents species A and E from growing fast enough to overcome B and D. Although rather little A is available to D, it is enough for D to gain a small selective advantage over C (both otherwise have the same fitness when D is free). D then becomes predominant over C and E.

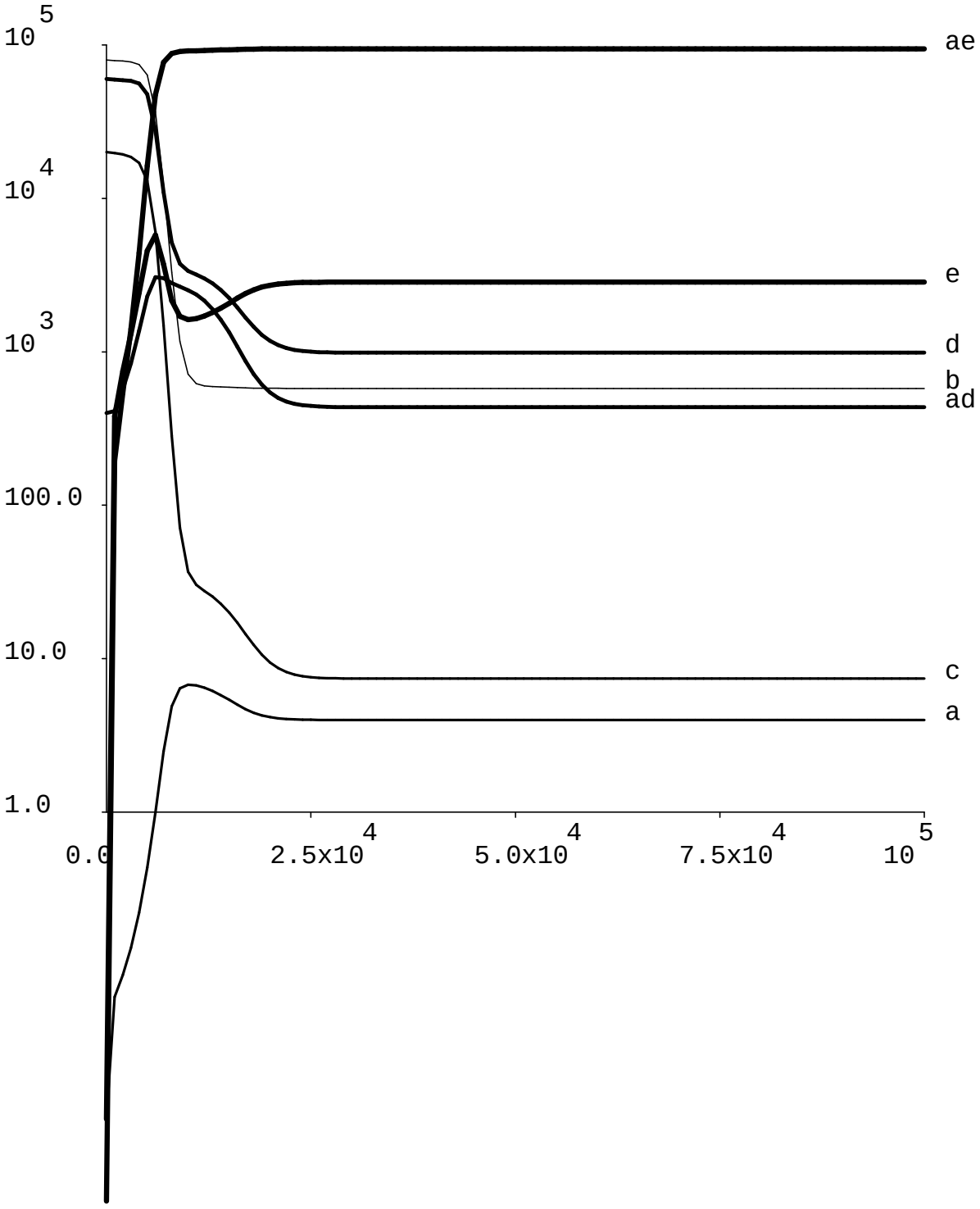

Figure 4 : Emergence of mutualism (ks_d=9)
Parameters and initial conditions are the same as in Figure 2, except ks_d=9. E can then predominantly invade A, which benefits from its presence. The dominant species is AE, which dominates above all other species in both trees.

