ON LEARNING AND DIVERGENCE FROM RATIONAL EXPECTATIONS IN AN OVERLAPPING GENERATIONS MODEL

by

Jean-Pascal BENASSY (*) and Michael C. BLAD (**)

N° 8524

June 1985

Revised September 1985

(*) C.N.R.S., CEPREMAP and E.N.S., Paris

(**) Department of Economics, The University of Sydney

"On Learning and Divergence from Rational Expectations

in an Overlapping Generations Model"

by Jean-Pascal BENASSY and Michael BLAD

A B S T R A C T

This paper investigates whether learning does actually lead to "rational expectations" in an overlapping generations model with fiat money. Expectations are formed according to a family of expectations functions containing a rational expectations scheme and indexed by a single parameter. This parameter is updated in every period to fit the latest observations. One finds that, unless the agents have rational expectations to start with, expectations will generally diverge from rational expectations in this simple model.

Journal of Economic Literature Classification Numbers : 020

"Apprentissage et absence de convergence vers les

anticipations rationnelles dans un modèle à générations imbriquées"

R E S U M E

Cet article étudie le problème de la convergence vers les anticipations rationnelles dans un modèle à générations imbriquées avec apprentissage sur les schémas d'anticipations. On considère pour cela une famille de fonctions d'anticipations indicée par un paramètre d'"apprentissage", et incluant un schéma d'anticipations rationnelles pour une certaine valeur de ce paramètre. Celui-ci est réestimé à chaque période sur la base des observations les plus récentes. L'étude dynamique du modèle montre qu'à moins d'avoir initialement des anticipations rationnelles, le processus d'apprentissage des agents ne convergera généralement pas vers celles-ci.

Mots clefs : Anticipations rationnelles, apprentissage, modèles à générations imbriquées.

ON LEARNING AND DIVERGENCE FROM RATIONAL EXPECTATIONS IN AN OVERLAPPING GENERATIONS MODEL

by

Jean-Pascal BENASSY and Michael C. BLAD (*)

1. INTRODUCTION

In studying dynamic models many economists nowadays use the hypothesis of rational expectations or perfect foresight. This assumption yields dynamic evolutions which may be dramatically different from those generated by other expectations schemes, as we shall clearly see in an example below. It is thus quite natural that one should inquire about the foundations of this particular assumption. Of course economic agents are not "born" with rational expectations. So the usual justification for rational expectations is that any other scheme of expectations formation would be eliminated by experience. The implicit presumption is thus that learning would lead to rational expectations.

So quite naturally the problem of whether learning actually does lead to rational expectations has lately attracted a good deal of attention (Cf. Blume-

(*) We are greatly indebted to Daniel Goroff for suggesting some mathematical tools used in Part 5 below. We also would like to thank P.A. Chiappori, R. Farmer, G. Fuchs, C. Gourieroux, P. Malgrange and participants in various seminars for useful discussions. Financial support from the Commissariat Général du Plan and the University of Copenhagen is gratefully acknowledged.

Easley [1982], Bray [1982], Bray-Savin [1984], Brock [1972], De Canio [1979], Fourgeaud-Gourieroux-Pradel [1984], Friedman [1979], Gourieroux-Laffont-Monfort [1983]). The above contributions generally start from a priori given relations linking current variables to expected ones, and establish conditions for convergence towards rational expectations. What we want to do here is to study learning in the framework of a completely specified, and fairly popular model, the overlapping generations model with fiat money introduced by Samuelson (1958) and further studied by Diamond (1965), Cass-Yaari (1966), Gale (1973), Grandmont-Laroque (1973), Cass-Okuno-Zilcha (1979), Balasko-Shell (1981), among others.

Learning in the overlapping generations model with fiat money has actually already been introduced in the pioneering contributions by G. Fuchs (1977a, 1977b, 1979a, 1979b), in which he studies the effects of learning on the stability of the system in models with several consumers.

We shall develop here a simple and slightly modified version of these models, and concentrate on the specific problem of the convergence of expectations towards a rational expectations scheme. The overlapping generations model is particularly attractive for such a study because its dynamic behavior is completely different depending on whether, for example, "naive expectations" or "perfect foresight" is assumed. We shall find that, under very reasonable and natural learning schemes, expectations usually do not converge towards rational expectations.

2. THE MODEL

The basic model we shall use here is the most simple "Samuelsonian" model of overlapping generations, in the temporary equilibrium version as developed by Grandmont-Laroque (1973).

The goods in the economy consist of a single perishable consumption good and fiat money. In each period t the good is exchanged against money at a price p_t. The agents in the economy are consumers of a single type living exactly two periods. At each date there are two consumers alive, one young, one old. The representative consumer has endowments of the good ω_1 and ω_2 in the two periods of his life, and a utility function :

$$U(x_1,x_2) = U(\omega_1 + z_1,\omega_2 + z_2)$$

where z_1 and z_2 are the net purchases in the two periods. There is in the economy a given quantity M of fiat money which is used to transfer wealth from one period to the next (this is actually the only way since the good is non-storable). The idea is that consumers save under the form of money when young, and spend their savings when old. In this scheme all the money is held at the beginning of each period by the old consumer. The net demand of good by the old consumer in period t is thus equal to M/p_t . The net demand of good by the young in period t is the solution in z_1 of the following program :

$$\text{Maximize } U(\omega_1 + z_1 , \omega_2 + z_2) \qquad \text{s.t.}$$

$$\begin{cases} p_t\, z_1 + m = 0 & m \geqslant 0 \\ p^e_{t+1}\, z_2 = m & \end{cases}$$

where p^e_{t+1} is the expected price for the following period and m the quantity of money saved. The restriction $m \geq 0$ is due to the fact that the young consumer cannot borrow. The optimal solution of this program depends only on the ratio p_t/p^e_{t+1} and is thus written

$$z_1\left(\frac{p_t}{p^e_{t+1}}\right)$$

The equilibrium condition on the goods market is written :

$$\frac{M}{p_t} + z_1\left(\frac{p_t}{p^e_{t+1}}\right) = 0$$

The assumptions on expectations formation we shall make below will ensure that an equilibrium always exists. For what follows it will be convenient to rewrite the equation above under the form :

$$\frac{M}{p_t} = \phi\left(\frac{M}{p^e_{t+1}}\right)$$

or, calling

$$q_t = M/p_t \qquad q^e_{t+1} = M/p^e_{t+1}$$

(1) $\qquad q_t = \phi(q^e_{t+1})$

This allows a simple and convenient graphical representation (Gale [1973], Cass-Okuno-Zilcha [1979]). Figure 1 has been drawn under the assumption, which we shall maintain in all that follows, that :

$$\frac{U_2(\omega_1, \omega_2)}{U_1(\omega_1, \omega_2)} > 1$$

where U_1 and U_2 denote the partial derivatives of U with respect to x_1 and x_2, respectively. This assumption guarantees that the case considered is the so-called "Samuelson" case (Cf. Gale [1973]), where agents would indeed be willing to save in a steady state with perfectly expected stationary prices.

