Stochastic average-cost control, with energy-related applications

Ksenia Chernysh

Submitted for the degree of
Doctor of Philosophy

Heriot-Watt University

Department of Actuarial Mathematics and Statistics,
School of Mathematical and Computer Sciences.

Abstract

In this thesis we present a new stochastic optimisation model arising from supply-side management of power networks. We provide the exact optimal solution under assumption that the environment is Markovian. For the semi-Markovian environment we establish existence of an optimal policy in an important subclass of policies. Finally, we solve the problem for a number of particular examples of environment.

Dedication

This thesis would never has happened if I was not lucky enough to study in Junior Mathematical school (YUMSH) in Saint-Petersburg. I would like to dedicate the thesis to all my teachers and fellow students who I shared my first mathematical excitement with.

Acknowledgements

The last four and half years were the most challenging in my life. I have changed not only my address, language I speak most of the time, but also my area of research.

My greatest thank you goes to my first supervisor Sergey Georgievich Foss. Thank you for supporting and teaching me all the way, even though it was sometimes difficult for you. I apologise for my stubbornness, and wish I used my time wiser. I am grateful for all the opportunities I had because of you, starting from my entire coming to Edinburgh, all the conferences and schools I participated in, and the most important the support I had when I started working in RBS. I hope I will find my ways of using everything I learned from you.

The second thanks goes to my second supervisor James Cruise. Thank you, for your support and advice. James, you were always a "good cop" compared to Sergey, which for sure kept me sane.

The Honorary mention goes to Stan Zachary. Stan, thank you for your great support during my studies, but, especially in the final months of my writing-up. Now, when I am feeling sentimental, I wish we worked more together.

I want to thank my friends in Edinburgh, Saint-Petersburg and other interesting places. I would be a very sad person if you were not here to support and entertain me. To name a few, Dasha, Dasha, Yaroslav, Anya, who I don't have pleasure to see often enough. My best Edinburgh pals who managed to finish their studies long before me, Efthalia and Jen, I wish we were less busy with our real-life jobs and meet more often. Thank you for proof-reading some of the thesis and making your valuable comments.

David, thank you for sharing days, food and flat with me. I owe you everything for the proof-reading of the entire text. Thanks for your love and support. I could not be more grateful.

A special thank you flies to Germany to Shurik, who started it all and sent my documents to Heriot-Watt, while I was away teaching some smart kids in a maths summer school.

My biggest and warmest thanks is, of course, to my mum, who was brave enough to let her only kid go to live in Scotland and do her PhD. I love you and miss you.

Table of contents

List of Figures

Chapter 1

Introduction

Throughout the recent decades collaboration between the power industry and academia deepened. Industry leaders have realised the importance of conducting rigorous research prior to implementation of new policies or technologies, to optimise the performance, minimise costs or prevent failures. Following this, a number of university based projects were funded by industry companies. In particular, Heriot-Watt University along with Durham and Cambridge Universities were involved in a joint programme supported by grant EP/I017954/1 provided by National Grid.

Applied problems have inspired a number of complex mathematical problems, which are of great scientific value on their own regardless of the applied setting. This mutually fruitful collaboration has already delivered great results and will continue to do so in the future.

The shift of energy supply strategies towards greener options brings complexity to the supply management due to the volatility of renewable power. This provides a great opportunity to apply probability and stochastic control theories to real life problems and contribute to society by reducing energy costs and preserving the environment.

The most used renewable power sources nowadays are wind and solar power. According to National Grid's Future Energy Scenarios yearly report of 2015, the target contribution of renewable power sources to the UK's electricity networks should hit 34% percent in 2020, with 18% input of wind power. Therefore, the questions of both forecasting renewable generation and of scheduling the power supply according to the prediction are highly important nowadays.

This thesis focuses on developing a stochastic optimisation framework for a particular power supply model with volatile net demand and also on establishing mathematical results within this framework.

1.1 Applied mathematics and energy problems

The majority of the problems originating from the energy industry may be split into the following classes:

1. supply-side management,
2. demand-side management,
3. storage management,
4. power grids management,
5. power arbitrage market,
6. forecasting.

This list is not exhaustive and represents only the major directions of research. Some complex problems might involve sub-parts belonging to different categories.

The applied problem which motivated this thesis and which we discuss in Section 1.2 falls into supply-side optimisation category. Optimisation problems related to power generation have been actively studied in recent times, for example see [19, 33, 41, 48].

The most famous problems falling into the supply-side management category are the economic dispatch of thermal units and the unit commitment problem. The economic dispatch problem (EDP) was first introduced and solved many years ago, but its more complex modifications and more efficient iterative algorithms have been widely studied since then. See [65] for the general overview of classical methods applied to treat the problem. Suppose that one faces a decision on how to split the generation between n thermal power units to meet demand D so that the cost associated with the system is minimal. The units have different (usually convex) production costs, and may be subject to constraints of different nature. The classical problem is static, but its more realistic dynamic version is widely studied, for example see [51] or one can have a look at the recent overview in [66]. The problem can be solved by different optimisation techniques. For example, one may use the method of Lagrange multipliers, if n is not too big and the cost functions are relatively simple, so that the equations can be easily solved. When this is not the case one may apply more practical iterative methods such as the λ-iteration, the gradient search or the Newton's method. These are classical optimisation techniques, for the reference the reader might see major numerical analysis textbooks such as [4, 12].

The unit commitment (UC) problem is very close to (EDP), but it is a much more complicated problem to solve. Suppose that one has M power units, each of which can be either on or off. The system administrator decides which power units to connect

to the grid to meet the demand and minimise the cost. There is a switching cost present in the system as well. The main difficulty comes from the large number of combinations (equal to 2^M) and from the discrete nature of the problem. In the last 20 years, with the increased importance of renewable power, the stochastic version of unit commitment attracted various researchers [56, 52, 43]. UC is usually being solved by various methods of stochastic programming. For a general introduction on the theory one may have a look in [55].

Another interesting example is [41], in which the theory of Markov decision processes is applied to scheduling of pool pumps in Florida. This example is in between supply and demand-side management.

1.2 The Power Supply Problem - our basic model

For simplicity, we consider only four types of power supply in this thesis: conventional generation, renewable generation, imported power and local storage. We summarize here the main properties which are of interest for this work.

- Conventional power plants require time to ramp up and down. Therefore, they must be scheduled in advance. If there is a shortage of power supply in the system, conventional generation cannot immediately provide the needed power.

- Imported power is more expensive than conventional generation. However in the event of a power shortage, imported power can be supplied immediately via interconnectors. Therefore, regardless of its high price, it is still purchased to prevent blackouts.

- The uncertainty in wind power prediction brings difficulty to the management of the power supply. We suppose that all wind power plants are switched on and uncontrollable. The entire amount of wind power produced is supplied to the system automatically. Therefore we are interested in the so-called *net demand*, which is the difference between the total demand and power produced by wind farms. In this thesis, the net demand is treated as a stochastic process.

- We suppose that the storage may be used locally for covering small shortages.

System operators, National Grid, predict the net demand. The power system acts autonomously to cover the mean of the predicted net demand, but it is unable to deal with any forecast errors. Therefore, in the event of a large power shortage imported power is purchased.

Figure 1.1 represents forecast errors (prediction minus actual wind generation) made in the period from April 2011 to October 2012. The prediction was made 4 hours in advance, the data was kindly provided by Dr. A. Richards, National Grid.

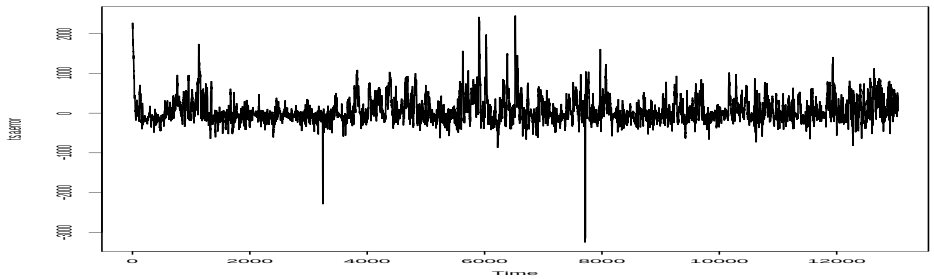


Figure 1.1: Errors in MW in wind power prediction, for 1.5 year period

When the error is negative there is no need to provide additional power. This is true since at the moment the power supply was scheduled, the prediction of the wind power was less optimistic than the actual outcome. Therefore demand is satisfied.

When the error is positive, one faces a situation, where the provided power supply fails to meet demand. This shortage in the power supply must be covered immediately which can be done with the help of imported power.

We assume that the smaller forecast errors may be covered by the local storage, therefore we restrict our interest to only the larger errors. Figure 1.2 is obtained from the previous one by removing the shortfalls below the level 1000MW.

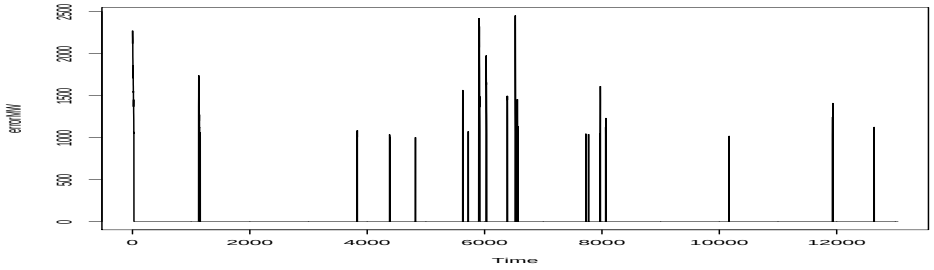


Figure 1.2: Only positive errors which are higher than 1000MW, for 1.5 year period

Hence, we model the error process as a collection of spikes (errors) of random heights with random times between two consecutive shortfalls. The stochastic assumptions for the error process are discussed in detail in Chapter 2.

The expenditure for running this system is the total price for imported power purchased at shortage moments to balance the system. Suppose that χ_k is the random variable standing for the height of the spike number k at time T_k. The average expenditure over period $(T_{k-1}, T_k]$ is then expressed as

$$c_1 \mathbb{E}\chi_k, \tag{1.1}$$

where c_1 is the unit cost of imported power.

To reduce the total expenditure 1.1 we come up with a different idea. We assume that we are allowed to provide additional conventional generation. The produced

power can be used to satisfy the sudden unpredicted changes in the net demand and possibly to reduce the expenditure connected with the expensive imported power. Suppose that $0 < c_2$ is the unit cost of conventional generation. In reality, the imported power is a lot more expensive than conventional power $c_1 \gg c_2$. The total cost is comprised of the cost for imported power at shortfall moments and the cost for the additional conventional generation. Suppose that $X(t)$ units of conventional generation are produced at moment t On period $(T_{k-1}, T_k]$ one spends

$$c_2 \int_{T_{k-1}}^{T_k} X(s)ds + c_1 \mathbb{E} \max(\chi_k - X(T_k), 0). \tag{1.2}$$

It is clear that the minimum in cost 1.2 is not larger than cost 1.1, since it is possible to take $X(t) = 0$. Hence, by providing additional conventional supply we might reduce the total expenditure.

We have mentioned before that the level of conventional generation cannot be changed instantly due to the physical constraints of power units. The upper and lower rates of a possible change are known as *ramp constraints*. The ramp constraints usually depend on the size of a power unit and on the combustible energy source (coal, oil, gas). For example, for the maximal increase in production rates one may look at Table 2.2 [65]. Presence of ramp constraints results in the need of scheduling in advance and in the continuous trajectories for level of conventional generation $X(t)$. Suppose for a moment that there are no ramp constraints, time instants T_k are known in advance and shortage χ_k is constant and known in advance. Clearly, in this scenario the optimal policy is to provide no additional generation at all the moments but T_k, at which we would ramp the production up immediately to level χ_k. However, in the presence of the ramp constraints it is impossible to be at level χ_k at time T_k, if for any $\delta > 0$ the following holds

$$X(T_k - \delta) < X(T_k) - U\delta,$$

where U is the positive ramp constraint. Therefore, the optimisation problem with the ramp constraints is more challenging and more interesting.

We are interested in solving the optimality problem for the average cost over infinite time horizon. From a probabilistic point of view, the most important topic for us is the stationary behaviour of optimal processes.

1.3 Optimisation under uncertainty

Optimisation under uncertainty started being popular in the 1950s with the famous works of Bellman [9, 10] and is still a challenging field in optimisation theory. It is very

difficult to cover the vast variety of optimisation models known nowadays. There are a number of classical books studying stochastic optimisation such as [11, 13, 35, 46]. A good mixture of recent advances were published as an outcome of the *Modern trends in Controlled Stochastic Processes* conferences in 2010 and 2015 [44, 45]. The author had the pleasure of being invited to the second one and to submit a paper [18], which contains partial results of Chapter 3.

To navigate in the world of stochastic optimisation one should be clear about the main components of the problem. These main components include the state space, the action space, the nature of randomness in the system. The rigorous definitions of the objects for our model are given in Chapter 2. State space can be considered as finite, denumerable or in the most general situation to be a Polish space. The same applies to the action space. The actions may be taken continuously or at discrete moments in time. The actions may be allowed to influence not only the controlled process but also the cost functional and the random environment. There are unbounded possibilities for the stochastic assumptions on the random environment.

The choice of optimisation criterion is another important aspect. To start with, the optimisation horizon can be finite or infinite. The existence of the optimal solution for finite horizon problems is usually obvious. It is easy to write a Bellman equation for them and then to solve it numerically by using dynamic programming, [55]. However, these methods are usually computationally difficult. From a probabilistic point of view the solutions are not of a high interest, because there is not enough time for them to start showing their limiting behaviour.

For infinite time horizon there are two possibilities: the β-discounted optimisation and the average cost optimisation. In contrast with the finite time horizon problems, the existence of the optimal solution is a challenging question. For most of the research papers on the topic their main contents are the existence of the optimal control and establishing Bellman equation, so that the optimal policy is its minimiser in a certain class. Usually β-discounted problems are easier to solve and sometimes the solution of the average cost problem is obtained as a limit of solutions for the β-discounted problems. This approach is known as the vanishing discount approach and the examples of its use can be found in [15, 53, 6]. A good survey on average cost optimisation for Markov decision processes in discrete time is presented in [2].

For our optimisation problem the state space of the controlled process is $\mathbb{R}^+$, actions are applied in continuous time and we are interested in average cost optimisation. We assume that the actions have a zero influence on the environment state. There are a number of models in a similar setting, such as Piecewise-Deterministic Markov Processes (PDMP) [23, 24], Markov Decision Drift Processes(MDDPs) [34] (predecessor of PDMPs, PDMP generalises the setting for MDDPs), and Stochastic Fluid Programs (SFPs) [7, 6].

The randomness is modelled in Markovian way, so that the environmental component is a Markov process with a general state space. The dynamics of the controlled process in all these models are given by two objects:

- a flow function, between the jump times of the underlying environment,
- a probability transition kernel that defines the value of the process at the jump times.

The setting of continuous control for PDMPs assumes that the controller can instantly change the underlying Markov process (by speeding it up or slowing it down), the transition kernel, and the flow function. Flow function control was considered for PDMPs in [30], however, to the best of our knowledge, there is no general theory on the average-cost optimisation for PDMPs in the case of the flow function control.

By contrast, similarly to our model, for SFPs the underlying Markov process is uncontrolled, while the flow function is controlled. The formalism of SFPs is the most appropriate to us, because it allows us to prove results for the power supply model without a pre-described flow function, taking into consideration a wider set of possible solutions.

For most of the problems usually only qualitative results on existence of the optimal solution are proven. In this thesis, we find the exact optimal policies in several particular cases.

1.4 Contribution and outline of the thesis

In Chapter 2 we formulate an appropriate stochastic optimisation framework for our problem. We provide rigorous definitions of its components and state the stochastic assumptions for the error process. This stochastic optimisation problem is new to the best of our knowledge, mostly due to the nature of the cost functional consisting of two different components.

In Chapter 3 we show that, under the assumption that the errors occur in accordance with a Poisson process and the errors' values form a Markov chain, our model can be treated as a Stochastic fluid program [7]. For this case we establish the exact shape of the solution for the average cost optimisation problem and discuss how to find it numerically. We also prove the continuity of the optimal processes with respect to the ramp constraints.

In Chapter 4 we establish limiting theory for the optimal processes from Chapter 3 under the semi-Markovian assumption on the inter-errors times. By using theory of regenerative and Markov processes, we prove that there exists a pseudo-optimal control minimising the cost in this particular subclass of controls.

In Chapter 5 we explore a number of particular examples for the error process. Firstly, we explore the deterministic case, in which inter-errors times are constant, and prove that the solution is periodic and find its closed form. Secondly, we explore the case, in which the inter-errors times are distributed as shifted exponential distribution and show that the solution inherits properties of both the exponential and the deterministic cases.

Chapter 2

Mathematical model

The primary goal of this chapter is to present the mathematical model we will work with throughout the thesis. We start with an informal discussion on how to introduce a stochastic optimisation problem in Section 2.1. We proceed by defining the main objects for our problem and stating the assumptions in Section 2.2. In Section 2.4 we show that the basic supply model from Section 1.2 falls into the introduced framework. Due to the applied nature of our problem we would like the controlled process to be bounded, and we establish sufficient conditions for this to hold in Section 2.5.

2.1 Optimisation problems and control: discussion

The first challenge in this thesis was to find an appropriate optimisation framework for the basic supply model from Section 1.2. When an optimisation problem arises in real life the first two questions that should be answered are:

- What to optimise?
  - What does the controlled process represent? Where does it take place? How is it possible to control it?
  - What is the cost associated with the system? What is the time horizon we work within and the optimality criterion?
  - What is the randomness in the system? Is it possible to influence it through our control? What is influenced by the random environment: controlled process and/or cost functional, etc.?
- What is the class of controls (policies) we restrict our search to?
  - Do they depend on the past?
  - When the decisions are taken and can be changed: continuously or at some specified epochs?

Only after the answers to all the above are clear one should proceed to the most important question

- How to optimise it?
  - Under which assumptions does the optimal solution exist?
  - How to prove that the optimal solution is unique?
  - How to find the optimal solution?

This chapter deals with the "What"-questions and the rest of the thesis is dedicated to the "How"-questions.

The following objects and the relationships between them are essential to the "What"-part. They will be formally defined in the next sections of this chapter. For now we provide a general understanding of them by making a link to the basic power supply model. The basic power supply model will be discussed in detail again in Section 2.4.

- A state space of the controlled process and the controlled process itself.

  For instance, for the basic power supply model, the controlled process is the level of additional conventional supply which at any point in time belongs to $\mathbb{R}^+$.

- A random environment.
  We need to define what it is influenced by and what is influenced by it.

  For the basic power supply model the random environment models the errors in wind power prediction. The charge at moments of power shortage depends on the random environment. The random environment is not influenced by any actions taken.

- A cost functional.
  We need to define when it is being charged and what it depends upon (controlled process, environment, action).

  For the basic power supply model the cost is comprised of two parts: the running cost which is charged instantly and the terminal charge which is applied at moments of power shortage in the system. The cost functional does not depend on the action taken.

- An action space.
  We need to understand what is controlled by it.

  For the basic power supply model, an action is the rate of warming up/cooling down of the combustible engine, which provides the system with additional

conventional supply. These rates are usually bounded by capability of the system and hence, is the action space. As previously mentioned, actions cannot change the random environment behaviour.

- Dynamics of the controlled process with respect to the action taken and the environment state.

  For the basic power supply model, the dynamics of the system can be given in integral form, because actions represent rate of change in the power production.

- An optimisation criterion.
  Some possibilities are: total cost over finite time horizon, discounted cost over infinite time horizon or average cost over infinite time horizon.

Once the objects above are specified, the next step is to choose

- a class of admissible policies.

The optimisation problem is considered only in this class.

The nature of these two steps is different. The first set of objects describes the properties of the model and the problem itself. Whereas the decision in regards to the class of admissible policies may be influenced by the purposes of research. Generally, enlarging the class of admissible policies makes it possible to prove the existence of the optimum. However, in a larger class it is more difficult to find the optimum.

In Section 2.2 we focus on the first step, with the second step analysed in Section 2.3.

2.2 The model

2.2.1 The state space, the random environment and the cost functional

In this thesis, we work only with controlled processes $X(t)$ taking values in *state space* $\mathbb{R}^+$.

Throughout the thesis we assume that all random variables and random processes are defined on the common probability space $(\Omega, \mathcal{F}, P)$.

For the model we consider charges of different types. The first is continuously charged for running the process $X(t)$, so it depends only on the level $X(t)$ and does not depend on any other variables. The second charge is applied at random times T_n and can also depend on the environment state Y_n. Therefore, we first introduce the random environment and then define the charges associated with the system.

Definition 2.2.1. The *random environment* is a sequence of pairs

$$E_n = (\tau_n, Y_n), n = \ldots, -1, 0, 1, 2, \ldots,$$

where τ_n are positive-valued random variables, and Y_n are random variables that take values in set $\mathcal{Y} \subset \mathbb{Z}$.

Let $\{T_n\}_{n\in\mathbb{Z}} = \sum_0^n \tau_n$, so that $\tau_n = T_n - T_{n-1}$. We require the following assumption to be satisfied:

Assumption 2.2.1. The mean of inter-arrival times τ_i is finite for each i,

$$\mathbb{E}\tau_i < \infty$$

.

We consider the following set of different stochastic assumptions in this thesis:

Assumptions 2.2.2.

1. (IRE - i.i.d. random environment)
   Sequences $\{\tau_n\}_{n\in\mathbb{Z}}$ and $\{Y_n\}_{n\in\mathbb{Z}}$ are mutually independent and each of them contains independent identically distributed (i.i.d) elements.

2. (SMRE-I - semi-Markovian environment-I)
   Sequence $\{Y_n\}_{n\in\mathbb{Z}}$ is a Markov chain taking values in $\mathcal{Y}$.
   Sequences $\{\tau_n\}_{n\in\mathbb{Z}}$ and $\{Y_n\}_{n\in\mathbb{Z}}$ are mutually independent.
   Sequence $\{\tau_n\}_{n\in\mathbb{Z}}$ is an i.i.d. sequence.

3. (SMRE-II - semi-Markovian environment-II)
   Sequence $\{Y_n\}_{n\in\mathbb{Z}}$ is a Markov chain taking values in $\mathcal{Y}$.
   There is a given set $\{\tau_{n,y}\}_{n\in\mathbb{Z},y\in\mathcal{Y}}$, the elements of which are independent.
   For each $y \in \mathcal{Y}$ elements of the sequence $\{\tau_{n,y}\}_{n\in\mathbb{Z}}$ are identically distributed with common distribution function F_y^τ.
   If $Y_n = y$ then $\tau_{n+1} = \tau_{n+1,y}$, and given Y_n the inter-arrival time τ_{n+1} is independent of everything else.

4. (MRE - Markovian random environment)
   Assumption (SMRE-I) holds and, in addition, elements of sequence $\{\tau_n\}_{n\in\mathbb{Z}}$ are exponentially distributed with parameter λ.

5. (SERE - stationary ergodic random environment)
   Sequence $\{E_n\}_{n\in\mathbb{Z}}$ is stationary ergodic.

Note that

$$(\text{IRE}) \subset (\text{MRE}) \subset (\text{SMRE-I}) \subset (\text{SMRE-II}).$$

If we further assume that the Markov chain $\{Y_n\}$ is Harris ergodic and, at time 0, it is distributed in accordance with its stationary distribution, then Assumptions (SMRE-I, II) and (MRE) are particular subcases of Assumption (SERE). Therefore, by default we suppose that Assumption (SERE) holds without mentioning it. If we require stronger assumptions in other parts of the thesis, we explicitly state them.

Next, we introduce *the terminal charge* (or *the terminal penalty*) f and *the running cost* g. Suppose that, $f : \mathbb{R}^+ \times \mathcal{Y} \to \mathbb{R}^+$ and $g : \mathbb{R}^+ \to \mathbb{R}^+$ are given. We assume that the following holds:

Assumption 2.2.3.

- for each fixed $y \in \mathcal{Y}$, function $f(x, y)$ is convex in x,
- g is convex and non-decreasing.

Convexity is a natural assumption for optimisation problems. It helps to establish existence of the optimal control. However, it generally does not guarantee uniqueness of the solution. The strict convexity is a more restrictive property. For example, even linear cost $g(x) = cx$ is not a strictly convex function. In Chapter 3, we need the strict convexity in a slightly weaker sense, only for a weighted sum of the two functions, to prove uniqueness of the optimal policy.

The cost functional we introduce below formalises the idea of the total charge with respect to both functions f and g within period $[0, t]$. Random variables τ_n stand for time between two consecutive terminal charges, whereas Y_n indicates the type of the next terminal charge at time T_{n+1}, so that the next charge is $f(X(T_{n+1}), Y_n)$.

Let

$$\nu_t := \max\{n : T_n \leq t\} \tag{2.1}$$

be a *counting process* for $\{T_n\}_{n\in\mathbb{N}}$.

Definition 2.2.2. For a realisation of a stochastic process $X(t)$ we define the sample-path cost functional on time interval $[0, t]$, $t > 0$, by

$$\mathcal{C}(X(\cdot), t) = \int_0^t g(X(s))ds + \sum_{i=1}^{\nu_t} f(X(T_i), Y_{i-1}). \tag{2.2}$$

As desired, cost functional (2.2) has the following meaning. Amount $g(X(s))$ is continuously charged for running the process $X(s)$ and amounts $f(X(T_n), Y_{n-1})$ are charged at random times T_n. Hence, the sample-path cost functional is the sum of all charges applied within time interval $[0, t]$.

We finish this subsection with a few examples of functions f and g. We start with examples of the running cost g.

Example 2.2.1.

1. Function $g(x) = cx$ is the cost of purchasing x units of the product at a unit price equal to c.

2. More generally, for $c_2 > c_1$ one can take

$$g(x) = \begin{cases} c_1 x & \text{for } x < x_1, \\ c_2 x + (c_1 - c_2)x_1 & \text{for } x \geq x_1. \end{cases}$$

   This cost function is piecewise-linear and continuous.

   This cost function models a scenario in which a client pays more for purchasing larger amounts of the product. For example, this type of cost is used with regards to power supply. Small amounts of power may be delivered from local reserve or by using renewable generation, whereas larger amounts of power should be produced using large conventional or nuclear power plants. The latter scenario incurs higher production cost and, moreover, delivery cost.

Note that neither of the functions above are strictly convex.
We now turn to examples of the terminal charge f.

Example 2.2.2.

1. Suppose that there is given function $z : \mathcal{Y} \to \mathbb{R}^+$. Then function $f(x, y) = c\max(z(y) - x, 0) = c(z(y) - x)^+$ may be viewed as a penalty cost in the following sense: the positive difference must be covered using a resource, which has a unit cost equal to c.

2. Another type of penalty might be of the form $f(x, y) = c|z(y) - x|$, which can be viewed as a difference penalty.

3. This example generalises the first case in Example 2.2.2 with a penalty for not reaching a random level instead of a fixed level $z(y)$. Assume that for each $y \in \mathcal{Y}$ a probability distribution F_y is given of some random variable χ_y. Let

$$f(x, y) = c\mathbb{E}(\chi_y - x)^+ = \int_0^\infty (z - x)^+ dF_y(z).$$

   We may regard it as the average penalty cost for not reaching the random level χ_y.

   Alternatively, one can take $f(x, y) = c\mathbb{E}|\chi_y - x|$, which is the average penalty for being distant from the random level χ_y.

2.2.2 The action space and the dynamics

Recall the discussion in the basic power supply model. Controlled process $X(t)$ stands for the level of additional conventional power. This power is produced by combustible engines and cannot be subject to a sudden change. There are physical constraints (ramp constraints) bounding the rate of change of production level. This can be expressed as Lipschitz continuity of sample paths of the controlled process $X(t)$. Hence, we would like to work with controlled processes satisfying the following.

Assumption 2.2.4. There are given $U, B > 0$ such that any trajectory of the controlled process $X(t)$ satisfies the following assumption

$$-B\delta \leq X(t+\delta) - X(t) \leq U\delta, \tag{2.3}$$

for all $t, \delta > 0$. Constants U and B are parameters of the system and we refer to them as the *ramp constraints*.

Let us introduce the action space and dynamics of the controlled process corresponding to these actions, so that the controlled process satisfies Assumption 2.2.4. The constant U(or B) then stands for the upper(lower) bound on the "derivative" of controlled process $X(t)$ describing how quickly it is possible to change the control. We write "derivative" understanding that Lipschitz functions are almost everywhere differentiable with respect to the Lebesgue measure. Therefore, in the next definition, we introduce the dynamics by formula 2.4 in integral form.

Definition 2.2.3. Let $A = [-B, U]$ be the *action space* of the system.
A *control* a is any measurable function $a : \mathbb{R}^+ \to A$.
Controlled process $X(t)$ corresponding to the control a is given by

$$X(t) = X(0) + \int_0^t a(s)ds. \tag{2.4}$$

Possible answers to the question of when to choose and change the controls $a(s)$ and what they depend upon result in definitions of different classes of policies in Section 2.3.

Due to the presence of the ramp constraints in the system, controls $a(s)$ should incorporate planning for the future. The controlled process cannot change its value instantly and is governed over time with the help of control $a(s)$.

Suppose that there are no ramp constraints in the system. This is equivalent to considering the limiting case, in which $U = B = \infty$, so the policies satisfying Assumption 2.2.4, are no longer necessarily continuous. This means that the controlled process $X(t)$ can be changed instantly and planning for the future is no longer needed.

Hence, the optimisation problem can be treated locally and the trajectories of the optimal controlled process consist of the local optimal controls for each time interval $[T_n, T_{n+1}]$. We consider the scenario, in which $U = B$ tend to infinity in Section 3.4 and prove that the optimal controls tend to those of the limiting case.

We note that in this thesis neither the cost functional nor the environment depend on the action taken. However, this dependence was considered for PDMPs, see [25] and references therein, Continuous Time Markov Decision Processes (CDMPs) [32, 26].

