On Inverse Problems and Machine Learning

Kirill V. Golubnichiy

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2022

Reading Committee:

Kenneth Bube, Chair

Michael Klibanov, Chair

Yu Yuan

Program Authorized to Offer Degree:
Mathematics

University of Washington

Abstract

On Inverse Problems and Machine Learning

Kirill V. Golubnichiy

Co-Chairs of the Supervisory Committee:
Professor Kenneth Bube
Department of Mathematics

Professor Michael Klibanov
Department of Mathematics and Statistics
University of North Carolina at Charlotte

This document is related to Ill-Posed and Inverse problems particularly focused on economic measurements. In 2015, I proposed to myself to work both analytically and numerically on a very fresh and surprising idea: to predict prices of stock options using the famous Black-Scholes equation. In mathematical finance, the Black–Scholes equation is a parabolic partial differential equation in both time and space that models the price of common financial assets. This equation when solved forwards in time to forecast prices of stock options is an ill-posed inverse problem. Note that standard techniques which were known at that time, do not provide any tools for predictions of prices. Besides, solving the Black-Scholes equation forwards in time is an ill-posed problem, which complicates the matter tremendously. On the other hand, it is intuitively clear that an accurate prediction of prices even for the next couple business days is exactly what the market dreams about.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

Let me express my deep gratitude to all those who took part in the preparation, presentation, public defense and discussion of my dissertation! First of all, let me express my sincere gratitude to my PhD advisors Professor Kenneth Bube and Professor Michael Klibanov for their help at all stages of the dissertation. I have received much valuable advice from Professor Keneth Bube to find balance between research, teaching, and family. I'm extremely grateful for Professor Michael Klibanov helping me grow as a specialist and as a person. Professor Klibanov is rightfully one of the world-class specialists in inverse problems, and it is a great honor for me to have him as my PhD advisor.

I also wanted to thank my undergraduate advisor Professor Mikhail Sukhinin who passed away a year ago. Thanks to Professor Sukhinin, I fell in love with inverse problems. Mikhail Sukhinin was not only an outstanding scientist, but also just a very good person, and I partly dedicate this dissertation to him.

I especially want to thank Professor Gunther Uhlmann for his trust and the opportunity to work at the University of Washington. Professor Uhlmann is an internationally recognized scientist in the field of inverse problems and my communication with him was an invaluable experience that I received.

Secondly, let me express my deep gratitude to the other members of my committee, Professor Yu Yuan, Professor Bobby Wilson and Professor Aleksandr Aravkin. Especially, I am very grateful to Professor Yu Yuan for his constant support and allowing me to participate in his PDE learning seminar. These mathematically challenging and intellectually stimulating experiences are invaluable preparation for my post-graduate career.

Thirdly, let me express my deep gratitude to all my friends from University of Washington for emotional support. These include Nikolaos Eptaminitakis, Andrea Ottolini, Josh Southerland and David Clancy. I would also like to thank my dear friend Ravi Shankar, Instructor at Princeton University, for engaging conversations and assisting with proofreading.

Especially, I would like to say a huge thank you to my best friend and collaborator Andrey Nikitin. Andrey is a great specialist in computer science, but most importantly he is a wonderful friend who has always supported me all this time. I am sincerely grateful that I have such a friend.

Finally, I want to thank my parents. There is no doubt that this is not only my dissertation but also theirs. I could not achieve without the support of my parents. This work is dedicated to my parents.

DEDICATION

to my Parents

Chapter 1

INTRODUCTION

The first question to address was about a realistic statement of an initial boundary value problem for the Black-Scholes equation. Indeed, it is well known that at each particular moment of time each stock option has only three prices: bid, ask and sell. On the other hand, the Black-Scholes equation requires an interval of stock prices for each particular moment of time. My thesis adviser (Professor Michael Klibanov) and I have invented an empirical initial boundary value problem for the Black-Scholes equation. Then this model has required a regularization-based numerical method which would solve the above ill-posed initial boundary value problem for the Black-Scholes equation forwards in time. This numerical method was also proposed and its convergence rate was established using a far non trivial tool of Carleman estimates.

Due to ill-posedness of the problem the solution is very sensitive to the noise in the initial data (stock and option prices for the three days preceding the day of forecast). My collaborator Andrey V. Nikitin (Amazon Web Services) and I applied Machine Learning to reduce the probability of non-profitable trades caused by wrong option price prognosis because of the noise in input data.

Here we discussed our numerical method based on minimization of Tikhonov-like functional $J_\beta(u)$ using conjugate gradient method. The minimization process was performed by Hyak Next Generation Supercomputer of the research computing club of University of Washington. The code was parallelized in order to maximize the performance on supercomputer clusters.

The historical data for stock options was collected from the Bloomberg terminal of University of Washington. From this data, we obtained about 177,000 minimizers.

In the previous paper (Inverse Problems, 32, 015010, 2016), a new heuristic mathematical model was proposed for accurate forecasting of prices of stock options for 1-2 trading days ahead of the present one. This new technique uses the Black-Scholes equation supplied by new intervals for the underlying stock and new initial and boundary conditions for option prices. The Black-Scholes equation was solved in the positive direction of the time variable, This ill-posed initial boundary value problem was solved by the so-called Quasi-Reversibility Method (QRM). This approach with an added trading strategy was tested on the market data for 368 stock options and good forecasting results were demonstrated. In the current paper, we use the geometric Brownian motion to provide an explanation of that effectivity using computationally simulated data for European call options. We also provide a convergence analysis for QRM. The key tool of that analysis is a Carleman estimate.

The next chapter presents a novel way to predict options price for one day in advance, utilizing the method of Quasi-Reversibility for solving the Black-Scholes equation. The Black-Scholes equation is solved forwards in time, which is an ill-posed problem. Thus, Tikhonov regularization via the Quasi-Reversibility Method is applied. This procedure allows to forecast stock option prices for one trading day ahead of the current one. To enhance these results, the Neural Network Machine Learning is applied on the second stage. Real market data are used. Results of Quasi-Reversibility Method and Machine Learning method are compared in terms of accuracy, precision and recall.

I conducted research on the convexification method for the Coefficient Inverse Problem, in collaboration with Professor Michael Klibanov (University of North Carolina at Charlotte) and Professor Kenneth Bube (University of Washington).The first step is to apply certain new changes of variables to the original problem to obtain a new Cauchy problem with the lateral Cauchy data for a quasilinear integro-differential equation with Volterra-like integrals in it. As soon as the solution of this problem is obtained, the target unknown coefficient can be computed by a simple backwards calculation. The second step is to obtain a new Carleman estimate for the linear part of the operator of that equation.

This chapter was devoted to inverse problems for nonlinear equations of the modified transfer, which can be regarded as a manageable problem. Various productions such problems for normal (unmodified) of the transport equation studied earlier by Aleksei I. Prilepko and most complete form in the dissertation of his follower Nikolai P. Volkov. If management is a factor in the coefficient of absorption or scattering indicatrix, even in the case of conventional linear transfer equation inverse problem becomes nonlinear (in the thesis N. P. Volkov overcome this difficulty).

Chapter 2

FORECASTING STOCK OPTIONS PRICES VIA THE SOLUTION OF AN ILL-POSED PROBLEM FOR THE BLACK-SCHOLES EQUATION

2.1 Introduction

A new heuristic mathematical algorithm designed to forecast prices of stock options was proposed in [17]. This algorithm is based on the so-called Quasi-Reversibility Method (QRM). QRM is a regularization method for an ill-posed problem for the Black-Scholes equation. The goal of this work is to address both analytically and numerically the following question: *Why this algorithm has worked well for real market data in [17, 16]?* Our explanations are based on our new analytical results in the probability theory and are supported by our numerical results for the computationally simulated data generated by the geometrical Brownian motion.

A significant advantage of the technique of [17] is that it uses historical data about stock and option prices only over short time intervals. This assumption is a practically valuable one since formations of those prices are random processes. This indicates that the information used in the algorithm possesses stable probabilistic characteristics.

The mathematical model of [17] was supplied by a trading strategy. Results of [17, Table 4] for real market data of [2] indicate that a combination of that mathematical model with that trading strategy has resulted in 72.83% profitable options out of 368 options for real market data. More recently, the model of [17] was used in [16] to forecast stock option prices in the case when results of QRM are enhanced by the machine learning approach, which was applied on the second stage of the price forecasting procedure. Market data of [2] for total 169,862 European call stock options were used in [16]. Following the machine learning approach, these data were divided in three sets [16, Table 1]: training (132,912 options), validation (13,401 options) and testing (23,549 options). Total 23,549 options were tested by QRM, and good results on predictions of options with profits were obtained in [16, first lines in Tables 2,3]. Later the authors of [16] have tested the performance of QRM for all those 169,862 options, and results were almost the same as ones of [16, first lines in Tables 2,3]. However, since the latter results are not yet published, then we do not discuss them here.

Remark 2.1.1: *Without further specifications, we consider in this paper only European call options. The mathematical model of [17] does not use neither the payment function at the expiry time nor the strike price.*

We now present in Tables 1,2 the most recent results of [16], which were obtained using the method of [17] for the data consisting of 23,549 historical trades collected in 2018. The same market data of [2] were used in Tables 1,2. Option prices for one trading day ahead of the present day were forecasted. Definitions of accuracy, precision and recall are well known, see, e.g. [14].

Table 1. Results of QRM for market data of [2] for 23,549 options [16, Table 2]

Method	Accuracy	Precision	Recall	Error
QRM	49.77%	55.77%	52.43%	12 %

In Table 1, "Error" means the average relative error of predictions of option prices, i.e.

$$\text{Error} = \frac{1}{N}\sum_{n=1}^{N}\left|\frac{p_{n,\text{corr}} - p_{n,\text{fc}}}{p_{n,\text{corr}}}\right| \cdot 100\%,$$

where $N = 23,549$ is the total number of tested options, $p_{n,\text{corr}}$ and $p_{n,\text{fc}}$ are correct and forecasted prices respectively of the option number n.

Table 2. The percentage of options with correct predictions of profits via the Quasi-Reversibility Method for the market data of [2] for 23,549 options [16, Table 3]

Method	Correctly Predicted Profitable Options
QRM	55.77%

A perfect financial market does not allow a winning strategy [10]. This means that to address the above question, we need to assume that the market is imperfect. The present article considers a model situation, in which there is a difference between the volatility σ of the underlying stock and traders' opinion $\hat{\sigma}$ of the volatility of an European call option generated by this stock. We prove analytically that, theoretically, this allows one to design a winning strategy. First, we back up this theory numerically for the ideal case when both volatilities are known. In practice, however, only $\hat{\sigma}$ is approximately known from [2], where implied volatility σ_{impl} of option prices is posted. It is reasonable to conjecture that $\hat{\sigma} \approx \sigma_{\text{impl}}$.

Second, to address the question posed in the first paragraph of this section for the non ideal case, we consider a mathematical model, in which the dynamics of the stock prices is generated by the stochastic differential equation of the geometric Brownian motion. This allows us to computationally generate the time series of stock prices. At the same time, we assume that the price of the corresponding stock option is governed by the Black-Scholes equation, in which the volatility coefficient is $\hat{\sigma}$. Hence, using that time series of stock prices, we apply the Black-Scholes formula to get the time series for prices of the corresponding options. Next, we apply the QRM to predict the prices of these options for one trading day ahead of the current one. Next, we formulate the winning strategy for the non ideal case.

Both the theory and the numerical studies of this paper support our two hyphotheses formulated in subsection 2.6.3. Our first hypothesis that the heuristic algorithm of [17] actually figures out in many cases the sign of the difference $\sigma - \hat{\sigma}$. Our second hypothesis is also based on our results below as well as on the "Precision" column of Table 1 and the second column of Table 2. More precisely, the second hypothesis is that probably about 56% of tested 23,549 options of [16] with the real market data had $\sigma - \hat{\sigma} < 0$.

This algorithm of [17] is based on the solution of a new initial boundary value problem (IBVP) for the Black-Scholes equation, see, e.g. [6, 38] for this equation. Since the Black-Scholes equation is actually a 1-D parabolic Partial Differential Equation (PDE) with the reversed time, then that IBVP is ill-posed, see, e.g. [17] for an example of a high instability of a similar problem. The ill-posedness of that IBVP is the main mathematical obstacle of that algorithm. Therefore, we solve that IBVP both here and in [17] by a specially designed version of QRM. QRM stably solves this problem forwards in time for two consecutive trading days after the current one. QRM is a version of the Tikhonov regularization method [35] being adapted to ill-posed problems for Partial Differential Equations (PDEs). We refer to [23] for the first publication on QRM as well as to [7, 8, 15, 20], [18] for some more recent ones.

We provide a convergence analysis for QRM being applied to the above problem. The main new element of this analysis is that we lift a restrictive assumption of [17] of a sufficiently small time interval. We note that the smallness assumption imposed on the time interval is a traditional one for initial boundary value problems for parabolic PDEs with the reversed time, see [15], [24, Theorem 1 of section 2 in Chapter 4], where a certain Carleman estimate was used. However, a new Carleman estimate was derived in [20] for a general parabolic operator of the second order with variable coefficients in the $n-$D case. This estimate enables one to lift that smallness assumption. We simplify here the Carleman estimate of [20] as well as some other results of [20] via adapting them to our simpler $1-$D case, as compared with the $n-$D case of [20].

The Black-Scholes equation describes the dependence of the price $v(s,t)$ of a stock option from the price of the underlying stock s and time t [5, 6, 38]. In fact, this is a parabolic Partial Differential Equation with the reversed time. Let $t = T$ be the maturity time and $t = 0$ is the present time [38]. Traditionally, initial boundary value problems for the Black-Scholes equation are solved backwards with respect to time $t \in (0,T)$ with the initial condition at $\{t = T\}$. The latter is a well posed problem, for which the classic theory of parabolic PDEs works, see, e.g. the book [22] for this theory.

However, the maturity time T is usually a few months away from the present time. It is obviously impossible to accurately predict the future behavior of the volatility coefficient of the Black-Scholes equation on such a large time interval. Since the formations of both stock and option prices are stochastic processes, then it is intuitively clear a good accuracy of forecasting of stock option prices for long time periods is unlikely.

Thus, we focus in this paper on forecasting of option prices for a short time period of just one trading day ahead of the current one. Let the time variable t counts trading days.

Since there are 255 trading days annually, then we introduce the dimensionless time t' as

$$t' = \frac{t}{255}. \tag{2.1.1}$$

Hence,

$$\text{one (1) dimensionless trading day } = 1/255 \approx 0.00392 << 1. \tag{2.1.2}$$

Remark 2.1.2. *To simplify notations, we still use everywhere below the notation t for the dimensionless time t' of (2.1.1).*

Remark 2.1.3. *There are many important questions about the technique of [17], which are not addressed in this paper, such as, e.g. the performance of this technique for some "stress" tests, its performance for significantly larger sets of market data, its performance for the case when the transaction cost is taken into account, and many others. However, addressing any of those questions would require a significant additional effort. Therefore, those questions are outside of the scope of this publication. Still, the question of the transaction cost might probably be addressed using a threshold number $\eta > 0$ in our trading strategy for the non-ideal case, see subsection 2.6.3.*

This paper is organized as follows. In section 2 we show that a winning strategy on an infinitesimal time interval might be possible if $\sigma \neq \hat{\sigma}$. In section 3 we present the heuristic mathematical model of [17]. In section 4 we present a convergence analysis for our version of QRM. In section 5 we use arguments of the probability theory to justify our trading strategy in the ideal case when both volatilities σ and $\hat{\sigma}$ are known. In section 6 we describe our numerical studies and end up with a trading strategy for the non ideal case when only the volatility $\hat{\sigma}$ is known. In addition, we formulate in section 6 our two hyphotheses mentioned above. Concluding remarks are given in section 7.

Disclaimer. *This document is written for academic purposes only. The authors do not provide any assurance that the technique of this paper would result in a successful trading on a real financial market.*

2.2 A Possible Winning Strategy

Let σ be the volatility of a certain stock and s be the price of this stock. Consider an option corresponding to this stock. Let $\hat{\sigma}$ be an idea of the volatility of that option, which has been developed among the agents, who trade this option on the market. If $\sigma \neq \hat{\sigma}$, then the financial market is imperfect, and an opportunity for designing a winning strategy exists.

At a given time t, the time until the maturity will occur is τ,

$$\tau = T - t. \tag{2.2.1}$$

Let s be the stock price and $f(s)$ be the payoff function of that option at the maturity time $t = T$. We assume that the risk-free interest rate is zero. Let $u(s, \tau)$ be the price of that option and the variable τ is the one defined in (2.2.1). We assume that the function $u(s, \tau)$ satisfies the Black-Scholes equation with the volatility coefficient $\hat{\sigma}$ [5, Chapter 7, Theorem 7.7]**:**

$$\frac{\partial u(s,\tau)}{\partial \tau} = \frac{\hat{\sigma}^2}{2} s^2 \frac{\partial^2 u(s,\tau)}{\partial s^2}, \quad s > 0, \tag{2.2.2}$$
$$u(s, 0) = f(s).$$

The specific formula for the payoff function is $f(s) = \max(s - K, 0)$, where K is the strike price [5]. Then the price function $u(s, \tau)$ of the option is given by the Black-Scholes formula [5]:

$$u(s,\tau) = s\Phi(\Theta_+(s,\tau)) - e^{-r\tau} K \Phi(\Theta_-(s,\tau)), \tag{2.2.3}$$

where $r = 0$ and

$$\Theta_\pm(s,\tau) = \frac{1}{\hat{\sigma}\sqrt{\tau}} \left[\ln\left(\frac{s}{K}\right) \pm \frac{\hat{\sigma}^2 \tau}{2} \right],$$
$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_z^\infty e^{-r^2/2} dr, \; z \in \mathbb{R}. \tag{2.2.4}$$

Let $v(s, t) = u(s, T - t)$. The stochastic equation of the geometric Brownian motion for the stock price s with the volatility σ has the form $ds = \sigma s dW$, where W is the Wiener process. The Itô formula implies

$$dv = \left(-\frac{\partial u(s, T-t)}{\partial \tau} + \frac{\sigma^2}{2} s^2 \frac{\partial^2 u(s, T-t)}{\partial s^2} \right) dt + \sigma s \frac{\partial u(s, T-t)}{\partial s} dW, \tag{2.2.5}$$

where dv is the option price change on an infinitesimal time interval and dW is the Wiener process.

Replacing in (2.2.5) $\partial_\tau u(s, T-t)$ with the right hand side of (2.2.2), we obtain

$$dv = \frac{(\sigma^2 - \hat{\sigma}^2)}{2} s^2 \frac{\partial^2 u(s, T-t)}{\partial s^2} dt + \sigma s \frac{\partial u(s, T-t)}{\partial s} dW. \tag{2.2.6}$$

The mathematical expectation of dW is zero [5, Chapter 4]. Therefore, we find that the expected value of the increment of the option price on an infinitesimal time interval is

$$\frac{(\sigma^2 - \hat{\sigma}^2)}{2} s^2 \frac{\partial^2 u(s, T-t)}{\partial s^2} dt. \tag{2.2.7}$$

In the mathematical finance, the second derivative

$$\frac{\partial^2 u(s, \tau)}{\partial s^2} \tag{2.2.8}$$

is called Greek $\Gamma(s, \tau)$. For an European call option [5, Chapter 9]

$$\Gamma(s, \tau) = \frac{1}{\hat{\sigma} s \sqrt{2\pi\tau}} \exp\left[-\frac{(\Theta_+(s, \tau))^2}{2} \right] > 0. \tag{2.2.9}$$

Therefore, it follows from (2.2.7)-(2.2.9) that the sign of the mathematical expectation of the increment of the option price on an infinitesimal time interval is determined by the sign of the difference $\sigma^2 - \hat{\sigma}^2$. Thus, if $\sigma^2 > \hat{\sigma}^2$, then a possible winning strategy involves buying an option at the present time and selling it in the next trading period. If $\sigma^2 < \hat{\sigma}^2$, then the winning strategy is to take the short position at the present time and to close the short position in the next trading period.

2.3 The Mathematical Model

2.3.1 The model

We now describe the mathematical model of [17]. We use this model here for computationally simulated data. Also, it was used in [16] for real market data to obtain the above Tables 1,2. We do not differentiate in this model between volatilities σ and $\hat{\sigma}$ and just use the time dependent volatility $\sigma(t)$.

Everywhere below, as the dimensionless time, we still use the notation t for t' in (2.1.1) for brevity. Let $\sigma(t)$ be the volatility of the option at the moment of time t. When working with the market data in [17, 16], we have used the historical implied volatility listed on the market data of [2]. Let $v_b(t)$ and $v_a(t)$ be respectively the bid and ask prices of the option and $s_b(t)$ and $s_a(t)$ be the bid and ask prices of the stock. It is known that

$$v_b(t) < v_a(t) \text{ and } s_b(t) < s_a(t).$$

For brevity, we simplify notations as $s_b = s_b(0)$, $s_a = s_a(0)$. We impose a natural assumption that $0 < s_b < s_a$.

It was observed on the market data in [17, formulas (2.3)-(2.6)] that the relative differences are usually small,

$$\left|\frac{s_a(t)}{s_b(t)} - 1\right| \leq 0.03, \quad \left|\frac{v_a(t)}{v_b(t)} - 1\right| \leq 0.27. \tag{2.3.1}$$

Hence, we define the initial condition $q(s)$ at $t = 0$ of the function $v(s,t)$ as the linear interpolation on the interval $s \in (s_b, s_a)$ between $v_b(0)$ and $v_a(0)$,

$$v(s,0) = q(s) = -\frac{s - s_a}{s_a - s_b} v_b(0) + \frac{s - s_b}{s_a - s_b} v_a(0), \ s \in (s_b, s_a). \tag{2.3.2}$$

Define the domain $Q_T = \{(s,t) \in (s_b, s_a) \times (0,T)\}$. We assume that the volatility of the option depends only on t, i.e. $\sigma = \sigma(t) \geq \sigma_0 = const. > 0$. Let L be the partial differential operator of the Black-Scholes equation,

$$Lv = \frac{\partial v}{\partial t} + \frac{\sigma^2(t)}{2} s^2 \frac{\partial^2 v}{\partial s^2} = 0 \text{ in } Q_T. \tag{2.3.3}$$

We impose the following initial and boundary conditions on the function $v(s,t)$:

$$v(s,0) = q(s), \ s \in (s_b, s_a), \tag{2.3.4}$$

$$v(s_b,t) = v_b(t), v(s_a,t) = v_a(t), \ t \in (0,T). \tag{2.3.5}$$

Conditions (2.3.3)-(2.3.5) represent the heuristic mathematical model of [17, formulas (2.3)-(2.6)]. Also, (2.3.3)-(2.3.5) is our IBVP for the Black-Scholes equation. We now formulate this as Problem 1:

Problem 1. *Find the function $v \in H^2(Q_T)$ satisfying conditions (2.3.3)-(2.3.5).*

Problem 1 is ill-posed since we need to solve equation (2.3.3) forwards in time.

Remarks 2.3.1:

1. *The conventional model for the Black-Scholes equation stresses on the maturity time T via considering the function $u(s,t) = v(s, T-t)$ instead of the function $v(s,t)$.*

Unlike this, we are not doing so in (2.3.3)-(2.3.5) since we do not need the maturity time, also, see Remark 1.1.

2. *As it is a conventional way in the theory of Ill-Posed problems, we increase here the required smoothness of the solution from* $H^{2,1}(Q_T)$ *to* $H^2(Q_T)$.

2.3.2 Three steps

In order to solve Problem 1, we need first to define the time dependent option's volatility $\sigma(t)$, and boundary conditions $v_b(t)$, $v_a(t)$. Then the initial condition $q(s)$ in (2.3.4) would be found via (2.3.4). We explain these in Steps 1,2 of this subsection 2.3.2.

In our computations of [17, 16] we have used the Implied Volatility of the options in the Last Trade Price (IVOL) of the day for $\sigma(t)$ [2]. As to s_b and s_a, we have used the End of the Day Underlying Price Ask and the End of Day Underlying Price Bid of [2]. Similarly for $v_b(t)$ and $v_a(t)$, in which case the End of the Day Option Price Ask and the End of Day Option Price Bid of [2] were used. The moment of time $\{t = 0\}$ is the End of the Present Day Time, and similarly for the following two trading days of $t = y, 2y$ and for the preceding two trading days $t = -y, -2y$. Naturally, the question can be raised here on how did we find future values of boundary conditions $v_b(t)$ and $v_a(t)$ for $t \in (0, 2y)$ in (2.3.5), and the same for $\sigma(t)$. This question is addressed in Step 2 below. Our method for the solution of Problem 1 consists of three steps:

Step 1 (introducing dimensionless variables)**.** First, we make equation (2.3.3) dimensionless. Recall that $s_b < s_a$. Introduce the dimensionless variable x for s as:

$$s \Leftrightarrow x = \frac{s - s_b}{s_a - s_b}.$$

Let y denotes one dimensionless trading day. By (2.1.2)

$$y = \frac{1}{255} \approx 0.00392. \tag{2.3.6}$$

By (2.3.2) the function $q(s)$ is transformed in the function $g(x)$,

$$g(x) = (1 - x) v_b(0) + x v_a(0). \tag{2.3.7}$$

And the operator L in (2.3.3) is transformed in the operator M,

$$Mv = v_t + \sigma^2 (t) A(x) v_{xx}, \tag{2.3.8}$$

$$A(x) = \frac{255}{2} \frac{[x(s_a - s_b) + s_b)]^2}{(s_a - s_b)^2}, \tag{2.3.9}$$

$$G_{2y} = \{(x,t) \in (0,1) \times (0,2y)\}. \tag{2.3.10}$$

Problem is transformed in Problem 2:

Problem 2. *Assume that functions*

$$v_b (t), v_a (t) \in H^2 [0,2y], \sigma (t) \in C^1 [0,2y]. \tag{2.3.11}$$

Find the solution $v \in H^2 (G_{2y})$ of the following initial boundary value problem:

$$Mv = 0 \text{ in } G_{2y}, \tag{2.3.12}$$

$$v (0,t) = v_b (t), v (1,t) = v_a (t), t \in (0,2y), \tag{2.3.13}$$

$$v (x,0) = g (x), x \in (0,1), \tag{2.3.14}$$

where the partial differential operator M is defined in (2.3.8), the function $A(x)$ is defined in (2.3.9), the initial condition $g(x)$ is defined in (2.3.7), and the domain G_{2y} is defined in (2.3.10).

Step 2 (interpolation and extrapolation). Having the historical market data for an option up to “today”, we forecast the option price for “tomorrow” and “the day after tomorrow”, with 255 trading days annually. “One day” corresponds to $y = 1/255$. “Today” means $t = 0$. “Tomorrow” means $t = y$. “The day after tomorrow” means $t = 2y$. We forecast these prices for the $s-$interval as $s \in [s_b(0), s_a(0)]$ via the solution of problem (2.3.12)-(2.3.14). To do this, however, we need to know functions $v_b(t)$, $v_a(t)$ and $\sigma(t)$ in the “future”, i.e. for $t \in (0,2y)$. We obtain approximate values of these functions via interpolation and extrapolation procedures described in the next paragraph.

Let $t = -2y$ be “the day before yesterday”, $t = -y$ be “yesterday” and $t = 0$ be “today”. Let $d(t)$ be any of three functions $v_b(t), v_a(t), \sigma(t)$. First, we interpolate the function $d(t)$ by the quadratic polynomial for $t \in [-2y, 0]$ using the values $d(-2y), d(-y), d(0)$. We obtain

$$d(t) = at^2 + bt + c \text{ for } t \in [-2y, 0]. \tag{2.3.15}$$

Next, we extrapolate (2.3.15) on the interval $t \in [0, 2y]$ via setting

$$d(t) = at^2 + bt + c \text{ for } t \in [0, 2y].$$

The so defined functions $v_b(t)$, $v_a(t), \sigma(t)$ were used to numerically solve problem (2.3.12)-(2.3.13) for both the computationally simulated data below and for real market data of Tables 1,2 above as well as in [17].

Step 3 (Numerical solution of Problem 2. Regularization). Since problem (2.3.3)-(2.3.5) is ill-posed, then we apply a regularization method to obtain an approximate solution of this problem. More precisely, we solve the following problem:

Minimization Problem 1. *Let $J_\alpha : H^2(G_{2y}) \to \mathbb{R}$ be the regularization Tikhonov functional defined as:*

$$J_\alpha(v) = \int_{G_{2y}} (Mv)^2 \, dxdt + \alpha \|v\|^2_{H^2(G_{2y})}, \tag{2.3.16}$$

where $\alpha \in (0,1)$ is the regularization parameter. Minimize functional (2.3.16) on the set S, where

$$S = \left\{ v \in H^2(G_{2y}) : v(0,t) = v_b(t), v(1,t) = v_a(t), v(x,0) = g(x) \right\}. \tag{2.3.17}$$

Minimization Problem 1 is a version of QRM for Problem 2. This version is an adaptation of the QRM for problem (2.3.12)-(2.3.13). In section 2.4 we present the theory of this specific version of the QRM. In particular, Theorem 2.4.2 of section 2.4 implies uniqueness of the solution $u \in H^2(G_{2y})$ of Problem 2 and provides an estimate of the stability of this solution with respect to the noise in the data. Theorem 2.4.3 of section 2.4 implies existence and uniqueness of the minimizer $v_\alpha \in H^{2,1}(G_{2y})$ of the functional $J_\alpha(v)$ on the set S defined in (2.3.17). Following the theory of Ill-Posed problems, we call such a minimizer "regularized solution" [35]. Theorem 2.4.4 estimates convergence rate of regularized solutions to the exact solution of Problem 2 with the noiseless data. These estimates depend on the noise level in the data.

