

# Matching and Saving in Continuous Time: Proofs

Christian Bayer<sup>(a)</sup> and Klaus Wälde<sup>(b)</sup>

<sup>(a)</sup> TU Berlin and

<sup>(b)</sup> University of Mainz, CESifo and Université catholique de Louvain<sup>1</sup>

April 2010

This paper provides the proofs to the analysis of a continuous time matching model with saving in Bayer and Wälde (2010a). The paper proves the results on consumption growth, provides an existence proof for optimal consumption and a detailed derivation of the Fokker-Planck equations.

JEL Codes: C62, C65

Keywords: continuous time uncertainty, Fokker-Planck equations, existence proof

## 1 Introduction

The paper by Bayer and Wälde (2010a) – BW10 in what follows – studies optimal saving in a matching model in continuous time. It states many results without proving them. All proofs are presented here.

This paper starts by recapitulating the essential equations of BW10 in the next section. The Keynes-Ramsey rule (Euler equation) in BW10 describes the impact of the interest rate, the time-preference rate and the precautionary savings term on consumption growth. Section 3 provides the proofs behind this description. This section also proves concavity of the value function, a result which is useful for numerical analysis. The only assumption we base our proofs on is continuous differentiability of the consumption function and a finite number of sign changes in a finite interval. These are very weak assumptions which are easily acceptable for our applications.

Once the Keynes-Ramsey rule is understood, BW10 continues by using a phase diagram to illustrate consumption-wealth dynamics. The existence proof of optimal consumption paths that corresponds to the phase-diagram (in some approximative sense) is provided in section 4. Such an existence proof is of importance as it establishes that intuitive reasoning is in fact true. What is more, as numerical identification

---

<sup>1</sup>Large parts of this paper were written while the authors were working at the Royal Institute of Technology (KTH) in Stockholm and the University of Glasgow, respectively. We are grateful to these institutions for their stimulating research environment. Christian Bayer: TU Berlin, Institute of Mathematics, Straße des 17. Juni 136, 10623 Berlin, bayer@math.tu-berlin.de, Klaus Wälde: University of Mainz, Mainz School of Management and Economics, Jakob-Welder-Weg 4, 55128 Mainz. klaus@waelde.com, www.waelde.com. We are very grateful to Walter Schachermayer for comments and guidance, Michael Graber, Giuseppe Moscarini, Sevi Rodríguez Mora, Josef Teichmann and Carlos Carrillo Tudela for comments and Jeremy Lise for discussions. The second author is indebted to Ken Sennewald and Christian Bauer for earlier collaboration on this topic.

of an optimal path is not always straightforward, knowing that the objective one looks for actually exists is very reassuring.

The analysis of BW10 becomes a general equilibrium analysis by using Fokker-Planck equations to describe the evolution of the density of the state variables over time. This density is required to close the model and obtain an endogenous interest and wage rate which is linked to the aggregate capital stock. The derivation of these equations is in section 5. This derivation is relatively detailed in order to show that applications in other contexts are possible as well.

Having obtained Fokker-Planck equations as a tool to describe the evolution of distributions raises the question about existence, uniqueness and stability of the limiting distribution for the state variables. A companion paper (Bayer and Wälde, 2010b) proves that the low interest rate case exhibits a unique invariant distribution and the system is ergodic: the distribution always converges to the invariant distribution as time approaches infinity.

Our proofs are related to various methods and strands in the literature. Tools for the Keynes-Ramsey proofs are elementary. The existence proof for optimal consumption builds on the classical theorem of continuous dependence of the solution to an ordinary differential equation on its initial value. The proof continues by showing that “hitting times” are continuous in initial values - which is far from obvious. We conclude by applying a version of the intermediate value theorem, specifically designed for our purpose. We note that the classical theory of ordinary differential equation, see, for instance, Mattheij and Molenaar (2002) is only applicable in our case after some modifications, because the right hand side satisfies the standard Lipschitz conditions only after a variable transformation.

The principles behind and the derivation of the Fokker-Planck equation for Brownian motion are treated e.g. in Friedman (1975, ch. 6.5) or Øksendal (1998, ch. 8.1). For our case of a stochastic differential equation driven by a Markov chain, we use the infinitesimal generator as presented e.g. in Protter (1995, ex. V.7). From general mathematical theory, we know that the density satisfies the corresponding Fokker-Planck equation  $\frac{\partial}{\partial t}p(t, x) = \mathcal{A}^*p(t, x)$ , where  $p$  denotes the density of the process at time  $t$  and  $\mathcal{A}^*$  is the adjoint operator of the infinitesimal generator  $\mathcal{A}$  of the process ( $x$  is the state variable). We follow this approach in our framework and obtain the Fokker-Planck equation for the law of the employment-wealth process.

Our approach constitutes a considerable generalization to the closed-form solution results used for structural estimation e.g. by Posch (2009) or Aït-Sahalia (2004). In these models, estimation is possible only if a closed-form solution for the density can be found. In many other structural estimation approaches, at least some structure for the density is known, e.g. the generalized exponential distribution in many duration models (Eckstein and Wolpin, 1995, Cahuc et al., 2006 or Launov and Wälde, 2010). In our setup, the density is only implicitly given, i.e. one can derive densities for estimating models with any dynamic property which is based on Brownian motions or Poisson processes (or Lévy processes) and not only with properties which allow for a (closed to) closed-form solution. As we present a very detailed derivation of the differential equations for the density, it should be easy to use this tool for structural estimation in other contexts as well. Lo (1988) derives a Fokker-Planck equation as

well for a one-dimensional stochastic process. We provide a setup with two processes which also makes clear how the Fokker-Planck approach can be generalized to three or more processes.

## 2 Background

In order to make this paper self-contained, we present all equations required for subsequent proofs in this section. The economic background of these equations is in Bayer and Wälde (2010a).

### 2.1 The model

Consider an individual that behaves optimally, according to some objective function and given certain constraints (see (1) and (2) below), by following a consumption function  $c(a(\tau), z(\tau))$ . Arguments of this function are the state variables wealth  $a(\tau)$  and the labour market status  $z(\tau) \in \{w, b\}$ . Wealth of such an individual follows the budget constraint

$$da(\tau) = \{ra(\tau) + z(\tau) - c(a(\tau), z(\tau))\} d\tau \quad (1)$$

where  $r$  is the constant but endogenous interest rate and  $z(\tau)$  simultaneously denotes labour market income. Labour market status and labour market income follow a process  $z(\tau)$  driven by two Poisson processes with state-dependent arrival rates,

$$dz(\tau) = \Delta dq_\mu - \Delta dq_s, \quad \Delta \equiv w - b. \quad (2)$$

Arrival rates are  $\mu(b) = \mu > 0$  and  $\mu(w) = 0$  for  $q_\mu$  and  $s(w) = s > 0$  and  $s(b) = 0$  for  $q_s$ . Since  $q_s$  is only active when  $z$  is in the state  $w$ , while  $q_\mu$  is only active when  $z$  is in the state  $b$ , this equation implies that  $z(\tau)$  jumps between the states  $w$  and  $b$ . Labour income  $w$  (net wage) and  $b$  (unemployment benefits) are constant but endogenous as well. This implies that  $z(\tau)$  is a time-homogeneous Markov chain (in continuous time) with state space  $\{w, b\}$ .

The instantaneous utility function is of the CRRA type,

$$u(c(\tau)) = \frac{c(\tau)^{1-\sigma} - 1}{1-\sigma}, \quad (3)$$

where all proofs will work with a positive  $\sigma \neq 1$ .

### 2.2 The reduced form

The reduced form in BW10 (sect. 4.2) describing optimal behaviour is expressed with wealth  $a$  as the exogenous variable (instead of time  $t$ ). Defining

$$x(a) \equiv c(a, w), \quad y(a) \equiv c(a, b), \quad (4)$$

the reduced form can be written as

$$\dot{x}(a) = \frac{r - \rho + s \left[ \left( \frac{x(a)}{y(a)} \right)^\sigma - 1 \right]}{ra + w - x(a)} \frac{x(a)}{\sigma}, \quad (5a)$$

$$\dot{y}(a) = \frac{r - \rho - \mu \left[ 1 - \left( \frac{y(a)}{x(a)} \right)^\sigma \right]}{ra + b - y(a)} \frac{y(a)}{\sigma}. \quad (5b)$$

Here,  $\dot{x}$  and  $\dot{y}$  denote the respective derivatives of  $x$  and  $y$  with respect to  $a$ . Additional parameters appearing here are the time preference rate  $\rho$ , the separation rate  $s$  and the matching rate  $\mu$  plus the degree of risk-aversion  $\sigma$ . We will assume for all proofs that  $r < \rho$ . This is a non-autonomous system due to the appearance of  $a$  in the denominators in both equations.

BW10 has shown that there is a natural borrowing limit which implies that any solution to (5) must satisfy

$$y(-b/r) = 0. \quad (6)$$

### 2.3 Boundary conditions

When graphically illustrating the full stochastic system in a phase diagram, BW10 introduced a point  $(a_w^*, c(a_w^*, w)) = (a_w^*, x(a_w^*))$  in  $\mathbb{R}^2$  called temporary steady state (TSS). At this point, the *stochastic* system  $(a(\tau), c(a(\tau), w))$  (the stochastic Keynes-Ramsey rules and the budget constraints, see (8) – (11) in BW10) is temporarily (until the next jump of  $q_s$ ) at rest. We know about the TSS from BW10 that

$$x(a_w^*) = ra_w^* + w, \quad (7)$$

and that the value of consumption after a jump, i.e.  $c(a_w^*, b) = y(a_w^*)$ , is given by  $y(a_w^*) = \psi x(a_w^*)$  where

$$\psi \equiv \left( 1 - \frac{r - \rho}{s} \right)^{-1/\sigma} < 1. \quad (8)$$

In all cases where  $\psi$  is used,  $r$  is sufficiently low such that  $\psi$  is a real number.

As the TSS must be part of any optimal consumption path  $c(a, z)$ , we are interested in solutions to the reduced form (5) which contain the point  $(a_w^*, x(a_w^*), y(a_w^*))$ . Note that this point is from  $\mathbb{R}^3$  in contrast to the TSS which is from  $\mathbb{R}^2$ . Note also that this point is not a steady state of this ODE system:  $\dot{y}(a_w^*)$  is negative and – what is “worse” in a sense –  $\dot{x}(a_w^*)$  is not defined as “ $\dot{x}(a_w^*) = 0/0$ ”.