2.2.3 Optimisation criteria

Generally, optimisation over infinite time horizon can be done in two different ways: via cost discounting or via time averaging. In this thesis, we are mainly interested in the time-average case. However, we also consider the discounted case, but mostly as an auxiliary step towards the main results. The approach we use is known in the literature as *vanishing discount approach*, see [15, 53, 6], for examples.

Suppose that $X(s)$ is any non-negative valued stochastic process defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ as the random environment $\{E_n\}$. We do not specify now how the process $X(s)$ is defined. Instead we take a simplistic convention that all the integrals and expectation in the definitions below exist. There is no loss of generality because our main goal is optimality, hence, we want the cost functionals to be correctly defined and to be finite.

Definition 2.2.4.

1. The *single period cost functional* is

$$\mathcal{C}(X(\cdot), \tau_0). \tag{2.5}$$

2. Suppose that $\beta > 0$ is a discount parameter. The *expected β-discounted cost over an infinite horizon* is

$$\mathcal{C}_\pi^\beta = \mathbb{E}\left[\int_0^\infty e^{-\beta t} g(X(s))ds + \sum_0^\infty e^{-\beta T_k} f(X(T_k), Y_{k-1})\right] \tag{2.6}$$

3. The *average cost functional* is defined by

$$\mathcal{C}_\pi = \limsup_{t\to\infty} \frac{1}{t}\mathbb{E}\left(\int_0^t g(X(s))ds + \sum_{i=1}^{\nu_t} f(X(T_i), Y_{i-1})\right). \tag{2.7}$$

It will be shown in Chapter 4 that under assumption (SMRE-II), for "sufficiently good" controlled processes the upper limit can be replaced with the limit

$$\limsup_{t\to\infty} \frac{\mathbb{E}\mathcal{C}_\pi(X(\cdot),t)}{t} = \lim_{t\to\infty} \frac{\mathbb{E}\mathcal{C}_\pi(X(\cdot),t)}{t}.$$

2.3 Classes of policies

In this section, we introduce the class of open-loop policies, the class of feedback policies and the class of fluid policies, which are of high interest in this work.

Recall that due to Assumption 2.2.4 the action space is $A = [-B, U]$.

Definition 2.3.1. Any measurable function $a : \mathbb{R}^+ \times \mathcal{Y} \times \mathbb{R}^+ \to A$ is called *an open-loop control*.
Assume that $Y_n = y$ and $X(T_n) = x$, then the dynamics of the process $X(t)$ until time T_{n+1} under the open-loop control a are

$$X(t) = X(T_n) + \int_0^{t-T_n} a(x, y, v)dv, \tag{2.8}$$

In terms of definition of action 2.4 it is equivalent to saying that action $a(x, y, v)$ is taken at time $v + T_n$. We consider only *admissible open-loop controls*, such that the process $X(t)$ never leaves $\mathbb{R}^+$.

Definition 2.3.2. A *policy* $\pi = (a_n)$ is an arbitrary sequence of admissible open-loop controls. A class of policies is denoted by Π.

Suppose that $(X(T_n), Y_n) = (x, y)$. Then action $a_n(x, y, v - T_n)$ is applied at the moment v. Random variables τ_n may have unbounded support, therefore, the controls $a_n(x, y, t)$ are defined for all $t \in \mathbb{R}^+$.

The definition of the open-loop policies can be easily explained. Suppose that at time T_n the controlled process is equal to $X(T_n)$. At this moment we get the value of Y_n, and (in case of Assumption SMRE-II) information about τ_{n+1}. Knowing the realisation of Y_n along with $X(T_n)$, we define the policy until time T_{n+1}. In other words, we choose the behaviour of the process $X(t)$ at the beginning of each time segment and do not change it until the end of this time segment.

Now we are ready to introduce optimal policies and optimal costs for both β-discounted and average cost functionals.

Definition 2.3.3. The *minimal* (β)*-discounted* and *the average costs* are

$$\mathcal{C}^\beta = \inf_{\pi\in\Pi} \mathcal{C}^\beta_\pi \quad \text{and} \quad \mathcal{C} = \inf_{\pi\in\Pi} \mathcal{C}_\pi.$$

Then the policy π is *optimal* in both cases if it attains the minimum.

If function a is left-continuous then the integral in the formula 2.8 has left derivative at each point and, in particular, has continuous sample paths.

We write $\mathbb{P}^u_{x,y}$ and $\mathbb{E}^u_{x,y}$, when we talk about a one step-transition under open-loop control u, and $\mathbb{P}^\pi_{x,y}$, $\mathbb{E}^\pi_{x,y}$ for the entire process.

Suppose that assumption (SMRE-II) holds, then the measures $\mathbb{P}^\pi_{x,y}$, $\mathbb{E}^\pi_{x,y}$ can be written in a closed form and the controlled process satisfies the following properties.

- Assume that the process $X(t)$ starts at point x at time 0 and the environment component is $Y_0 = y$ so that

$$\mathbb{P}^\pi_{x,y}(X(0) = x, Y_0 = y) = 1.$$

- The distribution of inter-arrival time τ_{n+1} depends only on Y_n and it results in

$$\mathbb{P}^\pi_{x,y}(T_{n+1} - T_n > t | T_0, \dots, T_n, X(T_0), \dots, X(T_n), Y_0, \dots, Y_n) = \mathbb{P}(T_{n+1} - T_n > t | Y_n).$$

- Moreover, given Y_n the value of Y_{n+1} is independent of the past and is independent of the value of τ_{n+1}. Hence, for the process $X(t)$ we have

$$\mathbb{P}^\pi_{x,y}\left((X(T_{n+1}), Y_{n+1}) \in B \times \{y_0\} | T_0, \dots, T_n, X(T_0), \dots, X(T_n), Y_0, \dots, Y_n\right)$$

$$= \mathbb{P}(Y_{n+1} = y_0 | Y_n) \times \mathbb{I}\left(\left[X(T_n) + \int_0^{T_{n+1}-T_n} a_n(X(T_n), Y_n, v)dv\right] \in B\right)$$

  for any $y_0 \in \mathcal{Y}$ and any measurable with respect to Lebesgue measure set B. Therefore, random sequence $\{(X(T_n), Y_n)\}$ forms a Markov chain.

The definition of open-loop controls was first introduced by Vermes (1985)[58]. This is the main class of optimal controls considered in most papers on optimisation since the 1990s. Another common option is *the class of feedback controls* where the control actions depend only on the position in control space and on the environment state.

Feedback controls formalise the idea of having "Markovian" controls, where the action does not depend neither on the history of the process nor, in particular, on the time.

We use the following agreement, we write

$$Y(t) := Y_{\nu_t},$$

where ν_t is the counting process for the sequence $\{T_n\}_{n\in\mathbb{N}}$.

Definition 2.3.4. Suppose that function $\psi : \mathbb{R}^+ \times \mathcal{Y} \to A$ is given. Then correspond-

ing process $X(t)$ with $X(0) = x$ is defined as

$$X(t) = x + \int_0^t \psi(X(s), Y(s))ds$$

We call the rule ψ the *feedback policy*.

As one can see, the class of feedback policies ψ is a proper subclass of the class of open-loop controls.

We present a number of important examples of open-loop controls and the corresponding processes $X(t)$.

Example 2.3.1.

1. Constant function $a_n(x, y, t) = 0$, results in constant deterministic process $X(t) = X(0)$.

2. Let function $k(x, z, t) : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ be

$$k(x, z, t) = \mathbb{I}_{\{x \le z\}} \min(x + Ut, z) + \mathbb{I}_{\{x > z\}} \max(x - Bt, z). \tag{2.9}$$

   This is a function, whose trajectory starts at level x at time 0 and moves with time t towards level z as quickly as possible. We adopt the convention that this function has a left derivative and write k' with the understanding that there are points where the real derivative does not exist, although the left derivative still does.

   Fix a positive number l and set

$$a_n(x, y, t) = k'_t(x, l, t)$$

   then the trajectory of the controlled process moves as quickly as possible towards level l and stays there forever once reached.

3. Alternatively, suppose there is a given sequence $\{l_n\}$ and

$$a_n(x, y, t) = k'_t(x, l_n, t).$$

   The corresponding process has trajectories moving towards level l_n in cycle number n.

4. Suppose now that $l : \mathcal{Y} \to \mathbb{R}^+$ and let

$$a_n(x, y, t) = k'_t(x, l(y), t) = \begin{cases} U, & \text{if} \quad x + Ut < l(y), \\ -B, & \text{if} \quad x - Bt > l(y), \\ 0, & \text{otherwise.} \end{cases} \tag{2.10}$$

A trajectory of the corresponding policy changes its destination at time moments T_n and moves towards a new level $l(Y_n)$ as quickly as possible. The level is determined by the environment state only.

Definition 2.3.5. Policies satisfying 2.10 form *the class of fluid policies.*

In Chapter 3 we show that under Assumpiton (MRE) the optimal policy exists and belongs to the class of fluid policies.

The name of this class goes back to fluid models [59, 39]. This is a set of models describing behaviour of a water tank of usually a finite capacity C, with different modes of water inflows and outflows. Suppose that water comes to the tank at constant rate U and level of water at time 0 is equal to x, then at time t the water level in the tank is equal to $k(x, C, t)$. If the water is being pumped out at a constant rate $-B$ then the level of water at time t is equal to $k(x, 0, t)$.

We work closely with fluid models in Chapter 3 and modify known results for stationary distribution of the water level in the tank for our model.

5. In fact, all the introduced controls in the above example but the third one are feedback controls.

2.4 Formal description of the basic power supply model

We formally define the objects from Section 2.2 one by one.

- The controlled process $X(t)$ stands for the additional level of conventional generation. It takes values in $\mathbb{R}^+$.

- Times τ_n stand for times between two consecutive shortfalls in power supply (spikes in the picture 1.2). Elements of the environment state space $\mathcal{Y}$ represent significantly different scenarios for the size of the next shortage (the height of the next spike). The pairs (τ_n, Y_n) should satisfy at least one of stochastic Assumptions 2.2.2.

  Example 2.4.1. One may take $\mathcal{Y} = \{1, 2\}$ and say that

– shortfalls smaller than 1500 MW are of type 1,

– otherwise they are of type 2.

This is illustrated in the picture below.

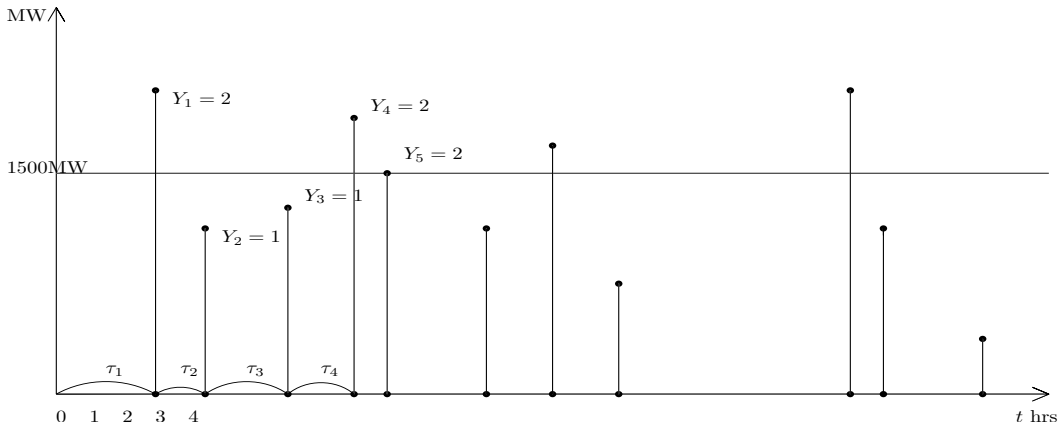


Figure 2.1: Example of the random environment 1

In this scenario random variables τ_n and Y_n are independent.

Example 2.4.2. Alternatively, one may take $\mathcal{Y} = \{1, 2, 3\}$ and assume that

– shortfalls smaller than 1500 MW are of type 1,

– shortfalls greater or equal than 1500 MW, where $\tau_n < 2$hrs are of type 2,

– shortfalls greater or equal than 1500 MW, where $\tau_n \geq 2$hrs are of type 3.

The figure below depicts this case.

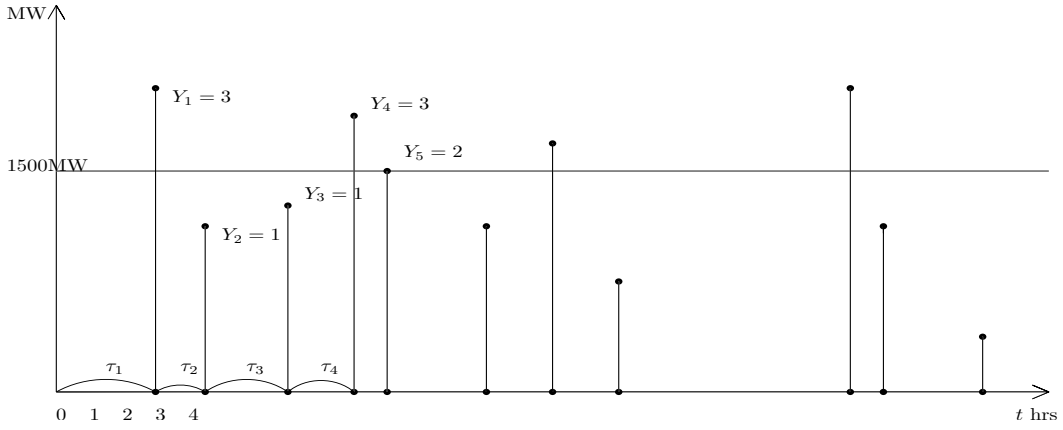


Figure 2.2: Example of the random environment 2

This is an example of a situation, in which random variables τ_n and Y_n are not independent.

The choice of the environment should depend on available statistical data and meteorological knowledge, as well as on the desired accuracy of the results.

It is easier to solve the problem numerically if space $\mathcal{Y}$ is small, see for example Section 3.6.1, although, additional elements in $\mathcal{Y}$ can provide a deeper insight to the behaviour of the spikes' height distribution and a higher level of accuracy of the solution.

- Positive constants $c_1 < c_2$ stand for the unit costs for conventional and imported power respectively. Running cost $g(x) = c_1 x$ stands for the cost of x units of conventional generation.

- The cost function $f(x, y)$ is modelled in the same manner as in Example 2.2.2.3. We repeat the definition with a few comments below.

  1. For each $y \in \mathcal{Y}$ there is given a probability distribution function F_y. This is a set of possible distributions for shortage amounts (height of spikes in picture 1.2).
  2. Suppose that there is given set $\{\chi_{n,y}\}_{n\in\mathbb{Z},y\in\mathcal{Y}}$, and its elements are independent. For each y elements of the sequence $\{\chi_{n,y}\}_{n\in\mathbb{Z}}$ are identically distributed with common distribution function F_y. If $Y_n = y$ then $\chi_{n+1} = \chi_{n+1,y}$, and given Y_n, it is independent of everything else:

  $$\mathbb{P}\{\chi_n > s | T_0, \ldots, T_n, X(T_0), \ldots, X(T_n), Y_0, \ldots, Y_n\} = \mathbb{P}\{\chi_n > s | Y_n\}.$$

  3. Therefore, function $f(x, y) = c_2\mathbb{E}(\chi_y - x)^+$ represents the expected loss at shortfall moment T_{n+1}, if $Y_n = y$ and $X(T_{n+1}) = x$. When the actual shortage in the power system is equal to χ_y and additional conventional supply is at level x, the net shortage $(\chi_y - x)^+$ must be covered by imported power at unit cost c_2.

  Recall from Example 2.2.2 that functions $f(x, y)$ and $g(x)$ are convex.

- Positive constants U and B stand for the ramp constraints of power generation plants. The power output cannot be changed immediately, because the procedure involves warming up or cooling down large combustion engines. This was also discussed in Section 1.2.

  The action space is $A = [-U, B]$. A single action a is the rate of change of conventional generation.

- The dynamics of the level of conventional generation $X(t)$ is represented through

the rate of change a given by

$$X(t) = X(0) + \int_0^t a(s)ds.$$

2.5 Class of bounded policies

We would like to take into consideration only bounded controlled processes $X(t)$. This is justified by applied nature of our problem ($X(t)$ stands for the level of additional conventional supply in the basic power supply model) and it makes the mathematical optimisation problem easier to solve, by bringing compactness to the state space of the controlled process. In this section, we discuss sufficient conditions on the boundedness of the optimal controlled process and summarise them in Assumption 2.5.1.

By definition, for any function $h : \mathbb{R}^+ \to \mathbb{R}^+$ we write $z = \operatorname{argmin} h(x)$ if for any $y \in \mathbb{R}^+$ holds $h(z) \leq h(y)$. Under assumption SMRE-I the random variables $\{\tau_n\}$ are i.i.d. For Assumption SMRE-II, under which the random variables $\{\tau_n\}$ are not necessarily i.i.d. we will assume that the reasoning holds for each of the distributions. Hence we only discuss i.i.d. case.

Suppose that each of the random variables $\{\tau_n\}$ is a mixture of a continuous random variable and a discrete random variable. Suppose that $t_1, t_2, \ldots, t_k$, are atoms with probabilities given by $\mathbb{P}(\tau_n = t_k) = p_k$ and $p(t)$ is the density function so that

$$\mathbb{P}(\tau_n \leq t) = \sum_{t_l \leq t} p_l + \int_0^t p(s)ds.$$

Let us focus on the single period optimisation problem. Suppose there exists a constant $K < \infty$ such that for any process $X(t)$ a new process $\widehat{X}(t) = \min(X(t), K)$ gives a non-larger value to the single period cost functional

$$\mathcal{C}(\widehat{X}(t), \tau_1) \leq \mathcal{C}(X(t), \tau_1).$$

This yields that for the average cost optimisation the new process $\widehat{X}(t)$ also improves the value of the cost functional, because the functional is represented as the averaged sum of functionals for single time periods. Therefore, the optimal controlled process cannot take values larger than K. We proceed by exploring assumptions which are sufficient for the existence of constant K.

Assume that the environment state is $Y_0 = y$ and the process $X(t)$ is observed on the interval $[0, \tau_1]$, then at τ_1 penalty $f(X(\tau_1), y)$ is charged. Let $\bar{F}(t) = \int_t^\infty p(s)ds$. If τ_1 is an absolutely continuous with respect to Lebesgue measure random variable then $\bar{F}$ is its tail distribution function. The charge associated with the running cost

may be expressed as

$$\mathbb{E}\int_0^{\tau_1} g(X(t))dt = \sum_{i=1}^{k} p_i \int_0^{t_i} g(X(t))dt + \int_0^{\infty} \left(\int_0^{s} g(X(t))dt\right) p(s)ds$$

$$= \sum_{i=1}^{k} p_i \int_0^{t_i} g(X(t))dt + \int_0^{\infty} \left(\int_t^{\infty} p(s)ds\right) g(X(t))dt$$

$$= \sum_{i=1}^{k} p_i \int_0^{t_i} g(X(t))dt + \int_0^{\infty} g(X(t))\bar{F}(t)dt,$$

The second equality is obtained by changing the integration order in accordance with Fubini's theorem. Hence, for the single period cost functional one may write

$$\mathcal{C}(X(t), \tau_1) = \mathbb{E}\left(\int_0^{\tau_1} g(X(t))dt + f(X(\tau_1), y)\right)$$

$$= \sum_{i=1}^{k} p_i \left(\int_0^{t_i} g(X(t))dt + f(X(t_i), y)\right) + \int_0^{\infty} \left(g(X(t))\bar{F}(t) + f(X(t), y)p(t)\right) dt.$$

We will tackle discrete and continuous parts separately.

2.5.1 Discrete case

Denote the part of the cost functional corresponding to the discrete part of the random variable by

$$\mathcal{C}_d = \sum_{i=1}^{k} p_i \left(\int_0^{t_i} g(X(t))dt + f(X(t_i), y)\right)$$

Instead of studying the minimum of $\mathcal{C}_d$ in the class of policies satisfying Assumption 2.2.4 we will have a look at each of the summands

$$\int_0^{t_i} g(X(t))dt + f(X(t_i), y).$$

Since g is increasing and U is the upper ramp constraint, minimum of the expression above is attained at a policy which satisfies $X(s) = X(t_i) - (t_i - s)U$. So one may look for a solution of

$$\min_{x \in \mathbb{R}^+} \int_0^{t_i} g(x - Ut)dt + f(x, y).$$

Denote the solution by $K_d(i, y)$ and let $K_d = \max_{i,y} K_d(i, y)$.

Lemma 2.5.1. *Suppose that $K_d < \infty$ then for any policy $X(t)$ holds*

$$\mathcal{C}_d(\min(X(t), K_d)) \leq \mathcal{C}_d(X(t)).$$

Proof. Functions $\int_0^{t_i} g(x - Ut)dt + f(x, y)$ are convex in x with their minimums at $K_d(i, y)$. Therefore for every summand we have

$$\int_0^{t_i} g(\min(X(t), K_d))dt + f(\min(X(t), K_d), y) \leq \int_0^{t_i} g(X(t))dt + f(X(t_i), y).$$

Hence, the similar holds for the convex combination of the summands and the lemma is proven. □

2.5.2 Continuous case

If the random variable is not pure discrete we may work with only the integral part from now on, because boundedness of the controlled process at all time moments, but t_k, implies boundedness at t_k due to continuity of the controlled process $X(t)$. One can divide both terms of the expression under the integral sign by $\bar{F}(t)$ to get

$$g(X(t)) + f(X(t), y)\frac{p(t)}{\bar{F}(t)}. \tag{2.11}$$

The expression is correctly defined, since $\bar{F}(t) = \mathbb{P}(\tau_1 > t) = 0$ yields $p(t) = 0$. Suppose now that

$$\sup_t \operatorname{argmin}_x \left(g(x) + f(x, y)\frac{p(t)}{\bar{F}(t)} \right) =: K(y) < \infty$$

Suppose that for some t holds $X(t) \geq K(y)$, then

$$g(X(t)) + f(X(t), y)\frac{p(t)}{\bar{F}(t)} \geq g(K) + f(K, y)\frac{p(t)}{\bar{F}(t))},$$

by convexity of the function and the definition of constant $K(y)$. Hence, the integral inequality holds

$$\begin{aligned}
\mathcal{C}(X, \tau_1) &= \int_0^{\infty} g(X(t)) + f(X(t), y)\frac{p(t)}{\bar{F}(t)}dt \\
&\geq \int_0^{\infty} g(\min(K(y), X(t))) + f(\min(K(y), X(t)))\frac{p(t)}{\bar{F}(t)}dt \\
&\geq \mathcal{C}(\min(X, K(y)), \tau_1).
\end{aligned}$$

The inequality is strict if the Lebesgue measure of the set of positive difference is positive

$$\mu\{t : X(t) > K(y)\} > 0.$$

If we assume that

$$\sup_{y\in\mathcal{Y}} K(Y) =: K < \infty. \tag{2.12}$$

then process $\min(X(t), K)$ gives a lower value to functional $\mathcal{C}$ than process $X(t)$. Therefore, if process $X(t)$ is a candidate for the optimal controlled process for $\mathcal{C}(X, \tau_0)$, then it cannot lie above level K.

The condition 2.12 is cumbersome and is difficult to check. Hence, we explore the sufficient conditions for it to hold, and summarise them in Assumption 2.5.1.

To proceed further, we split set $\mathcal{Y}$ into two subsets $\mathcal{Y}_1$ and $\mathcal{Y}_2$ such that

$$\mathcal{Y}_1 = \{y \in \mathcal{Y} : \arg\min_x f(x, y) < \infty\} \qquad \mathcal{Y}_2 = \mathcal{Y} \setminus \mathcal{Y}_1. \tag{2.13}$$

The other object we are interested in is *hazard rate function* $r(t) = \frac{p(t)}{\mathbb{P}(\tau_1 \geq t)}$. It is sometimes also called *failure rate*. The hazard rate function is one of the key definitions in insurance, finance and survival analysis. It is closely related with classification of tails for probability distributions. For the theory of heavy-tailed distributions one may look in [31].

We take into consideration two options:

1. $\sup r(t) < \infty$,
2. $\sup r(t) = \infty$.

Notice that, for the exponential distribution with parameter λ the hazard rate function $r(t) = \lambda$ and for Weibull distribution with parameter $k > 1$ holds $\sup r(t) = \infty$. Hazard rate functions can demonstrate very different behaviour. For example, monotone hazard rate functions are widely used in survival analysis and insurance. A distribution with a decreasing hazard rate function might be used to model infant mortality. A distribution with an increasing hazard rate function might be used to model the lifespan of species or products. Although, there exist distributions with non-monotone hazard rate functions.

It is a well known fact that for two strictly convex functions f_0 and f_1, their linear combination $\alpha_0 f_0 + \alpha_1 f_1$, where $\alpha_i \geq 0$ has the following property

$$\operatorname{argmin} \alpha_0 f_0 + \alpha_1 f_1 \in [\min_{i=0,1}(\operatorname{argmin} f_i), \max_{i=0,1}(\operatorname{argmin} f_i)].$$

Moreover, $\operatorname{argmin} \alpha_0 f_0 + \alpha_1 f_1 = \operatorname{argmin} f_i$ in the case $\alpha_i/\alpha_{1-i} = \infty$.

Assume that $K < \infty$. Let us show that $\mathcal{Y}_2 \neq \emptyset$ yields $\sup(r(t)) := R < \infty$. Function g is non-decreasing. Hence the expression

$$\operatorname{argmin}_x g(x) + sf(x, y)$$

can be arbitrarily large for large s, because $\operatorname{argmin} f(x,y) = \infty$ due to the definition of set $\mathcal{Y}_2$. Therefore, if set $\mathcal{Y}_2$ is non-empty, we have to assume that $\sup r(t) = R < \infty$. In this case, it is sufficient to have

$$\sup_{y \in \mathcal{Y}_2} \operatorname{argmin}_x \left(g(x) + Rf(x,y)\right) < \infty.$$

For set $\mathcal{Y}_1$ we consider both options $R = \infty$ and $R < \infty$. In the first case one has

$$\sup_{t,y} \operatorname{argmin}_x g(x) + r(t)f(x,y) = \sup_{y} \operatorname{argmin}_x f(x,y) < \infty.$$

For the second case it is enough to have a weaker condition satisfied

$$\sup_{y} \operatorname{argmin}_x g(x) + Rf(x,y) < \infty.$$

We now summarise all the findings in the following assumption. As we have shown above this assumptions yields that a candidate for the optimal controlled process is bounded from above.

Assumption 2.5.1.

- If random variable τ_1 is discrete with atoms in t_i then

$$\max_{i,y} \arg\min_{x \in \mathbb{R}^+} \int_0^{t_i} g(x - Ut)dt + f(x,y) \leq \infty.$$

- If random variable is absolutely continuous with respect to Lebesgue measure and the set $\mathcal{Y}_2 \neq \emptyset$, then $R < \infty$ and

$$\sup_{y \in \mathcal{Y}} \operatorname{argmin}_x \left(g(x) + Rf(x,y)\right) < \infty.$$

- If random variable is absolutely continuous with respect to Lebesgue measure and the set $\mathcal{Y}_2 = \emptyset$, and $R = \infty$ then

$$\sup_{y \in \mathcal{Y}} \operatorname{argmin}_x f(x,y) < \infty.$$

- If random variable is absolutely continuous with respect to Lebesgue measure and the set $\mathcal{Y}_2 = \emptyset$, and $R < \infty$ then

$$\sup_{y \in \mathcal{Y}} \operatorname{argmin}_x \left(g(x) + Rf(x,y)\right) < \infty.$$

From now on we will only consider models for which Assumption 2.5.1 holds.

Chapter 3

Exponential inter-arrival times

In this chapter, we consider the model under the Assumption (MRE). The cost functional given by formula 2.7 is comprised of two components: the terminal charges and the running cost. We show that under the assumption (MRE) these two components are similar in nature and the analysis can be done for both of them in the same manner.

To make the thesis self-contained, we describe the class of models known as Stochastic Fluid Programs (SFP), introduced by N.Bäuerle in [6] and [7] and summarise the main results and assumptions for it. We proceed by showing that under assumption (MRE) our model is a particular case of (SFP). The results from [6] and [7] are then used to show the existence of the optimal control. Our main result then states the uniqueness and describes the exact form of the solution. In Theorem (3.3.8) we prove that the unique optimal policy belongs to the class of fluid policies. The content of this part of the Chapter was presented at the *Modern trends in controlled stochastic processes* conference in Liverpool in July 2015 and it was later published in [18].

To find the optimal control numerically one may restrict their search to the class of fluid controls only. We recall that an element of this class is a policy, given by a number of levels $\{l(y)\}_{y\in\mathcal{Y}}$. If $Y_n = y$ then the controlled process moves towards the level $l(y)$ as quickly as possible and then stays at this level until the state variable changes to Y_{n+1}.

Section (3.5) provides a system of differential equations that describes stationary distributions for controls in the class of fluid policies (see Example 2.3.1.2). The functions in the equations are not continuous, which makes the problem more interesting. We discuss a possible way to bypass this complication and obtain the solution in Section 3.5.1. The time-average cost functional in the class of fluid policies can then be written as a function of levels $\{l(y)\}$. To conclude this chapter, we present some closed-form solutions for special cases with the environmental space $\mathcal{Y}$, such that $|\mathcal{Y}| = 2, 3$.

3.1 The cost functional in the exponential case

In this subsection, we explain why the average cost optimisation problem under Assumption (MRE) can be replaced by an equivalent one with a modified cost functional. This cost functional has only the component corresponding to the continuous charge.