2.4 Convergence Analysis

In this section, we provide convergence analysis for Problem 2 of subsection 2.3. This problem is the IBVP for parabolic equation (2.3.12) with the reversed time, see (2.3.8). The

QRM for this problem for a more general parabolic operator in $\mathbb{R}^n$ with arbitrary variable coefficients was proposed in [15] and convergence analysis was also carried out there. Then corresponding theorems were reformulated in [17]. Although a stability estimate was not a part of [17], such an estimate was proven in [15]. It was pointed out in Introduction, however, that traditional stability estimates for this problem were proven, using a certain Carleman estimate, only under the assumption that the time interval is sufficiently small. The same is true for the convergence theorems of QRM in [15, 17]. Unlike this, the smallness assumption was lifted in [20] via a new Carleman estimate. In this section, we significantly modify results of [20] for a simpler 1-D case. Recall (see Introduction) that this modification allows us to obtain more accurate estimates in the 1-D case, as compared with the $n-$D case of [20]. We note that even though we work in our computations below on a small time interval $(0, 2y) = (0, 0.00784)$ (see (2.3.6) and (2.3.10)), the smallness assumption of [15, 17], [24, Theorem 1 of section 2 in Chapter 4] might result in the requirement of even a smaller length of that interval.

2.4.1 Problem statement

Consider a number $T_1 > 0$ and denote

$$Q_{T_1} = \{(x,t) \in (0,1) \times (0,T_1)\}.$$

Let two numbers $a_0, a_1 > 0$ and $a_0 < a_1$. Let the function $a(x,t) \in C^1\left(\overline{Q}_{T_1}\right)$ satisfies:

$$\|a\|_{C^1\left(\overline{Q}_{T_1}\right)} \leq a_1, \ a(x,t) \geq a_0 \text{ in } Q_{T_1}. \tag{2.4.1}$$

Let functions $\varphi_0(t), \varphi_1(t) \in H^2(0,T_1)$. In the above case of subsection 2.3.2,

$$T_1 = 2y, a(x,t) = \sigma^2(t) A(x), \varphi_0(t) = v_b(t), \varphi_1(t) = v_a(t).$$

We now formulate Problem 3, which is a slight generalization of Problem 2.

Problem 3. *Find a solution $w \in H^2(Q_{T_1})$ of the following initial boundary value problem (IBVP):*

$$Pw = w_t + a(x,t) w_{xx} = 0 \text{ in } Q_{T_1}, \tag{2.4.2}$$

$$w(0,t) = \varphi_0(t), w(1,t) = \varphi_1(t), \ t \in (0,T_1), \tag{2.4.3}$$

$$w(x,0) = q(x) = \varphi_0(0)(1-x) + \varphi_1(0)x,\ x \in (0,1). \tag{2.4.4}$$

Remark 2.4.1. *Since Problem 3 is a more general one than Problem 2, then our convergence analysis for Problem 3, which we provide below, is also valid for Problem 2.*

The reason why we use the linear function for $w(x,0)$ in (2.4.4) is our desire to simplify the presentation by using the fact that, in the case of Problem 2, the initial condition in (2.3.14) is the linear function defined in (2.3.7). Problem 3 is an IBVP for the parabolic equation (2.4.2) with the reversed time. Therefore, this problem is ill-posed. Just as it is always the case in the theory of Ill-Posed problems [35], we assume that the boundary in (2.4.3) are given with a noise of the level $\delta > 0$, where δ is a sufficiently small number, i.e.

$$\|\varphi_0 - \varphi_0^*\|_{H^1(0,T_1)} < \delta, \|\varphi_1 - \varphi_1^*\|_{H^1(0,T_1)} < \delta, \tag{2.4.5}$$

where functions $\varphi_0^*, \varphi_1^* \in H^2(0,T_1)$ are "ideal" noiseless data. Following to one of postulates of the theory of Ill-Posed problems, we assume that there exists an exact solution $w^* \in H^2(Q_{T_1})$ of problem (2.4.2)-(2.4.4) with these noiseless data. We will estimate below how this noise affects the accuracy of the solution of Problem 3 (if this solution exists) and also will establish the convergence rate of numerical solutions obtained by QRM to the exact one as $\delta \to 0$.

Consider the following analog of functional (2.3.16):

$$I_\alpha(w) = \int_{Q_{T_1}} (Pw)^2\, dxdt + \alpha \|w\|^2_{H^2(Q_{T_1})}. \tag{2.4.6}$$

Introduce the set $Y \subset H^2(Q_{T_1})$,

$$Y = \left\{ w \in H^2(Q_{T_1}) : w(0,t) = \varphi_0(t), w(1,t) = \varphi_1(t), w(x,0) = q(x) \right\}. \tag{2.4.7}$$

We construct an approximate solution of Problem 3 via solving the following problem:

Minimization Problem 2. *Minimize the functional $I_\alpha(w)$ on the set Y given in (2.4.7).*

Similarly with the Minimization Problem1, Minimization Problem 2 means QRM for Problem 3.

2.4.2 Theorems

In this subsection, we formulate four theorems for Problem 3. Let $\lambda > 2$ be a parameter. Introduce the Carleman Weight Function $\psi_\lambda(t)$ for the operator $\partial_t + a(x,t)\partial_x^2$ as:

$$\psi_\lambda(t) = e^{(T_1+1-t)^\lambda},\ t \in (0, T_1). \tag{2.4.8}$$

Hence, the function $\psi_\lambda(t)$ is decreasing on $[0, T_1]$, $\psi'_\lambda(t) < 0$,

$$\max_{[0,T_1]} \psi_\lambda(t) = \psi_\lambda(0) = e^{(T_1+1)^\lambda},\ \ \min_{[0,T_1]} \psi_\lambda(t) = \psi_\lambda(T_1) = e. \tag{2.4.9}$$

Denote

$$H_0^2(Q_{T_1}) = \left\{u \in H^2(Q_{T_1}) : u(0,t) = u(1,t) = 0\right\}. \tag{2.4.10}$$

$$H_{0,0}^2(Q_{T_1}) = \left\{u \in H_0^2(Q_{T_1}) : u(x,0) = 0\right\}. \tag{2.4.11}$$

Theorem 2.4.1 (Carleman estimate). *Let the coefficient $a(x,t)$ of the operator P satisfies conditions (2.4.1). Then there exist a sufficiently large number $\lambda_0 = \lambda_0(T_1, a_0, a_1) > 2$ and a constant $C = C(T_1, a_0, a_1) > 0$, both depending only on listed parameters, such that the following Carleman estimate holds for the operator P :*

$$\int_{Q_{T_1}} (Pu)^2 \psi_\lambda^2 dxdt \geq C\sqrt{\lambda} \int_{Q_{T_1}} u_x^2 \psi_\lambda^2 dxdt + C\lambda^2 \int_{Q_{T_1}} u^2 \psi_\lambda^2 dxdt$$

$$-C\sqrt{\lambda}\,\|u\|_{H^2(Q_{T_1})}^2 - C\lambda(T_1+1)^\lambda e^{2(T_1+1)^\lambda} \|u(x,0)\|_{L_2(0,1)}^2, \tag{2.4.12}$$

$$\forall \lambda \geq \lambda_0, \forall u \in H_0^2(Q_{T_1}).$$

Carleman estimate (2.4.12) is the key to proofs of Theorems 2.4.2, 2.4.4.

Theorem 2.4.2 (Hölder stability estimate for Problem 3 and uniqueness). *Let the coefficient $a(x,t)$ of the operator P satisfies conditions (2.4.1). Assume that the functions $w \in H^2(Q_{T_1})$ and $w^* \in H^2(Q_{T_1})$ are solutions of Problem 3 with the vectors of data $(\varphi_0(t), \varphi_1(t))$ and $(\varphi_0^*(t), \varphi_1^*(t))$ respectively, where $\varphi_0, \varphi_1, \varphi_0^*, \varphi_1^* \in H^2(0, T_1)$. Also, assume that error estimates (2.4.5) of the boundary data hold. Choose an arbitrary number $\rho \in (0, T_1)$. Denote*

$$\mu = \mu(T_1, \rho) = \frac{\ln(T_1+1-\rho)}{\ln(T_1+1)} \in (0,1). \tag{2.4.13}$$

Then there exists a sufficiently small number $\delta_0 = \delta_0 (T_1, a_0, a_1) \in (0,1)$ *and a constant* $C_1 = C_1 (T_1, a_0, a_1, \rho) > 0$, *both depending only on listed parameters, such that the following stability estimate holds for all* $\delta \in (0, \delta_0)$:

$$\|w_x - w_x^*\|_{L_2(Q_{T_1-\rho})} + \|w - w^*\|_{L_2(Q_{T_1-\rho})} \leq \tag{2.4.14}$$

$$\leq C_1 \left(1 + \|w - w^*\|_{H^2(Q_{T_1})}\right) \exp\left[-\left(\ln \delta^{-1/2}\right)^{\mu}\right].$$

Below $C = C(T_1, a_0, a_1) > 0$ and $C_1 = C_1 (T_1, a_0, a_1) > 0$ denote different constants depending only on listed parameters.

Corollary 2.4.1 (uniqueness). *Let the coefficient* $a(x,t)$ *of the operator* P *satisfies conditions (2.4.1). Then Problem 3 has at most one solution (uniqueness).*

Proof. If $\delta = 0$, then (2.4.14) implies that $w(x,t) = w^*(x,t)$ in $Q_{T_1-\rho}$. Since $\rho \in (0, T_1)$ is an arbitrary number, then $w(x,t) \equiv w^*(x,t)$ in Q_{T_1}. □

Theorem 2.4.3 (existence and uniqueness of the minimizer). *Let functions* $\varphi_0(t), \varphi_1(t) \in H^2(0, T_1)$. *Let* Y *be the set defined in (2.4.7). Then there exists unique minimizer* $w_{\min} \in Y$ *of functional (2.4.6) and*

$$\|w_{\min}\|_{H^2(Q_{T_1})} \leq \frac{C}{\sqrt{\alpha}} \left(\|\varphi_0\|_{H^2(0,T_1)} + \|\varphi_1\|_{H^2(0,T_1)}\right). \tag{2.4.15}$$

In the theory of Ill-Posed Problems, this minimizer $w_{\min}$ is called "regularized solution" of Problem 3 [35]. According to the theory of Ill-Posed problems, it is important to establish convergence rate of regularized solutions to the exact one w^*. In doing so, one should always choose a dependence of the regularization parameter α on the noise level δ, i.e. $\alpha = \alpha(\delta) \in (0,1)$ [35].

Theorem 2.4.4 (convergence rate of regularized solutions). *Let* $w^* \in H^2(Q_{T_1})$ *be the solution of Problem 3 with the noiseless data* $(\varphi_0^*(t), \varphi_1^*(t))$. *Let functions* $\varphi_0, \varphi_1, \varphi_0^*, \varphi_1^* \in H^2(0, T_1)$. *Let* $w_{\min} \in Y$ *be the unique minimizer of functional (2.4.6) on the set* Y. *Assume that error estimates (2.4.5) hold. Choose an arbitrary number* $\rho \in (0, T_1)$. *Let* $\mu = \mu(T_1, \rho) \in (0,1)$ *be the number defined in (2.4.13) and let*

$$\alpha = \alpha(\delta) = \delta^2, \tag{2.4.16}$$

Then there exists a sufficiently small number $\delta_0 = \delta_0(T_1, a_0, a_1) \in (0,1)$ depending only on listed parameters such that the following convergence rate of regularized solutions $w_{\min}$ holds for all $\delta \in (0, \delta_0)$:

$$\|\partial_x w_{\min} - \partial_x w^*\|_{L_2(Q_{T_1-\rho})} + \|w_{\min} - w^*\|_{L_2(Q_{T_1-\rho})} \tag{2.4.17}$$

$$\leq C_1 \left(1 + \|w^*\|_{H^2(Q_{T_1})} + \|\varphi_0^*\|_{H^2(0,T_1)} + \|\varphi_1^*\|_{H^2(0,T_1)}\right) \exp\left[-\left(\ln \delta^{-1/2}\right)^{\mu}\right].$$

2.4.3 Proof of Theorem 2.4.1

We assume in this proof that $u \in C^2\left(\overline{Q}_{T_1}\right)$. The case $u \in H^2(Q_{T_1})$ can be obtained via density arguments. It is assumed in this proof that $\lambda \geq \lambda_0 = \lambda_0(T_1, a_0, a_1) > 2$ and λ_0 is sufficiently large. We remind that $C = C(T_1, a_0, a_1) > 0$ denotes different constants depending only on listed parameters. Change variables as

$$v(x,t) = u(x,t)\,\psi_\lambda(t) = u(x,t)\,e^{(T_1+1-t)^\lambda}. \tag{2.4.18}$$

Hence,

$$u(x,t) = v(x,t)\,e^{-(T_1+1-t)^\lambda},$$

$$u_t = \left(v_t + \lambda(T_1+1-t)^{\lambda-1} v\right) e^{-(T_1+1-t)^\lambda},$$

$$u_x = v_x e^{-(T_1+1-t)^\lambda},\ u_{xx} = v_{xx} e^{-(T_1+1-t)^\lambda}.$$

Hence,

$$(Pu)^2 \psi_\lambda^2 = \left[v_t + \left(a(x,t)\,v_{xx} + \lambda(T_1+1-t)^{\lambda-1} v\right)\right]^2$$

$$\geq v_t^2 + 2v_t\left(a(x,t)\,v_{xx} + \lambda(T_1+1-t)^{\lambda-1} v\right). \tag{2.4.19}$$

We have used here $(a+b)^2 \geq a^2 + 2ab$, $\forall a, b \in \mathbb{R}$. We now estimate from the below terms in the second line of (2.4.19).

Step 1. Estimate from the below $2a(x,t)\,v_{xx}v_t$. We have:

$$2a(x,t)\,v_{xx}v_t = (2a(x,t)\,v_xv_t)_x - 2a(x,t)\,v_xv_{xt} - 2a_x(x,t)\,v_xv_t$$

$$= (2a(x,t)\,v_xv_t)_x + \left(-a(x,t)\,v_x^2\right)_t - a_t(x,t)\,v_x^2 - 2a_x(x,t)\,v_xv_t.$$

Thus,

$$2a\left(x,t\right)v_{xx}v_t \geq \left(2a\left(x,t\right)v_xv_t\right)_x + \left(-a\left(x,t\right)v_x^2\right)_t - Cv_x^2 - C\left|v_x\right|\left|v_t\right|. \tag{2.4.20}$$

Step 2. Estimate from the below $2\lambda\left(T_1+1-t\right)^{\lambda-1}vv_t$. We have:

$$\begin{aligned}2\lambda\left(T_1+1-t\right)^{\lambda-1}vv_t &= \left(\lambda\left(T_1+1-t\right)^{\lambda-1}v^2\right)_t + \lambda\left(\lambda-1\right)\left(T_1+1-t\right)^{\lambda-2}v^2 \\ &\geq \left(\lambda\left(T_1+1-t\right)^{\lambda-1}v^2\right)_t + \frac{\lambda^2}{2}\left(T_1+1-t\right)^{\lambda-2}v^2.\end{aligned} \tag{2.4.21}$$

Step 3. Estimate from the below the entire second line of (2.4.19). Using (2.4.20), (2.4.21) and Cauchy-Schwarz inequality "with ε",

$$2ab \geq -\varepsilon a^2 - \frac{1}{\varepsilon}b^2,\ \forall a,b\in\mathbb{R},\ \forall\varepsilon>0, \tag{2.4.22}$$

we obtain

$$\begin{aligned}&v_t^2 + 2v_t\left(a\left(x,t\right)v_{xx} + \lambda\left(T_1+1-t\right)^{\lambda-1}v\right) \geq \\ &\geq v_t^2 - Cv_x^2 - C\left|v_x\right|\left|v_t\right| + \frac{\lambda^2}{2}\left(T_1+1-t\right)^{\lambda-2}v^2 \\ &+ \left(2a\left(x,t\right)v_xv_t\right)_x + \left(-a\left(x,t\right)v_x^2 + \lambda\left(T_1+1-t\right)^{\lambda-1}v^2\right)_t \\ &\geq \frac{1}{2}v_t^2 - Cv_x^2 + \frac{\lambda^2}{2}\left(T_1+1-t\right)^{\lambda-2}v^2 \\ &+ \left(2a\left(x,t\right)v_xv_t\right)_x + \left(-a\left(x,t\right)v_x^2 + \lambda\left(T_1+1-t\right)^{\lambda-1}v^2\right)_t.\end{aligned}$$

Thus, we have obtained that

$$\begin{aligned}&v_t^2 + 2v_t\left(a\left(x,t\right)v_{xx} + \lambda\left(T_1+1-t\right)^{\lambda-1}v\right) \geq \\ &\geq \frac{1}{2}v_t^2 - Cv_x^2 + \frac{\lambda^2}{2}\left(T_1+1-t\right)^{\lambda-2}v^2 \\ &+ \left(2a\left(x,t\right)v_xv_t\right)_x + \left(-a\left(x,t\right)v_x^2 + \lambda\left(T_1+1-t\right)^{\lambda-1}v^2\right)_t.\end{aligned} \tag{2.4.23}$$

Using (2.4.19) and (2.4.23) as well as dropping the non-negative term $v_t^2/2$ in the right hand side of (2.4.23), we obtain

$$\left(Pu\right)^2\psi_\lambda^2 \geq -Cv_x^2 + \frac{\lambda^2}{2}\left(T_1+1-t\right)^{\lambda-2}v^2 \tag{2.4.24}$$

$$+\left(2a\left(x,t\right)v_xv_t\right)_x+\left(-a\left(x,t\right)v_x^2+\lambda\left(T_1+1-t\right)^{\lambda-1}v^2\right)_t.$$

Step 4. Using (2.4.18), change variables in the right hand side of (2.4.24). We have $v^2=u^2\psi_\lambda^2, v_x^2=u_x^2\psi_\lambda^2$. Thus,

$$\left(Pu\right)^2\psi_\lambda^2\geq -Cu_x^2\psi_\lambda^2+\frac{\lambda^2}{2}\left(T_1+1-t\right)^{\lambda-2}u^2\psi_\lambda^2 \tag{2.4.25}$$

$$+\left(2a\left(x,t\right)u_x\left(u_t-\lambda\left(T_1+1-t\right)^{\lambda-2}u\right)\psi_\lambda^2\right)_x+\left(\left(-a\left(x,t\right)u_x^2+\lambda\left(T_1+1-t\right)^{\lambda-1}u^2\right)\psi_\lambda^2\right)_t.$$

Step 5. Estimate from the below $-Pu\cdot u\psi_\lambda^2$. We have

$$-Pu\cdot u\psi_\lambda^2=\left(-u_t-a\left(x,t\right)u_{xx}\right)ue^{2(T_1+1-t)^\lambda}$$

$$=\left(-\frac{1}{2}u^2e^{2(T_1+1-t)^\lambda}\right)_t-\lambda\left(T_1+1-t\right)^{\lambda-1}u^2e^{2(T_1+1-t)^\lambda} \tag{2.4.26}$$

$$+\left(-a\left(x,t\right)u_xue^{2(T_1+1-t)^\lambda}\right)_x+a\left(x,t\right)u_x^2e^{2(T_1+1-t)^\lambda}+a_x\left(x,t\right)u_xue^{2(T_1+1-t)^\lambda}.$$

Using (2.4.1) and (2.4.22), we obtain

$$a\left(x,t\right)u_x^2+a_x\left(x,t\right)u_xu\geq a_0u_x^2-a_1\left|u_x\right|\left|u\right|\geq\frac{a_0}{2}u_x^2-Cu^2$$

$$\geq\frac{a_0}{2}u_x^2-\lambda\left(T_1+1-t\right)^{\lambda-2}u^2.$$

Hence, multiplying (2.4.26) by $\sqrt{\lambda}$, we obtain

$$-\sqrt{\lambda}Pu\cdot u\psi_\lambda^2\geq\frac{a_0}{2}\sqrt{\lambda}u_x^2e^{2(T_1+1-t)^\lambda}-2\lambda^{3/2}\left(T_1+1-t\right)^{\lambda-2}u^2e^{2(T_1+1-t)^\lambda} \tag{2.4.27}$$

$$+\left(-\frac{\sqrt{\lambda}}{2}u^2e^{2(T_1+1-t)^\lambda}\right)_t+\left(-\sqrt{\lambda}a\left(x,t\right)u_xue^{2(T_1+1-t)^\lambda}\right)_x.$$

Step 6. Estimate from the below $\left(Pu\right)^2\psi_\lambda^2-\sqrt{\lambda}Pu\cdot u\psi_\lambda^2$. Using (2.4.25) and (2.4.27), we obtain

$$\left(Pu\right)^2\psi_\lambda^2-\sqrt{\lambda}Pu\cdot u\psi_\lambda^2\geq$$

$$\geq\frac{a_0}{2}\sqrt{\lambda}\left(1-\frac{2C}{\sqrt{\lambda}}\right)u_x^2\psi_\lambda^2+\frac{\lambda^2}{2}\left(T_1+1-t\right)^{\lambda-2}\left(1-\frac{4}{\sqrt{\lambda}}\right)u^2\psi_\lambda^2$$

$$+\frac{\partial}{\partial t}\left[\left(-a\left(x,t\right)u_x^2+\lambda\left(T_1+1-t\right)^{\lambda-1}u^2-\frac{\sqrt{\lambda}}{2}u^2\right)\psi_\lambda^2\right] \tag{2.4.28}$$

$$+\frac{\partial}{\partial x}\left[\left(2a\left(x,t\right)u_x\left(u_t-\lambda\left(T_1+1-t\right)^{\lambda-2}u\right)-\sqrt{\lambda}a\left(x,t\right)u_xu\right)\psi_\lambda^2\right].$$

Step 7. Estimate from the below

$$\int_{Q_{T_1}} (Pu)^2 \psi_\lambda^2 dxdt.$$

We have

$$(Pu)^2 \psi_\lambda^2 - \sqrt{\lambda} Pu \cdot u\psi_\lambda^2 \leq \frac{3}{2}(Pu)^2 \psi_\lambda^2 + \frac{1}{2}\sqrt{\lambda} u^2 \psi_\lambda^2.$$

Combining this with (2.4.28), we obtain

$$\begin{aligned}(Pu)^2 \psi_\lambda^2 &\geq C\sqrt{\lambda} u_x^2 \psi_\lambda^2 + C\lambda^2 u^2 \psi_\lambda^2 \\ &+\frac{\partial}{\partial t}\left[\left(-a(x,t)u_x^2 + \lambda (T_1+1-t)^{\lambda-1} u^2 - \frac{\sqrt{\lambda}}{2} u^2\right)\psi_\lambda^2\right] \\ &+\frac{\partial}{\partial x}\left[\left(2a(x,t)u_x\left(u_t - \lambda(T_1+1-t)^{\lambda-2}u\right) - \sqrt{\lambda} a(x,t) u_x u\right)\psi_\lambda^2\right].\end{aligned} \tag{2.4.29}$$

Integrate (2.4.29) using $u \in H_0^2(Q_{T_1})$ and also using (2.4.9). We obtain

$$\begin{aligned}\int_{Q_{T_1}} (Pu)^2 \psi_\lambda^2 dxdt &\geq C\sqrt{\lambda}\int_{Q_{T_1}} u_x^2 \psi_\lambda^2 dxdt + C\lambda^2 \int_{Q_{T_1}} u^2 \psi_\lambda^2 dxdt \\ &-C\sqrt{\lambda}\|u(x,T_1)\|_{H^1(0,1)}^2 - C\lambda(T_1+1)^\lambda e^{2(T_1+1)^\lambda}\|u(x,0)\|_{L_2(0,1)}^2.\end{aligned} \tag{2.4.30}$$

Finally, applying the trace theorem to the second line of (2.4.30), we obtain desired estimate (2.4.12) of this theorem. □

2.4.4 Proof of Theorem 2.4.2

Introduce the following functions:

$$\widetilde{\varphi}_0(t) = \varphi_0(t) - \varphi_0^*(t), \widetilde{\varphi}_1(t) = \varphi_1(t) - \varphi_1^*(t), \tag{2.4.31}$$

$$F(x,t) = \varphi_0(t)(1-x) + \varphi_1(t)x, \ F^*(x,t) = \varphi_0^*(t)(1-x) + \varphi_1^*(t)x, \tag{2.4.32}$$

$$\widetilde{F}(x,t) = F(x,t) - F^*(x,t) = \widetilde{\varphi}_0(t)(1-x) + \widetilde{\varphi}_1(t), \tag{2.4.33}$$

$$\widehat{w}(x,t) = w(x,t) - F(x,t), \widehat{w}^*(x,t) = w^*(x,t) - F^*(x,t), \tag{2.4.34}$$

$$\overline{w}(x,t) = \widehat{w}(x,t) - \widehat{w}^*(x,t). \tag{2.4.35}$$

It follows from (2.4.4), (2.4.5) and (2.4.31)-(2.4.35) that:

$$\widehat{w}(x,0) = \widehat{w}^*(x,0) = \overline{w}(x,0) = 0, \tag{2.4.36}$$

$$F_{xx}(x,t) = F^{*}_{xx}(x,t) = \widetilde{F}_{xx}(x,t) = 0, \tag{2.4.37}$$

$$\left\|\widetilde{F}_t\right\|_{L_2(Q_{T_1})}, \left\|\widetilde{F}\right\|_{L_2(Q_{T_1})} \leq C\delta. \tag{2.4.38}$$

By (2.4.2)-(2.4.4) and (2.4.31)-(2.4.35)

$$\overline{w}_t + a(x,t)\,\overline{w}_{xx} = -\widetilde{F}_t \text{ in } Q_{T_1}, \tag{2.4.39}$$

$$\overline{w}(0,t) = \overline{w}(1,t) = 0,\ t \in (0,T_1), \tag{2.4.40}$$

$$\overline{w}(x,0) = 0,\ x \in (0,1). \tag{2.4.41}$$

Also, by (2.4.10), (2.4.11) and (2.4.36)

$$\widehat{w}, \widehat{w}^{*}, \overline{w} \in H^{2}_{0,0}(Q_{T_1}). \tag{2.4.42}$$

Square both sides of equation (2.4.39), multiply by the function $\psi^2_\lambda(t)$ and integrate over the domain Q_{T_1}. Using (2.4.9) and (2.4.38), we obtain

$$\int_{Q_{T_1}} \left(\overline{w}_t + a(x,t)\,\overline{w}_{xx}\right)^2 \psi^2_\lambda(t)\,dxdt \leq C\delta^2 e^{2(T_1+1)^\lambda}. \tag{2.4.43}$$

Hence, applying Carleman estimate (2.4.12) to the left hand side of (2.4.43) and taking into account (2.4.9)-(2.4.11), we obtain

$$\int_{Q_{T_1}} \overline{w}^2_x \psi^2_\lambda dxdt + \lambda^{3/2} \int_{Q_{T_1}} \overline{w}^2 \psi^2_\lambda dxdt \leq \tag{2.4.44}$$

$$\leq C\delta^2 e^{2(T_1+1)^\lambda} + C\left\|\overline{w}\right\|^2_{H^2(Q_{T_1})},\ \forall\lambda \geq \lambda_0.$$

Since $Q_{T_1-\rho} \subset Q_{T_1}$ and also since by (2.4.8) $\psi^2_\lambda(t) \geq e^{2(T_1+1-\rho)^\lambda}$ in $Q_{T_1-\rho}$, then (2.4.44) implies

$$\left\|\overline{w}_x\right\|^2_{L_2(Q_{T_1-\rho})} + \left\|\overline{w}_x\right\|^2_{L_2(Q_{T_1-\rho})} \leq \tag{2.4.45}$$

$$\leq C\delta e^{(T_1+1)^\lambda} + Ce^{-(T_1+1-\rho)^\lambda}\left\|\overline{w}\right\|_{H^2(Q_{T_1})},\ \forall\lambda \geq \lambda_0.$$

Choose $\delta_0 = \delta_0(T_1, a_0, a_1) \in (0,1)$ so small that

$$\ln\left[\ln\left(\delta_0^{-1/2}\right)^{1/\ln(T_1+1)}\right] > \lambda_0. \tag{2.4.46}$$

Let $\delta \in (0, \delta_0)$. We now choose $\lambda = \lambda(\delta)$ so large that

$$e^{(T_1+1)^\lambda} = \frac{1}{\sqrt{\delta}}. \tag{2.4.47}$$

Hence,

$$\lambda = \lambda(\delta) = \ln\left[\ln\left(\delta^{-1/2}\right)^{1/\ln(T_1+1)}\right] > \lambda_0, \ \forall \delta \in (0, \delta_0). \tag{2.4.48}$$

Then

$$e^{-(T_1+1-\rho)^\lambda} = \exp\left[-\left(\ln \delta^{-1/2}\right)^\mu\right], \tag{2.4.49}$$

where the number $\mu \in (0,1)$ is defined in (2.4.13). We have

$$\frac{e^{-\left(\ln \delta^{-1/2}\right)^\mu}}{\sqrt{\delta}} = \exp\left[-\frac{1}{2}\ln \delta\left(1 + \frac{2\left(\ln \delta^{-1/2}\right)^\mu}{\ln \delta}\right)\right]. \tag{2.4.50}$$

Since $\mu \in (0,1)$, then the Hospital's rule implies

$$\lim_{\delta \to 0} \frac{2\left(\ln \delta^{-1/2}\right)^\mu}{\ln \delta} = \lim_{\delta \to 0}\left(-\mu\left(\ln \delta^{-1/2}\right)^{\mu-1}\right) = 0.$$

Hence,

$$\lim_{\delta \to 0}\left[-\frac{1}{2}\ln \delta\left(1 + \frac{2\left(\ln \delta^{-1/2}\right)^\mu}{\ln \delta}\right)\right] = \lim_{\delta \to 0}\left(\ln \delta^{-1/2}\right) = \infty. \tag{2.4.51}$$

It follows from (2.4.50) and (2.4.51) that

$$\lim_{\delta \to 0} \frac{e^{-\left(\ln \delta^{-1/2}\right)^\mu}}{\sqrt{\delta}} = \infty.$$

Hence,

$$\sqrt{\delta} \le C_1 e^{-\left(\ln \delta^{-1/2}\right)^\mu}, \ \forall \delta \in (0,1). \tag{2.4.52}$$

Using (2.4.45)-(2.4.49) and (2.4.52), we obtain

$$\|\overline{w}_x\|_{L_2(Q_{T_1-\rho})} + \|\overline{w}\|_{L_2(Q_{T_1-\rho})} \le \tag{2.4.53}$$

$$\le C_1\left(1 + \|\overline{w}\|_{H^2(Q_{T_1})}\right)\exp\left(-\left(\ln \delta^{-1/2}\right)^\mu\right), \ \forall \delta \in (0, \delta_0).$$

By (2.4.31)-(2.4.35) $\overline{w} = (w - w^*) - \widetilde{F}$. Hence, the triangle inequality, (2.4.5), (2.4.31)-(2.4.33), (2.4.38) and (2.4.53) imply (2.4.14), which is the target estimate of this theorem.