Figure 1

We shall also assume that the function ϕ has the "standard shape" (Cass, Okuno, Zilcha [1979]), depicted in Figure 1, i.e. that it is upward sloping. Note that the function ϕ is bounded above by ω_1, i.e. the endowment of the agent when young.

The model just defined can have two steady states, corresponding to the intersection of the ϕ function with the diagonal in the first quadrant. The "golden rule" q^*, corresponding to a stationary price $p^* = M/q^*$, and the "autarchy" state, corresponding to no exchange between generations and with a rate of inflation equal to :

$$\frac{U_2(\omega_1, \omega_2)}{U_1(\omega_1, \omega_2)} - 1$$

The "golden rule" is of course superior to the autarchy state. We can now see quite clearly how the issue of expectations is crucial for the dynamics of the model. Indeed if we assume "rational" expectations, i.e. $q^e_{t+1} = q_{t+1}$, then the dynamics is :

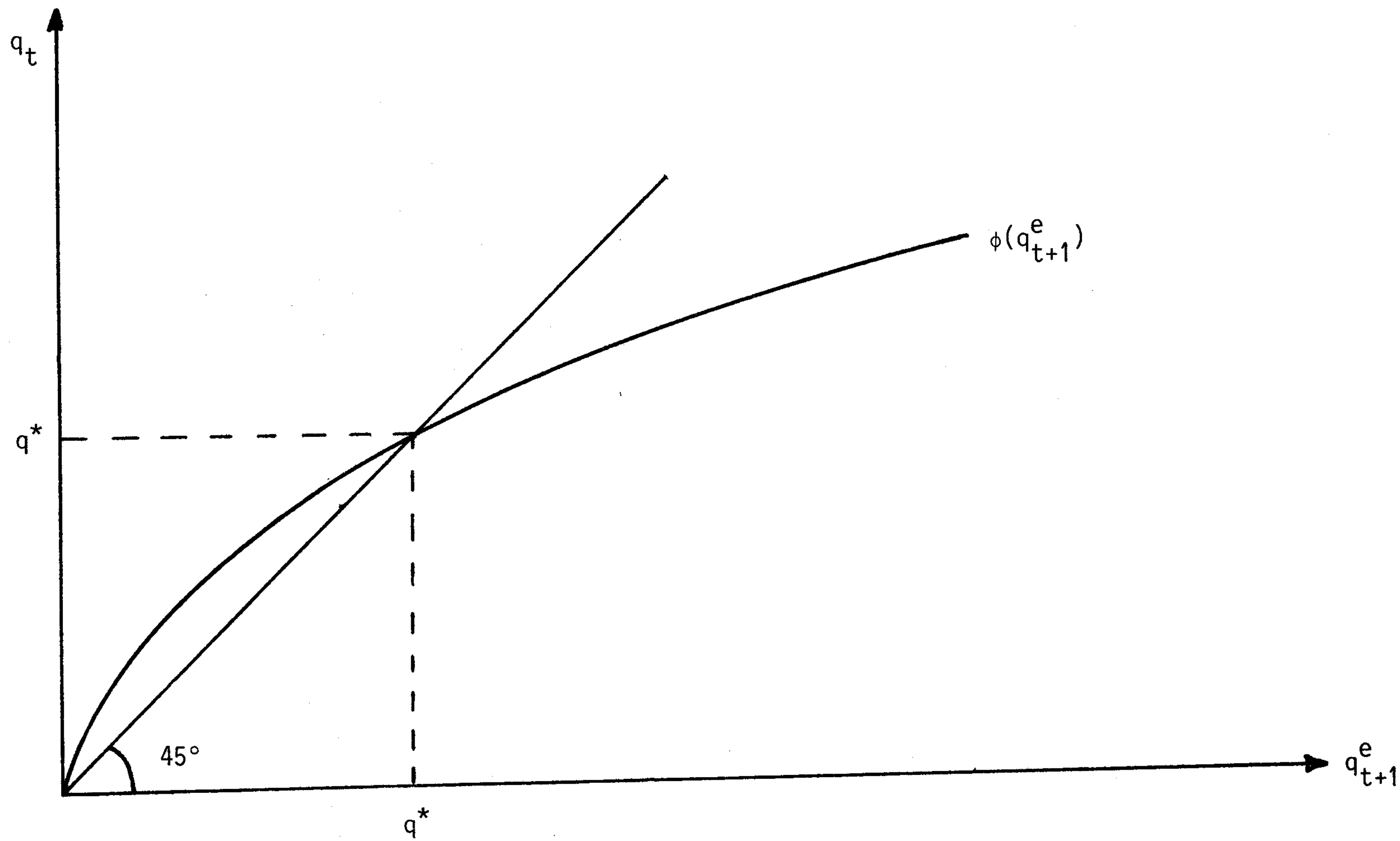

Figure 1

$$q_{t+1} = \phi^{-1}(q_t)$$

and the golden rule is always unstable (Figure 2). Moreover there is convergence to the autarchy state if one starts from a q below q^*. In the contrary if we assume "naïve" expectations, i.e. $q^e_{t+1} = q_{t-1}$, then the dynamics becomes

$$q_t = \phi(q_{t-1})$$

and the golden rule is always stable (Figure 3).

Figure 2

Figure 3

The question we shall address within the framework of this model is whether starting from "some" expectations function, learning will lead towards "rational expectations". A subsidiary question we may sometimes be able to answer is whether the system converges or not towards the golden rule or the autarchy state.

3. EXPECTATIONS AND LEARNING

We thus have a dynamical system where the expectations determine the current variables through the relation

(1) $$q_t = \phi(q^e_{t+1})$$

In order to define the dynamics fully, we need to specify the expectations formation function. We shall assume as in Fuchs [1977a ; 1977b, 1979a, 1979b] that expectations are given by a one parameter family of expectations functions, which we shall take in what follows to be of the form :