For the latter reason, BW10 introduced an auxiliary temporary steady state (aTSS) defined by  $(a_w^*, x_v(a_w^*))$  from  $\mathbb{R}^2$  where

$$x_v(a_w^*) = ra_w^* + w - v \quad (9)$$

and where  $v > 0$  is an arbitrarily small positive number. This value for  $x(a_w^*)$  replaces the value in (7). Obviously,  $(a_w^*, x_0(a_w^*)) = (a_w^*, x(a_w^*))$ . The value for  $y$  adjusts as well as it is given by

$$y_v(a_w^*) = \psi x_v(a_w^*) \quad (10)$$

where  $\psi$  is the value in (8). This aTSS corresponds to a point  $(a_w^*, x_v(a_w^*), y_v(a_w^*))$  for the ODE (5) from  $\mathbb{R}^3$ . This point, written as

$$\Phi \equiv \Phi_v(\hat{a}) = (\hat{a}, x_v(\hat{a}), y_v(\hat{a})) \quad (11)$$

will later serve as terminal values when solving our ODE systems. Note that  $\Phi$  depends on  $v$ .

## 2.4 The framework for the proofs

The only assumption we make for our proofs is an assumption on continuity.

**Assumption 1** *Relative consumption  $\chi(a) \equiv c(a, w)/c(a, b)$  is continuously differentiable in wealth  $a$ . The number of sign changes of  $\chi'(a)$  in any interval of finite length is finite.*<sup>2</sup>

This is obviously a very weak assumption which does not impose any economically relevant restriction on our solutions.

In addition to this assumption, we have no ambition in understanding the system (5) in all generality. We rather restrict ourselves to the domain where equilibrium dynamics take place. This region was identified in the fig. 1 in BW10 as the one delimited by  $-b/r$  and  $a_w^*$  for wealth and by the zero-motion lines for consumption levels, where we interpret the zero-motion lines as planes in the  $(a, x, y)$ -space. Moreover, we restrict our attention to cases where  $x \geq y$ , i.e. the consumption of the unemployed cannot be larger than the consumption of the employed.<sup>3</sup> Altogether, we consider the domain

$$Q_v = \{(a, x, y) \in \mathbb{R}^3 \mid a \geq -b/r, x \leq ra + w - v, y \geq ra + b, y \geq 0, x \geq y\}. \quad (12)$$

where  $v$  is the small positive constant used in (9) as well.

In the proofs, it will be convenient to restrict attention to a bounded set. Thus, we consider

$$R_{v, \Psi} = Q_v \cap \{(a, x, y) \in \mathbb{R}^3 \mid x \leq \Psi < \infty, a \leq (\Psi - w + v)/r\} \quad (13)$$

where  $\Psi$  only serves to make  $R_{v, \Psi} \subset \mathbb{R}^3$  a compact set, which we need to obtain global, uniform Lipschitz constants. We shall see below that  $\Psi$  has to be chosen larger than  $\Psi_0 = \frac{\psi w - b}{(1 - \psi)r}$ . In this case, however,  $\Psi$  does not interfere with the construction.

---

<sup>2</sup>The second sentence of this assumption is required to rule out “pathological cases”. One can construct continuously differentiable functions that change sign infinitely often in a finite neighborhood (think of  $x \sin(1/x)$  in a neighborhood of zero). None of these functions would be economically plausible in any way. This second sentence is used in the proof of prop. 4.

<sup>3</sup>We will prove later in lem. 3.10 that  $x > y$ . In this sense, the restriction for our domain  $Q_v$  is not binding in any economic sense.

### 3 Proof of the predictions of the Keynes-Ramsey rule

We will now prove prop. 3 of BW10. We repeat it here for reference, as our prop. 3.1.

**Proposition 3.1** *Consider a low interest rate, i.e.  $0 < r \leq \rho$ . Define a threshold level  $a_w^*$  by*

$$\frac{u'(c(a_w^*, b))}{u'(c(a_w^*, w))} \equiv 1 - \frac{r - \rho}{s}. \quad (14)$$

*For our instantaneous utility function (3), this definition reads*

$$c(a_w^*, b) = \psi c(a_w^*, w) \quad (15)$$

where  $\psi$  is from (8).

(i) *Consumption of employed workers increases if the worker owns a sufficiently low wealth level,  $a < a_w^*$ . Employed workers with  $a > a_w^*$  choose falling consumption paths.*

(ii) *Consumption of unemployed workers always decreases.*

(iii) *Consumption of employed workers exceeds consumption of unemployed workers at the threshold  $a_w^*$ , i.e.  $\psi \leq 1$  in (15) for  $r \leq \rho$ .*

Let us recall the findings in BW10 on consumption growth. Here, and for the remainder of this section, we use the notation as in BW10, i.e.  $a_w$  and  $a_b$  satisfy the ODE  $\dot{a}_z(\tau) = ra_z(\tau) + z - c(a_z(\tau), z)$  for fixed  $z \in \{w, b\}$ .

**Lemma 3.2** (BW10) *Individual consumption rises if and only if current consumption relative to consumption in the other state is sufficiently high.*

*For the employed worker, consumption rises if and only if  $c(a_w, w)$  relative to  $c(a_w, b)$  is sufficiently high,*

$$\frac{dc(a_w(\tau), w)}{d\tau} \geq 0 \Leftrightarrow \frac{u'(c(a_w(\tau), b))}{u'(c(a_w(\tau), w))} \geq 1 - \frac{r - \rho}{s} \Leftrightarrow \frac{c(a_w(\tau), w)}{c(a_w(\tau), b)} \geq 1/\psi, \quad (16)$$

where  $\psi$  is from (8).

*For the unemployed worker, consumption rises if and only if  $c(a_b, b)$  relative to  $c(a_b, w)$  is sufficiently high,*

$$\frac{dc(a_b(\tau), b)}{d\tau} \geq 0 \Leftrightarrow \frac{u'(c(a_b(\tau), w))}{u'(c(a_b(\tau), b))} \geq 1 - \frac{r - \rho}{\mu}. \quad (17)$$

#### 3.1 Proof of part (i)

##### 3.1.1 A local result

We first show that consumption  $c(a_w, w)$  rises in time for wealth smaller than but close to  $a_w^*$ .

Given that, by ass. 1, the number of sign changes of  $\chi'(a)$  in any interval for  $a$  of finite length is finite, for any  $a_0$  we can find an  $\varepsilon > 0$  such that  $\chi(a) \equiv c(a, w)/c(a, b)$  is monotonic in  $[a_0 - \varepsilon, a_0]$ . Exploiting this for  $a_w^*$ , whatever the properties of relative consumption, we can always find an  $\varepsilon$  such that one of the following three cases must hold for  $\Omega_\varepsilon \equiv [a_w^* - \varepsilon, a_w^*]$

$$\left. \begin{array}{l} \text{(i)} \\ \text{(ii)} \\ \text{(iii)} \end{array} \right\} \chi'(a)|_{a \in \Omega_\varepsilon} \left\{ \begin{array}{l} < \\ > \\ = \end{array} \right\} 0.$$

Note that we do not make any statement about the derivative in  $a_w^*$ . In fact, in case (i)  $\chi'(a)|_{a \in \Omega_\varepsilon}$  can be negative or zero, in case (ii), it can be positive or zero.

**Lemma 3.3** (a) *Consumption of employed workers rises over time for a wealth level  $a \in \Omega_\varepsilon$  if and only if case (i) holds,*

$$\frac{dc(a_w(\tau), w)}{d\tau} > 0 \text{ for } a_w(\tau) \in \Omega_\varepsilon \Leftrightarrow \text{case (i) holds.}$$

(b) *Consumption  $c(a_w(\tau), w)$  falls over time for  $a_w(\tau) \in \Omega_\varepsilon$  if and only if (ii) holds.*

**Proof.** (a) By (16),  $\frac{dc(a_w(\tau), w)}{d\tau} > 0 \Leftrightarrow c(a_w(\tau), w)/c(a_w(\tau), b) > 1/\psi$ . As  $c(a_w^*, w)/c(a_w^*, b) = 1/\psi$  at  $a_w^*$ , as  $w$  and  $b$  are parameters and using ass. 1, this is a condition on the derivative of relative consumption with respect to wealth  $a$  in  $\Omega_\varepsilon$ :  $dc(a_w(\tau), w)/d\tau$  is positive for  $a_w(\tau) \in \Omega_\varepsilon$  if and only if case (i) holds.

(b) By (16), consumption falls over time if relative consumption lies below  $1/\psi$ . This can be the case in  $\Omega_\varepsilon$  only if case (ii) holds. ■

**Lemma 3.4** *Relative consumption falls in wealth for  $a \in \Omega_\varepsilon$ ,  $\chi'(a)|_{a \in \Omega_\varepsilon} < 0$ , i.e. case (i) holds.*

**Proof.** a) Assume that case (ii) holds, i.e.  $\chi'(a)|_{a \in \Omega_\varepsilon} > 0$ . Then, by lem. 3.3,  $\frac{dc(a_w(\tau), w)}{d\tau} < 0$  for  $a_w(\tau) < a_w^*$ . Consumption of unemployed workers would still decrease in time for all wealth levels. In our set  $Q_v$ ,  $\frac{da_w(\tau)}{d\tau} > 0$  and therefore  $\frac{dc(a, w)}{da} < 0$ . As  $\frac{dc(a_b(\tau), b)}{d\tau} < 0$  and  $\frac{da_b(\tau)}{d\tau} < 0$  in  $Q_v$ , we know that  $\frac{dc(a, b)}{da} > 0$ . As a consequence,  $\chi'(a) < 0$ . This contradicts the assumption that case (ii) holds and case (ii) can be excluded.

b) Now assume that case (iii) holds, i.e. relative consumption is flat,  $\chi'(a)|_{a \in \Omega_\varepsilon \cup a_w^*} = 0$ . As  $c(a_w^*, w)/c(a_w^*, b) = 1/\psi$ ,  $dc(a_w(\tau), w)/d\tau = 0$  for  $a_w(\tau) \in \Omega_\varepsilon$ . As  $dc(a_b(\tau), b)/d\tau < 0$ , relative consumption is not constant – which contradicts the assumption that relative consumption is flat in wealth. As case (iii) is thereby excluded as well, the proof is complete. ■

### 3.1.2 A global result

We now complete the proof by a global result on consumption growth.