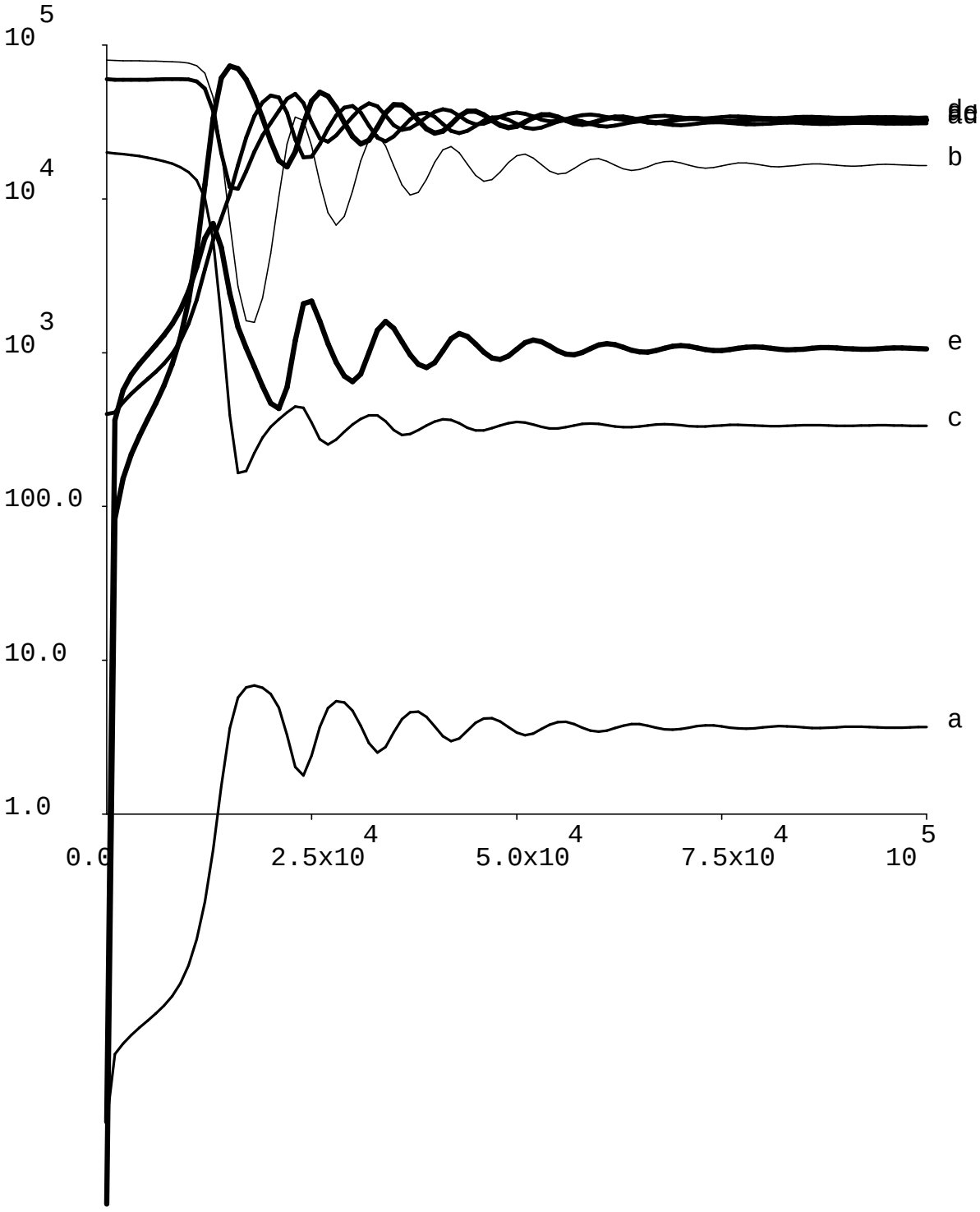

Figure 5 : Coexistence of mutualistic AE, egoist species AD and B (ks_d=7). The oscillations are an indication that this configuration is only weakly attractive.

