SJÄLVSTÄNDIGA ARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

A few inverse and optimal control problems stemming from Torricelli's law

av

Kristoffer Vinell

2009 - No 2

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 10691 STOCKHOLM

A few inverse and optimal control problems stemming from Torricelli's law

Kristoffer Vinell

Självständigt arbete i matematik 15 högskolepoäng, grundnivå

Handledare: Martin Tamm

2009

Abstract

Torricelli's law is often mentioned in the literature at the undergraduate level. Because of its generality and elegance, it can be applied to a wide range of outflow problems. However, in most contexts it is used only to describe the flow out of containers with constant cross-sectional area. In this thesis, the applications of Torricelli's law are extended to hold for almost any container of interest from a physical viewpoint.

In particular, we derive an ordinary differential equation that describes the flow from arbitrary containers. The ODE is applied to a few inverse problems, one of which involves finding the proper design for a water clock. In addition, we derive and solve an integral equation concerning flow rates by making use of the Laplace transform.

The ODE is also applied to a few optimal control problems, where the objective is to find minimal emptying times. We use Pontryagin's minimum principle to find necessary conditions for optimality, and prove that an optimal control is of the bang-bang type.

Acknowledgements

In retrospect, I can safely conclude that a lot of the work concerning the thesis was never carried out in a conventional way. At first, the problems to be solved were not known at hand. There was nothing more substantial than an idea, or a desire if you wish, to write about the generalization of a simple differential equation that I had found in a textbook. I would therefore like to thank my supervisor Martin Tamm for taking interest in my work at an early stage and for the encouragement along the way. I would also like to thank Yishao Zhou for valuable remarks and support throughout my work concerning optimal control theory.

Contents

1 Introduction

1.1 Inverse problems

Mathematics students at undergraduate level are frequently faced with *direct problems* as part of the course work. These problems are commonly designed to provide just enough information for the student to carry out a well defined process yielding a unique solution. Unfortunately, the situation is often far more complex in reality.

Any direct problem may be represented by an operator K modelling some process of interest. In control theory, this operator is called an *input-output mapping*. Along with the operator there is an input v in the domain of K with a corresponding output w. This relationship is depicted in Figure 1. Because K is a function, each input v in the domain of K yields a unique output w. Moreover, if K is continuous, as will frequently be the case in this thesis, then the output w depends continuously on v.

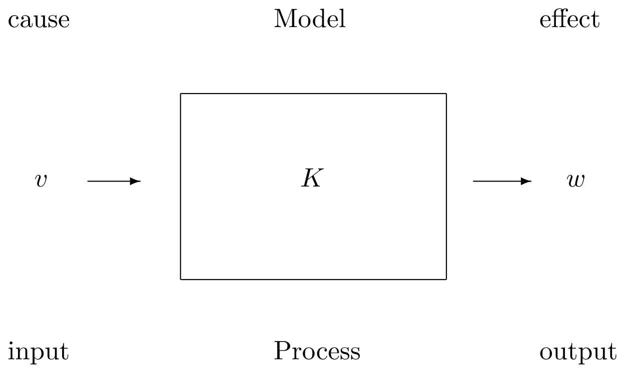

Figure 1: A direct problem

To every direct problem of this kind, two corresponding *inverse problems* arise naturally. One of them, known as the *causation* problem, consists of determining the cause v given both the effect w and the model K. An example of this kind that comes to mind is the problem of solving a linear system

$$Kv = w$$

for v, given that K is a finite dimensional linear operator. The other inverse problem, the *model identification* problem, involves determining the operator K given cause-effect pairs (v, w). In this context, some restrictions are generally imposed on the model, so as to allow only models from a given class of models.

In the study of inverse problems, two issues are of particular interest. The most immediate is the question of *existence*. For the causation problem, the question is whether there is a cause v (of a given class) that produces an observed effect w, for a certain model K. For the model identification problem the question of interest would be: given a number of cause-effect pairs (v, w), is there a model (in a given class) that can account for them? *Uniqueness* is also of importance. Is it possible that two distinct causes produce the same effect? Might different models account for the same cause-effect relationships?

1.2 Torricelli's law

A great deal of books on calculus at undergraduate level mention Torricelli's law at some point. Although this law arises mainly in fluid dynamics, there are some aspects of it that should be of interest to any student in mathematics. Beside the fact that it involves solving problems of physical nature, it is a good example of how to use the method of separating variables in solving ordinary differential equations (ODEs). Therefore, any mathematician should have come across it at least once or twice.

Evangelista Torricelli was an Italian physicist and mathematician during the 17th century. Although his work has had a great impact on a wide variety of fields in science, in this thesis we will be concerned with his achievements in fluid dynamics.

In 1643, Torricelli presented his main theorem concerning fluid flow, most commonly referred to simply as Torricelli's law. This theorem provides an elegant formula for the speed of a fluid flowing out of a tank through an opening under the force of gravity. The derivation of the formula requires only basic knowledge in physics. In fact, the simplicity of the theory behind the result gives us the opportunity to present the necessary ideas in a straight-forward manner.

Let us consider a tank of a given volume and height H initially filled with fluid. Suppose a small hole is made in the bottom of the tank, causing an outgoing flow. Consider a small amount of fluid with mass m flowing through the hole. The fluid leaving the tank will cause a loss of mgH in potential energy in the tank (g being the acceleration due to gravity). By the conservation law of energy, this amount of energy is converted into kinetic energy of the fluid leaving the tank. Hence, we have the relation

$$\frac{1}{2}mv^2 = mgH,$$

or equivalently,

$$v = \sqrt{2gH}, \tag{1.1}$$

where v denotes the speed of the fluid flow.

This is Torricelli's law in its natural form, which will be used in a number of inverse problems concerning fluid flow. However, although expression (1.1) holds in principle, the actual speed of the flow will be slightly less in practice, depending on which fluid we are concerned with and how the flow is affected by friction. In reality, the flow speed will actually be $v = \kappa\sqrt{2gH}$, for some constant κ, constrained by $0 < \kappa < 1$, depending on the situation. However, these issues are beyond the scope of this thesis and will not be dealt with. For generality, we will hence simply conclude that the flow speed is proportional to the square root of the fluid level.

1.3 Motivation

As noted earlier, Torricelli's law is often mentioned in literature at undergraduate level. However, in most contexts, the law is applied exclusively to efflux problems where the observed tank has constant cross-sectional area. This particular case is obviously of interest for various reasons, but constitutes only a fraction of the applications made possible by the generality of Torricelli's law.

Say for instance that we are interested in constructing a water clock, where it is desirable to achieve a constant decrease of the water level in the tank. As we shall see in the sequel, a tank with constant cross-sectional area will not do. Hence, we must construct a tank, with a more complex shape, that produces the desired flow.

In some cases, a certain behaviour of the flow is desired, such as in the water clock problem. To achieve this, the option is either to alter the area of the cross-section or that of the orifice; we are faced with an inverse problem. In other cases, it is desirable to steer a process towards an (in some sense) optimal behaviour. In this thesis, we will explore problems from both of these areas, all stemming from Torricelli's law.

1.4 Disposition

The thesis follows an application-driven approach. We begin with covering some important properties of the Laplace transform, which is in many cases an essential tool for solving differential and integral equations. Next, a brief introduction to control theory is presented. This is followed by a selection of important results from optimal control theory with emphasis on the minimum principle.

In the second part of the thesis a few inverse problems concerning flow rates are presented, one of which involves finding a proper design for a water clock. Both the method of Laplace transformation as well as concepts from control theory are applied here. Finally, we use optimal control theory to obtain a minimal depletion time, i.e. emptying time, among all tanks that are of interest from a physical viewpoint.

2 Laplace transforms

The use of Laplace transforms is an essential tool for solving differential as well as integral equations in many cases. In addition to several nice properties, the key feature of the Laplace transform lies in the fact that it makes it possible to transform linear differential and some integral equations into algebraic ones. This will prove to be crucial later on, when we establish a solution to a classical integral equation arising in an inverse problem stemming from Torricelli's law. For this purpose, it is necessary to provide some theory and results about the Laplace transform.

2.1 Definition and existence

The Laplace transform 2.1.1. The Laplace transform of a function $f(x)$, defined on $x \in [0, \infty)$ is the function

$$\mathscr{L}\left[f(x)\right] = \int_0^\infty e^{-sx} f(x)\, dx = F(s).$$

Thus, the operator $\mathscr{L}$ acts on any function $f(x)$ for which the integral exists, and produces the Laplace transform depending on the parameter s. However, it is not clear at first glance when this integral exists. Recall that the integral

$$\int_0^\infty f(x)\, dx$$

is defined as

$$\int_0^\infty f(x)\, dx = \lim_{b\to\infty} \int_0^b f(x)\, dx.$$

It is therefore reasonable to say that $f(x)$ has a Laplace transform whenever the limit

$$\lim_{b\to\infty} \int_0^b e^{-sx} f(x)\, dx$$

exists. Moreover, if we assume that $f(x)$ is piecewise continuous, then the existence (or non-existence) of its Laplace transform depends on the behaviour of $e^{-sx} f(x)$ for large x. To guarantee that the integrand diminishes rapidly enough for convergence, we shall further assume that $f(x)$ is of *exponential order*. This means that there are real positive constants M and c such that

$$|f(x)| \leq M e^{cx}. \tag{2.1}$$

Thus, if $f(x)$ satisfies (2.1), then we have

$$\left|e^{-sx} f(x)\right| \leq M e^{-(s-c)x}.$$

In addition, we note that

$$\begin{aligned} \int_0^\infty \left|e^{-sx} f(x)\right| dx &\leq M \int_0^\infty e^{-(s-c)x}\, dx \\ &= \frac{M}{s-c}, \end{aligned}$$

provided that $s > c$. Hence, the Laplace transform of $f(x)$ *converges absolutely* for all $s > c$, which implies that the Laplace transform converges itself. Under the conditions that $f(x)$ is piecewise continuous and of exponential order, we can therefore guarantee that $f(x)$ has a Laplace transform. However, it should be pointed out that these are sufficient conditions, not necessary. Take for example $f(x) = x^{-1/2}$, which is discontinuous at $x = 0$ but nevertheless has a Laplace transform, as will be shown later on.