Lemma 3.1.1. *Suppose that τ is an exponential random variable with parameter λ. Then for any function $g \geq 0$ and deterministic function $X : \mathbb{R}^+ \to \mathbb{R}^+$ the following holds*

$$\mathbb{E}\int_0^\tau g(X(t))dt = \mathbb{E}\frac{g(X(\tau))}{\lambda}.$$

Proof. Recall that the exponential distribution is the only probability distribution for which the tail probability and the density $p(t)$ are proportional

$$\lambda\mathbb{P}\{\tau \geq t\} = p(t).$$

Using this and the change of the integration order one has

$$\begin{aligned}\mathbb{E}\int_0^\tau g(X(t))dt &= \int_0^\infty \int_0^\tau \lambda g(X(t))dt\exp(-\lambda\tau)d\tau \\ &= \int_0^\infty \left(\int_t^\infty \lambda\exp(-\lambda\tau)d\tau\right) g(X(t))dt \\ &= \int_0^\infty g(X(t))\exp(-\lambda t)dt = \mathbb{E}\frac{g(X(\tau))}{\lambda}.\end{aligned}$$

□

Remark 3.1.2. The proposition still holds if the nonnegativity of g is replaced with the absolute integrability.

We write $Y(t)$ for a continuous version of Markov chain Y_n, so that $Y(t) = Y_n$ for $t \in [T_n, T_{n+1})$. Suppose that the process $Y(t)$ is a uniformised Markov process. This means that $T_{n+1} - T_n$ are i.i.d. exponentially distributed with parameter λ. As before $\nu_t = \max\{n : T_n \leq t\}$.

Suppose that the environment satisfies Assumption 2.5.1 and any candidate for the optimal controlled process $X(t)$ does not leave a compact set $[0, K]$ a.s. The time-average version of the cost functional 2.7 can be written as

$$\begin{aligned}\limsup_{t\to\infty}\frac{\mathbb{E}\mathcal{C}(X(\cdot),[0,t])}{t} &= \limsup_{t\to\infty}\mathbb{E}\frac{\int_0^t g(X(s))ds + \sum_{i=1}^{\nu_t} f(X(T_i), Y_{i-1})}{t} \\ &= \limsup_{t\to\infty}\mathbb{E}\frac{\int_0^t g(X(s)) + \lambda f(X(s), Y(s))ds}{t} - \mathbb{E}\frac{\int_{T_{\nu_t}}^t \lambda f(X(s), Y(s))ds}{t}\end{aligned}$$

$$= \limsup_{t\to\infty} \mathbb{E}\frac{\int_0^t g(X(s)) + \lambda f(X(s), Y(s))ds}{t}.$$

The last equality holds since

$$0 \le \mathbb{E}\frac{\int_{T_{\nu_t}}^t \lambda f(X(t), Y(t))dt}{t} \le \frac{\mathbb{E}(T_{\nu_t+1} - T_{\nu_t}) \max_{x\in[0,K], y\in\mathcal{Y}} f(x,y)}{t} \le \frac{\max_{x\in[0,K], y\in\mathcal{Y}} f(x,y)}{\lambda t}.$$

By a similar reasoning

$$\limsup_{t\to\infty} \frac{\mathbb{E}\mathcal{C}(X(\cdot), [0,t])}{t} = \limsup_{t\to\infty} \mathbb{E}\frac{\sum_{i=1}^{\nu_t} \lambda^{-1} g(X(T_i)) + f(X(T_i), Y_{i-1}) + \int_{T_{\nu_t}}^t g(X(t))dt}{t}$$

$$= \limsup_{t\to\infty} \mathbb{E}\frac{\sum_{i=1}^{\nu_t} \lambda^{-1} g(X(T_i)) + f(X(T_i), Y_{i-1})}{t}.$$

The similar relations hold for the infinite-time discounted cost functional as well. Therefore, one may take functions

$$\tilde{g}(x) = 0 \text{ and } \tilde{f}(x,y) = \lambda^{-1} g(x) + f(x,y) \tag{3.1}$$

and consider a new model, where the charges apply only at points T_n. Or similarly, by letting

$$\widehat{g}(x,y) = g(x) + \lambda f(x,y) \text{ and } \widehat{f}(x,y) = 0, \tag{3.2}$$

we reduce it to a model with the running cost only.

3.2 Stochastic fluid programs

We discuss the class of models, stochastic fluid programs (SFP), introduced in Bäuerle [6], [7] and show that under the assumption (MRE) our model might be viewed as a particular case of (SFP). Recall that the assumption (MRE) states that $\{T_n\}_{n\in\mathbb{Z}}$ is a Poisson process with a constant intensity and the sequence $\{Y_n\}_{n\in\mathbb{Z}}$ is a Markov chain. The notation in this section is different from the notation introduced in Bäuerle in order to keep the notation throughout the thesis consistent.

3.2.1 Definition

Let $\mathcal{Y}$ be a countable set and Q a generator for a Markov process $Y(t)$ that takes values in $\mathcal{Y}$. We refer to $Y(t)$ as *environment process*. Denote by T_n the jump times of the environment process $Y(t)$. We assume that the process $Y(t)$ is a uniformised Markov process, and elements of the sequence $\{T_{n+1} - T_n\}_{n\ge 0}$ are exponentially distributed

with a common parameter λ (see Section 3.7 for the definition).

Remark 3.2.1. In Bäuerle [6], the state space $\mathcal{Y}$ is assumed to be finite. Although, if the state space is countable and process $Y(t)$ is uniformisable (see definition 3.7.1), the main results of the paper still hold.

Let a closed set $S \subset \mathbb{R}^n$ be the *state space* of the controlled process and suppose that $\mathcal{B}(S)$ is the Borel σ-algebra on S. A compact and convex set $A \subset \mathbb{R}^k$ is the *action set* of the system.

The dynamics for SFP generalise dynamics for our problem. Suppose that for each $y \in \mathcal{Y}$ there is a linear function $b^y : A \to \mathbb{R}^n$. The controlled process $X(t)$ is defined as follows. Let $X(0) = x_0 \in S$. Suppose that at moment t action $a(t) \in A$ is taken, then the dynamics of the controlled process $X(t)$ on time interval $[0, T_1]$ are given by

$$X(t) = X(0) + \int_0^t b^y(a(s))ds. \tag{3.3}$$

We use the same terminology as in the last chapter.

Any measurable function $a : S \times \mathcal{Y} \times \mathbb{R}^+ \to A$ is called *an open-loop control*. Assume that $Y(s) = y$ for $s \in [T_n, T_{n+1}]$ and $X(T_n) = x$, then the dynamics of the process $X(t)$ in $[T_n, T_{n+1}]$ under the open-loop control a are

$$X(t) = X(T_n) + \int_0^{t-T_n} b^y(a(x, y, s))ds, \quad \textit{for} \quad T_n \le t \le T_{n+1}. \tag{3.4}$$

This is the same as to say that action $a(x, y, s)$ is taken at time $s + T_n$. If function a is left-continuous in s then the trajectories of the controlled process have a left derivative at each point and, in particular, have continuous sample paths. Sometimes we write $X_a(t)$ to underline the dependence of the process on the open-loop control a.

As before we are also interested in *the class of feedback controls*. For feedback controls the action taken depends only on the state (x, y) and neither on the history of the process nor, in particular, on the time.

Suppose that function $\psi : S \times \mathcal{Y} \to A$ is given. Then corresponding process $X_\psi(t)$ with $X(0) = x$ is defined as

$$X_\psi(t) = x + \int_0^t b^{Y(s)}(\psi(X(s), Y(s)))ds$$

We call the rule ψ *feedback policy*.

As one can see, the class of feedback policies ψ is a proper subclass of the class of open-loop controls.

We consider only *admissible controls*, such that the process $X(t)$ never leaves the

set S:

$$D(x,y) = \{a : \forall t \geq 0 \quad X(t) \in S, \text{ if } X(0) = x, Y(0) = y\}.$$

A *policy* $\pi = (a_n)$ is an arbitrary sequence of admissible open-loop controls. If $(X(T_n), Y(T_n)) = (x, y)$ then action $a_{\nu_s}(x, y, s - T_{\nu_s})$ is applied at the moment s.

Probability measures $\mathbb{P}^{\pi}_{x,y}$ and expectations $\mathbb{E}^{\pi}_{x,y}$ correspond to the process starting at (x, y) and controlled by policy π. We write $\mathbb{P}^{a}_{x,y}$ and $\mathbb{E}^{a}_{x,y}$, when we talk about a one step-transition under open-loop control a.

Since the exponential random variable $T_{n+1} - T_n$ can take arbitrary large values, the controls $a_n(x, y, t)$ should be defined for all $t \in \mathbb{R}^+$. This might explain their name open-loop controls.

Remark 3.2.2. It is convenient to have the same action set for all states (x, y), but different dynamics. The functions b^y make this possible. The dynamics of the model depend upon linear functions b^y, therefore the left-derivative of the controlled process is different for the same actions taken at different environment states. As sets $D(x, y)$ can be different, one should make sure that the controlled process does not leave the space S.

Finally, we introduce *the cost rate function*

$$c : S \times \mathcal{Y} \times A \to \mathbb{R}^+,$$

which is assumed to be convex in x and a for all $y \in \mathcal{Y}$.

Definition 3.2.1. Collection $(S, \mathcal{Y}, A, b, Q, c)$ defines a stochastic fluid program.

Remark 3.2.3. Our model is an SFP given by $(\mathbb{R}^+, \mathbb{Z}, [-B, U], b(x) = x, Q, \widehat{g})$, where $\widehat{g}$ is given by equation 3.2. Therefore, to use the results proved for SFP we need only to verify the assumptions required for them.

In the last part of this section we define the β-discounted and average cost functionals. The definitions are very close to the ones introduced for our model in Definition 2.2.4.

Definition 3.2.2. For a policy π let $\pi_s = a_{\nu_s}(X(T_{\nu_s}), Y(s), s - T_{\nu_s})$ be the continuous version of the policy.

1. Suppose that $\beta > 0$ is a discount parameter. The *expected β-discounted cost over an infinite horizon* under policy π starting at (x, y) is

$$\mathcal{C}^{\beta}_{\pi}(x, y) = \mathbb{E}^{\pi}_{x,y}\left[\int_0^{\infty} e^{-\beta t} c(X(s), Y(s), \pi_s) ds\right]$$

2. For a fixed initial value (x, y) the *average cost functional* under policy π is defined by
$$\mathcal{C}_\pi(x,y) = \limsup_{t\to\infty} \frac{1}{t}\mathbb{E}^\pi_{x,y}\int_0^t c(X(s), Y(s), \pi_s)ds.$$

3. The *minimal* (β)*-discounted* and *the average costs* are
$$\mathcal{C}^\beta(x,y) = \inf_\pi \mathcal{C}^\beta_\pi(x,y) \quad \text{and} \quad \mathcal{C}(x,y) = \inf_\pi \mathcal{C}_\pi(x,y).$$

Then the policy π is *optimal* in both cases if it attains the minimum.

3.2.2 Some remarks on SFP

In this subsection we discuss the importance of the linear functions b^y. It might appear that one could eliminate them by taking a larger action space, however, we show that this is not true. The presence of linear function b^y makes the problem more complicated and broadens the class of physical models where the formalism of SFP might be applied.

Let us attempt to enlarge the action space. Optimisation in a larger class is at least as efficient. However, the modified set of controls might make no physical sense. We present an example to prove this point.

Suppose we have a modification of the model. For a set $Z \subset \mathbb{R}^k$ we write $conv(Z)$ for its convex envelope, which is the minimal convex set containing Z. Let $\widetilde{A} = conv(\cup_{y\in\mathcal{Y}} b^y(A))$. Consider a model with linear functions $\hat{b}^y(u) = u$ defining the dynamics and action set $\widetilde{A}$. The open-loop controls for this model are $\tilde{a} : S\times\mathcal{Y}\times\mathbb{R}^+ \to \widetilde{A}$.

The new model allows more open-loop controls, because for each open-loop control a one can define $\tilde{a}(x, y, s) = b^y(a(x, y, s)) \in \widetilde{A}$. Therefore, the optimization in the new class should be at least as efficient. However, some of the new controls may have no physical explanation. In the next example, we show how functions b^y may help to introduce the natural restrictions in a model, and show that the proposed enlargement of the action space might lead to the loss of the physical sense.

Example 3.2.4. We discuss an example introduced in Sethi et al. [54]. Suppose there is a manufacturing machine, which is able to produce n different types of products. The machine can work in a number of modes with different production capacity. The environment space $\mathcal{Y}$ formalises the idea of these modes, and for each $y \in \mathcal{Y}$ there is a known finite production capacity $\beta(y)$.

The state space is $S = \mathbb{R}^n$. A vector $x \in S$ gives inventory/backlog levels for all the products. The decision space A stands for percentages of the total production

assigned to each of the products at particular time moment. It can be written as

$$A = \{a \in \mathbb{R}^n : a_i \geq 0, \sum a_i \leq 1\}.$$

The demand vector $\mu \in \mathbb{R}^k$ is assumed to be constant.

The rate of change of the inventory levels can be written as

$$b^y(a) = \beta(y)a - \mu.$$

So after integration one gets

$$X(t) = X(0) + t(\beta(y)a - \mu)$$

gives the vector of production levels at time t.

Suppose now that along with this model we introduce the model with extended action space $\widetilde{A}$. We can see that some of its elements violate the assumption that the sum of the productions should be less than the total possible production rate $\beta(y)$. In real life, sometimes it would be better to produce more, but it is not always possible due to the system constraints.

Hence, the formalism of the functions b^y is absolutely relevant in this case.

3.2.3 Main results and assumptions for SFP

This section summarises the results from Bäuerle [6], [7], which are relevant to our work. We start by stating assumptions and results for the β-discounted optimality. We proceed further by providing additional assumptions to formulate Theorem (3.2.7), which states the existence of the optimal solution for the average cost functional and provides the average cost optimality equation (ACOE 3.6).

For our model, one of these assumptions does not hold. Nevertheless, we explain why the results still hold. Finally, we take a closer look at ACOE and explain it in a less formal manner.

We use notation $C(x, y, a)$ for the expected cost between two jumps under open-loop control a. We can write

$$C(x, y, a) = \int_0^\infty \exp(-\lambda t) c(X_a(s), y, a(s)) ds.$$

According to Lemma (3.1.1) we can also use the representation

$$C(x, y, a) = \frac{\mathbb{E}c(X_a(T_1), Y_1)}{\lambda}.$$

According to Bäuerle [7], the following set of assumptions are needed to establish the

existence of a β-discounted optimal policy.

Assumptions 3.2.1.

(a) The state space S of the controlled process is closed. The action space A is convex and compact in the Euclidean norm.

(b) The function $c(x, y, a)$ is lower semicontinuous and convex in x and a for all $y \in \mathcal{Y}$;

(c) There exists a policy π^β such that $\mathcal{C}^\beta_{\pi^\beta}(x, y) < \infty$ for all $(x, y) \in S \times \mathcal{Y}$.

Remark 3.2.5. It is well known that a convex function is Lipshitz on open subsets of its domain (see Rockafellar [50]). Therefore, the assumption on the lower semicontinuity may be replaced with the continuity upto the boundary.

The following theorem states the existence of the optimal policy for $\beta > 0$ and provides the β-discounted cost optimality equation.

Theorem 3.2.6 (Theorem 4 and Lemma 5 in Bäuerle ([7]))**.**
Suppose that assumptions (3.2.1) hold, then

- $\mathcal{C}^\beta$ *is the minimal solution of*

$$h^\beta(x, y) = \min_{a \in D(x,y)} \left[C(x, y, a) + \mathbb{E}^a_{x,y} \exp(-(\beta + \lambda)T_1) h^\beta(X(T_1), Y(T_1)) \right]. \quad (3.5)$$

- *There exists a minimiser* $a^\beta : S \times \mathcal{Y} \times \mathbb{R}^+ \to A$ *of (3.5) and the stationary policy* $(a^\beta, a^\beta, a^\beta, \ldots)$ *is* β*-discounted optimal.*

- *Solution* $\mathcal{C}^\beta(x, y)$ *is convex in* x *for all* y*.*

Throughout the literature, equations alike to equation (3.5) are generally referred to as the *Bellman equation.*

Existence of the optimal solution in the average-cost problem is obtained by using a technique known as vanishing discount approach. In this approach, the solution is found as a limit of a sequence of the solutions of discounted problems along a certain subsequence $\beta_n \to 0$. Hence, Assumptions (3.2.1) are required to hold for any $\beta > 0$. The existence of β-discounted optimal stationary policy requires relatively mild assumptions, however, the existence of the time-average optimal stationary policy requires a more complicated and restrictive set of assumptions.

Fix $(x^*, y^*) \in S \times \mathcal{Y}$ and let

$$h^\beta(x, y) = \mathcal{C}^\beta(x, y) - \mathcal{C}^\beta(x^*, y^*).$$

Assumptions 3.2.2.

(a) The state space is $S = \mathbb{R}^n$ or $S = (\mathbb{R}^+)^n$ and the action space A is convex and compact in the Euclidean norm;

(b) The cost function $c(x, y, a)$ is convex and continuous in x and a for all $y \in \mathcal{Y}$;

(c) For all $\beta > 0$ there exists a policy π^β such that $\mathcal{C}^\beta_{\pi^\beta}(x, y) < \infty$ for all $(x, y) \in S \times \mathcal{Y}$;

(d) There exists a policy π such that $\mathcal{C}_\pi(x, y) < \infty$ for all $(x, y) \in S \times \mathcal{Y}$;

(e) There exist constants $L \in \mathbb{R}$ and $\hat{\beta} > 0$ and upper semicontinuous function $M : S \times \mathcal{Y} \to \mathbb{R}^+$ such that for all $(x, y) \in S \times \mathcal{Y}$ and $0 \leq \beta \leq \hat{\beta}$ the following holds

$$L \leq h^\beta(x, y) \leq M(x, y).$$

Moreover, for all admissible open-loop controls $a \in D(x, y)$ the following holds

$$\mathbb{E}^a_{x,y} M(X(T_1), Y(T_1)) < \infty.$$

(f) If $S = (\mathbb{R}^+)^n$ then for the β-discounted optimal policies π^β the following holds

$$\mathcal{C}^\beta_{\pi^\beta}(x, y) \quad \text{is increasing in } y.$$

Theorem 3.2.7 (Theorem 4,5, Bäuerle ([6]))**.** *Under assumptions (3.2.2) the following is true.*

1. *There exists a constant $\rho \geq 0$ and a convex function $h : S \times \mathcal{Y} \to \mathbb{R}$ such that the average optimality equation holds for all $(x, y) \in S \times \mathcal{Y}$:*

$$\lambda\rho + h(x, y) = \inf_{a \in D(x,y)} \left[C(x, y, a) + \mathbb{E}^a_{x,y} h(X(T_1), Y(T_1)) \right]. \tag{3.6}$$

2. *There exists a function $a^0 : S \times \mathcal{Y} \times \mathbb{R}^+ \to A$, such that the infimum of the right-hand side of equation (3.6) for (x, y) is attained for it. There also exists a sequence $\beta_n \to 0$ such that*

$$a^0(x, y) = \lim_{n \to \infty} a^{\beta_n}(x, y), \tag{3.7}$$

where a^{β_n} are solutions for the discounted problems with parameters β_n.

3. *Suppose that $a_0(x, y, t)$ can be given by a feedback control ψ_0 and $c(x, y, a)$ does not depend on a.*

Then a_0 is an average optimal policy and ρ is the minimal average cost.

Remark 3.2.8. Controls $a^\beta(x, y, t)$ are functions of t. Their convergence in 3.7 is understood in the Young topology (see Section 3.8).

For our problem, Assumption 3.2.2.(f) fails to hold. Indeed if it was true, then $\mathcal{C}^\beta(x, y)$ would attain its minimum at $x = 0$. Therefore the optimal controlled process would stay at level 0 for all β and y.

This assumption was used in Lemma 3 from [6] to show equi-continuity of functions h^β for small β. Then Lemma 3 was used in the proof of Theorem (3.2.7). It is easy to see that this assumption is not necessary for equi-continuity and was used for the means of the proof. Take a set of non-monotone equi-continuous functions defined on $\mathbb{R}^n$ and restrict them to $(\mathbb{R}^+)^n$. The restriction is not necessarily monotone.

Lemma 3 from [6] was also stated in Fernández-Gaucherand et al. [29] for convex $c(x, y, a)$ and $S = \mathbb{R}^n$, however there is no proof presented.

In Bäuerle [6] statement of the lemma was first proved for $S = \mathbb{R}^n$ and then for $S = (\mathbb{R}^+)^n$ by extending functions h^β in a convex manner on $\mathbb{R}^+$ and applying the first result to the extension.

In the following modification of the lemma we show that equi-continuity still holds if Assumption 3.2.2.(f) is removed from the list. Therefore, the statement of Lemma 3 and Theorem 4 from [6] still hold without Assumption 3.2.2.(f).

Lemma 3.2.9. *Suppose that $S = (\mathbb{R}^+)^n$ and that Assumptions 3.2.2.(a)-(e) hold. Then functions $h^\beta(x, y)$ are equi-Lipschitz in x on closed bounded subsets of S.*

To prove the lemma, we apply results on the extension of convex functions based on results in Yan [67].

Extension 3.2.10 (Theorem 3.1. Yan [67])**.** *Suppose f is a convex function on an arbitrary set $S \subset \mathbb{R}^n$. Then under either of the following assumptions, f can be extended to a convex function on the convex hull $conv(S)$:*

(a) f is bounded below,

(b) S contains a point in the relative interior of the convex hull $conv(S)$.

The new function is defined as

$$\hat{f}(x) = \inf\left\{\sum_k \lambda_k f(x_k), \text{ where } k \in \mathbb{N}, \lambda_k \geq 0, \sum_i \lambda_i = 1, x_k \in S, \sum \lambda_k x_k = x\right\}. \tag{3.8}$$

Remark 3.2.11. Note that both the assumptions hold for our problem.

Extension 3.2.12 (Theorem 4.1. Yan [67], Dragomirescu and Ivan [27]). *A convex function on a bounded convex subset can be extended to a convex function on the whole linear space if and only if it is a Lipschitz function.*
The new function is defined as

$$\hat{f}(x) = \sup\{\lambda f(y) + (1-\lambda)f(z), x = \lambda y + (1-\lambda)z, y, z \in S, \lambda \geq 1\}. \tag{3.9}$$

Remark 3.2.13. It is shown in the proof in [67] that $\hat{f}(x)$ takes only finite values.

Proof of Lemma 3.2.9. We fix $y \in \mathcal{Y}$ for the entire proof and suppress it in the notation for the proof. The proof consists of two parts. Firstly, we extend the convex functions $h^\beta(x)$ on $\widehat{S} = [-1, \infty)^n$ and show that they are still uniformly bounded by some function $\widehat{M}(x)$. Then, we apply Theorem 10.6 from Rockafellar [50], which states that any collection of uniformly bounded convex functions on an open bounded set S is equi-Lipschitz on S.
Note: if we apply Theorem 10.6 from Rockafellar [50] to the interior of S we show continuity only on those closed bounded sets, which do not intersect with the boundary of S. Hence, we enlarge the set S and then apply the theorem to the interior of the larger set $\widehat{S}$.

Note that, local convexity in a convex set yields global convexity in it. Therefore, we can construct extensions of functions $h^\beta(x)$ onto a sequence of increasing subsets $S^k \nearrow \widehat{S}$, step by step, making sure that they stay convex after every step.

The extension of functions $h^\beta(x)$ is described in an algorithmic way. Precisely, we show that any β function h^β on S can be extended to function $\widehat{h}^\beta$ on $\widehat{S}$ and that there exists a function $\widehat{M}$ (which does not depend on β) such that $\widehat{h}^\beta(x) \leq \widehat{M}(x)$. For the convenience of the proof we take $M(x) = \sup_\beta(h^\beta(x))$, which is finite due to Assumptions (3.2.2). It is convex as a supremum of convex functions.
Step 1

Extend h^β and M from $[0,1]^n$ to $[-1,1]^n$ by applying Extension 3.2.12. Let $S^1 = S \cup [-1,1]^n$. The extended function h^β is convex on S^1. By formula (3.9) we see that for extended functions $\widehat{f}(x) \leq \widehat{M}(x)$.
Step 2

Extend h^β from $S^1 \cap [-1,2]^n$ to $[-1,2]^n$ by applying Extension 3.2.10. Let $S^2 = S \cup [-1,2]^n$. The extended function h^β is convex on S^2. For $x \in [-1,2]^n$ holds

$$h^\beta(x) \leq \max_{z \in S^1 \cap [-1,2]^n} \widehat{M}(z),$$

thus the extended function h^β is bounded.
Step k

Extend f from $S^{k-1} \cap [-1,k]^n$ to $[-1,k]^n$ by applying the Extension 3.2.10. Let

$S^k = S \cup [-1, k]^n$. Function h^β is convex and bounded on S^k.

It is clear that S^k is a sequence of sets increasing to $\widehat{S}$. Hence we have obtained a convex extension of functions $h^\beta(x)$ on $\widehat{S}$ and a function $\widehat{M}(x)$ such that for all small β holds $h^\beta(x) \leq \widehat{M}(x)$. The last step is to apply Theorem 10.6 from Rockafellar [50] to the interior of $\widehat{S}$. □

Since in Theorem 3.2.7 the optimal rule is found as a limit of a certain subsequence of rules for the discounted problems with $\beta_k \to 0$, Assumptions 3.2.1 for large β can be eliminated. We prove a stronger result.

Proposition 3.2.14. *Assumption 3.2.2.(d) yields 3.2.2.(c) for small β.*

Proof. Corollary (1c) on page 183 of Widder [62] in our notation says that for any policy π and any initial state (x, y) the following holds

$$\limsup_{\beta \to 0} \beta \mathcal{C}_\pi^\beta(x, y) \leq \mathcal{C}_\pi(x, y).$$

By the definition of the upper limit for $\varepsilon = 1$ there exists $\bar{\beta}$ such that

$$\sup_{\beta \leq \bar{\beta}} \beta \mathcal{C}_\pi^\beta(x, y) \leq \mathcal{C}_\pi(x, y) + 1,$$

and, therefore, given the assumption, (3.2.2.(d)) this yields

$$\mathcal{C}_\pi^\beta(x, y) \leq \frac{\mathcal{C}_\pi(x, y) + 1}{\beta} < \infty \text{ for } \beta \leq \bar{\beta}.$$

□

Usually, the average optimality equation is used as a tool for establishing the existence of the optimal solution in optimization problems. On rare occasions it is possible to solve the equation explicitly. Sometimes it is possible to guess a solution and by using the equation (3.6) show that any other policy will be no better.

One should think of ρ, h and a_0 in the following manner. Policy a_0 is the optimal policy minimising $\mathcal{C}(x_0, y_0)$. It is stationary and does not depend on the initial point (x_0, y_0). Constant ρ is the optimal long time average cost corresponding to the optimal policy a_0. Recall that the length of a period $T_{n+1} - T_n$ is exponentially distributed with parameter λ. Hence, the average cost over interval $[T_n, T_{n+1}]$ under policy a_0 is equal to $\lambda\rho$. On the right hand side, $C(x, y, a_0)$ represents the cost for one single time period.

Function $h(x, y)$ is a little more difficult to explain. It is usually referred to as *relative value function*. One might think about a regeneration structure in the following manner. Consider two optimal strategies starting from (x, y) and (x^*, y^*). Then cost $h(x, y)$ is the difference we pay before the control trajectories couple.

3.3 Our main results

In this subsection we assume that $S \subset \mathbb{R}$. In all statements we assume that the action space A is convex and compact. Hence, it is a closed connected subset of $\mathbb{R}^n$. The images $b^y(\mathcal{A})$ are one-dimensional convex compact sets for all $y \in \mathcal{Y}$, thus they are closed intervals. For each $y \in \mathcal{Y}$ point 0 either belongs to it or the whole interval lies to one or the other side from 0. For our convenience we need the following.

Assumption 3.3.1. We assume that $0 \in b^y(A)$ for all $y \in \mathcal{Y}$, so the control process $X(t)$ is allowed to move up and down, and also stop, regardless of the state of the environment.

Under assumption (3.3.1) there exist functions $B, U : \mathcal{Y} \to \mathbb{R}^+$ such that

$$b^y(\mathcal{U}) = [-B(y), U(y)].$$

Recall function $k(x, z, t) : \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}^+$ given by formula 2.9 as

$$k(x, z, t) = \mathbb{I}_{\{x \le z\}} \min(x + Ut, z) + \mathbb{I}_{\{x > z\}} \max(x - Bt, z).$$

Its trajectory with respect to t starts at point x then increases(decreases) to level z at the maximal possible pace in accordance with the ramp constraints.

We modify functions k in accordance with the maximal possible constraints on the derivative $b^y(a)$ at the environment state y:

$$\bar{k}(x, y, z, t) = \mathbb{I}_{\{x \le z\}} \min(x + U(y)t, z) + \mathbb{I}_{\{x > z\}} \max(x - B(y)t, z). \tag{3.10}$$

Remark 3.3.1. Suppose that $A = [-B, U]$ and $b^y(a) = a$ then function $\bar{k}(x, y, z, t) = k(x, z, t)$. This is an important case when the ramp constraints do not depend on the environment.

In order to prove our main results we need an auxiliary lemma.

Lemma 3.3.2. *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a strictly convex function and $x^* \in \mathbb{R}$ is a point where the minimum of function $f(\cdot)$ attains*

$$f(x^*) < f(x), \qquad \forall x \neq x^*.$$

Assume further, that P is a probability measure on $\mathbb{R}^+$ absolutely continuous with respect to Lebesque measure. Consider a class of functions $X : \mathbb{R}^+ \to \mathbb{R}$, where $X(0) = x_0$ and X satisfies

$$-B\delta \le (X(t + \delta) - X(t)) \le U\delta,$$

where positive constants U and B are ramp constraints.

Then the optimization problem in the class of functions described above

$$L(X(t), x_0) := \int_0^\infty f(X(t))dP(t) \to \min_{X(t)}$$

has a unique solution given by

$$X^*(t) = k(x_0, x^*, t). \tag{3.11}$$

Proof. For the entire proof suppose that $x_0 < x^*$. The opposite case may be treated similarly. Firstly, we show that if $X(t)$ is the optimal policy then for all $t \geq 0$ holds $X(t) \leq x^*$. Consider a new policy $X_1(t) = \min(X(t), x^*)$ then the set $\{t : f(X_1(t)) < f(X(t))\}$ has positive measure due to the presence of constraints U and B. Therefore

$$L(X_1(t), x_0) < L(X(t), x_0).$$

Hence, the policy crossing level x^* cannot be optimal.