**Lemma 3.5** *Consumption  $c(a_w, w)$  (a) rises in time for all  $a < a_w^*$  and (b) decreases in time for all  $a > a_w^*$ .*

**Proof.** (a) Imagine to the contrary of “ $c(a_w, w)$  rises in time for all  $a < a_w^*$ ” that there is an interval  $]\Gamma_1, \Gamma_2[$  with  $\Gamma_2 < a_w^*$  such that this is the last interval before  $a_w^*$  where  $c(a_w, w)$  falls in time,

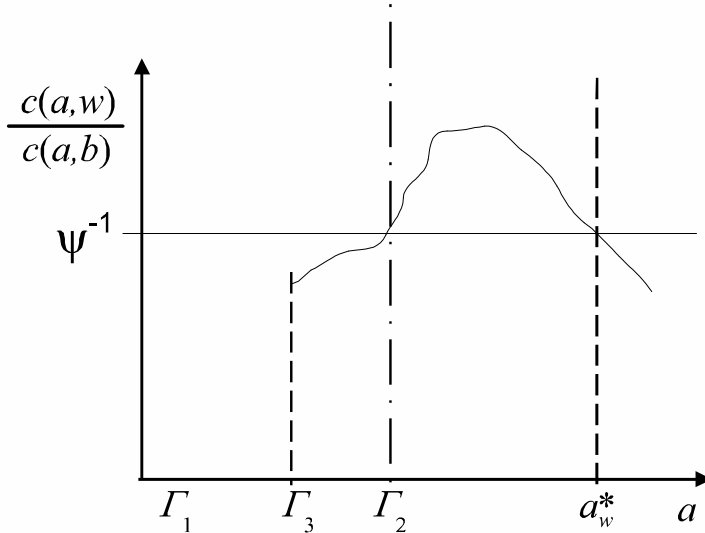
$$dc(a_w(\tau), w)/d\tau < 0, \quad \forall \Gamma_1 < a_w(\tau) < \Gamma_2 < a_w^*. \quad (18)$$

We now proceed as in the proof of lem. 3.4. As  $\frac{da_w(\tau)}{d\tau} > 0$  in  $Q_v$ , this would imply that  $\frac{dc(a, w)}{da} < 0$  for  $\Gamma_1 < a < \Gamma_2$ . We know that  $\frac{dc(a, b)}{da} > 0$  in  $Q_v$ . Hence, we would conclude that

$$\chi'(a) < 0, \quad \forall \Gamma_1 < a < \Gamma_2. \quad (19)$$

By (16), the assumption in (18) would hold if and only if relative consumption  $\frac{c(a_w, w)}{c(a_w, b)}$  is below  $1/\psi$  for  $\Gamma_1 < a < \Gamma_2$ :  $\frac{dc(a_w(\tau), w)}{d\tau} < 0 \Leftrightarrow \frac{c(a_w(\tau), w)}{c(a_w(\tau), b)} < 1/\psi$ . As  $\frac{c(a, w)}{c(a, b)}$  is continuous in wealth by ass. 1 and as case (i) holds by lem. 3.4,  $\frac{c(a, w)}{c(a, b)}$  can be smaller than  $1/\psi$  only if there is some range  $]\Gamma_3, \Gamma_2[$  in which  $\chi'(a) > 0$ . (An example of such a path is shown in fig. 1.) This is a contradiction to the conclusion in (19). Hence, consumption must rise in time for all  $a < a_w^*$ .

(b) This proof is in analogy to the proof of (a). ■



**Figure 1** *An example for relative consumption  $\chi(a) \equiv \frac{c(a, w)}{c(a, b)}$*

## 3.2 Intermediary steps

Before we prove the rest of prop. 3.1, we need some further intermediary results – which, however, are of some interest in their own right. Given that marginal utility from (3) is positive and decreasing,  $u'(c) > 0$  and  $u''(c) < 0$ , we can establish that  $x(a) > y(a)$ , i.e. consumption in the state of employment is larger than in the state of unemployment, keeping wealth constant. We prove in passing that the value functions  $V(a, z)$  are strictly concave in wealth  $a$ .

**Lemma 3.6** *Consumption rises in wealth,  $c_a(a, z) > 0$ .*

**Proof.** Prop. 3.1 (i) shows that  $dc(a_w(\tau), w)/d\tau > 0$  in  $Q_v$ . As  $da_w(\tau)/d\tau > 0$  as well, the derivative  $\dot{x}(a)$  in (5) is positive in  $Q_v$ . ■

**Lemma 3.7** *As marginal utility from consumption is positive, the value function  $V(a, z)$  rises in wealth,  $V_a(a, z) > 0$ .*

**Proof.** The first-order condition for optimal consumption is given by (47) in the appendix and reads

$$u'(c(a, z)) = V_a(a, z). \quad (20)$$

As marginal utility is positive by (3), the value function rises in wealth. ■

**Lemma 3.8** *As  $u''(c) < 0$  and as consumption rises in  $a$  by lemma 3.6, the value function is strictly concave in  $a$ .*

**Proof.** The partial derivative of the first-order condition with respect to wealth implies

$$u''(c(a, z)) c_a(a, z) = V_{aa}(a, z). \quad (21)$$

As  $u''(c(a, z)) < 0$  from the concavity of (3) and  $c_a(a, z)$  is positive by lem. 3.6,  $V_{aa}(a, z)$  must be negative. With lem. 3.7, the value function is strictly concave. ■

**Lemma 3.9** *The shadow price for wealth is higher in the state of unemployment,  $V_a(a, b) > V_a(a, w)$ .*

**Proof.** The derivation of the Keynes-Ramsey rule gives us (see app. A.1)

$$\begin{aligned} & (\rho - r) V_a(a, z) - s(z) [V_a(a, b) - V_a(a, w)] - \mu(z) [V_a(a, w) - V_a(a, b)] \\ & = [ra + z - c(a, z)] V_{aa}(a, z). \end{aligned}$$

In state  $z = w$ , this means

$$(\rho - r) V_a(a, w) - s(z) [V_a(a, b) - V_a(a, w)] = [ra + w - x(a)] V_{aa}(a, w). \quad (22)$$

Given the region we are interested in (where  $ra + w - x(a) > 0$ ) and given lemma 3.8, the right-hand side is negative. Hence, the left-hand side must be negative as well. As  $(\rho - r) V_a(a, w)$  is positive due to  $r < \rho$ , the second term must be negative. This is the case only for  $V_a(a, b) > V_a(a, w)$ . ■

**Lemma 3.10** *Consumption of the employed worker is higher than consumption of the unemployed worker,  $x(a) > y(a)$ .*

**Proof.** As  $V_a(a, b) > V_a(a, w)$ , the first-order condition implies  $u'(c(a, b)) > u'(c(a, w))$ . As the marginal utility is decreasing,  $c(a, w) > c(a, b) \Leftrightarrow x(a) > y(a)$ .

■

### 3.3 Proof of parts (ii) and (iii)

We now complete the final bits of the proof.

**Proof of prop. 3.1 – continuation.** (ii) By (17),  $dc(a_b(\tau), b)/d\tau < 0 \Leftrightarrow u'(c(a_b(\tau), w)) < \varkappa u'(c(a_b(\tau), b))$  where  $\varkappa \equiv 1 - \frac{r-\rho}{\mu} \geq 1$  as  $r \leq \rho$ . As  $u'(c(a_b(\tau), w)) < u'(c(a_b(\tau), b))$  with  $c(a_b(\tau), w) > c(a_b(\tau), b)$  from lem. 3.10, this condition always holds.

(iii) This follows from solving (14) for relative consumption. ■

## 4 The existence of an optimal consumption path

**Definition 4.1** (*Optimal consumption path*)

(i) *A consumption path is a solution<sup>4</sup>  $(a, x(a), y(a))$  of the ODE-system (5) for the range  $-b/r \leq a \leq \hat{a}$  in  $R_{v, \Psi}$  with terminal condition  $\Phi$  from (11).<sup>5</sup> The terminal condition  $\Phi$  satisfies (9), (10) for an arbitrary  $\hat{a} > -b/r$ .*

(ii) *An optimal consumption path is a consumption path which additionally satisfies  $y(-b/r) = 0$ . If such a path exists, we denote the wealth level  $\hat{a}$  of the terminal condition by  $a_w^*$ .*

We remark that the notion of an optimal consumption path depends on  $v$  (via  $\Phi$ ). The objective of this section consists in proving the following

**Theorem 4.2** *There is an optimal consumption path.*

### 4.1 Preliminaries

In what follows, we will use classical theorems for initial value problems for ODEs. Currently, we have formulated our system (5) as a terminal value problem, since we have set a value  $\Phi$ , which the system should attain at  $\hat{a} = a_w^*$ , i.e. at the end of the interval  $[-b/r, a_w^*]$  under consideration. For ease of notation and to help intuition, we shall now recast the problem into a classical initial value problem, i.e. we will require the value  $\Phi$  to be attained at the fixed beginning  $\tau = 0$  of an interval  $[0, \tau^*]$ , on which we study the problem.

<sup>4</sup>By a solution to (5), we here understand continuous maps  $x : [-b/r, \hat{a}] \rightarrow \mathbb{R}$ ,  $y : [-b/r, \hat{a}] \rightarrow \mathbb{R}$ , which are continuously differentiable in  $] -b/r, \hat{a}[$  and solve (5a) and (5b), respectively, in the open interval. In particular, we do not require the derivatives of  $x$  and  $y$  to converge for  $a \rightarrow -b/r$ .

<sup>5</sup>In the sense that an optimal consumption path satisfies  $x(\hat{a}) = x_v(\hat{a})$ ,  $y(\hat{a}) = y_v(\hat{a})$ .

