View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Publications at Bielefeld University

Learning to Play Best Response in Duopoly Games

By Jan Wenzelburger*

Fakultät für Wirtschaftswissenschaften

Universität Bielefeld, Postfach 100 131

D-33501 Bielefeld, Germany

jwenzelb@wiwi.uni-bielefeld.de

Discussion Paper No. 506

Abstract

We consider a quantity-setting duopoly market where firms lack perfect knowledge of the market demand function. They use estimated and therefore misspecified demand functions instead and determine their optimal strategies from the corresponding subjective payoff functions. The central issue of this paper is the question under which conditions a firm can learn the true demand function as well as the response function of its competitor from repeated estimations of historical market data. As soon as estimation errors are negligible, the correct response function is known and a firm is able to play best response in the usual game theoretic sense.

Keywords: Learning, bounded rationality, heterogeneity, duopoly games.

JEL Classification: C61, C62, D83, D84

First version: Jan. 2003, this version: Apr. 2004. Forthcoming International Game Theory Review.

---

*Acknowledgement. I am indebted to Hans Gersbach, Thorsten Pampel, and to two anonymous referees for valuable comments.

1 Introduction

Recently more and more game theorists and economists abandon the rather strong assumption that agents have perfect knowledge about the economic environment in which they live. Instead the focus has been turned to scenarios in which agents' views of the world may be erroneous. An open issue is what a plausible outcome of such scenarios with imperfect knowledge could be.

The literature on learning in games has investigated many situations in which economic agents have incomplete or erroneous information about the environment in which the game takes place, e.g., see Kirman & Salmon (1995), Marimon (1997), Fudenberg & Levine (1998), or Blume & Easley (1993, 1998). If an equilibrium of a game is thought of an outcome of a mental process that takes place in virtual time, the question whether the outcome is the same when such a process is made explicit naturally arises. Already Kirman (1975) provided an example of a duopoly game showing that, in general, this will not be the case. Several equilibrium concepts have been introduced to describe the long-run outcome of games in which agents lack full information, such as self-confirming equilibria or subjective equilibria, see Fudenberg & Levine (1993) and Kalai & Lehrer (1993, 1995).

Under the assumption that agents lack full information, one is left to specify first how informed agents are and second how agents try to retrieve the missing information. From an abstract point of view it is clear that possible long-run outcomes of the corresponding subjective game depend heavily on the underlying behavioral assumptions. The specification of a learning scheme according to which agents update their beliefs will thereby play a crucial role. For duopoly games, it has long been recognized that simple updating rules may lead to complex behavior of the system that may or may not converge to a desired equilibrium, e.g., see Rand (1978) or Dana & Montrucchio (1986) for an early contribution. Recent contributions have extensively investigated the resulting so-called learning dynamics, cf. Léonard & Nishimura (1999), Bischi, Mammana & Gardini (2000), Bischi, Chiarella & Kopel (2002) and others.

On the other hand, however, a common and still prevalent paradigm is that economic agents learn from errors they have made in the past and that they will try to eliminate all systematic forecast errors. From this perspective, the plausibility of behavioral assumptions concerning agents' capability to learn has to be treated with care. Any equilibrium concept based on incomplete information and subjective beliefs depends crucially these assumptions which for this reason should be justified.

The question arises to what extent a learning scheme enables an agent to eliminate all errors that are systematic from her perspective and what an agent will do if she finds out that her learning scheme fails to do so. Conceptually this amounts to defining a goal for a learning scheme before specifying the learning scheme itself. Depending on a presumed level of rationality, plausibility of a behavioral assumption may then be inferred from the success of a particular learning scheme to achieve the prescribed goal.

For the case of duopoly games whose evolution is driven by non-linear maps the question of which learning schemes find a desired subjective equilibrium in a systematic way remains to a large extent unresolved. Following Kirman (1975), Dana & Montrucchio (1986), Léonard & Nishimura (1999), Bischi, Chiarella & Kopel (2002) and others we consider a Cournot duopoly of two quantity setting firms that lack perfect knowledge of their rival's cost function and of the market demand function. Time is discrete and firms repeatedly play a subjective game with a one-period planning horizon. They use estimated and therefore misspecified demand functions and determine their optimal strategies from the corresponding subjective payoff functions. The central issue of this paper is under which conditions a firm can learn to play best response, if outputs of both firms and prices are observable quantities.

In order to be successful, such a firm will have to learn the true market demand function as well as the response function of its competitor through repeated estimations from historical market data. The present paper investigates the simple situation in which one firm behaves like a naive updater while the other firm is assumed to be sophisticated. For the latter firm we propose a learning scheme with two separate estimation routines that are carried out simultaneously. With the first routine the unknown market demand is estimated in each trading period. With the second one the competitor's response function and/or a self-confirming beliefs is estimated. While the first routine is a standard estimation problem, it is shown that finding a self-confirming beliefs equilibrium amounts to finding the zero of an unknown function. The whole response function will be estimated using spline approximations which can be made arbitrarily precise. As soon as estimation errors are negligible, the competitor's correct response function is known and a firm is able to play best response in the usual game theoretic sense.

2 Duopoly as a subjective game

Consider a discrete-time Cournot duopoly game, in which two firms indexed by $i = 1, 2$ offer homogeneous products on a common market. If p denotes the goods price and $q = q^{(1)} + q^{(2)}$ the industry output of the two firms, the aggregate market demand function D and its inverse f are given by

$$q = D(p),\ p \in \mathbb{R}_+ \quad \text{and} \quad p = f(q) := D^{-1}(q),\ q \in \mathbb{R}_+,$$

respectively. Firms do not know the exact specification of the inverse demand function f and instead use *subjective* estimates for f, given by the relationships

$$p = \widehat{f}^{(i)}(q), \quad i = 1, 2. \tag{1}$$

Moreover, firms have incomplete knowledge of the profit function of their rival. At the beginning of each trading period, they estimate the produced quantity of the rival firm. For an arbitrary period t, let $\widehat{q}_t^{(1)}$ ($\widehat{q}_t^{(2)}$) be the estimate of firm 1 (firm 2) for the output $q_t^{(2)}$ ($q_t^{(1)}$) of firm 2 (firm 1) in that period. The anticipated profit of firm 1 has to be based on two estimates, namely $\widehat{f}_t^{(1)}$ for the true inverse demand function f and $\widehat{q}_t^{(1)}$ for firm 2's output $q_t^{(2)}$ of period t. If $C^{(1)}(q^{(1)}, q^{(2)})$ denotes the cost function of firm 1, the *anticipated profit function* of firm 1 is given by

$$\Pi^{(1)}(q^{(1)}, \widehat{q}_t^{(1)}, \widehat{f}_t^{(1)}) := q^{(1)}\, \widehat{f}_t^{(1)}(q^{(1)} + \widehat{q}_t^{(1)}) - C^{(1)}(q^{(1)}, \widehat{q}_t^{(1)}). \tag{2}$$

Following Kopel (1996) and Bischi & Lamantia (2002), we allow for possible externalities in the cost functions of the firms.