2.2 Linearity, derivatives and integrals

As mentioned earlier, the Laplace transform is especially suited for dealing with linear differential as well as some integral equations. To see why this is so, consider the following simple consequences of the above definition concerning linearity, derivatives and integrals:

$$\begin{aligned}\mathscr{L}\left[\alpha f(x) + \beta g(x)\right] &= \int_0^\infty e^{-sx}\left(\alpha f(x) + \beta g(x)\right)\,dx \\ &= \alpha\int_0^\infty e^{-sx} f(x)\,dx + \beta\int_0^\infty e^{-sx} g(x)\,dx \\ &= \alpha\mathscr{L}\left[f(x)\right] + \beta\mathscr{L}\left[g(x)\right],\end{aligned}$$

for all constants α and β. Hence, the Laplace transform is a *linear* operator. Moreover,

$$\begin{aligned}\mathscr{L}\left[f'(x)\right] &= \int_0^\infty e^{-sx} f'(x)\,dx \\ &= \left[e^{-sx} f(x)\right]_0^\infty - \int_0^\infty (-s)e^{-sx} f(x)\,dx \\ &= -f(0) + s\int_0^\infty e^{-sx} f(x)\,dx \\ &= s\cdot\mathscr{L}\left[f(x)\right] - f(0)\end{aligned}$$

by use of partial integration, and so,

$$\mathscr{L}\left[f'(x)\right] = s\cdot\mathscr{L}\left[f(x)\right] - f(0). \tag{2.2}$$

Next, let

$$f(x) = \int_0^x g(t)\,dt. \tag{2.3}$$

Then by the fundamental theorem of calculus,

$$f'(x) = g(x),$$

and so from (2.2) follows that $\mathscr{L}\left[f'(x)\right] = s\cdot\mathscr{L}\left[f(x)\right] - f(0)$. But, clearly $f(0) = 0$ according to (2.3), and thus, $\mathscr{L}\left[f(x)\right] = \frac{1}{s}\mathscr{L}\left[f'(x)\right]$, which shows that

$$\mathscr{L}\left[\int_0^x g(t)\,dt\right] = \frac{1}{s}\mathscr{L}\left[g(x)\right].$$

2.3 The inverse Laplace transform and convolutions

In solving linear differential equations with the Laplace transform, the idea is to apply it to both sides of the equation of interest to obtain a purely algebraic equation. One can then express the Laplace transform of the unknown function in terms of the Laplace transforms of the known elements of the original equation. This procedure results in an equation of the form

$$\mathscr{L}[f(x)] = F(s),$$

where $f(x)$ is the unknown function. When $f(x)$ is continuous, the above equation is often written in the form

$$\mathscr{L}^{-1}[F(s)] = f(x).$$

It is customary to call $\mathscr{L}^{-1}$ the *inverse Laplace transformation*, and to refer to $f(x)$ as the *inverse Laplace transform* of $F(s)$. Hence, the inverse is used in the last stage of solving the equation, yielding an expression of the unknown function $f(x)$ explicitly. This is exactly the procedure carried out when solving an inverse problem including a differential or integral equation.

Having introduced the concept of the inverse, we are prepared to investigate the following problem: if $\mathscr{L}[f(x)] = F(s)$ and $\mathscr{L}[g(x)] = G(s)$, what is the inverse transform of $F(s)G(s)$?

To give an extensive answer to this question, we simply return to the definition, using dummy variables u and v in the integrals defining the transforms to obtain

$$\begin{aligned} F(s)G(s) &= \left[\int_0^\infty e^{-su} f(u)\,du\right]\left[\int_0^\infty e^{-sv} g(v)\,dv\right] \\ &= \int_0^\infty\int_0^\infty e^{-s(u+v)} f(u)g(v)\,dudv \\ &= \int_0^\infty \left[\int_0^\infty e^{-s(u+v)} f(u)\,du\right] g(v)\,dv, \end{aligned}$$

where the domain of integration is the first quadrant ($u \geq 0$, $v \geq 0$) in the uv-plane. We now introduce a new variable x in the inner integral of the last expression by putting $x = u + v$, so that $u = x - v$ and $du = dx$. This suggests that

$$\begin{aligned} F(s)G(s) &= \int_0^\infty \left[\int_v^\infty e^{-sx} f(x-v)\,dx\right] g(v)\,dv \\ &= \int_0^\infty\int_v^\infty e^{-sx} f(x-v)g(v)\,dxdv. \end{aligned}$$

This integration is extended over the first half of the first quadrant ($x \geq v$) in the xv-plane. By reversing the order of integration in the last expression, we get

$$
\begin{aligned}
F(s)G(s) &= \int_0^\infty \left[\int_0^x e^{-sx} f(x-v)g(v)\,dv\right] dx \\
&= \int_0^\infty e^{-sx} \left[\int_0^x f(x-v)g(v)\,dv\right] dx \\
&= \mathscr{L}\left[\int_0^x f(x-t)g(t)\,dt\right].
\end{aligned}
$$

The integral in the last expression is commonly referred to as the *convolution* of the functions $f(x)$ and $g(x)$. It can be regarded as a way of letting the function $f(x)$ "blend into" the function $g(x)$. The immediate consequence of the above equation is the important fact that

$$
\mathscr{L}[f(x)]\,\mathscr{L}[g(x)] = \mathscr{L}\left[\int_0^x f(x-t)g(t)\,dt\right], \tag{2.4}
$$

which is called the *convolution theorem* in the literature. It also follows that the convolution of two functions is commutative,

$$
f * g = g * f,
$$

by setting $y = x - t$ in the integral in (2.4). As we shall see, this result will serve as the ideal tool for solving certain types of integral equations, common in applications.

3 Optimal control theory

3.1 A few words on control theory

Mathematical control theory is, roughly speaking, the area of applied mathematics that deals with analysis and design of control systems. Control systems arise frequently in engineering applications, which rely on control theory. This theory is therefore incorporated in numerous modern devices, such as CD players, aircraft autopilots and robots, to name a few.

To *control* an object means to steer its behaviour to a desired state. For example, an airline pilot relies on the autopilot to follow a desired route. To achieve this, the autopilot must control several parts of the aircraft, including the engines and wings. Obviously, the change of route can be achieved in many ways, and some are more preferable than others. For instance, if maximum acceleration is applied at some instant, the flight may not be satisfactory to the passengers. There are clearly also physical limitations to the control, e.g. that the engines have a certain capacity. In reality, both of these aspects must be taken into account in finding a suitable control.

A further, but somewhat similar, example that illustrates these ideas is that of landing a spacecraft on a planetary surface. Assume that the spacecraft is at an altitude h from the surface at time $t = 0$, with initial downward velocity v. For simplicity we assume that gravitational forces can be disregarded and that we only consider vertical motion, upwards being the positive direction. Moreover, let x_1 denote the altitude, x_2 velocity and $u(t)$ the thrust applied by the rocket motor, subject to some capacity $|u(t)| \leq k$. The equations of motion are

$$\dot{x}_1 = x_2, \quad \dot{x}_2 = u,$$

with initial conditions

$$x_1(0) = h, \quad x_2(0) = v.$$

For a smooth landing at some time t_e, we require that

$$x_1(t_e) = 0, \quad x_2(t_e) = 0.$$

The problem is thus to determine the *control function* $u(t)$ so that both the initial and terminal conditions of the *state variables* x_1 and x_2 are satisfied, which is an objective at the core of control theory. As suggested above, this is most likely achieved by a number of controls, and many times it is desirable to choose one that in some sense optimizes the behaviour. The problem of finding such a control will be studied later on, when we try to establish the shape of a vessel that minimises the depletion time. So, without any further discussion on the properties of control systems or even defining what they are, we attend to the concept of optimality. All theory concerning control theory will be given when necessary.

3.2 Performance indices

Much like the equations above, we will consider systems with a general set of n nonlinear differential equations

$$\dot{x}(t) = f(x, u, t), \tag{3.1}$$

where $x \in \mathbf{R}^n$, subject to the initial condition

$$x(t_0) = x_0, \tag{3.2}$$

where the components of f are continuous and satisfy standard conditions, such as having continuous first partial derivatives so that the solution of (3.1) exists and is unique for condition (3.2).

By choosing different control functions $u(t)$, the system can be steered along many different paths, not all of which are equally desirable. Suppose that it is possible to measure the benefits associated with each path. Then we have the means to find an *optimal control* among all valid, *admissible*, controls. For this purpose, we introduce a *performance index*, or *objective functional*, a scalar which provides a measure by which the performance of the cost of the system can be judged.

It is understood that different performance indices are suitable in different situations. For example, reverting to the problem given above, we might be interested in minimising the total fuel consumption,

$$\int_0^{t_e} |u(t)|\, dt,$$

or, minimising the time of descent,

$$\int_0^{t_e} dt,$$

of the spacecraft. In other cases, the total amount of control effort that can be used to carry out a required task may be limited,

$$\int_{t_0}^{t_e} q(x, u, t)\, dt = c,$$

where c is a given constant. This is an *isoperimetric* constraint. It can be dealt with by defining a new variable

$$x_{n+1}(t) = \int_{t_0}^{t} q(x, u, s)\, ds,$$

so that

$$\dot{x}_{n+1}(t) = q(x, u, t). \tag{3.3}$$

Then (3.3) is simply added to the original set of differential equations (3.1), and the conditions

$$x_{n+1}(t_0) = 0, \quad x_{n+1}(t_e) = c,$$

are added to (3.2).

