Communications in Advanced Mathematical Sciences

VOLUME II ISSUE II

CAMS

ISSN 2651-4001

VOLUME II ISSUE II
ISSN 2651-4001

June 2019
http://dergipark.gov.tr/cams

COMMUNICATIONS IN ADVANCED MATHEMATICAL SCIENCES

Contents

Communications in Advanced Mathematical Sciences
Vol. II, No. 2, 75-104, 2019
Research Article
e-ISSN:2651-4001
DOI: 10.33434/cams.478999

KP-KdV Hierarchy and Pseudo-Differential Operators

Ahmed Lesfari[1] *

Abstract
The study of KP-KdV equations are the archetype of integrable systems and are one of the most fundamental equations of soliton phenomena and a topic of active mathematical research. Our purpose here is to give a motivated and a sketchy overview of this interesting subject. One of the objectives of this paper is to study the KdV equation and the inverse scattering method (based on Schrödinger and Gelfand-Levitan equations) used to solve it exactly. We study some generalities on the algebra of infinite order differential operators. The algebras of Virasoro and Heisenberg, nonlinear evolution equations such as the KdV, Boussinesq and KP play a crucial role in this study. We make a careful study of some connection between pseudo-differential operators, symplectic structures, KP hierarchy and tau functions based on the Sato-Date-Jimbo-Miwa-Kashiwara theory. A few other connections and ideas concerning the KdV and Boussinesq equations, the Gelfand-Dickey flows, the Heisenberg and Virasoro algebras are given.

Keywords: Gelfand-Levitan integral equation, Integrable systems, KdV equation, KP hierarchy, Schrödinger equation, Symplectic structures.
2010 AMS: Primary 70H06, Secondary 14H55, 14H70, 14K20

[1] *Department of Mathematics, Faculty of Sciences, University of Chouaïb Doukkali, B.P. 20, 24000 El Jadida, Morocco, ORCID: orcid.org/0000-0001-6213-4301*
*__Corresponding author__: lesfariahmed@yahoo.fr
Received: 5 November 2018, **Accepted:** 8 March 2019, **Available online:** 27 June 2019

1. Introduction

Korteweg and de Vries have established a nonlinear partial differential equation describing the gravitational wave propagating in a shallow channel [1] and possessing remarkable mathematical properties :

$$\frac{\partial u}{\partial t}-6u\frac{\partial u}{\partial x}+\frac{\partial^3 u}{\partial x^3}=0, \tag{1.1}$$

where $u(x,t)$ is the amplitude of the wave at the point x and the time t. The equation thus bearing their name (abbreviated KdV) admits a solution: the soliton or solitary wave. In fact, this model was obtained from Euler's equations (assuming irrotational flow) by Boussinesq around 1877 (see [2], p. 360) and rediscovered by Korteweg and de Vries in 1890. The solution to this equation was obtained and interpreted rigorously only in the early 1970s while a solitary wave was already observed in 1834 by engineer Scott Russell riding on the Edinburgh Glasgow Canal in Scotland; he described his observation of a hydrodynamic phenomenon as follows : " I was observing the motion of a boat which was rapidly drawn along a narrow channel by a pair of horses, when the boat suddenly stopped - not so the mass of watering the channel which it had put in motion; it accumulated round the prow of the vessel in a state of violent agitation, then suddenly leaving it behind, rolled forward with great velocity, assuming the form of a large solitary elevation, a rounded, smooth and well-defined heap of water, which continued its course along the channel apparently without change of form or diminution of speed. I followed it on horseback, and overtook it still rolling on at a rate of some eight or nine miles an hour, preserving its original figure some thirty feet long and a foot to a foot and a half in height. Its height gradually diminished, and after a chase of one or two miles I lost it in the windings of the channel. Such, in the month of August 1834, was my first chance interview with that singular and beautiful phenomenon which I have called the wave of translation ". Fascinated by this phenomenon, Scott-Russell built a wave pool in his garden and worked to generate and study these waves more carefully. This led to a paper [3] dubbed "The report on waves" published in 1844 by the British Association for the Advancement of Science.