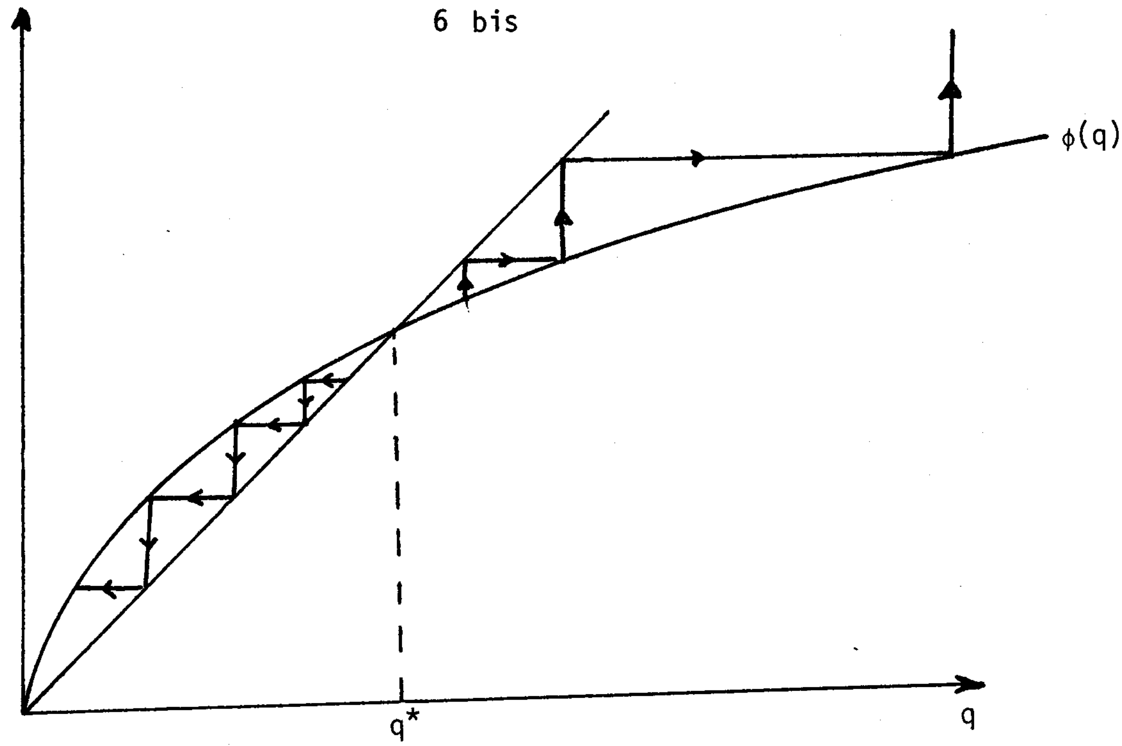

Figure 2

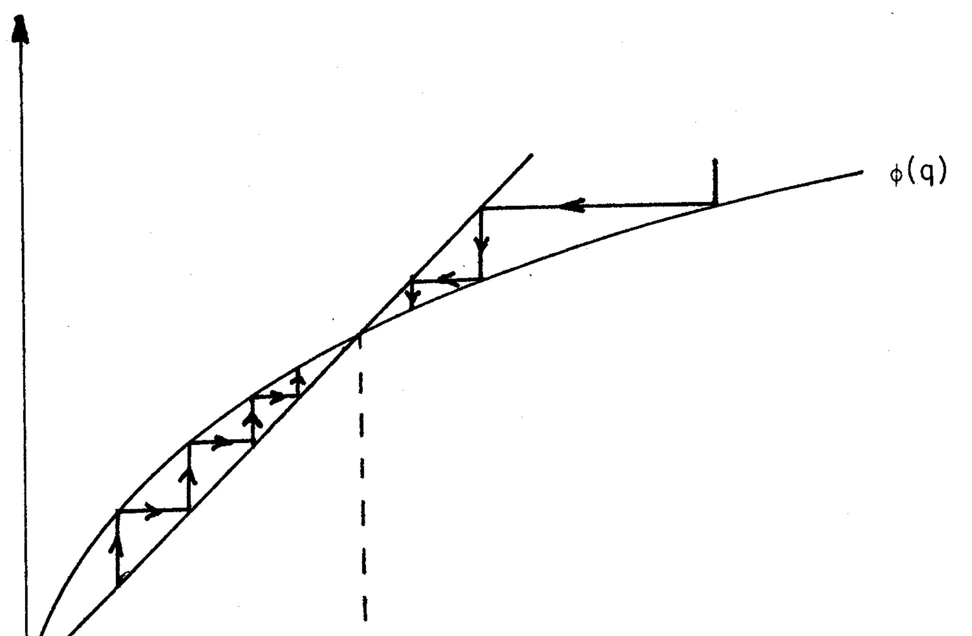

Figure 3

(2) $\quad q^e_{t+1} = \psi(q_{t-1}, \alpha_t)$

where α_t is a parameter which may vary from period to period through learning. Note that since expectations only depend on past values of q , we are sure that an equilibrium always exists (Grandmont-Laroque [1973]).

Since we shall study the problem of convergence towards rational expectations (which is here perfect foresight), we want further a "rational expectations function" to belong to the family , i.e we want that for some parameter value $\bar{\alpha}$, $\psi(.,\bar{\alpha})$ is a rational expectations function.
It is easy to see that on a perfect foresight path $q_{t+1} = \phi^{-2}(q_{t-1})$, so that :

$$\psi(q_{t-1}, \bar{\alpha}) = \phi^{-2}(q_{t-1})$$

From now on $\psi(.,\bar{\alpha})$ will be called "the" rational expectations function, as this is the only one in the family of expectations functions considered. Of course there would be others if current values or additional lagged values of q where included as arguments of ψ . But the complexity of the argument would be dramatically increased.

Learning

The dynamics will now be complete if we indicate how the parameter α_t changes in time due to learning. To describe this evolution we make the particularly simple and natural assumption that α_t is estimated exactly using the last pair of observations available. So the parameter in t + 2 will be fitted to the two observations in t - 1 and t + 1 and α_{t+2} is thus defined by :

(3) $$q_{t+1} = \psi(q_{t-1}, \alpha_{t+2})$$

Standard assumptions will ensure that this equation indeed has a solution in α_{t+2} for all observable q_{t-1} and q_{t+1} . In what follows we shall in fact use the learning rule under a different form, by making explicit the actual value of q_{t+1} as a function of q_{t-1} . Indeed equations (1) and (2) above yield the basic dynamic equation :

(4) $$q_t = \phi[\psi(q_{t-1}, \alpha_t)]$$

which, lagged one period, yields

$$q_{t+1} = \phi[\psi(q_t, \alpha_{t+1})]$$

Combining these last two equations we obtain :

$$q_{t+1} = \phi[\psi(\phi[\psi(q_{t-1}, \alpha_t)], \alpha_{t+1})]$$

Together with equation (3) above, this yields an equation defining implicitly α_{t+2} :

(5) $$\psi(q_{t-1}, \alpha_{t+2}) = \phi[\psi(\phi[\psi(q_{t-1}, \alpha_t)], \alpha_{t+1})]$$

which we shall sometimes write in an explicit form as :

(6) $$\alpha_{t+2} = \varphi(\alpha_{t+1}, \alpha_t, q_{t-1})$$

We thus have a three dimensional dynamical system which to each triple $(\alpha_{t+1}, \alpha_t, q_{t-1})$ associates a new triple $(\alpha_{t+2}, \alpha_{t+1}, q_t) = f(\alpha_{t+1}, \alpha_t, q_{t-1})$

where f is defined component wise by :

$$\begin{cases} \alpha_{t+2} = \varphi(\alpha_{t+1}, \alpha_t, q_{t-1}) \\ \alpha_{t+1} = \alpha_{t+1} \\ q_t = \phi[\psi(q_{t-1}, \alpha_t)] \end{cases}$$

The starting point will consist of an initial price at time zero, p_o, and thus an initial q_o, and two initial expectations functions corresponding to α_1 and α_2. The object of the study is to investigate the stability properties of the system, and notably whether α_t converges to $\bar{\alpha}$, which would correspond to convergence towards rational expectations.