□

2.4.5 Proof of Theorem 2.4.3

Denote $[,]$ the scalar product in the space $H^2(Q_{T_1})$. Let $\widehat{w} \in H^2_{0,0}(Q_{T_1})$ be the function defined in (2.4.34). Then, using (2.4.6) and (2.4.37), consider the functional

$$I_\alpha(\widehat{w}+F) = \int_{Q_{T_1}} (P\widehat{w}+F_t)^2\, dxdt + \alpha \|\widehat{w}+F\|^2_{H^2(Q_{T_1})}. \tag{2.4.54}$$

Suppose that the function $\widehat{w}_{\min} \in H^2_{0,0}(Q_{T_1})$ is a minimizer of the functional $I_\alpha(\widehat{w}+F)$ on the space $H^2_{0,0}(Q_{T_1})$, i.e.

$$I_\alpha(\widehat{w}_{\min}+F) \le I_\alpha(\widehat{w}+F),\ \forall \widehat{w} \in H^2_{0,0}(Q_{T_1}). \tag{2.4.55}$$

Consider the function $w_{\min} = \widehat{w}_{\min} + F$. Then it follows from (2.4.11) and (2.4.32) that $w_{\min} \in Y$, where the set Y is defined in (2.4.7). Also, $w = \widehat{w} + F \in Y$, $\forall \widehat{w} \in H^2_{0,0}(Q_{T_1})$. Hence, (2.4.55) implies that the function $w_{\min}$ is a minimizer of the functional $I_\alpha(w)$ on the set Y.

We now prove the reverse. Suppose that a function $w^{\min} \in Y$ is a minimizer of the functional $I_\alpha(w)$ on the set Y, i.e.

$$I_\alpha\left(w^{\min}\right) \le I_\alpha(w),\ \forall w \in Y. \tag{2.4.56}$$

Consider the function $\widehat{w}^{\min} = w^{\min} - F$. And for every function $\widehat{w} \in H^2_{0,0}(Q_{T_1})$ consider the function $w = \widehat{w} + F \in Y$. Then by (2.4.56)

$$I_\alpha\left(\widehat{w}^{\min}+F\right) = I_\alpha\left(w^{\min}\right) \le I_\alpha(w) = I_\alpha(\widehat{w}+F),\ \forall \widehat{w} \in H^2_{0,0}(Q_{T_1}).$$

Hence, $\widehat{w}^{\min}$ is a minimizer of functional (2.4.54) on the space $H^2_{0,0}(Q_{T_1})$. Therefore, it is sufficient to find a minimizer of the functional $I_\alpha(\widehat{w}+F)$ on the space $H^2_{0,0}(Q_{T_1})$.

By the variational principle the function $\widehat{w}_{\min} \in H^2_{0,0}(Q_{T_1})$ is a minimizer of the functional $I_\alpha(\widehat{w}+F)$ if and only if the following integral identity is satisfied:

$$\int_{Q_{T_1}} (P\widehat{w}_{\min} \cdot Ph)\, dxdt + \alpha [\widehat{w}_{\min}, h] = -\int_{Q_{T_1}} F_t \cdot Ph dxdt - \alpha [F,h], \tag{2.4.57}$$

$$\forall h \in H^2_{0,0}(Q_{T_1}).$$

Consider a new scalar product in $H_{0,0}^{2}(Q_{T_1})$ defined as

$$\{u,v\}=\int_{Q_{T_1}}(Pu\cdot Pv)\,dxdt+\alpha\,[u,v]\,,\ \forall u,v\in H_{0,0}^{2}(Q_{T_1})\,.$$

Recall that $\alpha\in(0,1)$. Obviously,

$$\alpha\,\|u\|_{H^2(Q_{T_1})}^2\leq\{u,u\}\leq C\,\|u\|_{H^2(Q_{T_1})}^2\,,\forall u\in H_{0,0}^{2}(Q_{T_1})\,.$$

Hence, norms $\sqrt{\{u,u\}}$ and $\|u\|_{H^2(Q_{T_1})}$ are equivalent. Hence, one can consider the scalar product $\{u,v\}$ as the scalar product in $H_{0,0}^{2}(Q_{T_1})$.

Hence, we can rewrite (2.4.57) as

$$\{\widehat{w}_{\min},h\}=-\int_{Q_{T_1}}F_t\cdot Phdxdt-\alpha\,[F,h]\,,\ \forall h\in H_{0,0}^{2}(Q_{T_1})\,. \tag{2.4.58}$$

The right hand side of (2.4.58) can be estimated as

$$\left|-\int_{Q_{T_1}}F_tPhdxdt-\alpha\,[F,h]\right|\leq C\,\|F\|_{H^2(Q_{T_1})}\,\{h\}\,,\ \forall h\in H_{0,0}^{2}(Q_{T_1})\,.$$

Hence, the right hand side of (2.4.58) can be considered as a bounded linear functional of $h\in H_{0,0}^{2}(Q_{T_1})$. Hence, by Riesz theorem there exists unique function $W\in H_{0,0}^{2}(Q_{T_1})$ such that

$$-\int_{Q_{T_1}}F_tPhdxdt-\alpha\,[F,h]=\{W,h\}\,,\ \forall h\in H_{0,0}^{2}(Q_{T_1})\,.$$

Comparing this with (2.4.58), we obtain

$$\{\widehat{w}_{\min},h\}=\{W,h\}\,,\ \forall h\in H_{0,0}^{2}(Q_{T_1})\,.$$

Therefore, $\widehat{w}_{\min}=W$. Thus, we have proven existence and uniqueness of the minimizer $\widehat{w}_{\min}$ of the functional $I_\alpha(\widehat{w}+F)$ on the space $H_{0,0}^{2}(Q_{T_1})$. Therefore, it follows from the discussion in the beginning of this proof that there exists unique minimizer of the functional $I_\alpha(w)$ on the set Y and this minimizer is $w_{\min}=\widehat{w}_{\min}+F$.

We now estimate the norm $\|w_{\min}\|_{H^2(Q_{T_1})}$. Setting in (2.4.57) $h=\widehat{w}_{\min}$ and using Cauchy-Schwarz inequality, we obtain

$$\int_{Q_{T_1}}(P\widehat{w}_{\min})^2\,dxdt+\alpha\,\|\widehat{w}_{\min}\|_{H^2(Q_{T_1})}^2\leq$$

$$\leq \frac{1}{2}\left\|F_t\right\|^2_{L_2\left(Q_{T_1}\right)} + \frac{1}{2}\left\|P\widehat{w}_{\min}\right\|^2_{H^2\left(Q_{T_1}\right)} + \frac{\alpha}{2}\left\|F\right\|^2_{H^2\left(Q_{T_1}\right)} + \frac{\alpha}{2}\left\|\widehat{w}_{\min}\right\|^2_{H^2\left(Q_{T_1}\right)}.$$

Hence,

$$\left\|\widehat{w}_{\min}\right\|_{H^2\left(Q_{T_1}\right)} \leq \frac{C}{\sqrt{\alpha}}\left\|F\right\|_{H^2\left(Q_{T_1}\right)}.$$

This estimate, triangle inequality and (2.4.32) imply the target estimate (2.4.15) of Theorem 2.4.3. □

2.4.6 Proof of Theorem 2.4.4

We still use notations (2.4.31)-(2.4.35). By Corollary 2.4.1 Problem 3 has at most one solution. Hence, there exists unique exact solution $w^* \in H^2\left(Q_{T_1}\right)$ of Problem 3 with the data $\varphi_0^*, \varphi_1^* \in H^2\left(0, T_1\right)$ in (2.4.3) and (2.4.4). Hence, we have the following analog of integral identity (2.4.57)

$$\int_{Q_{T_1}} \left(P\widehat{w}^* \cdot Ph\right) dxdt + \alpha\left[\widehat{w}^*, h\right] = \tag{2.4.59}$$

$$= -\int_{Q_{T_1}} F_t^* \cdot Phdxdt + \alpha\left[\widehat{w}^*, h\right],\ \forall h \in H^2_{0,0}\left(Q_{T_1}\right).$$

Subtract (2.4.59) from (2.4.57). Then, using (2.4.33), (2.4.35), we obtain

$$\int_{Q_{T_1}} \left(P\overline{w} \cdot Ph\right) dxdt + \alpha\left[\overline{w}, h\right] = \tag{2.4.60}$$

$$= -\int_{Q_{T_1}} \widetilde{F}_t \cdot Phdxdt + \alpha\left[\widehat{w}^*, h\right],\ \forall h \in H^2_{0,0}\left(Q_{T_1}\right).$$

Set in (2.4.60) $h = \overline{w}$. Then, using (2.4.38) and Cauchy-Schwarz inequality, we obtain

$$\int_{Q_{T_1}} \left(P\overline{w}\right)^2 dxdt \leq C\left(\delta^2 + \alpha\left\|\widehat{w}^*\right\|^2_{H^2\left(Q_{T_1}\right)}\right), \tag{2.4.61}$$

$$\left\|\overline{w}\right\|_{H^2\left(Q_{T_1}\right)} \leq C\left(\frac{\delta}{\sqrt{\alpha}} + \left\|\widehat{w}^*\right\|_{H^2\left(Q_{T_1}\right)}\right). \tag{2.4.62}$$

Inequality (2.4.61) is equivalent with

$$\int_{Q_{T_1}} \left(P\overline{w}\right)^2 \psi_\lambda^2 \psi_\lambda^{-2} dxdt \leq C\left(\delta^2 + \alpha\left\|\widehat{w}^*\right\|^2_{H^2\left(Q_{T_1}\right)}\right).$$

Since by (2.4.9) $\psi_\lambda^{-2}(t) \geq e^{-2(T_1+1)^\lambda}$ in Q_{T_1}, then (2.4.61) implies

$$\int_{Q_{T_1}} \left(P\overline{w}\right)^2 \psi_\lambda^2 dxdt \leq C\left(\delta^2 + \alpha\left\|\widehat{w}^*\right\|^2_{H^2\left(Q_{T_1}\right)}\right) e^{2(T_1+1)^\lambda}. \tag{2.4.63}$$

Hence, applying Carleman estimate (2.4.12) to the left hand side of (2.4.63) and recalling again that $\alpha \in (0,1)$, we obtain

$$\int_{Q_{T_1}} \overline{w}_x^2 \psi_\lambda^2 dxdt + \lambda^{3/2} \int_{Q_{T_1}} \overline{w}^2 \psi_\lambda^2 dxdt \leq$$

$$\leq C \left(\delta^2 + \alpha \left\| \widehat{w}^* \right\|_{H^2(Q_{T_1})}^2 \right) e^{2(T_1+1)^\lambda} + C \left\| \overline{w} \right\|_{H^2(Q_{T_1})}^2, \ \forall \lambda \geq \lambda_0.$$

Hence, we obtain similarly with (2.4.45)

$$\left\| \overline{w}_x \right\|_{L_2(Q_{T_1-\rho})}^2 + \left\| \overline{w} \right\|_{L_2(Q_{T_1-\rho})}^2 \leq C \left(\delta^2 + \alpha \left\| \widehat{w}^* \right\|_{H^2(Q_{T_1})}^2 \right) e^{2(T_1+1)^\lambda}$$

$$+ C e^{-2(T_1+1-\rho)^\lambda} \left\| \overline{w} \right\|_{H^2(Q_{T_1})}^2, \ \forall \lambda \geq \lambda_0.$$

Combining this with (2.4.62), we obtain

$$\left\| \overline{w}_x \right\|_{L_2(Q_{T_1-\rho})} + \left\| \overline{w} \right\|_{L_2(Q_{T_1-\rho})} \leq C \left(\delta + \sqrt{\alpha} \left\| \widehat{w}^* \right\|_{H^2(Q_{T_1})} \right) e^{(T_1+1)^\lambda} \tag{2.4.64}$$

$$+ C \frac{\delta}{\sqrt{\alpha}} e^{-(T_1+1-\rho)^\lambda} + \left\| \widehat{w}^* \right\|_{H^2(Q_{T_1})} e^{-(T_1+1-\rho)^\lambda}, \ \forall \lambda \geq \lambda_0.$$

Suppose now that $\alpha = \alpha(\delta) = \delta^2$, as stated in (2.4.16). Choose $\delta_0 = \delta_0(T_1, a_0, a_1) \in (0,1)$ as in (2.4.46) and $\lambda = \lambda(\delta)$ as in (2.4.48). Then (2.4.47), (2.4.49), (2.4.52) and (2.4.64) imply

$$\left\| \overline{w}_x \right\|_{L_2(Q_{T_1-\rho})} + \left\| \overline{w} \right\|_{L_2(Q_{T_1-\rho})} \leq \tag{2.4.65}$$

$$\leq C_1 \left(1 + \left\| \widehat{w}^* \right\|_{H^2(Q_{T_1})} \right) \exp \left[- \left(\ln \delta^{-1/2} \right)^\mu \right], \ \forall \delta \in (0, \delta_0).$$

The target estimate (2.4.17) of this theorem follows immediately from the triangle inequality, (2.4.5), (2.4.31)-(2.4.35) and (2.4.65). □

2.5 Probabilistic Arguments for a Trading Strategy

A heuristic algorithm of section 2.3 can be used as the basis for a trading strategy of options. The algorithm predicts the option price change relatively to the current price. The fact that this algorithm uses the information about stock and option prices only over a small time period makes realistic the assumptions of the model of Section 2.2 about the volatilities being independent on time. Formulas (2.2.6) and (2.2.9) indicate that the sign of the mathematical expectation of the option price increment should likely define the trading strategy. In addition to the mathematical expectation of the option price increment, it is necessary to take into account indicators that reflect the risk of using that trading strategy. This is because the option price dynamics is described by a random process. Based on the model of Section 2.2, we construct in this section such indicators for a "perfect" trading strategy, which always correctly predicts the sign of the mathematical expectation of the option price increment.

We assume in this section that both the volatility σ of the stock and the idea $\hat{\sigma}$ of the volatility of the call option, which has been developed among the participants involved in trading of this option, are known. Recall that we have assumed in Section 2 that the dynamics of the stock price is described by a stochastic differential equation of the geometric Brownian motion $ds = \sigma s dW$ with the initial condition $s(t_0) = s_0$, and also that the corresponding option price is $v(s(t),t) = u(s(t), T - \tau)$, where $\tau = T - t \in (0, T)$ and the function $u(s,\tau)$ can be found by the Black-Scholes formula (2.2.5). The option price expected by option market participants is described by a stochastic process $v(\hat{s}(t),t) = u(\hat{s}(t), T - t)$, where the expected stock price satisfies the stochastic differential equation of the geometric Brownian motion $d\hat{s} = \hat{\sigma}\hat{s}d\hat{W}_1$ with the initial condition

$$\hat{s}(t_0) = s_0 = s(t_0). \tag{2.5.1}$$

Here W_1 is a Wiener process, and the processes W and W_1 are independent.

Let $t_0 \in (0,T)$ be a certain moment of time and $\varepsilon > 0$ be a sufficiently small number. The true option price at the moment of time $t_0+\varepsilon$ is $v(s(t_0+\varepsilon), t_0+\varepsilon)$. On the other hand, at the same moment of time $t_0 + \varepsilon$ the price of this option expected by the participants of the market is $v(\hat{s}(t_0 + \varepsilon), t_0 + \varepsilon)$. It follows from (2.2.6) that, on the small time interval $(t_0, t_0+\varepsilon)$, a winning trading strategy of the options trading should be based on an estimate of the probability that $v(s(t_0+\varepsilon), t_0+\varepsilon) > v(\hat{s}(t_0), t_0)$. This probability is given in Theorem 2.5.1.

Theorem 2.5.1. *Let $\varepsilon > 0$ be a sufficiently small number and $\Phi(z), z \in \mathbb{R}$ be the function defined in (2.2.4). The probability that at the time $t_0 + \varepsilon$ the true option price $v(s(t_0+\varepsilon), t_0+\varepsilon)$ is greater than the price expected by the participants of the options market*

$v(\hat{s}(t_0+\varepsilon), t_0+\varepsilon)$ *is*

$$p = \Phi\left(\frac{(\hat{\sigma}^2 - \sigma^2)\sqrt{\varepsilon}}{2\sqrt{(\hat{\sigma}^2 + \sigma^2)}}\right). \tag{2.5.2}$$

Proof. The derivative

$$\frac{\partial u(s,\tau)}{\partial s}$$

is called the Greek delta. This parameter for a call option is

$$\Delta = \frac{\partial u(s,\tau)}{\partial s} = \Phi(\Theta_+(s,\tau)) > 0. \tag{2.5.3}$$

Since by (2.5.3) $\partial_s u(s,\tau) > 0$, then the inequality $v(s(t_0+\varepsilon), t_0+\varepsilon) > v(\hat{s}(t_0+\varepsilon), t_0+\varepsilon)$ is equivalent to the inequality $s(t_0+\varepsilon) > \hat{s}(t_0+\varepsilon)$. It follows from (2.5.1) that the latter inequality is equivalent with

$$\ln\left(\frac{s(t_0+\varepsilon)}{s(t_0)}\right) > \ln\left(\frac{\hat{s}(t_0+\varepsilon)}{\hat{s}(t_0)}\right).$$

It follows from the properties of the geometric Brownian motion, see, e.g. [28, Chapter 5, section 5.1] that the random variables

$$\ln\left(\frac{s(t_0+\varepsilon)}{s(t_0)}\right) \in N\left(-\frac{\sigma^2}{2}\varepsilon, \sigma^2\varepsilon\right),$$

$$\ln\left(\frac{\hat{s}(t_0+\varepsilon)}{\hat{s}(t_0)}\right) \in N\left(-\frac{\hat{\sigma}^2}{2}\varepsilon, \hat{\sigma}^2\varepsilon\right)$$

are normally distributed. Hence, the difference of these two random variables is also a normally distributed random variable, see, e.g. [21, Chapter 9, section 9.3], i.e.

$$\left[\ln\left(\frac{s(t_0+\varepsilon)}{s(t_0)}\right) - \ln\left(\frac{\hat{s}(t_0+\varepsilon)}{\hat{s}(t_0)}\right)\right] \in N\left(\frac{\sigma^2-\hat{\sigma}^2}{2}\varepsilon, (\hat{\sigma}^2+\sigma^2)\varepsilon\right).$$

Therefore, the value given by formula (2.5.2) is indeed the probability that the true option price $v(s(t_0+\varepsilon), t_0+\varepsilon))$ is greater than the expected option price $v(\hat{s}(t_0+\varepsilon), t_0+\varepsilon)$. $\square$

Theorem 2.5.1 implies that the operation of buying an option at the time moment t_0 and selling it at the time moment $t_0+\varepsilon$ will be profitable with the probability p given in (2.5.2), and this operation will be non-profitable with the probability $1-p$.

By (2.2.4) and (2.5.2)

$$p = \begin{cases} > 1/2 \text{ if } \sigma > \hat{\sigma}, \\ \leq 1/2 \text{ if } \sigma \leq \hat{\sigma}. \end{cases}$$

Hence, if $\sigma > \hat{\sigma}$, then it is reasonable to buy an option at the moment of time t_0 and sell it at the moment of time $t_0+\varepsilon$. Otherwise, it is reasonable to go in the short position on the option at $t = t_0$. Suppose that $p > 1/2$. A winning strategy, which takes into account risks, should involve the repetition operation multiple times with the same independent probabilities of outcomes. Consider n non-overlapping small time intervals $[t_j, t_j + \varepsilon]$, $j = 1, ..., n$, of the same duration $\varepsilon > 0$, on which the option purchase operations are carried out at the moment of time t_j with the subsequent sale at the moment of time $t_j + \varepsilon$.

Consider random variables $\{\xi_j\}_{j=1}^{n}$,

$$\xi_j = \begin{cases} 1, & \text{if } v(s(t_j+\varepsilon), t_j+\varepsilon) \geq v(\hat{s}(t_j), t_j), \\ 0, & \text{if } v(s(t_j+\varepsilon), t_j+\varepsilon) < v(\hat{s}(t_j), t_j), \end{cases} \quad j = 1, ..., n. \tag{2.5.4}$$

If $\xi_j = 1$, then the operation of buying that option at the moment of time t_j and selling it at the moment of time $t_j + \varepsilon$ was profitable. If $\xi_j = 0$, then that operation was non profitable. The random variables $\xi_1, ..., \xi_n$ are independent identically distributed random variables [21, Chapter 18, section 18.1]. The frequency of profitable trading operations is characterized by the random variable ζ,

$$\zeta = \frac{1}{n} \sum_{j=1}^{n} \xi_j. \tag{2.5.5}$$

It follows from [21, Chapter 9, section 9.3] that the variable ζ has a binomial distribution with the mathematical expectation p given in (2.5.2) and with the dispersion D, where

$$D = \frac{p(1-p)}{n}. \tag{2.5.6}$$

By the Central Limit Theorem of de Moivre-Laplace, the probability that more than half of trades are profitable is estimated as [21, Chapter 2, section 2.2]:

$$\sum_{k=[\frac{n}{2}+2]}^{n} \frac{n!}{(k-1)!(n-k)!} p^k (1-p)^{n-k} = \Phi\left(\frac{(1-2p)\sqrt{n}}{2\sqrt{p(1-p)}}\right) + \delta(n), \tag{2.5.7}$$

$$\lim_{n\to\infty} \delta(n) = 0. \tag{2.5.8}$$

In our trading strategy, we decide to make transactions if and only if the probability of the profitable trading is not less than a given value $\alpha > 1/2$. Hence, by (2.5.7) and (2.5.8)

$$\frac{(1-2p)\sqrt{n}}{2\sqrt{p(1-p)}} > \Phi^{-1}\left(\alpha - \delta(n)\right), \tag{2.5.9}$$

where Φ^{-1} is the inverse function of the function Φ of (2.2.4). By (2.5.9)

$$p^2 - p + \frac{n}{4\left\{n + [\Phi^{-1}(\alpha - \delta(n)]^2\right\}} > 0.$$

Thus, we should have

$$p \geq \frac{1}{2}\left(1 + \sqrt{\frac{[\Phi^{-1}(\alpha - \delta(n)]^2}{n + [\Phi^{-1}(\alpha - \delta(n)]^2}}\right). \tag{2.5.10}$$

To fulfill inequality (2.5.10), the imperfection of the stock market must be significant. More precisely, it follows from (2.5.2) and (2.5.10) that the difference between the volatilities σ and $\hat{\sigma}$ must satisfy the following inequality:

$$\sigma - \hat{\sigma} \geq \frac{2\sqrt{\sigma^2 + \hat{\sigma}^2}}{\sqrt{\varepsilon}(\sigma + \hat{\sigma})}\left|\Phi^{-1}\left(\frac{1}{2}\left(1 + \sqrt{\frac{[\Phi^{-1}(\alpha - \delta(n)]^2}{n + [\Phi^{-1}(\alpha - \delta(n)]^2}}\right)\right)\right|. \tag{2.5.11}$$

Based on this estimate of the difference $\sigma - \hat{\sigma}$, we design a trading strategy in the ideal case. "Ideal" means that we know both volatilities σ and $\hat{\sigma}$.

Trading Strategy for the Ideal Case:

Let $\beta_1 > 0$ and $\beta_2 < 0$ be two threshold numbers. Our trading strategy considers three possible scenarios:

1. If $\sigma - \hat{\sigma} \geq \beta_1 > 0$, then it is recommended to buy a call option at the current moment of time t_j with the subsequent sale at the next moment of time $t_j + \varepsilon$.

2. If $\sigma - \hat{\sigma} \leq \beta_2$, then then it is recommended to go short at the current moment of time t_j, followed by closing the short position at time $t_j + \varepsilon$.

3. If $\beta_2 < \sigma - \hat{\sigma} < \beta_1$, then it is recommended to refrain from trading.

The threshold values β_1 and β_2 might probably estimated via numerical simulations using the method of section 2.3, combined with formula (2.5.11).

2.6 Numerical Studies

2.6.1 Some numerical details for the algorithm of section 2.3

We have computationally simulated the market data as described in subsection 2.6.1 These data gave us initial and boundary conditions (2.3.2), (2.3.3) and (2.3.5), which, in turn, led us to (2.3.14), (2.3.13), see Steps 1,2 of subsection 2.3.2. Next, we have solved Minimization Problem (2.3.16), (2.3.17). To minimize functional (2.3.16), we wrote Mv and $\|v\|^2_{H^2(G_{2y})}$ in finite differences and, using the conjugate gradient method, have minimized the resulting discrete functional with respect to the values of the function v at grid points. The starting point of the minimization procedure was $v = 0$. The regularization parameter $\gamma = 0.01$ was the same as in [17], and it was chosen by trial and error. The step sizes h_x and h_t of the finite difference scheme with respect to x and t were $h_x = 0.01$ and $h_t = 0.0000784$ respectively. Since by (2.3.6) and (2.3.10) $G_{2y} = \{(x,t) \in (0,1) \times (0, 0.00784)\}$, then we had 100 grid points with respect to each variable x and t.

2.6.2 The data

We construct the stock price trajectory $s(t)$ as a solution to the stochastic differential equation $ds = \sigma s dW$ with the initial condition $s(0) = 100$, where $\sigma = 0.2$. We model the stock prices and then the prices of 90-days European call options of this stock during the life of the stock, assuming that the options are reissued many times with the same maturity date of 90 days. Thus, we obtain a time series $\{s(t_k)\}_{k=1}^N$, $N \geq n = 2000$. We set the payoff function for each option $f(s) = (s - 100)_+$. The generated stock price trajectory is shown on Figure 2.1.

Figure 2.1: *An example of the time dependent behavior of stock prices generated by the geometric Brownian motion with* $\sigma = 0.2$.

The probabilistic analysis of section 2.5 of the random variable ζ characterizes the effectiveness of an "ideal" trading strategy. Thus, we consider the ideal case first. Recall that "ideal" means here that this strategy is based on the knowledge of the information about the imperfection of the financial market, i.e. on the knowledge of both volatilities σ and $\hat{\sigma}$. However, in the real market data only approximate values of $\hat{\sigma}$ are available [2].

We test total thirty three (33) values of $\hat{\sigma}$. More precisely, in our computational simulations, we took the discrete values of $\hat{\sigma}$, where

$$\hat{\sigma} \in [0.05, 0.38] \text{ with the step size } h_{\hat{\sigma}} = 0.01. \tag{2.6.1}$$

We now generate the function, which describes the dependence of the mathematical expectation of the random variable ζ in (2.5.5) on $\hat{\sigma}$. Keeping in mind that $\sigma = 0.2$ in all cases, we compute for each of the discrete values of $\hat{\sigma}$ in (2.6.1) the mathematical expectation $p(\hat{\sigma})$ of the random variable ζ. We use formula (2.5.2) for $p(\hat{\sigma})$, also, see (2.2.4). This way we obtain the function $p(\hat{\sigma})$, which is the above dependence for the ideal case.

