EXPECTATIONS FORMATION AND STABILITY

OF LARGE SOCIOECONOMIC SYSTEMS *

by

Jean-Michel GRANDMONT **

N° 9424

October 1994

* The material of this paper was in part presented as the Presidential address at the World Meetings of the Econometric Society, Barcelona, 1990. I owe a specific debt to Guy Laroque : the results reviewed here are in part an outgrowth of joint work undertaken in the eighties. I had stimulating exchanges with many colleagues while this project was "in progress", in particular with Philippe Aghion, Jean-Pascal Benassy, Alan Blinder, Christian Gourieroux, Bruno Jullien, Martine Quinzii and Jorgen Weibull.

** CNRS and CEPREMAP, 142, rue du Chevaleret, 75013 PARIS.

EXPECTATIONS FORMATION AND STABILITY

OF LARGE SOCIOECONOMIC SYSTEMS

ABSTRACT : Analysis of the local stability of self-fulfilling expectations in a simple model, when many agents try, in a decentralized fashion, to learn the dynamics of the system, suggests a sort of general "uncertainty principle". If agents are rather uncertain about the local stability of the system, so that they are ready on average to extrapolate a significant range of regularities (trends) out of small past deviations from equilibrium, the actual temporary equilibrium dynamics with learning should be locally unstable, especially if the influence of expectations on the motion of the system is strong. Local stability seems to occur only when expectations don't matter much, or when agents are assumed to ignore, either by lack of ability, or by conscious choice if they are fairly confident about the local stability of the system, all locally divergent tendencies that are present in past deviations from equilibrium.

Key words : Learning, self-fulfilling expectations, temporary equilibrium.

JEL classification numbers : E10, E32, D84

-=-=-=-=-=-=-=-=-=-=-

FORMATION DES ANTICIPATIONS ET STABILITE DE

GRANDS SYSTEMES SOCIOECONOMIQUES

RESUME : L'analyse de la stabilité locale des anticipations auto-réalisatrices dans un modèle simple où un grand nombre d'agents essaie, de manière décentralisée, d'apprendre la dynamique du système, suggère une sorte de "principe d'incertitude" général. Si les agents sont plutôt incertains quant à la stabilité locale du système, de telle sorte qu'ils sont prêts en moyenne à extrapoler un ensemble étendu de régularités (trends) à partir de petites déviations passées par rapport à l'équilibre, la dynamique avec apprentissage devrait être localement instable, en particulier lorsque l'influence des anticipations sur la dynamique du système est forte. On ne semble obtenir la stabilité locale que lorsque le poids des anticipations est faible, ou quand les agents sont supposés, soit par manque de sophistication, soit délibérément s'ils sont convaincus de la stabilité locale du système, ignorer systématiquement les tendances localement divergents présentes dans les déviations passées par rapport à l'équilibre.

Mots Clés : Apprentissage, anticipations auto-réalisatrices, équilibre temporaire.

JEL numéros de classification : E10, E32, D84

EXPECTATIONS FORMATION AND STABILITY OF LARGE SOCIOECONOMIC SYSTEMS *

Jean-Michel Grandmont **

"You regard men as infinitely selfish and infinitely farsighted. The first hypothesis may perhaps be admitted in a first approximation, the second may call for some reservations".

Henri Poincaré to Léon Walras
(letter of October 1, 1901) [1]

1. INTRODUCTION

In many formal models that are currently used in economics or game theory, it is often assumed that individual agents have been endowed initially with enough information, enough analytical and computing abilities, so that every agent's forecast about the future (e.g. a probability distribution) is correct at any moment, conditionally upon his information. The informational requirements underlying the axiom of self-fulfilling expectations are, of course, extraordinarily demanding, and many of us have voiced doubts over the years about its practical relevance. One line of defense not unfrequently heard of, is that fulfilled expectations should be interpreted as the outcome of a "fast" learning process that has already converged. Framing the issue in this way does make necessary to analyze convergence to, or divergence from self-fulfilling expectations, in dynamic models in which *decentralized* learning processes are explicitly taken into account. Such an analysis does allow to test at least the internal consistency of the axiom of self-fulfilling expectations (as opposed to its empirical validity). That is the question to which this paper is devoted.

I wish to say honestly at the outset that I do not have a satisfactory general theory at this stage. What I do have, however, is a set of examples displaying enough similarities to suggest a kind of general "*uncertainty principle*" : learning, when agents are somewhat uncertain about the dynamics of the economic or social system, is bound to generate local instability of self-fulfilling expectations, if the influence of expectations on the dynamics is significant. Specifically, consider a particular intertemporal equilibrium with self-fulfilling expectations, let us say Λ (in what follows, I shall consider only the simple case of a deterministic stationary state, but I think the argument is much more general). Assume that the agents don't know the model [2], and that they try to find out not only the *position* in the state space of the equilibrium Λ (they may even know exactly where it is) but also the dynamics nearby, i.e. the *speed* with which trajectories may converge to or diverge from Λ . There is an ocean of small individual agents, so there are no strategic considerations involved here. By assumption, there is no possibility for all traders or part of them, to act collectively so as to try to learn the dynamics of the system and at the same time to control its stability. It turns out that what is important for the local stability or instability of the system, is not so much the particular "story" underlying the agents' learning process, but rather the range of regularities (e.g. trends) they are in effect ready to extrapolate out of past deviations from the equilibrium Λ . Under the circumstances described here, one should expect the agents to be rather uncertain about the local stability of the system, and thus to be willing to extrapolate on average a wide range of such regularities. The findings reported in this paper suggest that in such a case, if expectations matter significantly, one should get local instability in the resulting dynamics with learning. This lack of stability make take different forms depending upon the sort of assumptions made on the learning processes. If these are regular enough and involve a bounded memory, local instability occurs with "probability one", i.e. following almost every initial small perturbation away from the equilibrium Λ under consideration. The conclusion does carry over to learning processes using past forecasts, as in "error learning" models, in which the agents' memory is in effect infinite. In general, i.e. even if learning processes display discontinuities (as it is the case with least squares learning models involving past deviations from equilibrium), and a memory that increases over time, local divergence seems to occur with "positive probability", i.e. for an open set of initial small perturbations, the size of which becomes significant as the influence of expectations on the system is strong. One gets similar conclusions with Bayesian learning: local instability occurs in that case whenever the agents' initial prior about the dynamics near the equilibrium Λ is significantly uncertain (has a large variance).

In short, *learning generates local instability of self-fulfilling expectations whenever agents are on average uncertain about the local dynamics of the system, and thus ready to extrapolate a wide range of regularities (trends) out of past deviations from equilibrium, and when the influence of expectations on the dynamics of the system is significant.* Local instability of this sort may indeed explain why markets in which expectations are thought to play a significant role, such as markets for financial assets, durable goods, capital or inventories, actually display more volatility than others. *A contrario*, under the assumption of (or if one had convergence to) self-fulfilling expectations, we would have reached the reverse, counterfactual conclusion : in that case, for a given process of shocks to the "fundamentals" (or to expectations), the more expectations matter, the more stable the local dynamics would be.

These results should make apparent when learning may generate locally stable dynamics. Either expectations don't matter much, in which case learning has little impact on the dynamics of the system. Or expectations matter significantly, but traders are able, by lack of sophistication or by conscious choice (if they are fairly sure of the stability of the system), to extrapolate only a *restricted range of regularities* (in particular, of trends) out of past deviations from equilibrium. Many theoretical studies claiming that learning leads to stability and convergence to self-fulfilling expectations often rest, explicitly or implicitly, upon *a priori* restrictions of this kind. This is the case for instance when agents are postulated to form expectations through regressions on *stationary* exogenous processes. Another prominent example is provided by the sometimes indiscriminate use of a socalled *"projection facility"* in recent least squares learning models (see e.g. Marcet and Sargent (1988, 1989 a,b)). *Such restrictions typically make agents systematically ignore locally divergent tendencies that are present in past deviations from the target equilibrium.* These restrictions appear to reintroduce through the back door (although to a lesser degree), the very kind of forecasts coordination that made the notion of self-fulfilling expectations vulnerable to criticism and that was supposed to be in fact *explained* by the theory.

The mere possibility that self-fulfilling expectations may be unstable when agents are significantly uncertain about the dynamics of the system, raises interesting and rather old issues about how economists or game theorists look at socioeconomic processes. I should emphasize that the instability we are talking about is only *local*, and that plausible global nonlinearities, *originating from the agents' expectations schemes themselves*, may keep the motion of the system bounded. Thus even in the absence of shocks to the "fundamentals" (as in real business cycles models or stochastic games), or to expectations (as in self-fulfilling expectations endogenous business cycles theories), learning by itself might generate convergence to complex nonlinear ("chaotic") attractors, hence self-perpetuating endogenous fluctuations, *along which forecasting errors would never vanish*. I have already alluded to the fact that such learning induced endogenous fluctuations would help to understand for instance why markets in which expectations are thought to matter most, seem to display more volatility than others, a property that is not easily explained under the assumption of self-fulfilling expectations. Such complex "learning equilibria" might also give us an alternative explanation of the well documented fact that human beings keep making sizeable *recurrent* mistakes when trying to predict the fate of socioeconomic systems in which they participate, a fact that is hard to reconcile with self-fulfilling expectations unless one introduces (unfortunately, easily removable) asymmetries of information among agents about exogenous macroeconomic shocks (Lucas (1972)). The ultimate test that this approach will have to pass, however, is that such "learning equilibria" must, to be acceptable, exhibit a reasonable degree of consistency with the agents' beliefs. In this respect, one might envision situations in which agents think that they are living in a world that is relatively simple, although subjected to random (e.g. white noise) shocks, but in which deterministic "learning equilibria" are complex ("chaotic") enough to make the agents' forecasting mistakes still "self-fulfilling" in a well defined sense. For instance, the complexity of "learning equilibria" might induce these agents, who would have at their disposal a relatively wide, but nevertheless limited, battery of statistical tests ("*bounded rationality*"), not to reject the hypothesis that their recurrent forecasting errors are indeed attributable to white noise disturbances. That is unlikely to be an easy test, and it is not quite clear to me at this stage whether this program can generate operational results or even be formulated in a consistent way. But progress on this front, if possible, might provide an interesting alternative to our current paradigms, which rely very heavily on extreme, and often criticized, rationality axioms.

The conceptual framework of *sequential temporary equilibrium* will be used throughout this paper. This is not coincidental : the issue I am addressing here has been on the formal agenda of that research program since the early seventies, as examplified by the pioneering works of Fuchs and Laroque (1976), and of Fuchs (1976, 1977a,b, 1979a,b), within the framework of deterministic two-period general equilibrium overlapping generations models involving several goods, money and heterogenous consumers. Research on the dynamic stability of economic systems under the assumption of "adaptive" expectations had of course been very active in the fifties and sixties (see e.g. Cagan (1956), Nerlove (1958), Mincer (1969)). The literature on the subject has been in fact rapidly growing in the last decade. We now have quite a number of theoretical works analyzing the stability of self-fulfilling expectations equilibria (stationary states, deterministic cycles, finite Markov sunspot equilibria) in simple overlapping generations models (see for instance Benassy and Blad (1989), Bullard (1994), Grandmont (1985), Guesnerie and Woodford (1991), Guesnerie (1993), Marcet and Sargent (1989c), Tillman (1983, 1985), Woodford (1990)). There are also many studies of convergence to self-fulfilling expectations in stochastic linear or nonlinear macroeconomic models (for instance Bray (1982, 1983) , Bray and Savin (1986), Calvo (1988), Champsaur (1983), Chatterji (1994), Cyert and de Groot (1974), De Canio (1979), Evans (1983, 1985, 1989), Evans and Honkapohja (1992, 1993, 1994), Fourgeaud, Gourieroux and Pradel (1986), Frydman (1982), Gottfries (1985), Gourieroux, Laffont and Monfort (1983), Howitt (1990), Lucas (1986), Marcet and Sargent (1988, 1989 a,b), Vives (1993), Wickens (1982)). The issue has not been ignored by game theorists either : after early studies by Shapley (1964) and Miyasawa (1961) suggesting instability, researchers have been investigating convergence to particular equilibria in repeated games, when players revise adaptively their expectations and/or their strategies over time (e.g. Canning (1992), Crawford (1974, 1985, 1989), Fudenberg and Kreps (1988), Fudenberg and Levine (1993 a,b), Harsanyi and Selten (1988), Jordan (1985, 1991, 1992), Kalai and Lehrer (1993 a,b), Kandori, Mailath and Rob (1993), Milgrom and Roberts (1991), Nyarko (1991, 1992), Selten (1988)). Evolutionary games, on the other hand, can be viewed as dynamic games in which learning operates adaptively at the level of a whole population, through "Darwinian" selection, rather than at the individual level (for a neat review, see Weibull (1994)).

This paper is not intended to be a survey and I apologize in advance to the authors who may not find their work thoroughly discussed here. My goal is rather, by studying a simple, context free model, stripped of unnecessary technicalities, to try to go some way toward a better understanding of why one gets convergence in some cases, whereas one does not in others.

The paper is organized as follows. The central features of the analysis are presented in Section 2, when learning processes are smooth. The analysis is extended to error learning in Section 3. Learning involving least squares regressions on lagged deviations from equilibrium is taken up in Section 4, while Bayesian learning is considered in Section 5. A few concluding remarks are gathered in the final section.

2. SMOOTH LEARNING PROCESSES

I would like to be able to understand what happens in nonstochastic environments before taking into account the complications arising from randomness. I shall consider accordingly a deterministic (nonlinear) framework. Since the analysis will be local, the results should not change qualitatively if small white noise disturbances were added to the equations defining the system. The analogy I have in mind is that stability of a stochastic ARMA process, for instance, is entirely governed by the dynamics of its deterministic part.

Time is discrete. I shall focus on an extremely simple case, in which the state of the system is completely described at every date by a single real number x_t . [3] Depending upon the context, the state variable x may stand for a price, a rate of inflation, a real rate of interest, output or whatever you like. Traders plan one period ahead. Since I wish to abstract from all forms of uncertainty, the traders' forecasts are taken to be subjectively certain. We could assume equivalently that expectations are represented by probability distributions and that traders are risk neutral : only their mean forecasts matter in that case. I suppose that there are many traders, each of whom has a negligible influence on the whole system, in order to exclude any strategic considerations. To simplify notations, all traders forecasts at date t about the future state are assumed to be identical, the common forecast being noted ${}_t x^e_{t+1}$. The methods described below can nevertheless be made to bear upon the case of heterogenous beliefs. The forecast ${}_t x^e_{t+1}$ can then be reinterpreted as an *average* forecast, each individual forecast being weighted by its relative local contribution to the dynamics of the system (Remark 2.3).

The current equilibrium state x_t depends on the common (or average) forecast ${}_t x^e_{t+1}$ and on the past state x_{t-1} through the *temporary equilibrium relation*

(2.1) $$T(x_{t-1} , x_t , {}_t x^e_{t+1}) = 0 .$$

The analysis will be local, i.e. near a stationary state $x_t = \bar{x}$ defined by $T(\bar{x} , \bar{x} , \bar{x}) = 0$. The temporary equilibrium map T is supposed throughout to be well defined and *continuously differentiable* when its arguments are near $\bar{x}$. I shall note b_1 , b_0 and a the partial derivatives of T with respect to x_{t-1} , x_t and ${}_t x^e_{t+1}$, evaluated at the stationary state, and I shall assume that *expectations matter*, i.e. a $\neq$ 0 (otherwise the issue I want to look at would vanish). Equation (2.1) describes the "structural" characteristics of the system. It would arise for instance as a market clearing condition in simple competitive overlapping generations models, or in infinite horizon models with cash-in-advance constraints. The formulation is general enough to encompass imperfect competition or cases in which traders do not necessarily optimize and use for example "satisficing" strategies or other sorts of "adaptive" behavior.

In order to close the model, one must specify how traders forecast the future at any date, as a function of their information on the workings of the system and on past history. I shall deal primarily with the case in which traders base their forecasts on their observations of the current state x_t and finitely many past states $x_{t-1},\ldots,x_{t-L}$ (the case of error learning, in which traders also take into account past forecasts, making in effect the memory infinite, is considered in the next section, with essentially similar results). How they extract information from these data (given their knowledge of the overall system) can then be summarized by the forecasting rule, or *expectation function*

$$(2.2) \qquad {}_t x^e_{t+1} = \psi(x_t , \dots , x_{t-L}) .$$

Let us assume either that traders know where the stationary state $\bar{x}$ lies or that, if they don't know it, they are prepared to extrapolate constant sequences, i.e. $\psi(x,\dots,x) = x$ for x near $\bar{x}$. In all cases, one has $\psi(\bar{x},\dots,\bar{x}) = \bar{x}$. I shall assume moreover, here and in the next section, that the traders' expectation formation is regular enough, so that ψ is well defined and *continuously differentiable* in the immediate vicinity of the stationary state. In what follows, c_j will stand for the partial derivative of ψ with respect to x_{t-j} , evaluated at $\bar{x}$, for $j = 0,1,\dots,L$.

Before commenting on the mental processes that may lie behind (2.2), let us specify how the resulting *actual temporary equilibrium dynamics* is defined. Replacing the forecast in (2.1) by its expression (2.2) yields

$$(2.3) \qquad T(x_{t-1} , x_t , \psi(x_t , \dots , x_{t-L})) = 0 ,$$

which determines implicitly the current state x_t , given past history. Indeed, $x_t = \bar{x}$ for all t is a stationary solution of that equation. Let us assume now that the partial derivative of the left hand side of (2.3) with respect to x_t , evaluated at the stationary state, does not vanish, i.e. $b_0 + ac_0 \neq 0$. Then from the implicit function theorem, (2.3) can be solved uniquely in x_t near $\bar{x}$, which yields a local difference equation of the form

$$(2.4) \qquad x_t = W_{loc}(x_{t-1} , \dots , x_{t-L}) .$$

The issue is to analyze its stability, in relation to the local properties of the temporary equilibrium map T and of the expectation function ψ .