A little later, Boussinesq, then Korteweg and de Vries proposed equation (1.1) to explain this phenomenon. The KdV equation preserves mass, momentum, energy, and many other quantities. Many experiments have uncovered the astonishing properties of the solutions of this equation satisfying zero boundary conditions : when $|t| \longrightarrow \infty$, these solutions are decomposed into solitons, i.e., in waves of defined forms progressing at different speeds. These waves propagate over long distances without deformation and one of the remarkable characteristics of solitons is that they are exceptionally stable with respect to disturbances; the term $u\frac{\partial u}{\partial x}$ leads to shock waves while the term $\frac{\partial^3 u}{\partial x^3}$ produces a scattering effect. Everyone can contemplate solitons where the tide comes to die on the beaches. In the field of hydrodynamics for example, tsunamis (tidal waves) are manifestations of solitons. Generally, we group together under the term soliton solutions of nonlinear wave equations presenting the following characteristic properties : they are localized in space, last indefinitely and retain their amplitude and velocity even at the end of several collisions with other solitons. Solitons have become indispensable for the study of several phenomena. In particular, the study of wave propagation in hydrodynamics, the study of localized waves in astrophysical plasmas, They are involved in the study of signals in optical fibers, charge transport phenomena in conductive polymers, localized modes in magnetic crystals, etc. Industrialized societies have developed, after soliton studies, what may be called solitary lasers. The latter play an important role in the field of telecommunications. Ultra-short light signals sent in certain optical fibers made from a specific material can travel long distances without lengthening or fading. The construction of memories with ultra-fast communication time and low energy consumption, is based on the movement of magnetic vortices in the dielectric junction between two superconductors. At the molecular level, the theory of solitons can elucidate the contraction mechanism of striated muscles, the dynamics of biological macromolecules such as DNA and proteins. In the peptide and hydrogen chain of proteins, solitons arise from the marriage of dispersion due to intrapeptide vibrations and the non-linearity due to the interaction of these vibrations with the displacements of peptide groups around their position balanced. But also the theory of solitons had an impact on pure mathematics; for example, it provides the answer to the famous Schottky problem, posited a century ago, on the relations between the periods coming from a Riemann surface. Roughly, it is a question of finding criteria so that a matrix of the periods belonging to the Siegel half-space is the matrix of the periods of a Riemann surface. Geometrically, Schottky's problem consists in characterizing the Jacobians among all the Abelian mainly polarized varieties. In addition to the KdV equation, examples that may be mentioned among the nonlinear equations having soliton-type solutions are: the non-linear equation of Kadomtsev-Petviashvili, the nonlinear Schrödinger equation, the Sine Gordon equation, the Boussinesq equation, the Camassa-Holm equation, the Toda lattice consisting of vibrating masses arranged on a circle and interconnected by springs whose return force is exponential, the non-linear Klein Gordon equation, the Zabusky-Kruskal equation for the Fermi-Pasta-Ulam model of phonons in anharmonic lattice, and so on.

2. Stationary Schrödinger equation and integral Gelfand-Levitan equation

Since the method (discussed later) of solving the KdV equation is based on the idea of studying it in the form of an equation of a certain operator and using the analogy with quantum mechanic, we will expose certain mathematical notions of this mechanic. The terminology of the physicists will be used to describe the properties of the solutions of the stationary Schrödinger's equation,

$$\frac{\hbar}{2m}\psi'' + (\lambda - u(x))\psi = 0, \qquad ' \equiv \frac{d}{dx},$$

without stopping on the physical motivations of the introduced notions. We will see that the method of the inverse diffusion is reduced to the solution of a linear integral equation (Gelfand-Levitan equation). In the following, we will simplify the notation by using a system of units in which the Planck constant is $\hbar = 1$ and the mass of the particle is $m = \frac{1}{2}$. So consider the equation

$$\psi'' + (\lambda - u(x))\psi = 0, \qquad -\infty < x < \infty \tag{2.1}$$

where ψ (unknown) is the wave function of the particle, the spectral parameter λ is the energy of the particle, the function $u(x)$ is the potential or potential energy of the particle. This potential is assumed to have a compact support, i.e., is different from zero only in some domain. When the particle is free (i.e., $u = 0$) and has a positive energy (i.e., $\lambda = k^2$), then equation (2.1) is reduced to

$$\psi''(x)+k^2\psi=0, \tag{2.2}$$

and admits two linearly independent solutions e^{ikx} (describing the particle moving to the right) and e^{-ikx} (describing the particle moving to the left).