4. GLOBAL ANALYSIS : AN EXAMPLE

We shall now study the above dynamical system using a particular family of expectations functions ψ for which we shall be able to derive the full dynamics of the system. Consider indeed the following family of expectations functions

$$\psi(q_{t-1}, \alpha_t) = \phi^{\alpha_t}(q_{t-1})$$

where α_t is a positive or negative integer and ϕ^{α_t} denotes the α_t-th iterate of the function ϕ. For $\alpha_t = 0$ one has the "naïve" expectations function and for $\alpha_t = \bar{\alpha} = -2$ the rational expectations function. With this specification the above equation defining α_{t+2} (equation 5) becomes :

$$\phi^{\alpha_{t+2}}(q_{t-1}) = \phi^{2+\alpha_t+\alpha_{t+1}}(q_{t-1})$$

If one starts at $q_o = q^*$, then this equation admits any α_{t+2} as a solution, but this does not matter since the system will always stay at the golden rule. If however the starting point is different from the golden rule, then the above equation simplifies to :

(7) $\alpha_{t+2} = 2 + \alpha_t + \alpha_{t+1}$

This simple equation will allow us to make a complete study of the dynamics of the system. The results are the following : If one starts with $\alpha_1 = \alpha_2 = \bar{\alpha}$ (rational expectations), then one will always keep the rational expectations function, quite a natural result. But if the starting point is not rational expectations (i.e. $\alpha_1 \neq \bar{\alpha}$ or $\alpha_2 \neq \bar{\alpha}$) , then the system will always diverge from rational expectations. Two cases can actually occur : In one case, (which occurs in particular if one starts from naïve expectations, i.e. if $\alpha_1 = \alpha_2 = 0$) , the α_t 's go to plus infinity, and the system converges towards the golden rule. We should note in that case that, even if the system does not converge towards the rational expectations function, the equilibrium point (the golden rule) will be correctly forecasted. In the other case, the α_t 's go to minus infinity and the prices tend to infinity (i.e. the q_t 's tend to zero) so the system approaches the autarchy state. We shall now make the above statements more precise, and indicate the dividing line between the two cases through a theorem. Before that we only need to define the golden number γ , which we recall, is the positive solution of the equation, well-known to architects and mathematicians:

$$\gamma - 1 = \frac{1}{\gamma} \qquad \gamma > 1$$

We are now ready to prove :

<u>Theorem 1</u>. "<u>The Golden Number meets the Golden Rule</u>"

Consider an initial q_o between 0 and q^* and two initial values α_1 and α_2, then :

(a) Unless both α_1 and α_2 are equal to $\bar{\alpha} = -2$, the rational expectations value, α_t will diverge from this rational expectations value.

(b) If $\alpha_1 + 2 + \gamma(\alpha_2 + 2) > 0$, then α_t goes to plus infinity and q_t converges to the golden rule

(c) If $\alpha_1 + 2 + \gamma(\alpha_2 + 2) < 0$, then α_t goes to minus infinity and q_t goes to zero, i.e. the prices tend to infinity.

<u>Proof</u> : The starting point is to study the dynamics of the α_t's . Let us first call :

$$\beta_t = \alpha_t - \bar{\alpha} = \alpha_t + 2$$

Then equation (7) is rewritten as :

$$\beta_{t+2} = \beta_t + \beta_{t+1}$$

We recognize here the definition of a Fibonacci sequence. In order to study its evolution, let us decompose β_2 as follows :

$$\beta_2 = -\frac{\beta_1}{\gamma} + \varepsilon$$

Then it is trivial to check recursively that β_t can be written as :

$$\beta_t = \left(-\frac{1}{\gamma}\right)^{t-1} \beta_1 + \varepsilon f_t$$

where f_t is the t - th number in a Fibonacci sequence starting with $f_1 = 0$ and $f_2 = 1$. It is well-known that $f_t \to +\infty$ for $t \to \infty$, and thus the dynamics of β_t depends on the sign of ε . But, from the definition above :

$$\varepsilon = \gamma\beta_2 + \beta_1 = \gamma(\alpha_2 + 2) + \alpha_1 + 2$$

To prove the three statements in theorem 1 now becomes quite easy :

(a) Because α_1 and α_2 are integers and γ is not rational, ε cannot be zero unless both α_1 and α_2 are equal to $\bar{\alpha} = -2$.

(b) If $\alpha_1 + 2 + \gamma(\alpha_2 + 2) > 0$, then $\varepsilon > 0$ and thus α_t tends to infinity. As the dynamics is given by :

$$q_t = \phi^{1+\alpha_t}(q_{t-1})$$

it is clear that for all $q_0 > 0$, q_t tends towards q^* as t goes to infinity.

(c) If $\alpha_1 + 2 + \gamma(\alpha_2 + 2) < 0$, then $\varepsilon < 0$ and thus α_t tends to minus infinity. Again with the dynamics given by

$$q_t = \phi^{1+\alpha_t}(q_{t-1})$$

q_t tends to zero whenever q_0 is less than q^* .

Q.E.D.

5. *THE GENERAL CASE*

We shall now study the case of a general expectations function $\psi(q,\alpha)$. For this function we shall make the following assumptions, which naturally generalize the specific example discussed in the previous section :

A1 : $\psi_q > 0$ for all q , α

A2 : $\psi_\alpha > 0$ for all α and $0 < q < q^*$

A3 : $\psi(0,\alpha) = 0$ for all α

$\psi(q^*,\alpha) = q^*$ for all α .

Note that A3 leaves the possibility that the system described in section 3 may still converge towards the golden rule or the autarchy state even if expectations do not converge towards rational expectations. A3 is in fact not imperative, but it will make the discussion easier.

Let us briefly recall the general dynamical system from above :

$$\alpha_{t+2} = \varphi(\alpha_{t+1},\alpha_t,q_{t-1})$$

$$\alpha_{t+1} = \alpha_{t+1}$$

$$q_t = \phi[\psi(q_{t-1},\alpha_t)]$$

where the function φ is the explicit solution to the following implicit equation in α_{t+2} :

$$(5) \qquad \psi(q_{t-1},\alpha_{t+2}) = \phi[\psi(\phi[\psi(q_{t-1},\alpha_t)],\alpha_{t+1})]$$

It would be difficult to study in full generality the dynamics of the above system, but we shall be able to show that usually it does not converge to rational expectations in a sense we shall make precise in the theorems below. Before doing this, let us first make a few remarks.

First, if $q_o > q^*$, then the system will <u>never</u> converge towards rational expectations. Indeed A1 and A3 imply that q_t will always remain above q^*, and there is no rational expectations path with the sequence q_t always above q^*. Secondly if $q_0 = q^*$, again in view of A3 the system will always remain at q^*, while the α_t's will be undetermined. Consequently this case is not very interesting.