If all traders knew the characteristics of the system, i.e. equation (2.1), *and* if there was a central mechanism permitting to coordinate all forecasts, the problem I want to study would go away. All present and future traders could commit themselves to a common "correct" expectation function, which would then appear in (2.2). The resulting temporary equilibrium dynamics (2.4) or (2.3) would generate in that case perfect foresight paths (Grandmont and Laroque (1991, Section 1.1)).

When no such information is available, and/or no central coordination of all current and future traders' forecasting behaviours is possible, individual traders will have to try, in a *decentralized* fashion, to learn the dynamic laws of their environment by extracting possible regularities (e.g. trends, cyclical components) from past data. This learning process is summarized by the expectation function ψ . The formulation is general enough to cover the situation in which traders are endowed with a set of *a priori* beliefs about the dynamics of the system, indexed for instance by a vector of unknown parameters. They would update at date t their estimates of these parameters in the light of the observations $(x_t,\dots,x_{t-L})$ and use their updated estimates to establish their forecast. The resulting forecasting rule would then be of the form (2.2). If the estimation procedure involves regressions on exogenous variables (instrumental variables, sunspots), one can in fact generate in this way continuously differentiable (even linear) expectation functions (Grandmont and Laroque (1991, Section 1.2)). *The specification is thus compatible with changing beliefs of the agents, as long as the way in which they modify their views of the world on the basis of past observations does not itself vary over time.* The assumption of a bounded memory is made here to generate a simple time-independent dynamical system. When L is large, it should not yield too distorted a view of the case in which observations in the far distant past have a vanishing influence on current forecasts. We shall study in Section 4 an example in which the traders' memory may increase without bound over time. As noted earlier, the next section will be devoted to the case in which current expectations are based on past forecasts as well, making in effect the memory infinite. In all cases, one gets essentially similar results.

Local stability of the actual dynamics with learning is generically governed by the eigenvalues of the characteristic polynomial obtained by linearizing (2.4), or (2.3), near the stationary state $\bar{x}$

$$(2.5) \qquad Q_W(z) \equiv b_1\, z^{L-1} + b_0\, z^L + a\, \Sigma_0^L\, c_j\, z^{L-j} = 0 \ .$$

If all roots of (2.5) have modulus less than 1, the stationary state is locally stable. But as soon as one root of Q_W has modulus greater than 1, the stationary state is unstable : for all initial conditions $(x_{t-1},\dots,x_{t-L})$ near $\bar{x}$, except perhaps for a very "thin" set (of Lebesgue measure 0), the trajectories generated by the dynamics with learning (2.4) are pulled away from the stationary state.

We got the equation (2.3) defining the actual dynamics with learning by putting together (2.1) and (2.2). One should accordingly expect that all the information needed to assess local stability is embodied in the characteristic polynomials corresponding to these two equations. Linearizing (2.1) near the stationary state yields the polynomial

$$(2.6) \qquad Q_F(z) \equiv b_1 + b_0 z + a z^2 = 0 ,$$

whereas the same operation applied to the forecasting rule (2.2) gives rise to

$$(2.7) \qquad Q_\psi(z) \equiv z^{L+1} - \Sigma_0^L c_j z^{L-j} = 0 .$$

The polynomial Q_F is easy to interpret if we remark that trajectories with perfect foresight obey the twodimensional difference equation obtained by replacing the forecast ${}_t x^e_{t+1}$ in (2.1) by the actual value x_{t+1} . The two *perfect foresight roots* λ_1 , λ_2 of Q_F characterize, therefore, the local behavior of the perfect foresight dynamics near the stationary state.

The interpretation of the roots of Q_ψ is somewhat different. If traders are prepared to extrapolate constant sequences, i.e. $x \equiv \psi(x,\dots,x)$ for all x near $\bar{x}$, one gets by differentiation $1 - \Sigma_j c_j = 0$, which means that $z = 1$ is a solution of (2.7). Similarly, if traders extrapolate the particular trend r out of past deviations from the stationary state, i.e. $r^{L+1}(x - \bar{x}) + \bar{x} \equiv \psi(r^L(x - \bar{x}) + \bar{x},\dots,r(x - \bar{x}) + \bar{x},\ x)$, at least near $\bar{x}$, one gets again by differentiation that $z = r$ is solution of (2.7). If the traders extrapolate sequences of past deviations $x_{t-j} - \bar{x}$ that have period k , all k-th roots of unity are solutions of (2.7) (Grandmont and Laroque, 1986, 1990). More generally, the fact that $\mu = re^{i\theta}$ is a root of Q_ψ means that traders do extrapolate the trend r and the cyclical component corresponding to θ out of past deviations $x_{t-j} - \bar{x}$.

The L+1 roots $\mu_1, \ldots, \mu_{L+1}$ of the polynomial Q_ψ (the *local eigenvalues of the expectation function*) characterize the set of local regularities (trends, cyclical components) that the traders are, on average, ready to extrapolate out of small, past deviations from equilibrium. We reach a simple but important methodological conclusion, that will be a recurrent theme of this paper. *This set of local regularities is the only thing that matters in order to assess local stability of the actual dynamics with learning.* Indeed, one can rewrite the characteristic polynomial Q_W corresponding to the actual dynamics with learning by adding the term az^{L+1} to, and substracting it from, the expression (2.5). This yields

$$(2.8) \qquad Q_W(z) \equiv z^{L-1}\, Q_F(z) - aQ_\psi(z),$$

which is equivalent to

$$(2.9) \qquad Q_W(z) \equiv a[z^{L-1}(z - \lambda_1)(z - \lambda_2) - \prod_k (z - \mu_k)].$$

Local stability of the actual dynamics with learning depends only on how the local eigenvalues of the traders' average forecasting rule interact with the local characteristic roots of the perfect foresight dynamics.

Instability

Here comes the second main theme. The set of eigenvalues μ_k of the forecasting rule ψ describes in a sense the traders' beliefs about the local regularities they may have to face. If the memory L is large, for instance, the μ_k's can be distributed in the complex plane so as to approximate closely any probability distribution with a continuous density, that we may interpret as the traders' "prior". The following set of results (which is adapted from Grandmont-Laroque (1991)) implies that *if the agents are rather uncertain about the local dynamics of the system, and are thus ready, on average, to extrapolate a wide range of local regularities (trends) out of past deviations from the equilibrium, and if the influence of expectations on the dynamics of the system is significant, the actual dynamics with learning should be locally unstable.*

PROPOSITION 2.1. *Assume* $a \neq 0$ *(expectations matter) and* $b_0 + a\, c_0 \neq 0$ *(the actual dynamics with learning is locally well defined). Let the expectation function* ψ *have two local real eigenvalues* $\mu_1 < 0 < \mu_2$ *that differ from the perfect foresight roots* λ_1 *and* λ_2 .

1. *Let* n_F , n_W *be the number of real roots of the polynomials* Q_F , Q_W *that lie outside the interval* $[\mu_1 , \mu_2]$. *Then* n_W *is odd if and only if* n_F *is even (i.e. 0 or 2).*

This implies in particular

2. *If* $\mu_1 \leqq -1$, $\mu_2 \geqq 1$ *and if the interval* $[\mu_1 , \mu_2]$ *contains in its interior all perfect foresight characteristic roots that are real, the polynomial* Q_W *has a real root* r *that satisfies either* $r < \mu_1 \leqq -1$ *or* $r \geqq \mu_2 > 1$.

Proof. We shall use the following simple observation.

(2.10) Let $P(z)$ be a polynomial of degree n with real coefficients, the coefficient of the term of degree n being $\alpha \neq 0$. Let r be a real number with $P(r) \neq 0$. Then P has an odd number of real roots greater than r if and only if $\alpha\, P(r) < 0$, while P has an odd number of real roots less than r if and only if $(-1)^n \alpha\, P(r) < 0$. If r , s are real numbers with $r \leqq s$, $P(r)\, P(s) \neq 0$, then P has an odd number of real roots outside the interval $[r , s]$ if and only if $(-1)^n P(r)\, P(s) < 0$.

The proof is immediate if one remarks that $\alpha\, P(r) \neq 0$ has the same sign as $\prod_k (r - \rho_k)$, where the product runs over all real roots ρ_k of P . The sign of $\alpha\, P(r)$ is thus $(-1)^{n_r}$, where n_r is the number of real roots $\rho_k > r$. The sign of $(-1)^n \alpha\, P(r)$ is $(-1)^{m_r}$, where m_r is the number of real roots $\rho_k < r$. The sign of $(-1)^n P(r)\, P(s) \neq 0$ is $(-1)^{m_r + n_s}$.

To prove Part 1 of the Proposition, remark that Q_W has degree L , and that in view of (2.8), the expressions

$$A = (-1)^L\, Q_W(\mu_1)\, Q_W(\mu_2) \neq 0 \quad , \quad B = Q_F(\mu_1)\, Q_F(\mu_2) \neq 0 \, ,$$

have opposite signs. From (2.10) the sign of A is $(-1)^{n_W}$, while the sign of B is $(-1)^{n_F}$. Thus n_W is odd if and only if n_F is even.

Part 2 of the Proposition corresponds to the case $\mu_1 \leqq -1$, $\mu_2 \geqq 1$ and $n_F = 0$, in which case n_W is odd. Q.E.D.

As shown by the first part of the Proposition, the approach taken here involves a parity argument. Let the expectation function ψ have two real eigenvalues μ_1 , μ_2 , with 0 in between, that differ from the perfect foresight roots. Note that nothing prevents λ_1 and/or λ_2 to appear among the eigenvalues of ψ . The assumption simply means that traders stand ready to extrapolate at least two other (real) trends if they happen to observe them in past deviations from the stationary state. Now you ask yourself : how many real roots the polynomials Q_F and Q_W have outside the interval $[\mu_1,\mu_2]$? The Proposition tells you that these numbers change parity when you go from the perfect foresight dynamics to the actual dynamics with learning [4] . In particular, if the number of real perfect foresight roots outside $[\mu_1,\mu_2]$ is 0, and if $\mu_1 \leqq -1$, $\mu_2 \geqq 1$, the actual dynamics with learning must be locally unstable.

In the context I described, in which no coordination of forecasts is possible, one should expect the agents to be rather uncertain about the local dynamics of the system, so that the least and largest real eigenvalues of their (average) expectation function satisfy $\mu_1 \leqq -1$, $\mu_2 \geqq 1$, and span a large interval. The conditions of the second part of the Proposition will then be satisfied provided that the modulus of the two perfect foresight roots is not too great. Since their sum and their product is given by $\lambda_1 + \lambda_2 = - b_0 a^{-1}$ and $\lambda_1 \lambda_2 = b_1 a^{-1}$, this will occur if $|a|$ is not too small relatively to $|b_0|$ and $|b_1|$. Thus if the agents are ready to extrapolate a relatively wide range of local trends out of past deviations from the stationary state, and if the influence of expectations on the dynamics of the system is significant, one will get local instability. [5] The actual dynamics will in fact amplify the divergent tendencies that may be present in the agents' forecasting rule, since Q_W will have then a real characteristic root r that lies outside the interval $[\mu_1,\mu_2]$. [6]

Instability of this sort may explain why markets for financial assets, durable goods, capital or inventories, in which the role of expectations is thought to be important, display more volatility than others. It should be noted that we would have reached the opposite, counterfactual conclusion under the assumption of self-fulfilling expectations : the modulus of the two perfect foresight roots λ_1 and λ_2 is bound to be small when expectations matter a lot, i.e. when the parameter $|a|$ is large.

Stability

I now come to the third theme of the paper, which is a simple corollary of what precedes. *One can produce stability results in the present framework, but they will often involve systems in which the influence of expectations is weak, or strong restrictions on the range of regularities traders stand ready to extrapolate out of past deviations from the stationary state.* In the first case, learning has little impact on the dynamics of the system. On the other hand, imposing strong restrictions on the distribution of local eigenvalues of the traders' forecasting rule will achieve stability essentially by making agents systematically ignore, either by lack of ability or by conscious choice if they are fairly convinced of the stability of the system, locally divergent tendencies that show up in past deviations from equilibrium. A forerunner of the corresponding stability result can be found Fuchs and Laroque (1976). The following version is adapted from Grandmont (1985), Grandmont and Laroque (1986, 1990).

PROPOSITION 2.2. *Assume* $b_0 + ac_0 \neq 0$ *and let* α *be the maximum of* $|\Sigma_0^L c_j z^{L-j}|$ *when* $|z| = 1$ *. Then*

1. *The stationary state is locally stable in the actual dynamics with learning if* $|b_0| > |b_1| + |a|\,\alpha$.

In particular,

2. *Assume that traders extrapolate constant sequences, i.e.* $\psi(x,\dots,x) = x$ *for all* x *near* $\bar{x}$ *, and that* $c_j \geqq 0$ *for all* $j = 0,\dots,L$ *. Then* $\alpha = 1$ *, and the stationary state is locally stable in the actual dynamics with learning if* $|b_0| > |b_1| + |a|$.

Proof. This result is essentially a restatement of Grandmont and Laroque (1990, Section 2, Lemma and ensuing discussion). For the sake of completeness, I give here an alternative short proof.

Assume that Q_w has a root $|\bar{z}| \geqq 1$. In view of (2.5), this implies

$$|b_0| = |b_1\, \bar{z}^{-1} + a\, \Sigma_0^L c_j\, \bar{z}^{-j}| ,$$

in which case $|b_0| \leqq |b_1| + |a|\,\beta$, where β is the maximum of $|\Sigma_0^L c_j\, z^j|$ for

$|z| \leqq 1$. This maximum is bound to occur for $|z| = 1$, so β is also the maximum of $|\Sigma_0^L c_j z^{L-j}|$ when $|z| = 1$, i.e. $\beta = \alpha$. Thus if $|b_0| > |b_1| + |a| \alpha$, all roots of Q_w have modulus less than 1.

To prove part 2, remark that if the expectation function extrapolates constant sequences near $\bar{x}$, one has $\Sigma_0^L c_j = 1$, which implies $\alpha \geqq 1$. If in addition $c_j \geqq 0$ for all j , $|\Sigma_0^L c_j z^{L-j}| \leqq 1$ when $|z| = 1$ and thus $\alpha = 1$.

Q.E.D.

The first part of the Proposition presents a stability condition involving the parameter α, which measures in a sense the stability of the local eigenvalues of the forecasting rule and the width of their distribution. Indeed, let us remark that for every $|z| = 1$

$$\Sigma_0^L|c_j| \geqq \alpha \geqq |\Sigma_0^L c_j z^{L-j}| = |Q_\psi(z) - z^{L+1}| \geqq |Q_\psi(z)| - 1 .$$

So, if the distribution of the local eigenvalues of the expectation function has a wide support, the maximum of $|Q_\psi(z)| = \prod_k |z-\mu_k|$ for $|z| = 1$, hence α, will be large. On the other hand, if all local eigenvalues of the expectation function have a small modulus, the coefficients c_j, and therefore α, will also be small.

The stability condition $|b_0| > |b_1| + |a| \alpha$ cannot be satisfied unless the influence of the current state variable in the temporary equilibrium relation (2.1) is larger than the influence of the past. If we assume $|b_0| > |b_1|$, the actual dynamics with learning will be locally convergent when the product $|a| \alpha$ is small. This will occur in particular if expectations don't matter much ($|a|$ is small), or if the local eigenvalues of the agents' expectation function are close enough to 0 (the coefficient α is then small). [7]

The second part of the Proposition presents an example in which stability is indeed achieved along this line. One has $\Sigma_0^L c_j = 1$ when the expectation function extrapolates constant sequences. If in addition $c_j \geqq 0$ for all j, all local eigenvalues of the forecasting rule have modulus less than or equal to 1, as the reader will easily verify. All present and future traders are unable to, or have decided in advance not to, extrapolate any locally divergent trend out of past deviations from equilibrium.

The foregoing specification would arise if, for instance, traders formed their forecasts by regressing the state variable on a constant, which would yield ${}_t x^e_{t+1} = (x_t + \ldots + x_{t-L})/(L + 1)$. Formulations of this sort are common in the literature. Traders are often assumed to try to estimate a relation bewteen the endogenous variable x_t and a prespecified set of exogenous (instrumental, or sunspot) variables that are generated by stable processes (in stochastic settings, for instance, the exogenous variables are usually assumed to be covariance stationary). Again *such procedures generate forecasting rules that have all their local eigenvalues inside or on the unit circle in the complex plane.* [8] Assumptions of this sort (and the corresponding stability results) involve strong and unlikely restrictions on the range of regularities that agents are ready to extrapolate out of past deviations from equilibrium.

Remark 2.3 (Heterogenous beliefs). I now sketch briefly how the methods described in this section can be made to bear upon the case of heterogenous beliefs. Assume that there are m traders indexed by $i = 1,\ldots,m$, with m "large". The temporary equilibrium relation becomes

$$(2.11) \qquad T(x_{t-1}, x_t, {}_t x^e_{1,t+1}, \ldots, {}_t x^e_{m,t+1}) = 0 \ .$$

As before, T is continuously differentiable near the stationary state defined by $T(\bar{x},\ldots,\bar{x}) = 0$. Let b_1 , b_0 , a_i stand for the partial derivatives of T with respect x_{t-1} , x_t and ${}_t x^e_{i,t+1}$, evaluated at the stationary state. Expectations are supposed to matter in the aggregate, i.e. $a = \Sigma_i a_i \neq 0$, and $\alpha_i = a_i/a$ will denote the relative local contribution of the ith trader's forecast to the temporary equilibrium map. Associated to (2.11) is the perfect foresight polynomial Q_F which is defined as before by the expression (2.6).