Second, we consider a non-ideal case. More precisely, we test how our heuristic algorithm works for the computationally simulated data described in this subsection. We choose $n = 2000$ non-overlapping time intervals $[t_j, t_j + y]$, $j = 1, ..., n$, where $y = 1/255$ means one dimensionless trading day, see (2.1.2) and (2.3.6). We still use the dimensionless time as in (2.1.1), while keeping the same notation t for brevity. For every fixed value of $\hat{\sigma}$ indicated in (2.6.1), we calculate the option price $v(s(t_j), t_j) = u(s(T - t_j), T - t_j)$, where the function $u(s, \tau) = u(s, T - t)$ is given by Black-Scholes formula (2.2.3). Thus, numbers $v(s(t_j), t_j)$ form the option price trajectory. Figure 2 displays a sample of the trajectory of the option price for $\hat{\sigma} = 0.1$. Based on (2.3.1), we set bid and ask stock prices as well as corresponding bid and ask option prices as:

$$s_b(t_j) = 0.99 \cdot s(t_j) \text{ and } s_a(t_j) = 1.01 \cdot s(t_j),$$

$$v_b(s(t_j), t_j) = 0.99 \cdot v(s(t_j), t_j) \text{ and } v_a(s(t_j), t_j) = 1.01 \cdot v(s(t_j), t_j).$$

Next, we solve Problem 2 of section 2.3 for each j on the time interval $[t_j, t_j + 2y]$ by the algorithm of that section. When doing so, we take in (2.3.8) $\sigma^2(t) = \hat{\sigma}^2$ for $t \in [t_j, t_j + 2y]$ for all $j = 1, ..., 2000$. In particular, this solution via QRM gives us the function $v_{\text{comp}}(s, t_j + y)$, $s \in (s_b(t_j), s_a(t_j))$. We set the predicted price of the option at the moment of time $t_j + y$ as:

$$v_{\text{pred}}(t_j + y) = v_{\text{comp}}\left(\frac{s_b(t_j) + s_a(t_j)}{2}, t_j + y\right). \tag{2.6.2}$$

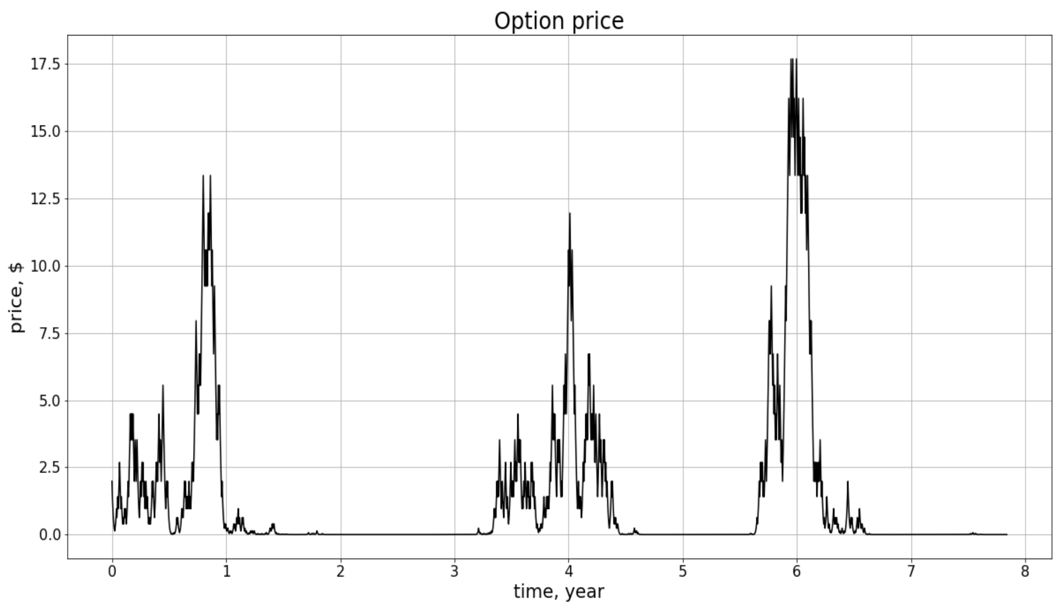

Figure 2.2: *The trajectory of the option price with* $\hat{\sigma} = 0.1$ *generated by the Black-Scholes formula (2.2.3) for the stock price trajectory of Figure 1. In this case,* σ *should be replaced with* $\hat{\sigma}$ *in (2.2.3).*

2.6.3 Results

For every discrete value of $\hat{\sigma}$ in (2.6.1), we introduce the sequence $\left\{\overline{\xi}_j\left(\hat{\sigma}\right)\right\}_{j=1}^n$. This sequence is similar with the sequence $\{\xi_j\}_{j=1}^n$ in (2.5.4). Recall that $v(s(t_j), t_j)$ and $v(s(t_j+y), t_j+y)$ are true prices of the option at the moments of time t_j and $t_j + y$ respectively. We set

$$\overline{\xi}_j\left(\hat{\sigma}\right) = \begin{cases} 1 \text{ if } v_{\text{pred}}\left(t_j+y\right) \geq v(s(t_j), t_j) \text{ and } v(s(t_j+y), t_j+y) \geq v(s(t_j), t_j), \\ 0 \text{ otherwise.} \end{cases} \tag{2.6.3}$$

Next, we introduce the function $\overline{\zeta}\left(\hat{\sigma}\right)$ of the discrete variable $\hat{\sigma}$ as:

$$\overline{\zeta}\left(\hat{\sigma}\right) = \frac{1}{n}\sum_{j=1}^n \overline{\xi}_j\left(\hat{\sigma}\right), \; n = 2000. \tag{2.6.4}$$

It follows from (2.6.3) and (2.6.4) that $\overline{\zeta}\left(\hat{\sigma}\right)$ is the frequency of correctly predicted profitable cases for trading of this option with the market's opinion $\hat{\sigma}$ of the volatility of the option.

Predictions are performed by our algorithm of section 3. Comparison of (2.5.4) and (2.5.5) with (2.6.3) and (2.6.4) shows that $\overline{\zeta}(\hat{\sigma})$ is similar with the ideal case of the random variable ζ. The bold faced curve on Figure 2.3 depicts the graph of the function $\overline{\zeta}(\hat{\sigma})$. The middle non-horizontal curve on Figure 2.3 depicts the graph of the function $p(\hat{\sigma})$, which was constructed in subsection 2.6.2 for the ideal case. The upper and the lowest curves on Figure 2.3 display the shifts of the ideal curve up and down by $\sqrt{D}$, where D is the dispersion of ζ and D is given in (2.5.6). In other words, there is a high probability chance that the values of ζ are contained in the trust corridor between these two curves. The vertical line indicates the "critical" value $\hat{\sigma} = \sigma = 0.2$, where $\sigma = 0.2$ is the volatility of the stock.

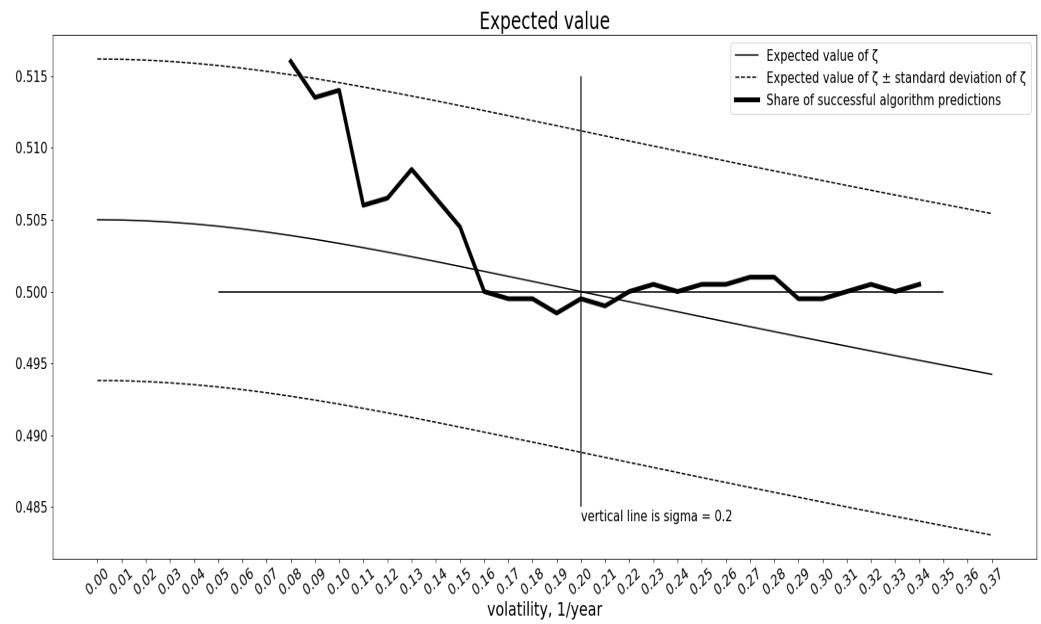

Figure 2.3: *Results of our computations. The vertical line indicates the value $\sigma = 0.2$ of the volatility of the stock in our numerical studies. The middle curve corresponds to the ideal case when both volatilities σ and $\hat{\sigma}$ are known. This curve depicts the dependence of the mathematical expectation $p(\hat{\sigma})$ of the random variable ζ on the market's opinion about the volatility $\hat{\sigma}$ of the option, see (2.5.2). Two curves, which are parallel to the middle one, are shifts of the latter by $\sqrt{D}$, where D is the dispersion of ζ, see (2.5.6). The bold faced curve represents the frequency $\overline{\zeta}(\hat{\sigma})$ of correctly predicted profitable cases for trading of this option with the market's opinion $\hat{\sigma}$ of the volatility of the option. Our algorithm of section 2.3 was used to compute $\overline{\zeta}(\hat{\sigma})$. The meaning of $\overline{\zeta}(\hat{\sigma})$ is similar with the meaning of the ideal case of the random variable ζ. Thus, it is natural that the bold faced curve lies in the trust corridor of ζ.*

One can see from the bold faced curve of Figure 2.3 that as long as $\hat{\sigma}$ is either rather close to σ or $\hat{\sigma} < \sigma$, the short position of this option represents a significant risk. However, when $\hat{\sigma}$ becomes less than $\approx 0.7\sigma$, the probability of the profit in the short position increases. This coincides with the prediction of our theory.

The bold faced curve is an analog of the middle curve of the mathematical expectation of the random variable ζ in the ideal case. Since the bold faced curve on Figure 2.3 lies within the trust corridor of the ideal algorithm, then we conclude that our prediction accuracy of profitable cases is comparable with the ideal one.

Unlike the ideal case, in a realistic scenario of the financial market data of, e.g. [2] only the information about the approximate values of $\hat{\sigma}$ is available. It is this information, which was used in [17, 16] and, in particular, in Tables 1,2.

Thus, our results support the following trading strategy in the non-ideal case:

Trading Strategy for the Non-Ideal Case:

Let $\eta > 0$ be a threshold number, which should be determined numerically by trial and error. For example, η might probably be linked with the transaction cost. Let $v_{\text{pred}}\left(t_j + y\right)$ be the number defined in (2.6.2).

1. If $v_{\text{pred}}\left(t_j + y\right) \geq v\left(s\left(t_j\right), t_j\right) + \eta$, then it is recommended to buy the option at the current trading day t_j and sell it on the next trading $t_j + y$.

2. If $v_{\text{pred}}\left(t_j + y\right) < v\left(s\left(t_j\right), t_j\right) - \eta$, then it is recommended to go short at the current trading day t_j, and to follow by closing the short position at the trading day $t_j + y$.

3. If $v\left(s\left(t_j\right), t_j\right) - \eta \leq v_{\text{pred}}\left(t_j + y\right) < v\left(s\left(t_j\right), t_j\right) + \eta$, then it is recommended to refrain from trading.

We believe that our results support the following two hypotheses:

Hyphothesis 1: *The reason why the heuristic algorithm of [17] and section 2.3 performs well is that it likely forecasts in many cases the signs of the differences $\sigma - \hat{\sigma}$ for the next trading day ahead of the current one.*

Hyphothesis 2: Since the maximal value of $\overline{\zeta}\left(\hat{\sigma}\right)$ in the bold faced curve of Figure 2.3 is 0.515, which is rather close to the value of 0.5577 in the "Precision" column of Table 1 and in the second column of Table 2, then we probably had in those tested real market data about 56% of options, in which $\sigma - \hat{\sigma} < 0$.

2.7 Concluding Remarks

We have considered a mathematical model, in which two markets are in place: the stock market and the options market. We have assumed that the market is imperfect. More precisely, we have assumed that agents of the option market have their own idea about the volatility $\hat{\sigma}$ of the option, and this idea might be different from the volatility σ of the stock. We have proven that if that $\sigma \neq \hat{\sigma}$, then there is an opportunity for a winning strategy. A rigorous probabilistic analysis was carried out. This analysis has shown that the mathematical expectation of the correctly guessed option price movements can be obtained, and it depends on the difference between σ and $\hat{\sigma}$.

We have considered both ideal and non-ideal cases. In the ideal case, both volatilities σ and $\hat{\sigma}$ are known. In the more realistic non-ideal case, however, only the volatility $\hat{\sigma} \approx \sigma_{\text{impl}}$ of the option is known from the market data, see, e.g. [2]. We have demonstrated in our numerical simulations that the accuracy of our prediction of profitable cases by the algorithm of [17] for the non-ideal case is comparable with that accuracy for the ideal case.

These results led us to two hypotheses. The first hypothesis is that our algorithm of [17] actually forecasts in many cases the signs of the differences $\sigma - \hat{\sigma}$ for the next trading day ahead of the current one. Our second hypothesis is based on our above results as well as on the "Precision" column in Table 1 and the second column in Table 2. This second hyphothesis tells one that probably about 56% out of tested 23,549 options of [16] with the real market data had $\sigma - \hat{\sigma} < 0$.

A new convergence analysis of our algorithm was carried out. To do this, the technique of [20] was modified and simplified for our specific case of the 1-D parabolic equation with the reversed time. We have lifted here the assumption of [17] that the time interval $(0, 2y)$ is sufficiently small. Indeed, even though we actually work with a small number $2y$ in our computations, that assumption might require even smaller values. In addition, we have derived a stability estimate for Problem 3 of subsection 2.4.1, which was not done in [17].

Chapter 3

QUASI-REVESIBILITY METHOD AND NEURAL NETWORK MACHINE LEARNING FOR FORECATING OF STOCK OPTION PRICES

3.1 Introduction

This chapter discusses a new empirical mathematical model for generating more accurate option trading strategy using initial and boundary conditions for the underlying stock. The idea was initially proposed in [17]. The basis for this idea is the Black-Scholes equation. In mathematical finance, the Black-Scholes equation is a parabolic partial differential equation that determines the dynamics of the price of European options [34].

The time at a given time t will occur is τ,

$$\tau = T - t. \tag{3.1.1}$$

$f(s)$ be the payoff function of that option at the maturity time $t = T$ and s is the stock price. Let's assume that the risk-free interest rate equals zero. The function $u\,(s, \tau)$ is the price of that option and the variable τ is the one defined in (3.1.1). Let's assume that this function $u\,(s, \tau)$ satisfies the Black-Scholes equation with the volatility coefficient σ [5, Chapter 7, Theorem 7.7]**:**

$$\begin{aligned} \frac{\partial u(s,\tau)}{\partial \tau} &= \frac{\sigma^2}{2} s^2 \frac{\partial^2 u(s,\tau)}{\partial s^2}, \\ u(s,0) &= f(s), \end{aligned} \tag{3.1.2}$$

The payoff function is $f(s) = \max\,(s - K, 0)$, where K is the strike price [5] and $s > 0$.

The option price function is defined by the Black-Scholes formula:

$$u(s.\tau) = s\Phi(\theta_+(s,\tau)) - e^{-r\tau} K\Phi(\theta_-(s,\tau)), \tag{3.1.3}$$

Based on the Itô formula, we have:

$$du = (-\frac{\partial u(s, T-t)}{\partial \tau} + \frac{\sigma^2}{2} s^2 \frac{\partial^2 u(s, T-t)}{\partial s^2})dt + \sigma s \frac{\partial u(s, T-t)}{\partial s} dW. \quad (3.1.4)$$

If equation (3.1.2) is solved forwards in time to forecast prices of stock options is an ill-posed inverse problem. For this reason, we used the Method of Quasi-Reversibility (QRM) that is a version of the Tikhonov regularization method. Uniqueness, stability and convergence theorems for this method were formulated in [17] and [20], also, see [19] for proofs.

We have four major questions that we raise in this paper:

1. What is the forecast interval of the options prices?

2. What are the boundary and initial conditions on the interval for the Black-Scholes equation?

3. What are the values of the volatility coefficient in the future?

4. How to solve the Black-Scholes equation forwards in time t?

The first three questions are addressed in our new mathematical model. We use the regularization method of [15] to address the fourth question. Theorems about stability and convergence of this method are formulated. These theorems were proven in [15] for a general parabolic equation of the second order where the main key of this method is based on the method of Carleman estimates.

Let the function $f \in L_2(0, \pi)$ and let $T = const. > 0$. To demonstrate that our problem is ill-posed, we consider the example of based on the problem for the heat equation with the reversed time

$$u_t + u_{xx} = 0, (x, t) \in (0, \pi) \times (0, T), \quad (3.1.5)$$

with Dirichlet boundary conditions

$$u(0, t) = 0, u(\pi, t) = 0, \quad (3.1.6)$$

and the initial condition

$$u(x,0) = f(x)\,. \tag{3.1.7}$$

The unique solution of this problem is:

$$u(x,t) = \sum_{n=1}^{\infty} f_n \sin(nx) e^{n^2 t}.$$

Consider

$$u_N(x,t) = \sum_{n=1}^{N} f_n \sin(nx) e^{n^2 t}.$$

Then

$$||u_N(x,T)||^2_{L_2(0,\pi)} = \sum_{n=1}^{N} f_n^2 e^{2n^2 T} \approx f_N^2 e^{2N^2 T} \to \infty$$

as $N, T \to \infty$. Hence, the problem (3.1.5)-(3.1.7) is severely unstable.

We conclude therefore that to obtain a more or less accurate solution of the Black-Scholes equation forwards in time, we need to solve it on a short time interval $(0,T)$. To get better accuracy, the regularization method works only for a short time interval.

Section 3.3 presents our mathematical model, the method of Quasi Reversibility as well as the trading strategy. The Quasi Reversibility Method is based on the minimization of a Tikhonov-like functional $J_\beta(u)$. We do this using conjugate gradient method. The minimization process was performed by Hyak Next Generation Supercomputer of the research computing club of University of Washington. The code was parallelized in order to maximize the performance on supercomputer clusters.

The historical data for stock options was collected from the Bloomberg terminal [2] of University of Washington. From this data, we obtained about 177,000 minimizers.

Due to ill-posedness of the problem the solution is very sensitive to the noise in the initial data (stock and option prices for the three days preceding the day of forecast). Given results of the Quasi Reversibility Method, we apply on the second stage Machine Learning to reduce the probability of non-profitable trades caused by wrong option price prognosis because of the noise in input data.

Section 3.4 is dedicated to application of binary classification and regression Neural Network Machine Learning.

Sections 3.5 and 3.6 present our results and the summary.

Python with the SciPy and Torch modules were used for implementation of the method of Quasi-Reversibility and Neural Network Machine Learning (binary classification and regression).

3.2 The new mathematical model and the method of Quasi-Reversibility

Let's denote s as the stock price, t as the time and $\sigma(t)$ as the volatility of the option. The historical implied volatility listed on the market data of [2] is used in our particular case. We assume that $\sigma = \sigma(t)$ to avoid other historical data for the volatility. Let's call $u_b(t)$ and $u_a(t)$ the bid and ask prices of the options at the moment of time t and $s_b(t)$ and $s_a(t)$ the bid and ask prices of the stock at the moment of time t. It is also known that

$$u_b(t) < u_a(t) \tag{3.2.1}$$

and

$$s_b(t) < s_a(t) \tag{3.2.2}$$

Let's introduce

$$f_s(t) = \frac{s_a(t)}{s_b(t)} - 1 \tag{3.2.3}$$

and

$$f_u(t) = \frac{u_a(t)}{u_b(t)} - 1 \tag{3.2.4}$$

Based on real market data we have observed that usually

$$0 \leq f_s(t) \leq 0.003 \tag{3.2.5}$$

and

$$0 \leq f_u(t) \leq 0.27 \tag{3.2.6}$$

The idea is to approximate the Black-Scholes equation solutions

$$Lu = u_t + \frac{\sigma^2(t)}{2} s^2 u_{ss} = 0, (s,t) \in (s_b(0), s_a(0)) \times (0, 2\tau) = X_{2\tau}, \tag{3.2.7}$$

with Dirichlet boundary conditions

$$u(s_b, t) = u_b(t), u(s_a, t) = u_a(t), \ t \in [0, 2\tau], \tag{3.2.8}$$

and the initial condition

$$u(s, 0) = f(s), \quad s \in [s_b(0), s_a(0)]. \tag{3.2.9}$$

Where L is the partial differential operator of the Black-Scholes equation. Based on Bloomerg terminal we used with End of Day Underlying Price Last, End of Day Underlying Price Bid, End of Day Underlying Price Ask, t is time, $\sigma(t)$ is the volatility of the stock option. It was used Implied Volatility Using Last Trade Price (IVOL).

$u(s,t)$ is the price of the stock option. End of Day Option Price Last, End of Day Option Price Bid and End of Day Option Price Ask are the notation that we applied in our algorithm.

Problem 1. *Find the function $u \in H^2(X_{2\tau})$ satisfying conditions (3.2.7)-(3.2.9).*

This problem considers as ill-posed since we solve equation (3.2.7) forwards in time.

Remarks 3.1: *We increase here the required smoothness of the solution from $H^{2,1}(X_{2\tau})$ to $H^2(X_{2\tau})$.*

Our algorithm based on solving the inverse problem for the Black-Scholes with reversed time equation has five steps:

Step 1 (Dimensionless variables).

We require to make our equation dimensionless. $s_b < s_a$. Let's denote $s_b = s_b(0)$, $s_a = s_a(0)$. Dimensionless variables were applied x, t' such that

$$x = \frac{s - s_b}{s_a - s_b} \tag{3.2.10}$$

$$t' = \frac{t}{255} \tag{3.2.11}$$

and now we can say that s is x and t is t'.

According to these substitutions, the equation becomes

$$Ru = u_t + \sigma^2 (t) A(x) u_{xx}, \tag{3.2.12}$$

where

$$A(x) = \frac{255}{2} \frac{[x(s_a - s_b) + s_b)]^2}{(s_a - s_b)^2} \tag{3.2.13}$$

$$X_{2\tau} = \{(x,t) \in (0,1) \times (0, 2\tau)\}. \tag{3.2.14}$$

$$u(x,0) = g(x), x \in (0,1) \tag{3.2.15}$$

$$u(0,t) = u_b(t), u(1,t) = u_a(t). \tag{3.2.16}$$

And the operator L in (3.2.7) is the operator R

Step 2 (Interpolation and extrapolation).

Our goal is to forecast option price from 'today' to 'tomorrow' and 'the day after tomorrow'. We do have 255 trading days annually. For this reason, let's introduce $\tau > 0$ as our unit of time for which we want to make our prediction the option price. Because we predict option prices having the information of these prices, as well as of other parameters for 'today', 'yesterday' and 'the day before yesterday', we consider τ is one trading day. 'One day' $\tau = 1/255$. 'Today' $t = 0$. 'Tomorrow' $t = \tau$. 'The day after tomorrow' $t = 2\tau$. The variable $s-$ is for interval, i.e $s \in [s_b(0), s_a(0)]$. We applied the idea associated with interpolation discrete values of functions $u_b(t)$, $u_a(t)$, and $\sigma(t)$ between these three points (the day before yesterday, yesterday and today) and then extrapolation functions $u_b(t)$, $u_a(t)$ between three points (today, tomorrow and the day after tomorrow). Where $t = -2\tau$ is "the day before yesterday", $t = -\tau$ is "yesterday" and $t = 0$ is "today". We used quadratic polynomials for both approximation and extrapolation of values of functions. Thus, these three functions $u_b(t)$, $u_a(t)$, and $\sigma(t)$ was obtained for a small future time interval, i.e $(0, 2\tau)$. ([17]). Where

$u_b(t)$, $u_a(t)$ were applied for boundary conditions and $\sigma(t)$ is coefficient function for our problem. The initial condition was set as $u(x,0) = g(x) = x(u_a(0) - u_b(0))x + u_b(0)$. This function is the result of approximation by linear function due to the fact that the interval between bid and ask prices is relatively small. The domain was $X_{2\tau} = \{(x,t) : x \in (0,1), t \in (0,2\tau)\}$.

Step 3 (Statement of the Problem).

Problem 2. *Assume that functions*

$$u_b(t), u_a(t) \in H^2[0,2\tau], \sigma(t) \in C^1[0,2\tau]. \tag{3.2.17}$$

Find the solution $u \in H^2(X_{2\tau})$ of the following initial boundary value problem:

$$Ru = 0 \text{ in } X_{2\tau}, \tag{3.2.18}$$

$$u(0,t) = u_b(t), u(1,t) = u_a(t), t \in (0,2\tau), \tag{3.2.19}$$

$$u(x,0) = g(x), x \in (0,1), \tag{3.2.20}$$

where the partial differential operator R is defined in (3.2.12), the function $A(x)$ is defined in (3.2.13), the initial condition $g(x)$ is defined in (3.2.15), and the domain $X_{2\tau}$ is defined in (3.2.14).

Theorem 3.2.1. *The following problem (3.2.12)-(3.2.15) has one solution $u \in H^{2,1}(X_{2\tau})$.*

The proof of this theorem is [20].

Step 4 (Numerical method of solving the problem. Regularization).

Due to the ill-posedness of the problem, we can not say about existence of the solution. Thus, it was applied the regularization method:

Let's consider function $F(x,t) = x(u_a(t) - u_b(t)) + u_b(t)$, $(x,t) \in X_{2\tau}$. This function $F \in H^2(X_{2\tau})$. It follows from (3.2.15) and (3.2.16) that

$$F(x,0) = g(x), \tag{3.2.21}$$

$$F(0,t) = u_b(t), F(1,t) = u_a(t). \tag{3.2.22}$$

We used an unbounded differential operator $R : H^{2,1}(X_{2\tau}) \to L^2(X_{2\tau})$, where $H^{2,1}(X_{2\tau})$ is a dense linear set in the space $L^2(X_{2\tau})$. Where

$$Ru = u_t + \sigma^2(t) A(x) u_{xx} \tag{3.2.23}$$

Let's introduce Tikhonov-like functional as:

$$J_\beta(u) = \int_{X_{2\tau}} (Ru)^2 \, dsdt + \beta \|u\|^2_{H^2(X_{2\tau})}, \tag{3.2.24}$$

where $\beta \in (0,1)$ is the parameter of regularization. To solve the problem, we minimized the functional $J_\beta(u)$ on the set

$$V = \left\{u \in H^2(X_{2\tau}) : u(0,t) = u_b(t), u(1,t) = u_a(t), u(x,0) = g(x)\right\}. \tag{3.2.25}$$

Step 5 (Minimization Problem).

Minimization Problem 1. $J_\beta : H^2(X_{2\tau}) \to \mathbb{R}$ *is the regularization Tikhonov functional.*

We have used the converting of our partial derivatives from (3.2.24) into finite differences. A finite difference grid was applied to cover the domain $X_{2\tau}$. The minimization process was to differentiate our functional $J_\beta(u)$ with respect to the values of the function $u(x,t)$ at each grid points via conjugate gradient method. The point $u = 0$ was used for the starting point. Based on computational study with simulated data we have realized that the optimal value of the regularization parameter would be $\beta = 0.01$.

Minimization Problem 1 is a QRM for Problem 2. This is an version of the QRM for problem (3.2.18)-(3.2.20). In section 3.3 we discuss the theory of this specific version of the QRM. In particular, Theorem 3.3.2 of section 3.3 presents uniqueness of the solution $u \in H^2(X_{2\tau})$ of Problem 2 and implies an estimate of the stability of this solution with respect to the noise in the data. Theorem 3.4.3 of section 3.3 shows existence and uniqueness of the minimizer $u_\beta \in H^{2,1}(X_{2\tau})$ of the functional $J_\beta(u)$ on the set V defined in (3.2.25).

We call such a minimizer "regularized solution" [35]. Theorem 3.3.4 estimates convergence rate of regularized solutions to the exact solution of Problem 2 with the noiseless data. Such estimates depend on the noise level in the data. All proof of these theorems are presented in [19].