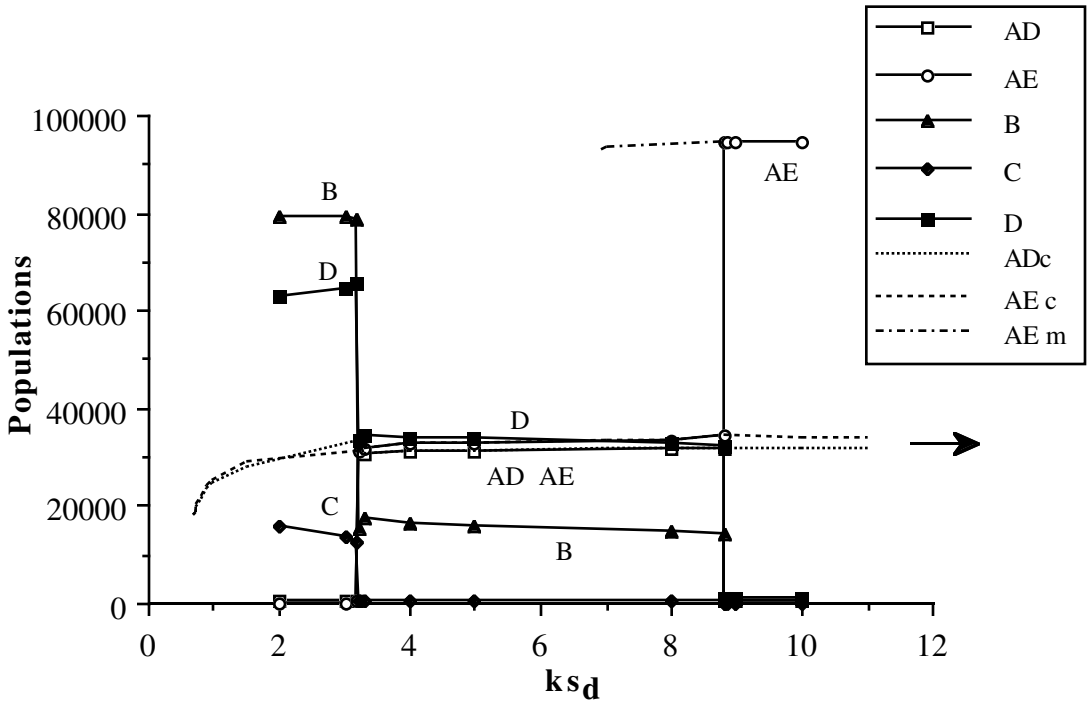

Figure 6: Equilibrium populations measured at time t=100000 as a function of ks_d for ε=10.
The three regimes, predominance of the ancestors, coexistence and emergence of mutualism are separated by sharp transitions. Populations of A and E, very small, are not represented on the diagram. Continuous lines were obtained from initial conditions in the absence of E species (emergence conditions). The dotted lines are obtained with initial populations of the attractors: AD_c and AE_c corresponding to the coexistence attractor and AE_m to the mutualism attractor. The arrow indicates that the coexistence attractor exists up to larger values of ks_d.