Let the strategy space of firm 1 be given by the compact interval $[q_{\min}^{(1)}, q_{\max}^{(1)}]$ and the strategy space of firm 2 be given by the compact interval $[q_{\min}^{(2)}, q_{\max}^{(2)}]$. Based on an estimate $\widehat{f}_t^{(1)}$ for the market demand, the *subjective best response* of firm 1 is a function

$$R^{(1)}(\cdot, \widehat{f}_t^{(1)}) : [\widehat{q}_{\min}^{(1)}, \widehat{q}_{\max}^{(1)}] \to [q_{\min}^{(1)}, q_{\max}^{(1)}], \quad \widehat{q}^{(1)} \mapsto R^{(1)}(\widehat{q}^{(1)}, \widehat{f}_t^{(1)}),$$

where

$$R^{(1)}(\widehat{q}^{(1)}, \widehat{f}_t^{(1)}) := \operatorname{argmax} \left\{ \Pi^{(1)}(q^{(1)}, \widehat{q}^{(1)}, \widehat{f}_t^{(1)}) \ : \ q^{(1)} \in [q_{\min}^{(1)}, q_{\max}^{(1)}] \right\}. \tag{3}$$

We assume that the compact interval $[\widehat{q}_{\min}^{(1)}, \widehat{q}_{\max}^{(1)}]$ of possible beliefs of firm 1 may be chosen large enough such that it includes firm 2's strategy space assumed to be unknown to firm 1.

Similarly, given the cost function $C^{(2)}(q^{(2)}, q^{(1)})$ of firm 2, the *anticipated profit function* of firm 2 is

$$\Pi^{(2)}(q^{(2)}, \widehat{q}^{(2)}, \widehat{f}_t^{(2)}) := q^{(2)}\, \widehat{f}_t^{(2)}(\widehat{q}^{(2)} + q^{(2)}) - C^{(2)}(q^{(2)}, \widehat{q}^{(2)}). \tag{4}$$

If the compact interval $[\widehat{q}^{(2)}_{\min}, \widehat{q}^{(2)}_{\max}]$ describes the beliefs firm 2 will choose from, then based on the estimate $\widehat{f}^{(2)}_t$, the *subjective best response* of firm 2 is

$$R^{(2)}(\cdot, \widehat{f}^{(2)}_t) : [\widehat{q}^{(2)}_{\min}, \widehat{q}^{(2)}_{\max}] \to [q^{(2)}_{\min}, q^{(2)}_{\max}], \quad \widehat{q}^{(2)} \mapsto R^{(2)}(\widehat{q}^{(2)}, \widehat{f}^{(2)}_t),$$

where

$$R^{(2)}(\widehat{q}^{(2)}, \widehat{f}^{(2)}_t) := \operatorname{argmax} \left\{ \Pi^{(2)}(q^{(2)}, \widehat{q}^{(2)}, \widehat{f}^{(2)}_t) \; : \; q^{(2)} \in [q^{(2)}_{\min}, q^{(2)}_{\max}] \right\}. \tag{5}$$

As above, we assume that the compact interval $[\widehat{q}^{(2)}_{\min}, \widehat{q}^{(2)}_{\max}]$ of possible beliefs of firm 2 may be chosen large enough such that it includes firm 1's strategy space assumed to be unknown to firm 2.

The time-line of actions within a typical period t is now as follows. At the beginning of period t, each firm $i = 1, 2$ forms estimates for the inverse demand function $\widehat{f}^{(i)}_t$ and for its rival's output $\widehat{q}^{(i)}_t$, $i = 1, 2$, respectively. After forming these estimates, they both maximize anticipated profit and produce their subjectively optimal quantities as given by the respective subjective response functions. Thus

$$\begin{cases} q^{(1)}_t &= R^{(1)}(\widehat{q}^{(1)}_t, \widehat{f}^{(1)}_t) \\ q^{(2)}_t &= R^{(2)}(\widehat{q}^{(2)}_t, \widehat{f}^{(2)}_t) \end{cases} \tag{6}$$

is the actual output of the economy in period t. Prices in period t are then determined by

$$p_t = f(q^{(1)}_t + q^{(2)}_t). \tag{7}$$

Notice that equations (6) and (7) are determined by subjective beliefs only. In particular, if none of the firms updates their estimates, then output and prices are constant over time. We will refer to eqs. (6) and (7) as the *economic law* of the subjective duopoly game. The time-line of the trading structure is displayed in Figure 1.

Figure 1 about here.

A special situation occurs if one or both of the firms are able to perfectly predict the rivals output and the (inverse) market demand function of the economy. Suppose firm 1 is able to do so. Then

$$\widehat{q}^{(1)}_t = q^{(2)}_t \quad \text{and} \quad \widehat{f}^{(1)}_t = f$$

and the subjective best response function (3) coincides with the *best response function* of firm 1. Similarly, if firm 2's estimates satisfy

$$\widehat{q}^{(2)}_t = q^{(1)}_t \quad \text{and} \quad \widehat{f}^{(2)}_t = f,$$

then she is able to play best response as well. A *Nash equilibrium* of the economy is then given by a solution

$$(q_{\star\star}^{(1)}, q_{\star\star}^{(2)}) \in [q_{\min}^{(1)}, q_{\max}^{(1)}] \times [q_{\min}^{(2)}, q_{\max}^{(2)}] \tag{8}$$

to the equations

$$\begin{cases} q^{(1)} &= R^{(1)}(q^{(2)}, f), \\ q^{(2)} &= R^{(2)}(q^{(1)}, f). \end{cases}$$

Concrete and worked-out examples that fit into the present formulation may be found in Bischi & Lamantia (2002) or Bischi, Chiarella & Kopel (2002).