3.3 The minimum principle

One convenient yet powerful way of dealing with a wide range of optimization problems is to use the minimum principle, which provides necessary conditions for optimality. Our objective is to derive these conditions, without presenting any formal proof of the principle.

Much like before, we consider the system (3.1) and initial conditions (3.2) along with a performance index

$$J(u) = Q[x(t_e), t_e] + \int_{t_0}^{t_e} q(x, u, t)\, dt, \tag{3.4}$$

where $x \in \mathbf{R}^n$ and $u \in \mathcal{U} \subseteq \mathbf{R}^m$, $\mathcal{U}$ being the set of admissible controls. Our objective is to minimise this cost.

In realistic problems, the controls are often constrained in some way, typically by $|u_i(t)| \leq k_i$, as will be the case in a problem proposed in the following section. We will therefore take this into account throughout the derivation. Moreover, we shall also assume that the functional $J(u)$ is differentiable. Then if u and $u + \delta u$ are two admissible controls for which J is defined then

$$\Delta J = J(u + \delta u) - J(u) = \delta J(u, \delta u) + j(u, \delta u)\|\delta u\|, \tag{3.5}$$

where δJ is a linear function in δu and $j(u, \delta u) \to 0$ as $\|\delta u\| \to 0$, using some suitable norm. Here, δJ is termed the *variation* in J corresponding to a variation δu in u. Moreover, we say that the control u^* is an *extremal*, and that J has a relative minimum at u^*, if there exists an $\epsilon > 0$ such that for all controls u satisfying $\|u - u^*\| < \epsilon$ the difference $J(u) - J(u^*)$ is nonnegative.

In the next theorem we revert briefly to the case when the controls are unbounded.

Theorem 3.3.1. *Assume there are no constraints on the control functions. Then a necessary condition for u^* to be an extremal is that $\delta J(u^*, \delta u) = 0$.*

We will not give any formal proof of this theorem. It is only stated to justify the derivation of the following result:

Theorem 3.3.2. *Assume the control functions are bounded. Then a necessary condition for u^* to be an extremal is that $\delta J(u^*, \delta u) \geq 0$.*

It should be noted that we will only consider variations such that $u + \delta u$ is admissible and $\|\delta u\|$ is sufficiently small so that the sign of ΔJ is determined by δJ in (3.5). This distinction will prove to be crucial to the derivation of the theorem, for which we will present the fundamentals.

To begin with, recall that the control function u^* minimises the functional J if

$$\Delta J = J(u) - J(u^*) \geq 0,$$

for all admissible controls u sufficiently close to u^*. If we let $u = u^* + \delta u$, then according to (3.5)

$$\Delta J(u^*, \delta u) = \delta J(u^*, \delta u) + j(u^*, \delta u)\|\delta u\|.$$

At this point, the assumption that the controls are constrained on their magnitudes complicates the necessary steps considerably. To see why, assume for now that the controls were unbounded. Then we could use the linearity of δJ with respect to δu, together with the fact that δu can vary arbitrarily to show that a necessary condition for u^* to be an extremal control is that the variation $\delta J(u^*, \delta u)$ must be zero for all admissible δu having sufficiently small norm. However, since we are in fact requiring that the admissible controls are bounded, δu may vary arbitrarily only if the extremal control is strictly within the boundary during all times in the interval $[t_0, t_e]$. If this is the case, then the constraints do not effect the solution of the problem. If, however, an extremal control attains a boundary during at least one subinterval $[t_1, t_2]$ of the interval $[t_0, t_e]$, then variations $\delta \hat{u}$ in an admissible control exist but its negatives (i.e. $-\delta \hat{u}$) are not admissible. If only these (admissible) variations are considered, then a necessary condition for u^* to minimise J is that $\delta J(u^*, \delta \hat{u}) \geq 0$. On the other hand, for variations $\delta \tilde{u}$ that are nonzero only for t outside the interval $[t_1, t_2]$, a necessary condition is that $\delta J(u^*, \delta \tilde{u}) = 0$, according to theorem 3.3.1. So, considering all admissible variations with $\|u\|$ small enough so that the sign of ΔJ is determined by δJ, we conclude that a necessary condition for u^* to minimise J is

$$\delta J(u^*, \delta u) \geq 0.$$

There is in fact a similar condition for minimum points in calculus. To illustrate this, we consider a differentiable function f on $[x_0, x_1]$. Its differential df is then the linear part of the increment Δf. We consider only admissible values of the increment Δx, which are small enough so that the sign of Δf is determined by the sign of df, as before. If f has a relative minimum at the end point x_0, then $df(x_0, \Delta x)$ must be nonnegative. The same requirement must hold for $f(x_1)$ to be a relative minimum. Thus, necessary conditions for f to have relative minimum at the end points are

$$\begin{aligned} df(x_0, \Delta x) &\geq 0, \quad \text{for admissible } \Delta x \geq 0 \\ df(x_1, \Delta x) &\geq 0, \quad \text{for admissible } \Delta x \leq 0 \end{aligned}$$

and a necessary condition for f to have a relative minimum at an interior point x, $x_0 < x < x_1$, is

$$df(x, \Delta x) = 0.$$

These are in accordance with the necessary conditions for the control problem;

$$\delta J(u^*, \delta u) \geq 0,$$

if u^* lies on the boundary during any portion of the time interval $[t_0, t_e]$, and

$$\delta J(u^*, \delta u) = 0,$$

if u^* lies within the boundary during the entire time interval $[t_0, t_e]$.

We are now ready to apply theorem 3.3.2 to our minimisation problem. We will however consider an altered form of the performance index, namely

$$\hat{J} = Q\left[x(t_e), t_e\right] + \int_{t_0}^{t_e} \left[q(x, u, t) + p'(f - \dot{x})\right] dt,$$

where p is a vector of Lagrange multipliers $p = [p_1, \ldots, p_n]'$. Note that this performance index is chosen so as to take the constraints into account as well, since $\dot{x} = f(x, u, t)$.

The *Hamiltonian function* is introduced as

$$\mathcal{H}(p, x, u, t) = q(x, u, t) + p' f(x, u, t)$$

to obtain

$$\hat{J} = Q\left[x(t_e), t_e\right] + \int_{t_0}^{t_e} \left[\mathcal{H} - p'\dot{x}\right] dt.$$

Integrating the last term by parts then yields

$$\hat{J} = Q\left[x(t_e), t_e\right] - \left[p'x\right]_{t_0}^{t_e} + \int_{t_0}^{t_e} \left[\mathcal{H} + (\dot{p})'x\right] dt.$$

Assume that t_0 and t_e are fixed and that u is differentiable on $t_0 \leq t \leq t_e$. The variation in $\hat{J}$ corresponding to a variation δu in u is

$$\delta\hat{J} = \left[\left(\frac{\partial Q}{\partial x} - p'\right)\delta x\right]_{t=t_e} + \int_{t_0}^{t_e} \left[\frac{\partial \mathcal{H}}{\partial x}\delta x + \frac{\partial \mathcal{H}}{\partial u}\delta u + (\dot{p})'\,\delta x\right] dt, \tag{3.6}$$

where δx is the variation in x in the differential equations (3.1) due to δu, and

$$\begin{aligned}
\frac{\partial \mathcal{H}}{\partial x} &= \left[\frac{\partial \mathcal{H}}{\partial x_1}, \ldots, \frac{\partial \mathcal{H}}{\partial x_n}\right] \\
\frac{\partial \mathcal{H}}{\partial u} &= \left[\frac{\partial \mathcal{H}}{\partial u_1}, \ldots, \frac{\partial \mathcal{H}}{\partial u_m}\right] \\
\frac{\partial Q}{\partial x} &= \left[\frac{\partial Q}{\partial x_1}, \ldots, \frac{\partial Q}{\partial x_n}\right]
\end{aligned}$$

Moreover, since $x(t_0)$ is specified from the initial condition, $(\delta x)_{t=t_0} = 0$ is implied. Notably, if we choose p suitably, all terms involving δx are removed from (3.6). Evidently, this happens if p satisfies the *adjoint equations*

$$\dot{p}_i = -\frac{\partial \mathcal{H}}{\partial x_i} \tag{3.7}$$

with corresponding terminal conditions

$$p_i(t_e) = \left(\frac{\partial Q}{\partial x_i}\right)_{t=t_e}, \qquad \text{for all } i = 1, \ldots, n. \tag{3.8}$$

If these conditions are met, then equation (3.6) reduces to

$$\delta\hat{J} = \int_{t_0}^{t_e} \left(\frac{\partial \mathcal{H}}{\partial u}\delta u\right) dt.$$

Recall that if $\mathcal{H}$ is differentiable, then

$$\mathcal{H}(p, x, u + \delta u, t) - \mathcal{H}(p, x, u, t) = \frac{\partial \mathcal{H}}{\partial u}\delta u + \varrho(u, \delta u)\|\delta u\|,$$

where $\varrho(u, \delta u) \to 0$ as $\|\delta u\| \to 0$. Hence, the expression for $\delta \hat{J}$ can be replaced by

$$\delta \hat{J}(u, \delta u) = \int_{t_0}^{t_e} [\mathcal{H}(p, x, u + \delta u, t) - \mathcal{H}(p, x, u, t) - \varrho(u, \delta u)\|\delta u\|] \, dt,$$

so by theorem 3.3.2 a necessary condition for $u = u^*$ to be a minimising control is that $\delta \hat{J}(u^*, \delta u)$ is nonnegative for all admissible variations δu. This in turn implies that

$$\mathcal{H}(p^*, x^*, u^* + \delta u, t) - \mathcal{H}(p^*, x^*, u^*, t) - \varrho(u, \delta u)\|\delta u\| \geq 0.$$