To this end, it is more useful to work with an autonomous system. Hence, we rewrite (5) by including  $m(a) = a$  as third variable which “replaces” wealth  $a$ , which now purely serves as a parameter, i.e. as the independent variable. This gives the system

$$\begin{aligned}\dot{m}(a) &= 1, \\ \dot{x}(a) &= \frac{r - \rho + s \left[ \left( \frac{x(a)}{y(a)} \right)^\sigma - 1 \right]}{rm(a) + w - x(a)} \frac{x(a)}{\sigma}, \\ \dot{y}(a) &= \frac{r - \rho - \mu \left[ 1 - \left( \frac{y(a)}{x(a)} \right)^\sigma \right]}{rm(a) + b - y(a)} \frac{y(a)}{\sigma}.\end{aligned}$$

Now define  $\tau \equiv \hat{a} - a$ ,  $x_1(\tau) \equiv m(\hat{a} - \tau)$ ,  $x_2(\tau) \equiv x(\hat{a} - \tau)$ ,  $x_3(\tau) \equiv y(\hat{a} - \tau)$ . Then,  $\frac{d}{d\tau}x_1(\tau) \equiv \dot{x}_1(\tau) = \frac{d}{d\tau}m(\hat{a} - \tau) = \frac{d}{d[\hat{a}-a]}m(a) = -\frac{d}{da}m(a) = -\dot{m}(a)$ . Doing the same for  $x$  and  $y$ , the “inverted” autonomous system therefore reads

$$\dot{x}_1(\tau) = -1, \tag{23a}$$

$$\dot{x}_2(\tau) = -\frac{r - \rho + s \left[ \left( \frac{x_2(\tau)}{x_3(\tau)} \right)^\sigma - 1 \right]}{rx_1(\tau) + w - x_2(\tau)} \frac{x_2(\tau)}{\sigma}, \tag{23b}$$

$$\dot{x}_3(\tau) = \frac{r - \rho - \mu \left[ 1 - \left( \frac{x_2(\tau)}{x_3(\tau)} \right)^\sigma \right]}{rx_2(\tau) + b - x_3(\tau)} \frac{x_3(\tau)}{\sigma}, \tag{23c}$$

where now  $\dot{x}_i$  denotes the derivative of  $x_i(\tau)$  with respect to  $\tau$ ,  $i = 1, 2, 3$ .

**Definition 4.3** *Given (23) and for  $\tau \geq 0$ , let  $X(\tau; \Phi) = (x_1(\tau), x_2(\tau), x_3(\tau))$  denote the solution of (23) started at  $X(0; \Phi) = \Phi \in R_{v, \Psi}$  from (11) where  $-b/r \leq \hat{a} \leq \frac{\Psi+v-w}{r}$ . For later use, we also introduce the notation  $x_i(\tau) = x_i(\tau; \Phi)$ ,  $i = 1, 2, 3$ .*

By passing from (5) to (23) we have reverted the time-direction – more precisely, in our setting, the wealth-direction – and turned a non-autonomous system into an autonomous one by including the independent variable as an additional component of the solution. Thus, the curve  $a \mapsto (a, x(a), y(a))$  with terminal value  $x(\hat{a}) = x_v(\hat{a})$ ,  $y(\hat{a}) = y_v(\hat{a})$  is equal to the curve  $\tau \mapsto X(\tau; \Phi)$  with  $\Phi = \Phi(\hat{a})$ , which is the solution of an initial value problem in the classical sense. However, the parametrization is reverted in the sense that in the former case we start at the left endpoint (“left” in the sense of the smallest value of the  $a$ -component) and end in the right endpoint, whereas in the latter case we start at the right endpoint and end in the left one. In particular, the absolute value of the speed long the curve is equal, but the direction is reversed.

## 4.2 Continuity of the solution in initial values

In order to be able to apply classical theorems, we need finite derivatives on the right-hand side of an ODE system. The right-hand side of the ODE (5), however, exhibits

singularities at the boundary  $y = ra + b$  of  $Q_v$ . This is of particular importance as the definition of the optimal consumption path in Definition 4.1 uses  $y(-b/r) = 0$  – which lies on this boundary. We obtain finite derivatives by (i) a coordinate transformation and by (ii) (temporarily) reducing the set on which we are interested in a solution by demanding that  $y \geq \varepsilon$ . We will later show how this reduction can then be removed again by passing  $\varepsilon \rightarrow 0$ .

**Lemma 4.4** (*Coordinate transformation*) *Let  $x(a)$  and  $y(a)$  be solutions of (5). The mapping  $a \mapsto y(a)$  is bijective. Change variables  $a = a(y)$  and consider  $x$  and  $a$  as functions of  $y$ . Then*

$$x'(y) \equiv \frac{dx(y)}{dy} = \frac{r - \rho + s \left[ \left( \frac{x(y)}{y} \right)^\sigma - 1 \right]}{r - \rho - \mu \left[ 1 - \left( \frac{y}{x(y)} \right)^\sigma \right]} \frac{x(y)}{y} \frac{ra(y) + b - y}{ra(y) + w - x(y)}, \quad (24a)$$

$$a'(y) \equiv \frac{da(y)}{dy} = \frac{ra(y) + b - y}{r - \rho - \mu \left[ 1 - \left( \frac{y}{x(y)} \right)^\sigma \right]} \frac{\sigma}{y}. \quad (24b)$$

**Proof.** Since  $\dot{y}(a) > 0$ ,  $y$  is a bijective function of  $a$ . As  $a'(y) = \frac{1}{\dot{y}(a)}$ , we obtain the second equation by inserting (5b). The first equation follows from “dividing (5a) by (5b)”. ■

We are going to avoid the singularity at  $y(-b/r) = 0$  by temporarily requiring these properties only to hold “up to an arbitrarily small number  $\varepsilon$ ”. We do this by considering the domain  $R_{\varepsilon, v, \Psi}$  as given in the following

**Definition 4.5** *Fix a numbers  $\varepsilon > 0$  and define*

$$R_{\varepsilon, v, \Psi} = R_{v, \Psi} \cap \{(a, x, y) \in \mathbb{R}^3 \mid y \geq \varepsilon\}. \quad (25)$$

This definition implies that we temporarily replace the requirement that  $y(-b/r) = 0$  by  $y(a) = \varepsilon$  for some  $-b/r \leq a \leq -b/r + \varepsilon/r$ .

**Lemma 4.6** *The right-hand side given in (24) is uniformly Lipschitz on  $R_{\varepsilon, v, \Psi}$ .*

**Proof.** Consider the right-hand side of (24a). The only possible points, where the Lipschitz constant can explode, are when the denominators in the right-hand side become 0 or when a term under a fractional power (i.e. with exponent  $\sigma$ ) becomes 0. In  $R = R_{\varepsilon, v, \Psi}$ ,  $y$  is uniformly bounded away from 0 and  $x$  is uniformly bounded away from  $ra + w$ . Moreover, note that  $r - \rho - \mu \left[ 1 - \left( \frac{y}{x} \right)^\sigma \right] = 0$  if and only if  $\left( \frac{y}{x} \right)^\sigma = 1 - \frac{r - \rho}{\mu}$ . Now  $1 - \frac{r - \rho}{\mu} > 1$  by the assumption that  $r < \rho$ . On the other hand,  $y < x$ , implying that  $\left( \frac{y}{x} \right)^\sigma < 1$ . Consequently, all the denominators are uniformly bounded away from 0.

For the fractional powers, note that  $x/y > 1$  is trivially uniformly bounded away from 0. As  $x \leq \Psi$ ,

$$\frac{y}{x} > \frac{\varepsilon}{\Psi}$$

is uniformly bounded away from 0 on  $R_{\varepsilon, v, \Psi}$ . This shows that (24a) is uniformly Lipschitz.

The same arguments show that the right-hand side of (24b) is uniformly Lipschitz, too. ■

**Remark 4.7** *Since the right hand side of (24) is uniformly Lipschitz, we can now apply the classical theory of ODEs. For instance, we have existence and uniqueness of the solution by the Picard-Lindelöf theorem, see Mattheij and Molenaar (2002, th. II.2.3, th. II.3.1). Moreover, the solution will be continuous as a function of the initial value, see, again, Mattheij and Molenaar (2002, th. II.4.7). In the lemma below, we will see how this even implies the corresponding properties for the non-transformed system (23).*

**Lemma 4.8** *(Continuity in initial values) Consider the set  $R = R_{\varepsilon, v, \Psi}$  from (25) and the solution  $X(\tau; \Phi)$  from Definition 4.3 with initial condition  $\Phi$  given in (11). The solution  $X(\tau; \Phi)$  depends continuously on its initial values  $\Phi$ . More precisely, there is a constant  $L > 0$  and an increasing map  $\kappa : [0, \infty[ \rightarrow [0, \infty[$  (a modulus of continuity) with  $\lim_{t \searrow 0} \kappa(t) = \kappa(0) = 0$  such that*

$$\|X(\tau_1; \Phi_1) - X(\tau_2; \Phi_2)\| \leq L\|\Phi_1 - \Phi_2\| + \kappa(|\tau_1 - \tau_2|),$$

*provided that  $\Phi_1, \Phi_2 \in R$  and  $X(\tau; \Phi_i) \in R$  for all  $0 \leq \tau \leq \max(\tau_1, \tau_2)$ ,  $i = 1, 2$ . Here,  $\|\cdot\|$  denotes the Euclidean norm on  $\mathbb{R}^3$ .*

**Proof.** By classical results from the theory of ordinary differential equations, see for instance Mattheij and Molenaar (2002, th. II.4.7), the solution of an ODE-system depends continuously on the initial data as long as the right-hand side is uniformly Lipschitz. More precisely, let  $Y(\tau; \Phi)$  denote the solution of an ODE with uniformly Lipschitz right-hand side (with Lipschitz constant  $C$ ), started at  $Y(\tau_0; \Phi) = \Phi$ , then

$$\|Y(\tau; \Phi_1) - Y(\tau; \Phi_2)\| \leq \exp(C(\tau - \tau_0)) \|\Phi_1 - \Phi_2\|.$$

Now consider the transformed system  $(a(y), x(y))$  from (24). By Lemma 4.6, the right-hand side is uniformly Lipschitz. The solution of (24) therefore depends continuously on its initial data  $(a_0, x_0)$ . It is then obvious that the trajectory  $(a(y), x(y), y)$  depends continuously on  $(a_0, x_0, y_0)$ . As system (24) is a reparameterized version of (5), the solution  $(a, x(a), y(a))$  to (5) from def. 4.1 is also continuous in its boundary conditions – even though the right hand side of (5) is not uniformly Lipschitz. Similarly, as (23) is just a reparameterization of (5), the solution  $X(\tau; \Phi)$  to (23) from def. 4.3 is also continuous in its initial condition  $\Phi$ .

In order to get the estimate, we now consider the ODE (23) and note that we only consider it on the compact set  $R_{\varepsilon, v, \Psi}$ . In the parametrization by  $y$  given in (24),  $y$  is the independent variable, i.e. plays the role of  $\tau$  in the above estimate. By compactness of  $R_{\varepsilon, v, \Psi}$ ,  $y$  only runs through a bounded set, therefore we can rewrite the constant in the above inequality as  $\exp(C(y - y_0)) \leq L$  for some suitable  $L > 0$ .