Let us denote by $E^{sr}_{(2)}$ (resp. $E^{sc}_{(2)}$) the space of real (or complex) solutions of equation (2.1) and by $E^{sr}_{(3)}$ (resp. $E^{sc}_{(3)}$) the space of the real (or complex) solutions of equation (2.2). The space $E^{sr}_{(3)}$ (resp. $E^{sc}_{(3)}$) is provided with the base $(\cos kx, \sin kx)$ (resp. (e^{ikx}, e^{-ikx})). Let $[\alpha,\beta]$ be the bounded support of u. The monodromy operator of equation (2.1) is a linear operator defined by

$$\mathscr{M}: E^{sr}_{(2)} \longrightarrow E^{sr}_{(2)}, \quad a\cos kx+b\sin kx \longmapsto \begin{cases} a\cos kx+b\sin kx & \text{si } x<\alpha \\ c\cos kx+d\sin kx & \text{si } x>\beta, \end{cases}$$

where a, b are constants and $(c,d)=\mathscr{M}_u(a,b)$. This means that for each solution of equation (2.2) is associated : (i) the solution of (2.1) which is to the left of α; in this region the solution of (2.2) coincides with that of (2.1). (ii) the solution of (2.1) which is to the right of β. Similarly, the complex monodromy operator of equation (2.1) is defined by

$$\mathscr{M}: E^{sc}_{(2)} \longrightarrow E^{sc}_{(2)}, \quad ae^{ikx}+be^{-ikx} \longmapsto \begin{cases} ae^{ikx}+be^{-ikx} & \text{si } x<\alpha \\ ce^{ikx}+de^{-ikx} & \text{si } x>\beta. \end{cases}$$

Recall that a particle propagating from $x=-\infty$, crosses a potential barrier with a transmission coefficient T and a reflection coefficient R if the equation (2.1) where $\lambda=k^2$ admits a solution ψ such that :

$$\psi=\begin{cases} Te^{ikx}, & \text{to the right of the barrier,} \\ e^{ikx}+Re^{-ikx}, & \text{to the left of the barrier.} \end{cases}$$

Theorem 2.1. *a) Let W be the phase plane formed by the representative points (ψ,ψ'). Let*

$$\mathscr{B}^{x_1}_{(2)}: E^{sr}_{(2)} \longrightarrow W, \quad \psi \longmapsto \mathscr{B}^{x_1}_{(2)}\psi=(\psi(x_1),\psi'(x_1)),$$

be an operator with ψ a solution of equation (2.1) whose initial conditions for $x=x_1\in\mathbb{R}$ are $(\psi(x_1),\psi'(x_1))$. Then the space $E^{sr}_{(2)}$ is isomorphic to W and the phase application of x_1 to x_2 defined by

$$g^{x_2}_{x_1} \equiv \mathscr{B}^{x_2}_{(2)}\left(\mathscr{B}^{x_1}_{(2)}\right)^{-1}: W \longrightarrow W, \quad (\psi(x_1),\psi'(x_1)) \longmapsto (\psi(x_2),\psi'(x_2)),$$

is a linear isomorphism.

b) If equation (2.1) where $\lambda=k^2$, has a confounded solution with ae^{ikx} for $x\ll 0$ and with be^{-ikx} for $x\gg 0$, then this solution is null. In addition, for all $k>0$ the ψ, T and R defined above exist and are unique.