But, if we continue to assume that $\|\delta u\|$ is sufficiently small so that the sign of $\Delta \mathcal{H}$ is not determined by it, then

$$\mathcal{H}(p^*, x^*, u^* + \delta u, t) - \mathcal{H}(p^*, x^*, u^*, t) \geq 0 \tag{3.9}$$

is a necessary condition, for all admissible δu and all $t \in [t_0, t_e]$. To see why (3.9) must hold true in the entire time interval, suppose it did not hold in some interval $t_1 \leq t \leq t_2$. Then, $\delta \hat{J}(u^*, \delta u)$ would become negative by choosing $\delta u = 0$ for all t outside this interval. Hence, equation (3.9) states that u^* minimises $\mathcal{H}$, so we have established the following fundamental result:

Pontryagin's minimum principle 3.3.3. *Necessary conditions for u^* to minimise (3.4) subject to (3.1) and (3.2) are that (3.7), (3.8) and (3.9) hold.*

Remark 3.3.4. It should be pointed out that there are also means of deriving this principle by using the well known Hamilton-Jacobi-Bellman equation. It is obtained by introducing a *value function* $V(x, t)$, which is the minimal future cost from time t onwards, given that the system is in state x at time t. Then, by using dynamic programming, one can show that a sufficient condition for optimality is given by the following partial differential equation

$$\frac{\partial V}{\partial t}(x, t) + \inf_{u \in \mathcal{U}} \left[q(x, u, t) + \frac{\partial V}{\partial x}(x, t) f(x, u, t) \right] = 0$$

with boundary condition $V(x, t_e) = Q(x, t_e)$.

However, this approach requires stronger assumptions compared to the minimum principle with variational calculus. Notably, the HJB equation requires that the value function has continuous partial derivatives in both x and t. Another drawback is that an explicit solution is often difficult to find, even implicit ones using numerical procedures. This is the main reason for why the variational calculus approach will be preferred in our study of optimization problems. On the positive side, the HJB equation gives a sufficient condition for optimality when the minimum principle only provides necessary conditions. Hence, we must always remember to prove that a candidate solution in fact is an optimal one when the minimum principle is used in favour of the HJB equation.

4 Examples of inverse problems

In this section we will use Torricelli's law to find a few inverse problems concerning fluid flow. We begin with deriving an ODE that describes the efflux out of a tank with a given shape. The generality of the equation makes it possible to apply it to almost any container that is of interest from a physical viewpoint. Next, we use this ODE to form a model identification problem, where the objective is to specify the geometry of a water clock. The second inverse problem, a causation problem, is that of finding a tank that produces a desired flow rate out of an irrigation canal.

4.1 The efflux equation

Let us consider a tank of volume V and height H initially filled with fluid. Suppose a small hole with cross-sectional area a is made in the bottom of the tank. Let $A(h)$, $0 \leq h \leq H$, be the area of the tank cross-section at height h.

Let h be the height of the fluid in the tank at time t. Moreover, let Δh denote the height drop during a small amount of time Δt that elapses from the moment t. Then the volume decrease ΔV is approximately $A(h)\Delta h$ during this period of time. If we let $v(h)$ denote the speed of the outgoing flow at height h, the amount of fluid leaving the tank during Δt is approximately equal to $av(h)\Delta t$. Combining these results, we conclude that $A(h)\Delta h = -av(h)\Delta t$. If we let $\Delta t \to 0$, the relationship turns into the differential equation

$$A(h)\frac{dh}{dt} = -av(h).$$

Applying Torricelli's law then yields

$$A(h)\frac{dh}{dt} = -a\kappa\sqrt{2gh}.$$

For convenience, we introduce a constant $\mu = a\kappa\sqrt{2g}$, which gives

$$A(h)\frac{dh}{dt} = -\mu\sqrt{h}. \tag{4.1}$$

Since we assume that the initial height is H, the initial condition is $h(0) = H$.

Remark 4.1.1. There is a restriction on the model. Due to physical considerations, equation (4.1) will not hold if a is large, because then one cannot ignore the kinetic energy gained by water moving toward the hole inside the tank. Therefore, we will assume that a is "small" in relation to the cross-sectional area, say $A(h) > \alpha$, for all $h \in [0, H]$, where α is some constant greater than a. For a thorough study of this phenomenon, one should consult a textbook on fluid dynamics. See for example [1].

Since we assume that the cross-sectional area is always positive, we can rewrite the equation as

$$\frac{dh}{dt} = -\mu\frac{\sqrt{h}}{A(h)}$$

or,

$$\dot{x}(t) = -\mu \frac{\sqrt{x(t)}}{u(x(t))}$$

by using the standard notation from control theory (the state variable x being the height and the control function u the cross-sectional area). By integrating from zero to t, and noting that $x(0) = H$, the following equivalent integral equation for x is obtained

$$x(t) = H - \mu \int_0^t \frac{\sqrt{x(s)}}{u(x(s))}\, ds \tag{4.2}$$

The equation has two aspects; the direct problem and the model identification problem. The former involves solving for $x(t)$ having $u(x)$ specified, and the latter determining $u(x)$ by knowledge of a given behaviour $x(t)$. In the sequel, our concern is the model identification problem, where t can be regarded as the input, x the output and u the input-output mapping.

The model identification problem evidently consists of solving a general nonlinear integral equation (4.2). Had the equation been linear in $x(t)$ and $u(t)$, then we could have applied the Laplace transform to express u in terms of the Laplace transform of x. However, there is no similar technique for nonlinear problems and hence no way of finding $u(x)$ for arbitrary $x(t)$. In the following, $x(t)$ must therefore be specified explicitly in order to obtain the corresponding input-output mapping u. An example of such a problem is dealt with next.

4.2 Designing a water clock

The water clock is, along with the sundial, probably the earliest time-measuring apparatus. It was the most accurate and widely used timekeeping device in Europe until it was replaced by the more accurate pendulum clock during the 17th century. However, the water clock is still used, but nowadays mainly for its elegance and simplicity rather than its accuracy.

There have been numerous examples of different constructions of water clocks over the years. What they all have in common, though, is that time is measured by means of a (possibly regulated) flow of liquid into or out of a container. These techniques can prove equally useful depending on the situation. However, we will pay attention only to the latter of these two, i.e. the outflow type.

Before Torricelli, there was no obvious way of constructing water clocks of outflow type, and one had to rely on other ideas. The Greeks, for instance, had problems with the decrease of outflow speed; one of the immediate consequences of Torricelli's law. However, thanks to Torricelli, there were means to construct an accurate water clock by specifying the shape of the vessel appropriately.

The objective here is to supply the mathematical model behind such a construction. The key feature of the desired vessel is the requirement of constant height decrease, i.e. that the height is linear in t. The next proposition contains the solution to this problem:

Proposition 4.2.1. *If $x(t) = H - ct$, for all t such that $x(t) > 0$ and some constant c, then $u(y) = \frac{\mu}{c}\sqrt{y}$ solves (4.2).*

Proof. By direct substitution into (4.2) we see that

$$
\begin{aligned}
x(t) &= H - \mu \int_0^t \frac{\sqrt{x(s)}}{u(x(s))}\, ds \\
&= H - \mu \int_0^t \frac{c\sqrt{x(s)}}{\mu\sqrt{x(s)}}\, ds, \qquad x(s) > 0 \\
&= H - ct
\end{aligned}
$$

□

Remark 4.2.2. The constant c is assumed to be some suitable positive number, so that $x(t)$ diminishes for increasing t in a desirable way. It must hence be chosen in accordance with the size of the container.

Remark 4.2.3. Geometrically, the area function u corresponds to the container obtained by revolving the curve $y = mx^4$ around the y-axis in the xy-plane, with $m = (\pi c/\mu)^2$. Symmetric shapes are most commonly used to design water clocks, for both artistic and historical reasons, and it is therefore instructive to point out this perspective of the problem.

4.3 Geometry via flow rates: an integral equation

Consider an irrigation canal of depth h. In the wall of the canal is a weir notch fitted with a sluice gate that is symmetric with respect to a central vertical axis through the gate, as in Figure 2. The shape of the notch is given by a function $x = u(y)$ as indicated in the figure. It is clear that the total rate of flow (i.e. the volume of flow per unit time) through the notch depends on this shape.

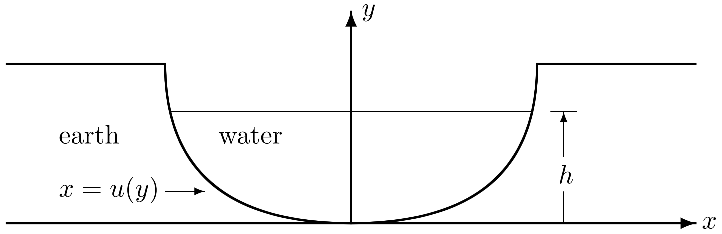

Figure 2: An irrigation canal

The relationship between the notch shape and the flow rate is obtained from Torricelli's law. If we consider a horizontal slab of water of thickness Δy at height y, then the volume of this slab passing through the notch per unit time is

$$2u(y)\Delta y\sqrt{2g\,(h-y)}.$$

Now, summing over all slabs and letting Δy tend to zero, we find that the total volume of flow per unit time through the notch, when the water level is h, is given by

$$f(h) = 2\sqrt{2g}\int_0^h u(y)\sqrt{h-y}\, dy. \tag{4.3}$$

The inverse problem, which is a causation problem, is to determine the geometry u having the flow rate f given. In this case, the function u should be regarded as the input-output mapping, the height h the input and f the output. Evidently, the problem involves the solution of the integral equation (4.3) that comes in the form of a convolution. It is therefore especially suited for methods involving Laplace transformation.