3.3 Analysis

This section is devoted to convergence analysis for Problem 2 of subsection 3.2. This problem is the initial boundary value problem for parabolic equation (3.2.18) with the reversed time. The QRM and convergence analysis for this problem for a more general parabolic operator in $\mathbb{R}^n$ with arbitrary variable coefficients was proposed in [15]. Then theorems were presented in [17]. However a stability estimate was not a part of [17], such an estimate was proven in [15]. The same is true for the convergence theorems of QRM in [15, 17]. The smallness assumption was lifted in [20] via a new Carleman estimate. Results of [20] for a 1-D case were significantly modified in this section. Our computations below on a small time interval $(0, 2\tau) = (0, 0.00784)$ (see [15, 17], [24, Theorem 1 of section 2 in Chapter 4] might result in the requirement of even a smaller length of that interval.

3.3.1 Problem statement

Let's consider a number $T > 0$ and introduce Q_T as:

$$Q_T = \{(x,t) \in (0,1) \times (0,T)\}.$$

Consider two numbers $b_0, b_1 > 0$ and $b_0 < b_1$. Let the function $b(x,t) \in C^1\left(\overline{Q}_T\right)$ satisfies:

$$\|b\|_{C^1\left(\overline{Q}_T\right)} \leq b_1, \; b(x,t) \geq b_0 \text{ in } Q_T. \tag{3.3.1}$$

We also have functions $\psi_0(t), \psi_1(t) \in H^2(0,T)$. In the above case of subsection 3.2,

$$T = 2\tau, b(x,t) = \sigma^2(t) A(x), \psi_0(t) = u_b(t), \psi_1(t) = u_a(t).$$

We now formulate Problem 3, which is a slight generalization of Problem 2.

Problem 3. *Find a solution $v \in H^2(Q_T)$ of the following (IBVP):*

$$Nv = v_t + b(x,t) v_{xx} = 0 \text{ in } Q_T, \tag{3.3.2}$$

$$v(0,t) = \psi_0(t), v(1,t) = \psi_1(t), \ t \in (0,T), \tag{3.3.3}$$

$$v(x,0) = z(x) = \psi_0(0)(1-x) + \psi_1(0)x, \ x \in (0,1). \tag{3.3.4}$$

Remark 3.3.1 *Because Problem 2 is less general than Problem 3, then this analysis of converegence for Problem 3 also works for Problem 2.*

We use the linear function for $v(x,0)$ in (3.3.4) is to simplify the initial condition in (3.2.20). Now problem 3 is an IBVP for the parabolic equation (3.3.2) with the reversed time. For this reason, the problem can be considered as ill-posed. Assume that the boundary with a noise of the level $\nu > 0$ in (3.3.3) are in place. Here ν is a sufficiently small number, i.e.

$$\|\psi_0 - \psi_0^*\|_{H^1(0,T)} < \nu, \|\psi_1 - \psi_1^*\|_{H^1(0,T)} < \nu, \tag{3.3.5}$$

where functions $\psi_0^*, \psi_1^* \in H^2(0,T)$ are "ideal" noiseless data. we assume that there exists an exact solution $v^* \in H^2(Q_T)$ of problem (3.3.2)-(3.3.4) with these noiseless data (based on on the theory of Ill-Posed problems). Below we present estimates how this noise affects the accuracy of the solution of Problem 3 and also discuss the convergence rate of numerical solutions obtained by QRM to the exact one as $\nu \to 0$.

Let's introduce the version of functional (3.2.24):

$$I_\beta(v) = \int_{Q_T} (Nv)^2 \, dxdt + \beta \|v\|^2_{H^2(Q_T)}. \tag{3.3.6}$$

We also have the set $W \subset H^2(Q_T)$,

$$W = \left\{ v \in H^2(Q_T) : v(0,t) = \psi_0(t), v(1,t) = \psi_1(t), v(x,0) = z(x) \right\}. \tag{3.3.7}$$

The solution of Problem 3 is approximate solution by solving the following problem:

Minimization Problem 2. *Minimize the functional $I_\beta(v)$ on the set W given in (3.3.7).*

Minimization Problem 2 is QRM for Problem 3.

3.3.2 Theorems

This subsection presents four theorems for Problem 3. All proofs might be found in [19]. First, let's introduce the Carleman Weight Function $\phi_\alpha(t)$ with $\alpha > 2$ for the operator

$\partial_t + b(x,t)\,\partial_x^2$ as:

$$\phi_\alpha(t) = e^{(T+1-t)^\alpha},\ t \in (0,T). \tag{3.3.8}$$

As a result, the function $\phi_\alpha(t)$ is decreasing on $[0,T]$, $\phi'_\alpha(t) < 0$,

$$\max_{[0,T]} \phi_\alpha(t) = \psi_\alpha(0) = e^{(T+1)^\alpha},\ \min_{[0,T]} \phi_\alpha(t) = \phi_\alpha(T) = e. \tag{3.3.9}$$

Denote

$$H_0^2(Q_T) = \left\{u \in H^2(Q_T) : u(0,t) = u(1,t) = 0\right\}. \tag{3.3.10}$$

$$H_{0,0}^2(Q_T) = \left\{u \in H_0^2(Q_T) : u(x,0) = 0\right\}. \tag{3.3.11}$$

Theorem 3.3.1 (Carleman estimate). *Let the coefficient $b(x,t)$ of the operator N satisfies conditions (3.3.1). Then there exist a sufficiently large number $\alpha_0 = \alpha_0(T, b_0, b_1) > 2$ and a constant $C = C(T, b_0, b_1) > 0$, both depending only on listed parameters, such that the following Carleman estimate holds for the operator N :*

$$\int_{Q_T} (Nu)^2 \phi_\alpha^2 dxdt \geq C\sqrt{\alpha} \int_{Q_T} u_x^2 \psi_\alpha^2 dxdt + C\alpha^2 \int_{Q_T} u^2 \phi_\alpha^2 dxdt$$
$$-C\sqrt{\alpha}\,\|u\|_{H^2(Q_T)}^2 - C\lambda (T+1)^\alpha e^{2(T+1)^\alpha} \|u(x,0)\|_{L_2(0,1)}^2, \tag{3.3.12}$$
$$\forall \alpha \geq \alpha_0, \forall u \in H_0^2(Q_T).$$

Carleman estimate (3.3.12) is the MAIN TOOL to proofs of Theorems 3.3.2, 3.3.4.

Theorem 3.3.2 (Hölder stability estimate for Problem 3 and uniqueness). *Let the coefficient $b(x,t)$ of the operator N satisfies conditions (3.3.1). Let's assume that the functions $v \in H^2(Q_T)$ and $v^* \in H^2(Q_T)$ are solutions of Problem 3 with the vectors of data $(\psi_0(t), \psi_1(t))$ and $(\psi_0^*(t), \psi_1^*(t))$ respectively, where $\psi_0, \psi_1, \psi_0^*, \psi_1^* \in H^2(0,T)$. Assume also that error estimates (3.3.5) of the boundary data is in place. Choose an arbitrary number $\epsilon \in (0,T)$. Denote*

$$\lambda = \lambda(T,\epsilon) = \frac{\ln(T+1-\epsilon)}{\ln(T+1)} \in (0,1). \tag{3.3.13}$$

Then there exists a sufficiently small number $\nu_0 = \nu_0(T_1, b_0, b_1) \in (0,1)$ and a constant $C_1 = C_1(T, b_0, b_1, \epsilon) > 0$, both depending only on listed parameters, such that the following stability estimate holds for all $\nu \in (0, \nu_0)$:

$$\|v_x - v_x^*\|_{L_2(Q_{T-\epsilon})} + \|v - v^*\|_{L_2(Q_{T-\epsilon})} \leq \tag{3.3.14}$$

$$\leq C_1 \left(1 + \|v - v^*\|_{H^2(Q_T)}\right) \exp\left[-\left(\ln \nu^{-1/2}\right)^{\lambda}\right].$$

Below $C = C(T, b_0, b_1) > 0$ and $C_1 = C_1(T, a_0, b_1) > 0$ denote different constants depending only on listed parameters.

Corollary 3.3.1 (uniqueness). *Let the coefficient $b(x,t)$ of the operator N satisfies conditions (3.3.1). Then Problem 3 has at most one solution (uniqueness).*

Proof. If $\nu = 0$, then (3.3.14) implies that $v(x,t) = v^*(x,t)$ in $Q_{T-\epsilon}$. Since $\epsilon \in (0,T)$ is an arbitrary number, then $v(x,t) \equiv v^*(x,t)$ in Q_T. □

Theorem 3.3.3 (existence and uniqueness of the minimizer)**.** *Let functions $\psi_0(t), \psi_1(t) \in H^2(0,T)$. Let W be the set defined in (3.3.7). Then there exists unique minimizer $v_{\min} \in W$ of functional (3.3.6) and*

$$\|v_{\min}\|_{H^2(Q_T)} \leq \frac{C}{\sqrt{\beta}} \left(\|\psi_0\|_{H^2(0,T)} + \|\psi_1\|_{H^2(0,T)}\right). \tag{3.3.15}$$

In the theory of Ill-Posed Problems, this minimizer $v_{\min}$ is called "regularized solution" of Problem 3 [35]. According to the theory of Ill-Posed problems, it is important to establish convergence rate of regularized solutions to the exact one v^*. In doing so, one should always choose a dependence of the regularization parameter β on the noise level ν, i.e. $\beta = \beta(\nu) \in (0,1)$ [35].

Theorem 3.3.4 (convergence rate of regularized solutions). *Let $v^* \in H^2(Q_T)$ be the solution of Problem 3 with the noiseless data $(\psi_0^*(t), \psi_1^*(t))$. Let functions $\psi_0, \psi_1, \psi_0^*, \psi_1^* \in H^2(0,T)$. Let $v_{\min} \in W$ be the unique minimizer of functional (3.3.6) on the set W. Assume that error estimates (3.3.5) hold. Choose an arbitrary number $\epsilon \in (0,T)$. Let $\lambda = \lambda(T,\epsilon) \in (0,1)$ be the number defined in (3.3.13) and let*

$$\beta = \beta(\nu) = \nu^2, \tag{3.3.16}$$

Then there exists a sufficiently small number $\nu_0 = \nu_0(T, b_0, b_1) \in (0,1)$ depending only on listed parameters such that the following convergence rate of regularized solutions $v_{\min}$ holds for all $\nu \in (0, \nu_0)$:

$$\|\partial_x v_{\min} - \partial_x v^*\|_{L_2(Q_{T-\epsilon})} + \|v_{\min} - v^*\|_{L_2(Q_{T-\epsilon})} \tag{3.3.17}$$

$$\leq C_1 \left(1 + \|v^*\|_{H^2(Q_T)} + \|\psi_0^*\|_{H^2(0,T)} + \|\psi_1^*\|_{H^2(0,T)}\right) \exp\left[-\left(\ln \nu^{-1/2}\right)^{\lambda}\right].$$

3.3.3 Trading Strategy:

We use minimizers obtained from the method of Quasi-Reversibility to build a strategy for trading options. Let's define

$$REAL(0) = \frac{u_a(0) + u_b(0)}{2} \tag{3.3.18}$$

$$REAL(\tau) = \frac{u_a(\tau) + u_b(\tau)}{2} \tag{3.3.19}$$

$$EST(\tau) = u_\beta(1/2, k\tau) \tag{3.3.20}$$

or if it was not applied dimensionless

$$EST(\tau) = u_\beta(\frac{s_a + s_b}{2}, k\tau) \tag{3.3.21}$$

where $k = 1$

Here $EST(\tau)$ means minimizer.

Let's buy an option if the following holds

$$EST(\tau) \geq REAL(0) \tag{3.3.22}$$

The predicted outcome of option trade is Positive if

$$EST(\tau) \geq REAL(0) \tag{3.3.23}$$

Definition 1.

It is True Positive if

$$EST(\tau) \geq REAL(0) \tag{3.3.24}$$

and

$$REAL(\tau) \geq REAL(0) \tag{3.3.25}$$

Definition 2.

It is True Negative if

$$EST(\tau) < REAL(0) \tag{3.3.26}$$

and

$$REAL(\tau) < REAL(0) \tag{3.3.27}$$

Definition 3.

It is False Positive if

$$EST(\tau) \geq REAL(0) \tag{3.3.28}$$

and

$$REAL(\tau) < REAL(0) \tag{3.3.29}$$

Definition 4.

It is False Negative if

$$EST(\tau) < REAL(0) \tag{3.3.30}$$

and

$$REAL(\tau) \geq REAL(0) \tag{3.3.31}$$

The accuracy of trading strategy is defined as

$$Accuracy = \frac{TP + TN}{\sum options} \tag{3.3.32}$$

where TP is a summation of True Positive and TN is a summation of True Negative and $\sum options$ is a summation of options in data set.

The precision of trading strategy is defined as

$$Precision = \frac{TP}{TP + FP} \tag{3.3.33}$$

where FP is a summation of False Positive.

The recall of trading strategy is defined as

$$Recall = \frac{TP}{TP + FN} \tag{3.3.34}$$

where FN is a summation of False Negative.

The average relative error of trading strategy is defined as

$$Error = \frac{1}{N} \sum \left| \frac{EST(\tau) - REAL(\tau)}{REAL(\tau)} \right| \tag{3.3.35}$$

3.4 Application of Neural Network Machine Learning

The Black-Scholes equation gives fair value of options in perfect market. However, real options prices contain some level of noise. We try to filter mispredictions (i.e. where minimizers result in False Positive or False Negative) caused by input noise using Machine Learning to improve accuracy, precision and recall of the trading strategy. We built a neural network with 13 element input vector and 3 fully connected hidden layers. (See Fig 1). Input vector consists of minimizers (for $t = \tau, 2\tau$) obtained from the method of Quasi-Reversibility, stock ask and bid price (for $t = 0$), option ask and bid price and volatility (for $t = -2\tau, -\tau, 0$).

Figure 3.1:

All vectors and labels are split into three parts: training, validation and test sets. The training set is used for weight learning. Validation set is used for tuning of the neural network hyper-parameters. Test set is for generating the outcomes of trading strategy.

We collected historical option and stock prices along with implied volatility on companies consisting of Russel 2000 index [1].

Table 3.

Set	Dates	Number of options
Training	2016/09/14-2018/05/31	132,912
Validation	2018/06/01-2018/06/29	13,401
Test	2018/07/02-2018/08/17	23,549

We compared the profitability of the trading strategy based on the original minimizer set with the profitability of the output of Machine Learning.

3.4.1 Machine Learning Input Vector Normalization

$$\mu = \frac{u_a(0) + u_a(-\tau) + u_a(-2\tau) + u_b(0) + u_b(-\tau) + u_b(-2\tau)}{6} \tag{3.4.1}$$

$$op_n = \frac{u(t) - \mu}{\sigma} \tag{3.4.2}$$

$$s_n = \frac{(s - st) - \mu}{\sigma} \tag{3.4.3}$$

where op_n is a normalized option price, s_n is a normalized stock price normalization, s is the stock, st is the strike and σ is the standard deviation.

3.4.2 Binary classification

Supervised Machine Learning has been applied to the neural network for the Cross Entropy Loss function with regularization:

$$L(\theta) = \frac{1}{m}\sum_{i=1}^{m}[-y^{(i)}\log(h_\theta(x^{(i)}) - (1 - y^{(i)})\log(1 - h_\theta(x^{(i)}))] + \frac{\lambda}{2m}\sum_{j=1}^{n}\theta_j^2 \tag{3.4.4}$$

Where θ are weights which are optimized by minimizing the loss function using the method of gradient descent. λ is a parameter of regularization. $x^{(i)}$ is our normalized 13 - dimensional vectors. h_θ is output of the neural network. m is the number of vectors in the training set. $y^{(i)}$ is our labels (the ground truth). The trading strategy is defined by

$$H_c = \begin{cases} 1, & \text{if } h_\theta > c \\ 0, & \text{otherwise.} \end{cases} \tag{3.4.5}$$

where c is the threshold obtained by maximizing accuracy on validation set. The labels are set to 1 for profitable trades and 0 otherwise.

3.4.3 Regression model

Similarly, instead of using binary classification, we can use the same ML architecture to predict the option price for tomorrow (τ). We have the same input features as the classification neural network. Regression learning uses mean squared error as the loss function:

$$L(\bar{h}_\theta, \bar{y}) = \frac{1}{m} \sum_n (\bar{h}_\theta - \bar{y}_n)^2 \tag{3.4.6}$$

Where m is the size of the data set, $\bar{h}_\theta$ is the predicted value and $\bar{y}$ is the real value ($Real(\tau)$).

3.5 Results

The following graph shows the accuracy of the results on validation set. We use it to determine the optimal value of hyper-parameter c (the threshold value of binary classification).

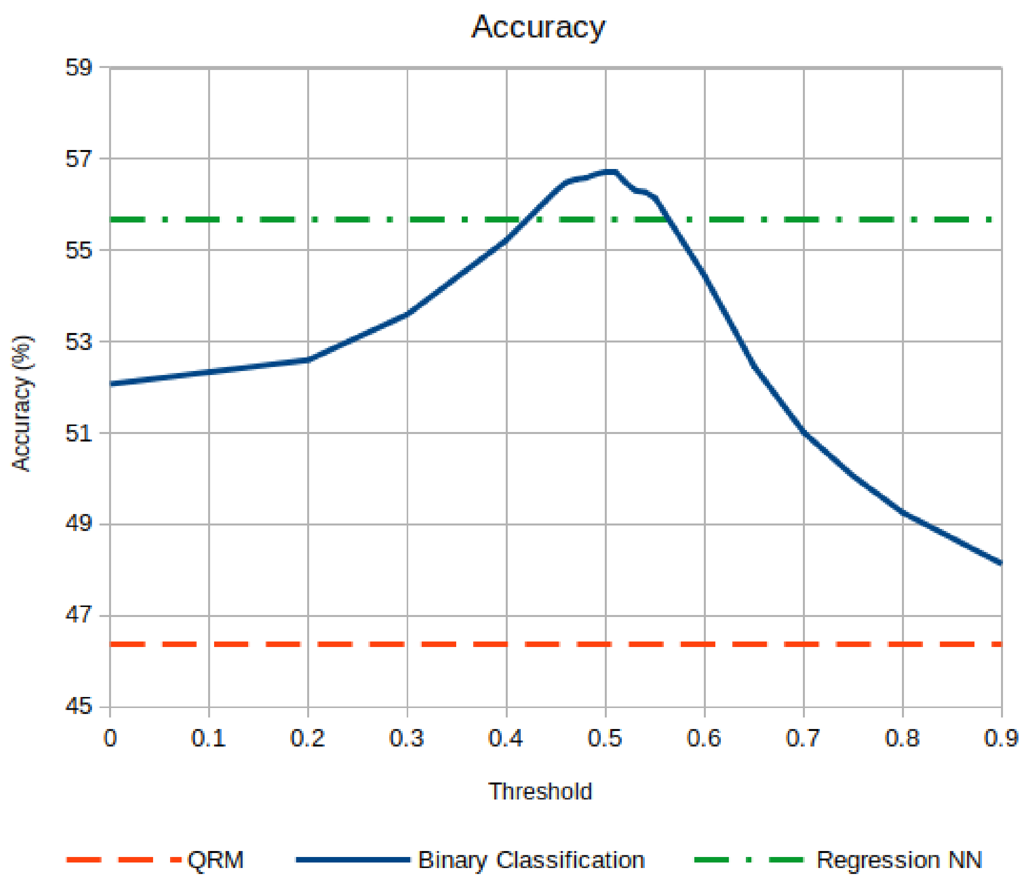

Figure 3.2: Accuracy. Threshold dependency

Observation 1.

The accuracy was improved by both Machine Learning methods compared to the method of Quasi-Reversibility. Based on this graph we set $c = 0.5$.

The next graph presents Recall and Precision diagram built on validation data set.

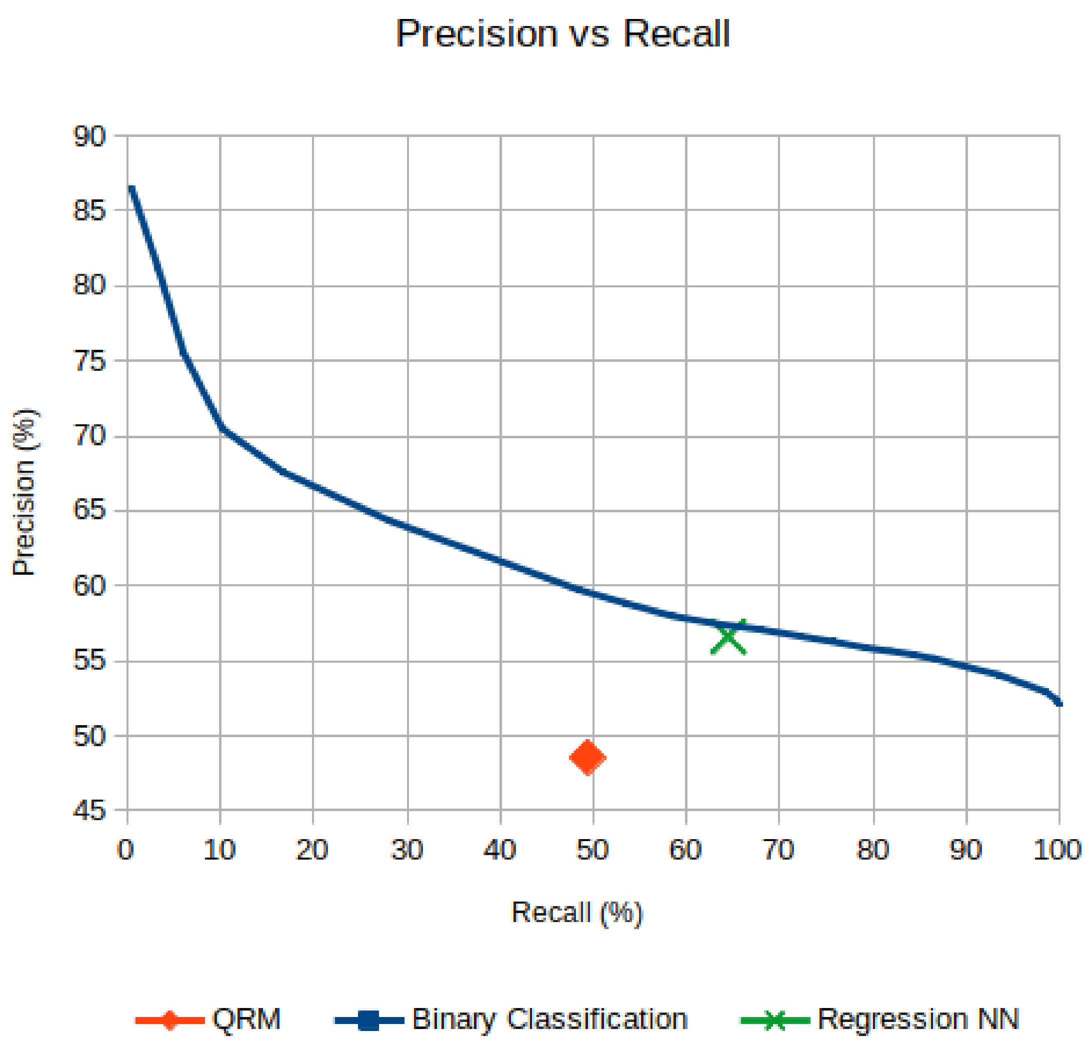

Figure 3.3: Precision vs Recall. Threshold dependency. Here **X** (Cross) indicates the position of Recall and Precision produced by Regression Neural Network and ◊ (Diamond) indicates the position of Recall and Precision produced by QRM.

Observation 2.

Binary Classification and Regression produced similar results that improved both precision and recall compared to the method of Quasi-Reversibility.

Further we divided our test data into bins where (horizontal axis, see Figure 3.4) each bin is determined by $\frac{s-st}{s}$ with step size 0.1. for each bin we calculated precision (see Figure 3.4).

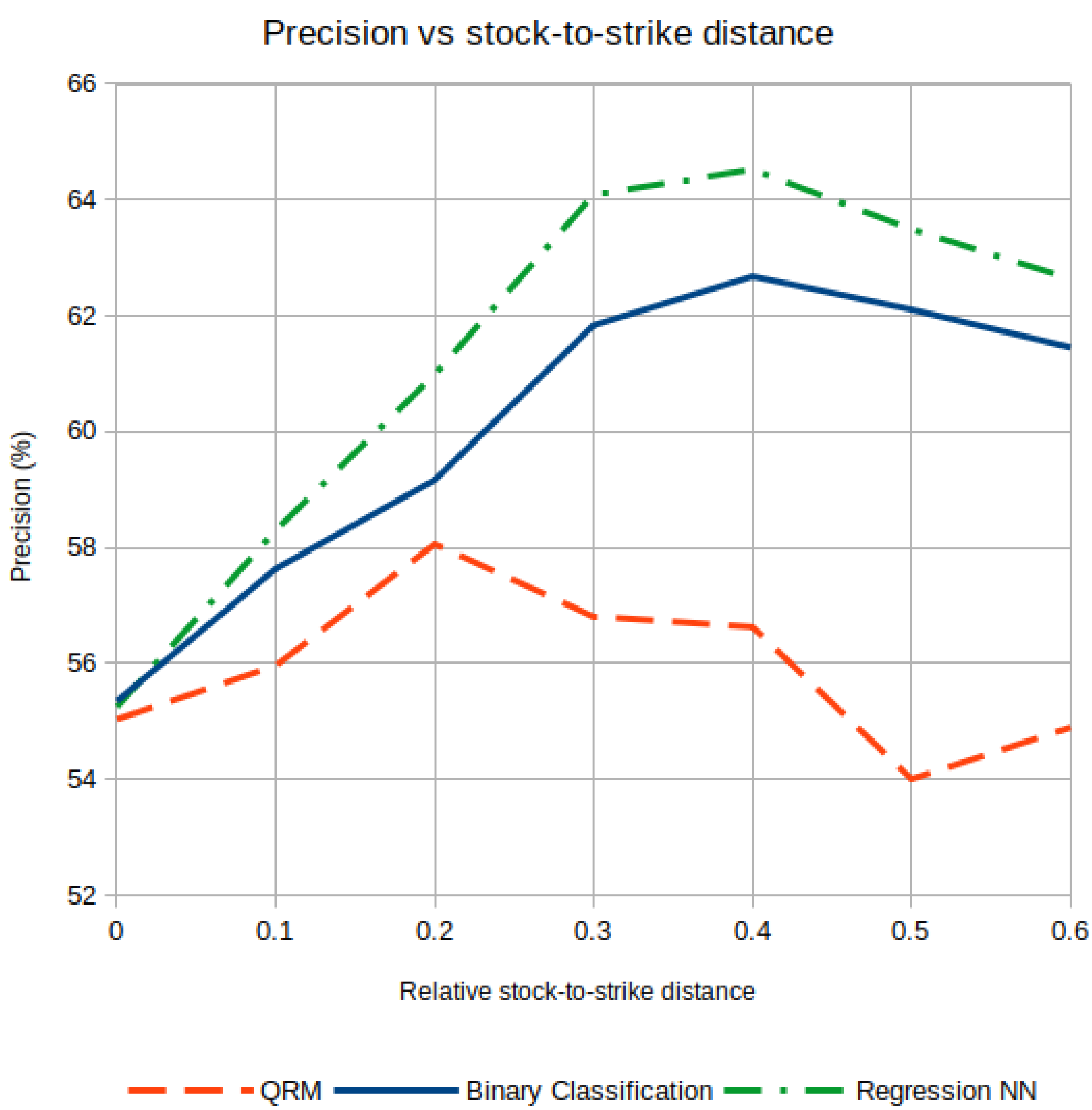

Figure 3.4: Binary Classification, Regression NN and the method of Quasi-Reversibility.

Observation 3.

To our surprise, when stock price was close to the strike price Machine Learning and the method of Quasi-Reversibility give similar precision (bin 0). With stock price diverging from the strike price Machine Learning produced better precision.

The following tables summarize the accuracy, precision and recall for all methods on test data.

Table 4. Final results on Test Data.

Method	Accuracy	Precision	Recall	Error
QRM	49.77%	55.77%	52.43%	12 %
Binary Classification	56.36%	59.56%	70.22%	NA
Regression NN	55.42%	60.32%	61.29%	NA

Table 5. Percentages of options with profits/losses for three different methods.

Method	Profitable options	Options with loss
QRM	55.77%	44.23 %
Binary Classification	59.56%	40.44%
Regression NN	60.32%	39.68%

3.6 Summary

To predict prices of stock options, we used two empirical mathematical models for Black-Scholes equation. The results achieved by solving the equation forwards in time (as an ill-posed problem) and applying Supervised Machine Learning (Binary Classification and Regression Neural Network, and using these methods with the real market data, show that this methodology produce promising results, potential applications within real-world trading and investment strategies.

The comparison of our methods resulted in the following two conclusions:

1. The predictions of the method of Quasi-Reversibility ended up being profitable for 55.77% of the options. Compare this to a 59.56% profitability rate for the Binary Classification method, and a 60.32% profitability rate for Regression Neural Network, used on the same data set and with the same trading strategy.

2. As shown by figures in section 3.5, option price forecasting using Machine Learning gives us significant accuracy and profit improvements over the method of Quasi-Reversibility. However, when stock price is close to the strike price both models give similar results.

The authors hypothesize that options traders can generate significant profits using trading strategies reliant on predictions generated with these methods.