The second step is to observe that if we have two controls $X_1(t) < X_2(t) \leq x^*$ for all $t \geq 0$ then clearly

$$L(X_2(t), x_0) < L(X_1(t), x_0).$$

This shows that at each point t the optimal control $X(t)$ should be as close to the optimal point x^* as possible. The fastest movement towards x^* is linear with speed U. The dynamics are formalised by function $k(x_0, x^*, t)$. □

We start with a solution for the discounted optimality problem.

Theorem 3.3.3. *Suppose that $S = \mathbb{R}^+$, $c(x, y, a)$ does not depend on a and is strictly convex in x, Assumptions 3.2.1 and 3.3.1 hold for all $\beta > 0$. Assume that the initial value is known $(X(0), Y(0)) = (x_0, y_0)$.*

Then there exists a unique function $l^\beta(y) : \mathcal{Y} \to \mathbb{R}^+$ such that the trajectory of the corresponding optimal process $X(t)$ can be found recursively in the following manner.

- $X(0) = x_0$;
- *for $t \in [0, T_1]$ let $X(t) = \bar{k}(x_0, y_0, l^\beta(y_0), t)$;*
- *for $t \in [T_n, T_{n+1}]$ let $X(t) = \bar{k}(X(T_n), Y(T_n), l^\beta(Y(T_n)), t - T_n)$.*

Remark 3.3.4. If assumption (3.3.1) fails then none of the controls given by formula (3.10) are feasible.

Remark 3.3.5. The sample path of the process $X(t)$ may be written in a close formula

$$X(t) = \bar{k}(X_{\nu_t}, Y_{\nu_t}, l(Y_{\nu_t}), t - T_{\nu_t}),$$

where as before $\nu_t = \max\{n : T_n < t\}$.

Remark 3.3.6. There exist a set of levels $\{l^\beta(y)\}$ labelled by the elements of the environment state $\mathcal{Y}$. Whenever the environment is in the state y the trajectory starts moving towards $l(y)$ at the maximal possible speed and stops if it reaches this level.

Remark 3.3.7. Another possibility to formulate the theorem is to explicitly give the control function $a_0(x, y, t)$, which is done in the proof of the theorem.

Proof. Due to Theorem (3.2.6) there exist $\rho \geq 0$, convex function $h^\beta : E \to \mathbb{R}$ and u_0, which satisfy the discounted optimality equation (3.5). Starting with the Bellman equation (3.5) we may write:

$$\mathcal{C}^\beta(x, y) = \min_{a \in D(x,y)} \left(C(x, y, a) + \mathbb{E}^a_{x,y} \exp(-(\beta + \lambda)T_1)\mathcal{C}^\beta((X(T_1)), Y(T_1), a) \right) \tag{3.12}$$

$$= \min_{a \in D(x,y)} \left(\int_0^\infty c(X(t), y) \exp(-(\beta + \lambda)t)dt + \right. \tag{3.13}$$

$$\left. + \mathbb{E}^a_{x,y} \exp(-(\beta + \lambda)T_1)\mathcal{C}^\beta(X(T_1), Y(T_1)) \right) \tag{3.14}$$

$$= \min_{a \in D(x,y)} \int_0^\infty \left(c(X(t), y) + \right. \tag{3.15}$$

$$\left. + \sum_{y' \in \mathcal{Y}, y' \neq y} \frac{-q_{yy'}}{q_{yy}} (\beta + \lambda)\mathcal{C}^\beta(X(t), y')) \right) \exp(-(\beta + \lambda)t)dt. \tag{3.16}$$

The function

$$c(x, y) + \sum_{y' \neq y} \frac{-q_{yy'}}{q_{yy}} (\beta + \lambda)\mathcal{C}^\beta(x, y') =: H(x, y) \tag{3.17}$$

is strictly convex in x due to convexity of $\mathcal{C}^\beta(x, y)$ for all y and strict convexity of $c(x, y)$. Therefore there exist points $l^\beta(y)$, where the minimum of function (3.17) attains.

Applying Lemma (3.3.2) to $H(x, y)$ and $dP(t) = \exp(-(\beta + \lambda)t)dt$ one gets that

$$X^*(t) = \bar{k}(x, y, l^\beta, t)(y)$$

provides the minimum in the right-hand side of equation 3.12.

Therefore, in terms of open-loop controls, we can write

$$a^{\beta}(x,y,t) = \begin{cases} \begin{cases} U(y) & \text{if } t < \frac{l^{\beta}(y)-x}{U(y)} \\ 0 & \text{if } t \geq \frac{l^{\beta}(y)-x}{U(y)} \end{cases} & \text{if } x < l^{\beta}(y) \\ \begin{cases} -B(y) & \text{if } t < \frac{x-l^{\beta}(y)}{B(y)} \\ 0 & \text{if } t \geq \frac{x-l^{\beta}(y)}{B(y)} \end{cases} & \text{if } x \geq l^{\beta}(y) \end{cases} \tag{3.18}$$

□

A similar result holds for the time-average optimality.

Theorem 3.3.8. *Suppose that $S = \mathbb{R}^+$, and $c(x,y,u)$ does not depend on u. Assume that assumptions (3.2.2) and (3.3.1) hold. Suppose that the process starts from the initial point (x_0, y_0).*

Then there exists a function $l^0(y) : \mathcal{Y} \to \mathbb{R}^+$ such that the trajectory of the corresponding optimal process $X(t)$ can be constructed recursively in the following manner.

- $X(0) = x_0$;
- *for $t \in [0, T_1]$ let $X(t) = \bar{k}(x_0, y_0, l^0(y_0), t)$;*
- *for $t \in [T_n, T_{n+1}]$ let $X(t) = \bar{k}(X(T_n), Y(T_n), l^0(Y(T_n)), t - T_n)$.*

Proof. The proof is similar to the proof of Theorem (3.3.3) with only a slight modification related to a different form of the optimality equation. Under the assumptions (3.2.2) there exist ρ and convex function $h(x,y)$ satisfying the equation (3.6), by the Theorem (3.2.7). Hence, starting with the ACOE one can write

$$\lambda\rho + h(x,y) = \inf_u \left[\int_0^{\infty} c(X(t),y)\exp(-\lambda t)dt + \right.$$

$$\left. \int_E h(x',y')P(x,y,u,dx',dy') \right] = \tag{3.19}$$

$$\inf_u \int_0^{\infty} \left(c(X(t),y) + \sum_{y' \in \mathcal{Y}, y' \neq y} \frac{q_{y_0 y'}}{q_{y_0 y_0}} \lambda h(X(t),y') \right) \exp(-\lambda t)dt.$$

Let

$$H(x,y) := \frac{c(x,y)}{\lambda} + \sum_{y' \neq y} q_{yy'} h(x,y').$$

Function $h(x,y)$ is convex in x due to Theorem 3.2.7. Therefore, function $H(x,y)$ is strictly convex in x due to convexity of all summands and strict convexity of $c(x,y)$. Hence, there exist points $l^0(y)$, where the minimum attains. We finish the proof by

using Lemma (3.3.2) with $H(x,y)$ and $dP(t) = \exp(-(\lambda)t)dt$. As before, the optimal control is given by

$$a^0(x,y,t) = \begin{cases} \begin{cases} U(y) & \text{if } t < \frac{l^0(y)-x}{U(y)} \\ 0 & \text{if } t \geq \frac{l^0(y)-x}{U(y)} \end{cases} & \text{if } x < l^0(y) \\ \\ \begin{cases} -B(y) & \text{if } t < \frac{x-l^0(y)}{B(y)} \\ 0 & \text{if } t \geq \frac{x-l^0(y)}{B(y)} \end{cases} & \text{if } x \geq l^0(y) \end{cases} \tag{3.20}$$

Trajectories $X_0(t)$ corresponding to u_0 are exactly

$$X_0(t) = \bar{k}(X_{\nu_t}, Y_{\nu_t} l(Y_{\nu_t}), t - T_{\nu_t})$$

Hence, due to the uniqueness of the optimal control in Lemma (3.3.2) we have shown that $a_0(x,y,t)$ minimises the left-hand side of the equation (3.19). The last step is to notice that the obtained controls are given by a feedback rule and, hence, due to the Theorem (3.2.7) part (3) it is indeed the optimal control. □

The uniqueness of the optimal policy allows us to show that the optimal solution in the time-average case can be obtained not only as a limit along a certain subsequence, but as a continuous limit.

Corollary 3.3.9. *Under the assumptions of Theorem (3.3.8) for all (x,y), we have*

1.

$$a^0(x,y) = \lim_{\beta \to 0} a^\beta(x,y), \tag{3.21}$$

2.

$$l^0(y) = \lim_{\beta \to 0} l^\beta(y). \tag{3.22}$$

Proof. 1. We prove that the sequential limit of a^{β_n} in Theorem (3.2.7) may be replaced with the continuous limit when β goes to 0. Due to the Theorem (3.2.7) we know that there exists a sequence β_n such that

$$a^{\beta_n}(x,y) \to a^0(x,y). \tag{3.23}$$

The convergence is understood in the Young topology (see Section 3.8). In the proof of Theorem (3.2.7) (see Theorem 4, Bäuerle [6]), the limiting function a_0 is obtained in the following manner. For an arbitrary sequence β_n, there exists a further subsequence β_{n_m} such that there is convergence

$$h^{\beta_{n_m}}(x) \to h^0(x) \tag{3.24}$$

uniformly on compacts (due to equi-continuity and uniform boundedness by Arcela-Askoli Theorem). Its limit $\hat{a}^0$ minimises the right-hand side of the equation (3.6). However, in Theorem (3.3.8) we have shown that the solution is unique. Therefore, each sequence β_n has a subsequence β_{n_m} such that

$$a^{\beta_{n_m}}(x,y) \to a^0(x,y). \tag{3.25}$$

We know that the sequence is convergent if, and only if, any subsequence has a further convergent subsequence and the limits are all the same. From this we deduce that

$$a^{\beta}(x,y) \to a^0(x,y). \tag{3.26}$$

2. Theorems (3.3.3) and (3.3.8) state that the policies are determined by the sets of levels $l^{\beta}(y)$ for $\beta \geq 0$. We fix (x,y) for now and drop it in the notation when it is convenient. Convergence $a^{\beta} \to a^0$ in the Young Topology is equivalent to convergence

$$\int_0^{\infty} g(t, a^{\beta}(t))dt \to \int_0^{\infty} g(t, a^0(t))dt, \tag{3.27}$$

for all $g(t,u) \in L_1(\mathbb{R}^+, C(U))$, where the latter is the space of functions measurable in t and continuous in u. Take $g(t,u) = e^{-t}u$, which belongs to $L_1(\mathbb{R}^+, C(U))$. We know that f^{β} can take only three different values U, $-B$ and 0 due to the formulas (3.18) and (3.20). Without loss of generality, assume that the initial value $x \leq l^0(y)$ then

$$\int_0^{\infty} e^{-t}a^0(t)dt = \int_0^{\frac{l(y)-x}{U}} e^{-t}Udt = U\left(1 - \exp\left(-\frac{l(y)-x}{U}\right)\right) \geq 0. \tag{3.28}$$

We first prove by contradiction that there exists $\hat{\beta}$ such that for all $\beta \leq \hat{\beta}$ the inequality $x \leq l^{\beta}(y)$ holds. Suppose there exist a sequence of β_k such that $x > l^{\beta_k}(y)$ for all $k \geq 0$. Then we can write

$$\int_0^{\infty} e^{-t}a^{\beta_k}(t)dt = \int_0^{\frac{x-l^{\beta_k}(y)}{B}} -e^{-t}Bdt = -B\left(1 - \exp\left(-\frac{x-l^{\beta_k}(y)}{B}\right)\right) \leq 0, \tag{3.29}$$

which should converge to the left-hand side of equation (3.30). As the expressions have different signs it must mean that both of them are equal to 0. Hence $l^0(y) = x$ and $l^{\beta_k}(y) \to x$, so $l^{\beta_k}(y) \to l^0(y)$ only if $l^0(y) = x$. Now if we take a different initial value $x' \neq l^0(y)$ we would obtain a contradiction.

Now take only $\beta < \hat{\beta}$ so that $x \le l^\beta(y)$, then

$$\int_0^\infty e^{-t} a^\beta(t) dt = \int_0^{\frac{l^\beta(y)-x}{U}} e^{-t} U dt = U\left(1 - \exp\left(-\frac{l^\beta(y)-x}{U}\right)\right) \tag{3.30}$$

converges to to the left-hand side of equation (3.30) if, and only, if $l^\beta(y) \to l^0(y)$. Therefore, we have proven the required convergence.

□

Theorem 3.3.10. *Suppose that $\mathcal{Y}$ is finite, Markov chain $Y(T_n)$ is irreducible, $R = \max_y \operatorname{argmin}_x c(x,y) < \infty$ and optimal levels $\{l^\beta(y)\}_{y\in\mathcal{Y}}$ exist for all $\beta > 0$.*

Then there exist integrable function $M(x,y)$ and constant L such that Assumption 3.2.2.(e) holds.

Proof. Suppose that $X(t)$ is an arbitrary controlled process, then a process defined by $\widehat{X}(t) = \min(X(t), R)$ gives a better value locally for β-discounted cost functional for all β. Hence, we obtain that the optimal policy should not lie above level R. This yields that $l^\beta(y) \le R$ for all $\beta > 0$ and $y \in \mathcal{Y}$. Note that if $X(0) = x_1 \le R$ then for any $t \ge 0$ holds $X(t) \le R$, so the process never leaves compact $[0, R]$ once reached.

Now fix $0 < x_0 < R$ and $y_0 \in \mathcal{Y}$ and let $h^\beta(x,y) = \mathcal{C}^\beta(x,y) - \mathcal{C}^\beta(x_0,y_0)$.

The rest of the proof is split into two parts for $x_1 \le R$ and $x_1 > R$. We write $(X_i(t), Y_i(t))$ for the process that starts at (x_i, y_i) for $i = 0, 1$.

Suppose that $x_1 \le x_0$. As before $\{T_k\}$ denote jumping times of the embedded Markov chain of the process $Y(t)$. Random time σ_1 is defined as

$$\sigma_1 = \min\{i : Y_0(T_i) = Y_1(T_i)\} \tag{3.31}$$

It is well known that time σ_1 is a.s. finite and, moreover, $\mathbb{E}\sigma_1 < \infty$, see for example [40]. Denote by σ_2 random time such as

$$\sigma_2 = \min\{i \ge 0 : X_0(T_{i+\sigma_1}) = X_1(T_{i+\sigma_1})\}. \tag{3.32}$$

We now show that $\mathbb{E}\sigma_2 < \infty$. Let $w = \min_y(\min(U(y), B(y))$. If for k holds $T_k - T_{k-1} \ge \frac{R}{w}$ then one has $X_1(T_k) = X_0(T_k)$, because there is enough time to reach level $l^\beta(Y(T_{k-1}))$ from any point $x \in [0, R]$.

Since $T_k - T_{k-1}$ is an exponential random variable with parameter λ we have

$$\mathbb{P}\left\{T_k - T_{k-1} \ge \frac{R}{w}\right\} = \exp\left(\frac{-\lambda R}{w}\right) =: p^*. \tag{3.33}$$

There exists a random variable $\Gamma \sim \text{Geom}(p^*)$ such that $\sigma_2 \le \sum_{j=\sigma_1+1}^{\Gamma} T_j$. Random variable Γ is independent of future values of $\{T_j\}_{j\ge\Gamma}$. Therefore, we may apply the

Wald's identity

$$\mathbb{E}\sigma_2 \leq \frac{1}{\lambda}\mathbb{E}\Gamma < \infty. \tag{3.34}$$

Hence,

$$|h^\beta(x_1, y_1)| = |\mathcal{C}^\beta(x_1, y_1) - \mathcal{C}^\beta(x_0, y_0)| \leq 2\mathbb{E}(\sigma_1 + \sigma_2) \max_{x\in[0,R],y} c(x, y) =: L. \tag{3.35}$$

Let $M(x, y) = L$ for $x_1 \leq x_0$.

Now suppose that $x_1 > R$. Then

$$h^\beta(x_1, y_1) = \big(\mathcal{C}^\beta(x_1, y_1) - \mathcal{C}^\beta(x_0, y_1)\big) + \big(\mathcal{C}^\beta(x_0, y_1) - \mathcal{C}^\beta(x_0, y_0)\big) \geq -L. \tag{3.36}$$

As was explained at the beginning of the proof, the first bracket on the right hand side is positive. The second bracket is greater than $-L$ as was shown before.

To establish the upper bound we note that for $t \leq \frac{x_1-R}{\min_y B(y)}$, the process $X_1(t)$ decreases to R. Let $\sigma_3 = \max\{\sigma_1, [\frac{x_1-R}{\min_y B(y)}]\}$, where $[x] = \min_{\mathbb{Z}}\{z \geq x\}$ is the ceiling function. Then for $j \geq \sigma_3$ hold $Y_1(T_j) = Y_0(T_j)$ and $X_1(T_j) \leq R$.

Let $\sigma_4 = \min\{j : X_1(T_{j+\sigma_3}) = X_0(T_{j+\sigma_3})\}$. Similarly to the previous case, we get $\mathbb{E}\sigma_4 < \infty$. Hence we have

$$\begin{aligned} |h^\beta(x_1, y_1)| =& \leq \max\left(\mathbb{E}\sigma_1, \frac{x_1 - R}{\min_y B(y)}\right) \max_y c(x_1, y) + \\ &+ \mathbb{E}\sigma_4 \max_{x\in[0,x_0],y} c(x, y) =: M(x_1, y_1) < \infty. \end{aligned} \tag{3.37}$$

Therefore we have constructed L and $M(x, y)$ as needed. □

To summarise the minimal assumptions required for the existence of the levels $l^0(y)$ we state a straight forward corollary of the results above.

Corollary 3.3.11. *Suppose that $S = \mathbb{R}^+$, $\mathcal{Y}$ is finite, Markov chain $Y(T_n)$ is irreducible and the strictly convex function $c(x, y, u)$ does not depend on u. Suppose that Assumptions (3.2.2.(d)) and (3.3.1) hold. Then there exists the set of optimal levels $l^0(y)$.*

3.3.1 The case of no environment

Suppose that there is no environment present, e.g. $\|\mathcal{Y}\| = 1$. Then there is no new information about the environment at any point in time and the optimal controlled process stays at the same level. This informal reasoning is formalised in the following corollary. This result with the informal reasoning was first suggested to us by Stan Zachary.

Corollary 3.3.12. *Suppose that $|\mathcal{Y}| = 1$ so $Y(T_n)$ are all identical (deterministic). Let $X(0) = x_0$ Then there exist a number l and the process corresponding to the optimal policy is*

$$X(t) = k(x_0, l, t).$$

For any initial value there exists $T > 0$ such that $X(t) = l$ for all $t \geq T$.

3.4 On the continuity of the optimal solution when $U, B \to \infty$

In this subsection we assume that the optimal levels $l^0(y)$ exist for all considered problems and that the ramp constraints do not depend on states of the process $Y(t)$. Now we would like to compare models with different values of ramp constraints U and B, whilst keeping the matrix Q and the cost functional $c(x, y)$ fixed. Then the optimal levels $l^0(y)$ might be regarded as functions of U and B. We denote them by $l^0_{U,B}(y)$. Let $m(y) = \operatorname{argmin}_x c(x, y)$, and renumber elements of $\mathcal{Y}$ so that $m(i) \leq m(j)$ if $i \leq j$. Assume that $|\mathcal{Y}| = n$. Let $\mathbf{m} = (m(1), m(2), \ldots, m(n))$.

We use the following notation in this section.

- For the set of levels $\ell = \{l(y)_{y\in\mathcal{Y}}\}$ we write $X^{\ell}_{U,B}(t)$ for the corresponding fluid process with levels ℓ and ramp constraints U and B.
- The optimal policy for ramp constraints U and B is then $X^{\ell^0_{U,B}}_{U,B}(t)$, where

  $$\ell^0_{U,B} = (l^0_{U,B}(1), l^0_{U,B}(2), \ldots, l^0_{U,B}(n))$$

  is the set of the optimal levels.
- Suppose that $U = B = \infty$, then for $t \in [T_n, T_{n+1}]$ and levels ℓ the trajectory $X^{\ell}_{\infty,\infty}(t)$ is defined by

  $$X^{\ell}_{\infty,\infty}(t) = l(Y_{T_n}).$$

  Note that the trajectory is not continuous.
- Let $\widetilde{X} = X^{\mathbf{m}}_{\infty,\infty}$.
- Let $\Delta(U, B) = (m(n) - m(1)) \max(1/U, 1/B)$.
  Note that for any pair $x_1, x_2 \in [m(1), m(n)]$ and ramp constraints U and B the time to reach level x_2 starting at level x_1 is not longer than $\Delta(U, B)$.

Consider the limiting case where $U = B = \infty$. In this situation there are no ramp constraints, so the trajectories are no longer necessarily continuous, and any positive

function on $\mathbb{R}^+$ may be considered as a possible trajectory. Then the trajectory $\widetilde{X}$ is average-optimal in the class of positive functions on $\mathbb{R}^+$ as it minimises the cost functional at every point:

$$\forall x, y \quad c(x,y) \geq c(m(y),y) \quad \Longrightarrow \tag{3.38}$$

$$\forall X(t), y \quad \mathbb{E}\int_0^\tau c(X(s),y)ds \geq \mathbb{E}\int_0^\tau c(\widetilde{X}(s),y)ds. \tag{3.39}$$

Intuition suggests that the larger the ramp constraints, the faster it is possible to move to the optimal levels $l^0(y)$. Thus the optimal levels $l^0_{U,B}(y)$ should approach $m(y)$ as $U, B \to \infty$. This is the content of the following theorem.

Theorem 3.4.1. *Suppose that $U, B \to \infty$ then $l^0_{U,B}(y) \to m(y)$, for all y.*
Moreover, there exists $\widetilde{C} > 0$ such that, for all y

$$|l^0_{U,B}(y) - m(y)| \leq \widetilde{C}\left(\max\left(\frac{1}{U}, \frac{1}{B}\right)\right). \tag{3.40}$$

To prove the theorem we need a simple auxiliary lemma.

Lemma 3.4.2. *1. There exists a constant $\widehat{C}$ such that*

$$\mathcal{C}(X^{\mathbf{m}}_{U,B}) - \mathcal{C}(\widetilde{X}) \leq \widehat{C} \times \Delta(U,B). \tag{3.41}$$

In particular,

$$\mathcal{C}(X^{\mathbf{m}}_{U,B}) \to \mathcal{C}(\widetilde{X}) \quad \textit{as} \quad U, B \to \infty. \tag{3.42}$$

2. Suppose that for the set of levels $\{l(y)\}_{y\in\mathcal{Y}}$ there exist y_0 such that

$$|l(y_0) - m(y_0)| \geq \delta.$$

Then

$$\mathcal{C}(X^{\ell}_{U,B}) - \mathcal{C}(\widetilde{X}) \geq \mathbb{P}\{Y(t) = y_0\} \times \delta \times \lambda \times \exp(-\Delta(U,B)\lambda). \tag{3.43}$$

Proof. 1. Let $M = \max_y \max_{x_1,x_2\in[m(1),m(n)]} |c(x_1,y) - c(x_2,y)|$. Within each time interval $[T_n, T_{n+1}]$ the maximal possible difference between the charges for both policies is

$$\mathbb{E}\int_{T_n}^{T_{n+1}} \left(c(X^{\mathbf{m}}_{U,B}(s),y) - c(\widetilde{X}(s),y)\right) ds \leq \Delta(U,B)M. \tag{3.44}$$

Therefore, for the average cost the following inequality holds

$$0 \leq \mathcal{C}(X^{\mathbf{m}}_{U,B}) - \mathcal{C}(\widetilde{X}) \leq \frac{\Delta(U,B)M}{\lambda}, \tag{3.45}$$

since $\mathbb{E}\tau_n = 1/\lambda$. Now if U and B tend to infinity we have $\frac{\Delta(U,B)M}{\lambda} \to 0$.

2. For the difference between the average cost functionals we have

$$\mathcal{C}(X^{\ell}_{U,B}) - \mathcal{C}(\widetilde{X}) = \limsup_{k\to\infty} \frac{\mathbb{E}\int_0^{T_k} \Big(c(X^{\ell}_{U,B}(s), Y(s)) - c(\widetilde{X}(s), Y(s))\Big)\, ds}{\frac{k}{\lambda}}$$

$$= \limsup_{k\to\infty} \frac{\sum_1^k \mathbb{P}\{Y(t) = y_0\} \int_{\Delta(U,B)}^{\infty} \delta \times (s - \Delta(U,B)) \times \lambda \exp(-\lambda s) ds}{\frac{k}{\lambda}}$$

$$\geq \mathbb{P}\{Y(t) = y_0\}\delta \times \lambda \exp(-\Delta(U,B)\lambda).$$

□

Now we are ready to prove Theorem (3.4.1).

Proof of Theorem (3.4.1). We prove the theorem by contradiction.

Suppose that there exists y such that the sequence of levels for average optimal policies $l^0(y)(U,B)$ does not converge to $m(y)$. Then there exist $\delta > 0$ and a subsequence (U_n, B_n) that

$$|l^0_{U_n,B_n}(y) - m(y)| \geq \delta. \tag{3.46}$$

Due to the first part of Lemma (3.4.2), we can take U_0, B_0 such that for $U > U_0, B > B_0$ the following holds

$$\mathcal{C}(X^{\mathbf{m}}_{U,B}) - \mathcal{C}(\widetilde{X}) \leq \frac{\delta\lambda\mathbb{P}\{Y(t) = y\}}{4} \text{ and} \tag{3.47}$$

$$\exp(-\Delta(U,B)\lambda) \geq \frac{1}{2}. \tag{3.48}$$

Due to inequality (3.43) one has

$$\mathcal{C}(X^{\ell_0}_{U,B}) - \mathcal{C}(\widetilde{X}) \geq \delta\mathbb{P}\{Y(t) = y\}\lambda \exp(-\Delta(U,B)\lambda) \geq \frac{\delta\lambda\mathbb{P}\{Y(t) = y\}}{2}. \tag{3.49}$$

Now by the definition of U_0, B_0 and by the triangle inequality one has

$$\mathcal{C}(X^{\ell_0}_{U,B}) - \mathcal{C}(X_{U,B}) \geq \frac{\delta\lambda\mathbb{P}\{Y(t) = y\}}{4}.$$

Therefore we get a contradiction with the optimality of strategy $X^{\ell_0}_{U_B}$. Hence the assumption that $l^0_{U,B}(y)$ does not converge to $m(y)$ is false.

Finally, the rate of convergence may be obtained as follows. For $U > U_0$ and $B > B_0$, we have

$$|\mathcal{C}(X^{\ell_0}_{U,B}) - \mathcal{C}(\tilde{X})| \le |\mathcal{C}(X^{\mathbf{m}}_{U,B}) - \mathcal{C}(\tilde{X})| \le \widehat{C}\Delta(U,B),$$

$$|\mathcal{C}(X^{\ell_0}_{U,B}) - \mathcal{C}(\tilde{X})| \ge \max_y |l^0_{U,B}(y) - m(y)| \min_y (\mathbb{P}\{Y(t) = y\}) \times \lambda \exp(-\Delta(U,B)\lambda)$$

$$\ge \max_y |l^0_{U,B}(y) - m(y)|\widehat{C_1},$$

where $\widehat{C_1} > 0$. Hence, $\max_y |l^0_{U,B}(y)(U,B) - m(y)| \le \widetilde{C}\Delta(U,B)$, where $\widetilde{C} = \widehat{C} \times \widehat{C_1}^{-1}$. □

The convergence of the optimal levels $l^0_{U,B}(y)$ to $m(y)$ has a useful application. For large U and B, instead of searching for the optimal levels $l^0_{U,B}(y)$, which is computationally difficult, one may take $m(y)$ as the set of levels for the trajectory, which is nearly optimal. Calculating $m(y)$ is simple, it is nothing more than finding the minimums of functions $c(x,y)$.

3.4.1 Computational example

We support the convergence result with an example obtained by running the Matlab code (see Appendix B). This code is based on the numeric algorithm described in Section 3.6 and the solution for each pair of ramp constraints is found on the grid with the parameter $h = 0.1$.

Suppose that $n = 3$,

$$Q = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

and

$$c(x,1) = x + 2\max\{1-x, 0\}, \quad m(1) = 1,$$

$$c(x,2) = x + 2\max\{2-x, 0\}, \quad m(2) = 2,$$

$$c(x,3) = x + 2\max\{5-x, 0\}, \quad m(3) = 5,$$

The code returns the numeric approximation of the solution oh the grid with parameter $h = 0.1$. We obtained the following results:

U=B=0.1	$l^0_{U,B}(1) = 1.9$	$l^0_{U,B}(2) = 2$	$l^0_{U,B}(3) = 2.1$
U=B=0.3	$l^0_{U,B}(1) = 1.8$	$l^0_{U,B}(2) = 2$	$l^0_{U,B}(3) = 2.2$
U=B=0.5	$l^0_{U,B}(1) = 1.7$	$l^0_{U,B}(2) = 2$	$l^0_{U,B}(3) = 2.3$
U=B=0.7	$l^0_{U,B}(1) = 1.5$	$l^0_{U,B}(2) = 2$	$l^0_{U,B}(3) = 2.5$
U=B=0.9	$l^0_{U,B}(1) = 1.4$	$l^0_{U,B}(2) = 2$	$l^0_{U,B}(3) = 2.6$
U=B=1	$l^0_{U,B}(1) = 1.3$	$l^0_{U,B}(2) = 2$	$l^0_{U,B}(3) = 2.7$
U=B=3	$l^0_{U,B}(1) = 1$	$l^0_{U,B}(2) = 2$	$l^0_{U,B}(3) = 4$
U=B=5	$l^0_{U,B}(1) = 1$	$l^0_{U,B}(2) = 2$	$l^0_{U,B}(3) = 5$
U=B=7	$l^0_{U,B}(1) = 1$	$l^0_{U,B}(2) = 2$	$l^0_{U,B}(3) = 5$
U=B=9	$l^0_{U,B}(1) = 1$	$l^0_{U,B}(2) = 2$	$l^0_{U,B}(3) = 5$

Table 3.1: Convergence of the levels $l^0_{U,B}(y)$

The table shows a clear convergence and this convergence is monotone.