The problem was proposed by Brenke in [2] and illustrated by Groetsch in [3]. As opposed to the former, Groetsch gives a solution to the integral equation when the flow rate f is given explicitly. Brenke, on the other hand, presents a solution for "arbitrary" flow rates, and does so by transforming (4.3) into a special case of *Abel's integral equation*. In the following, we will merge the ideas of Brenke and Groetsch to obtain a solution that holds in general, by applying the Laplace transform directly to the integral equation. However, to understand Brenke's approach, we start off with a few words on Abel's integral equation.

Abel's integral equation is a special case of an integral equation in the form of a convolution. It was studied extensively by Niels Henrik Abel during his work on an inverse problem in pure mechanics. The *generalized Abel's integral equation*, which is an extended version of the original equation, reads

$$g(x) = \int_0^x \frac{1}{(x-t)^\lambda} u(t)\, dt, \qquad 0 < \lambda < 1$$

Evidently, this is a singular equation, since the denominator $(x-t)^\lambda$ is zero for $t = x$. Moreover, it can be shown that derivation of (4.3) with respect to h yields

$$f'(h) = \sqrt{2g} \int_0^h \frac{u(y)}{\sqrt{h-y}}\, dy, \tag{4.4}$$

which is a special case of Abel's integral equation (with $\lambda = 1/2$). At this point, Brenke concludes that a solution of (4.4) is also a solution of the original equation (4.3). However, this differentiation yields an improper integral (with singularity at $y = h$). It is not clear, at first glance, whether or not this is a valid operation. This may however be asserted under the assumption that the derivative of u is continuous.

To see this, consider the right-hand side of (4.4). Using integration by parts, we find that

$$\begin{aligned}
\int_0^h \frac{u(y)}{\sqrt{h-y}}\, dy &= \left[-2\sqrt{h-y}\cdot u(y)\right]_0^h - \int_0^h -2\sqrt{h-y}\cdot u'(y)\, dy \\
&= -2\left(\sqrt{h-h}\cdot u(h) - \sqrt{h}\cdot u(0)\right) + 2\int_0^h u'(y)\sqrt{h-y}\, dy \\
&= 2u(0)\sqrt{h} + 2\int_0^h u'(y)\sqrt{h-y}\, dy \\
&= 2\int_0^h u'(y)\sqrt{h-y}\, dy,
\end{aligned}$$

upon using that $u(0) = 0$ is given by the geometrical setting. Thereby, we have built up the power of $(h-y)$ inside the integral sign. It is thus clear that if u' is continuous on $[0, h]$ then we can guarantee that the integral is nonsingular.

It is now time to return to the original equation (4.3) and try to solve it directly by means of Laplace transforms. It should be noted that we will assume that u is sufficiently smooth in $[0, h]$, so that the original equation (4.3) is well defined and that u has a Laplace transform. Before examining all the details we begin with stating the solution to the problem as a proposition:

Proposition 4.3.1. *The equation*

$$f(h) = 2\sqrt{2g}\int_0^h u(y)\sqrt{h-y}\,dy$$

has a continuous solution

$$u(h) = \frac{1}{\pi\sqrt{2g}}\int_0^h \frac{f''(y)}{\sqrt{h-y}}\,dy$$

provided that $f(x)$ is continuous and has finite first and second derivatives $f'(x)$ and $f''(x)$ with at most a finite number of discontinuities in the range of integration.

This statement could be proved by direct substitution. We will however carry out the complete derivation of the solution to fully illustrate the use of Laplace transformation. So, following standard procedure, we begin with applying the Laplace transform to both sides of the equation:

$$\mathscr{L}\left[f(h)\right] = 2\sqrt{2g}\mathscr{L}\left[\int_0^h u(y)\sqrt{h-y}\,dy\right]$$

which comes in the form of a convolution. Hence,

$$\mathscr{L}\left[f(h)\right] = 2\sqrt{2g}\left(\mathscr{L}\left[u(h)\right]\cdot\mathscr{L}\left[\sqrt{h}\right]\right)$$

by use of the convolution theorem. Next, we need the Laplace transform of $\sqrt{x}$. The required calculations are however quite cumbersome to carry out in detail and are therefore enclosed as a proposition.

Proposition 4.3.2. *The Laplace transform of $\sqrt{x}$ is given by*

$$\mathscr{L}\left[\sqrt{x}\right] = \frac{\sqrt{\pi}}{2}s^{-3/2}, \quad s > 0$$

Proof. By definition, we have that

$$\mathscr{L}\left[\sqrt{x}\right] = \int_0^\infty e^{-sx}\sqrt{x}\,dx$$

We make the substitution $u = \sqrt{sx}$, such that $x = u^2/s$ and $dx = 2u/s$. Noting that the range of integration is unchanged, the integral becomes

$$\int_0^\infty e^{-u^2}\frac{u}{\sqrt{s}}\frac{2u}{s}\,du = \frac{1}{s^{3/2}}\int_0^\infty u\cdot 2ue^{-u^2}\,du$$

By use of partial integration we find that

$$\begin{aligned}
\mathscr{L}\left[\sqrt{x}\right] &= \frac{1}{s^{3/2}}\left(\left[u\left(-e^{-u^2}\right)\right]_0^\infty - \int_0^\infty -e^{-u^2}\,du\right)\\
&= \frac{1}{s^{3/2}}\left(\left[-ue^{-u^2}\right]_0^\infty + \int_0^\infty e^{-u^2}\,du\right)\\
&= \frac{1}{s^{3/2}}\int_0^\infty e^{-u^2}\,du\\
&= \frac{1}{2s^{3/2}}\int_{-\infty}^\infty e^{-u^2}\,du,
\end{aligned}$$

since the integrand is an even function. It can be shown by elementary calculus that the integral in the last stage equals $\sqrt{\pi}$, and thus $\mathscr{L}\left[\sqrt{x}\right] = \frac{\sqrt{\pi}}{2}s^{-3/2}$. □

Using this result and rearranging the equation gives

$$\begin{aligned}
\mathscr{L}\left[u(h)\right] &= \frac{1}{2\sqrt{2g}}\left(\frac{2}{\sqrt{\pi}}s^{3/2}\cdot\mathscr{L}\left[f(h)\right]\right)\\
&= \frac{1}{\sqrt{2g\pi}}s^{-1/2}\left(s^2\cdot\mathscr{L}\left[f(h)\right]\right) \qquad (4.5)
\end{aligned}$$

The reason for rewriting the equation in this way is because there is no inverse Laplace transform of $s^{3/2}$ in explicit form. We will instead try to make something useful of the expression in brackets in the last expression. We have that

$$\begin{aligned}
s^2\cdot\mathscr{L}\left[f(x)\right] &= s^2\int_0^\infty e^{-sx}f(x)\,dx\\
&= s\int_0^\infty se^{-sx}f(x)\,dx\\
&= s\left(\left[-e^{-sx}f(x)\right]_0^\infty - \int_0^\infty -e^{-sx}f'(x)\,dx\right)\\
&= s\left(-f(0) + \int_0^\infty e^{-sx}f'(x)\,dx\right)\\
&= -sf(0) + s\int_0^\infty e^{-sx}f'(x)\,dx\\
&= -sf(0) + s\mathscr{L}\left[f'(x)\right]\\
&= -sf(0) + f'(0) + \mathscr{L}\left[f''(x)\right],
\end{aligned}$$

by use of repeated partial integration. From equation (4.3) follows directly that $f(0) = 0$. A similar argument can be given about $f'(0)$ by virtue of (4.4), so $f'(0) = 0$ as well. Inserting these results into (4.5) thus yields

$$\mathscr{L}\left[u(h)\right] = \frac{1}{\sqrt{2g\pi}}s^{-1/2}\cdot\mathscr{L}\left[f''(x)\right]. \qquad (4.6)$$

To find u explicitly we need the inverse Laplace transform of $s^{-1/2}$. It is established in the following proposition:

Proposition 4.3.3. *The inverse Laplace transform of $s^{-1/2}$ is given by*

$$\mathscr{L}^{-1}[s^{-1/2}] = \frac{1}{\sqrt{\pi}} x^{-1/2}, \quad x > 0$$

Proof. The inverse can be found by using the fact that the inverse Laplace transform of $s^{-3/2}$ is known. For this purpose, let $g(x) = \mathscr{L}^{-1}\left[s^{-1/2}\right]$. We then have the relation

$$\int_0^\infty e^{-sx} g(x)\, dx = s^{-1/2}.$$

Let G be a primitive function of g. Then, by partial integration

$$\left[e^{-sx}G(x)\right]_0^\infty + s\int_0^\infty e^{-sx}G(x)\, dx.$$

Now, if $G(x)$ is of exponential order, then

$$s\int_0^\infty e^{-sx}G(x)\, dx = s^{-1/2},$$

or equivalently,

$$\int_0^\infty e^{-sx}G(x)\, dx = s^{-3/2}.$$

From proposition 4.3.2 follows that $\mathscr{L}\left[\sqrt{x}\right] = \frac{\sqrt{\pi}}{2}s^{-3/2}$, and thus

$$\mathscr{L}\left[G(x)\right] = \frac{2}{\sqrt{\pi}}\mathscr{L}\left[\sqrt{x}\right].$$