Chapter 4

GLOBALLY STRICTLY CONVEX TIKHONOV-LIKE FUNCTIONAL FOR A COEFFICIENT INVERSE PROBLEM FOR A 1-D HYPERBOLIC EQUATION

4.1 Nonlinear Integro Differential Equation

All functions here are real valued ones.

The Forward Problem:

$$u_{tt} = u_{xx} + a(x)u, \ \ x \in R, t > 0 \tag{4.1.1}$$

$$u(x, 0) = 0, \ \ u_t(x, 0) = \delta(x). \tag{4.1.2}$$

The Inverse Problem: Suppose that

$$a(x) \in C^1(R) \tag{4.1.3}$$

$$a(x) \geq 0 \tag{4.1.4}$$

and that the following two functions $f_1(t)$, $f_2(t)$ are known

$$u(0, t) = f_1(t), \ \ u_x(0, t) = f_2(t) \tag{4.1.5}$$

Suppose also that the function $a(x)$ is known for $x \in (0, 1)$. Determine the function $a(x)$ for $x \in (0, 1)$ from functions $f_1(t)$, $f_2(t)$. The case of the equation

$$c(x)u_{tt} = u_{xx}, \ \ c(x) \geq 1 \tag{4.1.6}$$

can be considered similarly with the method presented below using the well known change of variables and assuming that $c(x) = 1$ for $x \in (0, \epsilon) \cup (1, \infty)$, where $\epsilon > 0$ is small number.

Let $H(y)$ be the Heaviside function that is 1 if $y > 0$ and 0 $y < 0$. The solution of the Forward Problem has the form

$$u(x,t) = \frac{1}{2}H(t-|x|) + \frac{1}{2}\int\int a(\xi)u(\xi,\tau)d\xi d\tau, \tag{4.1.7}$$

$$D(x,t) = \{(\xi,\tau) : |\xi| < \tau < t - |x-\xi|\} \tag{4.1.8}$$

$$u(x,t) > \frac{1}{2} \ \ for \ \ t > |x| \tag{4.1.9}$$

$$u, u_t \in C^2(t \geq |x|) \tag{4.1.10}$$

$$\lim_{t \to |x|^+} u(x,t) = \frac{1}{2}. \tag{4.1.11}$$

Consider the function

$$U(x,t) = u(x, t+x), \ \ x \geq 0, \ \ t \geq 0. \tag{4.1.12}$$

One can prove that it makes sense to consider the function $U(t,x)$ only the triangle Tr,

$$Tr = \{(x,t) : x \in (0,1), 0 < t < -2x + 2\} \tag{4.1.13}$$

Substituting $U(x,t) = u(x,t+x)$ in $u_{tt} = u_{xx} + a(x)u$ and using the initial and boundary conditions for inverse problem and $\lim_{t \to |x|^+} u(x,t) = \frac{1}{2}$, it can be obtained

$$U_{xx} - 2U_{xt} = -a(x)U, \ \ x > 0, \ \ t > 0, \tag{4.1.14}$$

$$U(x,0) = \frac{1}{2}, \tag{4.1.15}$$

$$U(0,t) = f_1(t), \ \ U_x(0,t) = f_1'(t) + f_2(t). \tag{4.1.16}$$

Since $U(x,t) \geq 1/2$ for $x \geq 0$, $t \geq 0$, then it can be considered the function

$$v(x,t) = \ln U(x,t) \tag{4.1.17}$$

and using the conditions

$$v_{xx} - 2v_{xt} + v_x^2 - 2v_x v_t = -a(x), \tag{4.1.18}$$

$$v(x,0) = -\ln 2, \tag{4.1.19}$$

$$v(0,t) = g_1(t), \;\; v_x(0,t) = g_2(t), \tag{4.1.20}$$

where

$$g_1(t) = \ln f_1(t), \;\; g_2(t) = \frac{f_1'(t) + f_2(t)}{f_1(t)}. \tag{4.1.21}$$

If it would be calculated the function $v(x,t)$, then it would be calculated the unknown coefficient $a(x)$ using $v_{xx} - 2v_{xt} + v_x^2 - 2v_x v_t = -a(x)$. Therefore, the effort below is focused on the calculation of the function $v(x,t)$.

Differentiating equation $v_{xx} - 2v_{xt} + v_x^2 - 2v_x v_t = -a(x)$ with respect to t and denote

$$q(x,t) = v_t(x,t). \tag{4.1.22}$$

Then by $v(x,0) = -\ln 2$ and $q(x,t) = v_t(x,t)$

$$v(x,t) = \int_0^t q(x,\tau)d\tau - \ln 2 \tag{4.1.23}$$

Hence, $v_{xx} - 2v_{xt} + v_x^2 - 2v_x v_t = -a(x)$ and $v(0,t) = g_1(t)$, $v_x(0,t) = g_2(t)$ lead to the folling problem for a nonlinear integro-differential equation

$$L(q) = q_{xx} - 2q_{xt} + 2q_x \int_0^t q_x(x,\tau)d\tau - 2q_x q - 2q_t \int_0^t q_x(x,\tau)d\tau = 0, \;\; x > 0, \;\; t > 0, \tag{4.1.24}$$

$$q(0,t) = g_1'(t), \;\; q_x(0,t) = g_2'(t), \;\; t > 0. \tag{4.1.25}$$

4.2 First Order Integro-Differential Equation

Introduce a new function

$$w(x,t) = q_x(x,t). \tag{4.2.1}$$

Then by (4.1.25) and (4.2.1)

$$q(x,t) = \int_0^x w(z,t)dz + g_1'(t) \tag{4.2.2}$$

Hence, we obtain an integro-differential equation in which only first order derivatives are involved. More precisely, (4.1.24) and (4.1.25) imply that

$$\begin{aligned} M(w) = w_x - 2w_t + 2w \int_0^t w(x,\tau)d\tau - \\ 2w(\int_0^x w(z,t)dz + g_1'(t)) - \\ 2(\int_0^x w_t(z,t)dz + g_1''(t)) \int_0^t w(x,\tau)d\tau = 0 \end{aligned} \tag{4.2.3}$$

$$w(0,t) = g_2'(t), \quad t > 0. \tag{4.2.4}$$

4.3 The Carleman Estimate for the Operator $\partial_x^2 - 2\partial_t\partial_x$

Let $\alpha = constant \in (0, \frac{1}{2})$. Denote

$$\psi(x,t) = -(x + \alpha t). \tag{4.3.1}$$

Since there is a need to find the function $a(x)$ for $x \in (0,1)$, it will be introduced the domain

$$\Omega_\alpha = \{(x,t) : x > 0, t > 0, -(x + \alpha t) > -1\} \tag{4.3.2}$$

$$\Omega_\alpha = \{(x,t) : x > 0, t > 0, -(x + \alpha t) < 1\} \tag{4.3.3}$$

Hence,

$$\sup_{\alpha \in (0,\frac{1}{2})} \{x : (x,t) \in \Omega_\alpha\} = 1, \tag{4.3.4}$$

$$\sup_{\alpha \in (0,\frac{1}{2})} \{t : (x,t) \in \Omega_\alpha\} = 2, \tag{4.3.5}$$

The boundary of the triangle Ω_α consists of three parts,

$$\partial\Omega_\alpha = \partial_1\Omega_\alpha \cup \partial_2\Omega_\alpha \cup \partial_2\Omega_\alpha \tag{4.3.6}$$

$$\partial_1\Omega_\alpha = \{x = 0, t \in (0, \frac{1}{\alpha})\}, \tag{4.3.7}$$

$$\partial_2\Omega_\alpha = \{x = 0 \in (0,1), t = 0\}, \tag{4.3.8}$$

$$\partial_3\Omega_\alpha = \{x > 0, t > 0, x + \alpha t = 1\}. \tag{4.3.9}$$

Lets note that $\partial_3\Omega_\alpha$ is the level curve of the function $\psi(x,t)$,

$$\psi(x,t)|_{\partial_3\Omega_\alpha} = \min_{\overline{\Omega_\alpha}} \psi(x,t) = -1. \tag{4.3.10}$$

Let $\lambda > 0$ be a parameter. Denote

$$\phi_\lambda(x,t) = \exp \lambda\psi(x,t) \tag{4.3.11}$$

Hence,

$$\phi(x,t)|_{\partial_3\Omega_\alpha} = \min_{\overline{\Omega_\alpha}} \psi(x,t) = e^{-\lambda} \tag{4.3.12}$$

Theorem 4.3.1. *There exists constants $C = C(\alpha) > 0$, $\lambda_0 = \lambda_0(\alpha) > 0$ depending only on the parameter α such that for every function $w \in H^1(\Omega_\alpha)$ such that $w(0,t) = 0$ the following Carleman estimate holds*

$$\begin{aligned} \int_{\Omega_\alpha} (w_x - 2w_t)^2 \phi_\lambda^2 dxdt \geq C\lambda^2(1-2\alpha)^2 \int_{\Omega_\alpha} w^2 \phi_\lambda^2 dxdt - \\ C\lambda e^{-2\lambda} \int_{\partial_3\Omega_\alpha} w^2 dS. \end{aligned} \tag{4.3.13}$$

Remark *By the trace theorem this Calerman estimate can be written*

$$\begin{aligned} \int_{\Omega_\alpha} (\omega_{xx} - 2\omega_{xt})^2 \phi_\lambda^2 dxdt \geq C\lambda \int_{\Omega_\alpha} (\omega_x^2 + \omega_t^2) \phi_\lambda^2 dxdt + \\ C\lambda^3 \int_{\Omega_\alpha} \omega^2 \phi_\lambda^2 dxdt - \\ C\lambda^3 \exp(-2\lambda) ||\omega||^2_{H^2(\Omega_\alpha)}. \end{aligned} \tag{4.3.14}$$

It is important for the method that an integral over $\partial_2\Omega_\alpha$ is positive.

4.4 Weighted Globally Strictly Convex Tikhonov-like Functional

First, we obtain zero Dirichlet boundary condition at $x = 0$ for the function $w(x, t)$. Denote

$$p(x, t) = w(x, t) - g_2'(t). \tag{4.4.1}$$

Then

$$p(0, t) = 0, \tag{4.4.2}$$

$$w(x, t) = p(x, t) + g_2'(t) \tag{4.4.3}$$

Hence

$$\begin{aligned} K(p) = p_x - 2p_t - 2g_2''(t) + 2(p + g_2')(\int_0^t p(x, \tau)d\tau + g_2(t) - g_2(0)) + \\ 2(p + g_2')(\int_0^x p(z, t)dz + g_1'(t) + xg_2'(t)) - \\ 2(\int_0^x p_t(z, t)dz + g_1''(t) + xg_2'(t))(\int_0^t p(x, \tau)d\tau + g_2(t)) = 0, \quad x > 0, t > 0 \end{aligned} \tag{4.4.4}$$

In addition,

$$p(0, t) = 0. \tag{4.4.5}$$

In particular, to prove the convexity of that functional, we will need to prove the following Lemma:

Lemma. *For any real valued function $f(x, t)$ the following inequalities hold:*

$$\int_{\Omega_\alpha} (\int_0^x p(z, t)dz)^2 \phi_\lambda^2(x, t)dxdt \le \frac{C}{\lambda} \int_{\Omega_\alpha} p^2(x, t)\phi_\lambda^2(x, t)dxdt, \tag{4.4.6}$$

$$\int_{\Omega_\alpha} (\int_0^t p(x, \tau)d\tau)^2 \phi_\lambda^2(x, t)dxdt \le \frac{C}{\lambda} \int_{\Omega_\alpha} p^2(x, t)\phi_\lambda^2(x, t)dxdt. \tag{4.4.7}$$

Let $T = constant > 2$. Consider the rectangle

$$G = \{(x,t) : x \in (0,1), t \in (0,T)\} \tag{4.4.8}$$

Denote

$$H_0^2(G) = \{r(x,t) \in H^2 : r(0,t) = r_x(0,t) = 0\}. \tag{4.4.9}$$

Let $R > 0$ be an arbitrary number. Consider the ball $B(R) \subset H_0^2(G)$,

$$B(R) = \{r \in H_0^2(G) : ||r||_{H^2(G)} < R\}. \tag{4.4.10}$$

Consider the following weighted Tikhonov-like functional

$$J_{\lambda,\beta}(p) = \int_G [K(p)]^2 \phi_\lambda^2 dxdt + \beta ||p||^2_{H^1(G)}, \tag{4.4.11}$$

where $\beta > 0$ is the regularization parameter.

Optimization Problem *Minimize this functional on $B(R)$.*

Let $\eta \in (0,1)$ be a sufficiently small number. Denote

$$\Omega_{\alpha,\eta} = \{(x,t) : x > 0, t > 0, x + \alpha t < 1 - \eta\}. \tag{4.4.12}$$

Hence $\Omega_{\alpha,\eta} \subset \Omega_\alpha$.

Theorem 4.4.1. *(Global strict convexity) For every $p \in H_0^2(G)$ the functional $J_{\lambda,\beta}(p)$ has its Frechét derivative $J'_{\lambda,\beta}(p)$. Let $\lambda_0 = \lambda_0(\alpha)$ be the number of previous theorem. There exists a number $\lambda_1 = \lambda_1(\alpha, R) \geq \lambda_0$ such that for every $\lambda \geq \lambda_1$ and for every $\beta \in (2e^{-\lambda}, 1)$ the functional $J_{\lambda,\beta}(p)$ is strictly convex on $\overline{B(R)}$. There exists a constant $C_1 = C_1(\alpha, R) > 0$ such that for every $\eta \in (0,1)$ and for any two points $p_1, p_2 \in \overline{B(R)}$ the following inequality holds*

$$\begin{gathered} J_{\lambda,\beta}(p_2) - J_{\lambda,\beta}(p_1) - J'_{\lambda,\beta}(p_1)(p_2 - p_1) \geq \\ C_1 e^{-2\lambda(1-\eta)} ||p_2 - p_1||^2_{H^1(\Omega_{\alpha,\eta})} + \frac{\beta}{2} ||p_2 - p_1||^2_{H^2(G)}. \end{gathered} \tag{4.4.13}$$

Next theorem follows from the previous theorem and results of [3].

Theorem 4.4.2. *Assume that conditions of previous theorem hold. Then for every $\lambda \geq \lambda_1$ and for every $\beta \in (2e^{-\lambda}, 1)$ there exists unique minimizer $p_{\lambda,\beta,\min} \in \overline{B(R)}$ of the functional $J_{\lambda,\beta}(p)$ on the ball $\overline{B(R)}$.*

In the standard way, it will be considered the data functions g_1', g_2', g_1'', g_2'', with the noise of the level σ as

$$||g_i' - (g_i')^*||_{L_2(0,T)}, \ \ ||g_i'' - (g_i'')^*||_{L_2(0,T)} \leq \sigma, \ \ i = 1, 2, ... \tag{4.4.14}$$

where "*" means noiseless data. Let $Q_{\overline{B}} : H_0^2(G) \to \overline{B(R)}$ be the operator of the projection of the space $H_0^2(G)$ on the closed ball $\overline{B(R)}$. Let $\gamma = const > 0$ and let $p^{(0)}$ be an arbitrary point of $\overline{B(R)}$. Consider the sequence of the gradient projection method,

$$p^{(n+1)} = Q_{\overline{\mathbb{B}}}(p^{(n)} - \beta J'_{\lambda,\beta}(p^{(n)}), \ \ n = 0, 1, 2, ... \tag{4.4.15}$$

Theorem 4.4.3. *Let $p^* \in B(R)$ be the exact solution of the problem which corresponds to the noiseless data. let $\lambda_1 > 0$ be the constant. Suppose that $\lambda_1 < \ln(\sigma^{-\epsilon})$ for a certain constant $\epsilon \in (0, 1)$. Choose $\lambda = \ln(\sigma^{-\epsilon})$ and $\beta = 2e^{-\lambda} = 2\sigma^{\epsilon}$. Then there exists a sufficiently small number $\gamma_0 = \gamma_0(\alpha, R, \sigma) \in (0, 1)$ such that for every $\gamma \in (0, \gamma_0)$ there exists a number $\mu = \mu()\gamma \in (0, 1)$ such that the following estimate holds*

$$||p^{(n)} - p^*||_{H^1(\Omega_{\alpha,\eta})} \leq C_2\sigma^{\nu} + \mu^n||p^0 - p_{\lambda,\beta,\min}||_{H^2(G)}, n = 1, 2, ... \tag{4.4.16}$$

where $\nu \in (0, \epsilon)$ is a certain number and the constant $C_2 = C_2(\alpha, R)$ is a constant depending only on listed parameters.

Remark *Since the starting point p^0 of the sequence is an arbitrary point of the ball $\overline{B(R)}$ and and since restrictions on R are not imposed, then the last theorem claims the global convergence to the exact solution of that sequence.*

Chapter 5

INVERSE PROBLEM FOR NONLINEAR MODIFIED TRANSPORT EQUATION WITH FINAL OVERDETERMINATION

5.1 Introduction

5.1.1 Preliminaries

The chapter investigates the local unique solvability of the inverse problem with a final overdetermenation for a nonlinear modified transport equation. These tasks can be considered as controllability problems, in which the control is, for example, a multiplier on the right side, or in some other coefficients, depending only on spatial variables in the case of the final overdetermenation and depending only on time in the case of an integral overdetermenation. The transport equation describes the processes of neutron transport in nuclear reactors [4], [25] the propagation of γ-radiation and light in the atmospheres of stars and planets [13], and also used in questions related to tomography and some other applied research. These processes are studied both from the point of view of direct and inverse problems. Historically, it has been that the direct problems of the linear transport theory were the first to be developed, which were especially intensively studied in the 60-70s of the last century in connection with the introduction of nuclear energy. The works devoted to these problems include, first of all, the works of K. Jorgens , G.I. Marchuk, V.S. Vladimirov, T.A. Germogenov K.M. Keyza , P.F. Zweifel, E.B. Shikhov, M.V. Maslennikova, Yu. L. Kuznetsov and a number of others.

The importance of studying inverse problems for differential equations of mathematical physics and, in particular, for the transport equation is explained by their pronounced applied orientation, noted above. Therefore, at present, the need to address the reverse problems of mathematical physics constantly appear both among theoreticians and applied scientists.

The study of inverse problems for linear transport equations began long ago. The first papers are considered to be [26], [29]. This was followed by a paper devoted, in particular, to the uniqueness of the solution of multidimensional inverse problems for a stationary one-velocity linear transport equation. For a nonstationary multivelocity transport equation.

$$\begin{aligned}&u_t(x,v,t)+(v,\nabla)u(x,v,t)+\sum(x,v,t)u(x,v,t)=\\&\int\limits_V J(x,v',t,v)u(x,v',t)dv'+F(x,v,t)\end{aligned} \tag{5.1.1}$$

the existence and uniqueness theorems for solutions of inverse problems in the class of functions that are continuous together with their derivatives $\frac{\partial u}{\partial t}$ and $(v,\nabla)u$ were obtained in [31], [30]. Similar theorems were proved by the semigroup method. In the questions of existence and uniqueness of generalized solutions of inverse problems for a nonstationary multivelocity linear transport equation with integral overdetermination are studied. As mentioned earlier, the correctness of the formulation of the corresponding direct problems is also studied. Referring to [37], we note that we consider inverse problems for a nonstationary multivelocity anisotropic kinetic modified transport equation (the linear case of the kinetic Boltzmann equation)

$$\begin{aligned}&u_t(x,v,t)+(v,\nabla)u(x,v,t)+\sum(x,v,t)u_0(x,v,t)=\\&\int\limits_V J(x,v',t,v)u(x,v',t)dv'+F(x,v,t),\quad \text{where}\quad (x,v,t)\in D=G\times V\times(0,T)\end{aligned} \tag{5.1.2}$$

Here the region G is the region of change of spatial coordinates, i.e. the area in which the process under study occurs, V - this is the area of change in the velocities of the particles under study, $(0,T)$ - time interval for monitoring the process. The function $u=u(x,v,t)$ characterizes the distribution density of particles in the phase space $G\times V$ at time $T\in(0,T)$, i.e., flying through point x with speed v at time t, the absorption coefficient $\Sigma=\Sigma(x,v,t)$ and the scattering indicatrix $J=J(x,v,t)$ characterize the properties of the medium in which the given process, $F=F(x,v,t)$ is a function of the radiation sources.

The transport equations are called: **1.** Stationary if the functions u,Σ,J,F do not depend on t, **2.** Single-velocity if V is the unit sphere, **3.** Isotropic if the coefficients Σ and J depend only on x and t.

An inverse problem is a problem where the function u and one of the parameters F, Σ or J are simultaneously determined from equation (1) of the initial condition

$$u(x, v, 0) = \varphi(x, v), \quad (x, v) \in \bar{G} \times V, \tag{5.1.3}$$

and boundary condition

$$u(x, v, t) = \mu(x, v, t), \quad (x, v, t) \in \gamma_- \times [0, T] \tag{5.1.4}$$

$\gamma_- = \{(x, v) \in \partial G \times V : (v, n_x) < 0\}$ and n_x — external normal to the boundary ∂G field of G at x.

The condition of final overdetermination:

$$u(x, v, T) = \psi(x, v), \quad (x, v) \in \bar{G} \times V \tag{5.1.5}$$

In this paper, we prove the local unique solvability of the inverse problem with a final redefinition for a nonlinear modified transport equation, which in this formulation can be considered as a local controllability problem, which consists in transferring the system from a zero state to a predetermined state in a fixed time interval. We write the function of radiation sources in the following form:

$$F(x, v, t) = f(x, v)\, g(x, v, t) + h(x, v, t), \tag{5.1.6}$$

the control here is $f(x, v)$, which depends on space variables, and g, h are given functions. In this case, the control is the multiplier on the right side, which depends only on spatial variables. A similar result for the linear problem of the transport equation was obtained earlier in the work of N.P. Volkov [2].

5.1.2 Preliminaries

In this subsection we recall some definitions and theorems that will be needed in what follows.

Definition 1. Let $f : X \to Y$. The mapping f is called strictly differentiable at the point x_0 if

$$f(x+h) - f(x) = f'(x_0)h + R(x,h)$$

where $\frac{||R(x,h)||}{||h||} \to 0$ as $x \to x_0$, $h \to 0$.

Theorem 5.1.1. *Let X, Y be Banach spaces, U be open that set in X, and the map $f : U \to Y$ is strictly differentiable on U and $f'(x_0) : X \to Y$ is an isomorphism of X onto Y for some point $x_0 \in U$. Then there is a neighborhood U' of the point x' such that f realizes a homeomorphism U' onto an open set $f(U')$, $f'(x)$ is an isomorphism of X onto Y for $x \in U'$. $f^{-1} : f(U') \to X$ is strictly differentiable on $f(U')$ and $(f^{-1})'f(x) = [f'(x)]^{-1}$ for $x \in U'$*

Theorem 5.1.2. *Let X be a Banach space, Y a separable topological vector space, $A : X \to Y$ a linear continuous operator, U the open unit ball in X, $PAU : AX \to [0, \infty[$ is the Minkowski functional of the set AU, and the mapping $\Psi : X \to AX$ satisfies the condition*

$$P_{AU}(\Psi(x) - \Psi(\bar{x})) \leq \nu(r)||x - \bar{x}||$$

for $||x - x_0|| \leq r$, and $||\bar{x} - x_0|| \leq r$,

for some $x_0 \in X$, where the function $\Theta : [0, \infty[\to [0, \infty[$ is non-decreasing. Let $b(r) = max(1 - \nu(r), 0)$ for $r \geq 0$. Let $\omega = \int_0^\infty b(r)dr \in]0, \infty]$, $r_* = \sup r \geq 0/b(r) > 0$, $\omega(r) = \int_0^t b(r)dt$ $(r \geq 0)$ and $f(x) = Ax + \Psi(x)$ for $x \in X$. Then $\forall r \in [0, r_*[$, $\forall y \in f(x_0) + \omega(r)AU$ $\exists x \in x_0 + rU : f(x) = y$.

Remark. *If in this theorem A is injective or $KerA$ has such some topological complement of E in X such that $A(E \cap U) = AU$, then its assertion follows from the contraction mapping principle. In particular, if A is injective, then the solution is unique.*

In what follows, we will also need Hadamard's theorem.

Theorem 5.1.3. *Let X, Y be Banach spaces, $A : X \to Y$ is an isomorphism of X onto Y $(AX = Y$, $||A|| < \infty$, $\exists A^{-1}$, $||A^{-1}|| < \infty)$, $f : X \to Y$ satisfies the Lipschitz condition with* *constant* $q/||A^{-1}||$, *where* $q < 1$. *Let* $F(x) = Ax + f(x)$. *Then* $F(x) = Y$ *and* F *is bijective and* $||F^{-1}(y) - F^{-1}(\bar{y})|| \leq \frac{||A^{-1}||}{1-q}||y - \bar{y}||$.

Introduce the following functional spaces and key:

Let the domain $G \subset \mathbb{R}^n$ is strictly convex, and limited closed set of V is contained in a spherical layer $\{v \in \mathbb{R}^n : 0 < v_0 \leqslant |v| \leqslant v_1 < \infty\}$.

$L_\infty\big(D, L_2(V)\big)$ — space classes essentially bounded functions on D with values in $L_2(V)$, where $D = G \times V \times (0, T)$.

The space $H_2(D) = \{u \in L_2(D) : u_t, (v, \nabla)u \in L_2(D), u|_{\Gamma_-} \in L_2(\Gamma_-)\}$ of functions u, summable in a square with their generalized derivatives $u_t, (v, \nabla)u$ on D and after $u|_{\Gamma_-}$ of $L_2(\Gamma_-)$ where $\Gamma_- = \gamma_- \times [0, T]$ and $\gamma_- = \{(x, v) \in \partial G \times V : (v, n_x) < 0\}$ and n_x — external normal to the boundary ∂G field of G at x. This space is a Banach norm on

$\|u\|_{H_2(D)} = [\|u\|^2_{2.D} + \|u_t\|^2_{2.D} + \|(v, \nabla)u\|^2_{2.D} + \|u|_{\Gamma_-}\|^2_{2.\Gamma_-}]^{1/2}$,

The space $H_\infty(D) = \{u \in L_\infty(D) : u_t, (v, \nabla)u \in L_\infty(D), u|_{\Gamma_-} \in L_\infty(\Gamma_-)\}$ essentially bounded on D functions u with their generalized derivatives $u_t, (v, \nabla)u$ and a trace $u|_{\Gamma_-}$ of $L_\infty(\Gamma_-)$, where $\Gamma_- = \gamma_- \times [0, T]$ and $\gamma_- = \{(x, v) \in \partial G \times V : (v, n_x) < 0\}$ and n_x — external normal to the boundary ∂G field of G at x. This space is a Banach norm on

$\|u\|_{H_\infty(D)} = [\|u\|_{\infty.D} + \|u_t\|_{\infty.D} + \|(v, \nabla)u\|_{\infty.D} + \|u|_{\Gamma_-}\|_{\infty.\Gamma_-}]$,

$W^t_2(D) =$
$\left\{F(x, v, t) \in L_2(D) : \dfrac{\partial F}{\partial t} \in L_2(D)\right\}$ with the norm $\|F\|_{W^t_2(D)} = \left[\|F\|^2_{L_2(D)} + \|\dfrac{\partial F}{\partial t}\|^2_{L_2(D)}\right]^{1/2}$,

$W^t_\infty(D) = \left\{F(x, v, t) \in L_\infty(D) : \dfrac{\partial F}{\partial t} \in L_\infty(D)\right\}$ with the norm $\|F\|_{W^t_\infty(D)}$
$= \left[\|F\|_{L_\infty(D)} + \|\dfrac{\partial F}{\partial t}\|_{L_\infty(D)}\right]$,

$h_2(G \times V) = \{\varphi \in L_2(G \times V) : (v, \nabla)\varphi \in L_2(G \times V), \varphi|_{\gamma_-} \in L_2(\gamma_-)\}$ with the norm

$\|\varphi\|_{h_2} = [\|\varphi\|^2_{L_2(G\times V)} + \|(v, \nabla)\varphi\|^2_{L_2(G\times V)} + \|\varphi|_{\gamma_-}\|^2_{L_2(\gamma_-)}]^{1/2}$, $C_{X\to Y}$ — smallest constant embedding X in Y, ie, $\|\cdot\|_Y \leqslant C_{X\to Y}\|\cdot\|_X$. $\mathcal{L}(X, Y)$ — many continuous linear operators from X in Y.

$h_\infty(G \times V) = \{\varphi \in L_\infty(G \times V) : (v, \nabla)\varphi \in L_\infty(G \times V), \varphi|_{\gamma_-} \in L_\infty(\gamma_-)\}$ with the norm

$\|\varphi\|_{h_\infty} = [\|\varphi\|_{L_\infty(G\times V)} + \|(v, \nabla)\varphi\|_{L_\infty(G\times V)} + \|\varphi|_{\gamma_-}\|_{L_\infty(\gamma_-)}]$, $C_{X\to Y}$ — smallest constant embedding X in Y, ie, $\|\cdot\|_Y \leqslant C_{X\to Y}\|\cdot\|_X$. $\mathcal{L}(X, Y)$ — set of linear continuous operators

X in Y.