In what follows we shall therefore only study the dynamics of the system when q_0 is in the open interval $]0,q^*[$. In this case we find that whatever the starting point, rational expectations are "self-reproducing" in the sense that if $\alpha_t = \alpha_{t+1} = \bar{\alpha}$, then $\alpha_{t+2} = \bar{\alpha}$ is the only solution to equation (5), independently of the level of q_{t-1}. However, if one does not start with rational expectations then one will generally not converge to it. If one specifically considers the quite natural assumption that $\alpha_1 = \alpha_2 \neq \bar{\alpha}$, then the system will never converge towards rational expectations. In fact we shall now first prove the stronger result that the system will <u>never</u> converge to rational expectations if $\alpha_1 \neq \bar{\alpha}$ or $\alpha_2 \neq \bar{\alpha}$, and $(\alpha_1 - \bar{\alpha})(\alpha_2 - \alpha) > 0$.

<u>Theorem 2</u>. Assume $q_0 \in]0,q^*[$. If $(\alpha_1 - \bar{\alpha})(\alpha_2 - \bar{\alpha}) > 0$, then the system never converges towards rational expectations (i.e α_t never converges to $\bar{\alpha}$) unless both $\alpha_1 = \bar{\alpha}$ and $\alpha_2 = \bar{\alpha}$.

<u>Proof</u> : The proof of the theorem is essentially based on the following lemma :

<u>Lemma 1</u>. Assume $q_{t-1} \in]0,q^*[$. Then :

$$\alpha_{t+1} > \bar{\alpha} \implies \alpha_{t+2} > \alpha_t$$

$$\alpha_{t+1} = \bar{\alpha} \implies \alpha_{t+2} = \alpha_t$$

$$\alpha_{t+1} < \bar{\alpha} \implies \alpha_{t+2} < \alpha_t$$

<u>Proof</u> : We shall prove this lemma in the case where $\alpha_{t+1} > \bar{\alpha}$, the proof in the other cases being similar. Using assumption A2 we have :

$$\alpha_{t+1} > \bar{\alpha} \implies \psi(q,\alpha_{t+1}) > \psi(q,\bar{\alpha}) = \phi^{-2}(q) \qquad 0 < q < q^*$$

Applying this inequality with the particular value $q = \phi[\psi(q_{t-1},\alpha_t)]$, we find:

$$\phi[\psi(\phi[\psi(q_{t-1}, \alpha_t)] , \alpha_{t+1})] > \phi[\phi^{-2}(\phi[\psi(q_{t-1},\alpha_t)])] = \psi(q_{t-1},\alpha_t)$$

By equation (5) the left hand side of this inequality is equal to $\psi(q_{t-1},\alpha_{t+2})$, so :

$$\psi(q_{t-1},\alpha_{t+2}) > \psi(q_{t-1},\alpha_t)$$

and by assumption A2 :

$$\alpha_{t+2} > \alpha_t$$

Q.E.D.

We can now complete the proof of Theorem 2. Assume first that $\alpha_1 \geq \bar{\alpha}$ and $\alpha_2 \geq \bar{\alpha}$ with not both being equal to $\bar{\alpha}$. Then a simple application of lemma 1 shows that both sequences $\langle\alpha_2,\alpha_4,\ldots,\alpha_{2k},\ldots\rangle$ and $\langle\alpha_3,\ldots,\alpha_{2k+1},\ldots\rangle$ are monotone increasing which of course implies that they diverge from $\bar{\alpha}$, the "rational expectations" value.

A symmetric reasoning can be done for $\alpha_1 \leq \bar{\alpha}$ and $\alpha_2 \leq \bar{\alpha}$, but not both equal to $\bar{\alpha}$.

Q.E.D.

We see that Theorem 2 ensures divergence for a reasonable set of initial parameters. But, though they are less likely to occur, we would like also to know what happens for initial parameters α_1 and α_2 such that $(\alpha_1-\bar{\alpha})(\alpha_2-\bar{\alpha}) < 0$.

Theorem 3 below will show that, at least in a neighborhood of $(\bar{\alpha},\bar{\alpha},0)$ the system "almost always" diverges. In order to get a better understanding of the dynamics, we shall however first characterize globally the phase portrait of the dynamical system when $(\alpha_1-\bar{\alpha})(\alpha_2-\bar{\alpha}) < 0$, i.e. when in (α_1,α_2) plane the initial point lies in the interior of the second or the fourth quadrant(*) (Recall that Theorem 2 characterized the dynamics in the complementary case). We shall thus first prove Proposition 1, which covers the case where (α_1,α_2) is in the interior of the second quadrant. The proposition is actually slightly more general characterizing the evolution when, for a given $t \geq 1$, (α_t,α_{t+1}) belongs to the interior of the second quadrant :

Proposition 1.

Assume $\alpha_t < \bar{\alpha}$ and $\alpha_{t+1} > \bar{\alpha}$. Then there are three possibilities :

(a) $(\alpha_{t+2},\alpha_{t+3})$ is in the first quadrant (origin excluded)

(b) $(\alpha_{t+2},\alpha_{t+3})$ is in the interior of the second quadrant, and :

$$\alpha_t < \alpha_{t+2} < \bar{\alpha}$$

$$\bar{\alpha} < \alpha_{t+3} < \alpha_{t+1}$$

(c) $(\alpha_{t+2},\alpha_{t+3})$ is in the third quadrant (origin excluded).

(*) Here and in what follows, the origin for the two α-coordinates is taken at $(\bar{\alpha}\ ,\ \bar{\alpha})$.

Proof : There are two possibilities for α_{t+2} :

$$\alpha_{t+2} > \bar{\alpha} \quad \text{or} \quad \alpha_{t+2} < \bar{\alpha}$$

If $\alpha_{t+2} > \bar{\alpha}$, by lemma 1 $\alpha_{t+3} > \alpha_{t+1} > \bar{\alpha}$ and thus $(\alpha_{t+2},\alpha_{t+3})$ is in the first quadrant, origin excluded, which is case (a) of Proposition 1.

If $\alpha_{t+2} < \bar{\alpha}$, then two cases can occur, depending on the value of α_{t+3} :

If $\alpha_{t+3} < \bar{\alpha}$, by definition $(\alpha_{t+2},\alpha_{t+3})$ is in the third quadrant (origin excluded), which is case (c) of Proposition 1.

If $\alpha_{t+3} > \bar{\alpha}$, then $(\alpha_{t+2},\alpha_{t+3})$ belongs to the interior of the second quadrant. Recall that we had assumed to start with that $\alpha_t < \bar{\alpha}$ and $\alpha_{t+1} > \bar{\alpha}$. Applying lemma 1 twice, we obtain immediately :

$$\alpha_t < \alpha_{t+2} < \bar{\alpha}$$

$$\bar{\alpha} < \alpha_{t+3} < \alpha_{t+1}$$

which is case (b) of Proposition 1. Q.E.D.