Given  $\Phi \in R_{\varepsilon, v, \Psi}$ . Then  $a_w^* \leq \frac{\Psi - w + v}{r}$ , which implies that the solution  $X(\tau; w)$  can only stay inside  $R_{\varepsilon, v, \Psi}$  until time  $\tau = \frac{\Psi - w + v + b}{r}$ , at most. Consider

$$D = \{(\tau, \Phi) \in [0, \infty[ \times R_{\varepsilon, v, \Psi} \mid X(\tau; \Phi) \in R_{\varepsilon, v, \Psi}\}.$$

Then  $D$  is a closed subset of  $[0, \frac{\Psi - w + v + b}{r}] \times R_{\varepsilon, v, \Psi}$ , implying that  $D$  is compact. Consequently,  $X : D \rightarrow R_{\varepsilon, v, \Psi}$  is uniformly continuous, which implies the existence of a modulus of continuity  $\kappa$  with

$$\|X(\tau_1; \Phi_1) - X(\tau_2; \Phi_2)\| \leq \kappa(|\tau_1 - \tau_2| + \|\Phi_1 - \Phi_2\|).$$

The inequality in the lemma then follows by the triangle inequality. ■

### 4.3 Continuity of the first hitting-wealth in initial values

While we have shown in the previous section that the solutions to all systems (5), (23) and (24) are continuous in initial values, this does not automatically imply that the solutions will be continuous on the boundary of the domain we are interested in, in the sense that the place where the solution leaves the domain  $R$  might not depend continuously on the initial data. This will now be proved in this section.

In the proofs and also in a later step, we will use the following

**Definition 4.9** (*First hitting-wealth*) Consider the set  $R_{\varepsilon, v, \Psi}$  from (25) and the solution  $X(\tau; \Phi)$  to the system (23). Consider the path  $y(a)$  that corresponds to  $x_2(\tau)$  of this solution. Then we define  $\hat{a}_{1st} = f(\hat{a})$  as the “first hitting-wealth” (in analogy to first hitting-time), i.e. the wealth level where the path  $y(a)$  hits any boundary of  $R_{\varepsilon, v, \Psi}$  for the first time. Similarly denote  $\tau(\Phi) \equiv \inf\{\tau \geq 0 \mid X(\tau; \Phi) \in \partial R_{\varepsilon, v, \Psi}\}$  and  $F(\Phi) \equiv X(\tau(\Phi); \Phi)$ .

We know that  $\hat{a}_{1st}$  exists because in the set  $R_{\varepsilon, v, \Psi}$  the derivatives in (23) are well-defined and a solution therefore exists. Notice that  $\hat{a}_{1st}$  equals the first component of  $F(\Phi(\hat{a}))$ .

We also need

**Definition 4.10** Let  $N \subset R_{\varepsilon, v, \Psi}$  with

$$N = \left\{ \Phi(\hat{a}) \mid \hat{a} \in \left[ -\frac{b}{r}, \frac{\psi[w - v] - b}{r[1 - \psi]} \right] \right\}$$

be the set of all potential initial conditions from (11) for a solution in the sense of def. 4.1. Here we implicitly assume that  $\Psi$  is large enough that indeed  $N \subset R_{\varepsilon, v, \Psi}$ .<sup>6</sup> Define  $M$  as

$$M = M_1 \cup M_2 \cup M_3 \subset R_{\varepsilon, v, \Psi} \tag{26}$$

---

<sup>6</sup>This is the only necessary condition on  $\Psi$  for the construction to work. In the sequel, we shall assume this condition without further notice.

with

$$\begin{aligned} M_1 &= \{(a, x, y) \in R_{\varepsilon, v, \Psi} \mid y = ra + b\}, \\ M_2 &= \{(a, x, y) \in R_{\varepsilon, v, \Psi} \mid a = -b/r\}, \\ M_3 &= \{(a, x, y) \in R_{\varepsilon, v, \Psi} \mid y = \varepsilon\}. \end{aligned}$$

This set will turn out to be the set of all potential first hitting-wealths.

Since we know that  $x > y$ , the trajectory will not hit the boundary of  $R$  at the part  $\{x = y\}$ . Therefore, we have the

**Corollary 4.11**  $F : N \rightarrow M$  is a well-defined map from  $\mathbb{R}^3$  to  $\mathbb{R}^3$ , i.e. for every  $\Phi \in N$ , the corresponding solution path  $X(\tau; \Phi)$  exists and stays in  $R_{\varepsilon, v, \Psi}$  until it finally hits  $M$  (and no other boundary of  $R_{\varepsilon, v, \Psi}$ ).

Before formulating the main lemma of this section, let us first derive a simple bound on the derivative  $\dot{y}(a)$  of the consumption of the unemployed.

**Lemma 4.12** For  $(a, x, y)$  in the interior of  $Q_v$  from (12), we have

$$\dot{y}(a) \geq \frac{r - \rho}{ra + b - y(a)} \frac{y(a)}{\sigma}.$$

**Proof.** By (5b) we have

$$\begin{aligned} \dot{y}(a) &= \frac{r - \rho - \mu \left[ 1 - \left( \frac{y(a)}{x(a)} \right)^\sigma \right]}{ra + b - y(a)} \frac{y(a)}{\sigma} \\ &= \left( \frac{r - \rho}{ra + b - y(a)} - \frac{\mu \left[ 1 - \left( \frac{y(a)}{x(a)} \right)^\sigma \right]}{ra + b - y(a)} \right) \frac{y(a)}{\sigma} > \frac{r - \rho}{ra + b - y(a)} \frac{y(a)}{\sigma}. \end{aligned}$$

The last inequality follows from the fact that  $\frac{\mu \left[ 1 - \left( \frac{y(a)}{x(a)} \right)^\sigma \right]}{ra + b - y(a)}$  is negative (and therefore  $-\frac{\mu \left[ 1 - \left( \frac{y(a)}{x(a)} \right)^\sigma \right]}{ra + b - y(a)}$  is positive) as  $ra + b - y(a)$  is negative in the interior of  $Q_v$ . ■

The key result in this section is presented in

**Lemma 4.13** The map  $F : N \rightarrow M$  is continuous.

**Proof.** We need to prove that for every  $\Phi \in N$  and every  $\delta > 0$  there is an  $\eta > 0$  such that

$$\|\Phi_0 - \Phi\| < \eta \implies \|F(\Phi_0) - F(\Phi)\| < \delta. \quad (27)$$

We start the proof by fixing  $\Phi_0, \Phi \in N$  such that  $\|\Phi_0 - \Phi\| < \eta$  for some  $\eta > 0$ . Let us first assume that  $\tau(\Phi_0) \leq \tau(\Phi)$ . By the triangle inequality and Lemma 4.8, we have

$$\begin{aligned} \|X(\tau(\Phi_0); \Phi_0) - X(\tau(\Phi); \Phi)\| &\leq \|X(\tau(\Phi_0); \Phi_0) - X(\tau(\Phi_0); \Phi)\| + \\ &\quad + \|X(\tau(\Phi_0); \Phi) - X(\tau(\Phi); \Phi)\| \\ &\leq L_1 \|\Phi_0 - \Phi\| + \kappa(|\tau(\Phi_0) - \tau(\Phi)|), \end{aligned} \quad (28)$$

for a constant  $L_1 > 0$  and the modulus of continuity  $\kappa$ . In order to get an estimate for  $|\tau(\Phi_0) - \tau(\Phi)|$ , we have to distinguish between three different cases.

*Case (i):*  $F(\Phi_0) \in M_1$ .

By Lemma 4.12, there are constants  $L_2, \ell_2 > 0$  such that  $\dot{y} \geq L_2$  for  $|y - (ra + b)| \leq \ell_2$ . More precisely, we can choose  $\ell_2 > 0$  freely and obtain the bound for  $L_2 = \frac{1}{\ell_2} \frac{(\rho-r)\varepsilon}{\sigma}$ . If  $L_1\eta \leq \ell_2$ , we can bound the absolute value of the derivative of  $x_3(\tau; \Phi)$  from below by  $L_2$  (for  $t \geq \tau(\Phi_0)$ ). This implies that the path  $X(\tau; \Phi)$  hits  $M_1$  before time  $\tau(\Phi_0) + \tau$  for

$$\tau(L_2 - r) = \ell_2 \iff \tau = \frac{\ell_2}{L_2 - r},$$

unless it hits another boundary of  $R_{\varepsilon, v, \Psi}$  before that. Inserting into (28), this gives the estimate

$$\|F(\Phi_0) - F(\Phi)\| \leq L_1\eta + \kappa\left(\frac{\ell_2}{L_2 - r}\right).$$

Choosing  $\ell_2 = L_1\eta$ , the bound is smaller than  $\delta$  provided that

$$\kappa\left(\frac{L_1}{\frac{C}{L_1\eta} - r}\eta\right) + L_1\eta < \delta, \quad (29)$$

where  $C \equiv \frac{(\rho-r)\varepsilon}{\sigma}$ . Note that the left hand side in (29) converges to zero for  $\eta \rightarrow 0$ , therefore we can find an  $\eta_0(\delta) > 0$  (only depending on the constants  $C, L_1$  and  $r$  and the modulus of continuity  $\kappa$ , but not on  $\Phi_0$  or  $\Phi$ ) such that the desired inequality (27) holds for  $\eta < \eta_0$ . We have tacitly assumed that  $L_2 = C/\ell_2 = \frac{C}{L_1\eta} > r$ , which can be realized by choosing  $\eta$  small enough.

*Case (ii):*  $F(\Phi_0) \in M_2$ .

Let  $\hat{a}$  denote the first component of  $\Phi$ , and  $\hat{a}_0$  the first component of  $\Phi_0$ . Note that  $x_1(\tau; \Phi) = \hat{a} - \tau$ , for every  $\tau \geq 0$ . Since  $X(\tau(\Phi_0); \Phi_0) \in M_2$ , we have  $-b/r = x_1(\tau(\Phi_0); \Phi_0) = \hat{a}_0 - \tau(\Phi_0)$ , implying that  $\tau(\Phi_0) = \hat{a}_0 + b/r$ . On the other hand,  $x_1(\tau(\Phi); \Phi) \geq -b/r$ , implying that  $\tau(\Phi) \leq \hat{a} + b/r$ . Combining these two results, we obtain

$$|\tau(\Phi_0) - \tau(\Phi)| = \tau(\Phi) - \tau(\Phi_0) \leq \hat{a} - \hat{a}_0 \leq \|\Phi_0 - \Phi\|.$$

Consequently, the inequality (28) implies

$$\|F(\Phi_0) - F(\Phi)\| \leq L_1 \|\Phi_0 - \Phi\| + \kappa(\|\Phi_0 - \Phi\|) \leq L_1\eta + \kappa(\eta),$$

and (27) holds for  $\eta$  small enough such that

$$L_1\eta + \kappa(\eta) < \delta. \quad (30)$$

*Case (iii):*  $F(\Phi_0) \in M_3$ .