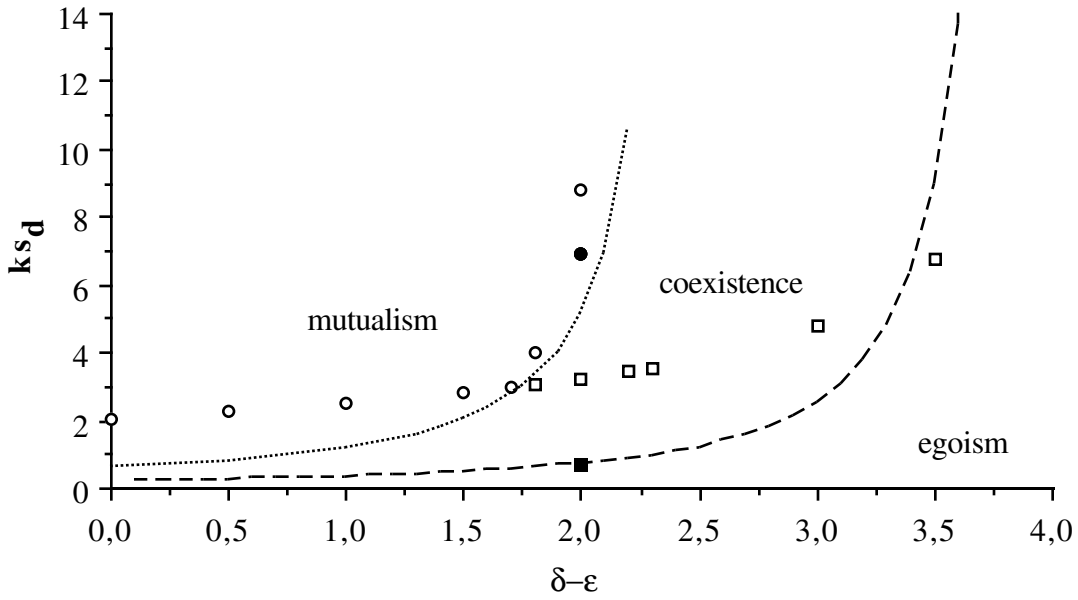

Figure 7 : "Phase diagram" for the emergence of mutualism in the parameter plane ($\delta-\varepsilon$, ks_d).
When the difference $\delta-\varepsilon$ increases, larger rejection rates of D are necessary for the emergence of mutualism. The two dotted lines were obtained by the slow manifold analysis (see section 4): they limit the domain of existence of the attractors. The open dots, obtained by numerical simulations, mark the transitions between the attractors that are reached from initial conditions in the absence of E (circles signal mutualism/ coexistence transition, squares coexistence/ egoism). To ensure consistency with theory, they should lie above the transition lines, which is indeed observed, except for one of them. The black dots correspond to the lower limits of the existence domains of the attractors determined by numerical simulations (round for mutualism, square for coexistence); they should lie on the dotted lines.

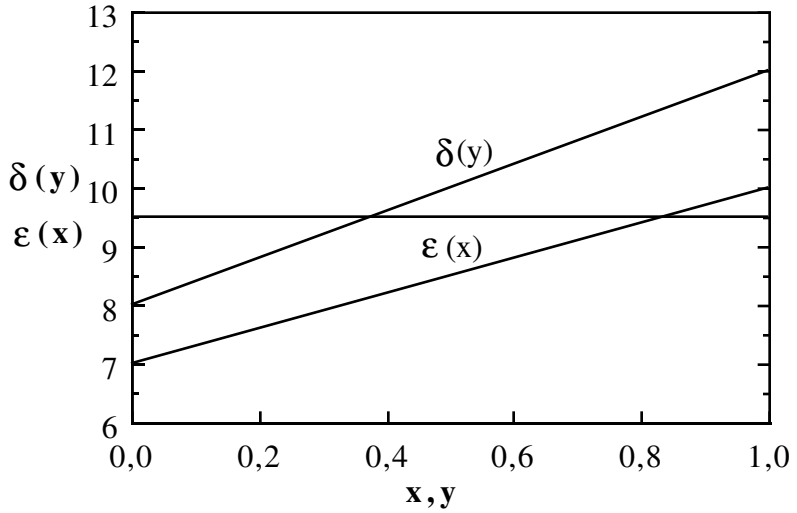

Figure 8 : Effective fitnesses and limits of stability
When x (resp. y), the fraction of associated true symbionts E (resp. parasites D) varies between 0 and 1, the effective fitness $\varepsilon(x)$ averaged on both associated and free forms varies between ε_F and ε (resp. $\delta_F < \delta(y) < \delta$). Equality of effective fitnesses (the transition condition) and the association equilibria require that the segments delimited on the horizontal line verify a ratio K defined in the text (see equation (4.11)).