Proposition 1 allows us to characterize in a very simple geometrical manner the trajectories of the α_t's when (α_1,α_2) is in the interior of the second quadrant. Consider indeed the sequence of pairs (α_1,α_2) ,..., $(\alpha_{2k+1},\alpha_{2k+2})$,... in the (α_t,α_{t+1}) plane. Proposition 1 says that, starting in the interior of the second quadrant, the corresponding points either move strictly towards the origin remaining in the second quadrant, or switch to the first or third quadrant (origin excluded) where by Theorem 2 they will diverge away from $(\bar{\alpha},\bar{\alpha})$. We should emphasize at this point that this characterization is a <u>global</u> one. So by Proposition 1 possible evolutions of the pairs

(α_1,α_2) ,..., $(\alpha_{2k+1},\alpha_{2k+2})$,..., will look as in Figure 4.

Figure 4

Note that a symmetric proposition, (and a similar Figure) is valid for the symmetric case where (α_1,α_2) belongs to the interior of the fourth quadrant.

Returning to the second quadrant case, we note that we have plotted in Figure 4 three types of evolutions, denoted by a , b and c . The geometry suggests a saddle-point property of the dynamics around the "equilibrium rational expectations" point, which means that a converging trajectory like b in Figure 4 is "exceptional". We shall make this intuition precise in the next theorem which characterizes the set S of initial values (α_1,α_2,q_o) such that the α_t's converge to $\bar{\alpha}$. Theorem 3 tells us that, at least locally, the dynamics look like a saddle-point dynamics, since the set S is of dimension 2 in a three dimensional space, which means that the system will "almost never" converge to rational expectations. Before stating and proving Theorem 3, let us define the set S rigorously and make a few remarks.

Definition. Let $S \subset R^3$ be defined as :

$$S = \{(\alpha_1,\alpha_2,q_o) \in R \times R \times]0, q^*[\mid \alpha_t \to \bar{\alpha} \text{ for } t \to \infty\}$$

We note for later reference that S contains the third axis below q^* , since for $\alpha_1 = \alpha_2 = \bar{\alpha}$ and $q_o < q^*$ the system is described by the "rational expectations dynamics" where q_t converges to zero.

Before stating the theorem we need to make one important observation.

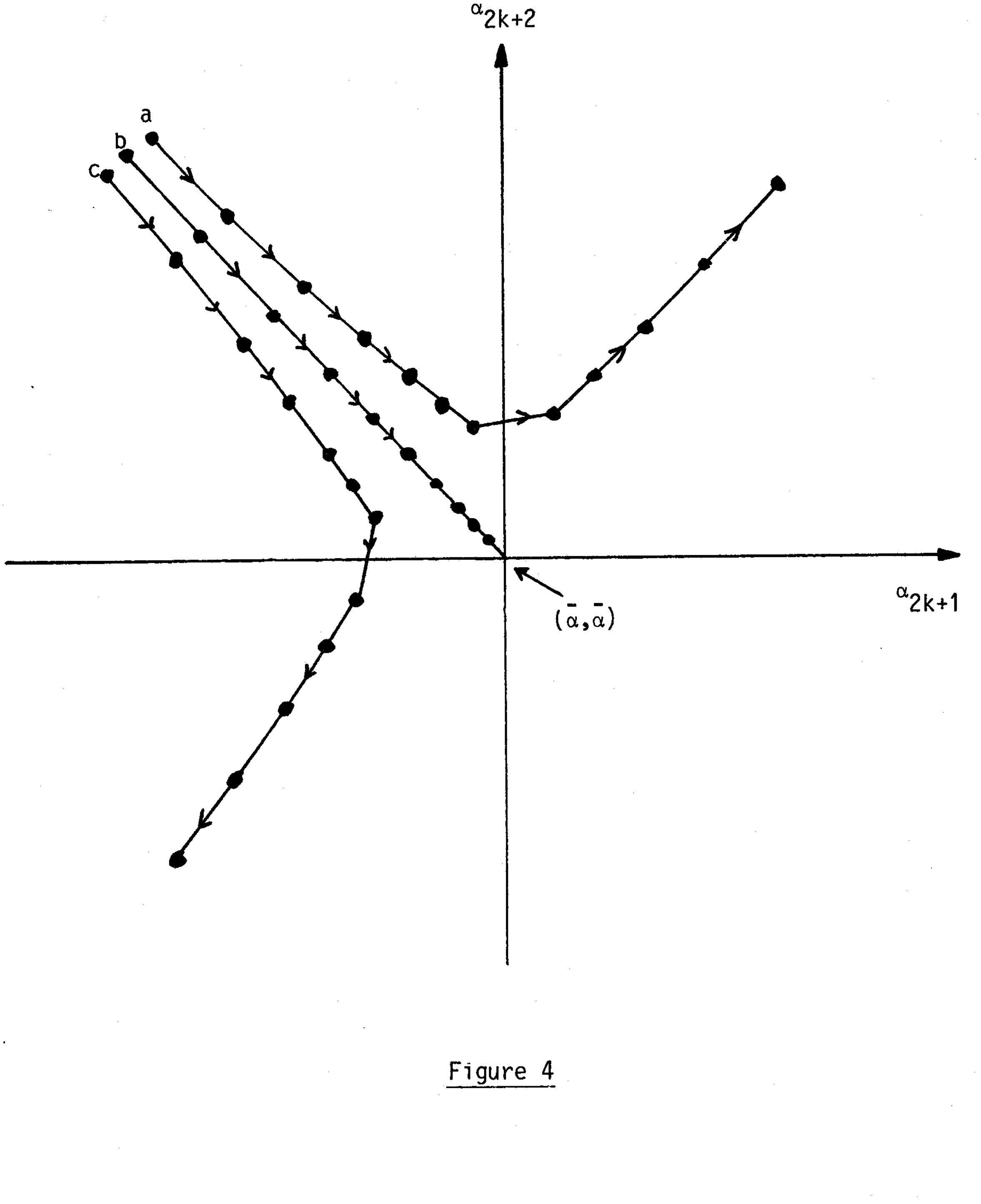

Figure 4

Consider a sequence $\langle \alpha_t, q_t \rangle$ such that $\alpha_t \to \bar{\alpha}$. Then for a sufficiently large t_o the path of q_t, $t > t_o$, will be "close" to a rational expectations path. But the only rational expectations paths are (1) stationary paths at the golden rule q^* (2) non-stationary paths between 0 and q^* converging towards 0. Since we know that q_t belongs to the open interval $]0,q^*[$, the rational expectations path must be of type (2) and so the triple $(\alpha_{t+1}, \alpha_t, q_{t-1})$ will converge towards $(\bar{\alpha},\bar{\alpha},0)$, which is a self-reproducing point. Theorem 3 will thus consider the behavior of the system around that point.