This implies that $G(x) = \frac{2}{\sqrt{\pi}}\sqrt{x}$. Finally, since $g(x) = G'(x)$, we find that

$$g(x) = \frac{1}{\sqrt{\pi}} x^{-1/2},$$

provided that $x > 0$. □

Finally, inserting into (4.6) yields

$$\begin{aligned}
\mathscr{L}\left[u(h)\right] &= \frac{1}{\sqrt{2g\pi}}\mathscr{L}\left[\frac{1}{\sqrt{\pi}}x^{-1/2}\right]\cdot\mathscr{L}\left[f''(x)\right] \\
&= \frac{1}{\pi\sqrt{2g}}\mathscr{L}\left[x^{-1/2}\right]\cdot\mathscr{L}\left[f''(x)\right]
\end{aligned}$$

By again invoking the convolution theorem and applying the inverse Laplace transform yields the continuous solution

$$u(h) = \frac{1}{\pi\sqrt{2g}}\int_0^h \frac{f''(y)}{\sqrt{h-y}}\, dy.$$

Remark 4.3.4. It should be noted that we have assumed that f is twice differentiable and that its derivatives are sufficiently smooth and have Laplace transforms. Moreover, we have also made the assumption that the Laplace transform of $x^{-1/2}$ exists, without argument. As noted earlier, this is not an obvious conclusion since it has a discontinuity at $x = 0$, and the corresponding Laplace transform therefore includes an improper integral. However, this can be remedied by partial integration,

$$\begin{aligned}\int_0^\infty e^{-sx}x^{-1/2}\,dx &= \left[2e^{-sx}\sqrt{x}\right]_0^\infty + 2s\int_0^\infty e^{-sx}\sqrt{x}\,dx \\ &= 2s\cdot\mathscr{L}\left[\sqrt{x}\right],\end{aligned}$$

which is well defined.

5 The optimal depletion time

In earlier sections, we have seen that the flow speed out of a container depends on its shape and the initial volume of fluid. Hence, there is reason to believe that the depletion time, i.e. the time it takes to empty a given container, is closely connected with these parameters as well.

In the following, we seek a theoretical minimum of the depletion time when the observed container has constant cross-sectional area and when the area may vary, respectively. In both cases it is imperative that the initial volume of fluid is the same, as well as the size of the exit hole, so as to have the possibility to compare the efficiency of different methods.

5.1 Problem specifications

We begin by defining the optimization problem. First of all, an expression for the depletion time is needed. It can be obtained from the efflux equation (4.1)

$$A(h)\frac{dh}{dt} = -\mu\sqrt{h}.$$

Since this is an autonomous ODE (with no explicit t dependence) the method of separating variables can be used to obtain an implicit solution

$$\int_{h_0}^{h} \frac{A(s)}{\sqrt{s}}\, ds = \int_{t_0}^{t} -\mu\, du.$$

Noting that $h(0) = H$ then yields

$$\int_{h}^{H} \frac{A(s)}{\sqrt{s}}\, ds = \mu t.$$

Moreover, if the depletion time is denoted by T, we have the boundary condition $h(T) = 0$. Thus,

$$\int_{0}^{H} \frac{A(s)}{\sqrt{s}}\, ds = \mu T,$$

or equivalently,

$$T = \frac{1}{\mu}\int_{0}^{H} \frac{A(s)}{\sqrt{s}}\, ds.$$

However, the integral has evidently a singularity at $s = 0$, so it is not obvious if this is well defined. To avoid dealing with this singularity, we shall instead consider the limit

$$T = \lim_{\epsilon\to 0^+} \frac{1}{\mu}\int_{\epsilon}^{H} \frac{A(s)}{\sqrt{s}}\, ds.$$

Since we require that the initial volume is constant, an expression for the volume of any given container filled to a height h is also of interest. Because $A(x)$ is the cross-sectional area at height x, the volume is given by

$$V(h) = \int_{0}^{h} A(s)\, ds, \tag{5.1}$$

for any tank, assuming that A is at least piecewise continuous on $[0, H]$.

Now, suppose that the initial volume is V_0, for some positive constant. We can then directly see from expression (5.1) that $V(H) = V_0$ is a boundary condition. This is in fact an isoperimetric constraint. It also follows that an additional boundary condition $V(0) = 0$. Thus, upon differentiation of (5.1) we obtain the problem of minimising

$$T = \lim_{\epsilon \to 0^+} \frac{1}{\mu} \int_{\epsilon}^{H} \frac{A(s)}{\sqrt{s}}\, ds$$

subject to

$$\begin{cases} V'(h) = A(h) \\ \lim_{\epsilon \to 0^+} V(\epsilon) = 0 \\ V(H) = V_0 \end{cases}$$

that holds for any tank with piecewise continuous area function. The reason for including the limit in the constraints is to avoid ambiguity in the calculations, although (5.1) is perfectly well defined. With this extension, the problem may be regarded as optimization over $[\epsilon, H]$ and then taking a limit.

If the volume V is regarded as the state equation and the cross-sectional area A is regarded as the control function, then the problem can be formulated as an optimal control problem in x and u as

$$T^* = \min \lim_{\epsilon \to 0^+} \frac{1}{\mu} \int_{\epsilon}^{H} \frac{u(t)}{\sqrt{t}}\, dt$$

subject to

$$\begin{cases} \dot{x}(t) = u(t) \\ \lim_{\epsilon \to 0^+} x(\epsilon) = 0 \\ x(H) = V_0 \end{cases} \tag{5.2}$$

using standard notation from control theory. This will prove convenient.

Moreover, as noted in remark 4.1.1, we must see to it that the cross-sectional area is large in comparison to the exit hole. On the other hand, we are not interested in tanks of infinite range and should therefore also put an upper bound on the size of the tank. It is hence reasonable to restrict this study solely to tanks with $\alpha \leq A \leq \beta$, for some suitable constants $\alpha > a$ and $\beta < \infty$. Consequently, the initial height H must be chosen such that $\alpha H \leq V_0 \leq \beta H$.

5.2 Constant cross-sectional area

Given the general specifications of the problem, we can easily transform this to hold in particular for tanks with constant cross-sectional area. Using $u(t) = c$, for some positive constant $c \in [\alpha, \beta]$, we obtain the optimization problem

$$T_c^* = \min \lim_{\epsilon \to 0^+} \frac{1}{\mu} \int_{\epsilon}^{H} \frac{c}{\sqrt{s}}\, ds \tag{5.3}$$

subject to

$$\begin{cases} \dot{x}(t) = c, \quad \alpha \le c \le \beta \\ \lim_{\epsilon \to 0^+} x(\epsilon) = 0 \\ x(H) = V_0 \end{cases}$$

We can directly see that the constraints are equivalent to the relationship

$$\begin{aligned} \lim_{\epsilon \to 0^+} c\,(H - \epsilon) &= V_0, \\ cH &= V_0 \end{aligned} \tag{5.4}$$

The problem can readily be solved without the methods of optimal control theory. We begin by computing the integral in (5.3),

$$\begin{aligned} \lim_{\epsilon \to 0^+} \frac{1}{\mu} \int_\epsilon^H \frac{c}{\sqrt{s}}\, ds &= \lim_{\epsilon \to 0^+} \frac{c}{\mu} \left[2\sqrt{s}\right]_\epsilon^H \\ &= \lim_{\epsilon \to 0^+} \frac{2c}{\mu} \left(\sqrt{H} - \sqrt{\epsilon}\right) \\ &= \frac{2c}{\mu}\sqrt{H} \end{aligned}$$

Finally, using (5.4) yields

$$T_c = \frac{2c}{\mu}\sqrt{\frac{V_0}{c}} = \frac{2}{\mu}\sqrt{V_0 c} \tag{5.5}$$

Clearly, this expression is minimal when $c = \alpha$ (and maximal when $c = \beta$, conversely). Hence, the minimal depletion time is given by

$$T_c^* = \frac{2}{\mu}\sqrt{V_0 \alpha}, \tag{5.6}$$

and is achieved when the cross-sectional area is α.

5.3 The general case

We now turn to the problem of finding the minimal depletion time for vessels of arbitrary shape. By this we mean vessels with cross-sectional areas that are constrained by $\alpha \le A(h) \le \beta$ in $[0, H]$.

A similar problem was proposed by Hanin in [4], who considered the possibility of finding a minimal depletion time for symmetric vessels. However, in the problem domain there were no restrictions given on the cross-sectional area, and therefore the depletion time could become arbitrarily small. To see this, let for example α tend to zero in expression (5.6). There are a few drawbacks with this approach. To begin with, if the cross-sectional area is allowed to become arbitrarily small, then the fluid flow cannot be modelled by means of Torricelli's law, for reasons already mentioned. Moreover, if we let α tend to zero, then the initial height H must tend to infinity, which is hardly realistic.

We will therefore restrict our attention solely to the set of admissible controls, $\mathcal{U}$, consisting of all piecewise continuous functions on $[0, H]$ that take on values in the control region $[\alpha, \beta]$, for some fixed constant H in $[V_0/\beta, V_0/\alpha]$.

We will use the minimum principle to find a solution. The objective functional to be minimised is

$$T(u) = \lim_{\epsilon \to 0^+} \frac{1}{\mu} \int_{\epsilon}^{H} \frac{u(t)}{\sqrt{t}} \, dt, \quad u \in \mathcal{U} \tag{5.7}$$

subject to the state equation, initial and terminal conditions according to (5.2). This should be regarded as optimization over $[\epsilon, H]$ and then taking a limit, as noted earlier.