Since  $x_3(\tau(\Phi_0); \Phi_0) = \varepsilon$ , we have  $0 \leq x_3(\tau(\Phi_0); \Phi) - \varepsilon \leq L_1\eta$ . By Lemma 4.12, we can find a constant  $L_3 > 0$  such that  $\dot{y} \geq L_3$  on  $R_{\varepsilon, v, \Psi}$  – note that  $L_3$  depends on  $\varepsilon$ . Thus,  $X(s; \Phi)$  will hit the boundary  $M_3$  before time  $\tau(\Phi_0) + \tau$  with  $\tau = L_1\eta/L_3$ , unless it hits another boundary of  $R_{\varepsilon, v, \Psi}$  before. In any case,  $|\tau(\Phi_0) - \tau(\Phi)| \leq L_1\eta/L_3$ , and we obtain

$$\|F(\Phi_0) - F(\Phi)\| \leq L_1\eta + \kappa\left(\frac{L_1}{L_3}\eta\right),$$

and (27) is satisfied for

$$L_1\eta + \kappa \left( \frac{L_1}{L_3}\eta \right) < \delta. \quad (31)$$

Choosing  $\eta$  small enough that both (29) and (30) and (31) are satisfied, settles the proof for  $\tau(\Phi_0) \leq \tau(\Phi)$ . Notice that none of the conditions (29), (30) and (31) depends on  $\Phi_0$ . Therefore, in the other case  $\tau(\Phi_0) \geq \tau(\Phi)$ , we can just revert the rôles of  $\Phi$  and  $\Phi_0$  and obtain the same results in cases (i), (ii) and (iii). ■

#### 4.4 Existence of a solution

This section proves our main result formulated in Theorem 4.2.

**Proof of Theorem 4.2.** Fix some  $\varepsilon > 0$  and consider  $R_{\varepsilon, v, \Psi}$ . By an intermediate value theorem applied to  $F : N \rightarrow M$ , we will obtain a point or points  $\Phi \in N$  such that  $F(\Phi) \in M_3$  as used in (26), i.e.  $x_3(\tau(\Phi); \Phi) = \varepsilon$  provided that we can show the existence of points (that could be called upper and lower bounds)  $\Phi_v^{\min}, \Phi_v^{\max} \in N$  with  $F(\Phi_v^{\min}) \in M_2$  and  $F(\Phi_v^{\max}) \in M_1$ . (Note that  $F = F_\varepsilon$  and all the  $M_i = M_i(\varepsilon)$ ,  $i = 1, 2, 3$ , depend on  $\varepsilon$  and  $v$ , but not on  $\Psi$ , provided that  $\Psi$  is large enough.)

Choose

$$\Phi_v^{\min} = \Phi(-b/r) = (-b/r, w - b - v, \psi[w - b - v]), \quad \Phi_v^{\max} = \Phi \left( \frac{\psi(w - v) - b}{(1 - \psi)r} \right).$$

By construction, both  $\Phi_v^{\min}$  and  $\Phi_v^{\max}$  are contained in  $N$ . Moreover, we trivially have  $F_\varepsilon(\Phi_v^{\min}) \in M_2(\varepsilon)$ ,  $F_\varepsilon(\Phi_v^{\max}) \in M_1(\varepsilon)$  for every  $\varepsilon > 0$  small enough. Note, in particular, that Lemma 4.13 also implies continuity of  $F$  in the boundary points  $\Phi_v^{\min}$  and  $\Phi_v^{\max}$  of  $N$ . Therefore, the image set  $F_\varepsilon(N)$  is a connected set, with non-empty intersection with both  $M_1$  and  $M_2$ . Since the distance

$$\text{dist}(M_1, M_2) = \inf \{ \|\Phi_1 - \Phi_2\| \mid \Phi_1 \in M_1, \Phi_2 \in M_2 \} = \frac{\varepsilon}{r} > 0,$$

we may conclude that  $F_\varepsilon(N) \cap M_3(\varepsilon) \neq \emptyset$ . This establishes that there must be a  $\Phi$  such that  $F_\varepsilon(\Phi) \in M_3$ . In words, there is an initial condition  $\Phi(\hat{a})$  such that the path  $(a, x(a), y(a))$  hits the boundary at  $y = \varepsilon$ .

Now define

$$N_3(\varepsilon) \equiv F_\varepsilon^{-1}(M_3(\varepsilon)) = \{ \Phi \in N \mid F_\varepsilon(\Phi) \in M_3(\varepsilon) \}.$$

By continuity of  $F_\varepsilon : N \rightarrow M(\varepsilon)$ , the set  $N_3(\varepsilon)$  is compact. Moreover, the family  $(N_3(\varepsilon))_{\varepsilon > 0}$  is directed in the sense that

$$0 < \varepsilon_2 < \varepsilon_1 \implies N_3(\varepsilon_2) \subset N_3(\varepsilon_1).$$

By standard results from topology, the intersection of a directed family of non-empty, compact sets is non-empty, i.e.

$$N_3(0) \equiv \bigcap_{\varepsilon > 0} N_3(\varepsilon) \neq \emptyset.$$

Indeed, take a decreasing sequence  $(\varepsilon_n)_{n \geq 1}$  of positive numbers converging to zero. For every  $n$  choose some  $\Phi_n \in N_3(\varepsilon_n)$ . By compactness of the largest set  $N_3(\varepsilon_1)$ , we can find a subsequence  $n_k$  such that  $(\Phi_{n_k})_{k \geq 1}$  converges to some  $\Phi$ . Note that  $\Phi \in N_3(\varepsilon_{n_k})$  for every  $k$ , since  $\Phi = \lim_{l \rightarrow \infty, l \geq k} \Phi_{n_l}$  and each such  $\Phi_{n_l}$  lies in the closed set  $N_3(\varepsilon_{n_k})$ . Now choose any  $\varepsilon > 0$  and pick a  $k$  such that  $\varepsilon_{n_k} < \varepsilon$ . Then  $\Phi \in N_3(\varepsilon_{n_k}) \subset N_3(\varepsilon)$ , implying that  $\Phi \in \bigcap_{\varepsilon > 0} N_3(\varepsilon)$ .

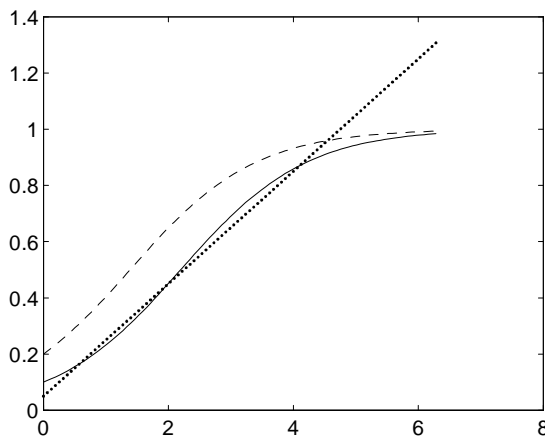
We claim that every element  $\Phi \in N_3(0)$  corresponds to an aTSS. Indeed, the path  $(a, x(a), y(a))$  with terminal value  $(\hat{a}, \hat{x}, \hat{y}) = \Phi$  (corresponding to the path  $X(\tau; \Phi)$ ) satisfies the ODE (5) on  $] -b/r, \hat{a}]$ . Moreover, it starts at  $N$  by construction, and for every  $\varepsilon > 0$ , it takes on the value  $\varepsilon$  somewhere on the interval  $] -b/r, -b/r + \varepsilon[$ . Thus, using monotonicity of  $y$ , we may conclude that

$$\lim_{a \searrow -b/r} y(a) = 0.$$

This establishes that there is an initial condition  $\Phi(\hat{a})$  such that the path  $y(a)$  hits the boundary at  $y = 0$  in the sense that  $y(-b/r) = 0$ . ■

**Remark 4.14** *It is essential for the proof of Theorem 4.2 that the trajectory  $X(\tau; \Phi)$  – or, equivalently,  $(a, x(a), y(a))$  – does not depend on  $\varepsilon$ , which only determines “how long” we observe the trajectory. This means that we observe the trajectory  $X(\tau; \Phi)$  for  $0 \leq \tau \leq \tau(\Phi)$ , with the hitting time  $\tau(\Phi)$  obviously depending on  $\varepsilon$ . Therefore, we can, for fixed  $\Phi \in N_3(0)$ , easily take the limit  $\varepsilon \rightarrow 0$ , which means that we take the limit in  $\tau(\Phi)$ , but do not change the trajectory itself. As a consequence, the ODE is automatically satisfied for the limit, at least for  $0 \leq \tau < \lim_{\varepsilon \rightarrow 0} \tau(\Phi)$ .*

Let us illustrate why we had to use the specific properties of the dynamic system (23) in the proof of lem. 4.13. Continuity in initial conditions does not imply continuity of “first hitting values” in general. Indeed, the first hitting times are inherently non-continuous functionals, even if both the paths and the set, which determines the hitting times, are smooth. This is sketched in ex. 4.15.



**Figure 2** *Non-continuity of the first hitting time in ex. 4.15.*

**Example 4.15** Consider the differential equation  $\dot{z}(t) = (1 - z(t))z(t)$  whose solution is  $z(t) = (1 + (z_0^{-1} - 1)e^{-t})^{-1}$ . This solution is continuous in the initial level  $z_0$  (for  $z_0 > 0$  which we assume) and the solution is plotted for  $z_0 \in \{0.1, 0.2\}$  in fig. 2. Now consider the first-hitting time on the straight line  $0.05 + t/5$  as drawn. Obviously, this time is not continuous in the initial values  $z_0$ .

## 5 Deriving the Fokker-Planck equations

This section derives the Fokker-Planck equations of the wealth-employment process  $(a(t), z(t))$  as described in Section 2, i.e. the partial differential equations which describe the dynamics in time of the density of these random variables. The derivation is in great detail as this facilitates applications for other purposes. Before we go through individual steps, here is the general idea. Step 1: We start with some function  $f$  having as arguments the variables whose density we would like to understand. We compute the differential of this function in the usual way. Step 2: The starting point here is Dynkin's formula. This formula, intuitively speaking, gives the expected value of some function  $f$ , whose arguments are the random variables we are interested in, as the sum of the current value of  $f$  plus the integral over expected future changes of  $f$ . The expected change of  $f$  is expressed by using the density of our random variables. The Dynkin formula is differentiated with respect to time. Step 3: By using integration by parts or the adjoint operator, we get an expression for the change of the expected value of  $f$ . Step 4: A different expression for this change of the expected value can be obtained by starting from the expected value and differentiating it. Step 5: Equating the two gives the differential equations for the density.