Proof. a) $\mathscr{B}^{x_1}_{(2)}$ is linear and for any representative point $(\psi,\psi')\in W$, there exists from the existence theorem (differential equations) a solution ψ satisfying the initial condition $(\psi(x_1),\psi'(x_1))$. Then Im $\mathscr{B}^{x_1}_{(2)} \equiv \{\mathscr{B}^{x_1}_{(2)}\psi : \psi\in E^{sr}_{(2)}\}=W$. Finally Ker $\mathscr{B}^{x_1}_{(2)} \equiv \{\psi : \psi\in E^{sr}_{(2)}, \mathscr{B}^{x_1}_{(2)}\psi=0\}=0$, follows from the uniqueness theorem because the solution satisfying the initial condition at the point x_1 is equal to zero. The result follows from the fact that the inverse of an isomorphism is one. If ψ_1 and ψ_2 are two solutions of equation (2.1), then $(\psi(x_1),\psi'(x_1))=\mathscr{B}^{x_1}_{(2)}\psi_1+\mathscr{B}^{x_1}_{(2)}\psi_2=(\psi_1(x_1),\psi_1'(x_1))+(\psi_2(x_1),\psi_2'(x_1))$, and this is equivalent to

$$\left(\mathscr{B}^{x_1}_{(2)}\right)^{-1}((\psi_1(x_1),\psi_1'(x_1))+(\psi_2(x_1),\psi_2'(x_1)))=\left(\mathscr{B}^{x_1}_{(2)}\right)^{-1}(\psi_1(x_1),\psi_1'(x_1))+\left(\mathscr{B}^{x_1}_{(2)}\right)^{-1}(\psi_2(x_1),\psi_2'(x_1)).$$

b) Let be $\langle ae^{ikx}, ae^{ikx}\rangle$, $\langle be^{-ikx}, be^{-ikx}\rangle$ and $\langle ae^{ikx}, ae^{-ikx}\rangle$ the hermitian forms in the space $E^{sc}_{(2)}$. Let's designate by $[.,.]$ the left scalar product, then

$$\langle ae^{ikx}, ae^{ikx}\rangle=\frac{i}{2}[ae^{ikx},\bar{a}e^{-ikx}]=\frac{i}{2}\begin{vmatrix} a & ia \\ \bar{a} & -i\bar{a} \end{vmatrix}=|a|^2.$$

Similarly, we have $\langle be^{-ikx}, be^{-ikx}\rangle=-|b|^2$ et $\langle ae^{ikx}, ae^{-ikx}\rangle=0$. By setting $z=z_1e^{ikx}+z_2e^{-ikx}$ where z_1 and z_2 are the coordinates of the vector z in the basis (e^{ikx},e^{-ikx}), we obtain $\langle z,z\rangle=|z_1|^2-|z_2|^2$, i.e., $\langle .,.\rangle$ is of type $(1,1)$. Since the monodromy operator retains this hermitian form, we deduce that $|a|^2 = -|b|^2$ and so $a = b = 0$. Consider now a particle going to $+\infty$ and let e^{ikx} be a solution to the right of the barrier. To the left of the barrier this solution becomes

$$e^{ikx} \curvearrowright ae^{ikx} + be^{-ikx}. \tag{2.3}$$

From what precedes, the coefficient a is nonzero. So to have the solution in question, simply divide the two members of (2.3) by a, $\frac{1}{a}e^{ikx} \curvearrowright e^{ikx} + \frac{b}{a}e^{-ikx}$. Taking $T = \frac{1}{a}$ and $R = \frac{b}{a}$, this shows that T and R are uniquely defined. □

In the same way, we can define an operator $\mathscr{B}^{x_1}_{(3)}$ of $E^{sr}_{(3)}$ in W that associates with each solution of equation (2.2), its initial condition at the point x_1. In this case, instead of "phase application", there will be "phase point".

We will now demonstrate a theorem that will be useful later.

Theorem 2.2. *(Liouville). Let $\frac{dx}{dt} = f(x)$, $x = (x_1, ..., x_n)$, be a system of ordinary differential equations whose solutions extend to the whole time axis. Let $\{g^t\}$ be the corresponding group of transformations : $g^t x = x + f(x)t + o(t^2)$, for t small. We denote by D a domain in phase space, $D(t) \equiv g^t D(0)$ and by $v(t)$ the volume of $D(t)$. If $\operatorname{div} f = \sum_{j=1}^{n} \frac{\partial f_j}{\partial x_j} = 0$, then $v(t) = v(0)$, i.e., g^t preserves the volume of any domain.*