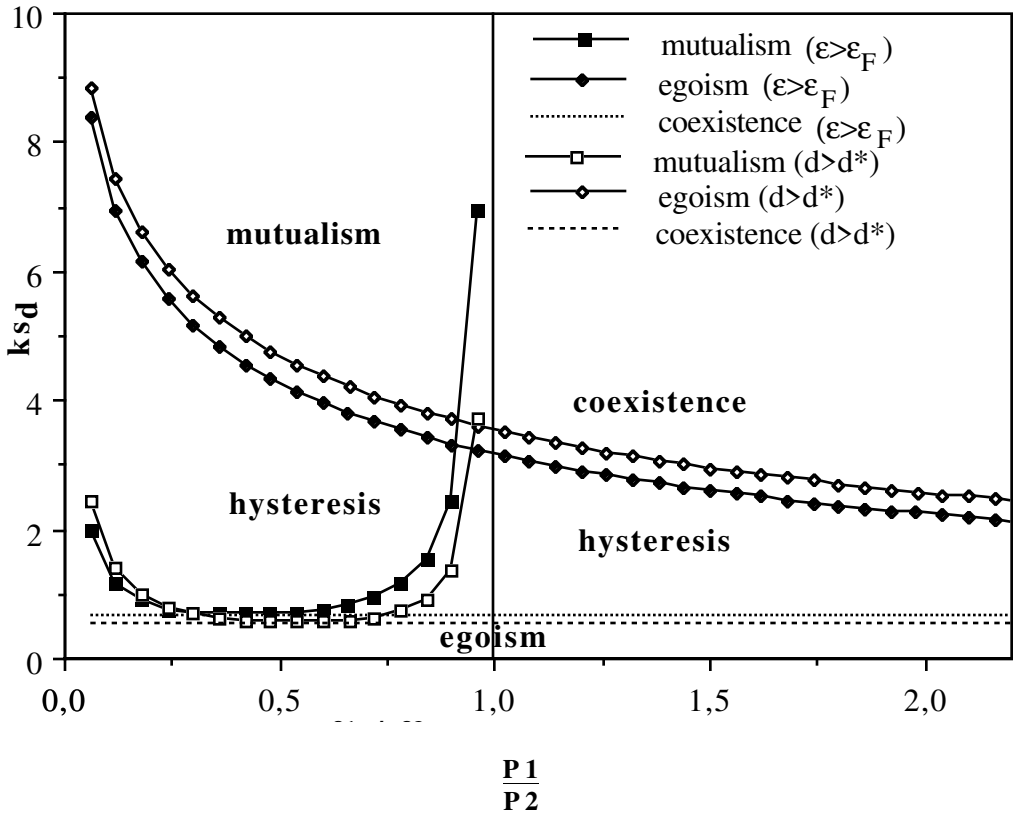

Figure 9: Limits of stability in the plane $(\frac{P1}{P2}, ks_d)$

The transition from egoism to coexistence or mutualism depends on the saturation ratio relating the maximum populations of A species, P1=f1.α.d*, to E species, P2=f2.ε.d, (equations. (4.18) and (6.5)). The vertical line at P1=P2 is a limit to the existence of mutualism which is only possible in the left part of the diagram. The limits of stability of the three regimes are drawn when the endosymbionts enjoys either an increase in fitness (black dots) or in lifetime (open dots). The upper decreasing curves are the upper stability limits of egoism (losanges), and the lower curves are the lower stability limits of mutualism (squares). The horizontal straight lines are the lower stability of coexistence. In the lower ks_d regions egoism, and in the upper regions mutualism or coexistence, according to the possibility of saturation of the host by the symbionts, are the unique possible attractors. In between, which attractor is reached depends from the initial conditions (hysteresis).

E E D E D E E D E E E D

|—|—|—|—|—|—|—|—|—|—|—|—|—|—|

Figure 10 : An organism A with p sites occupied or not by E's or D's.