Each trader's learning process is summarized, as in the text, by a continuously differentiable expectation function

$$(2.12) \qquad {}_t x^e_{i,t+1} = \psi_i(x_t,\ldots,x_{t-L}) \ , \qquad i = 1,\ldots,m \ ,$$

satisfying $\psi_i(\bar{x},\ldots,\bar{x}) = \bar{x}$. The partial derivative of ψ_i with respect to

x_{t-j} , evaluated at the stationary state, is noted c_{ij} . As a matter of fact, it is convenient to introduce an *average expectation function*, $\psi = \Sigma_i \alpha_i \psi_i$. Its partial derivative with respect to x_{t-j} , evaluated at the stationary state, is then $c_j = \Sigma_i \alpha_i c_{ij}$. The characteristic polynomial Q_ψ associated to the average forecasting rule is given by (2.7). Its roots describe the trends and cyclical regularities that, *on average*, traders are ready to extrapolate out of past local deviations from the stationary state.

Replacing each forecast in (2.11) by its expression (2.12) defines implicitly x_t as a function of past states. The resulting actual dynamics with learning is locally well defined if $b_0 + a c_0 \neq 0$. Simple calculations show that the corresponding characteristic polynomial Q_W is given as before by the expression (2.5). All results of the text apply accordingly if ψ is interpreted as the average expectation function $\Sigma_i \alpha_i \psi_i$. Note that taking into account the possible heterogeneity of beliefs opens new avenues that deserve further analysis. For instance, although every individual forecasting rule may involve notably divergent eigenvalues, these may "cancel out" in the aggregate, e.g. all $c_j = \Sigma_i c_{ij}$ may be small, and the actual dynamics with learning may thus be locally stable, if there is sufficient heterogeneity and symmetry (I am grateful to Alan Blinder for drawing my attention to this point). On the other hand, only a small group of traders with expectation functions involving significantly divergent local eigenvalues can in fact destabilize the whole system.

3. ERROR LEARNING

The smooth forecasting rule considered in the preceding section depended only on past states. It can be rationalized for instance as the outcome of regressions of the state variable on a prespecified set of exogenous (instrumental, sunspot) variables. I emphasized that beyond the specific story one may tell about how traders learn, local stability of the system rests on the local regularities traders end up extrapolating from past history. To illustrate the scope of the approach, I make now the forecasting rule depend also on previous forecasts. Thus the formulation covers the situation in which traders learn from forecasting errors made in the past [9] . Although the traders' mental processes may appear quite different at first sight, the application of the methods described in Section 2 will lead, as we are going to see shortly, to essentially similar results.

The structural behavior of the system is described as before by the continuously differentiable temporary equilibrium relation

$$(3.1) \qquad T(x_{t-1}\ ,\ x_t\ ,\ {}_t x^e_{t+1}) = 0\ .$$

The analysis is local, i.e. near a stationary state $\bar{x}$. With the same notation as before, the corresponding perfect foresight characteristic polynomial Q_F is given by (2.6).

The traders' (common) expectation function depends now on past forecasts ${}_{t-j}x^e_{t-j+1}$ and is noted

$$(3.2) \qquad {}_t x^e_{t+1} = \psi^*\begin{pmatrix} x_t\ ,\ x_{t-1}\ ,\dots,\ x_{t-L} \\ {}_{t-1}x^e_t\ ,\dots,\ {}_{t-L}x^e_{t-L+1} \end{pmatrix}$$

The value of ψ^* when all its arguments are equal to $\bar{x}$, is assumed again to be $\bar{x}$. I also suppose that ψ^* is continuously differentiable near the stationary state. In the sequel, c^*_j will stand for its partial derivative with respect to x_{t-j} , while γ^*_j will denote the partial derivative with respect to ${}_{t-j}x^e_{t-j+1}$, evaluated at the stationary state.

The actual dynamics with learning is defined by the two relations (3.1) and (3.2), which determine implicitly $(x_t, {}_t x^e_{t+1})$ as a function of $(x_{t-j}, {}_{t-j}x^e_{t-j+1})$ for $j = 1,\dots,L$. I shall assume that the (2x2) matrix of partial derivatives of (3.1), (3.2) with respect to $(x_t, {}_t x^e_{t+1})$, evaluated at the stationary state, is invertible — this is easily seen to be equivalent to $b_0 + a\, c^*_0 \neq 0$. By the implicit function theorem, the actual dynamics with learning is then locally well defined.

Stability or instability of the stationary state in this dynamics is governed as usual by the roots of the corresponding characteristic polynomial Q_W . The standard procedure to get this polynomial is to linearize (3.1), (3.2), to look for solutions of the linearized system that grow at a (possibly complex) gross rate z , and to equate the appropriate determinant to 0 . The reader will check without difficulty that the outcome of this procedure is a polynomial (of degree 2L) given by

$$(3.3) \qquad Q_W(z) \equiv (b_1 z^{L-1} + b_0 z^L)(z^L - \Sigma_1^L \gamma_j^* z^{L-j}) - az^L \Sigma_0^L c_j^* z^{L-j} = 0 .$$

This looks at first sight rather complicated. According to the principles outlined in Section 2, however, we should be able to get a more transparent picture by carefully analyzing the local characteristics of the expectation function.

The first thing to spell out is the set of local regularities that traders are able to extrapolate *correctly* from past deviations from equilibrium. To define this set properly, we have to look at what happens when there were indeed no forecasting errors in the past, and thus to consider the (fictitious) expectation function $\psi(x_t,\ldots,x_{t-L})$ obtained by setting in ψ^* every previous forecast ${}_{t-j}x^e_{t-j+1}$ equal to x_{t-j+1} for $j = 1,\ldots,L$. The local regularities traders can correctly extrapolate out of past deviations $x_{t-j} - \bar{x}$ from the stationary state are then summarized, as in Section 2, by the L+1 roots μ_1 ,..., μ_{L+1} of the characteristic polynomial Q_ψ associated to ψ (we call these roots the *local eigenvalues of the expectation function*), i.e.

$$(3.4) \qquad Q_\psi(z) \equiv z^{L+1} - \Sigma_0^L c_j z^{L-j} = 0 ,$$

where $c_j = c_j^* + \gamma_{j+1}^*$ for all j (with the convention $\gamma_{L+1}^* = 0$) .

Since expectations may involve error learning, *there is an additional important element to characterize here, which is whether and how fast traders learn (locally) from past mistakes*. To this end, it is convenient to set $x_t,\ldots,x_{t-L}$ equal to $\bar{x}$ in ψ^* and to consider the evolution of forecasts induced by the resulting difference equation

$$(3.5) \qquad {}_tx^e_{t+1} = \psi^*\begin{pmatrix} \bar{x} \;,\; \bar{x} \;,\ldots,\; \bar{x} & \\ & {}_{t-1}x^e_t \;,\ldots,\; {}_{t-L}x^e_{t-L+1} \end{pmatrix}$$

A natural assumption is that traders do learn locally from past mistakes. This means formally that the stationary state $\bar{x}$ is locally stable in the dynamics of forecasts induced by (3.5). How fast they learn is then determined by the modulus of the roots of the corresponding *error learning characteristic polynomial* Q_e, which is given by

$$(3.6) \qquad Q_e(z) \equiv z^L - \Sigma_1^L \gamma_j^* z^{L-j} = 0 .$$

I shall assume accordingly that all roots of Q_e have modulus less than 1 .

Adding $az^{L+1} Q_e(z)$ to (and substracting it from) the expression in (3.3) shows that the polynomial Q_W is given by

$$(3.7) \qquad Q_W(z) \equiv z^{L-1} Q_F(z) Q_e(z) - az^L Q_\psi(z) = 0 .$$

We reach therefore the same qualitative conclusion as in Section 2 : local stability of the stationary state in the actual dynamics with learning is governed by the interaction of the perfect foresight roots with the local eigenvalues of the expectation function, and also here (this is the novel feature in the present context) with the roots of the error learning polynomial Q_e . The perfect foresight roots and the roots of Q_e appear in fact to play similar roles in (3.7). As the latter roots were assumed to be inside the unit circle in the complex plane, one should expect error learning not to change qualitatively the results, by comparison to our previous analysis. I state first the counterpart of Proposition 2.1.

PROPOSITION 3.1. *Assume* $a \neq 0$ *(expectations matter),* $b_0 + ac_0^* \neq 0$ *(the actual dynamics with learning is locally well defined) and that all roots of the error learning polynomial* Q_e *have modulus less than* 1 *(traders do learn locally from past mistakes). Let the expectation function have two real eigenvalues* $\mu_1 < 0 < \mu_2$ *that differ from the perfect foresight roots* λ_1 *and* λ_2 *, and such that the interval* $[\mu_1, \mu_2]$ *contains in its interior all real roots of the error learning polynomial* Q_e .

Then Parts 1 *and* 2 *of Proposition* 2.1 *are valid without any change.*

Proof. From (2.10), one has $(-1)^L Q_e(\mu_1) Q_e(\mu_2) > 0$. In view of (3.7), this implies that the expressions

$$A = (-1)^{2L} Q_W(\mu_1) Q_W(\mu_2) \neq 0 , \quad B = Q_F(\mu_1) Q_F(\mu_2) \neq 0 .$$

have opposite signs. Thus the last four lines of the proof of Proposition 2.1 apply without any change. Q.E.D.

We get here the same sort of instability result as in Section 2. In a decentralized setup, one should expect the agents to be rather uncertain about the local dynamics of the system. So their (average) expectation function should tend to involve a wide net of local eigenvalues, and if expectations matter significantly, the actual dynamics with learning should be locally divergent.[10] As for local stability, the counterpart of Proposition 2.2 is

PROPOSITION 3.2. *Assume* $b_0 + ac^*_0 \neq 0$ *and that all roots of the error learning polynomial* Q_e *have modulus less than* 1 . *Let* α^* *be the maximum of* $|\Sigma_0^L c^*_j z^{L-j}/Q_e(z)|$ *when* $|z| = 1$. *Then the actual dynamics with learning is locally stable if* $|b_0| > |b_1| + |a| \alpha^*$.

Proof. Assume that Q_w has a root $|\bar{z}| \geqq 1$. From (3.3), one gets

$$|b_0| = |b_1 \bar{z}^{-1} + a [\Sigma_0^L c^*_j \bar{z}^{-j}/(1 - \Sigma_1^L \gamma^*_j \bar{z}^{-j})]| ,$$

in which case $|b_0| \leqq |b_1| + |a| \beta^*$, where β^* is the maximum of $|\Sigma_0^L c^*_j z^j/(1 - \Sigma_1^L \gamma^*_j z^j)|$ for $|z| \leqq 1$. Since the expression $(1 - \Sigma_1^L \gamma^*_j z^j)$ has no zero for $|z| \leqq 1$ (for otherwise Q_e would have a root with a modulus greater than or equal to 1) the maximum occurs for $|z| = 1$. Thus β^* is also the maximum of $|\Sigma_0^L c^*_j z^{L-j}/Q_e(z)|$ for $|z| = 1$, i.e. $\beta^* = \alpha^*$. If $|b_0| > |b_1| + |a| \alpha^*$ all roots of Q_w have modulus less than 1. Q.E.D.

To see how the parameter α^* appearing in the stability condition is related to the local characteristics of the expectation function, remark that for all $|z| = 1$

$$\alpha^* \geqq \left| \frac{\Sigma_0^L c^*_j z^{L-j}}{Q_e(z)} \right| = \left| \frac{Q_\psi(z)}{zQ_e(z)} - 1 \right| \geqq \left| \frac{Q_\psi(z)}{Q_e(z)} \right| - 1 .$$

Thus if the expectation function has a wide range of local eigenvalues and/or if traders tend to learn slowly from their past mistakes (Q_e has roots with modulus close to 1), the parameter α^* will tend to be large. On the other hand, if agents learn fast from past mistakes, all roots of Q_e, thus all coefficients γ_j^*, are small, which implies that $|Q_e(z)|$ is close to 1 when $|z| = 1$. If in addition all local eigenvalues of Q_ψ are clustered around 0, all coefficients c_j, hence all c_j^*, have a small modulus. In such a case, α^* is small. The parameter α^* is therefore a crude measure of how fast agents learn from past errors, and of the spread of the trends they are ready to extrapolate correctly out of past local deviations from the stationary state.

The sufficient stability condition stated in the Proposition requires that $|b_0| > |b_1|$. In that case, one gets local stability if the product $|a|\ \alpha^*$ is small, i.e. less than $|b_0| - |b_1|$. This will occur in particular when the influence of expectations is weak ($|a|$ is small). Or if people learn fast from past errors and extrapolate only small convergent trends out of past deviations from equilibrium (α^* is then small). Here again, these strong restrictions on the traders' learning schemes appear to involve a degree of forecasts coordination that seems unlikely in a decentralized framework.

Here is an example of such restrictions. Assume that ${}_tx^e_{t+1} = c\ x_{t-1} + (1-c)\ {}_{t-2}x^e_{t-1}$ with $0 < c < 1$. In this case, $L = 2$, $c_1^* = c$, $\gamma_2^* = (1-c)$, while the other partial derivatives of ψ^* vanish. It is readily checked that $\alpha^* = 1$ in this simple case, so the actual dynamics with learning is locally stable when $|b_0| > |b_1| + |a|$. The result relies heavily, however, on exceedingly strong restrictions on the range of regularities that traders are able to extrapolate correctly. Indeed the forecasting rule obtained by setting ${}_{t-2}x^e_{t-1}$ equal to x_{t-1} in ψ^* reduces to $\psi \equiv x_{t-1}$: the only things traders are assumed to extrapolate correctly from past data are sequences that are stationary or have period 2.

4. DISCONTINUOUS LEARNING PROCESSES

The smooth forecasting rules considered in Section 2 could be viewed as being generated, for instance, through the econometric estimation of relations that traders believe may exist between the endogenous state variable x (or its deviation $x - \bar{x}$ from equilibrium), and a prespecified set of exogenous (sunspot, instrumental) variables (Grandmont and Laroque

(1991), Section 1.2). Locally, such smooth forecasting rules act essentially as linear filters and, with a finite memory, can only extract a fixed, finite set of regularities (trends and frequencies) out of small past disturbances from the stationary state. I consider now the case in which agents try to ascertain their beliefs about the presence of a trend in past deviations from equilibrium, through the estimation of direct relations between current and past deviations of the endogenous state variable, i.e. between $x_t - \bar{x}$ and $x_{t-j} - \bar{x}$. The corresponding forecasting rules have then typically a *continuum* of local eigenvalues, but also display (in particular when estimation is made through least squares) significant *discontinuities* right at the stationary state. [11] Processes of this short have been extensively studied in the recent literature.

The local instability result we got in Section 2 will be also valid here : if traders are rather uncertain about the local dynamics of the system, and are thus prepared to extrapolate a wide range of trends out of past disturbances from the stationary state, and if the influence of expectations is significant, the actual dynamics with learning will be locally unstable for an open set, say A, of initial perturbations arbitrarily near the stationary state. Were expectation functions smooth, this fact would imply local instability for almost all small initial perturbations. A novel phenomenon will appear here, owing to the discontinuities involved in the case at hand : we shall see that there may be *another open set* of arbitrarily small perturbations, say B, for which one gets local convergence. In such a configuration, if initial perturbations away from equilibrium are drawn "at random", one will get local divergence with "positive probability" (when the initial conditions fall in the unstable set A), but one may get convergence also with "positive probability" (when the initial conditions fall in the stable set B). As we shall see, the instability result is extremely robust : it is preserved even if expectations functions change, in particular if the traders' memory increases without bound, as data accumulate over time. Moreover, the "probability" of divergence (the relative size of the set of initial perturbations for which there is local instability) is significant when the influence of expectations on the dynamics of the system is strong.

A byproduct of the analysis will be to make clear the reasons underlying some stability results that can be found in the literature, claiming that least squares learning yields convergence in models of this sort *for all (or almost all) initial perturbations* (with "probability one"). Such claims relie typically on the assumption that the range of trends which traders accept to extrapolate out of past deviations from equilibrium, is in effect restricted, usually through a socalled "*projection facility*", to lie within the foregoing open set B of initial perturbations for which there was indeed convergence without the restriction (see, e.g. Marcet and Sargent (1988, 1989 a,b)). When expectations do not matter much, the stable set B is large and the procedure may appear relatively innocent. But when expectations matter significantly, the stable set B may be fairly small, and the restriction that the "projection facility" lies in that set seems to require a degree of coordination among all present and future traders that is unlikely to prevail in a genuine decentralized setup.

To save notation, the stationary state is set at $\bar{x} = 0$, i.e. x *represents in this section a deviation from the steady state*. Since the analysis will be somewhat complicated by the discontinuity of the forecasting rule, I shall simplify matters by considering the following linearization of the temporary equilibrium relation

(4.1) $$b_1 x_{t-1} + b_0 x_t + a\ {}_t x^e_{t+1} = 0 ,$$

with $a \neq 0$ (expectations matter). This specification must be interpreted as a *local approximation* of what happens near the steady state. In particular, local instability, which would seem to imply that x_t diverges to $+\infty$ in the linear specification (4.1), only means that trajectories leave the immediate vicinity of the stationary state. To simplify even further the analysis, I shall also assume that the traders' forecasting rule does not depend on the current state

(4.2) $${}_t x^e_{t+1} = \psi(x_{t-1},\dots,x_{t-L}) ,$$

with $\psi(0,\dots,0) = 0$. Then if $b_0 \neq 0$, the actual dynamics with learning is locally defined by solving (4.1), (4.2) in x_t , which yields

(4.3) $$x_t = -b_0^{-1}[b_1 x_{t-1} + a\ \psi(x_{t-1},\dots,x_{t-L})] .$$

I introduce next the assumptions made on the forecasting rule. The case

I wish to look at arises when traders already know where is the steady state, but try to improve their forecasts out of the stationary trajectory by estimating a relation of the form $x_\tau = \beta\, x_{\tau-1}$ (perhaps up to some stochastic element), where β is an unknown parameter that is constrained to belong to some *a priori* given interval $[\mu_1,\mu_2]$. The interval stands for the range of real trends traders are prepared to extrapolate out of past deviations from the stationary state. It plays the role of the "*projection facility*" we were referring to above. If the traders' estimate at the beginning of period t is β_t , then their forecast is supposed to be obtained by "iterating twice" the relation $x_\tau = \beta\, x_{\tau-1}$, i.e. ${}_t x^e_{t+1} = \beta_t^2\, x_{t-1}$.