### 5.1 The expected change of some function $f$

Assume there is a function  $f$  having as arguments the state variables  $a$  and  $z$ . This function has a bounded support  $S$ , i.e.  $f(a, z) = 0$  outside this support.<sup>7</sup> Heuristically, the differential of this function, using a change of variable formula,<sup>8</sup> gives

$$\begin{aligned} df(a(\tau), z(\tau)) &= f_a(\cdot) \{ra(\tau) + z(\tau) - c(a(\tau), z(\tau))\} d\tau \\ &\quad + \{f(a(\tau), z(\tau) + \Delta) - f(a(\tau), z(\tau))\} dq_\mu \\ &\quad + \{f(a(\tau), z(\tau) - \Delta) - f(a(\tau), z(\tau))\} dq_s. \end{aligned}$$

Due to the state-dependent arrival rates (see table 1), only one Poisson process is active at a time. When we are interested in the expected change, we need to form expectations. We view  $a(\tau)$  and  $z(\tau)$  with  $\tau \geq t$  as two stochastic processes which start in  $t$  and where initial conditions  $a(t)$  and  $z(t)$  can be random variables. We therefore form expectations about  $df$  by using the unconditional expectations operator

<sup>7</sup>We can make this assumption without any restriction. As we will see below, this function will not play any role in the determination of the actual density.

<sup>8</sup>There are formal derivations of this equation in mathematical textbooks like Protter (1995). For a more elementary presentation, see Wälde (2009, part IV).

$E$  as the randomness of initial values are then also taken into account. This is useful for its generality and also when it comes to applications (see the discussion in BW10 on initial conditions and especially distributions).

Applying the conditional expectations operator  $E_\tau$  and dividing by  $d\tau$  yields the heuristic equation

$$\begin{aligned} \frac{E_\tau df(\cdot)}{d\tau} &= f_a(\cdot) \{ra(\tau) + z(\tau) - c(a(\tau), z(\tau))\} \\ &\quad + \mu [f(a(\tau), z(\tau) + \Delta) - f(a(\tau), z(\tau))] \mathbf{1}_{\{b\}}(z(\tau)) \\ &\quad + s [f(a(\tau), z(\tau) - \Delta) - f(a(\tau), z(\tau))] \mathbf{1}_{\{w\}}(z(\tau)) \end{aligned} \quad (32)$$

In what follows, we denote this expression by

$$\mathcal{A}f(a(\tau), z(\tau)) \equiv \frac{E_\tau df(a(\tau), z(\tau))}{d\tau} \quad (33)$$

which is, more precisely, the infinitesimal generator  $\mathcal{A}$  defined by

$$\mathcal{A}f(a, z) = \lim_{\epsilon \searrow 0} \frac{E(f(z(\tau + \epsilon), a(\tau + \epsilon)) | z(\tau) = z, a(\tau) = a) - f(a, z)}{\epsilon}.$$

Notice that  $\mathcal{A}f(a, z)$  does not depend on  $\tau$ , because the Markov-process  $(a(\tau), z(\tau))$  is time-homogeneous. We understand  $\mathcal{A}$  as an operator mapping functions (in  $a$  and  $z$ ) to other such functions. Moreover, note that all test-functions, i.e.  $C^\infty$  functions of bounded support, are in the domain of the operator  $\mathcal{A}$ , i.e. the domain of all functions  $f$  such that the above limit exists (for all  $a$  and  $z$ ).

## 5.2 Dynkin's formula and its manipulation

To abbreviate notation, we now define  $x(\tau) \equiv (a(\tau), z(\tau))$ . The expected value of our function  $f(x(\tau))$  is by Dynkin's formula (e.g. Yuan and Mao, 2003) given by

$$Ef(x(\tau)) = Ef(x(t)) + \int_t^\tau E(\mathcal{A}f(x(s))) ds. \quad (34)$$

To understand this equation, use the definition in (33) and formally write it as

$$Ef(x(\tau)) = Ef(x(t)) + \int_t^\tau \frac{Edf(x(s))}{ds} ds = Ef(x(t)) + \int_t^\tau Edf(x(s)).$$

Intuitively speaking, Dynkin's formula says that the expected value of  $f(x(\tau))$  is the expectation for the current value,  $Ef(x(t))$  (given that we allow for a random initial condition  $x(t)$ ), plus the "sum of" expected future changes,  $\int_t^\tau Edf(x(s))$ .

Let us now differentiate (34) with respect to time  $\tau$  and find

$$\frac{\partial}{\partial \tau} Ef(x(\tau)) = \frac{\partial}{\partial \tau} \int_t^\tau E(\mathcal{A}f(x(s))) ds = E(\mathcal{A}f(x(\tau))), \quad (35)$$

where the first equality used that  $Ef(x(t))$  is a constant and pulled the expectations operator into the integral. This equation says the following: We form expectations in  $t$  about  $f(x(\tau))$ . We now ask how this expectation changes when  $\tau$  moves further into the future, i.e. we look at  $\frac{\partial}{\partial \tau} E[f(x(\tau))]$ . We see that this change is given by the expected change of  $f(x(\tau))$ , where the change is  $\mathcal{A}f(x(\tau))$ .

Now define  $p(a, z, \tau)$  as the joint density of  $z(\tau)$  and  $a(\tau)$ . Notice that  $z$ ,  $a$  and  $\tau$  are independent variables now: as soon as we integrate with respect to the density  $p(a, z, \tau)$ , the whole time-dependence has been absorbed into  $\tau$ . The expectation operator  $E$  integrates over all possible states of  $x(\tau)$ . When we express this joint density as  $p(a, z, \tau) \equiv p(a, \tau|z)p_z(\tau)$ , we can write equation (35) as

$$\begin{aligned} \frac{\partial}{\partial \tau} Ef(x(\tau)) &= E(\mathcal{A}f(x(\tau))) \\ &= p_w(\tau) \int_{-\infty}^{\infty} \mathcal{A}f(a, w) p(a, \tau|w) da + p_b(\tau) \int_{-\infty}^{\infty} \mathcal{A}f(a, b) p(a, \tau|b) da. \end{aligned}$$

Now pull  $p_w(\tau)$  and  $p_b(\tau)$  back into the integral and use  $p(a, z, \tau) \equiv p(a, \tau|z)p_z(\tau)$  again for  $z = w$  and  $z = b$ . Then

$$\begin{aligned} \frac{\partial}{\partial \tau} Ef(x(\tau)) &= \int_{-\infty}^{\infty} \mathcal{A}f(a, w) p(a, w, \tau) da + \int_{-\infty}^{\infty} \mathcal{A}f(a, b) p(a, b, \tau) da \\ &\equiv \phi_w + \phi_b. \end{aligned} \tag{36}$$

**Remark 5.1** *The law of  $x(\tau)$  will not have a density, unless we already start with a density at the initial time  $t$ . In general,  $p$  will therefore be a measure or a distribution (in the sense of a generalized function, see (Yosida 1995), not a true function. Even in that case, the following derivation remains valid, though, using the usual calculus for generalized functions.*

### 5.3 The adjoint operator and integration by parts

This is now the crucial step in obtaining a differential equation for the density. It consists in applying an integration by parts formula which allows to move the derivatives in  $\mathcal{A}f(x(\tau))$  into the density  $p(x, \tau)$ . Let us briefly review this method, without getting into technical details. Given two functions  $f, g : \mathbb{R} \rightarrow \mathbb{R}$  and two fixed real numbers  $c < d$ , the factor rule of differentiation

$$d(f(x) \cdot g(x)) = df(x) \cdot g(x) + f(x) \cdot dg(x) \tag{37}$$

implies that  $f(d)g(d) - f(c)g(c) = \int_c^d f'(x)g(x)dx + \int_c^d f(x)g'(x)dx$ , a formula referred to as partial integration rule. In particular, it also holds for  $c = -\infty$  and  $d = +\infty$ , if the function evaluations are understood as limits for  $c \rightarrow -\infty$  and  $d \rightarrow +\infty$ , respectively. If  $f$  has bounded support, i.e. is equal to zero outside a fixed bounded set, then the function evaluations at  $\pm\infty$  vanish and we get

$$\int_{-\infty}^{+\infty} f'(x)g(x)dx = - \int_{-\infty}^{+\infty} f(x)g'(x)dx. \tag{38}$$

We now apply (38) to equation (36). We can do this as the expressions in (36) “lost” all stochastic features. To this end, insert the definition of  $\mathcal{A}$  given in (33) together with (32) into (36). To avoid getting lost in long expressions, we look at the both integrals in (36) in turn. For the second, observe that

$$\mathcal{A}f(a, b) = f_a(\cdot) \{ra + b - c(a, b)\} + \mu [f(a, w) - f(a, b)],$$

i.e. the term with  $s$  in (32) is missing given that we are in state  $b$ . Hence,

$$\begin{aligned} \phi_b &= \int_{-\infty}^{\infty} [f_a(a, b) \{ra + b - c(a, b)\} + \mu [f(a, w) - f(a, b)]] p(a, b, \tau) da \\ &= \int_{-\infty}^{\infty} f_a(a, b) \{ra + b - c(a, b)\} p(a, b, \tau) da \\ &\quad + \int_{-\infty}^{\infty} \mu [f(a, w) - f(a, b)] p(a, b, \tau) da. \end{aligned}$$

Now integrate by parts. As this integral shows, we only need to integrate by parts for the  $f_a$  term. The rest remains untouched. This gives with (38), where  $g(x)$  stands for  $\{ra + b - c(a, b)\} p(a, b, \tau)$  and  $x$  for  $a$ ,

$$\begin{aligned} \phi_b &= - \int_{-\infty}^{\infty} f(a, b) \left[ \left\{ r - \frac{\partial}{\partial a} c(a, b) \right\} p(a, b, \tau) + \{ra + b - c(a, b)\} \frac{\partial}{\partial a} p(a, b, \tau) \right] da \\ &\quad + \int_{-\infty}^{\infty} \mu [f(a, w) - f(a, b)] p(a, b, \tau) da. \end{aligned} \quad (39)$$

Now look at the first integral of (36). After similar steps (as the principle is the same, we replace  $b$  by  $w$  and the arrival rate  $\mu$  by  $s$  in the last equation), this reads

$$\begin{aligned} \phi_w &= - \int_{-\infty}^{\infty} f(a, w) \left[ \left\{ r - \frac{\partial}{\partial a} c(a, w) \right\} p(a, w, \tau) + \{ra + w - c(a, w)\} \frac{\partial}{\partial a} p(a, w, \tau) \right] da \\ &\quad + \int_{-\infty}^{\infty} s [f(a, b) - f(a, w)] p(a, w, \tau) da. \end{aligned} \quad (40)$$