3 Information, forecasting, and dynamics

As pointed out in the last section, the evolution of outputs and prices of the duopoly game described by (6) and (7) is exclusively driven by the forecasting technology that firms use. We therefore have to discuss the information set on which firms may base their subjective evaluation of the future. We distinguish between two kinds of information, *public* information and *private* information. For an arbitrary period t, we assume that past prices and realized outputs are observable, such that

$$\{q_s^{(1)}, q_s^{(2)}, p_s\}_{s=0}^{t-1} \tag{9}$$

constitutes the set of publicly observable quantities in period t. The private information of a firm consists of its cost function and past estimates for the rival's output and for the market demand function. That is, apart from the knowledge of its cost function firm i's private information consists of the set

$$\{\widehat{q}_s^{(i)}, \widehat{f}_s^{(i)}\}_{s=0}^{t-1}, \quad i = 1, 2. \tag{10}$$

The information set on which firm i will base its decision in period t then is

$$I_t^{(i)} := \{q_s^{(1)}, q_s^{(2)}, p_s, \widehat{q}_s^{(i)}, \widehat{f}_s^{(i)}\}_{s=0}^{t-1}, \quad i = 1, 2. \tag{11}$$

Forecasting rules for firm i in period t are formally functions, say $\psi_t^{(i)}$ and $\phi_t^{(i)}$, that map historical data known at date t into forecasts

$$\begin{cases} \widehat{q}_t^{(i)} &= \psi_t^{(i)}(I_t^{(i)}), \\ \widehat{f}_t^{(i)} &= \phi_t^{(i)}(I_t^{(i)}). \end{cases} \tag{12}$$

Inserting the forecasting rules (12) into the economic law (6) yields a (non-autonomous) dynamical system

$$\begin{cases} q_t^{(1)} &= R^{(1)}\big(\psi_t^{(1)}(I_t^{(1)}), \phi_t^{(1)}(I_t^{(1)})\big) \\ q_t^{(2)} &= R^{(2)}\big(\psi_t^{(2)}(I_t^{(2)}), \phi_t^{(2)}(I_t^{(2)})\big) \end{cases} \tag{13}$$

that describes the evolution of the Cournot duopoly game. Many contributions have investigated the dynamics of the case with forecasting rules that are independent of time, e.g., see Rand (1978), Dana & Montrucchio (1986), Léonard & Nishimura (1999), Bischi, Mammana & Gardini (2000), Bischi, Chiarella & Kopel (2002) and others. Observe, however, that as in the classical cobweb model, the evolution of the outputs is exclusively driven by subjective beliefs of the two firms and thus by a behavioral assumption on the rationality of firms.

4 Dynamics with perfect knowledge

Consider the case with a *'naive'* and a *'sophisticated'* firm. Let firm 1 be the naive one that never updates her subjective demand function and has naive expectations concerning the future output of firm 2. The two forecasting functions of firm 1 at any date $t \in \mathbb{N}$ then take the simple form

$$(14) \qquad \begin{cases} \widehat{q}_t^{(1)} &= \psi_t^{(1)}(I_t^{(1)}) := q_{t-1}^{(2)}, \\ \widehat{f}_t^{(1)} &= \phi_t^{(1)}(I_t^{(1)}) \equiv \widehat{f}^{(1)}. \end{cases}$$

Notice that both forecasting rules in (14) are independent of time. Accordingly, the subjective best response function (3) of firm 1 becomes

$$(15) \qquad q_t^{(1)} = R^{(1)}(q_{t-1}^{(2)}, \widehat{f}^{(1)}).$$

The response function (15) is similar to the one used by Dana & Montrucchio (1986), Léonard & Nishimura (1999), Bischi, Chiarella & Kopel (2002) and others.

Consider now the sophisticated firm 2. Suppose for the moment that firm 2 has no forecasting errors on the demand function f such that

$$(16) \qquad \widehat{f}_t^{(2)} = \phi_t(I_t^{(2)}) \equiv f.$$

Assuming that firms use the forecasting rules (14) and (16), the economic law (6) takes the form

$$(17) \qquad \begin{cases} q_t^{(1)} &= R^{(1)}(q_{t-1}^{(2)}, \widehat{f}^{(1)}), \\ q_t^{(2)} &= R^{(2)}(\widehat{q}_t^{(2)}, f). \end{cases}$$

The forecast errors ζ_t of firm 2 are given by the relationship

$$(18) \qquad \zeta_t = q_t^{(1)} - \widehat{q}_t^{(2)} = R^{(1)}(q_{t-1}^{(2)}, \widehat{f}^{(1)}) - \widehat{q}_t^{(2)}, \quad (q_{t-1}^{(2)}, \widehat{q}_t^{(2)}) \in \mathbb{R}_+^2.$$

The *forecast error function* (18) confirms the intuition that in order to have perfect foresight on the rival's output for all times t, firm 2 has to know the subjective response function of firm 1. The forecasting rule that provides firm 2 with these correct forecasts is formally given by the function

$$\psi_\star^{(2)} : [q_{\min}^{(2)}, q_{\max}^{(2)}] \to \mathbb{R}_+,$$

where

$$\widehat{q}_t^{(2)} = \psi_\star^{(2)}(q_{t-1}^{(2)}) := R^{(1)}(q_{t-1}^{(2)}, \widehat{f}^{(1)}). \tag{19}$$

This forecasting rule may be referred to as *perfect forecasting rule.* If firm 2 knows the perfect forecasting rule and the market demand function f, she is able to play best response for all times t. Inserting the perfect forecasting rule (19) into the economic law (17) gives

$$\begin{cases} q_t^{(1)} &= R^{(1)}(q_{t-1}^{(2)}, \widehat{f}^{(1)}), \\ q_t^{(2)} &= R^{(2)}\big(R^{(1)}(q_{t-1}^{(2)}, \widehat{f}^{(1)}), f\big). \end{cases} \tag{20}$$

This shows that the dynamics under perfect foresight for firm 1 is generated by the second function in (20) alone and hence are essentially one dimensional. The fixed point $(q_\star^{(1)}, q_\star^{(2)})$ of the perfect-foresight dynamics generated by (20) satisfies the equations

$$\begin{cases} q_\star^{(1)} &= R^{(1)}(q_\star^{(2)}, \widehat{f}^{(1)}), \\ q_\star^{(2)} &= R^{(2)}(q_\star^{(1)}, f). \end{cases} \tag{21}$$

In comparison with the Nash equilibrium (8), this fixed point will be referred to as a *(stationary) self-confirming beliefs equilibrium (SBE)* of the duopoly game. It is easily seen that the SBE coincides with the Nash equilibrium, if firm 1 had the correct specification of the market demand function f. In a self-confirming beliefs equilibrium, both firms correctly predict their rivals output, whereas only firm 1 knows the correct market demand function.[1]

5 Adaptive learning of SBE

In this section we investigate the problem of how firm 2 can learn to play best response if she has incomplete knowledge of the market demand function and the response function of the rival firm 1 is unknown. In order to be successful, firm 2 will have to learn true market demand function as well as the response function of its competitor from repeated estimations of historical market data. As proposed above, we will separate

[1] Our notion of a self-confirming beliefs equilibrium is closely related to the notion of a self-confirming equilibrium of Fudenberg & Levine (1993) but not the same, because $\widehat{f}^{(1)}$ may be arbitrarily wrong.

the estimation of the market demand and the rival's response function as follows. In each trading period, a first routine is used to estimate the unknown market demand. A second one approximates a self-confirming beliefs equilibrium from an estimate of the competitor's response function. While the first problem is a standard estimation problem, we will show that finding a a self-confirming beliefs equilibrium amounts to finding the zero of an unknown function.