The actual estimate β_t may depend upon the particular story traders have in mind. For instance, if they believe in a model of the form $x_\tau = \beta\, x_{\tau-1} + \eta_\tau$, in which η_τ is white noise, the unconstrained ordinary least squares (OLS) estimate of β , made at the onset of period t by using the past deviations $x_{t-1},\ldots,x_{t-L}$, is given by

$$(4.4)\qquad \beta^*_t = \Sigma_2^L\, x_{t-j+1}\, x_{t-j}/\Sigma_2^L\, x^2_{t-j} ,$$

which is well defined when $\Sigma_2^L\, x^2_{t-j} \neq 0$. If the traders'beliefs are of the form $x_\tau/x_{\tau-1} = \beta + \eta_\tau$, then the unconstrained OLS estimate β^*_t is well defined whenever $x_{t-2} \neq 0,\ldots,x_{t-L} \neq 0$, and is given by

$$(4.5)\qquad \beta^*_t = \frac{1}{L-1}\, \Sigma_2^L\, x_{t-j+1}/x_{t-j} .$$

In either case, the actual estimate β_t will be set by making the "projection facility" come into play. If the unconstrained estimate β^*_t belongs to $[\mu_1,\mu_2]$, the agents will accept it and set $\beta_t = \beta^*_t$. On the other hand if past data tell the traders that the unconstrained estimate should lie outside the interval, they simply discard it and replace it by the closest element of $[\mu_1,\mu_2]$, i.e. $\beta_t = \mu_1$ if $\beta^*_t < \mu_1$ and $\beta_t = \mu_2$ if $\beta^*_t > \mu_2$ (of course, the "projection facility" plays no role if μ_1 and μ_2 are infinite).

Here as in the previous sections, the details of the story one tells about how traders learn do not matter much. *What is important is the set of local regularities traders end up extrapolating from past deviations.* A closer look at (4.4), (4.5), shows that in both cases, the unconstrained estimate is in fact a weighted average of the past ratios x_{t-j+1}/x_{t-j} , $j = 2,\ldots,L$, the weights being $x_{t-j}^2/\Sigma_2^L x_{t-j}^2$ in one case and $1/(L-1)$ in the other, and therefore that it lies between their maximum and minimum. This is the only qualitative property of the forecasting rule that will be needed in the sequel.

(4.a) (Traders average out growth rates of past deviations from the steady state). The forecasting rule satisfies $\psi(0,\ldots,0) = 0$ and is a well defined function when $x_{t-2} \neq 0$, ..., $x_{t-L} \neq 0$. It has then the form

$$\psi(x_{t-1},\ldots,x_{t-L}) = [\beta(x_{t-1},\ldots,x_{t-L})]^2 x_{t-1} .$$

The estimate $\beta(x_{t-1},\ldots,x_{t-L})$ belongs to the intersection of the prespecified interval $[\mu_1,\mu_2]$, and of the interval $[m,M]$ defined by the minimum m and the maximum M of the past ratios x_{t-j+1}/x_{t-j} for $j = 2,\ldots,L$, whenever the intersection of these two intervals is nonvoid. When $M \leqq \mu_1$, the estimate is equal to μ_1 . It is equal to μ_2 when $\mu_2 \leqq m$.

A consequence of the assumption is that $\psi(r^{L-1} x,\ldots,r x,x) = r^{L+1} x$ for every $x \neq 0$ and every $\mu_1 \leqq r \leqq \mu_2$: *the range of real "local eigenvalues" of ψ is the whole interval $[\mu_1,\mu_2]$* . The counterpart of the good news is of course that even if the forecasting rule is assumed to be continuous for $x_{t-2} \neq 0$,...,$x_{t-L} \neq 0$ (which we needed not assume in *(4.a)*), it is bound to be highly discontinuous at the stationary state $x = 0$.

The specification *(4.a)* of the forecasting rule generates an interesting structure for the actual dynamics with learning. Indeed, one can rewrite (4.3) whenever $x_{t-1} \neq 0,\ldots,x_{t-L} \neq 0$

$$(4.6) \qquad x_t/x_{t-1} = \Omega(\beta(x_{t-1},\ldots,x_{t-L})) \text{ , with } \Omega(\beta) \equiv - b_0^{-1}[b_1 + a\beta^2].$$

The map Ω, introduced in the literature (in a multidimensional stochastic framework) by Marcet and Sargent (1988, 1989a), has a remarkable interpretation. It associates to the estimate $\beta_t = \beta(x_{t-1},\ldots,x_{t-L})$, made by the traders at the onset of period t, about the average trend they think is present in past deviations from equilibrium, the *actual* (gross) growth rate that is observed in that period. The fixed points of Ω coincide with the perfect foresight roots whenever they are real, since $\Omega(\beta) = \beta$ is equivalent to $Q_F(\beta) = 0$, where Q_F is the perfect foresight characteristic polynomial

(2.6). The graph of Ω is a parabola, the asymptotic branches of which go up when the sum of the two local perfect foresight roots, i.e. $\lambda_1 + \lambda_2 = - b_0 a^{-1}$, is positive, down otherwise.

As mentioned at the beginning of the Section, if traders are prepared to extrapolate a wide range of real trends out of past deviations from the steady state (if the "projection facility" $[\mu_1,\mu_2]$ is relatively large), and if expectations matter significantly, then one should get, as in the smooth case, local instability of the actual dynamics with learning for an open set of small initial perturbations. Owing to the discontinuity of the forecasting rule, however, there may also exist here another open set of small initial perturbations generating local convergence. The following simple example should make clear why one must expect findings of this sort.

Example 4.1. Assume that traders use only the past two deviations x_{t-1}, x_{t-2} when computing β_t. With obvious notation, let the estimate $\beta(x_{t-1}/x_{t-2})$ of assumption *(4.a)* be equal to x_{t-1}/x_{t-2} if the ratio belongs to $[\mu_1,\mu_2]$, to μ_1 if $x_{t-1}/x_{t-2} \leqq \mu_1$, and to μ_2 if the ratio exceeds μ_2 . In view of (4.6), the actual dynamics with learning is then given by $x_t/x_{t-1} = \Omega(\beta(x_{t-1}/x_{t-2}))$, provided that these ratios are well defined. A case in which the two local perfect foresight roots are real and their sum is positive is represented in Fig. 1.a. One sees there that the consequence of composing the maps Ω and β is to "chop off" the parts of the parabola representing Ω that lie outside the "projection facility", and to replace these parts by horizontal lines. The Figure is drawn under the assumption that the given interval $[\mu_1,\mu_2]$ contains in its interior the two perfect foresight roots, and that $\Omega(\mu_2) > 1$. There is local divergence when the initial perturbations are such that $x_{-1}/x_{-2} > \lambda_2$, since then the sequence x_t/x_{t-1} tends to $\Omega(\mu_2)$. In fact, the argument goes through even when the two perfect foresight roots are complex (so that the parabola lies above the 45° line). All what is needed to get local divergence is that μ_2 is larger than both perfect foresight roots when they are real, and that it satisfies $\Omega(\mu_2) > 1$. The position of μ_1 is indeed of minor importance. Proposition 4.2 below states a general instability result along this line for estimators satisfying *(4.a)*. The result is valid even these estimators depend on time, and in particular when the agents' memory increases without bound as data accumulate.

On the other hand, if the map Ω is supposed to be contracting at λ_1 in the Fig. 1.a, the ratios x_t/x_{t-1} (and therefore the estimates β_t) go to λ_1 when the initial perturbations are such that x_{-1}/x_{-2} is indeed close to λ_1. In that case, one gets in addition local convergence of the deviations x_t to 0 whenever $|\lambda_1| < 1$. In such a configuration, the system is stable, *and* the agents do learn the true value of the stable perfect foresight root λ_1. Proposition 4.3 states a general convergence result along this line for estimators satisfying *(4.a)*. There, these estimators may also vary over time, but the assumption of a bounded memory seems to play an important role. One can, however, still get stability of the system (i.e. convergence of the deviations x_t to 0), even with time dependent estimators involving an unbounded memory, and even if the map Ω is not contracting at λ_1, provided that one gives up the requirement that agents eventually learn the true value of λ_1 (Corollary 4.5).

<u>Fig. 1.a</u> <u>Fig. 1.b</u>

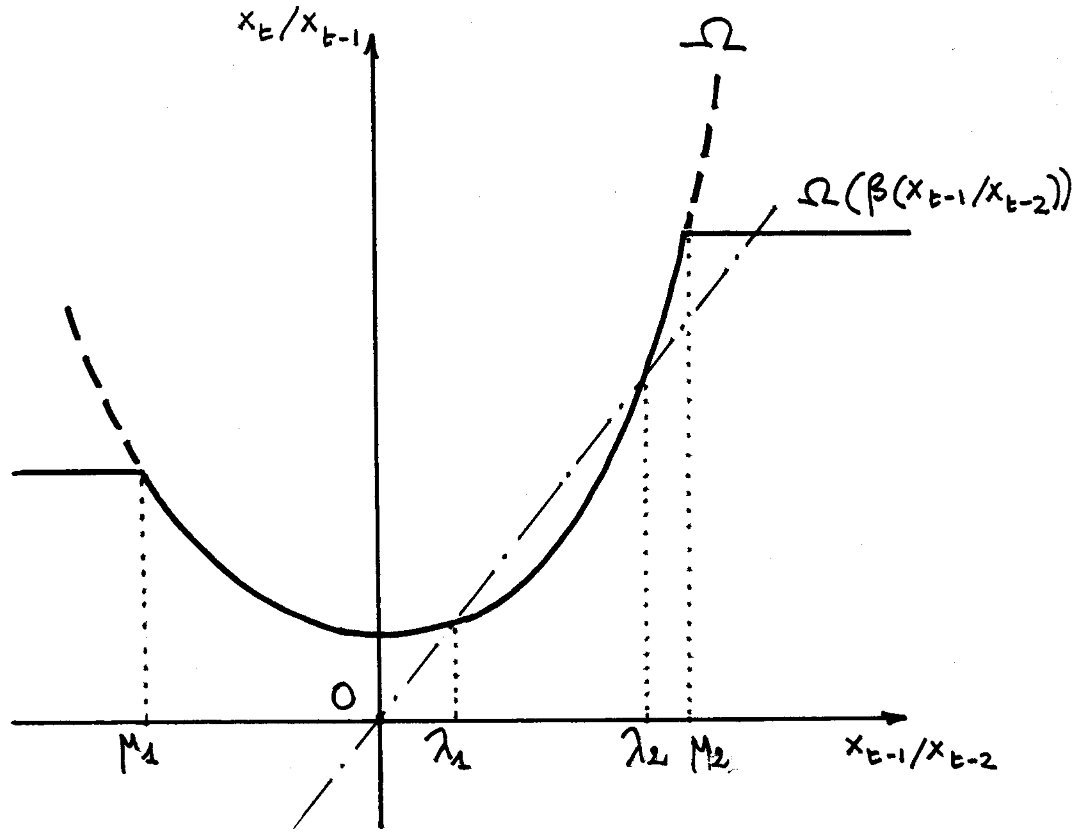
Instability

We first state a general instability result.[12]

PROPOSITION 4.2. *Assume that* $a \neq 0$, $b_0 \neq 0$ *have opposite signs (the sum of the two perfect foresight roots, i.e.* $\lambda_1 + \lambda_2 = -b_0 a^{-1}$, *is positive) and let assumption (4.a) hold (agents average out trends in past deviations from equilibrium). Let* $x_{-1} \neq 0 , \ldots, x_{-L} \neq 0$ *be arbitrary initial small perturbations from the steady state prior to date* $t = 0$.

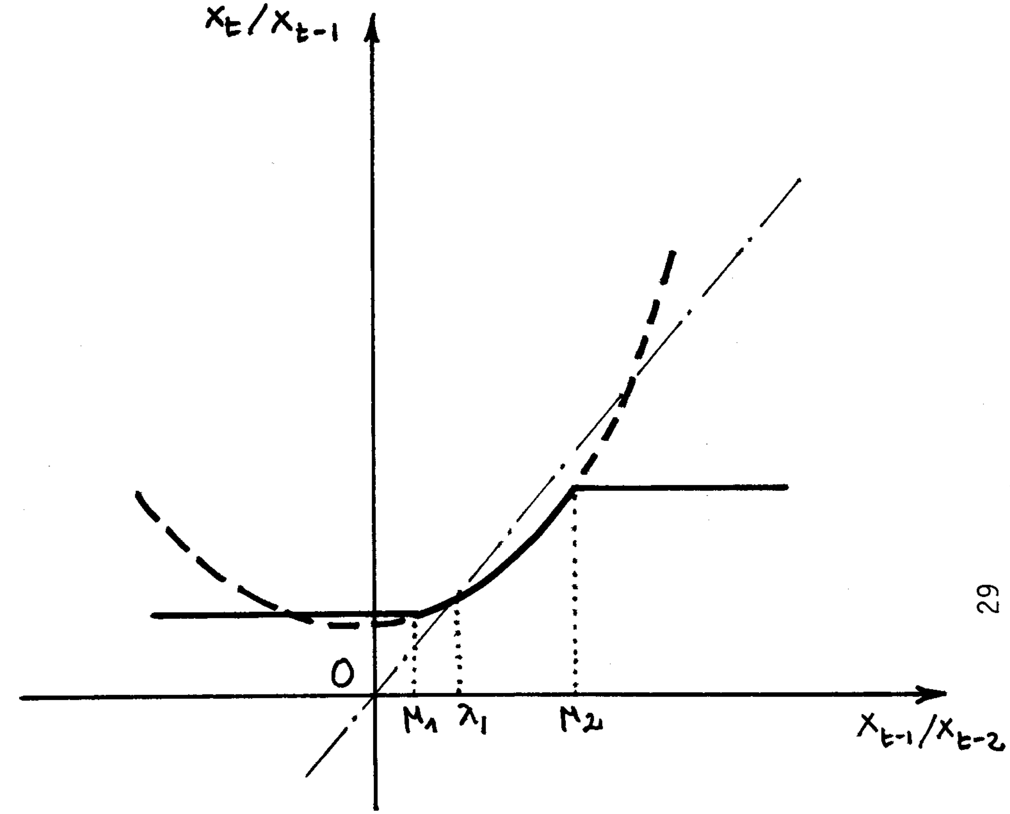
Assume that $\mu_2 > 0$ *is larger than all perfect foresight roots that are real and satisfies* $\mu_2 \geqq 1$, *or more generally* $\Omega(\mu_2) > 1$. *Then the actual dynamics with learning is locally divergent for an open set of initial perturbations. Specifically, if* $x_{-1}/x_{-2} \geqq \mu^*, \ldots, x_{-L+1}/x_{-L} \geqq \mu^*$, *where* $\mu^* > 0$ *belongs to* $[\mu_1, \mu_2]$ *and exceeds all real perfect foresight roots, then the sequence of deviations from equilibrium generated by the actual dynamics with learning satisfies* $x_t \neq 0$ *and* $x_t/x_{t-1} \geqq \Omega(\mu^*) > \mu^*$ *for all* $t \geqq 0$, *and is thus locally divergent if* $\Omega(\mu^*) > 1$.

The result still hold if the estimator $\beta(x_{t-1}, \ldots, x_{t-L})$ *appearing in (4.a) depends on time and is of the form* $\beta_t(x_{t-1}, \ldots, x_{-1}, \ldots, x_{-L})$, *and in particular if the agents' memory increases without bound as data accumulate.*



x_t/x_{t-1}
Ω
$\Omega(\beta(x_{t-1}/x_{t-2}))$
0
μ_1
λ_1
λ_2
μ_2
x_{t-1}/x_{t-2}

Fig. 1.a



x_t/x_{t-1}
0
μ_1
λ_1
μ_2
x_{t-1}/x_{t-2}

Fig.1.b

The proof of this statement is almost trivial, in view of Fig. 1.a (in the more general case considered here, the parabola may lie above the 45° line since the perfect foresight roots may be complex). Indeed, if μ^* is chosen as in the Proposition, one has $\Omega(\mu^*) > \mu^*$. Then at no date can the estimates fall below μ^* , and so one must have for all $t \geqq 0$

$$x_t/x_{t-1} = \Omega(\beta(x_{t-1},\ldots,x_{t-L}) \geqq \Omega(\mu^*) > \mu^* ,$$

and the result follows. The argument clearly goes through even when the estimator in *(4.a)* is of the form $\beta_t(x_{t-1},\ldots,x_{-1},\ldots,x_{-L})$.

The case in which a and b_0 have the same sign is handled similarly by focussing attention instead on the lower trend $\mu_1 < 0$, which should be then less than all real perfect foresight roots and satisfy $\mu_1 \leqq -1$, or $\Omega(\mu_1) < -1$. So we have here the analogue of Proposition 2.1 in the smooth case : *if* $\mu_1 \leqq -1$, $\mu_2 \geqq 1$ *(or more generally, if* $\mu_1 < 0$, $\mu_2 > 0$ *are such that* $|\Omega(\mu_1)| > 1$, $|\Omega(\mu_2)| > 1$), *and if the interval* $[\mu_1,\mu_2]$ *contains in its interior all perfect foresight roots that are real, there is local divergence for an open set of initial conditions.* [13]

The discussion of Section 2 on the plausibility of getting local divergence in the actual dynamics with learning is relevant here. In a decentralized framework, one should expect agents to be uncertain about the dynamics of the system, so the range of regularities they are prepared to extrapolate out of past deviations from equilibrium, here the "projection facility" $[\mu_1,\mu_2]$, should be large and the conditions of Proposition 4.2 should hold, especially when the influence of expectations is significant (if $|a|$ is large, the perfect foresight roots are small, while $|\Omega(\mu_1)|$ and $|\Omega(\mu_2)|$ are large).