Theorem 3. There exists a neighborhood U of $(\bar{\alpha},\bar{\alpha},0)$ such that, in that neighborhood, S is topologically homeomorphic to a hyperplane of dimension 2. Moreover all $(\alpha_1, \alpha_2, q_o)$ in S fulfill :

$$(\alpha_1 - \bar{\alpha})\ (\alpha_2 - \bar{\alpha}) < 0 \quad \text{unless} \quad \alpha_1 = \bar{\alpha} \text{ and } \alpha_2 = \bar{\alpha}$$

Proof : By the observation just above Theorem 3, the set S is actually equal to $W^s(\bar{\alpha},\bar{\alpha},0)$, the stable manifold corresponding to the fixed point $(\bar{\alpha},\bar{\alpha},0)$. We shall now demonstrate that this set has the announced properties. The proof will proceed in a number of steps.

The first step is to characterize the linearized system around $(\bar{\alpha},\bar{\alpha},0)$ which is a self-reproducing point. We get the following :

$$\begin{bmatrix} \alpha_{t+2} - \bar{\alpha} \\ \alpha_{t+1} - \bar{\alpha} \\ q_t - 0 \end{bmatrix} = \begin{bmatrix} \varphi_1 & \varphi_o & \varphi_q \\ 1 & 0 & 0 \\ 0 & \phi'\psi_\alpha & \phi'\psi_q \end{bmatrix} \begin{bmatrix} \alpha_{t+1} - \bar{\alpha} \\ \alpha_t - \bar{\alpha} \\ q_{t-1} - 0 \end{bmatrix}$$

where :

$$\varphi_1 = \frac{\partial\varphi}{\partial\alpha_{t+1}} \quad \varphi_o = \frac{\partial\varphi}{\partial\alpha_t} \quad \varphi_q = \frac{\partial\varphi}{\partial q_{t-1}}$$

Moreover, at the point $(\bar{\alpha},\bar{\alpha},0)$:

$$\psi_q = \frac{1}{\phi'^2}, \quad \psi_\alpha = 0 \quad \text{(because of A3)}$$

$$\varphi_1 = \phi', \quad \varphi_o = \psi_q \phi'^2 = 1$$

So the linearized system around $(\bar{\alpha},\bar{\alpha},0)$ simplifies to :

$$\begin{bmatrix} \alpha_{t+2} - \bar{\alpha} \\ \alpha_{t+1} - \bar{\alpha} \\ q_t - 0 \end{bmatrix} = \begin{bmatrix} \phi' & 1 & \varphi_q \\ 1 & 0 & 0 \\ 0 & 0 & 1/\phi' \end{bmatrix} \begin{bmatrix} \alpha_{t+1} - \bar{\alpha} \\ \alpha_t - \bar{\alpha} \\ q_{t-1} - 0 \end{bmatrix}$$

In order to determine the corresponding eigenvalues, we consider the characteristic polynomial of the above matrix :

$$\pi(\lambda) = (\lambda - 1/\phi')(\lambda^2 - \phi'\lambda - 1)$$

The three roots of this polynomial are :

$$\lambda_1 = \frac{\phi' + \sqrt{\phi'^2+4}}{2} \qquad \lambda_2 = \frac{\phi' - \sqrt{\phi'^2+4}}{2} \qquad \lambda_3 = 1/\phi'$$

with the corresponding eigenvectors :

$$V_1 = \begin{pmatrix} \lambda_1 \\ 1 \\ 0 \end{pmatrix} \qquad V_2 = \begin{pmatrix} \lambda_2 \\ 1 \\ 0 \end{pmatrix} \qquad V_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Note that since $\phi'(0) > 1$, we have :

$$\lambda_1 > 1 \qquad -1 < \lambda_2 < 0 \qquad 0 < \lambda_3 < 1$$

So we have a saddle-point type dynamics around the point $(\bar{\alpha},\bar{\alpha},0)$. Since no eigenvalue has modulus one, the point $(\bar{\alpha},\bar{\alpha},0)$ is a hyperbolic fixed point of the map f defined by the given dynamical system : (*)

$$f : R \times R \times]0,q^*[\to R \times R \times]0,q^*[$$

$$f(\alpha_{t+1}, \alpha_t, q_{t-1}) = (\alpha_{t+2}, \alpha_{t+1}, q_t)$$

Further the stable manifold E^s of the linearized system around $(\bar{\alpha},\bar{\alpha},0)$, $Df(\bar{\alpha},\bar{\alpha},0)$, is the hyperplane spanned by the two eigenvectors V_2 and V_3 and is thus of dimension 2 .

In the second step of the proof we shall characterize the local stable manifold around $(\bar{\alpha},\bar{\alpha},0)$. Before that let us make notations a little more compact, writing the dynamics :

(*) Note that by nature the function f is defined only for strictly positive values of q . In order to apply the relevant mathematical theorems around $(\bar{\alpha},\bar{\alpha},0)$, f can be suitably extended in a neighborhood of that point by standard methods. The values of such an extended function for negative q's would be of course economically irrelevant.

$$x_t = f(x_{t-1})$$

with $$x_{t-1} = (\alpha_{t+1}, \alpha_t, q_{t-1})$$

Now the local stable manifold is defined for a neighborhood U of $\bar{x} = (\bar{\alpha},\bar{\alpha},0)$ as :

$$W_u^s(\bar{x}) = \{x \mid f^t(x) \to \bar{x} \text{ as } t \to \infty \text{ and } f^t(x) \in U \quad \forall\, t \geq 0\}$$

Note at this stage that $W_u^s(\bar{x})$ is a priori smaller than the intersection of U with $W^s(\bar{x}) = W^s(\bar{\alpha},\bar{\alpha},0)$, since this last set is defined by :

$$W^s(\bar{x}) = \{x \mid f^t(x) \to \bar{x} \text{ as } t \to \infty\}$$

Now we shall use the Hartman-Grobman theorem and the stable manifold theorem in its local version (Cf. for example Guckenheimer-Holmes [1983] theorems 1.4.1 and 1.4.2 p. 18) to characterize the local stable manifold $W_u^s(\bar{x})$.

The Stable Manifold Theorem says that for an hyperbolic fixed point $\bar{x} = (\bar{\alpha},\bar{\alpha},0)$ the local stable manifold $W_u^s(\bar{x})$ is tangent to E^s at $\bar{x}$ and has the same dimension, that is 2 in this case.

We can further characterize $W_u^s(\bar{x})$ through the Hartman-Grobman theorem which states that in a neighborhood U of the hyperbolic fixed point $\bar{x}$, the original dynamics described by f is topologically conjugate to that of the linearized system. More specifically there exists an homeomorphism h defined on U such that :

$$h \circ f = Df \circ h$$

where Df is the linearized system at $\bar{x} = (\bar{\alpha},\bar{\alpha},0)$. This implies that, at least locally, E^s and $W^s_u(\bar{x})$ are homeomorphic, $W^s_u(\bar{x})$ being the local image of E^s through h^{-1} .