3.5 Markov-modulated fluid model with multiple marks on the tank and differential equations for its joint stationary distribution

We have shown in Theorems 3.3.3 and 3.3.8 that the optimal controlled processes $X(t)$ are of a certain form: if $Y(T_k) = y$ then it goes to level $l^\beta(y)$ at the maximal pace. In this subsection we derive a differential equation along with the boundary conditions, the solution of which descibes the time-stationary distribution of these processes. We then discuss how to solve this system.

Different fluid models (also knows as fluid queues or dam models) have been widely studied since the 1960s. For the classical model one may look in [1], [42] or [39], and for the most advanced models in [8], [21] or [22].

We describe a new fluid model below. We are given a water tank of capacity l_n with marks for levels $l_1 < l_2 < \ldots < l_n$. The continuous time Markov process $Y(t)$ takes values in $\{1, \ldots, n\}$. Its Q-matrix is given by $(q_{ij})_{1 \le i,j \le n}$. Suppose that $U(i), B(i) > 0$ are maximal rates to fill and empty the water tank, when $Y(t)$ is in state i.

Process $X(t)$ stands for the water level in the tank. It evolves in accordance with $Y(t)$ and $\{l_i\}_{i=1}^n$. Given that $Y(t) = j$ the process is defined by

$$\frac{dX(t)}{dt} = \begin{cases} 0, \text{ if } X(t) = l_j, \\ U(j), \text{ if } X(t) < l_j, \\ -B(j) \text{ if } X(t) > l_j. \end{cases} \tag{3.50}$$

This is a situation in which the level of water in the tank tries to reach level l_j as

quickly as possible, whilst $Y(t) = j$. We assume that the dynamics of the process $Y(s)$ in $[t, t+h]$ are independent of $\{X(t_1)\}_{t_1 \le t}$ given that $Y(t)$ is known. In particular, the number of jumps of $Y(s)$ in $[t, t+h]$ are also independent of past values of process $X(t)$.

$$\mathbb{P}(\ \#\text{ of jumps of } Y(s) \in [t, t+h] = k | X(t) = x_0, Y(t)) \tag{3.51}$$

$$= \mathbb{P}(\#\text{ of jumps of } Y(s) \in [t, t+h] = k | Y(t)). \tag{3.52}$$

We refer to the stochastic process $(X(t), Y(t))$ as $FL(l_1, \ldots, l_n)$.

Remark 3.5.1.
Control processes $X(t)$ from Theorems (3.3.3) and (3.3.8) are $FL(l^\beta(1), \ldots, l^\beta(n))$.

Suppose that $(\tilde{X}, \tilde{Y})$ denotes the unique stationary regime of process $(X(t), Y(t))$. Existence of the limiting distribution for Markov process $(X(t), Y(t))$ (which coincides with the stationary distribution in embedded moments) is discussed for a class of fluid policies for semi-Markov random environment in Chapter 4. We use notation

$$F_j(x,t) = \mathbb{P}(X(t) < x, Y(t) = j), \quad G_j(x,t) = \mathbb{P}(X(t) \ge x, Y(t) = j) \tag{3.53}$$

$$F_j(x) = \lim_{t \to +\infty} F_j(x,t), \quad G_j(x) = \lim_{t \to +\infty} G_j(x,t). \tag{3.54}$$

Theorem 3.5.2. *The joint probability distribution functions $F_j(x)$ (and $G_j(x)$ as well) satisfy the system of equations*

$$\frac{dF_j(x)}{dx} = \begin{cases} \frac{1}{U(j)} \sum_{1 \le i \le n} F_i(x) q_{ij}, & \text{if} \quad l_1 < x < l_j, \\ \frac{-1}{B(j)} \sum_{1 \le i \le n} F_i(x) q_{ij}, & \text{if} \quad l_n > x > l_j, \end{cases} \tag{3.55}$$

along with the set of boundary conditions

$$\begin{cases} F_j(l_1) = 0, & \text{if} \quad 1 < j \le n, \\ F_j(l_n) = \pi_j, & \text{if} \quad 1 \le j < n, \end{cases} \tag{3.56}$$

where $\pi_j = \mathbb{P}(\tilde{Y} = j)$ is the stationary distribution for process $Y(t)$.

Functions $F_j(x)$ are continuous everywhere but in l_j, so distribution of stationary version $\tilde{X}$ has positive atoms in $l_1, \ldots, l_n$.

Remark 3.5.3. There are $2n-2$ boundary conditions and $2n-2$ differential equations. If we had fewer boundary conditions, the system could have had multiple solutions. If we had more boundary conditions, the system could have been unsolvable.

Proof. We start by showing that the boundary conditions satisfy Equation 3.56. Suppose that $X(0) = x_0 > l_n$ then for $t > \frac{x_0 - l_n}{\min_i B(i)}$ one has $X(t) \in [l_1, l_n]$. This and the

similar reasoning for $x_0 < l_1$ yields that in stationary regime $\tilde{X} \in [l_1, l_n]$. Moreover, $\mathbb{P}(\tilde{X} \in \mathcal{B}) > 0$ for any open non-empty set $\mathcal{B} \subset [l_1, l_n]$. These two observations imply the result.

The rest of the proof is influenced by the proof of Theorem (1.2.1) in A.de Silva Soares [20], which was inspired by the proof of Theorem (1.3.7) Barbot [5]. Suppose that, T_k is the time of the last jump of the process $Y(t)$ and $Y(T_k) = j$. Since $Y(T_{k-1}) \neq j$, we obtain that the probability of event $X(T_k) = l_j$ is equal to 0. Thus for small $\delta > 0$ process $X(T_k + \delta)$ is moving towards level l_j at either pace $U(j)$ if $X(T_k) < l_j$ or at $-B(j)$ if $X(T_k) > l_j$.

With no loss of generality, we denote the rate by r. One has

$$\frac{\partial}{\partial t} G_j(x,t) = \lim_{h \to 0^+} \frac{G_j(x, t+h) - G_j(x,t)}{h}$$

$$= \lim_{h \to 0^+} \frac{\mathbb{P}(X(t+h) > x, Y(t+h) = j) - \mathbb{P}(X(t) > x, Y(t) = j)}{h}$$

$$= \lim_{h \to 0^+} \frac{\mathbb{P}(X(t+h) > x, Y(t+h) = j, 0 \text{ jumps in } [t, t+h])}{h}$$

$$+ \lim_{h \to 0^+} \frac{\mathbb{P}(X(t+h) > x, Y(t+h) = j, 1 \text{ jumps in } [t, t+h]) - \mathbb{P}(X(t) > x), Y(t) = j)}{h}$$

$$+ \lim_{h \to 0^+} \frac{o(h)}{h}$$

$$=: \lim_{h \to 0^+} \frac{P_0 + P_1 - \mathbb{P}(X(t) > x), Y(t) = j)}{h}. \tag{3.57}$$

We may consider only 0 or 1 jumps in $[t, t+h]$ due to the fact that

$$\mathbb{P}(\# \text{ of jumps in } [t, t+h] \geq 2) = o(h). \tag{3.58}$$

The first term P_0 is easy to tackle. If there was no jump in the period $[t, t+h]$ then

$$X(t+h) = X(t) + rh \text{ for small } h \text{ such that } X(t) + rh \neq l_j.$$

Denote by $A_0 = \{0 \text{ jumps in } [t, t+h]\}$ then one has

$$\mathbb{P}(X(t+h) > x), Y(t+h) = j, A_0) = \mathbb{P}(X(t) > x - rh, Y(t) = j, A_0) \tag{3.59}$$

$$= \mathbb{P}(A_0 | X(t) > x - rh, Y(t) = j) \times \mathbb{P}(X(t) > x - rh, Y(t) = j) \tag{3.60}$$

$$= \mathbb{P}(A_0 | Y(t) = j) \times G_j(x - rh, t) = (1 + q_{jj}h) G_j(x - rh, t), \tag{3.61}$$

where the first equality in the third line follows from formula (3.51).

To proceed with the expression for P_1 we note that

$$X(t) > x + r_1 h \Longrightarrow X(t+h) > x \Longrightarrow X(t) > x - r_2 h, \tag{3.62}$$

where $r_1 = \min_j(\min(U(j), B(j))$ and $r_2 = \max_j(\max(U(j), B(j))$ Therefore, for any event A one obtains

$$\mathbb{P}(X(t) > x + r_1 h, A) \leq \mathbb{P}(X(t+h) > x, A) \leq \mathbb{P}(X(t) > x - r_2 h, A). \tag{3.63}$$

We show that for event $A_1 = \{Y(t+h) = j, 1 \text{ jump in } (t, t+h]\}$ and for any $y \in [l_1, l_n]$ holds

$$\mathbb{P}(X(t) > y, A_1) = \sum_{i \neq j} q_{ij} h G_j(y, t) + o(h). \tag{3.64}$$

By the law of total probability we may write

$$\begin{aligned}
\mathbb{P}(X(t) > y, A_1) &= \sum_{i \neq j} \mathbb{P}(X(t) > y, Y(t) = i, A_1) \\
&= \sum_{i \neq j} \mathbb{P}(A_1 | Y(t) = i, X(t) > y) \times G_i(y, t) \\
&= \sum_{i \neq j} \mathbb{P}(A_1 | Y(t) = i) \times G_i(y, t) \\
&= \sum_{i \neq j} q_{ij} h G_i(y, t) + o(h).
\end{aligned} \tag{3.65}$$

Hence, by continuity of $G_j(y, t)$ everywhere but at point l_j and the formula (3.63) we get

$$\lim_{h \to 0} \frac{1}{h} \mathbb{P}(X(t+h) > x, Y(t+h) = j, 1 \text{ jump }) = \sum_{i \neq j} q_{ij} G_i(x, t). \tag{3.66}$$

Thus, substituting formulas for P_0 and P_1 into formula 3.57 one has

$$\begin{aligned}
\frac{\partial}{\partial t} G_j(x, t) &= \lim_{h \to 0^+} \frac{1}{h} \left((1 + q_{jj} h) G_j(x - rh, t) - G_j(x, t) \right) + \sum_{i \neq j} q_{ij} G_i(x, t) \\
&= -r \frac{\partial}{\partial x} G_j(x, t) + q_{jj} G_j(x, t) + \sum_{i \neq j} q_{ij} G_i(x, t) \\
&= -r \frac{\partial}{\partial x} G_j(x, t) + \sum_{i} q_{ij} G_i(x, t).
\end{aligned} \tag{3.67}$$

Since there is a convergence in distribution to the time-stationary distribution we

have

$$\lim_{t\to\infty}\frac{\partial}{\partial t}G_j(x,t)=0.$$

Therefore,

$$r\frac{\partial}{\partial x}G_j(x)=\sum_i q_{ij}G_i(x) \tag{3.68}$$

The last step is to return from functions $G_j(x)$ to $F_j(x)=\pi_j-G_j(x)$. □

3.5.1 How to solve the system of equations given the boundary conditions

The system of equations 3.55 is discontinuous at points (l_j, j). Although in every segment $S_j=[l_j, l_{j+1})$ is a homogeneous linear differential equation. It would be possible to solve it if we knew correct boundary conditions at points l_j and l_{j+1}. The boundary conditions we have are concentrated in l_1 and l_n. Below, we present an algorithm allowing us to overcome this difficulty.

Let $S_j=[l_j,l_{j+1})$ for $j=1\dots n-1$ and $S_n=l_n$. Then S_j forms a partition of the state space S of process $X(t)$. In each segment S_j the system (3.55) is a homogeneous matrix differential equation. If we use notation $\mathbb{F}$ for vector $(F_1(x),\dots,F_n(x))$ one may write

$$\frac{d\mathbb{F}}{dx}=\mathbb{F}Q_j, x\in S_j, \tag{3.69}$$

where Q_j is given by its entries $q^j_{i,l}$

$$q^j_{i,l}=\begin{cases}\frac{q_{il}}{U(j)}, & \text{if} \quad i>j\\ -\frac{q_{il}}{B(j)}, & \text{if} \quad i\le j,\end{cases} \tag{3.70}$$

If the boundary conditions are known, the solution of the system (3.69) in each S_j can be obtained via spectral analysis. The solution of the system satisfying the boundary conditions can then be expressed as a linear combination of eigenvectors corresponding to the system. For the theory of classic differential equations one may look at [49] or [28].

Each of the probability distribution functions $F_j(x)$ are continuous everywhere except at a single point l_j. Let $p_j=\mathbb{P}(X(t)=l_j, Y(t)=j)$. Values p_j are unknown variables, however we use them to obtain the solution. The solution is described algorithmically step by step.

Algorithm

- **Step one**

  Solve system 3.69 in S_1 with boundary conditions $F_j(l_1)=0$ if $j>1$ and

$F_1(l_1) = p_1$.

- **Step** i
  Solve the system in S_i. All the functions except $F_i(x)$ are continuous at l_i. We use their previously defined values $F_j(l_i)(p_1, \dots, p_{i-1})$ and $F_i(l_i) = p_i$ as boundary conditions. Notation $F_j(l_i)(p_1, \dots, p_{i-1})$ underlines that constructed functions depend on unknown values $p_1, \dots p_{i-1}$.

- **Define values p_j to satisfy the boundary conditions at l_n**
  We obtained that functions $\mathbb{F}(p_1, \dots, p_{n-1})$ are now defined everywhere except at point b_n. We know that $F_j(l_n) = \pi_j$ for $j < n$. These are $n-1$ numeric equations for $n-1$ unknown values $p_1, \dots, p_{n-1}$. Value p_n can then be expressed as follows
  $$p_n = \pi_n - F_n(l_n-),$$
  where $F_n(l_n-) = \mathbb{P}(X(t) < l_n, Y(t) = n)$ is obtained as a solution of the system.

We have defined functions $\mathbb{F}(x)$ for all x and they satisfy the boundary conditions (3.56).

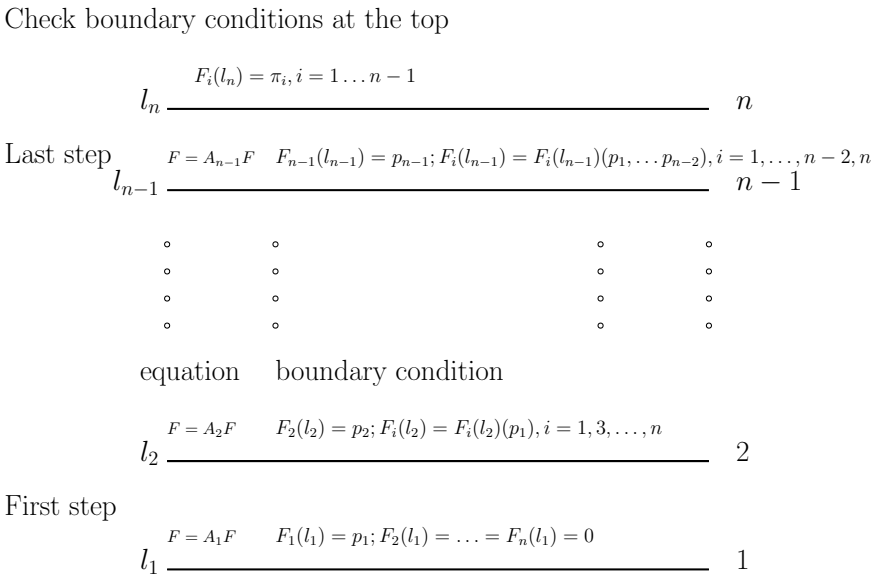


Figure 3.1: Step-by-step solution of the system

3.6 Optimization in the introduced class of fluid models

We have shown in Theorems (3.3.3) and (3.3.8) that the trajectories of the optimal control processes $X(t)$ correspond to a water level of a fluid model with levels $l^\beta(y)$. Therefore, if the model satisfies the assumptions of the theorems, one may restrict

the area of search from the set of all open-loop controls to the set of fluid models with levels $l_1, \dots, l_n$, which are treated here as variables. The average cost functional under control $FL(l_1, \dots, l_n)$ is a function of levels $l_1, \dots, l_n$ and can be expressed as

$$L(l_1, \dots, l_n) = \sum_{i=1}^{n} \left(\int_{l_1}^{l_n} c(x, i) dF_i(x)(l_1, \dots, l_n) + p_i c(l_i, i) \right). \tag{3.71}$$

Suppose that $L(l_1, \dots, l_n)$ is differentiable with respect to its parameters, then one can use the method of Lagrangian multipliers to find its minimal value.

For small values of n it is possible to obtain an explicit equation for the stationary distribution of $FL(l_1, \dots, l_n)$. It then is possible to get the exact values $(l_i)_{i=1}^n$ by solving the corresponding equations numerically. For example, we have a written a Matlab code (see Appendix B) to find the optimal solution, in the case when Markov chain $\mathcal{Y}$ has only three states. It calculates the stationary distribution as a function of parameters (l_1, l_2, l_3) by applying the algorithm from Subsection 3.5.1 and spectral analysis. Then it numerically finds a pseudo optimal solution on a grid. We have used the outcomes to support the convergence rate of the optimal solutions with respect to the ramp constraints in Section 3.4.

3.6.1 An exact solution of the system for $|\mathcal{Y}| = 2$

Consider the easiest scenario when $|\mathcal{Y}| = 2$, $U(1) = U(2) = U$, $B(1) = B(2) = B$. Assume that

$$Q = \begin{pmatrix} -\lambda_1 & \lambda_1 \\ \lambda_2 & -\lambda_2 \end{pmatrix}.$$

Then a fluid model $Fl(l_1, l_2)$ is an example of the simplest single buffer fluid model. It is well known that stationary distribution of the process $(X(t), Y(t))$ exists and the equations were given in multiple papers, for example in [1]. It is a particular subcase of Theorem 3.5.2. In this case is possible to solve the system explicitly and the next lemma gives a closed form solution.

Lemma 3.6.1. *The stationary distribution of model $Fl(l_1, l_2))$, ($l_1 < l_2$) can be given by the following formulas*

- *If $\lambda_2 U \neq \lambda_1 B$ the stationary distribution is*

$$F_1(x) = \frac{\lambda_2}{\lambda_1 + \lambda_2} \cdot \frac{U' z^x - U' z^{l_1}}{B' z^{l_2} - U' z^{l_1}}, \tag{3.72}$$

$$F_2(x) = \frac{\lambda_1}{\lambda_1 + \lambda_2} \cdot \frac{B' z^x - U' z^{l_1}}{B' z^{l_2} - U' z^{l_1}}, \tag{3.73}$$

$$\mathbb{P}(X(0) = l_1, Y(0) = 2) = F_2(l_1), \tag{3.74}$$

$$\mathbb{P}(X(0) = l_2, Y(0) = 1) = \frac{\lambda_2}{\lambda_1 + \lambda_2} - F_1(l_2), \tag{3.75}$$

$$\tag{3.76}$$

where $U' = \lambda_1/U$, $B' = \lambda_2/B$ *and* $z = \exp(B' - U')$.

- *If* $\lambda_2 U = \lambda_1 B$ *one has*

$$F_1(x) = \lambda_2 z(x - l_1) + zU \tag{3.77}$$

$$F_2(x) = \frac{\lambda_2 B z}{U}(x - l_1), \tag{3.78}$$

$$\mathbb{P}(X(0) = l_1, Y(0) = 1) = F_1(l_1) = zU, \tag{3.79}$$

$$\mathbb{P}(X(0) = l_2, Y(0) = 2) = \frac{\lambda_1}{\lambda_1 + \lambda_2} - F_2(l_2), \ \textit{where} \tag{3.80}$$

$$z = \frac{\lambda_2}{\lambda_1 + \lambda_2} \times \frac{1}{\lambda_2(l_2 - l_1) + U}. \tag{3.81}$$

Proof. To start with, the stationary distribution for Markov process $Y(t)$ is a solution of the equation $\pi Q = 0$, which results in

$$\pi = \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}, \frac{\lambda_1}{\lambda_1 + \lambda_2}\right)$$

The system 3.55 along with the boundary conditions 3.56 can be rewritten as

$$\begin{cases} F_1(x)' &= \frac{\lambda_1}{B} F_1(x) + \frac{-\lambda_2}{B} F_2(x), \\ F_2(x)' &= \frac{\lambda_1}{U} F_1(x) - \frac{-\lambda_2}{U} F_2(x), \\ F_2(l_1) &= 0, \\ F_1(l_2) &= \frac{\lambda_2}{\lambda_1 + \lambda_2}. \end{cases} \tag{3.82}$$

Multiplying the first equation by B, the second by U and then subtracting them results in

$$BF_1'(x) = UF_2'(x).$$

By integrating it one gets

$$F_2(x) = \frac{B}{U} F_1(x) + c_1, \tag{3.83}$$

where c_1 is a currently unknown constant. Now substitute the expression 3.83 in the

first differential equation of 3.82 to obtain

$$F_1(x)' = F_1(x)\left(\frac{\lambda_1}{B} - \frac{\lambda_2}{U}\right) - \frac{\lambda_2 c_1}{B}. \tag{3.84}$$

Now suppose that $\lambda_1 U = \lambda_2 B$, then the first term disappears and we get that

$$F_1(x) = -\frac{\lambda_2 c_1}{B}x + c_2,$$

and the constants c_1 and c_2 can be found from the equations on the boundary conditions.

Now suppose that $\lambda_2 U \neq \lambda_1 B$, then the equation (3.84) is a linear differential equation of the form $f'(x) = Af(x) + D$, which has solutions of the form $f(x) = c/A\exp(Ax) - D/A$, where c is a constant. Therefore, we obtain

$$F_1(x) = \frac{c_2}{\frac{\lambda_1}{B} - \frac{\lambda_2}{U}}\exp\left(\frac{\lambda_1}{B} - \frac{\lambda_2}{U}\right)x + \frac{\lambda_2 c_1}{B\left(\frac{\lambda_1}{B} - \frac{\lambda_2}{U}\right)}. \tag{3.85}$$

By solving the equations on the boundary conditions one obtains values for constants c_1 and c_2. □

3.6.2 Optimisation for a particular problem with $|\mathcal{Y}| = 2$

We suppose that $U = B = 1$. Assume that the infinitesimal operator of Markov process $Y(t)$ is given by

$$Q = \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix}.$$

Then the stationary distribution for Markov process $Y(t)$ is $\pi = (\pi_1, \pi_2) = (1/2, 1/2)$.

Therefore due to Lemma (3.6.1), the time-stationary distribution of the fluid model $Fl(a, b)$ driven by process $Y(t)$ may be expresses in the form

$$\begin{cases} F_1(x) = \frac{x-a+2}{2(b-a)+4}, \\ F_2(x) = \frac{x-a}{2(b-a)+4}, \\ \mathbb{P}(X(0) = a, Y(0) = 1) = \mathbb{P}(X(0) = b, Y(0) = 2) = \frac{1}{(b-a)+2}. \end{cases} \tag{3.86}$$

Using Proposition (3.1.1) we can write the average cost functional, defined by the triple of functions $(g(x), f(x, 1), f(x, 2))$, as

$$H(a, b) = \mathcal{C}(FL(a, b)) = \mathbb{E}g(x) + 2f(x, y) =$$

$$= \int_a^b g(x)d(F_1(x) + F_2(x)) + 2\int_a^b f(x, 1)dF_1(x) + 2\int_a^b f(x, 2)dF_2(x) +$$

$$+g(a)F_1(a) + g(b)(\frac{1}{2} - F_2(b))$$

$$+2f(a,1)F_1(a) + 2f(b,2)(\frac{1}{2} - F_2(b)) =$$

$$=\frac{\int_a^b g(x) + f(x,1) + f(x,2)dx + 2f(a,1) + 2f(b,2) + g(a) + g(b)}{b-a+2}. \quad (3.87)$$

Suppose that $a \leq b \leq c$, then we have an interesting property

$$H(a,c) = \frac{H(a,b)(b-a+2) + H(b,c)(c-b+2) - 2H(b,b)}{c-a+2}. \quad (3.88)$$

Given numbers $c \geq 0$ and $A \leq D$ assume that the cost functional is defined by

$$g(x) = cx, \qquad f(x,1) = |A-x| \qquad \text{and } f(x,2) = |D-x|. \quad (3.89)$$

We classify the optimal policies for the introduced model with respect to the parameters.

Theorem 3.6.2. *1. If $c = 2$ then any couple $0 \leq a \leq b \leq A$ is optimal.*

2. If $c > 2$ then the optimal couple is $a = b = 0$.

3. If $c < 2$ and $\sqrt{8/c} + A - 2 \leq D$
then $(a,b) = (A, \sqrt{8/c} + A - 2)$.

4. If $c < 2$, $\sqrt{8/c} + A - 2 > D$ and $\frac{4}{1+\sqrt{4-2c}} \geq D - A + 2$
then $(a,b) = (A, D)$.

5. If $c < 2$, $\sqrt{8/c} + A - 2 > D$ and $\frac{4}{1+\sqrt{4-2c}} < D - A + 2$
then $(a,b) = ((D+2) - (D-A)\sqrt{2/(2-c)}, D)$.

Remark 3.6.3. For $c = 2$ the cost functional $g(x) + 2f(x,1)$ is not strictly convex, therefore there is no contradiction with the theorem 3.3.8 stating the uniqueness of the optimal policy.

The proof of the theorem is technical and involves long, basic mathematical computations. For the sake of simplicity of the text, the proof is presented in Appendix A.

3.7 Appendix I- Uniformisable Markov processes

Suppose that $X(t)$ is a continuous-time Markov process (CTMP) with a countable state space S and generator matrix $Q = \{q_{ij}\}_{i,j \in S}$.

Definition 3.7.1. If the elements of the diagonal satisfy

$$\sup_{i\in S}(-q_{ii}) < \infty$$

we say that the Markov process is *uniformisable*.

Let $\gamma \geq \sup_i(-q_{ii})$ and define matrix P by

$$P = I + \frac{1}{\gamma}Q.$$

Suppose that Markov chain $\bar{X}_n$ has state space S and transition probability matrix P. Suppose that sequence $\{T_n\}_{n\geq 0}$ represents arrival times of a Poisson process with intensity γ and process $\widehat{X(t)}$ changes its state at times T_n so that

$$\widehat{X(T_n)} = \bar{X}_n.$$

Theorem 3.7.1. *The process $\widehat{X(t)}$ is a CTMP and its transition law coincides with the one of process $X(t)$.*

The proof of the theorem might be found in many sources for example in ([36]).

Therefore, a uniformisable CTMP can be viewed as a discrete-time Markov chain with matrix P, where changes occur in accordance with Poisson process $\{T_n\}$ with intensity λ.

3.8 Appendix II - Young topology

This section follows section 43, Davis [23]. We introduce Young topology. Suppose that $\mathcal{U} \subset \mathbb{R}^k$ is compact and $\mathcal{P}(\mathcal{U})$ is the space of probability measures on it. The set $\mathcal{R}$ of relaxed controls is the set of measurable functions $\nu : \mathbb{R}^+ \to \mathcal{P}(\mathcal{U})$. To introduce topology on $\mathcal{R}$ we need an auxiliary space X. Let $X = L_1(\mathbb{R}^+, C(U))$ the space of functions $f(t,u)$ measurable in t, continuous in u and satisfy

$$\|f\| = \int_0^\infty \max_{u\in\mathcal{U}} |f(t,u)|dt < \infty.$$

X is a Banach space under this norm. Its dual space is $X^* = L_\infty(\mathbb{R}^+, C^*(\mathcal{U}))$. Space $C^*(\mathcal{U})$ consists of the set of signed measures on $\mathcal{U}$ under the total variation norm. Therefore, X^* consists of measurable functions $\nu : \mathbb{R}^+ \to C^*(\mathcal{U}))$ such that

$$\|\nu\|_* = \text{esssup}_{t\in\mathbb{R}^+} \|\nu_t\|_{C^*} < \infty.$$

The weak$*$ topology on X^* is the topology which is equivalent to the following notion of convergence:

$$\nu_n \to \nu \iff (f, \nu_n) \to (f, \nu) \text{ for all } f \in X,$$

where

$$(f, \nu) = \int_0^\infty \int_{\mathcal{U}} f(t, u) d\nu_t(u) dt.$$

By Alaoglu Theorem unit ball B_1 in X^* is compact in the weak$*$ topology. The last thing to say is that $\mathcal{R}$ is a closed subset of B_1, hence is compact.

Definition 3.8.1. The Young topology $\mathcal{Y}$ on $\mathcal{R}$ is the relative weak$*$ topology of $\mathcal{R}$ considered as a subset of B_1.

Thus $(\mathcal{R}, \mathcal{Y})$ is a compact space.

Chapter 4

Fluid policies in semi-Markovian environment

In this chapter we assume that the random environment $\{E_n\} = \{\tau_n, Y_n\}$ satisfies Assumption (SMRE-II) 2.2.2.3. Recall that this is a situation, in which sequence $\{Y_n\}$ forms a Markov chain and inter-arrival times $\{\tau_{n+1}\}_{n\geq 0} = \{\tau_{(n+1,Y_n)}\}_{n\geq 0}$ depend on the value of Y_n only.

In Chapter 3 we showed that under Assumption (MRE) 2.2.2.2 the optimal control exists and belongs to the class of fluid policies. We have also proposed how to find the optimal fluid control numerically. For the case of a semi-Markovian environment, we cannot guarantee that the optimal control always belongs to the class of fluid policies. Moreover, in Chapter 5, we provide a number of examples of random environment, and establish the exact shape of the optimal controls, and it is obvious that these controls do not belong to the class of fluid policies.