Stability

I now look at conditions ensuring stability (for an open set of initial deviations) of the actual dynamics with learning, first when one requires that agents do eventually learn the true value of a stable perfect foresight root, and next when the requirement is relaxed.

PROPOSITION 4.3. *Let* $a \neq 0$, $b_0 \neq 0$ *and let assumption (4.a) hold.*

Assume that both perfect foresight roots are real with $0 < |\lambda_1| < |\lambda_2|$, *that* λ_1 *belongs to the interior of the interval* $[\mu_1,\mu_2]$ *and that the map* Ω *is contracting at* λ_1 , *i.e.* $|\Omega'(\lambda_1)| < 1$. *Then, for an open cone of initial deviations, the agents will eventually learn the true value of* λ_1, *and the deviations* x_t *will go to 0 if the root* λ_1 *is stable.*

Specifically, let $\varepsilon > 0$ *be small enough so that* $(\lambda_1-\varepsilon\ ,\ \lambda_1+\varepsilon)$ *does not contain 0 or* λ_2, *and* $|\Omega'(\beta)| \leqq k$ *for some* $k < 1$ *for all* β *in that interval. Then if the initial ratios* $x_{-1}/x_{-2},\ldots,x_{-L+1}/x_{-L}$ *all lie within* ε *of* λ_1 , *the sequence of deviations generated by the actual dynamics with learning satisfies* $x_t \neq 0$ *for all* $t \geqq 0$, $\lim_{t\to\infty} x_t/x_{t-1} = \lim_{t\to\infty} \beta(x_{t-1},\ldots,x_{t-L}) = \lambda_1$, *and is thus locally convergent to 0 when* $|\lambda_1| < 1$.

This is merely a restatement of the local convergence result of Grandmont and Laroque (1991, Part 2 of Proposition 2 and Corollary 1). The principle of the proof, which I will only sketch to save space, should be fairly clear. If the initial ratios are within ε of λ_1 , so will be the estimate *(4.a)* and thus x_0/x_{-1} , since Ω is locally contracting. By induction, all ratios x_t/x_{t-1} and all estimates β_t will be trapped within ε of λ_1 , for $t \geqq 0$. Convergence to λ_1 of these ratios and of the estimates then follows from the fact that Ω is locally contracting and that the memory L is finite. When in addition $|\lambda_1| < 1$, this implies that the sequence of deviations x_t tend to 0. The proof obviously is valid even if the estimator $\beta(.)$ appearing in the forecasting rule *(4.a)* is time-dependent, but it relies essentially on the assumption of a bounded memory.

One may remark that the only property of the interval $(\lambda_1 - \varepsilon\ ,\ \lambda_1 + \varepsilon)$ that was actually used in the above argument, is that any sequence of estimates generated by the difference equation $\beta_t = \Omega(\beta_{t-1})$, remains trapped in that interval once it starts there, and then converges to the perfect foresight root λ_1. It turns out that when the map Ω is contracting at λ_1 , the largest interval having this property is $(-\lambda_2,\lambda_2)$. Indeed, if we focus, for the sake of concreteness, on the case in which the sum of the two real perfect foresight roots is positive (b_0 and a have opposite signs), this is obvious if the root λ_1 of smallest modulus is nonnegative (Fig. 1.a). In the case where λ_1 is negative (Fig. 2.a), we observe that the image by Ω of the interval $(-\lambda_2,\lambda_2)$ is $(-b_0^{-1} b_1,\lambda_2)$, and that the fact that Ω is contracting at λ_1 , i.e. $\Omega'(\lambda_1) = |2\lambda_1/(\lambda_1+\lambda_2)| < 1$, is equivalent to $|b_0^{-1} b_1| = |\lambda_1 \lambda_2/(\lambda_1+\lambda_2)| < \lambda_2/2$. Thus all trajectories generated by $\beta_t = \Omega(\beta_{t-1})$ stay in $(-\lambda_2,\lambda_2)$ once they start there, and it is not difficult to see that they all then converge to λ_1 (one needs only to verify that $\Omega^2(0) = \Omega(\Omega(0)) < 0$, as shown by the arrows in Fig. 2.a).

So one can replace the interval $(\lambda_1 - \varepsilon , \lambda_1 + \varepsilon)$ by $(-\lambda_2,\lambda_2)$ in Proposition 4.3. The only thing we have to worry about then, is that a sequence of deviations x_t may hit at some (finite) time the steady state $\bar{x} = 0$, in which case the agents' estimate may not be defined. If the estimator $\beta(.)$ in *(4.a)* is regular enough, however, the set of initial deviations $x_{-1},\ldots,x_{-L}$ for which this may happen, should be "thin" (of Lebesgue measure zero), in which case this event could be considered having "probability zero".

COROLLARY 4.4. *Under the assumptions of Proposition 4.3, let the initial perturbations* $x_{-1} \neq 0,\ldots,x_{-L} \neq 0$ *be such that the ratios* $x_{-1}/x_{-2},\ldots,x_{-L+1}/x_{-L}$ *all belong to the open interval* $(-\lambda_2,\lambda_2)$. *Then if the sequence of deviations generated by the actual dynamics with learning is well defined (i.e. satisfies* $x_t \neq 0$ *for all* $t \geq 0$*), the agents will eventually learn the true value of the perfect foresight root* λ_1 , *i.e.* $\lim_{t\to\infty} \beta(x_{t-1},\ldots,x_{t-L}) = \lambda_1$, *and the sequence* x_t *will converge to* 0 *if* $|\lambda_1| < 1$.

The result is valid even if the estimators $\beta(.)$ *in (4.a) depend on time, but involve a bounded memory.*

<u>Fig. 2.a</u> <u>Fig. 2.b</u>

The approach I have adopted so far is to show the possibility of stability for an open cone of initial deviations under the requirement that the agents eventually learn the value of a stable perfect foresight root λ_1. More generally, one may analyze when the system is stable (the deviations x_t go to 0), *without the agents necessarily discovering the value of* λ_1 : after all, this is the route we took in the previous two sections. If we go back to the proof of Proposition 4.3 (or Corollary 4.4), we see that the only thing we need to implement this program, is to find this time an interval

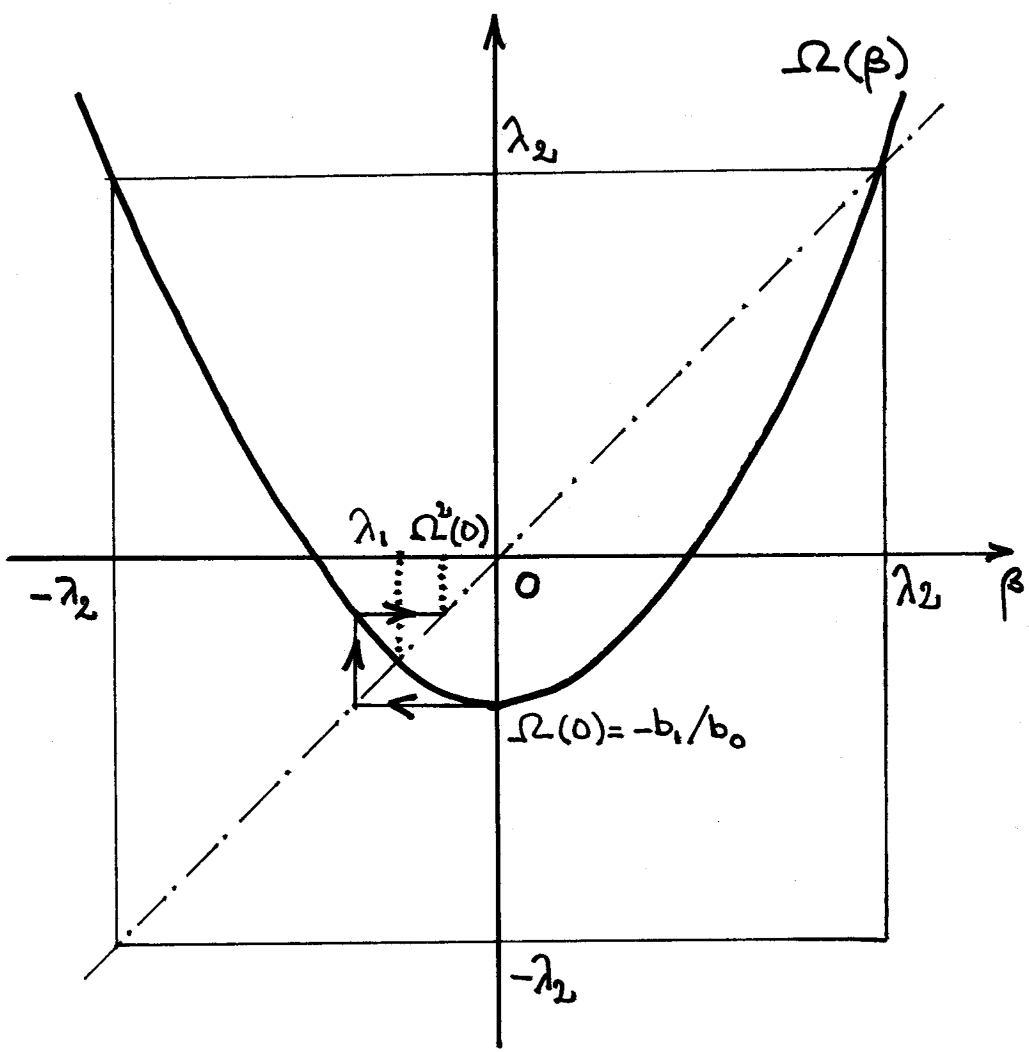

$\Omega(\beta)$
λ_2
$-\lambda_2$
λ_1
$\Omega^2(0)$
0
β
$\Omega(0)=-b_1/b_0$

Fig. 2.a

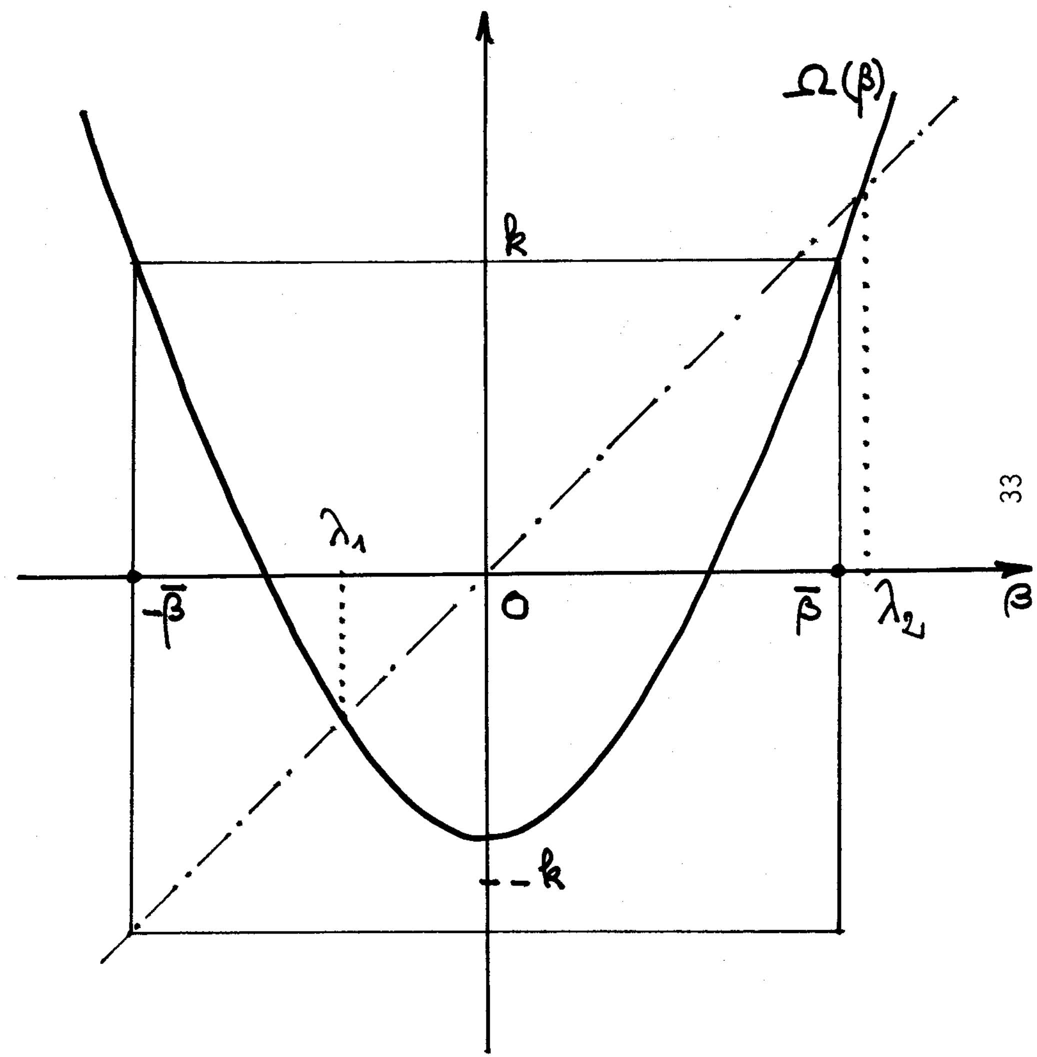

$\Omega(\beta)$
k
$-k$
λ_1
$-\bar{\beta}$
0
$\bar{\beta}$
λ_2
β

Fig. 2.b

$B = (\beta_1, \beta_2)$ that is invariant by Ω, such that $|\Omega(\beta)| \leqq k < 1$ for all β in B, and that has a nonempty intersection with $[\mu_1, \mu_2]$. Then if the initial ratios belong to B, all ratios x_t/x_{t-1} and estimates β_t will be trapped in B for all $t \geqq 0$. They will not necessary converge to λ_1 as Ω may not be contracting at λ_1 (Fig. 2.b), but the sequence of deviations x_t, if it does not hit the equilibrium $\bar{x} = 0$, will go to 0, since $|x_t| = |\Omega(\beta_t) x_{t-1}| \leqq k|x_{t-1}|$. One will then get stability of the system, in fact even when the estimators in *(4.a)* depend on time and involve an unbounded memory.

For such an invariant interval to exist, it is necessary that the two perfect foresight roots be real, and that $|\Omega(0)| < 1$, or equivalently $|b_0| > |b_1|$. Then if the root λ_2 of largest modulus is stable, i.e. $|\lambda_2| < 1$, the largest candidate for such interval is $B = (-\lambda_2, \lambda_2)$. It is indeed actually invariant whenever $|\Omega(0)| = |b_1/b_0| = |\lambda_1\lambda_2/(\lambda_1+\lambda_2)| \leqq |\lambda_2|$ (or equivalently whenever the slope of Ω at the smallest root λ_1 is larger than or equal to -2) : one has then $|\Omega(\beta)| \leqq |\lambda_2|$ for all β in B. If on the other hand $|\lambda_2| \geqq 1$, one can take any invariant interval of the form $B = (-\bar{\beta}, \bar{\beta})$, where $\bar{\beta}$ is solution of $|\Omega(\beta)| = k$ with $|b_1/b_0| \leqq k < 1$ (Fig. 2.b).

COROLLARY 4.5. *Let* $a \neq 0$, $b \neq 0$ *and assume that the estimator in (4.a) is of the form* $\beta_t(x_{t-1}, \ldots, x_{-1}, \ldots, x_{t-L})$, *so that the agents' memory may increase without bound as data accumulate. Let* B *be an open interval that is invariant by* Ω , *such that* $|\Omega(\beta)| \leqq k < 1$ *for all* β *in* B, *and that has a nonempty intersection with the "projection facility"* $[\mu_1, \mu_2]$.

Then for any initial perturbations $x_{-1} \neq 0, \ldots, x_{-L} \neq 0$ *such that the ratios* $x_{-1}/x_{-2}, \ldots, x_{-L+1}/x_{-L}$ *all belong to* B, *the sequence of deviations* x_t *generated by the actual dynamics with learning converges to* 0 *whenever it satisfies* $x_t \neq 0$ *for all* $t \geqq 0$. *The agents may not, however, learn the value of a perfect foresight root.*

The foregoing results exhibit open sets (cones) of initial perturbations for which the actual dynamics with learning is stable whenever it is well defined. These sets (or the corresponding intervals that are invariant by Ω) may be of fairly significant size when expectations don't matter much. Indeed, if the parameter $|a|$ is small, given b_0 and b_1, the parabola representing Ω is "close" to the horizontal line of equation $f(\beta) = -b_1/b_0$. So if the influence of expectations on the system is small, the two perfect foresight roots are both real ; the root of smallest modulus λ_1 is then close to $-b_1/b_0$ (and thus stable if $|b_0| > |b_1|$), with $|\Omega'(\lambda_1)|$ small, while $|\lambda_2|$ is large. Then if $|b_0| > |b_1|$, one can apply Corollary 4.4 and conclude that the actual dynamics with learning, whenever it is well defined, is locally stable, with the agents eventually learning the true value of λ_1, if all initial ratios $x_{-1}/x_{-2},\dots,x_{-L+1}/x_{-L}$ lie in the fairly large interval $(-\lambda_2,\lambda_2)$, and if the agents' estimation procedure involves a finite memory. If the agents' memory is unbounded, the invariant interval B in Corollary 4.5 can also be chosen as fairly large, and one gets local stability of the actual dynamics with learning when the initial ratios lie in B (but the agents may not learn the true value of the stable root λ_1). This conclusions are not quite surprising since, after all, when expectations don't matter much, the dynamics of the system is "not far" from $x_t = -b_0^{-1} b_1 x_{t-1}$, which is stable whenever $|b_0| > |b_1|$.