An adaptive learning scheme that instead of a SBE approximates a whole response function of the rival firm will be introduced in Section 6. There are many reasons for focusing on SBE. First of all, if a SBE is asymptotically stable under the perfect-foresight dynamics, then all successful learning may in long run end up in that SBE. A method that searches directly for a SBE might then be more efficient. On the contrary, if a SBE is unstable under perfect foresight, one is likely to end up with complex dynamics implying that the whole response function of the rival has to be learned over time. In order to reduce the learning speed it again might be advantageous to first identify and at the same time stabilize a SBE. In either case it therefore is a good strategy to first localize a SBE and from there on approximate a response function in a second step, if necessary.

Following Dana & Montrucchio (1986), Léonard & Nishimura (1999), Bischi, Chiarella & Kopel (2002) and others, we adopt the popular assumption that the market demand function f is linear and given by

$$f(q) = a - bq, \quad q \in \mathbb{R}_+,$$

where q is the aggregate output and $a, b > 0$ are unknown parameters. Suppose that two different aggregate outputs $q_1 \neq q_2$ along with the corresponding market prices p_1 and p_2 have been observed, such that

$$p_1 = a - bq_1 \quad \text{and} \quad p_2 = a - bq_2. \tag{22}$$

Since $q_1 \neq q_2$, these two equations can be solved for the unknown coefficients a and b, yielding

$$a = \frac{q_1 p_2 - q_2 p_1}{q_1 - q_2} \quad \text{and} \quad b = \frac{p_2 - p_1}{q_1 - q_2}.$$

As a consequence, the linear market demand function (22) is perfectly known after two different observations of the aggregate output. It is readily seen from (17) that two different aggregate outputs can easily be generated by firm 2. In the deterministic setting with a linear market demand function f, as adopted for the remainder of this paper, it is therefore justified to assume that the market demand function is perfectly known to firm 2.

Before we complete our learning scheme we make the following observation. Replacing the realization $q_{t-1}^{(2)}$ in first equation of (17) gives

$$q_t^{(1)} = R^{(1)}\big(R^{(2)}(\widehat{q}_{t-1}^{(2)}, f), \widehat{f}^{(1)}\big). \tag{23}$$

This implies that the dynamics of the repeatedly played duopoly is now exclusively driven by the forecasting rules $\psi_t^{(2)}$, $t \in \mathbb{N}$ that firm 2 uses to predict firm 1's output. In particular, if firm 2 does not update the forecast such that $\widehat{q}_{t-1}^{(2)} \equiv \widehat{q}_0^{(2)}$ for all times t, then we see again from (23) together with the economic law (17) that the output of both firms is constant over time. Observe that $(q_\star^{(1)}, q_\star^{(2)})$ is a SBE, iff

$$(24) \qquad \begin{cases} q_\star^{(1)} &= R^{(1)}\big(R^{(2)}(q_\star^{(1)}, f), \widehat{f}^{(1)}\big), \\ q_\star^{(2)} &= R^{(2)}(q_\star^{(1)}, f). \end{cases}$$

It follows from (24) that any SBE defines a fixed point of the map (23) and vice versa any fixed point $\widehat{q}_\star^{(2)} = q_\star^{(1)}$ of (23) defines a SBE. A learning scheme that attempts to learn a SBE must therefore try to find a fixed point of the univariate map (23).

These considerations lead to the following behavioral assumptions concerning the structural knowledge of the two firms.

Assumption 5.1 *The information of the two firms encompasses the following:*

(i) Prices and outputs of both firms are observable for both firms.

(ii) The market demand function is unknown.

(iii) Firm 1 *is a naive updater whose behavior is described by (15).*

(iv) Firm 2 *correctly anticipates firm* 1*'s behavior in the sense that she knows the functional form of (15) but not its correct specification. She is aware that the market demand function along with the response function has to be estimated.*

We will now introduce a learning scheme for firm 2 that finds a SBE. Such a learning scheme will have to search for a fixed point $\widehat{q}_\star^{(2)}$ of the map (23) in some *uncertainty interval* $[\widehat{q}_{\min}^{(2)}, \widehat{q}_{\max}^{(2)}]$ that is assumed to contain a SBE. For simplicity of notation we rewrite (23) as a map[2]

$$F : [\widehat{q}_{\min}^{(2)}, \widehat{q}_{\max}^{(2)}] \to [q_{\min}^{(1)}, q_{\max}^{(1)}]$$

defined by

$$(25) \qquad q_t^{(1)} = F(\widehat{q}_{t-1}^{(2)}) := R^{(1)}\big(R^{(2)}(\widehat{q}_{t-1}^{(2)}, f), \widehat{f}^{(1)}\big).$$

Consider now linear forecasting rules of the form

$$(26) \qquad \psi_L(\widehat{q}^{(2)}; \alpha) = \widehat{q}^{(2)} + \alpha, \quad \widehat{q}^{(2)} \in [\widehat{q}_{\min}^{(2)}, \widehat{q}_{\max}^{(2)}]$$

[2]Note that for the purpose of this paper, it is only relevant that the function F depends on one variable only. Thus simple updating of $\widehat{f}^{(1)}$ that depends on the same variable, albeit somewhat unrealistic in this context, would not affect the arguments that follow.

where $\alpha \in \mathbb{R}$ is some parameter. Within the uncertainty interval, $\alpha > 0$ will induce a sequence $\widehat{q}_t^{(2)} = \psi_L(\widehat{q}_{t-1}^{(2)}; \alpha)$, $t \in \mathbb{N}$ of increasing forecasts, whereas $\alpha < 0$ induces a sequence of decreasing forecasts. With the help of Assumption 5.1, we now introduce the following naive learning scheme for firm 2.