Although optimality in the class of fluid policies does not guarantee optimality in the larger class of open-loop policies, finding the optimal control in the class of fluid policies may be attractive for multiple reasons. Firstly, the approach is computationally tractable. To find the solution one needs to solve a system of differential equations alike 3.55. Secondly, the fluid policies are simple and, therefore, easy to implement in practice. One needs to have a system controller with only three modes: "up" $(a(s) = U)$, "down" $(a(s) = -B)$ and "off" $(a(s) = 0)$.

The idea behind the proof of the existence of the optimal solution in the class of fluid policies follows from the basic topological fact: A continuous function attains its minimum in a compact set. We show that the set of fluid policies is compact in ℓ_∞. The main challenge then is to show that the average cost functional is continuous as a function of the set of levels, which define the fluid policy. To achieve this we apply the theory of regenerative processes and Markov chains to show that the limiting and the stationary at embedded moments distributions exist and are continuous as functions of control. The relevant definitions and theorems on both theories can be

found respectively in Sections 4.8 and 4.9. Continuity is not a trivial property. For example, it was studied for a single server queue in [37] and [60] and for various stochastic models in [47].

4.1 Controlled process associated with a fluid policy is regenerative

The theory of regenerative processes is a powerful tool to establish the existence of limiting distributions of stochastic processes under very mild assumptions. The definition 4.8.2 of regenerative process formalises the idea that the process might be split into i.i.d. cycles, so that the process starts over again in a new cycle and does not depend completely on the past.

Recall that function k is defined by formula 2.9 as

$$k(x, y, t) = \mathbb{I}_{\{x \leq y\}} \min(x + Ut, y) + \mathbb{I}_{\{x > y\}} \max(x - Bt, y).$$

Then the controlled process corresponding to the fluid policy starting at x_0 and given by function $l : \mathcal{Y} \to \mathbb{R}^+$ is

$$X(0) = x_0,$$

$$X(T_k + t) = k(X(T_k), l(Y_k), t) \quad \textit{for} \quad 0 < t < \tau_{k+1}, k \geq 1.$$

To prove the main results we need a further assumption on the random environment. Denote by $\Delta = \max(K/U, K/B)$, where K is the reasonable upper bound defined in Section 2.5.

Assumption 4.1.1. Markov chain Y_n is positive recurrent with atom $\{y_0\} \subset \mathcal{Y}$ so that

$$\mathbb{E}_{y_0} \sigma(y_0) < \infty, \text{ where} \tag{4.1}$$

$$\sigma(y_0) = \min\{k \geq 1 : Y_k = y_0\} \tag{4.2}$$

Moreover, inter-arrival times $\tau_{(n,Y_n)}$ are greater than Δ with positive probability

$$\mathbb{P}(\tau_{n,y_0} > \Delta) =: \delta_2 > 0. \tag{4.3}$$

Remark 4.1.1. Since the state space $\mathcal{Y}$ of the Markov chain $\{Y_n\}_{n \geq 0}$ is finite, the assumption on having a positive recurrent atom can be replaced by the assumption on the chain to be irreducible, because an irreducible finite-state Markov chain is

always positive recurrent.

Fix element $y_1 \in \mathcal{Y}$ such that $\mathbb{P}(Y_{n+1} = y_1 | Y_n = y_0) =: \delta_1 > 0$.

Theorem 4.1.2. *Suppose that Assumption 4.1.1 holds. Then $X(t)$ is a regenerative process. Namely, there exists a renewal process $S_n = Z_1 + \ldots + Z_n, \quad n = 1, 2, \ldots$ such that*

- *The post-S_n process $\{X_{S_n+t}\}_{t\geq 0}$ is independent of $\{S_k\}_{k=1}^n$;*
- *The distribution of the post-S_n process is independent of n. In particular, elements $\{Z_k\}_{k\geq 2}$ are identically distributed;*
- *Increments $Z_{n+1}, Z_{n+2}, \ldots$ are independent of $\{S_k\}_{k=1}^n$. In particular, elements $\{Z_k\}_{k\geq 1}$ are independent.*

Moreover, the length of a typical cycle $Z_n, (n \geq 2)$ has finite mean

$$\mathbb{E}Z_n < \infty.$$

Proof. Due to the independence of τ_n of Y_n and Markov property of $\{Y_n\}_{n\geq 0}$ we have

$$\mathbb{P}(\tau_n > \Delta, Y_n = y_1 | Y_{n-1} = y_0) = \delta_1 \times \delta_2 = \gamma > 0.$$

Consider a sequence of events

$$\{A_n\}_{n\geq 0} = \{\omega : Y_{n+1} = y_1, Y_n = y_0, \tau_{n+1} > \Delta\}_{n\geq 0}.$$

Note that event A_n implies that the level $l(y_0)$ will be reached regardless of value $X(T_n)$, and therefore $(X(T_{n+1}), Y_{n+1}) = (l(y_0), y_1)$. Hence, if event A_n occurs, then the post-T_n process does not depend on the past.

Denote by $\sigma_1, \sigma_2, \ldots$ consecutive occurrences of the events A_n:

$$\sigma_1(\omega) = \min(k : I(A_k) = 1),$$

$$\sigma_i(\omega) = \min(k > \sigma_{i-1} : I(A_k) = 1).$$

We also need

$$\eta_{l,0}(\omega) = \sigma_l(\omega),$$

$$\eta_{l,1}(\omega) = \min(k > \sigma_l : Y_k = y_0),$$

$$\eta_{l,i}(\omega) = \min(k > \eta_{l,i-1} : Y_k = y_0).$$

Denote by $\alpha_k = \sigma_k - \sigma_{k-1}$ the length of k-th cycle. Then sequence $\{\alpha_k\}_k \geq 0$ is i.i.d. because

$$\alpha_k =_d \sum_{i=1}^{\Gamma_k} \eta_{l,i},$$

where Γ_k is a geometric random variable with parameter γ independent of future values $(\tau_i, Y_i)_{i>\alpha_k}$.

The structure of the cycles is depicted in the figure below.

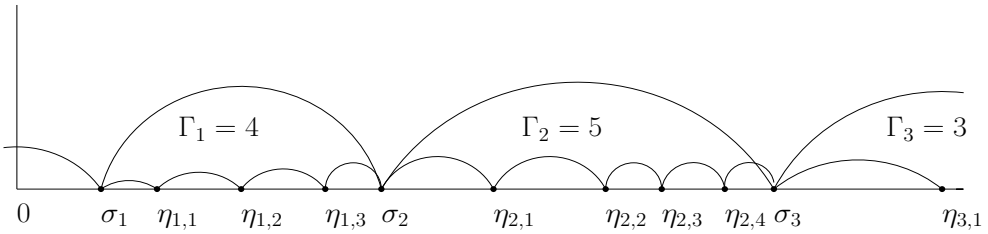


Figure 4.1: Regenerative cycles

Since $\{Y_n\}$ is Harris ergodic with positive recurrent atom $\{y_0\}$ we have $\mathbb{E}(\eta_{i,j+1} - \eta_{i,j}) < \infty$. Due to the independence of Γ_k of future values $(\tau_i, Y_i)_{i>\alpha_k}$ one may apply the Wald's identity

$$\mathbb{E}\alpha_k \leq \frac{1}{\gamma}\mathbb{E}\eta_{1,1} < \infty. \tag{4.4}$$

Consider sequence $\{Z_i\}_{i\geq 1} = \{T_{\sigma_i} - T_{\sigma_{i-1}}\}_{i\geq 1}$. To show that $\{Z_i\}_{i\geq 2}$ is i.i.d. we write

$$Z_i = \tau_{\sigma_{i-1}+1} + \tau_{\sigma_{i-1}+2} + \ldots + \tau_{\sigma_i - 1} + \tau_{\sigma_i} \tag{4.5}$$

$$= \sum_{k=1}^{\Gamma_i} \sum_{j=\eta_{i,k-1}+1}^{\eta_{i,k}} \tau_j. \tag{4.6}$$

Cycles $(\tau_{\eta_{i,k-1}+1}, \tau_{\eta_{i,k-1}+2}, \tau_{\eta_{i,k}})_{i\geq 1}$ are i.i.d. Their number is equal to Γ_i and it is independent of i.

Moreover, since Γ_i is independent of future values $(\tau_i, Y_i)_{i>\alpha_k}$ we can apply the Wald's identity

$$\mathbb{E}Z_i = \mathbb{E}\sum_{k=1}^{\Gamma_i} \sum_{j=\eta_{i,k-1}+1}^{\eta_{i,k}} \tau_j + \tau_{\sigma_i}$$

$$= \mathbb{E}\Gamma_i \times \mathbb{E}\left(\sum_{j=\eta_{i,k-1}+1}^{\eta_{i,k}} \tau_j\right)$$

$$\leq \frac{1}{\gamma} \times \mathbb{E}\alpha_k \times \max_y \mathbb{E}\tau_{1,y} < \infty. \tag{4.7}$$

Hence T_{σ_k} is a delayed renewal process $T_{\sigma_k} = T_{\sigma_1} + Z_2 + \ldots + Z_k$. The process $X(t)$ is regenerative and T_{σ_k} is its imbedded renewal process. □

4.2 Existence of the limiting distribution

To establish existence of the limiting distribution for regenerative processes $X(t)$ we require the following

Assumption 4.2.1. Elements $\{\tau_n\}$ have a spread-out distribution.

Definition 4.8.3 recalls the definition of the spread-out distribution. The assumption about the spread-out distribution is a generalisation of the assumption of non-periodicity, which is required for the stability of a Markov Chain.

Due to the formula 4.5, the assumption implies that the increments of the imbedded renewal process Z_n also have a spread-out distribution. Furthermore, the formula 4.7 shows that the length of the usual cycle has finite mean. The straightforward application of Theorem 4.8.1 results in the following:

Proposition 4.2.1. *Suppose that Assumptions 4.2.1 and 4.1.1 hold. Then the limiting distribution of the process $X(t)$ exists and is independent of the starting point x_0. The convergence holds in the sense of total variation distance.*

We will refer to the limiting distribution as π_1.

4.3 Imbedded Markov process and stationary distribution

The terminal charges $f(x, y)$ are applied at imbedded epochs $\{T_n\}_{n\geq 0}$. The stationary distribution of the imbedded process $\{X(T_n)\}_{n\geq 0}$ does not necessarily coincide with the limiting distribution in continuous time π_1. The property, when both distributions are the same, is widely known as PASTA (Poisson arrivals see time averages) and was first studied in [63]. Later it was studied for a variety of stochastic models, see [57], [38], for example. In Chapter 3, we used reasoning similar to PASTA to establish that the terminal and the running charges can be treated in a similar way, if the environment is Markovian (satisfies Assumption MRE). In the semi-Markov case PASTA property generally does not hold. For example, in papers [64], [17] it was shown that for certain models the fact that the limiting and the stationary at imbedded moments distributions coincide implies exponentiality of inter-arrival times and this property is known as ANTI-PASTA.

In this section we establish sufficient conditions for the stationary distribution at imbedded moments to exist.

We will apply the following well known theorem, for the proof see [16].

Theorem 4.3.1. • *Suppose that X_n is a discrete time-homogeneous Markov chain defined on state space $\mathcal{X}$ with finite generated σ-algebra $\mathcal{B}(\mathcal{X})$.*

Then there exists a function $W : \mathcal{X} \times [0,1] \to \mathcal{X}$ and an i.i.d. sequence χ of random variables uniformly distributed on the interval $[0,1]$ (the distribution is denoted by U[0,1]), such that the representation holds almost surely

$$X_{n+1} = W(X_n, \chi) \quad \textit{for all} \quad n \geq 0.$$

• *Suppose that a sequence of random variables $\{X_n\}_{n\geq 0}$ is given by*

$$X_{n+1} = W(X_n, \chi),$$

where elements $\chi_{n\geq 0}$ are an i.i.d sequence.

Then $\{X_n\}$ is a Markov chain.

By the definition of a fluid policy, we have

$$X(T_{n+1}) = k(X(T_n), l(Y_n), \tau_{n,Y_n}).$$

We know that $\tau_n = \tau_{n,Y_n}$ is independent of $\{\tau_k, Y_k\}_{k\neq n}$ conditionally on Y_n. Hence, there exists a sequence $\{\chi_{n,1}\}_{n\geq 0}$ of i.i.d $U[0,1]$ variables and a function W_1 such that the representation holds

$$\tau_n = W_1(Y_n, \chi_{n,1}).$$

Markov property for chain $\{Y_n\}$ implies that there exists a function W_2 and an i.i.d. sequence $\{\chi_{n,2}\}$ of $U[0,1]$ variables such that

$$Y_{n+1} = W_2(Y_n, \chi_{n,2}).$$

Therefore,

$$X(T_{n+1}, Y_{n+1}) = (k(X(T_n), l(Y_n), W_1(Y_n, \chi_n)), W_2(Y_n, \chi_2)).$$

Hence, by applying theorem 4.3.1 we obtain that the chain $(X(T_n), Y_n)$ is Markov.

We explore necessary conditions for the process to be Harris ergodic. The Harris ergodicity is a sufficient condition for the stationary probability measure to exist and be unique. Moreover, the process converges to its stationary distribution in total variation norm. The main definitions and statements on Harris chains can be found in Section 4.9.

Recall elements y_0 and y_1 from Section 4.1. To show that atom $R = (l(y_0), y_1)$ is a regenerative set for Markov chain $\{X(T_n), Y_n\}_{n\geq 0}$ it is enough to show that this set is positive recurrent, since any one-point recurrent set is regenerative (see e.q. Asmussen [3]). Let $\sigma(R) = \min(n > 0 : (X(T_n), Y_n) \in R, (X(0), Y_0) \in R)$. We show that $\mathbb{E}\sigma(R) < \infty$ holds.

Notation of the previous subsection allows us to write

$$\mathbb{E}\sigma(R) = \mathbb{E}(\sigma_2 - \sigma_1) < \infty,$$

where boundedness of the latter was established in Theorem 4.1.2.

Therefore, $(X(T_n), Y_n)$ is a Harris chain and the stationary measure exists and is unique. We refer to it as π_2.

To show that the convergence to the stationary distribution is in total variation norm, one needs aperiodicity of the process. In our case, aperiodicity of chain Y_n is a sufficient condition.

4.4 Average cost functional represented through stationary distributions

In this section we show that the average cost functional may be regarded as a function of limiting distribution π_1 and stationary distribution in imbedded moments π_2. This yields that the upper limit in definition of the average cost functional may be replaced by the ordinary limit and does not depend on the initial state. We prove that the convergence is in fact the a.s. convergence and not only the convergence of means.

Recall that $\nu_t = \max\{n : T_n \leq t\}$.

Theorem 4.4.1. *Suppose that Assumptions 4.1.1 and 4.2.1 hold.*
Then for any x_0 we have

- *The a.s. convergence*

$$\frac{1}{t}\left(\int_0^t g(X(s))ds + \sum_{i=1}^{\nu_t} f(X(T_i), Y_{i-1})\right) \xrightarrow{a.s.} \mathbb{E}_{\pi_1} g(X(0)) + \frac{\mathbb{E}_{\pi_2} f(X(0), Y_0)}{\mathbb{E}\tau_1};$$

- *The L_1 convergence*

$$\mathbb{E}\left|\frac{1}{t}\left(\int_0^t g(X(s))ds + \sum_{i=1}^{\nu_t} f(X(T_i), Y_{i-1})\right) - \left(\mathbb{E}_{\pi_1} g(X(0)) + \frac{\mathbb{E}_{\pi_2} f(X(0), Y_0)}{\mathbb{E}\tau_1}\right)\right| \to 0;$$

- *The convergence of means*

$$\mathbb{E}\frac{1}{t}\left(\int_0^t g(X(s))ds + \sum_{i=1}^{\nu_t} f(X(T_i), Y_{i-1})\right) \to \mathbb{E}_{\pi_1} g(X(0)) + \frac{\mathbb{E}_{\pi_2} f(X(0), Y_0)}{\mathbb{E}\tau_1}.$$

In particular we have,

$$\mathcal{C}(X) = \lim_{t\to\infty} \mathbb{E}\frac{1}{t}\left(\int_0^t g(X(s))ds + \sum_{i=1}^{\nu_t} f(X(T_i), Y_{i-1})\right)$$

$$= \mathbb{E}_{\pi_1} g(X(0)) + \frac{\mathbb{E}_{\pi_2} f(X(0), Y_0)}{\mathbb{E}\tau_1},$$

so that the limit exists and does not depend on the initial state $X(0)$.

Proof. Both components can be treated in the same way since Markov chain $(X(T_n), Y_n)$ is regenerative in discrete time. Therefore, we proceed with the component corresponding to the continuous time only.

- The strong law of large numbers and the central limit theorem were proven for averages of functions of regenerative processes in Whitt [61] and a proof is also provided in Asmussen [3]. To apply the result, we need that to be sure that the length of a usual cycle has finite mean

$$\mathbb{E}Z_i < \infty,$$

and that the maximal difference in values of the function f, namely random variable
$\max_{S_{n-1}\le t<S_n}(f(X(t)) - f(X_{S_n(t)})) =: V_n$ has finite mean

$$\mathbb{E}V_n < \infty.$$

The first was shown by formula 4.7. The second holds since the controlled process does not leave compact set $[0, K]$.

- Almost sure convergence and uniform integrability result in convergence in L_1-norm. Here, uniform integrability holds in even stronger sense of uniform dominance, since process $X(t)$ does not leave compact set $[0, K]$ due to Assumption 2.5.1.

- Convergence of means follows from convergence in L_1-norm.

□

4.5 Topology on the set of level functions l

Each level function $l : \mathcal{Y} \to [0, K]$ can be regarded as an element of space $[0, K]^{\mathcal{Y}}$. Suppose that each of the copies $[0, K]$ is embedded with standard Euclidian topology, then due to Tikhonov's theorem the product space is compact in the sense of the product topology. The norm on the product space can be arbitrarily chosen from a set of ℓ_p-norms

$$d_p(l_1, l_2) = \sum_{y \in \mathcal{Y}} (|l_1(y) - l_2(y)|^p)^{1/p}, 1 \leq p < \infty \tag{4.8}$$

$$d_\infty(l_1, l_2) = \max_y |l_1(y) - l_2(y)|, \tag{4.9}$$

though for our purposes the most convenient is d_∞-metric.

4.6 Continuity of the limiting and the stationary distribution with respect to the levels function

To establish the continuity we need to introduce the Lévy metric. We show that convergence of the sets of levels in d_∞-metric results in convergence of the corresponding limiting (and stationary) distributions in the Lévy metric and also in the weak sense. This holds as convergence in Lévy metric is equivalent to the weak convergence, see [14], for example.

Definition 4.6.1. Let $F, G : \mathbb{R} \to [0, 1]$ be two probability distribution functions. Define the Lévy distance between them to be

$$L(F, G) := \inf\{\varepsilon > 0 | F(x - \varepsilon) - \varepsilon \leq G(x) \leq F(x + \varepsilon) + \varepsilon, \text{for all } x \in \mathbb{R}\}.$$

We use the same notation for random variables $L(X_1, X_2) := L(F, G)$, where F and G are their distribution functions.

The space of real-valued measures with Lévy metric is a complete metric space, see for example [14].

Theorem 4.6.1. *Suppose that controlled processes X_1 and X_2 are defined by sets of levels $\{l_1(y)\}$ and $\{l_2(y)\}$ respectively and $d_\infty(l_1, l_2) = \delta$. Then*

1. *If $|X_1(0)(\omega) - X_2(0)(\omega)| \leq \delta$ a.s. then for any $t > 0$ the following holds*

$$|X_1(t)(\omega) - X_2(t)(\omega)| \leq \delta \text{ a.s.}$$

2. *For any* $X_1(0) = x_1$ *and* $X_2(0) = x_2$ *there exists* $t^* > 0$ *such that, for* $t > t^*$, *the following holds*

$$|X_1(t)(\omega) - X_2(t)(\omega)| \leq \delta \text{ a.s. .}$$

If random variables $\tau_{i,y}$ *have unbounded support, then we can guarantee*

$$\mathbb{E}t^* < \infty.$$

3. *Suppose that* π_1^1 *and* π_1^2 *are limiting distributions of the processes* $X_1(0)$ *and* $X_2(0)$, *then*

$$L(\pi_1^1, \pi_1^2) \leq \delta.$$

4. *Suppose* π_2^1 *and* π_2^2 *are stationary distributions of the first components of the imbedded Markov chains* $\{(X_1(T_k), Y_k)\}$ *and* $\{(X_2(T_k), Y_k)\}$, *then*

$$L(\pi_2^1, \pi_2^2) \leq \delta.$$

Proof. 1. For $t :\in [T_k, T_{k+1}]$ the following is true

$$|X_1(t)(\omega) - X_2(t)(\omega)| \leq \max\left(l_1(Y_k) - l_2(Y_k), X_1(T_k)(\omega) - X_2(T_k)(\omega)\right) \leq \delta.$$

Therefore,

$$|X_1(T_k)(\omega) - X_2(T_k)(\omega)| \leq \delta \text{ a.s.}$$

yields

$$|X_1(T_{k+1})(\omega) - X_2(T_{k+1})(\omega)| \leq \delta \text{ a.s.}$$

and then the statement will follow by induction.

2. Consider a sequence of events

$$C_k = \{\tau_k > \max(K, x_1, x_2)/\min(U, B)\}.$$

If C_k occurs then $X_1(T_k) = X_2(T_k)$ and $Y_1(k-1) = Y_2(k-1)$ so

$$|X_1(T_k)(\omega) - X_2(T_k)(\omega)| = |l_1(Y_k) - l_2(Y_k)| \leq \delta.$$

Using the induction argument from the previous part of the theorem we obtain that the similar holds for all $t \geq T_k$.

Let $\sigma(C) = \min\{K : I(A_k) = 1\}$ for the number of the first successful event A_k. Then $\sigma(C)$ is a geometric random variable with mean $\frac{2}{p(x_1,x_2)}$, where $p(x_1, x_2) = \inf_y \mathbb{P}(\tau_{k,y} > \max(K, x_1, x_2)/\min(U, B))$. Therefore, $\mathbb{E}\sigma(C) < \infty$.

3. Proof of two last statements is similar. Convergence in the Lévy metric is equivalent to the weak convergence. Since the differences between corresponding elements of two sequences are uniformly bounded, the same applies to their weak limits.

□

The straightforward corollary takes place

Corollary 4.6.2. *Limiting distribution $\pi_1(l)$ and stationary distribution $\pi_2(l)$ are continuous functions with respect to Lévy-Prokhorov metric.*

Convergence of $d_\infty(l_n, l) \to 0$ results in the weak convergence of the corresponding distributions.

4.7 Continuity of the average cost functional with respect to the levels function and the existence of the optimal fluid policy

The very last step in this chapter is to combine the convergence results from Theorem 4.6.1 with the convergence result of Theorem 4.4.1 saying that the cost functional is a function of the limiting and the stationary distributions. By superposition this results in the following

Proposition 4.7.1. *Average cost functional $\mathcal{C}$ is continuous as a function of levels $l(y)$ in ℓ_∞ topology.*

Proof. Suppose that $l_n \to l$ in d_∞. Denote by $\pi_1(l_n)$ and $\pi_2(l_n)$ the corresponding limiting and stationary processes of l_n and by $\pi_1(l)$ and $\pi_2(l)$ those of l.

Then for any continuous function $H : [0, K] \to \mathbb{R}^+$, for any $\varepsilon > 0$ there exists $N > 0$ such that for $n > N$

$$\left| \int_0^K H(x) d\pi_i(l_n)(x) - \int_0^K H(x) d\pi_i(l)(x) \right| \leq \varepsilon.$$

Recall the representation of the cost functional from Theorem 4.4.1 as an integral function of the limiting and the stationary distributions. Combining the last two arguments we get that the average cost functional is continuous as a function of the set of levels l. □

The existence of the optimal fluid policy is a consequence of the compactness of the set of level functions and the continuity of the cost functional with respect to this topology.

We summarise our main findings of this chapter in the following theorem.

Theorem 4.7.2. *If Assumptions 4.1.1 and 4.2.1 hold, then the optimal policy in the class of fluid polices exist.*

4.8 Appendix I - Regenerative processes

Definition 4.8.1.
Suppose that $(Y_i)_{i\in\mathbb{Z}^+}$ are independent positive valued random variables and that $(Y_i)_{i>0}$ are identically distributed. Let $S_0 = Y_0$ and $S_n = S_{n-1} + Y_n$. Then $\{S_n\}$ is called a renewal process.

Definition 4.8.2.
We call stochastic process X_t regenerative if there exists a renewal process $S_n = Y_0 + \cdots + Y_n$ such that the post-S_n process $((Y_{n+k})_{k\in\mathbb{N}}, (X_{S_n+t})_{t>0})$ is independent of $Y_0, \ldots Y_n$ and its distribution does not depend upon n. For $Y_0 = 0$ we will call by P_0 and $\mathbb{E}_0$ the probability and expectation functionals of the corresponding renewal process and refer to this scenario as 0-delayed case. If $Y_0 \neq 0$ then we refer to it as delayed case.

Definition 4.8.3.
We say that distribution F is spread out if there exists an integer n and an absolutely continuous distribution G so that

$$F^{*n}(A) \geq G(A).$$

We refer to distribution G as to the absolutely continuous component of distribution F.

Theorem 4.8.1.
Assume that a (possibly delayed) regenerative process X_t has metric state space, right-continuous paths and non-lattice cycle length distribution F with finite mean μ. Then the limiting distribution, say P_e, of X_t exists and is given by

$$\mathbb{E}_e f(X_t) = \frac{1}{\mu}\mathbb{E}_0 \int_0^Y f(X_s)ds$$

Suppose that F is a spread-out distribution and the paths $X_t(\omega)$ are measurable then the above convergence also holds in total variation distance sense.

4.9 Appendix II - Markov processes

This section summarises main definitions and theorems on the theory of Markov chains, along the lines of [3]. Suppose that there is a given sequence of random

variables $\{X_n\}_{n\geq 0}$ on a common probability space Ω with Borel sigma-algebra $\mathcal{B}(\Omega)$ taking values in $(S, \mathcal{B}(S))$. Denote by $\mathcal{F}_n$ the sigma-algebra generated by random variables $X_0, \ldots, X_n$.

Definition 4.9.1.
Sequence X_n forms a Markov chain if

$$\mathbb{P}\{X_{n+1} \in A | \mathcal{F}_n\} = \mathbb{P}\{X_{n+1} \in A | X_n\}.$$

We write $\mathbb{P}^r(x, A) = \mathbb{P}_x(X_r \in A)$, for $r \geq 0$.

Definition 4.9.2.
Suppose that X_n is a Markov chain. Let $\tau(R) = \inf\{k \geq 1 : X_k \in R\}$.

- Then a set $R \in \mathcal{B}(S)$ is called recurrent if $\tau(R)$ is measurable and $\mathbb{P}_x(\tau(R) < \infty) = 1$.
- We call R a regeneration set if R is recurrent and for some $r > 0$ there exist $\varepsilon \in (0, 1)$ and probability measure λ on $(\Omega, \mathcal{B}(\Omega))$ such that for all $x \in R$ and $B \in \mathcal{B}(\Omega)$ the following holds

$$\mathbb{P}^r(x, B) \geq \varepsilon \lambda(B).$$

- We call a Markov chain X_n with a regeneration set Harris recurrent.
- It is shown in Asmussen (p199, [3]) that a Harris chain is a regenerative process. Denote by Y the length of the first cycle of the 0-delayed process. We call the chain positive recurrent if $\mathbb{E}Y < \infty$. We call the chain aperiodic if distribution of Y is aperiodic.
- Aperiodic positive recurrent Harris chain is called Harris ergodic.
- A σ-finite measure ν is called stationary if $\nu \geq 0$, $\nu \neq 0$ and

$$\nu \mathbb{P}^r = \nu, \text{ for all } r \geq 0.$$

Theorem 4.9.1 (Theorems 3.5-6, [3])**.**

- *For a Harris recurrent Markov chain the stationary probability measure exists and is unique.*
- *For a Harris ergodic chain the $\mathbb{P}_x$ distribution of X_n converges to its stationary measure π in total variation norm.*

Chapter 5

Other environments

In this chapter we consider three different options for inter-arrival times τ_n in a particular case, where the only randomness in the system comes from the inter-arrival times. Thus, in terms of our model we assume that $|\mathcal{Y}| = 1$. The primary goal of this chapter is to present a multitude of possible behaviours of optimal policies depending on the inter-arrival times. In all scenarios we assume that $\{\tau_n\}$ is i.i.d. sequence.

The scenarios are:

1. Markovian randomness, $\tau_n \sim \exp(\lambda)$, a subcase of MRE stochastic assumptions (this case was studied thoroughly in Chapter 3);

2. Full predictability, $\tau_n \equiv T$, where T is a constant known in advance;

3. Blend of predictability and randomness, $\tau_n = T + \exp(\lambda)$, so-called shifted exponential distribution.

We will describe optimal policies under all three scenarios, and establish a link between them.

In the basic energy supply model 1.2 these scenarios model different behaviour of the error process for the wind power prediction. We analyse them separately below.

Scenario 1: Markovian randomness

The memorylessness property of the exponential distribution suggests that regardless of the time passed since the last shortfall, the distribution of the remaining waiting time is unchanged. On the other hand, for all other distributions the following is true: the time passed since the last shortfall provides an additional information about the remaining waiting time. In theory, knowing this information, we might adapt the policy with time. Although, under the considered scenario, there is no new information available with time, which would suggest to adjust the policy. The intuitive explanation was suggested by Stan Zachary and led to the result that the constant policy is optimal in this case.

The rigorous proof of this result is a straightforward corollary of Theorem 3.3.8 to the case where $|\mathcal{Y}| = 1$ and we summarised it as Corollary 3.3.12. The constant level l^* will be found in Section 5.1.