Notably, the objective functional is time-varying, since it depends explicitly on t. Moreover, since there is one system equation $\dot{x} = u(t)$, only one Lagrange multiplier is necessary to incorporate in the Hamiltonian. Thus,

$$\begin{aligned} \mathcal{H}(p, x, u, t) &= p(t)u(t) + u(t)/\sqrt{t} \\ &= u(t)\left(p(t) + \frac{1}{\sqrt{t}}\right) \end{aligned}$$

According to the necessary conditions of optimality in the minimum principle, an optimal control $u^*(t)$ must minimise the Hamiltonian. Since this function is linear in u, we see directly that the minimising argument is

$$u^*(t) = \begin{cases} \alpha, & \text{if } p^*(t) < 1/\sqrt{t} \\ \beta, & \text{if } p^*(t) \geq 1/\sqrt{t} \end{cases}$$

along with the adjoint equation

$$\dot{p}^*(t) = -\frac{\partial \mathcal{H}}{\partial x} = 0.$$

From this equation follows that $p^*(t) = c_1$, for some constant c_1. Because there is no function outside the integral in our objective functional (5.7), condition (3.8) vanishes in the minimum principle and does not provide any terminal condition for p^*. Hence, our conclusion so far is that the optimal control is given by

$$u^*(t) = \begin{cases} \alpha, & \text{if } \epsilon \leq t < 1/c_1^2 \\ \beta, & \text{if } 1/c_1^2 \leq t \leq H \end{cases}$$

which is clearly well defined regardless of the value of c_1. An optimal control like this one is known as a *bang-bang* control, since the control alternates between extremes, in this case between α and β.

The optimal control changes values at the switching point $t_s = 1/c_1^2$. This is uniquely determined by the terminal condition in (5.2), which implies that

$$\lim_{\epsilon \to 0^+} \int_{\epsilon}^{H} u^*(t) \, dt = V_0.$$

Moreover, there can only be one switching time that satisfies this condition, since $u^*(t)$ is piecewise constant. So,

$$
\begin{aligned}
V_0 &= \lim_{\epsilon \to 0^+} \int_{\epsilon}^{H} u^*(t)\, dt \\
&= \lim_{\epsilon \to 0^+} \int_{\epsilon}^{t_s} \alpha\, dt + \int_{t_s}^{H} \beta\, dt \\
&= \lim_{\epsilon \to 0^+} \alpha\,(t_s - \epsilon) + \beta\,(H - t_s) \\
&= \alpha t_s + \beta\,(H - t_s)
\end{aligned}
$$

Hence, $t_s = (\beta H - V_0)\,/\,(\beta - \alpha)$. Reverting to the original notations, the optimal area function is thus

$$
A(h) = \begin{cases} \alpha, & \text{if } 0 \le h < t_s \\ \beta, & \text{if } t_s \le h \le H \end{cases}
$$

and the minimal depletion time is

$$
T_g^* = \frac{2}{\mu}\left(\beta\sqrt{H} - \sqrt{\beta H - V_0}\sqrt{\beta - \alpha}\right). \tag{5.8}
$$

This is well defined since β is the maximal cross-sectional area allowed, and it is thus understood that $\beta H \ge V_0$.

5.4 Verification of optimality

As noted earlier, the minimum principle provides only necessary conditions for optimality. To show that our bang-bang control is in fact an optimal control, we need sufficient conditions. The HJB equation would have provided these unless the control function had been discontinuous. Instead we must use a different approach to show that the bang-bang control is optimal. This is provided by the following lemma:

Sufficient conditions for optimality 5.4.1. *Consider the problem of minimising*

$$
\int_{t_0}^{t_e} q(x(t), u(t))\, dt \tag{5.9}
$$

subject to

$$
\dot{x} = f(x(t), u(t)), \quad x(t_0) = x_0, x(t_e) = x_e \tag{5.10}
$$

where u belongs to the set of admissible controls $\mathcal{U}$, here the set of bounded, piecewise continuous functions defined on a bounded interval. Further, let X be the set of initial states from which the terminal state x_e is reachable.

Let $u^*(t)$ be a bang-bang control with one switching point t_s, and let $x^*(t)$, $t_0 \le t \le t_e$, be the corresponding solution. Then, $u^*(t)$ is optimal with respect to all the admissible controls $u(t)$ with corresponding solutions $x(t)$, $t_0 \le t \le t_e$, such that $x(t_0) = x_0$, $x(t_e) = x_e$ and $x(t) \in X$, for all $t \in (t_0, t_e)$, if there exists a piecewise continuously differentiable function V^* on X such that

(i) $q(x^*(t), u^*(t)) + \text{grad}\, V^*(x^*(t)) \cdot f(x^*(t), u^*(t)) = 0, \quad$ *for all* $t \in [t_0, t_e)$

(ii) $q(x,u) + \text{grad } V^*(x) \cdot f(x,u) \geq 0, \quad$ *for all* $u \in \mathcal{U}, x \in X$

Proof. Suppose $u^*(t)$ is a bang-bang control and that $V^*(x)$ and the gradient of $V^*(x)$ are discontinuous at $x(t_s) = x_s$. Then there are functions $V_1^*(x)$ and $V_2^*(x)$ on X with values in $\mathbf{R}$ such that

$$V^*(x) = \begin{cases} V_1^*(x(t)), & t_0 \leq t < t_s \\ V_2^*(x(t)), & t_s \leq t \leq t_e \end{cases}$$

and v_1^* and v_2^* on X with values in $\mathbf{R}^2$ such that

$$\text{grad } V^*(x) = \begin{cases} v_1^*(x(t)), & t_0 \leq t < t_s \\ v_2^*(x(t)), & t_s \leq t \leq t_e \end{cases}$$

Now, define

$$\begin{aligned} I_1 &= \int_{x_0}^{x_e} \text{grad } V^*(x)\, dx \\ &= \int_{x_0}^{x_s} v_1^*(x)\, dx + \int_{x_s}^{x_e} v_2^*(x)\, dx \\ &= V_1^*(x_s) - V_1^*(x_0) + V_2^*(x_e) - V_2^*(x_s) \end{aligned} \tag{5.11}$$

By (i) of the lemma, together with (5.10) and (5.11),

$$\begin{aligned} I_2 &= \int_{t_0}^{t_e} q(x^*(t), u^*(t))\, dt \\ &= \int_{t_0}^{t_e} -\text{grad } V^*(x^*(t)) \cdot f(x^*(t), u^*(t))\, dt \\ &= \int_{t_0}^{t_e} -\text{grad } V^*(x^*(t)) \cdot \frac{dx^*}{dt}\, dt \\ &= \int_{x_0}^{x_e} -\text{grad } V^*(x^*)\, dx^* \\ &= -I_1 \end{aligned}$$

Now, consider the difference in the value of the performance index,

$$\Delta = \int_{t_0}^{t_e} q(x(t), u(t))\, dt - \int_{t_0}^{t_e} q(x^*(t), u^*(t))\, dt \tag{5.12}$$

We may add the null term $I_1 + I_2$ to the right-hand side of (5.12) without changing its value. In view of (5.10),

$$I_1 = \int_{t_0}^{t_e} \text{grad } V^*(x(t)) \cdot f(x(t), u(t))\, dt,$$

so that

$$
\begin{aligned}
\Delta &= \int_{t_0}^{t_e} q(x(t),u(t))\,dt - \int_{t_0}^{t_e} q(x^*(t),u^*(t))\,dt + I_1 + I_2 \\
&= \int_{t_0}^{t_e} \left[q(x(t),u(t)) + \text{grad } V^*(x(t))\cdot f(x(t),u(t))\right]\,dt - \int_{t_0}^{t_e} q(x^*(t),u^*(t))\,dt + I_2 \\
&= \int_{t_0}^{t_e} \left[q(x(t),u(t)) + \text{grad } V^*(x(t))\cdot f(x(t),u(t))\right]\,dt
\end{aligned}
$$

Thus, by (ii) of the lemma, we have

$$\Delta \geq 0$$

which concludes the proof. □

This is a special case of a sufficiency theorem presented in [5]. It is adjusted especially to meet the request for sufficient optimality conditions for problems with bang-bang control. Note that we have assumed that there was only a single switching point. This can be extended to hold for controls with several switching points, but this will however complicate the calculations considerably. Another restriction is that the lemma is applicable only to autonomous problems, without explicit t dependence. It can however easily be extended to hold for nonautonomous ones, by introducing an additional state variable.

In our case we let $x_2 = t$. Then we obtain the equivalent problem of minimising

$$\lim_{\epsilon\to 0^+} \frac{1}{\mu}\int_{\epsilon}^{H} \frac{u(t)}{\sqrt{x_2(t)}}\,dt$$

subject to

$$\begin{cases} \dot{x} = \begin{bmatrix} u(t) & 1\end{bmatrix}' \\ \lim_{\epsilon\to 0^+} x(\epsilon) = \begin{bmatrix} 0 & 0\end{bmatrix}' \\ x(H) = \begin{bmatrix} V_0 & H\end{bmatrix}' \end{cases}$$

If $u^*(t)$ is the control defined above, then the corresponding solution is

$$x^*(t) = \begin{cases} \begin{bmatrix}\alpha t & t\end{bmatrix}', & \epsilon \leq t < t_s \\ \begin{bmatrix}\beta t + (\alpha-\beta)\,t_s & t\end{bmatrix}', & t_s \leq t \leq H \end{cases}$$

So, $u^*(t)$ is the optimal control according to the lemma if there is a function V^* such that

(i) $u^*(t)/\sqrt{t} + \text{grad } V^*(x^*(t))\cdot\begin{bmatrix}u^*(t) & 1\end{bmatrix}' = 0, \quad \text{for all } t \in [\epsilon, H)$

(ii) $u/\sqrt{x_2} + \text{grad } V^*(x^*)\cdot\begin{bmatrix}u & 1\end{bmatrix}' \geq 0, \quad \text{for all } u \in \mathcal{U}, x \in X$