Proof. We have $v(t) = \int_{D(t)} dx = \int_{D(0)} \frac{\partial g^t x}{\partial x} dx$, where $\frac{\partial g^t x}{\partial x}$ is the Jacobian matrix, $\frac{\partial g^t x}{\partial x} = I + \frac{\partial f}{\partial x} t + o(t^2)$. The determinant of the operator $I + \frac{\partial f}{\partial x} t$ is equal to the product of the eigenvalues. These (taking into account their multiplicities) are equal to $1 + t\frac{\partial f_j}{\partial x_j}$ where $\frac{\partial f_j}{\partial x_j}$ are the eigenvalues of $\frac{\partial f}{\partial x}$. Then

$$\det \frac{\partial g^t x}{\partial x} = 1 + t\sum_{j=1}^{n} \frac{\partial f_j}{\partial x_j} + o(t^2) = 1 + t \operatorname{div} f + o(t^2).$$

Therefore, $v(t) = \int_{D(0)} (1 + t \operatorname{div} f + o(t^2))dx$, and $\left.\frac{dv(t)}{dt}\right|_{t=0} = \int_{D(0)} \operatorname{div} f dx$. Since $t = t_0$ is not worse than $t = 0$, we also have

$$\left.\frac{dv(t)}{dt}\right|_{t=t_0} = \int_{D(t_0)} \operatorname{div} f dx,$$

and the proof of the theorem follows. □

Note that the Liouville's theorem is easily generalized to the case of non autonomous systems $(f = f(x,t))$. Indeed, the terms of first degree in the expression of $\frac{\partial g^t x}{\partial x}$ remain the same. But the terms of degree greater than one do not intervene in the proof. In other words, Liouville's theorem is a first order theorem.

Let $SL(2,\mathbb{R})$ be the real unimodular group, i.e., the set of all real 2×2 matrices with determinant one. In other words, $SL(2,\mathbb{R})$ is the group of all linear transformations of $\mathbb{R}^2$ that preserve oriented area $[.,.]$ (see the notation used in the proof of theorem 2.1). Consider the group $SU(1,1)$ of $(1,1)$-unitary unimodular matrices. This is the set of all complex 2×2 matrices with determinant one preserving the hermitian form $|z_1|^2 - |z_2|^2$ (see again the notation used in the proof of theorem 2.1). In other words, they are matrices of the form $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ for which $|a|^2 - |b|^2 = |c|^2 - |d|^2 = 1$, $a\overline{c} - b\overline{d} = 0$, $ad - bc = 1$.

Theorem 2.3. *The matrix of the monodromy operator $\mathscr{M}$ in the basis $(\cos kx, \sin kx)$ (resp. (e^{ikx}, e^{-ikx})) belongs to the group $SL(2,\mathbb{R})$ (resp. $SU(1,1)$).*

Proof. We show that the determinant of the monodromy operator of the Schrödinger equation is equal to one. Note that $(\cos kx, \sin kx)$ is a basis on the space $E^{sr}_{(3)}$. As $\mathscr{B}^x_{(3)} \cos kx = (\cos kx, -k\sin kx)$ and $\mathscr{B}^x_{(3)} \sin kx = (\sin kx, k\cos kx)$, so W is provided with a basis in which the matrix of the operator (we use here the same notation for the operator and the matrix) is written

$$\mathscr{B}^x_{(3)} = \begin{pmatrix} \cos kx & \sin kx \\ -k\sin kx & k\cos kx \end{pmatrix},$$

hence $\det \mathscr{B}^x_{(3)} = k$, independent of x. Let us denote by x^+ the point x to the left of the support of the potential and by x^- the one on the right. We have the following situation :

$$\mathscr{M} : E^{sr}_{(3)} \longrightarrow E^{sr}_{(3)}, \quad a\cos kx + b\sin kx \longmapsto c\cos kx + d\sin kx, \quad (c,d) = \mathscr{M}_u(a,b),$$

$$\mathscr{B}^{x^-}_{(3)} : E^{sr}_{(3)} \longrightarrow W, \quad a\cos kx + b\sin kx \longmapsto (a\cos kx^- + b\sin kx^-, -ak\sin kx^- + bk\cos kx^-),$$

$$\mathscr{B}^{x^+}_{(3)} : E^{sr}_{(3)} \longrightarrow W, \quad c\cos kx + d\sin kx \longmapsto (a\cos kx^+ + b\sin kx^+, -ak\sin kx^+ + bk\cos kx^+),$$

$$g^{x^+}_{x^-} : W \longrightarrow W, \quad (a\cos kx^- + b\sin kx^-, -ak\sin kx^- + bk\cos kx^-) \longmapsto (a\cos kx^+ + b\sin kx^+, -ak\sin kx^+ + bk\cos kx^+).$$