Scenario 2: Full predictability

This models the situation, in which the times between two consecutive shortfalls in power supply (the errors in the wind power prediction) are known in advance.

Although the situation is not very realistic, we consider it for the sake of mathematical completeness of the study. Chronologically, the results for deterministic inter arrival times were amongst the first results obtained for the thesis.

Intuition suggests, that due to the lack of randomness the information available to a decision maker at times T_n does not change with time. Hence, the policy should be periodic. This intuition is formalised in Section 5.2. Moreover the exact equation for the optimal policy is found.

Scenario 3: Blend of predictability and randomness

Under this scenario the time between shortfalls is known to be greater than T, but after time T the waiting time does not provide any additional information. This might model an English idiom "the calm before the storm" or if we look form the other side (and change the idiom a little bit) "the calm after the storm". The author is neither expert in meteorology nor in the scientific base of English idioms, although still believes that such a situation might be close to reality.

It is natural to expect a combination of behaviours of the two previous scenarios for the optimal policy. In Section 5.3 we show that the intuition is indeed correct.

Three sections in this chapter deal are dedicated to scenarios (1-3) respectively.

5.1 Markovian randomness

The main purpose of this section is to find an expression for the constant l^* from Corollary 3.3.9. Recall that it is the constant, such that policy $X(t) \equiv l^*$ is time average optimal policy. We show that such a policy is locally optimal and, hence, is time average optimal. Since $|\mathcal{Y}| = 1$ we may suppress y from notation and write $f(x)$ meaning $f(x, y)$.

Theorem 5.1.1. *The policy optimizing the time average cost functional is a constant function l^*, where*

$$l^* = \arg\min_x \left(g(x) + \lambda f(x)\right).$$

Proof. Due to Lemma 3.1.1 the one time-period cost functional can be expressed as

$$\mathbb{E}\frac{g(X(\tau))}{\lambda} + f(X(\tau)).$$

Therefore $l^* = \arg\min_x (g(x) + \lambda f(x))$ minimises the cost functional locally, and, hence, is the optimal solution. □

Note that the level l^* does not depend on the ramp constraints. As we have just shown the optimal level l^* is the argminimum of the convex combination of convex functions, hence the following corollary holds.

Corollary 5.1.2.
- *If* $\lambda \to 0$ *then* $l^* \to \arg\min_x g(x) = 0$.
- *If* $\lambda \to \infty$ *then* $l^* \to \arg\min_x f(x)$.

5.2 Full predictability

In this section we assume that the charges $f(x, y)$ are applied at time instants $\{kT\}_{k\geq 0}$ and that there is only one possible value of y so we may write $f(x, y) = f(x)$. We start by showing that the class of feasible policies can be replaced with a smaller class, so that the optimal policy belongs to the new class. The search will be then done in this class.

Suppose that $s(v)$ is an arbitrary function. With the knowledge of values $s(iT)$, $i = 1, 2, \ldots$ we construct a new function $\bar{s}(y)$, such that the value of the cost functional not larger. Consider a set of non-negative valued continuous functions $\tilde{s}$ such that $\tilde{s}(iT) = s(iT)$ for all i and the left derivative of $\tilde{s}$ on each interval $[iT, (i+1)T]$ is an increasing function with values in $\{-B, 0, U\}$. Take $\bar{s}$ as the minimal function in this set, so that

$$\tilde{s}(v) \geq \bar{s}(v) \qquad \text{for all } \tilde{s}(v) \text{ and } y.$$

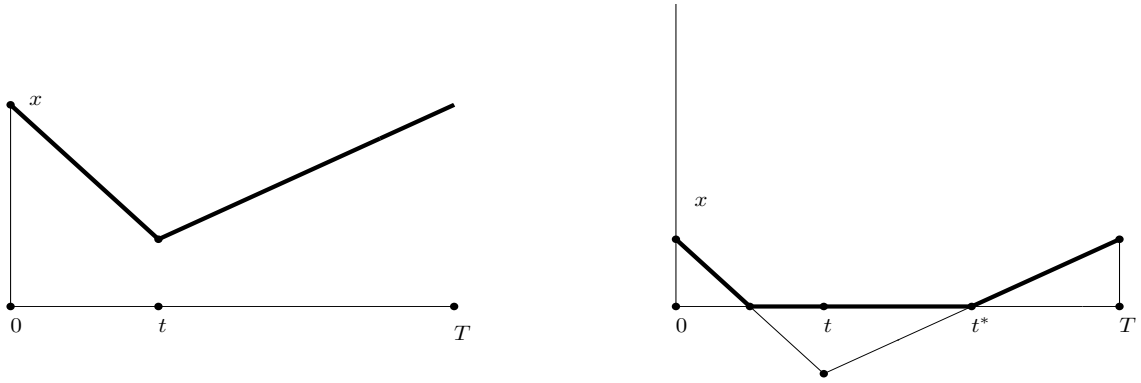
It is easy to see that the minimal function exists and is unique. The value of the cost functional is smaller for $\bar{s}$ than for s, because the part corresponding to the area beneath the plot of the function is smaller, since $\bar{s}(v) \leq s(v)$ for all v.

Therefore we restrict our search if the optimal policy to the class of functions

$$\mathcal{S} = \{s(v) : s'(v) \text{ is increasing }, s'(v) \in \{-B, 0, U\} \text{ and } s'(v) = 0 \text{ only if } s(v) = 0\}.$$

As we have just shown, the functions from class $\mathcal{S}$ can be uniquely defined and hence parametrised by the sequence $\{s(iT)_{i\in\mathbb{Z}}\}$.

Now we proceed with exploring properties of the cost functional on a single period of time $[0, T]$. Due to the last observation, we work only with the functions from class $\mathcal{S}$ restricted on interval $[0, T]$. An alternative way to parametrise these functions by (x, t), where $s(0) = x$, is presented in the picture below.



An easy computation shows that

$$s(x,t,y) = \begin{cases} (x - By)^+, \text{ if } y < t \\ (x - Bt + U(y-t))^+, \text{ if } y \geq t. \end{cases} \tag{5.1}$$

Recall the classic definition from analysis, see [50] for the theory of convex functions.

Definition 5.2.1. Suppose that $S \subset \mathbb{R}^k$. A function $f : S \to \mathbb{R}$ is convex if for any point $x, y \in S$ and $p \in [0,1]$ so that $px + (1-p)y \in S$ holds

$$f(px + (1-p)y) \leq pf(x) + (1-p)f(y).$$

Lemma 5.2.1. *Function $I(x,t) = \int_0^T g(s(x,t,y))dy$ is convex.*

Proof. To prove that a twice-differentiable function of two variables is convex, it is enough to show that its Hessian is positive semidefinite. By definition

$$H(I) = \det \begin{pmatrix} I''_{xx} I''_{xt} \\ I''_{tx} I''_{tt} \end{pmatrix} = I''_{xx} \times I''_{tt} - I''^2_{tx}$$

Using the property of differentiation under the integral sign one gets

$$I'_x(x,t) = \int_0^T g'(s(x,t,y))s'_x(x,t,y)dy,$$

$$I''_{xx}(x,t) = \int_0^T g''(s(x,t,y))(s'_x(x,t,y))^2 + g'(s(x,t,y))s''_{xx}dy =$$

$$= \int_0^T g''(s(x,t,y))(s'_x(x,t,y))^2 dy,$$

$$I'_t(x,t) = \int_0^T g'(s(x,t,y))s'_t(x,t,y)dy,$$

$$I''_{tt}(x,t) = \int_0^T g''(s(x,t,y))(s'_t(x,t,y))^2 + g'(s(x,t,y))s''_{tt}dy =$$

$$= \int_0^T g''(s(x,t,y))(s'_t(x,t,y))^2 dy,$$

$$I''_{tx}(x,t) = \int_0^T g''(s(x,t,y))s'_t(x,t,y)s'_x(x,t,y) + g'(s(x,t,y))s''_{tx}dy =$$

$$= \int_0^T g''(s(x,t,y))s'_t(x,t,y)s'_x(x,t,y).$$

To obtain the inequality for the Hessian, we apply Hölder inequality to the functions $h_1 = g''(s(x,t,y))^{1/2}(s'_x(x,t,y))$ and $h_2 = -g''(s(x,t,y))^{1/2}(s'_t(x,t,y))$. Functions h_1 and h_2 are correctly defined, because function g is convex. Therefore,

$$I''_{xx} \times I''_{tt} - I''^2_{tx} = \int_0^T g''(s(x,t,y))(s'_x(x,t,y))^2 dy \times \int_0^T g''(s(x,t,y))(s'_t(x,t,y))^2 dy -$$

$$- \left(\int_0^T g''(s(x,t,y))s'_t(x,t,y)s'_x(x,t,y) \right)^2 =$$

$$\int_0^T h_1^2(y)dy \times \int_0^T h_2^2(y)dy - \left(\int_0^T h_1(y)h_2(y)dy \right)^2 \geq 0.$$

□

5.2.1 Optimisation for a single period problem

For further investigation, we introduce a family of additional one-period cost functionals.

Definition 5.2.2. Suppose that $p \in [0,1]$ and let $q = 1 - p$. We define p-cost functional by the equation

$$\mathcal{C}_p(x,t) = pf(x) + \int_0^T g(s(x,t,y))dy + qf(x + UT - (U+B)t).$$

The main idea behind the definition can be easily explained. We would like to show that the solution of the average cost problem can be constructed as a solution to a single period problem. Although, clearly for the single period cost functional the optimal solution is to be at level 0 for as long as possible and then to increase at maximal pace U. The original cost functional treats values $x(0)$ and $x(T)$ significantly different.

On the other hand, the optimal policy for the average cost optimisation problem is shift preserving ($X(s) = X(T+s)$), so the end point for one period is the start one for the next one.

Thus, we try to balance the impact of the start and end points to the cost functional. The p-cost functionals formalise this idea

Let us find a point t^* such that $s(x, t^*, T) = x$. We have $s(x, t^*, T) = x + UT - (U + B)t^* = x$, and hence $t^* := \frac{UT}{U+B}$.

Theorem 5.2.2. *1. Function $\mathcal{C}_p(x, t)$ is strictly convex.*

2. For each $p \in [0, 1]$ there exist a unique pair $(x(p), t(p))$ minimising $C(p, x, t)$.

3. Let $I(x, t) = \int_0^T g(s(x, t, y))dy$. Suppose that

$$f'(0) + I'_x(0, t^*) \leq 0. \tag{5.2}$$

(a) There exists x^* such that for all $p \in [0, 1]$ holds $(C_p(x^*, t^*))'_x = 0$.

(b) There exists p^* such that $(C_{p^*}(x^*, t^*))'_t = 0$

(c) Policy $s(x^*, t^*, y)$ is p^*-optimal.

4. Suppose that

$$f'(0) + I'_x(0, t^*) > 0.$$

Then a policy $s(x^*, t^*, y) = s(0, t^*, y) \equiv 0$ is optimal for functional $C_p(x, t)$ for any p.

5. Suppose that $S_{p^*}(t)$ is a periodic extension of the function $s(x(p^*), t(p^*), y)$ from $[0, T]$ on $\mathbb{R}$. If for another policy $\hat{S}(t)$ there exists $\varepsilon > 0$ such that

$$\lim_{n\to\infty} \frac{|\{n : |\hat{S}(nT) - S_{p^*}(nT)| \geq \varepsilon\}|}{n} \neq 0$$

then for the average cost functional $C(S)$ holds

$$C(\hat{S}) > C(S_{p^*}).$$

Proof. 1. Function $C_p(x, t)$ is strictly convex as a convex combination of strictly convex function $f(x)$ and convex functions $I(x, t)$ and $f(s(x, t, T))$.

2. Strictly convex function $C_p(x, t)$ attains its minimum on compact set $[0, \arg\min f(x)]$. The optimal policy cannot lie above level $\arg\min f(x)$, because g is an increasing function.

3. (a) For the derivative with respect to t at point (x, t^*) one has

$$(\mathcal{C}_p(x, t^*))'_x = pf'(x) + \int_0^T g'(s(x, t^*, y))s'_x(x, t^*, y)dy.$$

This function is increasing in x. By taking its value at $x_0 = \arg\min f(x)$

$$(\mathcal{C}_p(x_0, t^*))'_x = \int_0^T g'(s(x, t^*, y))s'_x(x, t^*, y)dy \geq 0.$$

Therefore, if equation 5.2 holds then there exists x^* such that $(\mathcal{C}_p(x^*, t^*))'_x = 0$.

(b) Now by taking a derivative with respect to t we have

$$(C_p(x^*, t^*))'_t = I'_t(x^*, t^*) - (1-p)(U+B)f'(x^*).$$

The derivative is decreasing with respect to p. We now substitute values $p = 0$ and $p = 1$ to get

$$(C_0(x^*, t^*))'_t = -(U+B)f'(x^*) > 0$$

$$(C_1(x^*, t^*))'_t = I'_t(x^*, t^*) = \int_0^T g'(s(x, t^*, y))s'_t(x, t^*, y) \leq 0.$$

Therefore, there exists p^* such that $(C^*_p(x^*, t^*))'_t = 0$.

(c) The statement follows immediately, as we have found a zero of the gradient of a strictly convex function.

4. Suppose that $|\hat{S}(nT) - S_{p^*}(nT)| \geq \varepsilon$ then let

$$\delta = \min(f(x^* - \varepsilon) - f(x^*), f(x^*) - f(x^* + \varepsilon)) = f(x^*) - f(x^* + \varepsilon),$$

since f is convex.

Denote by S_n the restriction of policy $\hat{S}$ on $[nT, (n+1)T]$, then

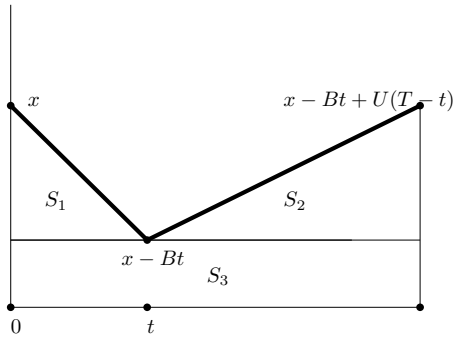
$$\begin{aligned}
\mathcal{C}(\hat{S}) &= \lim\inf_{n\to\infty} \frac{\sum_0^n C_{p^*}(S_n)}{n} \\
&\geq \lim\inf_{n\to\infty} \frac{\sum_0^n C_{p^*}(s(x^*, t^*)) + \delta\mathbb{I}(|S(nT) - x^*| \geq \varepsilon)}{n} \\
&\geq \mathcal{C}_{p^*}(x^*, t^*) + \delta \lim\inf_{n\to\infty} \frac{|\{n : |\hat{S}(nT) - S_{p^*}(nT)| \geq \varepsilon\}|}{n} \\
&> \mathcal{C}_{p^*}(x^*, t^*).
\end{aligned}$$

□

5.2.2 Numeric example

In this subsections we find an exact solution for the optimal control for the basic example. Recall that the terminal charge f is associated with the cost of purchasing of nuclear power to cover the shortage in power supply and the continuous charge corresponds to the additional level of conventional generation. It was discussed in Section 2.4 that it is reasonable to take $g(x) = c_1 x$ and $f(x) = c_2 \mathbb{E}(\chi - x)^+$ for this model.

We start by establishing formulas for area beneath function $s(x, t, y)$ representing the integral part of the cost functional. We explore different cases for the location of the plot of the function $s(x, t, y)$ and x-axis. They might intersect or may not. A graphic proof is presented in figures below and then the result is summarised in Table 5.1.



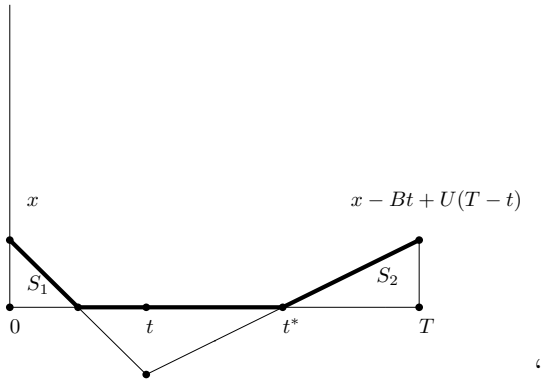


Table 5.1: Areas for functions $s(x, t, y)$

-	$x \geq Bt$	$x \leq Bt$
S_1	$\frac{Bt^2}{2}$	$\frac{x^2}{2B}$
S_2	$\frac{U(T-t)^2}{2}$	$\frac{(x+UT-(U+B)t)^2}{2U}$
S_3	$(x - Bt)T$	-

Recall a very useful representation

$$\mathbb{E}(\chi - x)^+ = \int_x^{+\infty} \bar{F}(y)dy.$$

Hence, the one period cost functional $C(x, t)$ can be written as

$$C(x,t) = c_2 \int_{x+UT-(U+B)t}^{+\infty} \bar{F}(y)dy + c_1 \begin{cases} \frac{Bt^2}{2} + \frac{U(T-t)^2}{2} + (x - (Bt)T) & \text{, for } x \geq Bt \\ \frac{x^2}{2B} + \frac{(x+UT-(U+B)t)^2}{2U} & \text{, otherwise.} \end{cases} \tag{5.3}$$

To find the optimal policy it is sufficient to find p^* and x^* satisfying the assumptions of Theorem 5.2.2.

Theorem 5.2.3. *Assume that* $p^* = \frac{U}{U+B}$ *and* $x^* \geq 0$ *satisfies the equation*

$$c_2\overline{F}(x) = c_1 \min\left(T, \frac{x(U+B)}{UB}\right). \tag{5.4}$$

Then cost functional $\mathcal{C}_{p^*}(x,t)$ *attains its minimum at point* (x^*, t^*) *and* $s(x^*, t^*, y)$ *(which was given by the equation 5.1) is the optimal policy for the time average cost functional.*

Proof. The function $\mathcal{C}_p(x,t)$ is differentiable and strictly convex as a function of x and t. To find its minimum it is sufficient to find a zero of its gradient.

Firstly, we take a derivative with respect to t:

$$\mathcal{C}_p(x,t)'_t = c_2(1-p)(U+B)\overline{F}(x+UT-(U+B)t)+c_1 \begin{cases} (U+B)(t-T), & x \geq Bt \\ \frac{(t(U+B)-UT-x)(U+B)}{U}, & x \leq Bt. \end{cases} \tag{5.5}$$

At point $(x, \frac{UT}{U+B})$ one gets:

$$\begin{aligned} \mathcal{C}\left(x, \frac{UT}{U+B}\right)'_t &= c_2(1-p)(U+B)\overline{F}(x) + c_1 \begin{cases} -BT & x \geq \frac{BUT}{U+B} \\ -\frac{x(U+B)}{U} & x \leq \frac{BUT}{U+B} \end{cases} \\ &= c_2(1-p)(U+B)\overline{F}(x) - c_1 \min\left(BT, \frac{x(U+B)}{U}\right). \end{aligned}$$

Secondly, by taking a derivative with respect to x one gets

$$\mathcal{C}\left(x, \frac{UT}{U+B}\right)'_x = -c_2\overline{F}(x) + c_1 \begin{cases} T & x \geq Bt \\ \frac{x(U+B)}{BU} & x \leq Bt \end{cases} \tag{5.6}$$

Hence, at point $(x, \frac{UT}{U+B})$ we have

$$\mathcal{C}_p\left(x, \frac{UT}{U+B}\right)'_x = -c_2\overline{F}(x) + c_1 \min\left(T, \frac{x(U+B)}{UB}\right).$$

If x^* is a zero of derivative $\mathcal{C}_p(x, \frac{UT}{U+B})'_x = 0$, when x^* satisfies equation 5.4. If we take $p^* = \frac{U}{U+B}$, then

$$\begin{aligned} \mathcal{C}_{p^*}\left(x^*, \frac{UT}{U+B}\right)'_t &= c_2(1-p)(U+B)\overline{F}(x) - c_1 \min\left(BT, \frac{x(U+B)}{U}\right) \\ &= -(U+B)(1-p)\mathcal{C}_{p^*}\left(x^*, \frac{UT}{U+B}\right)'_x = 0. \end{aligned}$$

Hence, we found p^* such that p^*-cost functional attains its minimum at point (x^*, t^*). So the policy corresponding to it is optimal. □

5.3 Blend of predictability and randomness

Suppose that $T_{i+1} = T_i + T + \chi_{i+1}$, where $\{\chi_i\}_{i\geq 0}$ are i.i.d. with increments distributed as $\exp(\lambda)$. We refer to $[T_i, T_i + T]$ as deterministic phase and to $[T_i + T, T_{i+1}]$ as exponential phase. Recall class S introduced in Section 1

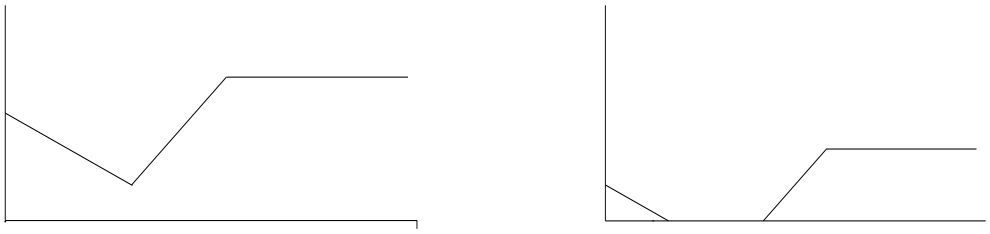
$$S = \{s(v) : s'(v) \text{ is increasing }, s'(v) \in \{-B, 0, U\} \text{ and } s'(v) = 0 \text{ only if } s(v) = 0\}.$$

In this section we restrict our search only to the policies such that the corresponding control process belongs to the class described below.

Definition 5.3.1.
A process X belongs to the class $\mathbb{S}$ if

- For any i holds $X|_{[T_i, T_i+T]} \in S$ and
- There exists a constant level $l \geq 0$ such that for all $i \geq 0$ and for all $s \in [T_i + T, T_{i+1}]$ holds $X(s) = k(X(T_i + T), l, s - T_i)$, where k is given by formula 2.9.

The figure below depicts controls of the class on a single period. Level l may or may not be reached at times $\{T_i\}_{i\geq 0}$.



Remark 5.3.1. The assumption to restrict our consideration to the class $\mathbb{S}$ is made in order to simplify our analysis. However, there are strong arguments to support the conjecture that the optimal policy belongs to the class $\mathbb{S}$. First, it was proven earlier in the Markovian case, where $T = 0$. Second, it will be justified in a few pages for sufficiently large values of T. Third, suppose that an analogue of ACOE 3.6 holds not only for Markovian environment, but also for the semi-Markovian setting with $T_{i+1} = T_i + T + \exp(\lambda)$. In our particular subcase there is no environment, so one may write

$$\lambda\rho + h(x) = \inf_{a\in A} \left[C(x, a) + \mathbb{E}_x h(X_a(T + \chi_1))\right], \tag{5.7}$$

where A is the class of feasible policies. Then there exists a convex relative value function h minimising both sides of the equation. Recall that $C(x, a)$ stands for the average single period cost for a process which starts at x and evolves according to

policy a. Since $T_{i+1} - T_i \sim T + \exp(\lambda)$ one may expand the expression for $C(x, a)$ as

$$C(x,a) = \int_0^T g(X_a(t))dt + \mathbb{E}\left(\lambda^{-1} g(X_a(T+\chi_1)) + f(X_a(T+\chi_1))\right).$$

Therefore, the right-hand side of ACOE can be rewritten as

$$\inf_{a\in A} \int_0^T g(X_a(t))dt + \left(\mathbb{E}_x \lambda^{-1} g(X_a(T+\chi_1)) + f(X_a(T+\chi_1)) + h(X_a(T+\chi_1))\right).$$

Function $H(y) := \lambda^{-1} g(y) + f(y) + h(y)$ is strictly convex and independent of the starting point x. Hence, due to Lemma 3.3.2, the optimal trajectory goes to level $\operatorname{argmin}_y H(y)$ as fast as possible, and stays there if the level is reached before the beginning of the new deterministic period.

Finally, we expect that the proof of the conjecture is lengthy, but routine.

Time-homogeneous policies such that the controlled processes corresponding to them belong to the class $\mathbb{S}$ can be parametrised in several ways. The most convenient parametrisation for us is the following. The parametrisation is given by function $z : \mathbb{R}^+ \to \mathbb{R}^+$ and the level l such that for any x_0 the policy starting at x_0 at time T_i is at level $z(x_0)$ at time $T_i + T$. We will work with subclass $\mathbb{S}_z$ of $\mathbb{S}$, which is introduced in the definition below. Its important properties are given following the definition.

Definition 5.3.2.
Class $\mathbb{S}_z$ contains all the processes of $\mathbb{S}$ for which $z(x) \equiv c$, where c is a positive constant.

There are multiple benefits of solving the problem in the class $\mathbb{S}_z$.

- Any policy of the subclass is parametrised by only two values (z, l). Hence, finding pseudo-optimal controls is an easier problem to solve.
- The pseudo-optimal controls are very easy to implement in real-life.
- For any policy of class $\mathbb{S}_z$, sequence $\{X(T_i + T)\}_{k\geq 0}$ is stationary.

The outline for the rest of this section is as follows. In Subsection 5.3.1 we find an exact solution for sufficiently large T. This solution belongs to the class $\mathbb{S}_z$. Hence, for this case the assumption regarding the existence of the unique level l for all time periods is justified. In Subsection 5.3.2 we explain how to solve the problem in class $\mathbb{S}_z$ even for smaller T. The problem reduces to a single period problem, which can be easily solved. Finally, in Subsection 5.3.3, we show that policies of the class $\mathbb{S}_z$ can only be optimal if the controlled process touches 0 level at every deterministic phase $[T_i, T_i + T]$.

5.3.1 Solution for large T

Intuition suggests that when T goes to infinity, the optimal process will stay at level 0 for long periods of time, leaving it only a short time prior to the beginning of an exponential cycle and then returning to 0 as quickly as possible. To solve the problem for large T we need a simple auxiliary lemma.

Lemma 5.3.2. *Suppose that G is a strictly convex function with its minimum at l, then for any $s > 0$ function $G(k(x, l, s))$ is a convex function of x.*

Proof. For a fixed s one has

$$k(x,l,s) = \begin{cases} x + Us & \text{, if } x \le l - Us \\ l & \text{, if } l - Us \le x \le l + Bs \\ x - Bs & \text{, if } x \ge l + Bs \end{cases}$$

Since convexity is local property and is preserved with respect to linear changes of variables, we only need to check the convexity of the superposition $G(k)$ at $l - Us$ and $l + Bs$. Suppose that $x_1 < x_2$ and p are such that $px_1 + (1-p)x_2 = l - Us$ and $x_1 < l - Us$ and $x_2 < l + Bs$. Then

$$f(x_1 + Us) > f(l),$$

since l is the function's argminimum and, hence, the convexity holds

$$f(l) = f(k(l - Us, l, s)) < p(f(x_1 + Us)) + qf(l).$$

□

In the following proposition we establish a lower bound for T such that the intuition described in the beginning of the subsection holds. For this T we find the optimal level l and prove the uniqueness of the optimal solution.

Proposition 5.3.3. *Suppose that $g(x) = x$. Let $L := \operatorname{argmin}_x f(x) + \lambda^{-1}x$. Then the following holds.*

1. *The optimal level $l \le L$.*
2. *For $T \ge \frac{L}{B} + \frac{L}{U}$ the optimal level is $l = \operatorname{argmin}_x \left(f(x) + \lambda^{-1}x + \frac{x^2}{2U} \right)$ and there exists a unique constant $t(x) = t$.*

Proof. 1. For any process $X(t)$ the process $X(t) = \min(X(t), L)$ gives a lower value for the single-period cost functional, and, hence for the average-cost functional as well.

2. Firstly we show that, for any process from the class $\mathbb{S}$ such that $X(0) \leq L$ and $X(T) \leq L$, its trajectory hits level 0 during time interval $[0, T]$. Suppose that the process is given by $(t(x), l)$ then if the process does not hit level 0 one has

$$X(T) = x - (U + B)t(x) + UT \geq UT \geq U\left(\frac{L}{B} + \frac{L}{U}\right) > L.$$

Now consider a slightly different optimisation problem. As before χ_1 is exponentially distributed with parameter λ. Consider the optimisation problem on the single interval $[-L/U, \chi + L/B]$. Clearly, the trajectory of the optimal process starts at 0 at unknown time $-t$ and goes towards unknown level l at speed U, then remains at level l, if the level is reached, before time χ_1. At time χ_1, the trajectory changes its direction and goes towards 0 at pace $-B$. As a function of t and l the cost functional can be written as

$$\widehat{C}(t, l) = \frac{Ut^2}{2} + \mathbb{E}\left(f(k(Ut, l, \chi_1)) + \lambda^{-1}k(Ut, l, \chi_1) + \frac{(k(Ut, l, \chi_1))^2}{2B}\right). \quad (5.8)$$

Function $f + \lambda^{-1}x + x^2/2B$ is strictly convex. Let $l^* = \operatorname{argmin}_x(f(x) + \lambda^{-1}x + x^2/2B)$. Hence, we can apply Lemma 3.3.2 to obtain that for any fixed t the minimum of the functional attains for process $X(s) = k(Ut, l^*, s)$. Function $\widehat{C}(t, l^*)$ is strictly convex with respect to t. This holds due to Lemma 5.3.2 and strict convexity of $\frac{Ut^2}{2}$. Therefore, there exists a unique t^*, such that (Ut^*, l) is the optimal policy for the average cost functional.