The set $W^s_u(\bar{x})$ has therefore the properties stated in Theorem 3, and in step three we will show that in U the sets $W^s(\bar{x})$ and $W^s_u(\bar{x})$ are actually identical. For that we shall show that the trajectory of any point belonging to U will never return to this neighborhood once it has left it. To prove that, it suffices to show that points in U can leave it only by entering either the positive orthant in R^3 , or the orthant with the two α-axes negative and the q-axis positive, both cases implying divergence of the α's . Without loss of generality we may assume U to be a ball of centre $(\bar{\alpha},\bar{\alpha},0)$ and radius $\delta > 0$. Note first that since the q-axis' intersection with U belongs to the stable manifold $W^s(\bar{\alpha},\bar{\alpha},0)$, all q's in a (sufficiently small) neighborhood of this intersection will move towards 0 by the dynamics, i.e. the q-values decrease along trajectories in U . Thus :

$$(\alpha_t,\alpha_{t+1},q_{t-1}) \in U \implies q_t < q_{t-1}$$

Now consider a point such that (α_t,α_{t+1}) belongs to the second quadrant (the proof for the fourth quadrant is symmetrical). Then either $(\alpha_{t+1},\alpha_{t+2})$ belongs to the first quadrant, in which case Theorem 2 implies that the system will diverge without returning to U , or $(\alpha_{t+1},\alpha_{t+2})$ belongs to the fourth quadrant, in which case we shall now show that $(\alpha_{t+1},\alpha_{t+2},q_t) \in U$. Indeed from lemma 1 we obtain that :

$$\alpha_t < \alpha_{t+2} < \bar{\alpha}$$

and thus :

$$(\alpha_{t+1}-\bar{\alpha})^2 + (\alpha_{t+2}-\bar{\alpha})^2 + q_t^2 < (\alpha_t-\bar{\alpha})^2 + (\alpha_{t+1}-\bar{\alpha})^2 + q_{t-1}^2$$

which implies that $(\alpha_{t+1},\alpha_{t+2},q_t) \in U$ because U is a ball.

To finish the proof in step four we apply Theorem 2 : If $(\alpha_1,\alpha_2,q_o) \in S$, then $(\alpha_1-\bar{\alpha})(\alpha_2-\bar{\alpha}) < 0$ unless $\alpha_1 = \bar{\alpha}$ and $\alpha_2 = \bar{\alpha}$.

Q.E.D.

6. CONCLUSIONS

The above analysis stresses the need for questioning the theoretical foundation of the widely used hypothesis that agents behave as if they were endowed with rational expectations. Indeed, in the framework of an overlapping generations model with a simple learning procedure for expectations, we have shown that unless agents have rational expectations (that is, perfect foresight) at the outset, their expectations schemes will almost never converge towards rational expectations.

We used a one-parameter family of expectations functions, so the lack of convergence is neither caused by a situation with "too many" parameters to estimate, nor is it due to a misspecified family of expectations functions, as the family contains a rational expectations function. Moreover the demand-supply functions of the model are not postulated, as is often done in similar exercises, but they are derived from standard maximization behavior. Finally the basic model employed, the overlapping generations model with fiat money, is a widely used one. Thus the problem of non convergence to rational expectations cannot be considered as purely "pathological" or as a mere curiosum, and certainly deserves some further scrutiny.

Even though all the basic ingredients have been purposely kept highly simple, the intrinsic nonlinearities of the model give rise to a rather complicated three-dimensional dynamical system. In spite of these difficulties, this line of approach seems worth pursuing in various directions. First, if this proves to be technically feasible, one should investigate more complex expectations functions and learning procedures in such nonlinear models with rigorously derived demand and supply functions. Secondly, though in this paper we mainly focused on the behavior of the expectational variable α , the example of section 4 shows that this type of model can generate quite interesting dynamics in prices and quantities, yielding more possibilities than either the naïve or rational expectations assumptions. A more general study of such dynamics should be the subject of a forthcoming paper.

B I B L I O G R A P H Y

BALASKO Y. and SHELL K.(1981), "The Overlapping-Generations Model II. The case of Pure Exchange with Money", Journal of Economic Theory, 24, pp 112-142.

BLUME L.E. and EASLEY D.(1982),"Learning to be Rational", Journal of Economic Theory, 26, pp 340-351.

BRAY M. (1982), "Learning, Estimation and the Stability of Rational Expectations", Journal of Economic Theory, 26, pp 318-340.

BRAY M. and SAVIN N.E. (1984), "Rational Expectations Equilibria, Learning and Model Specification", University of Cambridge D.P.

BROCK W.A. (1972), "On Models of Expectations that arise from Maximizing Behavior of Economic Agents Over Time", Journal of Economic Theory, 5, pp 348-376.

CASS D., OKUNO M. and ZILCHA I. (1979), "The Role of Money in Supporting the Pareto Optimality of Competitive Equilibrium in Consumption-Loan Type Models", Journal of Economic Theory, 20, pp 41-80.

CASS D. and YAARI M. (1966), "A Reexamination of the Pure Consumption Loans Model", Journal of Political Economy, 74, pp 353-367.

De CANIO S.J. (1979), "Rational Expectations and Learning from Experience", Quarterly Journal of Economics, 92, pp 47-57.

DIAMOND P.A. (1965), "National Debt in a Neoclassical Growth Model", American Economic Review, 55, pp 1126-1150.

FOURGEAUD C., GOURIEROUX C. and PRADEL J. (1984), "Learning Procedures and Convergence to Rationality", CEPREMAP D.P., forthcoming in Econometrica.

FRIEDMAN B.M. (1979), "Optimal Expectations and the Extreme Information Assumptions of "Rational Expectations" Macro Models", Journal of Monetary Economics, 5, pp 23-41.

FUCHS G. (1977a), "Formation of Expectations : A Model in Temporary General Equilibrium Theory", Journal of Mathematical Economics, 4, pp 167-187.

FUCHS G. (1977b), "Dynamical Role and Evolution of Expectations", in Systèmes dynamiques et modèles économiques, Colloques internationaux du C.N.R.S., Vol. 259, Paris.

FUCHS G. (1979a), "Dynamics of Expectations in Temporary General Equilibrium Theory", Journal of Mathematical Economics, 6, pp 229-251.

FUCHS G. (1979b), "Is Error Learning Behaviour Stabilizing ?", Journal of Economic Theory, 20, pp 300-317.

GALE D. (1973), "Pure Exchange Equilibrium of Dynamic Economic Models", Journal of Economic Theory, 6, pp 12-36.

GOURIEROUX C., LAFFONT J.J. and MONFORT A. (1983), "Révision adaptative des anticipations et convergence vers les anticipations rationnelles", Revue d'Economie Appliquée, 36.

GRANDMONT J.M. and LAROQUE G. (1973), "Money in the Pure Consumption Loan Model", Journal of Economic Theory, 6, pp 382-395.

GUCKENHEIMER J. and HOLMES P. (1983) Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields, Springer-Verlag, New York.

SAMUELSON P.A. (1958), "An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money", Journal of Political Economy, 66, pp 467-482.