Algorithm 5.1 *Let $\widehat{q}_0^{(2)} \in \{\widehat{q}_{min}^{(2)}, \widehat{q}_{max}^{(2)}\}$ be arbitrary, $m \geq 2$ be an integer and $\epsilon > 0$ be a given tolerance level.*

1. Orientation stage. *Set $k = m$.*
   (a) *If $|F(\widehat{q}_0^{(2)}) - \widehat{q}_0^{(2)}| \leq \epsilon$, then stop.*
   (b) *If $\widehat{q}_0^{(2)} = \widehat{q}_{min}^{(2)}$, then set $\alpha = [\widehat{q}_{max}^{(2)} - \widehat{q}_{min}^{(2)}]/k$.*
   (c) *If $\widehat{q}_0^{(2)} = \widehat{q}_{max}^{(2)}$, then set $\alpha = -[\widehat{q}_{max}^{(2)} - \widehat{q}_{min}^{(2)}]/k$.*
2. Iteration stage. *Set $\widehat{q}_t^{(2)} = \psi_L^t(\widehat{q}_0^{(2)}; \alpha) = \widehat{q}_0^{(2)} + t\alpha$ with $\psi_L^t(\cdot; \alpha)$ denoting the t-th iterate of $\psi_L(\cdot; \alpha)$, until $[F(\widehat{q}_t^{(2)}) - \widehat{q}_t^{(2)}] \cdot [F(\widehat{q}_0^{(2)}) - \widehat{q}_0^{(2)}] \leq 0$ or $t = k$.*
3. Updating stage.
   (a) *If $[F(\widehat{q}_t^{(2)}) - \widehat{q}_t^{(2)}] \cdot [F(\widehat{q}_0^{(2)}) - \widehat{q}_0^{(2)}] > 0$ for all $t \leq k$, then set $m = 2k$.*
   (b) *Let τ denote the first time for which $[F(\widehat{q}_\tau^{(2)}) - \widehat{q}_\tau^{(2)}] \cdot [F(\widehat{q}_0^{(2)})) - \widehat{q}_0^{(2)}] \leq 0$. Set*
   $$\widehat{q}_{min}^{(2)} := \min\{\widehat{q}_{\tau-1}^{(2)}, \widehat{q}_\tau^{(2)}\}, \quad \widehat{q}_{max}^{(2)} := \max\{\widehat{q}_{\tau-1}^{(2)}, \widehat{q}_\tau^{(2)}\}, \qquad \widehat{q}_0^{(2)} := \widehat{q}_\tau^{(2)},$$
   and continue with stage 1.

Figure 2 about here.

The idea of Algorithm 5.1 is to reduce the length of an *uncertainty interval* $[\widehat{q}_{\min}^{(2)}, \widehat{q}_{\max}^{(2)}]$ in which a possible fixed point of the function (25) lies. The search may be started from either end of the interval $[\widehat{q}_{\min}^{(2)}, \widehat{q}_{\max}^{(2)}]$. Whenever the sign of the forecast error changes, one obtains a smaller interval that contains at least one fixed point of (25). If no sign changes were observed, the step size is reduced in order to increase precision. This procedure is repeated until the length of the uncertainty interval is below the tolerance level.

An example for the economic intuition of the learning scheme proposed in Algorithm 5.1 is provided in Figure 2. As long as there are positive (negative) forecast errors, increase (reduce) the current forecast by some quantity α. A positive (negative) forecast error means that the expected output of firm 1 was too low (high). As soon as a negative (positive) forecast error is obtained, check how close to a SBE the current forecast is and reduce the uncertainty interval $[\widehat{q}^{(2)}_{\min}, \widehat{q}^{(2)}_{\max}]$ whenever possible.

Before starting the learning scheme (5.1), an initial uncertainty interval $[\widehat{q}^{(2)}_{\min}, \widehat{q}^{(2)}_{\max}]$ can be chosen as follows. Setting

$$\begin{aligned} \underline{q}^{(2)} &:= \min\left\{R^{(2)}(\widehat{q}^{(2)}, f) : \ \widehat{q}^{(2)} \in [\widehat{q}^{(2)}_{\min}, \widehat{q}^{(2)}_{\max}]\right\}, \\ \overline{q}^{(2)} &:= \max\left\{R^{(2)}(\widehat{q}^{(2)}, f) : \ \widehat{q}^{(2)} \in [\widehat{q}^{(2)}_{\min}, \widehat{q}^{(2)}_{\max}]\right\}, \end{aligned} \tag{27}$$

for the range of firm 2's response function, we see that $[\underline{q}^{(2)}, \overline{q}^{(2)}] \subset [q^{(2)}_{\min}, q^{(2)}_{\max}]$ for any choice of the interval $[\widehat{q}^{(2)}_{\min}, \widehat{q}^{(2)}_{\max}]$. Thus, initially $[\widehat{q}^{(2)}_{\min}, \widehat{q}^{(2)}_{\max}]$ should be taken as to minimize $[q^{(2)}_{\min}, q^{(2)}_{\max}] \setminus [\underline{q}^{(2)}, \overline{q}^{(2)}]$. In this way, $[\widehat{q}^{(2)}_{\min}, \widehat{q}^{(2)}_{\max}]$ should be made large enough to cover firm 1's strategy space $[q^{(1)}_{\min}, q^{(1)}_{\max}]$ as much as possible.

The next theorem shows that the forecasts generated by repeated application of Algorithm 5.1 converge to a SBE.

Theorem 5.1 *Let Assumption 5.1 be satisfied and assume that $[\widehat{q}^{(2)}_{min}, \widehat{q}^{(2)}_{max}]$ contains at least one fixed point of (25). Let $\widehat{q}^{(2)}_0 \in \{\widehat{q}^{(2)}_{min}, \widehat{q}^{(2)}_{max}\}$ be an arbitrary initial forecast and $\{\epsilon_n\}_{n=0}^{\infty}$ be sequence of tolerance levels with $\epsilon_n \to 0$ as $n \to \infty$. Then a repeated application of Algorithm 5.1 yields a sequence of forecasts $\{\widehat{q}^{(2)}_t\}_{t\in\mathbb{N}}$ which converges to a SBE.*

Proof. By construction, for each ϵ_n the Algorithm 5.1 either yields a new uncertainty interval if a sign change occurs or it ends if no sign change occurs. In the first case, a new uncertainty interval $[\widehat{q}^{(2)\prime}_{\min}, \widehat{q}^{(2)\prime}_{\max}]$ is obtained that contains at least one SBE. This uncertainty interval is strictly contained in the initial uncertainty interval $[\widehat{q}^{(2)}_{\min}, \widehat{q}^{(2)}_{\max}]$ and thus smaller. Since ϵ_n becomes arbitrary small, for sufficiently large n an uncertainty interval that contains a SBE will be obtained. Since $\epsilon_n \to 0$, the length of this interval converges to zero. This yields the theorem. *Q.E.D.*

Algorithm 5.1 is inspired by a minimization scheme introduced by Berman (1966) and can easily be generalized to practically all one-dimensional models of the Cobweb type, cf. Wenzelburger (2002a). Notice that it stabilizes a SBE even for cases in which that steady state is unstable under the perfect-foresight dynamics (20).