The picture is almost opposite, however, when the influence of expectations on the dynamics is significant. When $|a|$ increases, given b_0 and b_1, the value of $|\Omega(\beta)|$ becomes large for any given $\beta \neq 0$. In that case, either the perfect foresight roots are complex, or when they are real, they have a small modulus, with $|\Omega'(\lambda_1)|$ being large. *When expectations matter a lot, the conditions ensuring stability in Proposition 4.3 and its Corollaries, are bound to be violated for, as the reader will easily verify, the invariant intervals appearing there become empty.*

Here again, one *can* get local stability for (almost) all initial perturbations in the present model, but that will often mean either that the influence of expectations is small, or that the range of the "projection facility" has been appropriately restricted. Fig. 1.b should make clear how the argument would go in the simple framework of Example 4.1. The trick is straightforward enough : it suffices to restrict the "projection facility" so that the interval $[\mu_1,\mu_2]$ is small enough and contains in its interior the target perfect foresight root λ_1 ! It is easy to see that the same procedure works in the general case. Assume that the conditions of Proposition 4.3 are satisfied and that, in addition, the "projection facility" $[\mu_1,\mu_2]$ has been restricted to lie within ε of λ_1. Then no matter what are the (nonzero) initial perturbations, the estimate $\beta(x_{-1},\dots,x_{-L})$ at date 0 will belong to $(\lambda_1-\varepsilon\ ,\ \lambda_1+\varepsilon)$, and so will $x_0/x_{-1} = \Omega(\beta(x_{-1},\dots,x_{-L}))$ since Ω is contracting on that interval. It should be clear that when the process is repeated, the first L ratios $x_{L-1}/x_{L-2},\ldots,x_0/x_{-1}$ will all be trapped within ε of λ_1. From $t = L$ on, Proposition 4.3 comes into play, and one gets convergence.

COROLLARY 4.6. *Under the assumptions of Proposition 4.3, suppose that the interval* $[\mu_1,\mu_2]$ *lies within* ε *of the perfect foresight root* λ_1. *Then for every initial perturbations* $x_{-1} \neq 0,\ldots,x_{-L} \neq 0$, *the sequence of state variables generated by the actual dynamics with learning satisfies* $x_t \neq 0$ *for all* $t \geqq 0$, $\lim_{t\to\infty} x_t/x_{t-1} = \lim_{t\to\infty} \beta(x_{t-1},\ldots,x_{t-L}) = \lambda_1$, *and is thus locally convergent when* $|\lambda_1| < 1$.

This is the procedure that underlies many results found in the literature claiming stability for almost all initial perturbations. A prominent example is Marcet and Sargent (1988, 1989a,b). Of course, one can play with the Corollaries instead of Proposition 4.3, and get stability by requiring the "projection facility" to lie within the various invariant intervals appearing there. As noted earlier, these invariant intervals may be fairly large if expectations don't matter much, so the procedure is relatively innocent in this case. But the size of these invariant intervals, and thus of the "projection facility", will be rather small (they may even vanish) when the influence of expectations is significant.

As a matter of fact, the most general procedure to "force" stability is simply to require, in the spirit of Corollary 4.5, that $|\Omega(\beta)| \leqq k < 1$ for all β in $[\mu_1,\mu_2]$. Then the actual dynamics with learning will be automatically stable for all nonzero initial perturbations, since it will satisfy $|x_t| = |\Omega(\beta_t)\, x_{t-1}| \leqq k\,|x_{t-1}|$ for all t, no matter how estimates are formed (the estimators in *(4.a)* may depend on time and involve an unbounded memory). However, the agents typically will not learn the value of a stable perfect foresight root. *Since* $|\Omega(\beta)| \leqq (|b_1| + |a|\,\alpha)/|b_0|$, *where* α *is the maximum of* μ_1^2 *and of* μ_2^2, *one will get this sort of stable configuration if* $|b_0| > |b_1| + |a|\,\alpha$, *which is the exact analogue of Propositions* 2.2 *and* 3.2 *in the smooth case.*

All these procedures have in common to achieve stability at the cost of assuming that agents systematically ignore the locally divergent tendencies that appear in past deviations from equilibrium, an unlikely feature in a truly decentralized framework. In all cases, the restrictions involved are exceedingly strong when the influence of expectations is large.

Unbounded memory : recursive least squares

The foregoing stability results (Corollary 4.5) suggest that when the agents' memory is unbounded, one may get convergence to 0 of the deviations from equilibrium for an open set of initial perturbations, but that the agents may not learn the true value of a stable perfect foresight root.

To get a little more information on this issue, I consider here the specific case of the unconstrained OLS learning scheme (4.4), in which agents are now assumed to keep track of all previous observations. The estimate made at the onset of period t is then given by

$$(4.7) \qquad \beta_t = \Sigma_2^{t+L} x_{t-j+1}\, x_{t-j} / \Sigma_2^{t+L} x_{t-j}^2 ,$$

which is well defined whenever the denominator does not vanish. As is well known (see Wickens (1982), Marcet and Sargent (1989a)), the above time dependent estimator can be restated in a nice recursive form, by setting the denominator in (4.7) equal to $1/\omega_t^2$

$$(4.8) \qquad 1/\omega_{t+1}^2 = x_{t-1}^2 + 1/\omega_t^2 \qquad , \qquad \beta_{t+1}/\omega_{t+1}^2 = x_t x_{t-1} + \beta_t/\omega_t^2 .$$

Thus if the sequence of deviations $x_t \neq 0$ from equilibrium is known, the dynamic evolution of the estimates can be determined by the following difference equation, in which we note $m(\xi) = 1/(1 + \xi^2)$

$$(4.9) \quad \omega_{t+1}^2 = m(\omega_t x_{t-1})\, \omega_t^2 \, , \; \beta_{t+1} = m(\omega_t x_{t-1})\, \beta_t + [1-m(\omega_t x_{t-1})](x_t/x_{t-1}) ,$$

for all $t \geqq 0$, subject to the initial conditions

$$(4.10) \qquad 1/\omega_0^2 = \Sigma_2^L x_{-j}^2 \qquad , \qquad \beta_0 = \Sigma_2^L x_{-j+1}\, x_{-j}/\Sigma_2^L x_{-j}^2 .$$

An interesting feature of this learning scheme is that agents revise their estimates "adaptively", the estimate β_{t+1} at the end of period t being a convex combination of the previous estimate β_t and of the actual ratio x_t/x_{t-1} observed in that period.

To complete the description of the evolution of the system under this learning scheme, we have to use (4.6), which defines the actual deviation in period t, i.e. $x_t = \Omega(\beta_t)\, x_{t-1}$, from the estimate made at the onset of the period. The actual dynamics is thus described, out of equilibrium (for $x_t \neq 0$), by the following three-dimensional, time-independent difference equation, which is obtained by replacing in (4.9) the actual ratio by $\Omega(\beta_t)$

$$(4.11)\qquad \left| \begin{array}{l} x_t = \Omega(\beta_t)\, x_{t-1} \quad , \qquad \omega^2_{t+1} = m(\omega_t x_{t-1})\omega^2_t \ , \\ \\ \beta_{t+1} = m(\omega_t x_{t-1})\beta_t + [1-m(\omega_t x_{t-1})]\ \Omega(\beta_t), \text{ with } m(\xi) = 1/(1+\xi^2), \end{array} \right.$$

for all $t \geqq 0$, again subject to the initial conditions (4.10).

Although the dynamical system (4.11) is smooth everywhere in $(x_t, \omega_{t+1}, \beta_{t+1})$, its trajectories still depend in a very discontinuous fashion on the deviations $x_{-1} \neq 0, \ldots, x_{-L} \neq 0$, through their influence on the initial estimate β_0. We know that Proposition 4.2 applies to the case at hand (μ_1 and μ_2 are infinite here since there is no "projection facility"), so there is local divergence for an open set (cone) of initial deviations. This can be very easily seen directly from (4.11). Let us focus for the sake of concreteness on the case in which the parameters $a \neq 0$, $b_0 \neq 0$, have opposite signs, so that the parabola representing Ω has it asymptotic branches going up. If the initial ratios $x_{-1}/x_{-2}, \ldots, x_{-L+1}/x_{-L}$ are all greater than $\mu^* > 0$, where μ^* exceeds all perfect foresight roots that are real, the initial estimate β_0 is also by construction larger than μ^* , while the actual ratio $x_0/x_{-1} = \Omega(\beta_0)$ will exceed β_0 . But (4.11) then shows that β_1 , being a convex combination of β_0 and $\Omega(\beta_0)$, must lie in between these two values. Thus by induction, for all $t \geqq 0$, the sequence of estimates β_t is increasing, satisfies $\beta_t < \beta_{t+1} < \Omega(\beta_t) < \Omega(\beta_{t+1})$ and exceeds μ^* , while the sequence of actual ratios $x_t/x_{t-1} = \Omega(\beta_t)$ is also increasing and is larger that $\Omega(\mu^*) > \mu^*$. So one gets local divergence if μ^* is large enough to ensure $\mu^* \geqq 1$, or more generally $\Omega(\mu^*) > 1$.

Corollary 4.5, which is valid here, suggests that one may have stability of the dynamics with learning for an open cone of initial perturbations, but that the agents may not succeed in learning the value of a stable perfect foresight root. That this may be true is hinted at by the fact that the system (4.11) has a whole set of stationary solutions of the form $x_t = 0$, $\omega_t = \omega^* \geqq 0$, $\beta_t = \beta^*$. So any of these points might be stable. The issue is easily settled by taking advantage of the smoothness of the dynamical system (4.11), and by studying its local stability by linearizing it near an arbitrary stationary solution $(0,\omega^*,\beta^*)$. This yields

$$(4.12) \qquad dx_t = \Omega(\beta^*)dx_{t-1} \ , \ d\omega_{t+1} = d\omega_t \ , \ d\beta_{t+1} = d\beta_t \ .$$

It follows from a standard "center manifold reduction" (see e.g. Grandmont (1988b), Theorem B.5.3) that there is a small open neighborhood V of $(0,\omega^*,\beta^*)$ such that the difference equation (4.11), restricted to that neighborhood, is equivalent (through a nonlinear change of variables) to (4.12). Therefore, if the initial perturbations generate through (4.10) an initial estimate β_0 in that neighborhood V (this will occur for an open cone of initial perturbations), the dynamics with learning will stay in V and will be stable (x_t will go to 0) *if and only if* $|\Omega(\beta^*|$ < 1 . Then (ω_t,β_t) will converge to some $(\bar{\omega},\bar{\beta})$, which may not be (ω^*,β^*), and will surely differ from any perfect foresight root if $\beta^* \neq \lambda_1$, λ_2 and if V is small enough. So the study of this specific case of learning through recursive least squares confirms what we suspected from our general stability analysis. When the agents' memory is unbounded, what matters for stability is whether the *values* of $|\Omega(\beta)|$, rather than its derivative, are less than 1, whereas stability does not necessarily mean that the agents learn the value of a stable perfect foresight root.

5. BAYESIAN LEARNING

In the learning schemes considered so far, the "prior" of the agents (the set of regularities, such as trends, they are willing to extrapolate out of past deviations from equilibrium) was fixed, and one may wonder how the results depend upon this particular feature. I study in this section a case in which agents use Bayesian learning to ascertain the presence of a trend in past disturbances from equilibrium. In that case, the agents' prior, in particular the degree of confidence with which they hold their beliefs about this hypothetical trend, is indeed updated over time in the light of new observations. We shall see that the same general "uncertainty principle" as before emerges from the analysis. If the agents initially believe that the system will be locally unstable, or if they are fairly uncertain about its local stability (the variance of their initial prior is large), the actual dynamics with learning will be locally unstable. If on the other hand they are initially fairly sure about the local stability of the system, this prophecy will be self-fulfilling. Agents may not, however, learn the value of a stable perfect foresight root. Moreover, the range of circumstances (i.e. of the agents' initial priors) under which one gets local instability (resp. stability) becomes fairly significant (resp. small) when the influence of expectations on the system is strong. [14]

I will stick here to the linear local approximation of the temporary equilibrium relation (4.1)

$$(5.1) \qquad b_1\, x_{t-1} + b_0\, x_t + a\, {}_t x^e_{t+1} = 0 \ ,$$

with $a \neq 0$, $b_0 \neq 0$, in which x is a deviation from the steady state $\bar{x} = 0$. This specification will also facilitate comparison with the recursive least squares learning scheme considered previously. I suppose here, as in the preceding section, that the agents know that the steady state lies at $\bar{x} = 0$, but try to make better forecasts by ascertaining the presence of a possible trend in deviations from equilibrium. Their belief about the local dynamics of the system is described by $x_\tau = \beta\, x_{\tau-1} + \eta_\tau$, where β is an unknown parameter, and η_τ is a sequence of i.i.d. normal random variables with mean 0 and variance $\sigma^2_\eta > 0$. The agents are uncertain about the value of β : their prior at the beginning of period t is assumed to be represented by a normal random variable with mean β_t and variance σ^2_t , which is uncorrelated with η_t . The agents' forecast about $x_{t+1} = \beta\, x_t + \eta_{t+1} = \beta^2\, x_{t-1} + \beta\, \eta_t + \eta_{t+1}$, is then obtained (assuming here also risk neutrality) by taking the mathematical expectation of that random variable, conditionally on the information available at the beginning of period t (i.e. not using x_t). This yields the forecasting rule

$$(5.2) \qquad {}_t x^e_{t+1} = (\beta^2_t + \sigma^2_t)\, x_{t-1} \ .$$

The main difference with least squares learning that was studied earlier, is that the variance of the agents' prior now explicitly enters the forecasting rule. It is essentially through this channel that the agents' subjective uncertainty about the local dynamics will influence the motion of the system. Indeed, if we combine (5.1) and (5.2), (5.3) $$x_t = -b_0^{-1}[b_1 + a(\beta_t^2 + \sigma_t^2)]\, x_{t-1} \; .$$

In order to close the model, we need to specify how the agents revise their beliefs. The standard updating formula for Bayesian learning in the case of normal distributions as here (see e.g. Wickens (1982)), tells us that the agents' posterior at the end of period t (or the beginning of period t+1) about the unknown parameter β ,is a normal distribution with mean β_{t+1} and variance σ_{t+1}^2 , given by

(5.4) $$\sigma_\eta^2/\sigma_{t+1}^2 = x_{t-1}^2 + \sigma_\eta^2/\sigma_t^2 \;, \quad \beta_{t+1}\, \sigma_\eta^2/\sigma_{t+1}^2 = x_t\, x_{t-1} + \beta_t\, \sigma_\eta^2/\sigma_{t+1}^2 \; .$$

Comparison with the updating formula (4.8) for recursive least squares shows that the two learning schemes are in fact formally *identical* if one sets $\omega_t = \sigma_t/\sigma_\eta$. The only difference between the two schemes (for a given sequence of deviations x_t), is that the initial estimate β_0 (and ω_0) depends (discontinuously) on initial observations through (4.10) in recursive least squares, whereas the initial prior β_0 , σ_0 (or ω_0) is in principle an independent variable in the case of Bayesian updating (the two schemes become completely identical if the initial prior is assumed to be formed through (4.10) in the Bayesian case, a remark made in particular by Wickens (1982)).

This observation enables us to write down almost directly the equations governing the motion of the dynamics with learning, since it must be identical, through the relation $\omega_t = \sigma_t/\sigma_\eta$, to the dynamics (4.11) in the case of recursive least squares, the only difference being that the actual dynamics (5.3) involves here the variance of the agents' prior. So if $\omega_t = \sigma_t/\sigma_\eta$, and $m(\xi) = 1/(1 + \xi^2)$,

(5.5) $$\left|\begin{array}{ll} x_t & = -b_0^{-1}[b_1 + a(\beta_t^2 + \sigma_\eta^2\, \omega_t^2)]\, x_{t-1} = \Omega(\beta_t\, ,\, \omega_t)x_{t-1} \; , \\ \omega_{t+1}^2 & = m(\omega_t x_{t-1})\omega_t^2 \; , \\ \beta_{t+1} & = m(\omega_t x_{t-1})\beta_t + [1 - m(\omega_t x_{t-1})]\; \Omega(\beta_t, \omega_t) \; . \end{array}\right.$$

Let us assume for the sake of concreteness that $a \neq 0$, $b_0 \neq 0$ have opposite signs, so that the parabola representing Ω has its asymptotic branches going up. Then the curve representing $\Omega(\beta,\omega_t)$ for fixed ω_t, is obtained by translating upward Ω by the positive amount $- b_0^{-1} a \sigma_\eta^2 \omega_t^2 = - b_0^{-1} a \sigma_t^2$ (Fig. 3). By analogy with the case of recursive least squares learning, one should expect the dynamics with learning to be unstable, if the mean $\beta_0 > 0$ of the initial prior is larger than both perfect foresight roots (whenever they are real), and satisfies $\Omega(\beta_0) > 1$. Indeed, in that case, one has $|x_0| \geqq \Omega(\beta_0)|x_{-1}|$, while β_1, being a convex combination of β_0 and of $\Omega(\beta_0,\omega_0) \geqq \Omega(\beta_0)$, must satisfy $\beta_1 \geqq \beta_0$. By induction, the sequence β_t is non decreasing, while the sequence $|x_t|$ is increasing and diverges to infinity when $x_{-1} \neq 0$, since $|x_t| \geqq \Omega(\beta_t)|x_{t-1}| \geqq \Omega(\beta_0)|x_{t-1}|$.

The interesting feature here is that one should get also local divergence when the initial variance is large. Indeed, given $x_{-1} \neq 0$, one can choose $\omega_0 = \sigma_0/\sigma_\eta$ large enough so that $m(\omega_0 x_{-1})$ is small, while $\Omega(\beta,\omega_0)$ is so large (uniformly in β) as to ensure that β_1 is positive, satisfies $\Omega(\beta_1) > 1$, and exceeds all real perfect foresight roots, and this independently of the initial mean β_0. Since then $|x_0| > 0$, one is back to the previous case for $t = 1$ on, and one has local divergence.