Summarizing, we find

$$\begin{aligned} \frac{\partial}{\partial \tau} E f(x(\tau)) &= \phi_w + \phi_b \\ &= \int_{-\infty}^{\infty} f(a, w) \left[ - \left\{ r - \frac{\partial}{\partial a} c(a, w) \right\} p(a, w, \tau) - \{ra + w - c(a, w)\} \frac{\partial}{\partial a} p(a, w, \tau) \right] da \\ &\quad + \int_{-\infty}^{\infty} s [f(a, b) - f(a, w)] p(a, w, \tau) da \\ &\quad + \int_{-\infty}^{\infty} f(a, b) \left[ - \left\{ r - \frac{\partial}{\partial a} c(a, b) \right\} p(a, b, \tau) - \{ra + b - c(a, b)\} \frac{\partial}{\partial a} p(a, b, \tau) \right] da \\ &\quad + \int_{-\infty}^{\infty} \mu [f(a, w) - f(a, b)] p(a, b, \tau) da. \end{aligned} \quad (41)$$

## 5.4 The expected value again and finish

- The expected value again

Let us now derive the second expression for the change in the expected value. By definition, and as an alternative to the Dynkin formula (34), we have

$$Ef(x(\tau)) = \int_{-\infty}^{\infty} f(a, b) p(a, b, \tau) da + \int_{-\infty}^{\infty} f(a, w) p(a, w, \tau) da. \quad (42)$$

When we differentiate this expression with respect to time, we get

$$\begin{aligned} \frac{\partial}{\partial \tau} Ef(x(\tau)) &= \int_{-\infty}^{\infty} f(a, b) \frac{\partial}{\partial \tau} p(a, b, \tau) da \\ &+ \int_{-\infty}^{\infty} f(a, w) \frac{\partial}{\partial \tau} p(a, w, \tau) da. \end{aligned} \quad (43)$$

Note that we can use

$$\frac{\partial}{\partial \tau} \int_{-\infty}^{\infty} f(a, z) p(a, z, \tau) da = \int_{-\infty}^{\infty} f(a, z) \frac{\partial}{\partial \tau} p(a, z, \tau) da$$

as  $z$  and  $a$  inside this integral are no longer functions of time.

- Step 5 - Equating the two expressions

We now equate (41) with (43). Collecting terms belonging to  $f(a, w)$  and  $f(a, b)$  gives

$$\int_{-\infty}^{\infty} f(a, w) \varphi_w da + \int_{-\infty}^{\infty} f(a, b) \varphi_b da = 0, \quad (44)$$

where

$$\begin{aligned} \varphi_w &\equiv - \left\{ r - \frac{\partial}{\partial a} c(a, w) + s \right\} p(a, w, \tau) - \{ ra + w - c(a, w) \} \frac{\partial}{\partial a} p(a, w, \tau) \\ &+ \mu p(a, b, \tau) - \frac{\partial}{\partial \tau} p(a, w, \tau) \end{aligned}$$

and

$$\begin{aligned} \varphi_b &\equiv - \left\{ r - \frac{\partial}{\partial a} c(a, b) + \mu \right\} p(a, b, \tau) - \{ ra + b - c(a, b) \} \frac{\partial}{\partial a} p(a, b, \tau) \\ &+ sp(a, w, \tau) - \frac{\partial}{\partial \tau} p(a, b, \tau). \end{aligned}$$

Obviously, the above equation is satisfied if

$$\varphi_b = \varphi_w = 0. \quad (45)$$

These are the Fokker-Planck equations used in BW10.

It is easy to see that the integral equation can only be satisfied for all functions  $f$ , if these Fokker-Planck equations are satisfied. Indeed, assume that  $\varphi_b > 0$  on an interval  $I = [d - \epsilon, d + \epsilon]$ . One can find a non-negative function  $f$  smooth in  $a$  such that  $f(a, w) = 0$  for all  $a$  and

$$f(a, b) = \begin{cases} 1, & a \in [d - \epsilon/2, d + \epsilon/2], \\ 0, & a \in ] - \infty, d - \epsilon] \cup [d + \epsilon, \infty[. \end{cases}$$

Inserting this test function into the integral equation gives

$$\int_{-\infty}^{\infty} f(a, w) \varphi_w da + \int_{-\infty}^{\infty} f(a, b) \varphi_b da = 0 + \int_{d-\epsilon}^{d+\epsilon} f(a, b) \varphi_b da > 0$$

by construction. Therefore,  $\varphi_b = 0$  has to hold for all  $a \in \mathbb{R}$ , and similarly for  $\varphi_w$ .

## 6 Conclusion

This paper proves the results stated in Bayer and Wälde (2010a). In section 3, we prove concavity of the value function and the link between the interest rate and consumption growth. This proof is of interest as it is based only on very weak assumptions about properties of optimal consumption.

In section 4, we give an existence proof for an optimal consumption path given an auxiliary temporary steady state (aTSS). The aTSS is characterized by the fact that we only end  $v$ -close to the zero-motion line of the employed, for any  $v > 0$ , however small. It has proven to be very simple to work with this aTSS numerically. Due to the non-continuity of hitting times of dynamic systems, see fig. 2, the proof is designed for our purposes. We use the specific form of the dynamic system (5) at hand in order to establish continuity in this special case. If this proof is to be used for other systems, one always would have to make sure that lem. 4.13 or a corresponding variant of it will hold.

In section 5, we give an easily accessible exposition into the mathematical methods used to establish Fokker-Planck partial differential equations for the dynamics of densities of stochastic processes given by stochastic differential equations. While the derivation is based on the particular model introduced in section 2, we have tried to keep it rather general. For systems driven by different stochastic differential equations, the infinitesimal generator  $\mathcal{A}$  will have a different form, thus requiring different integrations by parts. In some abstract sense, these integrations by parts serve to compute the *adjoint operator* to  $\mathcal{A}$ , which in turn determines the Fokker-Planck equation. Therefore, the derivation used in section 5 can be applied for a wide class of stochastic models.

Concerning future work, it would be desirable from a theoretical perspective to extend the proof of the auxiliary temporary steady state to an existence proof for a temporary steady state of the wealth dynamics. In the notation used in section 4, this means that we would like to take the limit  $v \rightarrow 0$ , like we took the limit  $\epsilon \rightarrow 0$ . There are, however, certain complications as compared with the latter situation. For

once, note that the trajectory  $X(\tau; \Phi)$ ,  $\Phi = \Phi_v$ , does not depend on  $\varepsilon$ . Only the time  $\tau(\Phi)$ , when  $X(\tau; \Phi)$  hits the boundary  $\partial R_{\varepsilon, v, K}$  does. Therefore, we only had to establish the existence of an initial value  $\Phi_v$ , such that the corresponding trajectory satisfies  $x_3(-b/r) = 0$ . In contrast,  $X(\tau; \Phi_v)$  obviously does depend on  $v$ . Thus, taking a limit  $v \rightarrow 0$  involves taking a limit of solutions to the ODE, and we would have to additionally show that the limiting trajectory again solves the ODE. We could avoid this problem, if we knew how to choose  $\Phi_v = \Phi_v(\hat{a})$ , such that two trajectories  $X(\tau; \Phi_v(\hat{a}))$  and  $X(\tau; \Phi_0(\hat{a}'))$  coincide (on  $R_{\varepsilon, v, K}$ ). This already indicates the dilemma, because we do not know how to solve the ODE for  $v = 0$  because of the singularity in (5a) at the zero-motion line of the employed. If we could overcome this problem, we would still face a similar kind of “shooting” problem as in theo. 4.2. This time, however, we would have to hit a line  $\{(a, ra + w, \psi(ra + w)) \mid -b/r \leq a \leq a^{\max}\}$  in  $(a, x, y)$ -space instead of the line  $\{(-b/r, x, 0) \mid x \geq 0\}$ .

## 7 Appendix

There is a short appendix which is available upon request.

## References

- Aït-Sahalia, Y. (2004): “Disentangling diffusion from jumps,” *Journal of Financial Economics*, 74, 487–528.
- Bayer, C., and K. Wälde (2010a): “Matching and Saving in Continuous Time : Theory,” CESifo Working Paper 3026.
- (2010b): “Matching and Saving in Continuous Time: Stability,” mimeo - available at [www.waelde.com](http://www.waelde.com).
- Cahuc, P., F. Postel-Vinay, and J.-M. Robin (2006): “Wage Bargaining with On-the-job Search: Theory and Evidence,” *Econometrica*, 74, 323–365.
- Eckstein, Z., and K. I. Wolpin (1995): “Duration to First Job and the Return to Schooling: Estimates from a Search-Matching Model,” *Review of Economic Studies*, 62, 263–286.
- Friedman, A. (1975): *Stochastic differential equations and applications*. Vol. 1. Academic Press [Harcourt Brace Jovanovich Publishers], New York, Probability and Mathematical Statistics, Vol. 28.
- Launov, A., and K. Wälde (2010): “Estimating Incentive and Welfare Effects of Non-Stationary Unemployment Benefits,” mimeo - available at [www.waelde.com](http://www.waelde.com).
- Lo, A. W. (1988): “Maximum Likelihood Estimation of Generalised Ito processes with discretely sampled data,” *Econometric Theory*, 4, 321–247.

- Mattheij, R., and J. Molenaar (2002): Ordinary differential equations in theory and practice, vol. 43 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, Reprint of the 1996 original.
- Øksendal, B. (1998): *Stochastic Differential Equations*. Springer, Fifth Edition, Berlin.
- Posch, O. (2009): “Structural estimation of jump-diffusion processes in macroeconomics,” *Journal of Econometrics*, forthcoming.
- Protter, P. (1995): *Stochastic Integration and Differential Equations*. Springer-Verlag, Berlin.
- Wälde, K. (2009): *Applied Intertemporal Optimization*. Lecture notes, University of Mainz, Available at [www.waelde.com/aio](http://www.waelde.com/aio).
- Yosida, K. (1995): *Functional analysis, Classics in Mathematics*. Springer-Verlag, Berlin, Reprint of the sixth (1980) edition.
- Yuan, C., and X. Mao (2003): “Asymptotic stability in distribution of stochastic differential equations with Markovian switching,” *Stochastic Processes and Their Applications*, 103, 277–291.