It follows from condition (i) that

$$\begin{cases} \alpha/\sqrt{t} + \text{grad } V^*(x^*(t))\cdot\begin{bmatrix}\alpha & 1\end{bmatrix}' = 0, & \epsilon \leq t < t_s \\ \beta/\sqrt{t} + \text{grad } V^*(x^*(t))\cdot\begin{bmatrix}\beta & 1\end{bmatrix}' = 0, & t_s \leq t < H \end{cases}$$

which is evidently equivalent to the set of equations

$$\begin{cases} \alpha/\sqrt{t} + \dot{V}^*(x^*(t)) = 0, & \epsilon \le t < t_s \\ \beta/\sqrt{t} + \dot{V}^*(x^*(t)) = 0, & t_s \le t < H \end{cases}$$

Hence, V^* must satisfy $V^*(x^*(t)) = -2\alpha\sqrt{t} + c_2$ on $\epsilon \le t < t_s$ and $V^*(x^*(t)) = -2\beta\sqrt{t} + c_2$ on $t_s \le t < H$. Notably, it has a discontinuity at $t = t_s$. We must also take condition (ii) into account when choosing V^*. Since $x_2 = t$, it seems reasonable to try

$$V^*(x) = \begin{cases} -2\alpha\sqrt{x_2}, & \epsilon \le t < t_s \\ -2\beta\sqrt{x_2}, & t_s \le t < H \end{cases}$$

It has already been shown that this choice holds for (i), so it remains to be seen that condition (ii) holds. To begin with, when grad $V^*(x) = \begin{bmatrix} 0 & -\alpha/\sqrt{x_2} \end{bmatrix}$ we must show that

$$\frac{u}{\sqrt{x_2}} - \frac{\alpha}{\sqrt{x_2}} \ge 0, \quad \text{for all } u \in \mathcal{U}, x \in X$$

Note first that this is well defined since $x_2 = t > \epsilon$. Moreover, the inequality must hold since $u \ge \alpha$, for all $u \in \mathcal{U}$. Finally, when grad $V^*(x) = \begin{bmatrix} 0 & -\beta/\sqrt{x_2} \end{bmatrix}$ we must show that

$$\frac{u}{\sqrt{x_2}} - \frac{\beta}{\sqrt{x_2}} \ge 0, \quad \text{for all } u \in \mathcal{U}, x \in X$$

But it can easily be seen that $u = \beta$ is the only admissible control that transfers the state variable from x_s to x_e, so this expression must be zero, and we have thus shown that u^* in fact is an optimal control.

5.5 A comparison

It should be noted that there is a conceptually important difference between the constant area case and the general case. In the former, the initial height H is allowed to vary in order to find the minimal depletion time for a specific (fixed) initial volume V_0. In the latter, both H and V_0 are fixed, and the optimal depletion time is then determined by these parameters, as seen in (5.8). Since these methods rely on different assumptions, it is crucial to keep both H and V_0 fixed to compare the efficiency of the methods.

We must first point out that it is still required that H is restricted by $V_0/\beta \le H \le V_0/\alpha$. So, if both H and V_0 are fixed, then the constant cross-sectional area is uniquely determined, $c = V_0/H$, as well as the depletion time,

$$T_c = \frac{2}{\mu}\frac{V_0}{\sqrt{H}},$$

in accordance with (5.5). Then the improvement is

$$\begin{aligned} \Delta T = T_c - T_g^* &= \frac{2}{\mu}\frac{V_0}{\sqrt{H}} - \frac{2}{\mu}\left(\beta\sqrt{H} - \sqrt{\beta H - V_0}\sqrt{\beta - \alpha}\right) \\ &= \frac{2}{\mu}\left(\frac{V_0}{\sqrt{H}} - \beta\sqrt{H} + \sqrt{\beta H - V_0}\sqrt{\beta - \alpha}\right) \end{aligned}$$

A plot of the improvement in depletion time as the initial height H increases is included below (Figure 3). The curve is based on generic values $\alpha = 0.5$, $\beta = 10$ and $V_0 = 10$ as H is taken over the interval $[1, 20]$.

An additional plot shows the height decrease over time for both the constant cross-sectional area case and when optimal control is applied (Figure 4). In the former case, the cross-sectional area c is set to the mean of alfa and beta. The graphs are otherwise based on parameter values $\alpha = 0.5$, $\beta = 10$, $V_0 = 10$ as above, consequently yielding $H \approx 1.9$.

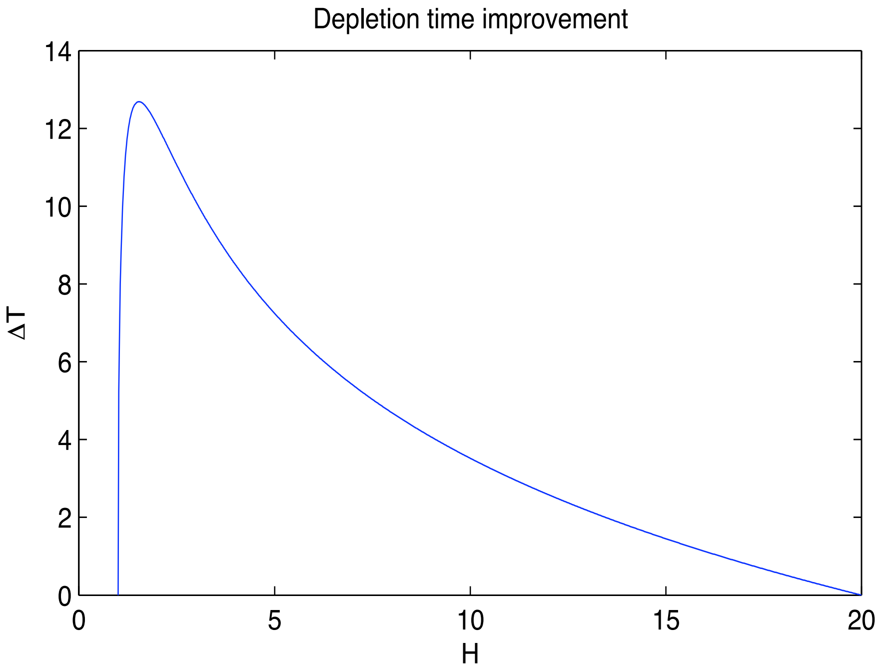

Figure 3: Depletion time improvement

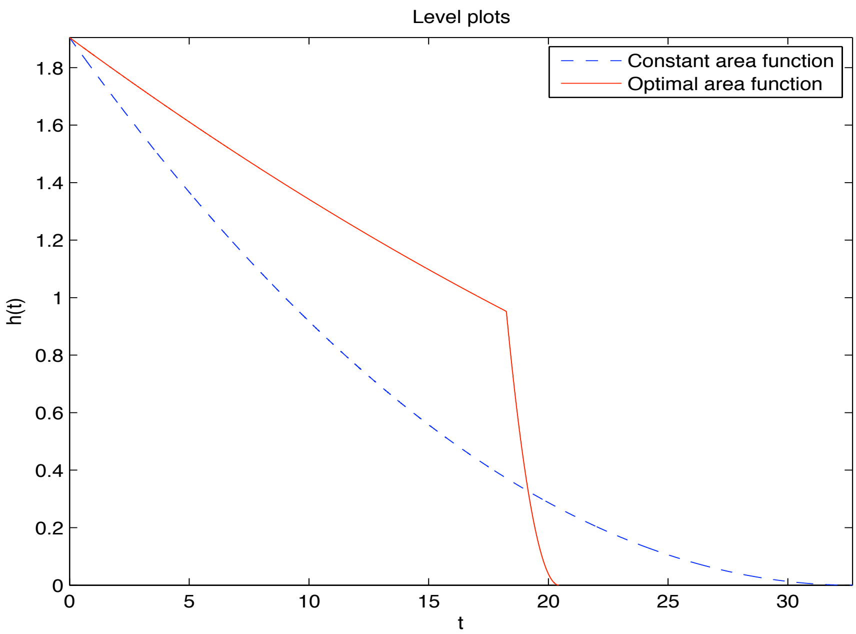

Figure 4: Outflow in the optimal versus the constant area case

6 Conclusion

We have studied outflow problems derived from Torricelli's law with emphasis on inverse problems and optimal control. In particular, we have shown that the outflow from arbitrary containers can be described by an autonomous nonlinear ODE. When the shape of the container is given by the curve $y = cx^4$ revolved around the y-axis, a proper design for a water clock is obtained, producing constant height decrease. An additional problem of interest, that is closely related to this one, but has not been addressed in this thesis, is to what extent the outflow can be controlled by choosing the cross-sectional area appropriately.

Moreover, we have derived and solved an inverse problem involving an integral equation that comes in the form of a convolution. We have shown that the solution is also given by a convolution, and that it can be obtained by applying the Laplace transform to the equation.

In the final section of the thesis, we have used the above ODE to derive two optimal control problems, which both involve finding minimal depletion times for arbitrary containers with known initial volume. In the first of these problems, we found that the minimum is attained if the container has minimal cross-sectional area. In the latter problem, where the initial height of the fluid is also specified, we found that the optimum is given by applying a bang-bang control. When comparing the methods (for a given fixed initial height), we found that there is a considerable difference between their performance. In particular, the bang-bang control resulted in a significant improvement compared to the method using a constant control.

References

[1] Vernard, J. K. *Elementary Fluid Mechanics* (Wiley, 1962)

[2] Brenke, W. C. An Application of Abel's Integral Equation, *The American Mathematical Monthly* **29** (1922) 58-60

[3] Groetsch, C. W. Inverse Problems and Torricelli's Law, *The College Mathematics Journal* **24** (1993) 215-217

[4] Hanin, L. G. Which tanks empty faster?, *The American Mathematical Monthly* **106** (1999) 943-947

[5] Leitmann, G. Sufficiency Theorems for Optimal Control, *Journal of Optimization Theory and Applications* **5** (1968) 285-292