In this section, we study generalized solutions of the direct problem in the class $H_2(D)$, taking instead of the modified equation the more general equation

$$\frac{\partial u}{\partial t} + (v, \nabla)u(x, v, t) = (Pu)(x, v, t) + F(x, v, t) \tag{5.1.7}$$

where the operator P satisfies the following conditions:

1. $P : X \to Y$, where X, Y are real Banach spaces,
2. P is a Lipschitz-continuous operator uniformly with respect to $t \in (0, T)$, that is, there exists a Lipschitz constant p independent of t such that for any $u_1, u_2 \in X$ the condition

$$||(Pu_1 - Pu_2)(*, *, t)|| \leq p||(u_1 - u_2)(*, *, t)||$$

here by $|| * ||$ in the case $X = H_2(D)$ we mean the norm $||u(*, *, t)||_{W_2^t(G\times V)}$.

Note that the operator P in the considered problem can be non-linear as long as the initial and boundary conditions are satisfied for it. In order to avoid piling up in the future, we will consider instead of a neighborhood of the point u_0 a neighborhood of zero.

Subtracting

$$\begin{aligned} &u_t(x, v, t) + (v, \nabla)u(x, v, t) + \sum(x, v, t)u_0(x, v, t) = \\ &\int_V J(x, v', t, v)u(x, v', t)dv' + F(x, v, t) \end{aligned} \tag{5.1.8}$$

and

$$\begin{aligned} &\tilde{u}_t(x, v, t) + (v, \nabla)\tilde{u}(x, v, t) + \sum(x, v, t)u_0(x, v, t) = \\ &\int_V J(x, v', t, v)\tilde{u}(x, v', t)dv' + \tilde{F}(x, v, t) \end{aligned} \tag{5.1.9}$$

we get

$$\begin{aligned} &\hat{u}_t(x, v, t) + (v, \nabla)\hat{u}(x, v, t) = \\ &\int_V J(x, v', t, v)\hat{u}(x, v', t)dv' + \hat{F}(x, v, t) \end{aligned} \tag{5.1.10}$$

Theorem 5.1.4. ***(Forward Problem for the Modified Transport Equation)*** *Let* $X = H_2(D)$, $Y = W_2^t(D)$, $F \in Y$, $\mu \in W_2^t(\Gamma_-)$ *and* $\phi \in h_2(G \times V)$ *and the matching conditions are met* $\phi(x,v) = \mu(x,v,0)$ *for almost all* $(x,v) \in \gamma_-$. *Then there is a unique generalized solution* $u \in H_2(D)$ *of the problem (5.1.2)-(5.1.4) satisfying a priori estimate*

$$||u||_{H_2(D)} \leq C(||F||_{W_2^t(D)} + ||\phi||_{h_2} + ||\mu||_{W_2^t(\Gamma_-)}) \tag{5.1.11}$$

Theorem 5.1.5. ***(Forward Problem for the Modified Transport Equation)*** *Let* $X = H_2(D)$, $Y = W_2^t(D)$, $F \in Y$, $\mu \in W_2^t(\Gamma_-)$ *and* $\phi \in h_2(G \times V)$ *and the matching conditions are met* $\phi(x,v) = \mu(x,v,0)$ *for almost all* $(x,v) \in \gamma_-$. *Then there is a unique generalized solution* $\tilde{u} \in H_2(D)$ *of the problem (5.1.9), (5.1.3) and (5.1.4) satisfying a priori estimate*

$$||\tilde{u}||_{H_2(D)} \leq C(||F||_{W_2^t(D)} + ||\phi||_{h_2} + ||\mu||_{W_2^t(\Gamma_-)}) \tag{5.1.12}$$

Theorem 5.1.6. ***(Forward Problem for the Modified Transport Equation)*** *Let* $X = H_2(D)$, $Y = W_2^t(D)$, $F \in Y$, $\mu \in W_2^t(\Gamma_-)$ *and* $\phi \in h_2(G \times V)$ *and the matching conditions are met* $\phi(x,v) = \mu(x,v,0)$ *for almost all* $(x,v) \in \gamma_-$. *Then there is a unique generalized solution* $\hat{u} \in H_2(D)$ *of the problem (5.1.10), (5.1.3) and (5.1.4) satisfying a priori estimate*

$$||\hat{u}||_{H_2(D)} \leq C(||F||_{W_2^t(D)} + ||\phi||_{h_2} + ||\mu||_{W_2^t(\Gamma_-)}) \tag{5.1.13}$$

The method of proving theorems on the direct problem for the modified transport equation does not essentially differ from the proof of the theorem of the direct problem of the transport equation N.P. Volkov [2].

5.2 Production of the inverse problem for the modified nonlinear transfer equation in the case where $F(x,v,t) = f(x,v)\,g(x,v,t)$ *and controllability is* $f(x,v)$ *in* $H_2(D) \times L_2(G \times V)$.

We consider the inverse problem

$$\begin{gathered} u_t(x,v,t) + (v,\nabla)u(x,v,t) + \sum(x,v,t)u_0(x,v,t) + [S(u)](x,v,t) = \\ \int\limits_V J(x,v',t,v)u(x,v',t)dv' + F(x,v,t), \quad \text{where} \quad (x,v,t) \in D \end{gathered} \tag{5.2.1}$$

where the nonlinear part

$$[S(u)](x,v,t) = \int\limits_0^T \int\limits_{G\times V} Q(x,v,x',v',t)\,\alpha\big(u(x',v',t')\big)\,dx'\,dv'\,dt'\,, \tag{5.2.2}$$

with conditions

$$u(x,v,t) = 0, \quad (x,v,t) \in \gamma_- \times [0,T], \tag{5.2.3}$$

$$u(x,v,0) = 0, \quad (x,v) \in \bar{G} \times V \tag{5.2.4}$$

$$u(x,v,T) = \psi(x,v), \quad (x,v) \in \bar{G} \times V \tag{5.2.5}$$

In order to avoid further accumulation will consider, instead of around the point u_0 neighborhood of zero. Namely, subtracting from the equation

$$u_t(x,v,t) + (v,\triangledown)u(x,v,t) + \Sigma(x,v,t)\,u_0(x,v,t) =$$
$$= \int_V J\,(x,v,t,v')\,u(x,v',t)\,dv' + F(x,v,t),$$

equation

$$\tilde{u}_t(x,v,t) + (v,\triangledown)\tilde{u}(x,v,t) + \Sigma(x,v,t)\,u_0(x,v,t) =$$
$$= \int_V J\,(x,v,t,v')\,\tilde{u}(x,v',t)\,dv' + \tilde{F}(x,v,t),$$

get

$$\hat{u}_t(x,v,t) + (v,\triangledown)\hat{u}(x,v,t) =$$
$$= \int_V J\,(x,v,t,v')\,\hat{u}(x,v',t)\,dv' + \hat{F}(x,v,t).$$

Let the source function is represented as a

$$F(x,v,t) = f(x,v)\,g(x,v,t), \tag{5.2.6}$$

where u and f — unknown function and g, ψ — apriori defined function, which can be viewed as a manageable task, namely, translate the system of $u(x,v,0) = 0$ in the final overdetermination condition $u(x, v, T) = \psi(x, v)$ through controllability of $f(x, v)$. In our work we prove that there exists a unique solution (u, f) this problem in the neighborhood of the point u_0 in the corresponding functional space (all the symbols and terms will be introduced below) for sufficiently small in the norm $\psi(x, v)$. Here and further $\nabla = \nabla_x$.

In our case, [37] describes the process of mass transfer multi anisotropic modified kinetic equation:

$$\begin{aligned} u_t(x, v, t) + (v, \nabla)\, u(x, v, t) + \sum(x, v, t)\, u_0(x, v, t) = \\ = \int\limits_V J(x, v, t, v')\, u(x, v', t)\, dv' + F(x, v, t), \\ (x, v, t) \in G \times V \times (0, T), \end{aligned} \tag{5.2.7}$$

in which the function $u(x, v, t)$ describes the density distribution particles in the phase space $G \times V$ at time $t \in (0, T)$, a function $\sum(x, v, t), J(x, v, t, v')$ and $F(x, v, t)$ – environment in which this process proceeds, as with coefficient of absorption, scattering indicatrix, and as a source respectively. The function u_0 is fixed. The same equation cause the problem of radiation of charged particles, as well as spread of γ-radiation. As shown in the [32], if you asked all the characteristics of the environment Σ, J, F, and as " incoming flow", i.e

$$u(x, v, t) = \mu(x, v, t), \quad (x, v, t) \in \gamma_- \times [0, T], \tag{5.2.8}$$

and the initial state of the process, ie.

$$u(x, v, 0) = \varphi(x, v), \quad (x, v) \in \bar{G} \times V, \tag{5.2.9}$$

the state of $u(x, v, t)$ can be defined unambiguously at any time t, and there is derived estimate accuracy

$$\begin{aligned} \|u\|_{H_2(D)} \leqslant C\Big(\|F\|_{W_2^t(D)} + \|\varphi\|_{h_2(G\times V)} + \|\mu\|_{W_2^t(\Gamma_-)} + \\ + M(V)\|\sigma\|_{L_2(G\times V)} + M(V)\|u_0\|_{H_2(D)}\Big). \end{aligned} \tag{5.2.10}$$

where $M(V)$ - action sets V. Consider the inverse problem for a linear derivative of the transport equation using the sources. An interesting case where the controls are allowed stationary, ie do not change over time. Suppose that

$$F(x, v, t) = f(x, v)\, g(x, v, t) + h(x, v, t), \tag{5.2.11}$$

where f — desired, and g, h — priori defined functions. From a physical point of view, the task of controllability is as follows: Can you find such a function $f(x,v)$ and the corresponding function $u(x,v,t)$, satisfying the conditions (5.2.7)–(5.2.9), so that at time $t=T$ distribution function was equal to $\psi(x,v)$. In such a setting, this task is the task of managing non-transfer using the source of F. A generalized solution of class $H_2(D)\times L_2(G\times V)$ of the inverse problem (5.2.7) will understand a pair of functions (u,f): $u\in H_2(D)$ and $f\in L_2(G\times V)$, satisfying almost everywhere the conditions (5.2.7)–(5.2.9),(5.2.5).

Theorem 5.2.1. ***(Unique solvability of the inverse problem for linear modified transport equation in*** $H_2(D)\times L_2(G\times V)$*).*

Let $\sum,\sum_t\in L_\infty(D)$*;* $J,J_t\in L_\infty\big(D,L_2(V)\big)$*;* $u_0\in H_2(D)$*,* $u_{0_t}\in L_2(D)$*,* $g,g_t\in L_\infty(G\times V,L_2(0,T))$*,* $|g(x,v,T)|\geq g_0>0$*;* $h,h_t\in L_2(D)$*;* $\mu,\mu_t\in L_2(\Gamma_-)$*;* $\varphi,(v,\nabla)\varphi,\psi,(v,\nabla)\psi\in L_2(G\times V)$*,* $\varphi|_{\gamma_-},\psi|_{\gamma_-}\in L_2(\gamma_-)$*;* $v_0^{-1}diamG<a$*, and harmonization of the conditions* $\varphi(x,v)=\mu(x,v,0)$ *and* $\psi(x,v)=\mu(x,v,T)$ *for almost all* $(x,v)\in\gamma_-$*. Then the controllability of the problem has a unique solution* $\{u,f\}\in H_2(D)\times L_2(G\times V)$*.*

As noted above, to facilitate the presentation will consider, instead of around the point u_0 neighborhood of zero. The method of proof of this theorem is not fundamentally different from the evidence Theorems on unique solvability of the inverse problem for a linear transport equation N.P.Volkov [36].

5.2.1 Statement of the main result.

Since the operator

$$\hat{\hat{\mathcal{A}}}:L_2(G\times V)\longrightarrow h_2(G\times V),$$

$$f\mapsto\psi,$$

linear and bijective (because the solution is unique), the Banach theorem $\hat{\hat{\mathcal{A}}}$ — isomorphism, i.e

$$\hat{\hat{\mathcal{A}}}^{-1}:h_2(G\times V)\longrightarrow L_2(G\times V),$$

is continuous and, thus, $\forall\,\psi\ \exists\,!$ solution f linear inverse problem, the

$$\|f\|_{L_2(G\times V)}\leqslant\bar{C}\|\psi\|_{h_2(G\times V)},\tag{5.2.12}$$

for some $\bar{C} > 0$ ($\bar{C}$ depends on $\sum, J, g, h, \mu, G, V$, as well as of the norms of differential and integral operators, introduced in the proof of the theorem).

Let $\mu(x,v,t) = 0, (x,v,t) \in \gamma_- \times [0,T]$; Let $\mu(x,v,t) = 0, (x,v,t) \in \gamma_- \times [0,T]$; $\varphi(x,v) = 0,\ (x,v) \in \bar{G} \times V$; $h(x,v,t) \equiv 0$ on D.

Let $Lu = u_t + (v,\nabla)\,u - \int\limits_V Ju\,dv' = -\sum u_0 + F,\ H = \{u \in H_2(D) | \exists F \in W_2^t(D)$: u — solution of the problem (5.2.7)–(5.2.9) with the norm $\|u\|_H = \|Lu\|_{W_2^t(D)}$. Then $L : H \to W_2^t(D)$ — isometric isomorphism of H on $W_2^t(D)$ [32] (5.2.10), H — fully and continuously invested in the $H_2(D)$. From there, it should be that $\exists \hat{C} > 0,\ \forall t_1 \in]0,T],\ \forall u \in H : \|u|_{t=t_1}\|_{L_2(G\times V)} \leqslant \hat{C}\|u\|_H$. Note also that the theorem on the trace for $H_2(D)$ the operator

$$\begin{gathered} \Lambda_u :\ H \longrightarrow L_2(G\times V), \\ u \mapsto u_t|_{t=t_1}, \end{gathered} \tag{5.2.13}$$

continuous because of the continuity of embedding H in $H_2(D)$. Let $\chi(x,v) = f(x,v)\,g(x,v,T)$.

Everywhere in the future will be taken prior designations. Let $\alpha : \mathbb{R} \to \mathbb{R}$ twice continuously differentiable on $\mathbb{R}$, with

$$\begin{cases} |\alpha(u)| \leqslant C_1 |u|\,, \\ |\alpha'(u)| \leqslant C_1\,, \quad \alpha'(0) = 0\,, \\ |\alpha''(u)| \leqslant C_2\,. \end{cases} \tag{5.2.14}$$

Then, $Q(x,v,x',v',t)$ continuous on $\bar{G} \times \bar{V} \times \bar{G} \times \bar{V} \times [0,T]$.

Introduce the operator S: $H_2(D) \to W_2^t(D)$,

$[S(u)](x,v,t) = \int\limits_0^T \int\limits_{G\times V} Q(x,v,x',v',t)\,\alpha\big(u(x',v',t')\big)\,dx'dv'dt'$.

Theorem 5.2.2. ***(Unique solvability of the inverse problem for the nonlinear modified transport equation in*** $H_2(D) \times L_2(G\times V)$***).***

Let $\sum, \sum_t \in L_\infty(D),\ |\sum(x,v,T)| \geq \sum_0 > 0,\ (x,v) \in G \times V; J, J_t \in L_\infty\big(D, L_2(V)\big)$*;* $u_0 \in H_2(D)\ u_{0_t} \in L_2(D),\ g, g_t \in L_\infty(G\times V, L_2(0,T)),\ |g(x,v,T)| \geq g_0 > 0$*;* $\psi, (v,\nabla)\psi \in L_2(G\times V),\ \psi|_{\gamma_-} \in L_2(\gamma_-)$*;* $v_0^{-1} diam G < a$ *and condition* $\psi(x,v) = 0$ *when* $(x,v) \in \gamma_-$*. Then the problem*

$$
\begin{aligned}
&u_t(x,v,t) + (v,\nabla)u(x,v,t) + \sum(x,v,t)u_0(x,v,t) + [S(u)](x,v,t) = \\
&\int\limits_V J(x,v',t,v)u(x,v',t)dv' + F(x,v,t), \quad \textit{where} \quad (x,v,t) \in D
\end{aligned}
\tag{5.2.15}
$$

$$
u(x,v,t) = 0, \quad (x,v,t) \in \gamma_- \times [0,T], \tag{5.2.16}
$$

$$
u(x,v,0) = 0, \quad (x,v) \in \bar{G} \times V \tag{5.2.17}
$$

$$
u(x,v,T) = \psi(x,v), \quad (x,v) \in \bar{G} \times V \tag{5.2.18}
$$

$$
F(x,v,t) = f(x,v)\, g(x,v,t), \tag{5.2.19}
$$

where f — desired, and g — apriori given function, has unique solution (u, f) in the neighborhood of the point u_0 in $H_2(D) \times L_2(G \times V)$ for sufficiently small in the norm $\psi \in L_2(G \times V)$.

Remark 5.1. *Proof of Theorem (5.2.2) is reduced to checking the feasibility conditions of the theorem on the inverse functions. The main points are the verification of the strict differentiability of the corresponding mappings and the choice of function spaces associated with problem (5.1.2)-(5.1.5) .*

5.2.2 Proof.

According to a theorem about the inverse function (5.1.1) there exist open neighborhood U_1 of H and V_1 of $W_2^t(D)$ points of u_0 and $\xi(u_0)$, respectively, the operator $\xi : u \mapsto Lu + S(u)$ is implementing diffeomorphism U_1 on V_1 of class C^1. In doing so, the operator $\eta = \xi^{-1} : V_1 \to U_1$ is strictly differentiate on V_1 as a mapping on H, and hence in $H_2(D)$ and $\eta'(F) = \left[\xi'\big(\eta(F)\big)\right]^{-1}$ at $F \in V_1$. In this case, by virtue of linearity and continuity operator $\chi(x,v) \mapsto \chi(x,v)g(x,v,t)/g(x,v,T)$ of $L_2(G \times V)$ in $W_2^t(D)$ it follows that the operator $P : L_2(G \times V) \to H_2(D)$, $\chi(x,v) \mapsto P(\chi) = \eta\left(\chi(x,v)g(x,v,t)/g(x,v,T)\right)$ is strictly for Frechet differentiate in neighborhood of $V_2 \subset L_2(G \times V)$ the point $P(u_0)$, and then the operators

$$\begin{aligned}
&O_1 : V_2 \to L_2(G \times V),\\
&\qquad \chi \mapsto \big(P(\chi)\big)_t\Big|_{t=T}, \qquad (5.2.20)\\
&O_2 : V_2 \to L_2(G \times V),\\
&\qquad \chi \mapsto (v, \nabla)\big(P(\chi)\big)\Big|_{t=T}, \qquad (5.2.21)\\
&O_3 : V_2 \to L_2(G \times V),\\
&\qquad \chi \mapsto \int\limits_V JP(\chi)\, dv'\Big|_{t=T}, \qquad (5.2.22)\\
&O_4 : V_2 \to L_2(G \times V),\\
&\qquad \chi \mapsto S\big(P(\chi)\big)\Big|_{t=T}, \qquad (5.2.23)
\end{aligned}$$

also strongly differentiable on V_2. Next, we will seek a solution nonlinear problem (5.1.2)-(5.1.5) in the form of $u = P(\chi)$, where u — solution of nonlinear problem (5.1.2)-(5.1.5) on the right part of $F(x, v, t) = \chi(x, v)g(x, v, t)/g(x, v, T)$. Introduce the operator

$$M(\chi) = \Bigg\{\chi - \big(P(\chi)\big)_t\Big|_{t=T} - (v, \nabla)\big(P(\chi)\big)\Big|_{t=T} - S\big(P(\chi)\big)\Big|_{t=T} + \\ + \Bigg(\int\limits_V JP(\chi)\, dv'\Bigg)\Bigg|_{t=T}\Bigg\} \times \left(\frac{(P(\chi))\big|_{t=T}}{\sum(x, v, T)u_0}\right), \quad (5.2.24)$$

$$M'(u_0)\chi = \Bigg\{\chi - \big(P'(u_0)\chi\big)_t\Big|_{t=T} - (v, \nabla)\big(P'(u_0)\chi\big)\Big|_{t=T} + \\ + \Bigg(\int\limits_V J\big(P'(u_0)\chi\big)\, dv'\Bigg)\Bigg|_{t=T}\Bigg\} \times \left(\frac{(P(\chi))\big|_{t=T}}{\sum(x, v, T)u_0}\right) + \\ + \Bigg\{\chi - \big(P(\chi)\big)_t\Big|_{t=T} - (v, \nabla)\big(P(\chi)\big)\Big|_{t=T} - S\big(P(\chi)\big)\Big|_{t=T} + \\ + \Bigg(\int\limits_V JP(\chi)\, dv'\Bigg)\Bigg|_{t=T}\Bigg\} \times \left(\frac{(P'(u_0)\chi)\big|_{t=T}}{\sum(x, v, T)u_0}\right), \quad (5.2.25)$$

where $P'(u_0)\chi = \eta'(u_0)\,(\chi(x,v)g(x,v,t)/g(x,v,T))$, i.e direct solution of linear problems with the right part of $F(x,v,t) = \chi(x,v)g(x,v,t)/g(x,v,T)$ (solution exists, as shown in [32], and the operator $\eta'(u_0)$ is continuous, because is estimation accuracy). Substituting in (5.1.2) $u|_{t=T}$ of (5.1.5) and $F(x,v,t) = \chi(x,v)g(x,v,t)/g(x,v,T)$, we arrive to the equation

$$M(\chi) = \psi(x,v), \tag{5.2.26}$$

which by virtue of the continuity of $M'(u_0)$, as well as the existence and boundedness of $[M'(u_0)]^{-1}$ (from the solvability linear inverse problem) can be applied theorem (5.1.1). According to the theorem (5.1.1) there exist an open neighborhood U_* in $L_2(G\times V)$ and V_* in $h_2(G\times V)$ of points of u_0 and $M(u_0)$, respectively, that $M:\chi\to\psi$ is implementing diffeomorphism U_* on V_* class C^1. Therefore, $\forall\,\psi\in V_*\ \exists!\ \chi\in U_* : M(\chi)=\psi$. The theorem is proved.

A proof of the local unique solvability in $H_2(D)\times L_2(G\times V)$ inverse problem with final overdetermination for the modified nonlinear equations transport in the case where $F(x,v,t) = f(x,v)\,g(x,v,t)$ and controllability is $f(x,v)$.

5.3 Production of the inverse problem for the modified nonlinear transfer equation in the case where $F(x,v,t) = f(x,v)\,g(x,v,t)$ and controllability is $f(x,v)$ in $H_\infty(D)\times L_\infty(G\times V)$.

We consider the inverse problem

$$\begin{aligned} &u_t(x,v,t) + (v,\nabla)u(x,v,t) + \sum(x,v,t)u_0(x,v,t) + [S(u)](x,v,t) = \\ &\int_V J(x,v',t,v)u(x,v',t)dv' + F(x,v,t), \quad \text{where} \quad (x,v,t)\in D \end{aligned} \tag{5.3.1}$$

where the nonlinear part

$$[S(u)](x,v,t) = \int_0^T \int_{G\times V} Q(x,v,x',v',t)\,\alpha\big(u(x',v',t')\big)\,dx'\,dv'\,dt'\,, \tag{5.3.2}$$

with conditions

$$u(x,v,t) = 0, \quad (x,v,t)\in\gamma_-\times[0,T], \tag{5.3.3}$$

$$u(x,v,0) = 0, \quad (x,v)\in\bar{G}\times V \tag{5.3.4}$$

$$u(x,v,T) = \psi(x,v), \quad (x,v) \in \bar{G} \times V \tag{5.3.5}$$

In order to avoid further accumulation will consider, instead of around the point u_0 neighborhood of zero. Namely, subtracting from the equation

$$u_t(x,v,t) + (v,\triangledown)u(x,v,t) + \Sigma(x,v,t)\,u_0(x,v,t) =$$
$$= \int_V J\,(x,v,t,v')\,u(x,v',t)\,dv' + F(x,v,t),$$

equation

$$\tilde{u}_t(x,v,t) + (v,\triangledown)\tilde{u}(x,v,t) + \Sigma(x,v,t)\,u_0(x,v,t) =$$
$$= \int_V J\,(x,v,t,v')\,\tilde{u}(x,v',t)\,dv' + \tilde{F}(x,v,t),$$

get

$$\hat{u}_t(x,v,t) + (v,\triangledown)\hat{u}(x,v,t) =$$
$$= \int_V J\,(x,v,t,v')\,\hat{u}(x,v',t)\,dv' + \hat{F}(x,v,t).$$

Let the source function is represented as a

$$F(x,v,t) = f(x,v)\,g(x,v,t), \tag{5.3.6}$$

where u and f — unknown function and g, ψ — apriori defined function, which can be viewed as a manageable task, namely, translate the system of $u(x,v,0) = 0$ in the final overdetermination condition $u(x,v,T) = \psi(x,v)$ through controllability of $f(x,v)$. In our work we prove that there exists a unique solution (u,f) this problem in the neighborhood of the point u_0 in the corresponding functional space (all the symbols and terms will be introduced below) for sufficiently small in the norm $\psi(x,v)$. Here and further $\nabla = \nabla_x$.

In our case, [37] describes the process of mass transfer multi anisotropic modified kinetic equation:

$$u_t(x,v,t) + (v,\nabla)\,u(x,v,t) + \sum(x,v,t)\,u_0(x,v,t) =$$
$$= \int\limits_V J(x,v,t,v')\,u(x,v',t)\,dv' + F(x,v,t), \tag{5.3.7}$$
$$(x,v,t) \in G \times V \times (0,T),$$

in which the function $u(x,v,t)$ describes the density distribution particles in the phase space $G\times V$ at time $t\in(0,T)$, a function $\sum(x,v,t), J(x,v,t,v')$ and $F(x,v,t)$ – environment in which this process proceeds, as with coefficient of absorption, scattering indicatrix, and as a source respectively. The function u_0 is fixed. The same equation cause the problem of radiation of charged particles, as well as spread of γ-radiation. As shown in the [32], if you asked all the characteristics of the environment Σ, J, F, and as " incoming flow", i.e

$$u(x,v,t)=\mu(x,v,t),\quad (x,v,t)\in\gamma_-\times[0,T], \tag{5.3.8}$$

and the initial state of the process, ie.

$$u(x,v,0)=\varphi(x,v),\quad (x,v)\in\bar{G}\times V, \tag{5.3.9}$$

the state of $u(x,v,t)$ can be defined unambiguously at any time t, and there is derived estimate accuracy

$$\begin{aligned}&\|u\|_{H_\infty(D)}\leqslant C\Big(\|F\|_{W^t_\infty(D)}+\|\varphi\|_{h_\infty(G\times V)}+\|\mu\|_{W^t_\infty(\Gamma_-)}+\\&+M(V)\|\sigma\|_{L_\infty(G\times V)}+M(V)\|u_0\|_{H_\infty(D)}\Big).\end{aligned} \tag{5.3.10}$$

where $M(V)$ - action sets V. Consider the inverse problem for a linear derivative of the transport equation using the sources. An interesting case where the controls are allowed stationary, ie do not change over time. Suppose that

$$F(x,v,t)=f(x,v)\,g(x,v,t)+h(x,v,t), \tag{5.3.11}$$

where f — desired, and g, h — priori defined functions. From a physical point of view, the task of controllability is as follows: Can you find such a function $f(x,v)$ and the corresponding function $u(x,v,t)$, satisfying the conditions (5.3.7)–(5.3.9) so that at time $t=T$ distribution function was equal to $\psi(x,v)$. In such a setting, this task is the task of managing non-transfer using the source of F. A generalized solution of class $H_\infty(D)\times L_\infty(G\times V)$ of the inverse problem (5.3.7) will understand a pair of functions (u,f): $u\in H_\infty(D)$ and $f\in L_\infty(G\times V)$, satisfying almost everywhere the conditions (5.3.7)–(5.3.9),(5.2.5).

Theorem 5.3.1. ***(Unique solvability of the inverse problem for linear modified transport equation in*** $H_\infty(D)\times L_\infty(G\times V)$***).***

Let $\sum, \sum_t \in L_\infty(D)$*;* $J, J_t \in L_\infty\big(D, L_2(V)\big)$*;* $u_0 \in H_\infty(D)$*,* $u_{0_t} \in L_\infty(D)$*,* $g, g_t \in L_\infty(G\times V, L_\infty(0,T))$*,* $|g(x,v,T)| \geq g_0 > 0$*;* $h, h_t \in L_\infty(D)$*;* $\mu, \mu_t \in L_\infty(\Gamma_-)$*;* $\varphi, (v,\nabla)\varphi, \psi, (v,\nabla)\psi \in L_\infty(G\times V)$*,* $\varphi|_{\gamma_-}, \psi|_{\gamma_-} \in L_\infty(\gamma_-)$*;* $v_0^{-1} diam G < a$*, and harmonization of the conditions* $\varphi(x,v) = \mu(x,v,0)$ *and* $\psi(x,v) = \mu(x,v,T)$ *for almost all* $(x,v) \in \gamma_-$*. Then the controllability of the problem has a unique solution* $\{u,f\} \in H_\infty(D)\times L_\infty(G\times V)$*.*

As noted above, to facilitate the presentation will consider, instead of around the point u_0 neighborhood of zero. The method of proof of this theorem is not fundamentally different from the evidence Theorems on unique solvability of the inverse problem for a linear transport equation N.P.Volkov [36].