<u>Fig. 3</u>

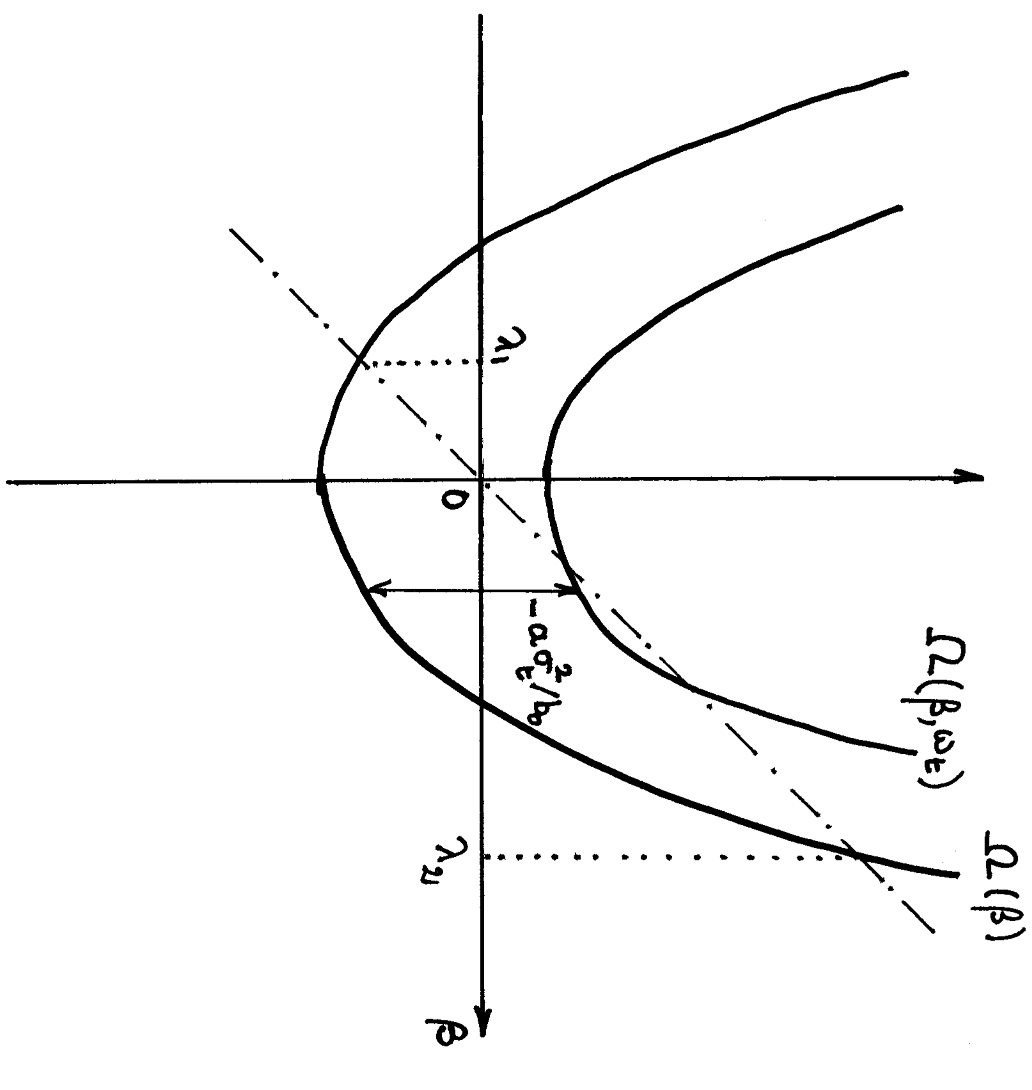
PROPOSITION 5.1. *Assume that* $a \neq 0$, $b_0 \neq 0$ *have opposite signs. Let* $\mu^* > 0$ *be larger than the two perfect foresight roots whenever they are real, satisfying* $\Omega(\mu^*) > 1$. *For any initial state* $(x_{-1},\omega_0,\beta_0)$ *with* $x_{-1} \neq 0$, *let* $(x_t,\omega_{t+1},\beta_{t+1})$ *be the corresponding sequence generated for* $t \geqq 0$ *by the dynamics with Bayesian learning* (5.5).

1. *If* $\beta_0 \geqq \mu^*$, *the sequence* β_t *is nondecreasing, whereas the sequence* $|x_t|$ *is increasing and diverges to infinity.*

2. *There exists* $\bar{\sigma}(\mu^*,x_{-1}) > 0$ *such that a large initial subjective uncertainty, i.e.* $\sigma_0 \geqq \bar{\sigma}(\mu^*,x_{-1})$, *implies* $\beta_1 \geqq \mu^*$ *and* $x_0 \neq 0$, *independently of* β_0. *In that case too, the sequence* $|x_t|$ *is increasing for* $t \geqq 0$ *and diverges to infinity.*

In order to study the range of initial priors (β_0,σ_0) leading to a stable dynamics with Bayesian learning, we can apply exactly the same method we used in the previous section for recursive least squares. The difference

Fig. 3



equation (5.5) has a whole set of stationary solutions of the form $\bar{x} = 0$, $\omega^* = \sigma^*/\sigma_\eta \geqq 0$, β^*. Linearizing (5.5) near such an equilibrium yields

$$(5.6) \qquad dx_t = \Omega(\beta^*,\omega^*)dx_{t-1} \; , \; d\omega_{t+1} = d\omega_t \; , \; d\beta_{t+1} = d\beta_t \; .$$

So, there exists a neighborhood V of $(0,\omega^*,\beta^*)$ such that if the initial conditions $(x_{-1},\omega_0,\beta_0)$ are in V , the sequence generated by the dynamics with Bayesian learning (5.5) will stay in V and will converge to some $(\bar{x} = 0 \; , \; \bar{\omega} \; , \; \bar{\beta})$ — where $(\bar{\omega},\bar{\beta})$ may differ from (ω^*,β^*) — if and only if $|\Omega(\beta^*,\omega^*)| = |b_0^{-1}[b_1 + a(\beta^{*2} + \sigma^{*2})]| < 1$. This will be the case in particular if $|b_0| > |b_1| + |a|(\beta^{*2} + \sigma^{*2})$, a condition that is the exact analogue of the sufficient stability conditions we obtained when studying smooth learning processes with a bounded memory (Propositions 2.2 and 3.2). The condition requires $|b_0| > |b_1|$, and one gets local stability if the parameter $|a|$ measuring the influence of expectations is small, or if $(\beta^{*2} + \sigma^{*2})$ is small, i.e. if the agents think the system is quite stable (β_0 is close to β^* and is thus small) and if that belief is held with enough confidence (σ_0 is close to σ^* and thus is small). As a final remark, although one gets local stability in such circumstances, the agents will typically never learn the value of a stable perfect foresight root.

PROPOSITION 5.2. *Consider a prior* (β^*,σ^*) *such that* $|b_0| > |b_1 + a(\beta^{*2} + \sigma^{*2})|$. *Then if the initial perturbation* x_{-1} *is sufficiently small, and if the agents' initial prior* (β_0,σ_0) *is close enough to* (β^*,σ^*), *the dynamics with Bayesian learning* (5.5) *will generate a sequence* $(x_t,\omega_{t+1} = \sigma_{t+1}/\sigma_\eta \; , \; \beta_{t+1})$ *that converges to some* $(0,\bar{\sigma}/\sigma_\eta,\bar{\beta})$.

The sufficient local stability condition will be satisfied in particular when $|b_0| > |b_1| + |a|(\beta^{*2} + \sigma^{*2})$, *i.e. if* $|b_0| > |b_1|$, *and if the influence of expectations is small, or if the agents initially think that the system is quite stable and if they hold that belief with enough confidence.*

6. CONCLUSION

A sort of general "*uncertainty principle*" seems to emerge from the findings reviewed here. If agents care not only about the position of an equilibrium with self-fulfilling expectations, but also about the dynamics nearby, and if they are rather uncertain about the stability of the system, so that they are ready on average to extrapolate a significant range of regularities (trends) out of past deviations from equilibrium, the actual temporary equilibrium dynamics with learning should be locally unstable (at least for an open set of initial perturbations, i.e. with "positive probability"), especially if the influence of expectations on the motion of the system is strong. Convergence for most (or almost all) initial perturbations (with "probability close or equal to 1") seems to occur only when expectations don't matter much, or when agents are assumed to ignore, either by lack of ability, or by conscious choice if they are fairly confident about the local stability of the system, all locally divergent tendencies that are present in past deviations from equilibrium.

It remains to see whether this sort of conclusion stands in more general systems involving, in particular, multidimensional state variables as well as multiple leads and lags. The analysis appears nevertheless to suggest already two possible avenues to explore.

Since it seems hopeless, short of putting overly strong restrictions on the agents' learning processes, to expect convergence to self-fulfilling expectations with "probability 1" (or even high "probability"), it would seem wise to redirect research in this area toward the evaluation, and if possible the classification, of the relative frequencies of divergence from (and of convergence to) self-fulfilling expectations equilibria, as a function of their local characteristics (perfect foresight roots) and of those of the learning processes (in particular their "eigenvalues"). Evans and Honkapohja (1993) have introduced new techniques, drawn from the work of Benveniste, Métivier and Priouret (1990) on adaptive algorithms, that do not employ "projection facilities" and might accordingly be useful in this respect.

On the other hand, the mere possibility that self-fulfilling expectations may be locally unstable raises interesting issues about the sort of long run equilibria generated by learning dynamics and their consistency with the agents' beliefs. Indeed, global nonlinearities, originating in the agents' expectations formation processes themselves, may keep the motion bounded, and lead to convergence to complex nonlinear "learning equilibria", along which forecasting errors would never vanish.

That such complex "learning equilibria" may occur has been already documented by some researchers (see Bullard (1994), Hommes (1991 a,b), Sorger (1994)), and is easily illustrated for instance in the simple formulation presented in Example 4.1. Let $b_1 x_{t-1} + b_0 x_t + a\, {}_t x^e_{t+1} = 0$ be a linear specification of the temporary equilibrium map, where x is a deviation from the equilibrium $\bar{x} = 0$, and assume to simplify matters that the agents' forecast is of the form ${}_t x^e_{t+1} = \beta_t^2 x_{t-1}$, with $\beta_t = x_{t-1}/x_{t-2}$. The resulting temporary equilibrium dynamics is then given by

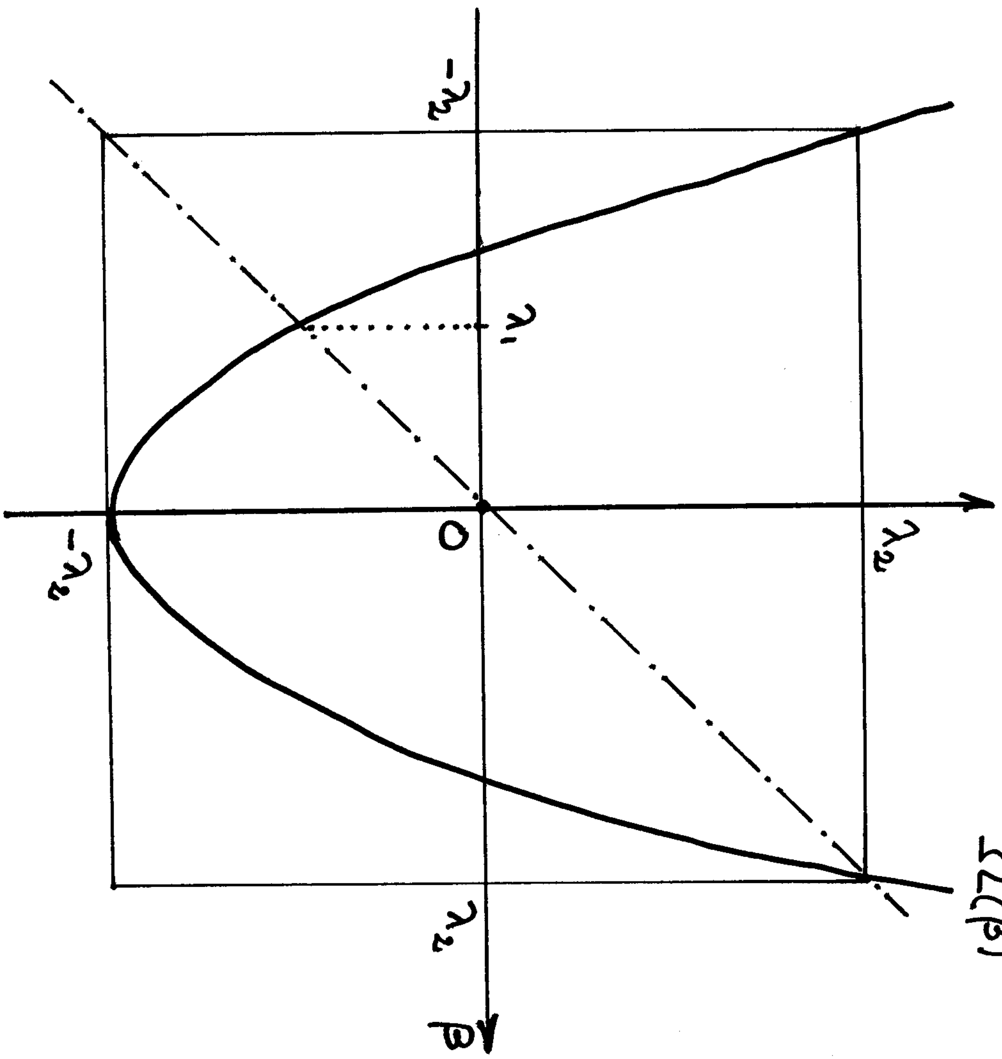
$$(6.1) \qquad x_t/x_{t-1} = -b_0^{-1}[b_1 + a\,(x_{t-1}/x_{t-2})^2] = \Omega(x_{t-1}/x_{t-2}),\ \beta_t = x_{t-1}/x_{t-2}.$$

If one considers the "resonant" case represented in Fig. 4, in which the two perfect foresight roots are real and satisfy $b_1/b_0 = \lambda_2$ (or $2\lambda_1 + \lambda_2 = 0$), it is known that the temporary equilibrium dynamics generated by (6.1), when restricted to the invariant interval $(-\lambda_2, \lambda_2)$, is indeed "chaotic". For almost any initial ratio x_{-1}/x_{-2} in that interval, the trajectory will be aperiodic and eventually fill in the whole interval $(-\lambda_2, \lambda_2)$.

<u>Fig. 4</u>

Such complex "learning equilibria" may be at first sight good candidates to explain why agents keep making significant and recurrent mistakes when trying to predict the fate of the socioeconomic systems they participate in. To be acceptable, however, the observed pattern along such "learning equilibria" should display some reasonable degree of consistency with the agents' beliefs. One might envision situations in which agents do believe (wrongly) that the world is relatively simple (e.g. linear) but subjected to random shocks, and in which the corresponding (deterministic) "learning equilbria" are complex enough to make the agents' forecasting mistakes "self-fulfilling" in a well defined sense. For instance, the agents might be assumed to have at their disposal a reasonably wide, but nevertheless limited, battery of statistical tests ("bounded rationality") which would not allow them to reject the hypothesis that their recurrent forecasting mistakes are attributable to random disturbances (the above example blatantly fails that sort of test. The agents' belief is, say, $x_\tau = \beta x_{\tau-1} + \eta_\tau$ or $x_\tau = (\beta + \eta_\tau)\, x_{\tau-1}$, where η_τ is white noise and β is an unknown but *constant* parameter. Yet their estimate β_t fluctuates wildly along the "learning equilibrium", and that observation should cause the

Fig. 4



agents to conclude that the class of models they believe in is wrong). That is not likely to be an easy test, and it is not quite clear to me at this stage whether such a program can actually generate operational results or is even feasible (for a first step, see Sorger (1994)). Yet progress on this front, if possible, might provide an interesting alternative to our current paradigms, which rely very heavily on extreme, and often criticized, rationality axioms.

REFERENCES

BARNETT, W., B. CORNET, C. d'ASPREMONT, J. GABSZEWICZ, and A. MAS-COLELL, Eds. (1991) : *Equilibrium Theory and Applications*. Cambridge, UK : Cambridge University Press.

BARNETT, W., J. GEWEKE and K. SHELL, Eds. (1989) : *Economic Complexity : Chaos, Sunspots, Bubbles and Nonlinearity*. Cambridge, UK : Cambridge University Press.

BENASSY, J.P., and M. BLAD (1989) : "On Learning and Rational Expectations in an Overlapping Generations Model", *Journal of Economic Dynamics and Control*, 13, 379-400.

BENEVISTE, A., M. METIVIER and P. PRIOURET (1990) : *Adaptive Algorithms and Stochastic Approximations*. New York : Springer-Verlag.

BLUME, L.E., and D. EASLEY (1984) : "Rational Expectations Equilibrium : An Alternative Approach", *Journal of Economic Theory*, 34, 116-129.

——— (1993), "Rational Expectations and Rational Learning", Cornell University.

BRAY, M. (1982) : "Learning, Estimation and the Stability of Rational Expectations", *Journal of Economic Theory*, 26, 318-339.

——— (1983) : "Convergence to Rational Expectations Equilibrium", in R. Frydman and E. Phelps, Eds., 1983.

BRAY, M., and D. KREPS (1987) : "Rational Learning and Rational Expectations", in *Arrow and the Ascent of Modern Economic Theory*, ed. by G. Feiwel. New York : New York University Press.

BRAY, M., and N. SAVIN (1986) : "Rational Expectations Equilibria, Learning and Model Specification", *Econometrica*, 54, 1129-1160.

BULLARD, J. (1994) : "Learning Equilibria", Federal Reserve Bank of St Louis, forthcoming in *Journal of Economic Theory*.