Under the conditions of Assumption 5.1 it is, in principle, no problem to find all SBE of the nonlinear duopoly game and hence to choose among multiple SBE. This can be seen as follows. Any uncertainty interval $[\widehat{q}^{(2)}_{\min}, \widehat{q}^{(2)}_{\max}]$ with $[F(\widehat{q}^{(2)}_{\min}) - \widehat{q}^{(2)}_{\min}] \cdot [F(\widehat{q}^{(2)}_{\max}) - \widehat{q}^{(2)}_{\max}] < 0$ contains at least one SBE of the duopoly game. Choosing a sufficiently small step size (i.e. take a large m), one obtains a fine partition of $[\widehat{q}^{(2)}_{\min}, \widehat{q}^{(2)}_{\max}]$ by extending the updating stage of Algorithm 5.1 to the whole interval. Whenever the function $\widehat{q}^{(2)} \mapsto F(\widehat{q}^{(2)}) - \widehat{q}^{(2)}$ changes its sign, a SBE is found.

This naive learning scheme is clearly not the most efficient one but nevertheless surprisingly successful. Any other more refined numerical method that approximates zeros of a function could be applied as well. To see this, we show how Newton's method can be incorporated in an adaptive scheme searching for a SBE. Since the derivatives of the function F defined in (25) are unknown to firm 2, we apply Newton's secant method (see Ortega & Rheinboldt 1970) to find a fixed point of the map F. A Newton step is now given by

$$\widehat{q}^{(2)}_{\tau+1} = \widehat{q}^{(2)}_{\tau} - \left[\frac{F(\widehat{q}^{(2)}_{\tau-1}) - \widehat{q}^{(2)}_{\tau-1} - F(\widehat{q}^{(2)}_{\tau}) + \widehat{q}^{(2)}_{\tau}}{\widehat{q}^{(2)}_{\tau-1} - \widehat{q}^{(2)}_{\tau}} \right]^{-1} [F(\widehat{q}^{(2)}_{\tau}) - \widehat{q}^{(2)}_{\tau}], \quad \tau \in \mathbb{N}. \tag{28}$$

Newton's secant method is useful, because it is known for its fast convergence. However, global convergence of the Newton method and its refinements requires the uniqueness of the fixed point.

6 Adaptive learning of best response functions

Recently, Nonaka & Matsumoto (2004) have argued that the long-run average profit for chaotic output fluctuation may be strictly higher than the profit of stationary outputs.[3] In such a case, a firm may be interested to learn the perfect forecasting rule (19). In this section we therefore propose a learning scheme which finds an arbitrarily precise approximation of a perfect forecasting rule (19) for firm 2 using ideas analogous to Wenzelburger (2002a). We assume throughout the remainder of this paper that firm 1 behaves like a naive updater and that the conditions of Assumption 5.1 hold. In this case the learning scheme for a sophisticated firm 2 amounts to estimate the whole response function of firm 1 from historical data.[4] In order to focus on the estimation of the rival's response function we abstract from estimation errors concerning the market demand function f and assume that firm 2 knows f perfectly.

Since the specific functional form of the subjective best response function (15) is unknown to firm 2, our approximations will be chosen from the class of cubic spline functions which are well known for their good approximation properties, see e.g. Watson

[3]In such a situation a firm has an incentive to maximize discounted future profits instead of short-term outputs. We leave the interesting question whether it is advantageous for firms to change their objective function for future research.

[4]As soon as firm 1 uses more complicated forecasting rules or updating schemes, it will be impossible for firm 2 to distinguish between the response function and the forecasting rule without observing firm 2's beliefs. In such cases the composition of response function and forecasting rule has to be estimated all together, using more general techniques, e.g., see White, Gallant, Hornik, Stinchcombe & Woolridge (1992).

(1980).[5] To fix notation, let

$$\Delta = \{\underline{q}^{(2)} = a_0 < a_1 < \cdots < a_n = \overline{q}^{(2)}\} \tag{29}$$

denote a partition of the compact subinterval $[\underline{q}^{(2)}, \overline{q}^{(2)}] \subset [q^{(2)}_{\min}, q^{(2)}_{\max}]$ of firm 2's strategy space. This subinterval is conveniently taken as the range of firm 2's response function.

Definition 6.1 *A (cubic) spline function ψ_Δ associated with Δ is a real-valued function $\psi_\Delta : [\underline{q}^{(2)}, \overline{q}^{(2)}] \to \mathbb{R}$ with the following properties:*

(i) ψ_Δ is two times differentiable with continuous second derivatives,

$$\psi_\Delta \in C^2[\underline{q}^{(2)}, \overline{q}^{(2)}];$$

(ii) ψ_Δ is a polynomial of degree 3 on each interval $[a_i, a_{i+1}]$, $i = 0, \ldots, n$.

The original idea of spline interpolation was to construct a smooth curve through a prescribed set of points. The following result is standard, see e.g., Watson (1980) or Stoer (1979).

Proposition 6.2 *Let Δ be given. Then for any prescribed set of values b_i, $i = 0, \ldots, n$, there exists a unique cubic spline function ψ_Δ, such that*

$$(i) \quad \psi_\Delta(a_i) = b_i,\ i = 0, \ldots, n \quad \textit{and} \quad (ii) \quad D^2\psi_\Delta(\underline{q}^{(2)}) = D^2\psi_\Delta(\overline{q}^{(2)}) = 0.$$

Spline functions can also be used as approximations of continuous functions. To this end, let $\|g\|_\infty$ denote the supremum norm of a real-valued continuous function g, $b_i = g(a_i)$, $i = 0, \ldots, n$ be a prescribed set of values at the knots a_i, $i = 0, \ldots, n$, and replace Condition *(ii)* by $D\psi_\Delta(a_i) = Dg(a_i)$, $i = 0, n$. The following result is due to Carlson & Hall (1973).

Proposition 6.3 *Let $g \in C^m[\underline{q}^{(2)}, \overline{q}^{(2)}]$ with $m = 1, 2, 3,$ or 4, $\|\cdot\|_\infty$ denote the supremum norm on $[\underline{q}^{(2)}, \overline{q}^{(2)}]$, and ψ_Δ be the unique cubic spline approximation of g such that*

$$(i) \quad \psi_\Delta(a_i) = g(a_i),\ i = 0, \ldots, n \quad \textit{and} \quad (ii) \quad D\psi_\Delta(a_i) = Dg(a_i),\ i = 0, n.$$

Then there exist constants $C_{m,r}$, such that

$$\|D^r\psi_\Delta - D^r g\|_\infty \le C_{m,r}\, \|D^m g\|_\infty\, \|\Delta\|^{m-r}, \quad 0 \le r \le \min\{m, 3\},$$

where $\|\Delta\| := \max_i (a_{i+1} - a_i)$. For $r < 3$, the constants $C_{m,r}$ are independent of the partition Δ.