□

5.3.2 Optimisation in class $\mathbb{S}_z$

Recall that this class of policies is parametrised by the value $z = X(T_i + T)$ and the level l. The process should evolve optimally in $[T_i + T, T_{i+1} + T]$ so that the total cost on this period is the smallest, given that the border values are fixed. Hence, the average cost optimisation in class $\mathbb{S}_z$ can be replaced with the optimisation for a single-period problem for interval $[T_i + T, T_i + T + \chi_{i+1} + T]$ in the class of policies with $X(T_i + T) = X(T_i + T + \chi_{i+1} + T)$. Denote by $G(z_1, z_2)$ the minimal charges of a functional

$$\int_0^T X(s)ds, \text{ given } X(0) = z_1 \text{ and } X(T) = z_2.$$

Then, the charges on a single period are:

$$\mathbb{E}\int_T^{T+\chi_1} g(k(z, l, s))ds + f(k(z, l, T + \chi_1)) + \int_{T+\chi_1}^{T+\chi_1+T} G(k(z, l, T + \chi_1), z)ds. \quad (5.9)$$

We may find the distribution at time T_{i+1} as a function of (z, l).

Proposition 5.3.4. *For any $z < v < l$ holds*

$$\mathbb{P}(X_{T_{i+1}} < v | X_{T_i+T} = z) = 1 - \exp\left(-\lambda \frac{v-z}{U}\right)$$

and

$$\mathbb{P}(X_{T_{i+1}} = l | X_{T_i+T} = z) = \exp\left(-\lambda \frac{l-z}{U}\right).$$

Proof. One has

$$\mathbb{P}(X_{T_{i+1}} < v | X_{T_i+T} = z) = \mathbb{P}(x + U\chi_{i+1} < v | X_{T_i+T} = z)$$

$$= \mathbb{P}\left(\chi_{i+1} < \frac{v-z}{U}\right) = 1 - \exp\left(-\lambda \frac{v-z}{U}\right).$$

The second equality is obtained in a similar way. □

Take a random variable θ having distribution from the last proposition. Then the total cost on this single period is equal to

$$\mathbb{E}\left(f(\theta) + \lambda^{-1}\theta)\right) + \mathbb{E}G(\theta, z)ds.$$

This is a function of two variables (z, l) and its minimum may be found by using direct differentiation or by applying any of numeric methods.

5.3.3 Properties of the optimal solution

In this section we suppose that the version of ACOE holds for our problem and that there exists a convex relative value function h.

Lemma 5.3.5. *Consider the optimisation problem on an interval of length T with strictly convex terminal charge $H(x)$ and continuous charge x so that the cost functional is*

$$L(X) = \int_0^T X(s)ds + H(X(T)).$$

Suppose that $X(0) = x$ and find the optimal "turning point" $t(x)$. Then

1. *$t(x)$ is a non-decreasing function;*
2. *For the optimal solutions given by $(z(x), l)$, starting at x_1 and x_2 holds*

$$|z_2 - z_1| \leq |x_2 - x_1|.$$

3. *Suppose that for some $x < l$ the optimal trajectory does not touch 0, then $z(x)$ cannot be constant.*

Proof. 1. Suppose that for some $x = 0$ the optimal policy stays at 0 till time t. Then for all $x \geq Ut$ the optimal policy will start going upwards only at time t. So $t(x)$ is constant for $x \geq Ut$. Consider cost functional $L(X)$ as a function of (x, t). It was calculated in Table 5.1 that for policies not touching level 0 the cost functional is equal to

$$L(x,t) = \frac{Bt^2}{2} + \frac{U(T-t)^2}{2} + (x - Bt)T + H(x + UT - (U+B)t).$$

Turning point $t(x)$ is the point where the derivative $L'_t(x,t) = 0$. Note that $X(T) \leq \operatorname{argmin}_s H(s)$. One has for $x_1 < x_2$ and the same t

$$L'_t(x_1,t) = -(B+U)T + (B+U)t - (U+B)H'(x_1 + UT - (U+B)t) \quad (5.10)$$

$$> -(B+U)T + (B+U)t - (U+B)H'(x_2 + UT - (U+B)t),$$

because the derivative H' is an increasing function since H is convex. Then the derivative $L'_t(x,t)$ is an increasing function of t with x fixed. Therefore if $L'_t(x_1, t(x_1)) = 0$ then $L'_t(x_2, t(x_1)) < 0$ and hence for $t(x_2)$ we have $t(x_2) > t(x_1)$.

2. Suppose that $x_1 < x_2$ then as we have just shown $t(x_1) \leq t(x_2)$.

$$z_1 = x_1 + UT - (U+B)t(x_1)$$

$$z_2 = x_2 + UT - (U+B)t(x_2)$$

$$z_2 - z_1 = (x_2 - x_1) + (U+B)(t(x_1) - t(x_2)) \geq (x_2 - x_1).$$

The last inequality is strict if $t(x_1) < t(x_2)$.

3. Suppose that z is constant, then due to the last statement, for any $x \geq l$ the optimal policy touches zero.

□

To apply the last statement of the theorem in our case we should take $H(z)$ as the remaining part of the cost for the process crossing level z at moment T, plus the average of the relative value function h with respect to the value $X(T + \chi_1)$.

Appendix A

Proof of Theorem 3.6.2

Proof. The cost functional can be written as

$$H(a,b) = \frac{\int_a^b cx + |A-x| + |D-x|dx + c(a+b) + 2|A-a| + 2|D-b|}{b-2+2} =$$

$$= \frac{c(a+b)}{2} + \frac{\int_a^b |A-x| + |D-x|dx + 2|A-a| + 2|D-b|}{b-a+2}.$$

To open the absolute values we consider the following cases

1. $a \le b \le A \le D$,
2. $A \le a \le b \le D$,
3. $a \le A \le b \le D$,
4. $a \le A \le D \le b$,
5. $A \le a \le D \le b$,
6. $A \le D \le a \le b$.

First of all, notice that the optimal solution cannot be of types (4)-(6).

For a couple (a,b) of type (5) consider couple $a^* = A-(b-D)$ and $b^* = D-(a-A)$ of type (3). We can see that $b-a = b^*-a^*$ and

$$\frac{\int_a^b |A-x| + |D-x|dx + 2|A-a| + 2|D-b|}{b-a+2} =$$

$$= \frac{\int_{a^*}^{b^*} |A-x| + |D-x|dx + 2|A-a^*| + 2|D-b^*|}{b^*-a^*+2},$$

but $a^*+b^* \le a+b$. Therefore $H(a^*,b^*) \le H(a,b)$.

For couple (a,b) of the case (6) one takes couple $a^* = A-(b-D)$ and $b^* = A-(a-D)$ of type (1), which gives a smaller value of the functional H. Finally, for a couple (a,b) of type (4) can be replaced with couple $(a^*,b^*) = (a-\min(a,D-b), b-\min(a,D-b))$ of type (3).

We use notation M_i for the minimal value in the case (i).

1. In the case when $a \le b \le A \le D$ the cost functional is equal to

$$H(a,b) = \frac{c(a+b)}{2} + \frac{\int_a^b (A+D-2x)dx + 2(A+D) - 2(a+b)}{b-a+2} =$$

$$= \frac{c(a+b)}{2} + \frac{(b-a+2)(A+D) - (b+a)(b-a+2)}{b-a+2} =$$

$$= A + D + \left(\frac{c}{2} - 1\right)(a+b).$$

The minimum value of the expression depends on the sign of $c/2 - 1$.

- If $c < 2$ we get $a = b = A$ and $M_1 = (c-1)A + D$,
- If $c = 2$ then an arbitrary couple $0 \le a \le b \le A$ is optimal and $M_1 = A+D$,
- else if $c > 2$ then $a = b = 0$ and $M_1 = A + D$.

2. If $A \le a \le b \le D$ the cost functional is equal to

$$H(a,b) = \frac{c(a+b)}{2} + \frac{\int_a^b (D-A)dx + 2(D-A) - 2(b-a)}{b-a+2} =$$

$$= \frac{c(a+b)}{2} + \frac{(b-a+2)(D-A) - 2(b-a+2) + 4}{b-a+2} =$$

$$= D - A - 2 + \frac{c(a+b)}{2} + \frac{4}{b-a+2}.$$

Function $H(a,b)$ increases in a if $b-a$ is fixed, therefore $a = A$. Taking a derivative with respect to b one gets

$$\frac{\partial H}{\partial b} = \frac{c}{2} - \frac{4}{(b-A+2)^2}.$$

The derivative is an increasing function of b. At point $b = A$ the sign of the derivative coincides with the sign of $c-2$. Therefore, if $c \ge 2$, then the optimal $b = A$, and since pair (A, A) belongs to the case (1) we have $M_2 \ge M_1$. If $c = 2$ one can see that $M_1 = M_2$.

If $c < 2$ we need to look for the zero of the derivative which attains at $b = \sqrt{8/c} + A - 2)$. It yields that $b = \min(D, \sqrt{8/c} + A - 2)$. As we have shown before the optimal pair $a = b = A$ for the case (1) given $c < 2$ belongs to the case (2), hence $M_2 \le M_1$.

3. In the case (3) the cost functional is equal to

$$H(a,b) = \frac{c(a+b)}{2} + \frac{\int_a^A A + D - 2x dx + \int_A^b D - A dx + 2(D-b) + 2(A-a)}{b-a+2} =$$

$$= \frac{c(a+b)}{2} + \frac{(A+D)(A-a) - A^2 + a^2 + (D-A)(b-A) + 2(A+D) - 2(a+b)}{b-a+2} =$$

$$= D - A - 2 + \frac{c(a+b)}{2} + \frac{(A-a+2)^2}{b-a+2}.$$

The derivative with respect to b is

$$\frac{\partial H}{\partial b} = \frac{c}{2} - \frac{(A-a+2)^2}{(b-a+2)^2}.$$

If $c \geq 2$ then $\frac{\partial H}{\partial b} \geq 0$ and the function is everywhere increasing in b, thus optimal $b = A$. Hence, the optimal couple belongs to the case (1) and the minimum is not improved: $M_3 \geq M_1$.

If $c < 2$ then the derivative $\frac{\partial H}{\partial b}$ increases in a and b and $\frac{\partial H}{\partial b}(A, A) \leq 0$. Therefore, for each a there exists $b(a) = \sqrt{8/c} + A - 2 > A$ such that $\frac{\partial H}{\partial b}(a, b(a)) = 0$. Notice, that it is possible to have $b(a) > D$, which doesn't belong to the case (3). Now we would like to choose the best of the couples $(a, b(a))$. Taking a derivative with respect to a one gets

$$\frac{\partial H}{\partial a}(a, b(a)) = \frac{c}{2} + \frac{-2(A-a+2)(b(a)-a+2) + (A-a+2)^2}{(b(a)-a+2)^2} =$$

$$= \frac{(A-a+2)^2}{(b(a)-a+2)^2} - 1 + \frac{(A-a+2-(b(a)-a+2))^2}{(b(a)-a+2)^2} =$$

$$= \frac{(A-a+2)^2 + (b(a)-A)^2}{(b(a)-a+2)^2} - 1 =$$

$$= -\frac{2(A-a+2)(b(a)-A)}{(b(a)-a+2)^2} \leq 0.$$

Therefore, from the couples $(a, b(a))$ the biggest one would be the best. If $b(A) \leq D$ then couple $(A, b(A))$ is the best one because for $b > b^*(A)$ holds $\frac{\partial H}{\partial b}(A, b) > 0$.

If $b(A) > D$, let $\hat{a}$ be such that $b(\hat{a}) = D$. The optimal couple (a, b) satisfies $a \geq \hat{a}$ and $b = D$, since the derivative with respect to b is increasing when $b \leq b(\hat{a}) = D$. Finally,

$$\frac{\partial H}{\partial a}(a, D) = \frac{c}{2} + \frac{-2(A-a+2)(D-a+2) + (A-a+2)^2}{(D-a+2)^2} =$$

$$= \frac{c}{2} - 1 + \frac{(D-A)^2}{(D-a+2)^2}.$$

The derivative increases in a. Denote the zero of the derivative by a^*, then then for optimal $a = \min(a^*, A)$. Let $x^* = \frac{2}{(D-a+2)}$, then one can obtain the sign of $\frac{\partial H}{\partial a}(A, D)$ by finding a positive solution of the equation

$$\frac{c}{2} - 1 + (1-x)^2 = 0 \Longleftrightarrow x = \frac{1+\sqrt{4-2c}}{2}$$

Therefore, if $x^* \geq x$ then $a = A$, otherwise $a = a^*$.

To finish we list all possible variants for optimal couples (a, b).

- If $c \geq 2$ then $M_3 \geq M_1$ and the optimal solution belongs to case (1).

  otherwise for $c < 2$:

- if $\sqrt{8/c} + A - 2 \leq D$ then $(a, b) = (A, b(A))$ and $M_3 = M_2 \leq M_1$

  for $c < 2$ and $\sqrt{8/c} + A - 2 \geq D$ can be two possible scenarios:

- if $\frac{4}{1+\sqrt{4-2c}} \geq D - A + 2$ then $(a, b) = (A, D)$ and $M_3 = M_2 \leq M_1$,
- or if $\frac{4}{1+\sqrt{4-2c}} < D - A + 2$ then $(a, b) = (a^*, D)$ and $M_3 \leq M_2 \leq M_1$.

By checking all possible options we obtained that the global minimum is either M_1 if $c \geq 2$, or M_3 otherwise. □

Appendix B

Matlab code

```
% INSERT TRANSITION MATRIX
Q=[-2  1  1;   1 -2  1;  1  1  -2];
 Qh=[Q  [1;1;1]];
b=[0  0  0  1]';
pi=Qh'\b;
% INSERT TERMINAL CHARGES
A1=1;
A2=2;
A3=5;
cost1=@(x)  x + 2.*max(A1-x,0);
cost2=@(x)  x + 2.*max(A2-x,0);
cost3=@(x)  x + 2.*max(A3-x,0);
for  k=0.3:0.1:0.3
B=k;
U=k;

%-------------------------
%-----------------------------------------
m1=[-1/B  0  0;  0  1/U  0;  0  0  1/U];
m2=[-1/B  0  0;  0  -1/B  0;  0  0  1/U];
Q1=m1*Q';
Q2=m2*Q';
%-----------------------------------------
[ve1,va1]=eig(Q1);    % eigenvectors are columns
[ve2,va2]=eig(Q2);
%-----------------------------------------

Y11=@(x)  exp((va1(1,1))*x);
```

```
29 Y12=@(x) exp((va1(2,2))*x);
30 Y13=@(x) exp((va1(3,3))*x);
31
32 Y21=@(x) exp((va2(1,1))*x);
33 Y22=@(x) exp((va2(2,2))*x);
34 Y23=@(x) exp((va2(3,3))*x);
35
36 %-----------------------------------------
37 m=10000;
38 sol =[A1 A1 A1];
39 for l1=A1:10^(-1):A3
40     for l2=l1:10^(-1):A3
41         for l3=l2:10^(-1):A3
42         l=[l1 l2 l3];
43         bot=[ve1(:,1).*Y11(l1) ve1(:,2).*Y12(l1) ve1(:,3).*
               Y13(l1)];
44 top=[ve2(:,1).*Y21(l3) ve2(:,2).*Y22(l3) ve2(:,3).*Y23(l3)];
45 mid1=[ve1(:,1).*Y11(l2) ve1(:,2).*Y12(l2) ve1(:,3).*Y13(l2)];
     % BC's
46 mid2=[ve2(:,1).*Y21(l2) ve2(:,2).*Y22(l2) ve2(:,3).*Y23(l2)];
47 %-----------------------------------------
48 six=[bot zeros(3,3); zeros(3,3) top];
49 col1=zeros(5,1);
50 eight_six=[six [-1; col1] [col1;-1]];
51 bc1=[ mid1(1,:); mid1(3,:)];
52 bc2=[ mid2(1,:); mid2(3,:)];
53 eight=[eight_six; bc1 -bc2 zeros(2,2)];
54
55 y=[0 0 0 pi(1) pi(2) 0 0 0]';
56
57 c=eight\y;
58 % c11=c(1), c12=c(2), c13=c(3), c21=c(4), c22=c(5), c23=c(6),
     p1=c(7)
59 % q3=pi(3)-p3=c(8)=>
60 p1=c(7);
61 p3=pi(3)-c(8);
62 p2=-(mid1(2,1)*c(1)+mid1(2,2)*c(2)+mid1(2,3)*c(3)-mid2(2,1)*c
     (4)-mid2(2,2)*c(5)-mid2(2,3)*c(6));
63
```

```
64 F11=@(x) c(1)*ve1(1,1)*Y11(x)+c(2)*ve1(1,2)*Y12(x)+c(3)*ve1
       (1,3)*Y13(x);
65 F12=@(x) c(1)*ve1(2,1)*Y11(x)+c(2)*ve1(2,2)*Y12(x)+c(3)*ve1
       (2,3)*Y13(x);
66 F13=@(x) c(1)*ve1(3,1)*Y11(x)+c(2)*ve1(3,2)*Y12(x)+c(3)*ve1
       (3,3)*Y13(x);
67 F21=@(x) c(4)*ve2(1,1)*Y21(x)+c(5)*ve2(1,2)*Y22(x)+c(6)*ve2
       (1,3)*Y23(x);
68 F22=@(x) c(4)*ve2(2,1)*Y21(x)+c(5)*ve2(2,2)*Y22(x)+c(6)*ve2
       (2,3)*Y23(x);
69 F23=@(x) c(4)*ve2(3,1)*Y21(x)+c(5)*ve2(3,2)*Y22(x)+c(6)*ve2
       (3,3)*Y23(x);
70 %-------------------------------------------
71 G11=@(x) c(1)*ve1(1,1)*va1(1,1)*Y11(x)+c(2)*ve1(1,2)*va1(2,2)
       *Y12(x)+c(3)*ve1(1,3)*va1(3,3)*Y13(x);
72 G12=@(x) c(1)*ve1(2,1)*va1(1,1)*Y11(x)+c(2)*ve1(2,2)*va1(2,2)
       *Y12(x)+c(3)*ve1(2,3)*va1(3,3)*Y13(x);
73 G13=@(x) c(1)*ve1(3,1)*va1(1,1)*Y11(x)+c(2)*ve1(3,2)*va1(2,2)
       *Y12(x)+c(3)*ve1(3,3)*va1(3,3)*Y13(x);
74 G21=@(x) c(4)*ve2(1,1)*va2(1,1)*Y21(x)+c(5)*ve2(1,2)*va2(2,2)
       *Y22(x)+c(6)*ve2(1,3)*va2(3,3)*Y23(x);
75 G22=@(x) c(4)*ve2(2,1)*va2(1,1)*Y21(x)+c(5)*ve2(2,2)*va2(2,2)
       *Y22(x)+c(6)*ve2(2,3)*va2(3,3)*Y23(x);
76 G23=@(x) c(4)*ve2(3,1)*va2(1,1)*Y21(x)+c(5)*ve2(3,2)*va2(2,2)
       *Y22(x)+c(6)*ve2(3,3)*va2(3,3)*Y23(x);
77 %-------------------------------------------
78
79
80 %x=l1:10^(-6):l3;
81 F1=@(x) (x>l2).*(F21(x)-F11(x))+ F11(x);
82 F2=@(x) (x>l2).*(F22(x)-F12(x))+ F12(x);
83 F3=@(x) (x>l2).*(F23(x)-F13(x))+ F13(x);
84
85 %-------------------------------
86 part11=@(x) cost1(x).*G11(x);
87 part12=@(x) cost2(x).*G12(x);
88 part13=@(x) cost3(x).*G13(x);
89 part21=@(x) cost1(x).*G21(x);
90 part22=@(x) cost2(x).*G22(x);
```

```
91  part23=@(x) cost3(x).*G23(x);
92  %-------------------------------
93  % plot(x,part11(x))
94  i11=integral(part11, l1,l2);
95  i12=integral(part12, l1,l2);
96  i13=integral(part13, l1,l2);
97  i21=integral(part21, l2,l3);
98  i22=integral(part22, l2,l3);
99  i23=integral(part23, l2,l3);
100 %-------------------------------
101 res=i11+i12+i13+i21+i22+i23+p1.*cost1(l1) +p2.*cost2(l2)+p3.*
       cost3(l3);
102 %str = 'The value of the current iteration is:\n ';
103 %fprintf(str)
104 if res < m
105     m=res;
106     sol = [l1 l2 l3];
107     %fprintf('%d, %d, %d, %d, \n' ,l1, l2, l3, m)
108 end
109         end
110
111     end
112 end
113 %str = 'The value for the current B is: ';
114 fprintf('%d, %d, %d, %d, \n', k, sol(1), sol(2), sol(3))
115 end
```

Bibliography

[1] Anick, D., Mitra, D., and Sondhi, M. Stochastic theory of a data-handling system with multiple sources. *Bell System Technical Journal, The 61*, 8 (1982), 1871–1894.

[2] Arapostathis, A., Borkar, V. S., Fernández-Gaucherand, E., Ghosh, M. K., and Marcus, S. I. Discrete-time controlled markov processes with average cost criterion: a survey. *SIAM Journal on Control and Optimization 31*, 2 (1993), 282–344.

[3] Asmussen, S. *Applied probability and queues*, vol. 51. Springer Science & Business Media, 2008.

[4] Atkinson, K. E. *An introduction to numerical analysis*. John Wiley & Sons, 2008.

[5] Barbot, N. *Files d'attente fluides en environnement markovien*. PhD thesis, Rennes 1, 2002.

[6] Bäuerle, N. Convex stochastic fluid programs with average cost. *Journal of mathematical analysis and applications 259*, 1 (2001), 137–156.

[7] Bäuerle, N. Discounted stochastic fluid programs. *Mathematics of Operations Research 26*, 2 (2001), 401–420.

[8] Bean, N. G., and OReilly, M. M. Performance measures of a multi-layer markovian fluid model. *Annals of Operations Research 160*, 1 (2008), 99–120.

[9] Bellman, R. Dynamic programming and lagrange multipliers. *Proceedings of the National Academy of Sciences 42*, 10 (1956), 767–769.

[10] Bellman, R. A markovian decision process. Tech. rep., DTIC Document, 1957.

[11] Bertsekas, D. P. *Dynamic programming and optimal control*, vol. 1. Athena Scientific Belmont, MA, 1995.

[12] Bertsekas, D. P. Nonlinear programming.

[13] Bertsekas, D. P., Shreve, S. E., and Scientific, A. Stochastic optimal central: T he discretetime case.

[14] Billingsley, P. *Convergence of probability measures*. John Wiley & Sons, 2013.

[15] Blackwell, D. Discounted dynamic programming. *The Annals of Mathematical Statistics 36*, 1 (1965), 226–235.

[16] Borovkov, A. A., and Foss, S. G. Stochastically recursive sequences and their generalizations. *Siberian Advances in Mathematics 2*, 1 (1992), 16–81.

[17] Brémaud, P., Kannurpatti, R., and Mazumdar, R. Event and time averages: a review. *Advances in applied probability* (1992), 377–411.

[18] Chernysh, K. Average cost optimisation for a power supply management model. *Modern Trends in Controlled Stochastic Processes,Volume II* (2015), 307–323.

[19] Cruise, J., Flatley, L., Gibbens, R., and Zachary, S. Optimal control of storage incorporating market impact and with energy applications. *arXiv preprint arXiv:1406.3653* (2014).

[20] da Silva Soares, A., and Latouche, G. Fluid queues: building upon the analogy with qbd processes.

[21] da Silva Soares, A., and Latouche, G. Matrix-analytic methods for fluid queues with finite buffers. *Performance Evaluation 63*, 4 (2006), 295–314.

[22] da Silva Soares, A., and Latouche, G. Fluid queues with level dependent evolution. *European Journal of Operational Research 196*, 3 (2009), 1041–1048.

[23] Davis, M. H. Piecewise-deterministic markov processes: A general class of non-diffusion stochastic models. *Journal of the Royal Statistical Society. Series B (Methodological)* (1984), 353–388.

[24] Davis, M. H. *Markov Models & Optimization*, vol. 49. CRC Press, 1993.

[25] do Valle Costa, O. L., and Dufour, F. *Continuous average control of piecewise deterministic Markov processes*. Springer, 2013.

[26] Doshi, B. T. Continuous time control of markov processes on an arbitrary state space: average return criterion. *Stochastic Processes and their Applications 4*, 1 (1976), 55–77.

[27] Dragomirescu, F., and Ivan, C. The smallest convex extensions of a convex function. *Optimization 24*, 3-4 (1992), 193–206.

[28] Evans, L. C. Partial differential equations. *Graduate Studies in Mathematics 19* (1998).

[29] Fernández-Gaucherand, E., Arapostathis, A., and Marcus, S. I. Convex stochastic control problems. In *Decision and Control, 1992., Proceedings of the 31st IEEE Conference on* (1992), IEEE, pp. 2179–2180.

[30] Forwick, L., Schäl, M., and Schmitz, M. Piecewise deterministic markov control processes with feedback controls and unbounded costs. *Acta Applicandae Mathematica 82*, 3 (2004), 239–267.

[31] Foss, S., Korshunov, D., and Zachary, S. *An introduction to heavy-tailed and subexponential distributions*. Springer, 2011.

[32] Guo, X., and Rieder, U. Average optimality for continuous-time markov decision processes in polish spaces. *The Annals of Applied Probability* (2006), 730–756.

[33] Harsha, P., and Dahleh, M. Optimal management and sizing of energy storage under dynamic pricing for the efficient integration of renewable energy. *Power Systems, IEEE Transactions on 30*, 3 (2015), 1164–1181.

[34] Hordijk, A., and Schouten, F. A. V. D. D. Average optimal policies in markov decision drift processes with applications to a queueing and a replacement model. *Advances in Applied Probability* (1983), 274–303.

[35] Howard, R. A. Dynamic programming. *Management Science 12*, 5 (1966), 317–348.

[36] Keilson, J. *Markov chain modelsrarity and exponentiality*, vol. 28. Springer Science & Business Media, 2012.

[37] Kennedy, D. P. The continuity of the single server queue. *Journal of Applied Probability* (1972), 370–381.

[38] König, D., and Schmidt, V. Extended and conditional versions of the pasta property. *Advances in Applied Probability* (1990), 510–512.

[39] Kulkarni, V. G. Fluid models for single buffer systems. *Frontiers in queueing: Models and applications in science and engineering* (1997), 321–338.

[40] Lindvall, T. *Lectures on the coupling method*. Courier Corporation, 2002.

[41] Meyn, S., Barooah, P., Busic, A., and Ehren, J. Ancillary service to the grid from deferrable loads: the case for intelligent pool pumps in florida. In *Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on* (2013), IEEE, pp. 6946–6953.

[42] Mitra, D. Stochastic theory of a fluid model of producers and consumers coupled by a buffer. *Advances in Applied Probability* (1988), 646–676.

[43] Papavasiliou, A., and Oren, S. S. Multiarea stochastic unit commitment for high wind penetration in a transmission constrained network. *Operations Research 61*, 3 (2013), 578–592.

[44] Piunovskiy, A. B. *Modern Trends in Controlled Stochastic Processes*. Luniver Press, 2010.

[45] Piunovskiy, A. B. *Modern Trends in Controlled Stochastic Processes,Volume II*. Luniver Press, 2015.

[46] Puterman, M. L. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[47] Rachev, S. The problem of stability in queueing theory. *Queueing Systems 4*, 4 (1989), 287–317.

[48] Richmond, N., Jacko, P., and Makowski, A. M. Optimal planning of slow-ramping power production in energy systems with renewables forecasts and limited storage. In *Probabilistic Methods Applied to Power Systems (PMAPS), 2014 International Conference on* (2014), IEEE, pp. 1–6.

[49] Robinson, J. C. *An introduction to ordinary differential equations*. Cambridge University Press, 2004.

[50] Rockafellar, R. T. *Convex analysis*. Princeton university press, 2015.

[51] Ross, D. W., and Kim, S. Dynamic economic dispatch of generation. *Power Apparatus and Systems, IEEE Transactions on*, 6 (1980), 2060–2068.

[52] Ruiz, P. A., Philbrick, C. R., and Sauer, P. W. Wind power day-ahead uncertainty management through stochastic unit commitment policies. In *Power Systems Conference and Exposition, 2009. PSCE'09. IEEE/PES* (2009), IEEE, pp. 1–9.

[53] Sennott, L. I. Average cost optimal stationary policies in infinite state markov decision processes with unbounded costs. *Operations Research 37*, 4 (1989), 626–633.

[54] Sethi, S., Suo, W., Taksar, M. I., and Zhang, Q. Optimal production planning in a stochastic manufacturing system with long-run average cost. *Journal of Optimization Theory and Applications 92*, 1 (1997), 161–188.

[55] Shapiro, A., Dentcheva, D., et al. *Lectures on stochastic programming: modeling and theory*, vol. 16. SIAM, 2014.

[56] Takriti, S., Birge, J. R., and Long, E. A stochastic model for the unit commitment problem. *Power Systems, IEEE Transactions on 11*, 3 (1996), 1497–1508.

[57] Van Doorn, E. A., and Regterschot, G. Conditional pasta. *Operations Research Letters 7*, 5 (1988), 229–232.

[58] Verms, D. Optimal control of piecewise deterministic markov process. *Stochastics: An International Journal of Probability and Stochastic Processes 14*, 3 (1985), 165–207.

[59] Virtamo, J., and Norros, I. Fluid queue driven by anm/m/1 queue. *Queueing Systems 16*, 3-4 (1994), 373–386.

[60] Whitt, W. The continuity of queues. *Advances in Applied Probability* (1974), 175–183.

[61] Whitt, W. Continuity of generalized semi-markov processes. *Mathematics of operations research 5*, 4 (1980), 494–501.

[62] Widder, D. V. The laplace transform. 1946. *Zbl0139 29504* (1959).

[63] Wolff, R. W. Poisson arrivals see time averages. *Operations Research 30*, 2 (1982), 223–231.

[64] Wolff, R. W. A note on pasta and anti-pasta for continuous-time markov chains. *Operations research 38*, 1 (1990), 176–177.

[65] Wood, A. J., and Wollenberg, B. F. *Power generation, operation, and control*. John Wiley & Sons, 2012.

[66] Xia, X., and Elaiw, A. Optimal dynamic economic dispatch of generation: a review. *Electric Power Systems Research 80*, 8 (2010), 975–986.

[67] Yan, M. Extension of convex function. *arXiv preprint arXiv:1207.0944* (2012).