5.3.1 Statement of the main result.

Since the operator

$$\hat{\hat{\mathcal{A}}} : L_\infty(G\times V) \longrightarrow h_\infty(G\times V),$$

$$f \mapsto \psi,$$

linear and bijective (because the solution is unique), the Banach theorem $\hat{\hat{\mathcal{A}}}$ — isomorphism, i.e

$$\hat{\hat{\mathcal{A}}}^{-1} : h_\infty(G\times V) \longrightarrow L_\infty(G\times V),$$

is continuous and, thus, $\forall\, \psi\ \exists\,!$ solution f linear inverse problem, the

$$\|f\|_{L_\infty(G\times V)} \leqslant \bar{C}\|\psi\|_{h_\infty(G\times V)}, \tag{5.3.12}$$

for some $\bar{C} > 0$ ($\bar{C}$ depends on $\sum, J, g, h, \mu, G, V$, as well as of the norms of differential and integral operators, introduced in the proof of the theorem).

Let $\mu(x,v,t) = 0, (x,v,t) \in \gamma_-\times[0,T]$; Let $\mu(x,v,t) = 0, (x,v,t) \in \gamma_-\times[0,T]$; $\varphi(x,v) = 0,\ (x,v) \in \bar{G}\times V;\ \ h(x,v,t) \equiv 0$ on D.

Let $Lu = u_t + (v,\nabla)\,u - \int\limits_V Ju\,dv' = -\sum u_0 + F$, $H = \{u \in H_\infty(D) | \exists F \in W^t_\infty(D)$: u — solution of the problem (5.3.7)–(5.3.9)$\}$ with the norm $\|u\|_H = \|Lu\|_{W^t_\infty(D)}$. Then $L : H \to W^t_\infty(D)$ — isometric isomorphism of H on $W^t_\infty(D)$ [32] (5.3.10), H — fully and continuously invested in the $H_\infty(D)$. From there, it should be that $\exists \hat{C} > 0,\ \forall\, t_1 \in$

$]0,T]$, $\forall u \in H : \|u|_{t=t_1}\|_{L_\infty(G\times V)} \leqslant \hat{C}\|u\|_H$. Note also that the theorem on the trace for $H_\infty(D)$ the operator

$$\begin{gathered}\Lambda_u :\ H \longrightarrow L_\infty(G\times V),\\ u \mapsto u_t|_{t=t_1},\end{gathered} \tag{5.3.13}$$

continuous because of the continuity of embedding H in $H_\infty(D)$. Let $\chi(x,v) = f(x,v)\, g(x,v,T)$.

Everywhere in the future will be taken prior designations. Let $\alpha : \mathbb{R} \to \mathbb{R}$ twice continuously differentiable on $\mathbb{R}$, with

$$\begin{cases} |\alpha(u)| \leqslant C_1|u|\,, \\ |\alpha'(u)| \leqslant C_1\,, \quad \alpha'(0) = 0\,, \\ |\alpha''(u)| \leqslant C_2\,. \end{cases} \tag{5.3.14}$$

Then, $Q(x,v,x',v',t)$ continuous on $\bar{G}\times\bar{V}\times\bar{G}\times\bar{V}\times[0,T]$.
Introduce the operator S: $H_\infty(D) \to W^t_\infty(D)$,
$[S(u)](x,v,t) = \int\limits_0^T \int\limits_{G\times V} Q(x,v,x',v',t)\,\alpha\big(u(x',v',t')\big)\,dx'dv'dt'$.

Theorem 5.3.2. *(**Unique solvability of the inverse problem for the nonlinear modified transport equation in** $H_\infty(D)\times L_\infty(G\times V)$).*

Let $\sum, \sum_t \in L_\infty(D)$, $|\sum(x,v,T)| \geq \sum_0 > 0$, $(x,v)\in G\times V$;$J, J_t \in\ L_\infty\big(D, L_2(V)\big)$; $u_0 \in H_\infty(D)$ $u_{0_t} \in L_\infty(D)$, $g, g_t \in L_\infty(G\times V, L_\infty(0,T))$, $|g(x,v,T)| \geq g_0 > 0$; $\psi, (v,\nabla)\psi \in L_\infty(G\times V)$, $\psi|_{\gamma_-} \in L_\infty(\gamma_-)$; $v_0^{-1} diamG < a$ and condition $\psi(x,v) = 0$ when $(x,v)\in\gamma_-$. Then the problem

$$\begin{gathered} u_t(x,v,t) + (v,\nabla)u(x,v,t) + \textstyle\sum(x,v,t)u_0(x,v,t) + [S(u)](x,v,t) = \\ \int\limits_V J(x,v',t,v)u(x,v',t)dv' + F(x,v,t), \quad \textit{where} \quad (x,v,t)\in D \end{gathered} \tag{5.3.15}$$

$$u(x,v,t) = 0, \quad (x,v,t)\in\gamma_-\times[0,T], \tag{5.3.16}$$

$$u(x,v,0) = 0, \quad (x,v)\in\bar{G}\times V \tag{5.3.17}$$

$$u(x,v,T) = \psi(x,v), \quad (x,v)\in\bar{G}\times V \tag{5.3.18}$$

$$F(x,v,t) = f(x,v)\, g(x,v,t), \tag{5.3.19}$$

where f — desired, and g — apriori given function, has unique solution (u, f) in the neighborhood of the point u_0 in $H_\infty(D) \times L_\infty(G \times V)$ for sufficiently small in the norm $\psi \in L_\infty(G \times V)$.

Remark 5.2. *Proof of Theorem (5.3.2) is reduced to checking the feasibility conditions of the theorem on the inverse functions. The main points are the verification of the strict differentiability of the corresponding mappings and the choice of function spaces associated with problem (5.1.2)-(5.1.5) .*

5.3.2 Proof.

According to a theorem about the inverse function (4.1.1) there exist open neighborhood U_1 of H and V_1 of $W^t_\infty(D)$ points of u_0 and $\xi(u_0)$, respectively, the operator $\xi : u \mapsto Lu + S(u)$ is implementing diffeomorphism U_1 on V_1 of class C^1. In doing so, the operator $\eta = \xi^{-1} : V_1 \to U_1$ is strictly differentiate on V_1 as a mapping on H, and hence in $H_\infty(D)$ and $\eta'(F) = \left[\xi'\bigl(\eta(F)\bigr)\right]^{-1}$ at $F \in V_1$. In this case, by virtue of linearity and continuity operator $\chi(x, v) \mapsto \chi(x, v)g(x, v, t)/g(x, v, T)$ of $L_\infty(G \times V)$ in $W^t_\infty(D)$ it follows that the operator $P : L_\infty(G \times V) \to H_\infty(D)$, $\chi(x, v) \mapsto P(\chi) = \eta\left(\chi(x, v)g(x, v, t)/g(x, v, T)\right)$ is strictly for Frechet differentiate in neighborhood of $V_2 \subset L_\infty(G \times V)$ the point $P(u_0)$, and then the operators

$$O_1 : V_2 \to L_\infty(G \times V),$$
$$\chi \mapsto \bigl(P(\chi)\bigr)_t\Big|_{t=T}, \tag{5.3.20}$$
$$O_2 : V_2 \to L_\infty(G \times V),$$
$$\chi \mapsto (v, \nabla)\bigl(P(\chi)\bigr)\Big|_{t=T}, \tag{5.3.21}$$
$$O_3 : V_2 \to L_\infty(G \times V),$$
$$\chi \mapsto \int\limits_V JP(\chi)\, dv'\Big|_{t=T}, \tag{5.3.22}$$
$$O_4 : V_2 \to L_\infty(G \times V),$$
$$\chi \mapsto S\bigl(P(\chi)\bigr)\Big|_{t=T}, \tag{5.3.23}$$

also strongly differentiable on V_2. Next, we will seek a solution nonlinear problem (5.1.2)-(5.1.5) in the form of $u = P(\chi)$, where u — solution of nonlinear problem (5.1.2)-(5.1.5) on the right part of $F(x,v,t) = \chi(x,v)g(x,v,t)/g(x,v,T)$. Introduce the operator

$$M(\chi) = \Big\{\chi - \big(P(\chi)\big)_t\Big|_{t=T} - (v,\nabla)\big(P(\chi)\big)\Big|_{t=T} - S\big(P(\chi)\big)\Big|_{t=T} + \\ + \Big(\int\limits_V JP(\chi)\,dv'\Big)\Big|_{t=T}\Big\} \times \left(\frac{\big(P(\chi)\big)\Big|_{t=T}}{\sum(x,v,T)u_0}\right), \quad (5.3.24)$$

$$M'(u_0)\chi = \Big\{\chi - \big(P'(u_0)\chi\big)_t\Big|_{t=T} - (v,\nabla)\big(P'(u_0)\chi\big)\Big|_{t=T} + \\ + \Big(\int\limits_V J\big(P'(u_0)\chi\big)\,dv'\Big)\Big|_{t=T}\Big\} \times \left(\frac{\big(P(\chi)\big)\Big|_{t=T}}{\sum(x,v,T)u_0}\right) + \\ + \Big\{\chi - \big(P(\chi)\big)_t\Big|_{t=T} - (v,\nabla)\big(P(\chi)\big)\Big|_{t=T} - S\big(P(\chi)\big)\Big|_{t=T} + \\ + \Big(\int\limits_V JP(\chi)\,dv'\Big)\Big|_{t=T}\Big\} \times \left(\frac{\big(P'(u_0)\chi\big)\Big|_{t=T}}{\sum(x,v,T)u_0}\right), \quad (5.3.25)$$

where $P'(u_0)\chi = \eta'(u_0)\,(\chi(x,v)g(x,v,t)/g(x,v,T))$, i.e direct solution of linear problems with the right part of $F(x,v,t) = \chi(x,v)g(x,v,t)/g(x,v,T)$ (solution exists, as shown in [32], and the operator $\eta'(u_0)$ is continuous, because is estimation accuracy). Substituting in (5.1.2) $u|_{t=T}$ of (5.1.5) and $F(x,v,t) = \chi(x,v)g(x,v,t)/g(x,v,T)$, we arrive to the equation

$$M(\chi) = \psi(x,v), \quad (5.3.26)$$

which by virtue of the continuity of $M'(u_0)$, as well as the existence and boundedness of $[M'(u_0)]^{-1}$ (from the solvability linear inverse problem) can be applied theorem (5.1.1). According to the theorem (5.1.1) there exist an open neighborhood U_* in $L_\infty(G \times V)$ and V_* in $h_\infty(G \times V)$ of points of u_0 and $M(u_0)$, respectively, that $M : \chi \to \psi$ is implementing diffeomorphism U_* on V_* class C^1. Therefore, $\forall\, \psi \in V_*\ \exists!\ \chi \in U_* : M(\chi) = \psi$. The theorem is proved.

A proof of the local unique solvability in $H_\infty(D)\times L_\infty(G\times V)$ inverse problem with final overdetermination for the modified nonlinear equations transport in the case where $F(x,v,t)=f(x,v)\,g(x,v,t)$ and controllability is $f(x,v)$.

5.4 *Production of the inverse problem for the modified nonlinear transfer equation in the case where $\Sigma(x,v,t)=\sigma(x,v)\,g(x,v,t)$ and controllability is $\sigma(x,v)$ in $H_\infty(D)\times L_\infty(G\times V)$.*

We prove local unique solvability in $H_\infty(D)\times L_\infty(G\times V)$ inverse problem with final overdetermination for the modified nonlinear equations transport in the case where $\Sigma(x,v,t)=\sigma(x,v)\,g(x,v,t)$ and controllability is $\sigma(x,v)$. In this paper the local unique solvability inverse problem with final overdetermination modified equation for nonlinear transport. These tasks can be viewed as the problem of controllability, which management is, for example, the multiplier on the right side, or in any other factors, depending only on spatial variables in the case of the final overdetermination and independent only occasionally in the case of integral overdetermination. We consider the direct and inverse problems for integro-differential equation, a modified equation:

$$u_t(x,v,t)+(v,\nabla)u(x,v,t)+\Sigma(x,v,t)\,u_0(x,v,t)=$$

$$=\int_V J\,(x,v,t,v')\,u(x,v',t)\,dv'+F(x,v,t)$$

(in the usual transport equation, instead of u_0 is u [36]). This equation more convenient for the study of some of its non-linear perturbations. Absorption coefficient is presented in the form

$$\Sigma\,(x,v,t)=\sigma\,(x,v)\,g\,(x,v,t)+\,h\,(x,v,t),$$

Multiplier $\sigma(x,v)$ plays the role of government. Inverse problem is a pair of $(u,\sigma)\in H_\infty(D)\times L_\infty(G\times V)$ (actually, only σ), and the meaning of the problem lies in the translation through σ of the system initial state $\varphi(x,v)$ in the final state $\psi(x,v)$. In order to avoid further accumulation will consider, instead of around the point u_0 neighborhood of zero. Namely, subtracting from the equation

$$u_t(x,v,t)+(v,\nabla)u(x,v,t)+\Sigma(x,v,t)\,u_0(x,v,t)=$$

$$= \int_V J\,(x,v,t,v')\,u(x,v',t)\,dv' + F(x,v,t),$$

equation

$$\tilde{u}_t(x,v,t) + (v,\triangledown)\tilde{u}(x,v,t) + \tilde{\Sigma}(x,v,t)\,u_0(x,v,t) =$$
$$= \int_V J\,(x,v,t,v')\,\tilde{u}(x,v',t)\,dv' + F(x,v,t),$$

get

$$\hat{u}_t(x,v,t) + (v,\triangledown)\hat{u}(x,v,t) + \hat{\Sigma}(x,v,t)\,u_0(x,v,t) =$$
$$= \int_V J\,(x,v,t,v')\,\hat{u}(x,v',t)\,dv'.$$

5.4.1 Proof.

Everywhere in the future will be taken prior designations. Let $\alpha : \mathbb{R} \to \mathbb{R}$ twice continuously differentiable on $\mathbb{R}$, with

$$\begin{cases} |\alpha(u)| \leqslant C_1 |u|\,, \\ |\alpha'(u)| \leqslant C_1\,, \quad \alpha'(0) = 0\,, \\ |\alpha''(u)| \leqslant C_2\,. \end{cases} \tag{5.4.1}$$

Then, $Q(x,v,x',v',t)$ continuous on $\bar{G} \times \bar{V} \times \bar{G} \times \bar{V} \times [0,T]$. Introduce the operator S: $H_\infty(D) \to W^t_\infty(D)$, $[S(u)](x,v,t) = \int\limits_0^T \int\limits_{G\times V} Q(x,v,x',v',t)\,\alpha\big(u(x',v',t')\big)\,dx'dv'dt'$.

Theorem 5.4.1. ***(Unique solvability of the inverse problem for the nonlinear modified transport equation in*** $H_\infty(D) \times L_\infty(G\times V)$***).***

Let $F, F_t \in L_\infty(D)$, $|\sum(x,v,T)| \geq \sum_0 > 0$, $(x,v) \in G \times V$;$J, J_t \in L_\infty\big(D, L_2(V)\big)$; $u_0 \in H_\infty(D)$ $u_{0_t} \in L_\infty(D)$, $g, g_t \in L_\infty\big(G \times V, L_2(0,T)\big)$, $|g(x,v,T)| \geq g_0 > 0$; $\psi, (v,\nabla)\psi \in L_\infty(G\times V)$, $\psi|_{\gamma_-} \in L_\infty(\gamma_-)$; $v_0^{-1} diam G < b$ *and condition* $\psi(x,v) = 0$ *when* $(x,v) \in \gamma_-$. *Then the problem*

$$u_t(x,v,t) + (v,\nabla)u(x,v,t) + \sum(x,v,t)u_0(x,v,t) + [S(u)](x,v,t) = \int\limits_V J(x,v',t,v)u(x,v',t)dv', \quad \textit{where} \quad (x,v,t) \in D \tag{5.4.2}$$

$$u(x,v,t) = 0, \quad (x,v,t) \in \gamma_- \times [0,T], \tag{5.4.3}$$

$$u(x,v,0) = 0, \quad (x,v) \in \bar{G} \times V \tag{5.4.4}$$

$$u(x,v,T) = \psi(x,v), \quad (x,v) \in \bar{G} \times V \tag{5.4.5}$$

$$\Sigma(x,v,t) = \sigma(x,v)\, g(x,v,t), \tag{5.4.6}$$

where σ — desired, and g — apriori given function, has unique solution (u,σ) in a neighborhood of zero in $H_2(D) \times L_\infty(G \times V)$ for sufficiently small in the norm $\psi \in L_\infty(G \times V)$.

Remark. Proof of Theorem (5.4.1) boils down to checking the feasibility conditions of the theorem on the inverse functions.

Proof.

According to a theorem about the inverse function (5.1.1) there exist open zero neighborhood U_1 of H and V_1 of $W^t_\infty(D)$ respectively, the operator $\xi : u \mapsto Lu + S(u)$ is implementing diffeomorphism U_1 on V_1 of class C^1. In doing so, the operator $\eta = \xi^{-1} : V_1 \to U_1$ is strictly differentiate on V_1 as a mapping on H, and hence in $H_\infty(D)$ and $\eta'(\Sigma) = \left[\xi'\big(\eta(\Sigma)\big)\right]^{-1}$ at $\Sigma \in V_1$. In this case, by virtue of linearity and continuity operator $\chi(x,v) \mapsto \chi(x,v)g(x,v,t)/g(x,v,T)$ of $L_\infty(G \times V)$ in $W^t_\infty(D)$ it follows that the operator $P : L_\infty(G \times V) \to H_\infty(D)$, $\chi(x,v) \mapsto P(\chi) = \eta\left(\chi(x,v)g(x,v,t)/g(x,v,T)\right)$ is strictly for Frechet differentiate in around zero $V_2 \subset L_\infty(G \times V)$ the point $P(0)$, and then the operators

$$O_1 : V_2 \to L_\infty(G \times V),$$
$$\chi \mapsto \big(P(\chi)\big)_t\Big|_{t=T}, \tag{5.4.7}$$

$$O_2 : V_2 \to L_\infty(G \times V),$$
$$\chi \mapsto (v,\nabla)\big(P(\chi)\big)\Big|_{t=T}, \tag{5.4.8}$$

$$O_3 : V_2 \to L_\infty(G \times V),$$
$$\chi \mapsto \int\limits_V JP(\chi)\, dv'\Big|_{t=T}, \tag{5.4.9}$$

$$O_4 : V_2 \to L_\infty(G \times V),$$
$$\chi \mapsto S\big(P(\chi)\big)\Big|_{t=T}, \tag{5.4.10}$$

also strongly differentiable on V_2. Next, we will seek a solution nonlinear problem (5.1.2)–(5.1.5) in the form of $u = P(\chi)$, where u — solution of nonlinear problem (5.1.2)–(5.1.5) with a coefficient of absorption (weakening) $\Sigma(x,v,t) = \chi(x,v)g(x,v,t)/g(x,v,T)$. Introduce the operator

$$M(\chi) = \Big\{ -\big(P(\chi)\big)_t\Big|_{t=T} - (v,\nabla)\big(P(\chi)\big)\Big|_{t=T} - S\big(P(\chi)\big)\Big|_{t=T} + \\ + \Big(\int\limits_V JP(\chi)\,dv'\Big)\Big|_{t=T}\Big\} \times \left(\frac{(P(\chi))\Big|_{t=T}}{\chi_0(1+\dfrac{\delta(\chi)}{\chi_0}+...)u_0} \right), \quad (5.4.11)$$

$$M'(0)\chi = \Big\{ -\big(P'(0)\chi\big)_t\Big|_{t=T} - (v,\nabla)\big(P'(0)\chi\big)\Big|_{t=T} + \\ + \Big(\int\limits_V J\big(P'(0)\chi\big)\,dv'\Big)\Big|_{t=T}\Big\} \times \left(\frac{(P(\chi))\Big|_{t=T}}{\chi_0(1+\dfrac{\delta(\chi)}{\chi_0}+...)u_0} \right) + \\ + \Big\{ -\big(P(\chi)\big)_t\Big|_{t=T} - (v,\nabla)\big(P(\chi)\big)\Big|_{t=T} - S\big(P(\chi)\big)\Big|_{t=T} + \\ + \Big(\int\limits_V JP(\chi)\,dv'\Big)\Big|_{t=T}\Big\} \times \left(\frac{\big(P'(0)\chi\big)\Big|_{t=T}}{\chi_0(1+\dfrac{\delta(\chi)}{\chi_0}+...)u_0} \right) + \\ + \Big\{ -\big(P(\chi)\big)_t\Big|_{t=T} - (v,\nabla)\big(P(\chi)\big)\Big|_{t=T} - S\big(P(\chi)\big)\Big|_{t=T} + \\ + \Big(\int\limits_V JP(\chi)\,dv'\Big)\Big|_{t=T}\Big\} \times \left(-\frac{(P(\chi))\Big|_{t=T}}{(\chi_0(1+\dfrac{\delta(\chi)}{\chi_0}+...))^2 u_0}\frac{\delta(\chi)}{\chi_0} \right), \quad (5.4.12)$$

where $P'(0)\chi = \eta'(0)\,(\chi(x,v)g(x,v,t)/g(x,v,T))$, ie direct solution of linear problems with a right-hand side $\Sigma(x,v,t) = \chi(x,v)g(x,v,t)/g(x,v,T)$ (solution exists, as shown in [32]). Substituting in (5.1.2) $u|_{t=T}$ of (5.1.5) and $\Sigma(x,v,t) = \chi(x,v)g(x,v,t)/g(x,v,T)$, we arrive to the equation

$$M(\chi) = \psi(x,v), \quad (5.4.13)$$

which by virtue of the continuity of $M'(0)$, as well as the existence and boundedness of $[M'(0)]^{-1}$, can be applied theorem, *3*. According to the theorem *3* there exist an open neighborhood of zero U_* in $L_\infty(G\times V)$ and V_* in $h_\infty(G\times V)$ respectively, that $M:\chi\to\psi$ is implementing diffeomorphism U_* on V_* class C^1. Therefore, $\forall\,\psi\in V_*\;\exists!\;\chi\in U_*:M(\chi)=\psi$. The theorem is proved.

The author expresses his gratitude to the participants of the seminar on mathematical physics of Keldysh Institute of Applied Mathematics (Russian Academy of Sciences) for your interest in this topic. He also thanked his scientific supervisor of PFUR (Peoples' Friendship University of Russia, Moscow) M.F. Sukhinin for useful discussions.

5.4.2 Main results of the chapter

1. We proved the local unique solvability of the inverse problem with a final redefinition for the nonlinear modified transport equation in the case when $F(x,v,t)=f(x,v)g(x,v,t)$ and the control is $f(x,v)$ in $H_2(D)\times L_2(G\times V)$.

2. We proved the local unique solvability of the inverse problem for the nonlinear modified transport equation with final redefinition in the case when $F(x,v,t)=f(x,v)g(x,v,t)$ and the control is $f(x,v)$ in the class bounded functions $f(x,v)$ in $H_\infty(D)\times L_\infty(G\times V)$.

3. We proved the local unique solvability of the inverse problem for the nonlinear modified transport equation with final redefinition in the case when $\Sigma(x,v,t)=\sigma(x,v)g(x,v,t)$ and the control is $\sigma(x,v)$ in the class bounded functions $\sigma(x,v)$ in $H_\infty(D)\times L_\infty(G\times V)$.

The obtained results are published in [11] and [12].

BIBLIOGRAPHY

[1] *https://money.cnn.com/data/markets/russell.*

[2] https://bloomberg.com.

[3] M.V. Klibanov A.B. Bakushinskii and N.A. Koshev. Carleman weight functions for a globally convergent numerical method for ill-posed Cauchy problems for some quasi-linear PDEs. *Nonlinear Analysis: Real World Applications*, 34:201–224, 2017.

[4] G. I. Bell and S. Glasstone. Nuclear reactor theory. 1970.

[5] T. Björk. Arbitrage theory in continuous time. *Oxford University Press*, 1999.

[6] F. Black and M. Scholes. The pricing of options and corporate liabilities. *J. of Political Economy*, (81):637–654, 1973.

[7] L. Bourgeois. Convergence rates for the quasi-reversibility method to solve the cauchy problem for laplace's equation. *Inverse Problems*, (22):413–430, 2006.

[8] L. Bourgeois and J. Darde. *A duality-based method of quasi-reversibility to solve the Cauchy problem in the presence of noisy data*, volume 26. 2010.

[9] J. C. Cox, S. A. Ross, and M. Rubinstein. Option pricing: a simplified approach. *J. of Political Economy*, 7:229–263, 1979.

[10] E. F. Fama and K. R. French. The capital asset pricing model: theory and evidence. *J. of Economic Perspective*, 18(3):25–46, 2004.

[11] K. V. Golubnichiy. The controllability problems for modified transfer equation. // international conference "differential equation and related topics",dedicated to the memory of i.g. petrovskii. *Book of Abstracts, Moscow, MSU*, page 106, 2007.

[12] K. V. Golubnichiy. On one the controllability problems for modified transfer equation. *Proceedings of the Voronez Winter School of Mathematics under S.G. Crain. VSU, Voronez*, pages 87–95, 2008.

[13] E. Hopf. Mathematical problems of radioactiv equilibrium. *Cambridge Tracts*, page 103, 1934.

[14] R. A. Irizarry. Introduction to data science: Data analysis and prediction algorithms with r. *Chapman & Hall/CRC Data Science Series, Taylor & Francis Group*, 2020.

[15] M. V. Klibanov. Carleman estimates for the regularization of ill-posed cauchy problems. *Appl. Numer. Math.*, 94:46–74, 2015.

[16] M. V. Klibanov, K. V. Golubnichiy, and A. V Nikitin. Application of neural network machine learning to solution of black-scholes equation. *arXiv: 2111.06642*, 2021.

[17] M. V. Klibanov, A. V. Kuzhuget, and K. V. Golubnichiy. An ill-posed problem for the black-scholes equation for a profitable forecast of prices of stock options on real market data. *Inverse Problems*, 32(015010), 2016.

[18] M. V. Klibanov and J. Li. Inverse problems and carleman estimates: Global uniqueness, global convergence and experimental data. 2021.

[19] M. V. Klibanov, A. A. Shananin, K. V. Golubnichiy, and S. M. Kravchenko. Forecasting stock options prices via the solution of an ill-posed problem for the black-scholes equation. *arXiv preprint arXiv:2202.07174*, 2022.

[20] M. V. Klibanov and A. G. Yagola. Convergent numerical methods for parabolic equations with reversed time via a new carleman estimate. *Inverse Problems*, 35(115012), 2019.

[21] L. Koralov and Y. G. Sinai. Theory of probability and random processes. 2007.

[22] O. A. Ladyzhenskaya, V. A Solonnikov, and N. N. Uralceva. Linear and quasilinear equations of parabolic type. 1968.

[23] R. Lattes and J. L. Lions. The method of quasireversibility: Applications to partial differential equations. 1969.

[24] M. M. Lavrentiev, V. G Romanov, and S. P. Shishatskii. Ill-posed problems of mathematical physics and analysis. 1986.

[25] G. Marchuk and V. Lebedev. Numerical methods in the theory of neutron transport. 1986.

[26] G. I. Marchuk. *Dokladi AN USSR*, 1964.

[27] R. Merton. Theory of rational option pricing. *Bell J. of Economics and Management Science*, ,(4):141–183, 1973.

[28] B. Øksendal. Stochastic differential equations. *Calculus of Variations and Partial Differential Equations*, 2000.

[29] A. I. Prilepko. Inverse problems of potential theory (elliptic and hyperbolic equations and transport equations). *Math notes*, 14(5):755–767, 1973.

[30] A. I. Prilepko and A. A. Ivankov. Inverse problems of determining the coefficient cient, scattering indicatrix and the right side of the non-stationary multiplier state-velocity transport equation. *Differential Equations*, 21(5):870–885, 1985.

[31] A. I. Prilepko and A. A. Ivankov. Inverse problems of determining the coefficient on the right side of the non-stationary multirate equation transfer by override at point. *Differential Equations*, 21(1):109–119, 1985.

[32] A. I. Prilepko and Volkov N. P. Inverse problems determine the parameters of the unsteady transport equation for redefinition of the integral type. *Differential. Equations*, 23, 1987.

[33] A. N Shiryaev. Essentials of stochastic finance: Facts, models, theory. 1999.

[34] S. E. Shreve. Stochastic calculus for finance ii. continuous - time models. 2003.

[35] A. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov, and A. G. Yagola. Numerical methods for the solution of ill-posed problems. 1995.

[36] N. P. Volkov. Inverse problems for non-kinetic transport equation. thesis phd. *Thesis PhD. Faculty of Computational Mathematics and Cybernetics of Moscow State University*, page 111, 1987.

[37] N. P. Volkov. Sufficient conditions of solvability some inverse problems (tasks) for mass. In *Inverse problems for mathematical models physical processes. Moscow Engineering Physics Institute (State University)*, page 16, 1991.

[38] P. Wilmott, S. Howison, and J. Dewyne. The mathematics of financial derivatives. *Communications on Pure and Applied Mathematics.*