[5] Spline functions could as well be replaced here by other classes of approximating functions such as wavelets.

The precise values of the constants $C_{m,r}$ are found in Carlson & Hall (1973).[6] While it is well-known how to compute a spline approximation for a known function g, firm 2 in our model faces three basic problems. First, firm 1's strategy space as well as the domain of $R^{(1)}(\cdot, \widehat{f}^{(1)})$ are a priori unknown to firm 2. Second, since $R^{(1)}(\cdot, \widehat{f}^{(1)})$ is unknown, so are the knots (a_i, b_i) with $b_i = R^{(1)}(a_i, \widehat{f}^{(1)})$ associated with a partition Δ. These are needed to construct the spline approximation. Third, the required partition Δ is a partition on the strategy space $[q^{(2)}_{\min}, q^{(2)}_{\max}]$ of firm 2. This partition will be obtained only indirectly via the response function $R^{(2)}(\cdot, f)$ of firm 2 and a partition $\widehat{\Delta}$ on a suitably chosen interval $[\widehat{q}^{(2)}_{\min}, \widehat{q}^{(2)}_{\max}]$ for forecasts.

The learning scheme which approximates a locally perfect forecasting rule will therefore involve the following steps.

Algorithm 6.1 *The basic steps of the learning scheme to approximate locally perfect forecasting rules are the following:*

1. *Choose an interval* $[\widehat{q}^{(2)}_{min}, \widehat{q}^{(2)}_{max}]$ *along with a suitable partition* $\widehat{\Delta}$*;*
2. *Determine knots* $(a_i, R^{(1)}(a_i, \widehat{f}^{(1)}))$ *for the induced partition* Δ*;*
3. *Compute spline approximation* ψ_Δ *of* $R^{(1)}(\cdot, \widehat{f}^{(1)})$ *associated with* Δ*.*

1. Choosing an interval. Analogously to Section 5, $[\widehat{q}^{(2)}_{\min}, \widehat{q}^{(2)}_{\max}]$ should be taken as to maximize $[\underline{q}^{(2)}, \overline{q}^{(2)}]$, where the bounds are given in (27). In view of (15), $[\underline{q}^{(2)}, \overline{q}^{(2)}]$ is the relevant part of the domain of firm 1's response function.

2. Computing knots. Let

$$\widehat{\Delta} = \{\widehat{q}^{(2)}_{\min} = \widehat{a}_0 < \widehat{a}_1 < \cdots < \widehat{a}_n = \widehat{q}^{(2)}_{\max}\}$$

denote a partition of the compact interval $[\widehat{q}^{(2)}_{\min}, \widehat{q}^{(2)}_{\max}]$. Then the points

$$a_i := R^{(2)}(\widehat{a}_i, f),\ b_i := R^{(1)}(a_i, \widehat{f}^{(1)}), \quad i = 0, \dots, n$$

determine the knots of the induced partition

$$\Delta = \{\underline{q}^{(2)} = a_0 < a_1 < \cdots < a_n = \overline{q}^{(2)}\},$$

after suitably renumbering the $\{a_i\}$ according to size if necessary.

3. Computing the spline function. Having determined all knots (a_i, b_i), $i = 1, \dots, n$, the computation of the spline function ψ_Δ with $\psi_\Delta(a_i) = b_i$ is a routine calculation in numerical mathematics, see e.g. Watson (1980).

As an immediate consequence of Proposition 6.3, we obtain the following result.

[6]It is also relatively straightforward to see that Condition *(ii)* in Proposition 6.3 can be replaced by Condition *(ii)* of Proposition 6.2, see Stoer (1979).

Theorem 6.4 *Let $\psi_\star(\cdot) = R^{(1)}(\cdot, f)$ be continuously differentiable. Then for each partition Δ of $[\underline{q}^{(2)}, \overline{q}^{(2)}]$ there exists a spline approximation ψ_Δ and a constant $C > 0$ such that*

$$\|\psi_\Delta - \psi_\star\|_\infty \leq C\,\|D\psi_\star\|_\infty \|\Delta\|_\infty.$$

The constant C is independent of Δ.

Assuming that $R^{(2)}(\cdot, f)$ is continuously differentiable,

$$\|\Delta\|_\infty \leq \|DR^{(2)}(\cdot, f)\|_\infty \|\widehat{\Delta}\|_\infty.$$

Theorem 6.4 then shows that the spline approximation ψ_Δ of $\psi_\star$ can be made arbitrarily precise by choosing a fine partition $\widehat{\Delta}$ with sufficiently small $\|\widehat{\Delta}\|_\infty$.

The stability of a self-confirming beliefs equilibrium $(q_\star^{(1)}, q_\star^{(2)})$ with $q_\star^{(1)} = R^{(1)}(q_\star^{(2)}, \widehat{f}^{(1)})$ and $q_\star^{(2)} \in [\underline{q}^{(2)}, \overline{q}^{(2)}]$, can be seen from the slope of ψ_Δ, provided the partition was fine enough. If ψ_Δ is monotonically increasing locally around its fixed point, then the perfect-foresight dynamics may become complex. However, whenever firm 2 realizes this fact she may direct the system to a subjective beliefs equilibrium using the methods presented in Section 5.

7 Conclusions

The analysis of adaptive learning in a dynamic non-linear duopoly game showed that the correct use of the structural information about the market mechanism enables a firm to learn self-confirming beliefs equilibria of the game. The idea was to estimate the market demand function along with the response function of the rival firm from historical data. While linear demand functions are easily identified, it was shown that along an orbit of the system, a firm receives more and more information about the shape of the rival's response function and thus of the location of the stationary self-confirming beliefs equilibria.

The learning scheme proposed in the present paper was based on a simple geometric intuition. It converges globally for all initial conditions and all parameterizations which guarantee the existence of solutions of the corresponding subjective games. The crucial assumption for the success of the learning scheme was that the output of the rival firm and prices are observable quantities.

The approach of this paper may be generalized in several respects. First, competitors who are more sophisticated and use more intricate updating schemes should be considered. Second, using methods presented in Wenzelburger (2002b) the estimation techniques of the paper should appropriately be altered as to take in to account stochastic perturbations of possibly nonlinear market demand functions.

References

Berman, G. (1966): "Minimization by successive approximation", *SIAM J. Numer. Anal.*, 3(1), 123–133.

Bischi, G. I., C. Chiarella & M. Kopel (2002): "On market games with misspecified demand functions: Long run outcomes and global dynamics", mimeo.