CAGAN, P. (1956) : "The Monetary Dynamics of Hyperinflation", in *Studies in the Quantity of Money*, ed. by M. Friedman. Chicago : University of Chicago Press.

CALVO, G. (1988) : "Passive Money and Idiosyncratic Expectations", University of Pennsylvania, Philadelphia.

CANNING, D. (1992) : "Average Behavior in Learning Models", *Journal of Economic Theory*, 57, 442-472.

CHAMPSAUR, P. (1983) : "On the Stability of Rational Expectations Equilibria", CORE DP 8324, Université Catholique de Louvain. French version in *Cahiers du Séminaire d'Econométrie*, 26 (1984), 47-65.

CHAMPSAUR, P., *et alii*, Eds. (1990) : *Essays in Honor of E. Malinvaud*. Cambridge, MA : The MIT Press.

CHATTERJI, S. (1994) : "Temporary Equilibrium Dynamics with Bayesian Learning", Indian Statistical Institute, New Delhi, forthcoming *Journal of Economic Theory*.

CRAWFORD, V. (1974) : "Learning the Optimal Strategy in Zero-Sum Games", *Econometrica*, 42, 885-891.

——————— (1985) : "Learning Behavior and Mixed Strategy Nash Equilibria", *Journal of Economic Behavior and Organization*, 6, 69-78.

——————— (1989) : "Learning and Mixed Strategy Equilibria in Evolutionary Games", *Journal of Theoretical Biology*, 140, 537-550.

CYERT, R.M., and M.H. De GROOT (1974) : "Rational Expectations and Bayesian Analysis", *Journal of Political Economy*, 82, 521-536.

De CANIO, S. (1979) : "Rational Expectations and Learning from Experience", *Quarterly Journal of Economics*, 93, 47-57.

EVANS, G. (1983) : "The Stability of Rational Expectations in Macroeconomic Models", in R. Frydman and E. Phelps, Eds., 1983.

————— (1985) : "Expectational Stability and the Multiple Equilibria Problem in Linear Rational Expectations Models", *Quarterly Journal of Economics*, 100, 1217-1233.

————— (1989) : "The Fragility of Sunspots and Bubbles", *Journal of Monetary Economics*, 23, 297-318.

EVANS, G., and S. HONKAPOHJA (1992) : "On the Robustness of Bubbles in Linear RE Models", *International Economic Review*, 33, 1-14.

—————————————— (1994) : "On the Local Stability of Sunspot Equilibria under Adaptive Learning Rules", *Journal of Economic Theory*, 64, 142-161.

—————————————— (1993) : "Economic Dynamics with Learning : New Stability Results", University of Edinburgh and University of Helsinki.

FELDMAN, M. (1987) : "An Example of Convergence to Rational Expectations with Heterogenous Beliefs", *International Economic Review*, 28, 635-650.

FOURGEAUD, C. GOURIEROUX, and J. PRADEL (1986) : "Learning Procedures and Convergence to Rationality", *Econometrica*, 54, 854-868.

FRYDMAN, R. (1982) : "Toward an Understanding of Market Processes", *American Economic Review*, 72, 652-658.

FRYDMAN, R., and E. PHELPS, Eds. (1983) : *Individual Forecasting and Aggregate Outcomes*. Cambridge, UK : Cambridge University Press.

FUCHS, G. (1976) : "Asymptotic Stability of Stationary Temporary Equilibria and Changes in Expectations", *Journal of Economic Theory*, 13, 201-216.

——— (1977a) : "Formation of Expectations : A Model in Temporary General Equilibrium Theory", *Journal of Mathematical Economics*, 4, 167-188.

——— (1977b) : "Dynamic Role and Evolution of Expectations", in *Systèmes Dynamiques et Modèles Economiques*. Paris : Editions du CNRS.

——— (1979a) : "Dynamics of Expectations in Temporary General Equilibrium Theory", *Journal of Mathematical Economics*, 6, 229-252.

——— (1979b) : "Is Error Learning Behavior Stabilizing ?", *Journal of Economic Theory*, 20, 300-317. Reprinted in J.M. Grandmont, Ed., 1988a.

FUCHS, G., and G. LAROQUE (1976) : "Dynamics of Temporary Equilibria and Expectations", *Econometrica*, 44, 1157-1178. Reprinted in J.M. Grandmont, Ed., 1988a.

FUDENBERG, D., and D. KREPS (1988) : "A Theory of Learning, Experimentation and Equilibrium in Games", Stanford University.

FUDENBERG, D., and D. LEVINE (1993a) : "Steady State Learning and Nash Equilibrium", *Econometrica*, 61, 541-573.

——— (1993b) : "Self-Confirming Equilibrium", *Econometrica*, 61, 523-545.

GOTTFRIES, N. (1985) : "Multiple Perfect Foresight Equilibria and Convergence of Learning Processes", *Journal of Money, Credit and Banking*, 17, 111-117.

GOURIEROUX, C., J.J. LAFFONT, and A. MONFORT (1983) : "Révision Adaptative des Anticipations et Convergence vers les Anticipations Rationnelles", *Economie Appliquée*, 36, 9-26.

GRANDMONT, J.M. (1985) : "On Endogenous Competitive Business Cycles", *Econometrica*, 53, 995-1045.

———, Ed. (1988a) : *Temporary Equilibrium. Selected Readings*. San Diego : Academic Press.

——— (1988b) : "Nonlinear Difference Equations, Bifurcations and Chaos : An Introduction", CEPREMAP, Paris. Also available in Stanford

Lecture Note Series, Stanford University.

GRANDMONT, J.M., and G. LAROQUE (1986) : "Stability of Cycles and Expectations", *Journal of Economic Theory*, 40, 138-151.

——————— (1990) : "Stability, Expectations and Predetermined Variables", in P. Champsaur *et alii*, Eds., 1990.

——————— (1991) : "Economic Dynamics with Learning : Some Instability Examples", in W. Barnett *et alii*, Eds., 1991.

GUESNERIE, R. (1993) : "Successes and Failures in Coordinating Expectations", *European Economic Review*, 37, 243-268.

GUESNERIE, R., and M. WOODFORD (1991) : "Stability of Cycles with Adaptive Learning Rules", in W. Barnett *et alii*, Eds., 1991.

HARSANYI, J., and R. SELTEN (1988) : *A General Theory of Equilibrium Selection in Games*. Cambridge, MA : MIT Press.

HOMMES, C. (1991a) : "Adaptive Learning and Roads to Chaos", *Economic Letters*, 36, 127-132.

——— (1991b) : *Chaotic Dynamics in Economic Models. Some Simple Case Studies*, Groningen Theses in Economics Management and Organization. Groningen : Wolters-Noordhoff.

HOWITT, P. (1990) : "Interest Rate Control and Nonconvergence to Rational Expectations", *Journal of Political Economy*, 100, 776-800.

INGRAO, B., and G. ISRAEL (1990) : *The Invisible Hand. Economic Equilibrium in the History of Science*. Cambridge, MA : MIT Press.

JORDAN, J. (1985) : "Learning Rational Expectations : The Finite State Case", *Journal of Economic Theory*, 36, 257-276.

——— (1991) : "Bayesian Learning in Normal Form Games", *Games and Economic Behavior*, 3, 60-81.

——— (1992) : "The Exponential Convergence of Bayesian Learning in Normal Form Games", *Games and Economic Behavior*, 4, 202-217.

KALAI, E., and E. LEHRER (1993a) : "Rational Learning Leads to Nash Equilibrium", *Econometrica*, 61, 1019-1045.

——————— (1993b) : "Bounded Learning Leads to Correlated Equilibrium", Northwestern University.

KANDORI, M., G. MAILATH, and R. ROB (1993) : "Learning, Mutation and Long Run Equilibria in Games", *Econometrica*, 61, 29-56.

LUCAS, R.E., Jr. (1972) : "Expectations and the Neutrality of Money", *Journal of Economic Theory*, 4, 103-124.

———— (1986), "Adaptive Behavior and Economic Theory", *Journal of*

Business, 59, 401-426.

MARCET, A., and T. SARGENT (1988) : "The Fate of Systems with Adaptive Expectations", *American Economic Review, Papers and Proceedings*, 78, 168-172.

——————————— (1989a) : "Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models", *Journal of Economic Theory*, 48, 337-368.

——————————— (1989b) : "Convergence of Least Squares Learning in Environments with Hidden State Variables and Private Information", *Journal of Political Economy*, 97, 1306-1322.

——————————— (1989c) : "Least Squares Learning and the Dynamics of Hyperinflation", in W. Barnett *et alii*, Eds., 1989.

MILGROM, P., and J. ROBERTS (1991) : "Adaptive and Sophisticated Learning in Normal Form Games", *Games and Economic Behavior*, 3, 82-100.

MINCER, J. (1969) : "Models of Adaptive Forecasting", in *Economic Forecasts and Expectations*, ed. by J. Mincer. New York : National Bureau of Economic Research.

MIYASAWA, K. (1961) : "On the Convergence of the Learning Process in a 2x2 Non-Zero Sum Two-Person Game", Economic Research Program, Res. Mem. 33, Princeton University.

NERLOVE, M. (1958) : "Adaptive Expectations and Cobweb Phenomena", *Quarterly Journal of Economics*, 73, 227-240.

NYARKO, Y. (1991) : "Learning in Mis-Specified Models and the Possibility of Cycles", *Journal of Economic Theory*, 55, 416-427.

——————— (1992) : "Bayesian Learning in Repeated Games Leads to Correlated Equilibria", New York University.

SELTEN, R. (1988) : "Adaptive Learning in Two-Person Games", University of Bonn.

SHAPLEY, L. (1964) : "Some Topics in Two-Person Games", in *Advances in Game Theory*, ed. by M. Dresher, L. Shapley and A. Tucker. Princeton : Princeton University Press.

SORGER, G. (1994) : "Boundedly Rational Chaos : An Example of a Self-Fulfilling Mistake", University of Vienna.

TILLMAN, G. (1983) : "Stability in a Pure Consumption Loan Model", *Journal of Economic Theory*, 30, 315-329. Reprinted in J.M. Grandmont, Ed., 1988a.

——— (1985) : "Existence and Stability of Rational Expectations Equilibria in a Simple Overlapping Generations Model", *Journal of Economic Theory*, 36, 333-351.

TOWNSEND, R. (1978) : "Market Anticipations, Rational Expectations and Bayesian Analysis", *International Economic Review*, 19, 481-494.

VIVES, X. (1993) : "How Fast Do Rational Agents Learn", *Review of Economic Studies*, 60, 329-348.

WEIBULL, J. (1994) : *Evolutionary Game Theory*, University of Stockholm.

WICKENS, M. (1982) : "Are Time Series Forecasts Self-Falsifying ? A Study of Expectations Formation", University of Southampton.

WOODFORD, M. (1990) : "Learning to Believe in Sunspots", *Econometrica*, 58, 277-307.

FOOTNOTES

* The material of this paper was in part presented as the Presidential address at the World Meetings of the Econometric Society, Barcelona, 1990. I owe a specific debt to Guy Laroque : the results reviewed here are in part an outgrowth of joint work undertaken in the eighties. I had stimulating exchanges with many colleagues while this project was "in progress", in particular with Philippe Aghion, Jean-Pascal Benassy, Alan Blinder, Christian Gourieroux, Bruno Jullien, Martine Quinzii and Jorgen Weibull.

** CNRS and CEPREMAP, 142, rue du Chevaleret, 75013 PARIS.

[1] Front quotation of Chap. 1 in Ingrao and Israel (1990).

[2] The approach used here is thus distinct from socalled "rational" learning, as considered for instance by Blume and Easley (1984, 1993), Bray and Kreps (1987), Feldman (1987) or Townsend (1978), which are in effect self-fulfilling expectations models where agents are initially incompletely informed about some parameter(s), but may eventually learn it (them) through Bayesian updating. By contrast, in what follows, the agents' beliefs are typically incorrect out of a self-fulfilling expectations equilibrium.

[3] Some of the results (in particular Proposition 2.1 below) can be generalized to a multidimensional framework, by adapting the methods of Grandmont and Laroque (1990).

[4] When applied to a framework with a multidimensional state variable, along the lines of Grandmont and Laroque (1990), the parity argument depends upon whether the rank of the matrix of partial derivatives of the temporary equilibrium map with respect to expectations, is odd or even.

[5] The interval $[\mu_1,\mu_2]$ need not be very large if expectations matter a lot. For instance, let the expectation function ψ be independent of x_t , so that $c_0 = 0$. If $|a|$ is large, Q_W is dominated by the term $a \Sigma_1^L c_j z^{L-j}$, which is a polynomial of degree L-1. One of the roots of Q_W has a large modulus, and is therefore unstable.

[6] Less symmetric statements can be obtained by application of (2.10) to cover, for instance, cases in which agents have prior information about the sign of the perfect foresight roots. Such statements involve the comparison of the sum $\sigma_\lambda = - b_0 a^{-1}$ of the two perfect foresight roots, and of the sum $\sigma_\mu = c_0$ of the local eigenvalues of the expectation function ψ. Then if $\sigma_\mu < \sigma_\lambda$, and if ψ has an eigenvalue $\mu_2 > 0$ that is greater than all perfect foresight roots that are real, the polynomial Q_W has a real root $r > \mu_2$, so that the actual dynamics with learning is unstable whenever $\mu_2 \geqq 1$. Or if $\sigma_\lambda < \sigma_\mu$, and if ψ has an eigenvalue $\mu_1 < 0$ that is less than all perfect foresight roots that are real, Q_W has a real root $r < \mu_1$, which is thus unstable when $\mu_1 \leqq -1$. These two statements imply that in Proposition 2.1, the root r of Q_W that lies outside $[\mu_1,\mu_2]$ is less than μ_1 if $\sigma_\lambda < \sigma_\mu$, greater than μ_2 if $\sigma_\mu < \sigma_\lambda$.

[7] The following simple argument shows what happens in that sort of configuration. If a and/or the coefficients c_j are small, the polynomial Q_W essentially behaves as $b_1 z^{L-1} + b_0 z^L = z^{L-1}(b_1 + b_0 z)$, so it has L-1 roots with a small modulus, while the last one is close to $-b_1/b_0$.

[8] An early example can be found in Fuchs and Laroque (1976, Corollary 1) when the temporary equilibrium map does not depend on x_{t-1}. A similar argument appears in Grandmont (1985), where it is argued, in a simple overlapping generations OLG model, that forecasting rules of this sort may make the Golden Rule (or cycles) stable in the temporary equilibrium dynamics with learning when it is unstable in the forward perfect foresight dynamics (see also Grandmont and Laroque (1986)). Essentially the same argument reappears in Lucas (1986), following the seminal contribution of Bray (1982), in a simple (stochastic) OLG model when learning involves an unbounded memory, as well as in Marcet and Sargent (1989c).

[9] Error learning has been considered by Fuchs (1979b), Guesnerie and Woodford (1991), among others.

[10] The statements made in footnote 6 apply equally well here. If $\sigma_\mu < \sigma_\lambda$ and if the expectation function ψ has an eigenvalue $\mu_2 > 0$ that is greater than all perfect foresight roots that are real, and larger than all real roots of the learning polynomial Q_e, then the polynomial Q_W has a real root $r > \mu_2$, which is thus unstable whenever $\mu_2 \geqq 1$. A symmetric statement holds for the case $\sigma_\lambda < \sigma_\mu$. So in Proposition 3.1, the root r of Q_W that lies outside the interval $[\mu_1,\mu_2]$ is less than μ_1 if $\sigma_\lambda < \sigma_\mu$, greater than μ_2 if $\sigma_\mu < \sigma_\lambda$.

[11] It is important to emphasize that such discontinuities arise only when agents try to extrapolate a continuum of trends out of *deviations* $x_t - \bar{x}$ from equilibrium. By contrast, it is easy to generate smooth forecasting rules that extrapolate a continuum of trends appearing in the *levels* of the state variable. For instance, assume that agents believe the law of motion to be $x_\tau = \beta\, x_{\tau-1} + \eta_\tau$, in which η_τ in white noise, that their estimate β_t of the unknown parameter β is made at the beginning of period t through ordinary least squares using the observations $x_{t-1},\ldots,x_{t-L}$, and that their forecast is then obtained by "iterating twice" the law of motion they believe in, i.e.

$$ {}_t x^e_{t+1} = \beta^2_t\, x_{t-1} = (\Sigma^L_2\, x_{t-j+1}\, x_{t-j} / \Sigma^L_2\, x^2_{t-j})^2\, x_{t-1} \,. $$

That forecasting rule is smooth at a stationary state $\bar{x} \neq 0$. By construction, it extrapolates correctly *any* real trend present in past *levels* x_{t-j} of the state variable. Yet it can only extrapolate a discrete (and severely restricted) set of trends (frequencies) out of past *deviations* $x_{t-j} - \bar{x}$ from equilibrium. The partial derivatives at $\bar{x} \neq 0$ are $c_0 = 0$, $c_1 = 1 + 1/(L-1)$, $c_2 = \ldots = c_{L-1} = 0$, $c_L = -\, 2/(L-1)$. The characteristic polynomial Q_W has a root equal to 1 (and another equal to -1 when L is odd). It is not difficult to verify (by contradiction) that for L large, all *local* eigenvalues of the above forecasting rule have a modulus that is close to 1, a very strong restriction indeed.

[12] The statement is adapted from Grandmont and Laroque (1991, Part 1 of Proposition 2). Forerunners of this type of instability results can be found in Champsaur (1983) or Benassy and Blad (1989).

[13] The "asymmetric" statement in Proposition 4.2 is the analogue of the remark made in footnote 6 for the case $\sigma_\mu < \sigma_\lambda$, since here the sum of the two perfect foresight roots is positive while the forecasting rule does not depend on x_t (which would imply $\sigma_\mu = c_0 = 0$ in the smooth case).

[14] The results reviewed in this section are adapted from S.Chatterji (1994).