Bischi, G.-I. & F. Lamantia (2002): "Nonlinear Duopoly Games with Positive Cost Externalities Due to Spillover Effects", *Chaos, Solitons & Fractals*, 13, 2031–2048.

Bischi, G.-I., C. Mammana & L. Gardini (2000): "Multistability and Cyclic Attractors in Duopoly Games", *Chaos, Solitons & Fractals*, 11, 543–654.

Blume, L. & D. Easley (1993): "What Has the Rational Learning Literature Taught Us?", in *Essays in Learning and Rationality in Economics*, ed. by A. Kirman & M. Salmon. Blackwell, Oxford.

——— (1998): "Rational Expectations and Rational Learning", in *Organizations with Incomplete Information: Essays in Economic Analysis: A Tribute to Roy Radner*, ed. by M. Majumdar. Cambridge University Press, Cambridge.

Carlson, R. & C. Hall (1973): "Error Bounds for Bicubic Spline Interpolation", *Journal of Approximation Theory*, 7, 41–47.

Dana, R. A. & L. Montrucchio (1986): "Dynamic complexity in duopoly games", *Journal of Economic Theory*, 40, 40–56.

Fudenberg, D. & D. K. Levine (1993): "Self-Confirming Equilibrium", *Econometrica*, 61(3), 523–545.

——— (1998): *The Theory of Learning in Games*. MIT Press.

Kalai, E. & E. Lehrer (1993): "Subjective Equilibrium in Repeated Games", *Econometrica*, 61, 1231–1240.

——— (1995): "Subjective Games and Equilibria", *Games and Economic Behavior*, 8, 123–163.

Kirman, A. & M. Salmon (eds.) (1995): *Learning and Rationality in Economics.* Blackwell Publishers, Oxford a.o.

Kirman, A. P. (1975): "Learning by Firms about Demand Conditions", in *Adaptive Economic Models*, ed. by R. Day & T. Groves, pp. 137–156. Academic Press.

Kopel, M. (1996): "Simple and Complex Adjustment Dynamics in Cournot Duopoly Models", *Chaos, Solitons & Fractals*, 7(12), 2031–2048.

Krawczyk, J. & F. Szidarovszky (2003): "On Stable Learning in Dynamic Oligopolies", Mimeo.

Léonard, D. & K. Nishimura (1999): "Nonlinear dynamics in the Cournot model without full information", *Annals of Operations Research*, 89, 165–173.

Marimon, R. (1997): "Learning for Learning in Economics", in *Advances in Economics and Econometrics: Theory and Applications*, ed. by D. Kreps & K. Wallis, vol. I of *Seventh World Congress of the Econometric Society*. Cambridge University Press.

Nonaka, Y. & A. Matsumoto (2004): "Statistical Dynamics in Chaotic Cournot Model with Complentary Goods", Discussion Paper No. 59, Chuo University.

Ortega, J. & W. Rheinboldt (1970): *Iterative Solution of Nonlinear Equations in Several Variables*. Academic Press, New York a.o.

Rand, D. (1978): "Exotic Phenomena in Games and Duopoly Models", *Journal of Mathematical Economics*, 5, 173–184.

Stoer, J. (1979): *Einführung in die Numerische Mathematik I*, Heidelberger Taschenbücher. Springer-Verlag, Berlin a.o., 3rd edn.

Szidarovszky, F. (2003): "Global Stability Analysis of a Special Learning Process in Dynamic Oligopolies", Mimeo.

Watson, G. A. (1980): *Approximation Theory and Numerical Methods*. Wiley, New York a.o.

Wenzelburger, J. (2002a): "Global Convergence of Adaptive Learning in Models of Pure Exchange", *Economic Theory*, 19(4), 649–672.

——— (2002b): "Learning in Random Economic Systems with Expectations Feedback", Habilitationsschrift, University of Bielefeld.

White, H., A. Gallant, K. Hornik, M. Stinchcombe & J. Woolridge (1992): *Artificial Neural Networks: Approximation and Learning Theory*. Blackwell, Cambrigde (Mass.) a.o.

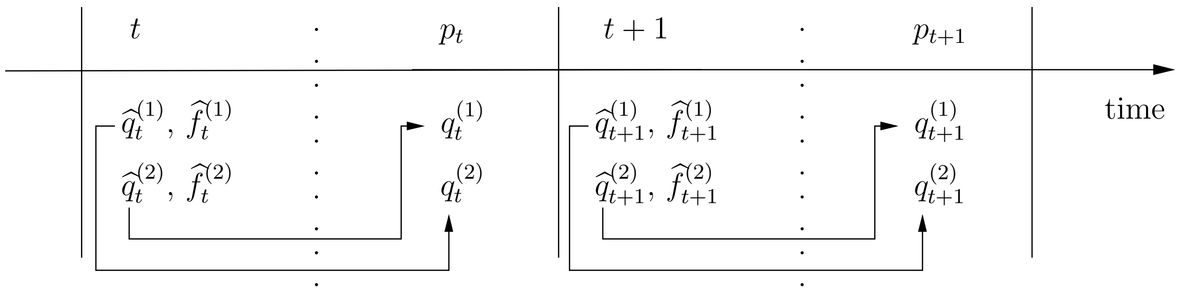

Figure 1: Time-line of actions.

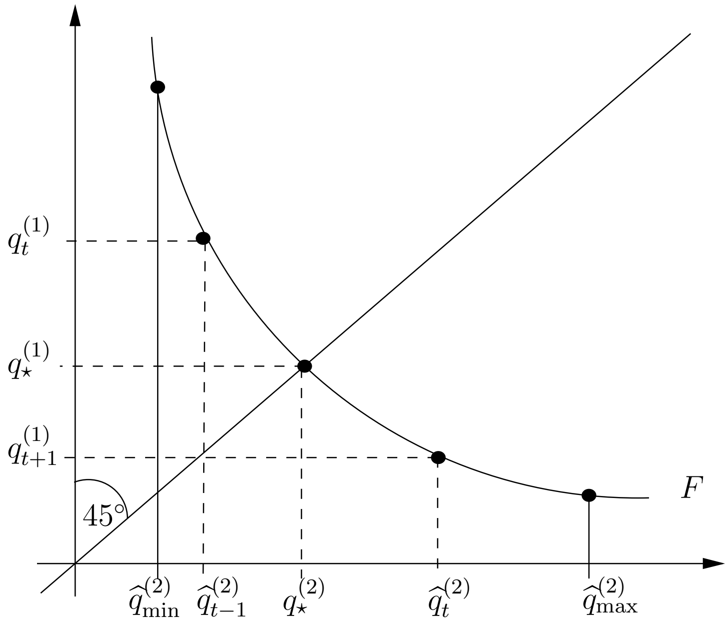

Figure 2: Adaptive learning scheme.