We verify directly that : $g^{x^+}_{x^-} o \mathscr{B}^{x^-}_{(3)} = \mathscr{B}^{x^+}_{(3)} o \mathscr{M}$, and since $\det \mathscr{B}^{x^+}_{(3)} = \mathscr{B}^{x^-}_{(3)}$, so we have $\det \mathscr{M} = \det g^{x^+}_{x^-}$. Now g^x preserves the areas according to Liouville's theorem (indeed, by putting $\psi_1 = \psi$, $\psi_2 = \psi'$, we rewrite equation (2.1) under form

$$\psi_1' = \psi_2 \equiv f_1, \qquad \psi_2' = (u(x) - \lambda)\psi_1 \equiv f_2.$$

Here we have $f = (f_1, f_2)$, $t = x$ and $\operatorname{div} f = \dfrac{\partial \psi_2}{\partial \psi_1} + \dfrac{\partial (u(x) - \lambda)\psi_1}{\partial \psi_2} = 0$). Therefore, $\det g^{x^+}_{x^-} = 1$ and consequently $\det \mathscr{M} = 1$. For the case of $SU(1,1)$, we will show that the matrix (also denoted $\mathscr{M}$) of an operator is real and unimodular in the basis $(\cos kx, \sin kx)$ if and only if it is special $(1,1)$-unitary in complex conjugate basis (e^{ikx}, e^{-ikx}). By setting as in the proof of theorem 1, $z = z_1 e^{ikx} + z_2 e^{-ikx}$ where z_1 and z_2 are the coordinates of the vector z in the basis (e^{ikx}, e^{-ikx}), we obtain $\langle z, z \rangle = |z_1|^2 - |z_2|^2$, i.e., $\langle ., . \rangle$ is of type $(1,1)$. The monodromy operator conserves this hermitian form. Say that $\mathscr{M}$ is real and unimodular in the basis $(\cos kx, \sin kx)$ is equivalent to $\mathscr{M} \in GL(2,\mathbb{R}) \cap SL(2,\mathbb{C})$ or what amounts to the same $\mathscr{M} \in SU(1,1)$ or what is equivalent $\mathscr{M}$ is $(1,1)$-unitary and unimodular in the basis (e^{ikx}, e^{-ikx}). □

Define the solutions $\psi_1(x,\lambda)$ and $\psi_2(x,\lambda)$ of equation (2.1) by the initial conditions : $\psi_1(0,\lambda) = 1$, $\psi_1'(0,\lambda) = 0$, $\psi_2(0,\lambda) = 0$, $\psi_2'(0,\lambda) = 1$. For the simple case $u(x) = 0$, we obviously have

$$\psi_1(x,\lambda) = \cos\sqrt{\lambda}x = 1 + \left(-\frac{1}{2}\lambda\right)x^2 + \left(\frac{1}{24}\lambda^2\right)x^4 + O\left(x^6\right), \tag{2.4}$$

$$\psi_2(x,\lambda) = \frac{1}{\sqrt{\lambda}}\sin\sqrt{\lambda}x = x + \left(-\frac{1}{6}\lambda\right)x^3 + \left(\frac{1}{120}\lambda^2\right)x^5 + O\left(x^7\right).$$

For $\sqrt{\lambda}$, we can choose for example the determination $\sqrt{\lambda} = \sqrt{r}e^{i\frac{\theta}{2}}$ where $\lambda = re^{i\theta}$ with $r > 0$ and $-\pi < \theta < \pi$. Let α be an arbitrary real number. The function $\psi(x,\lambda) = \psi_1(x,\lambda) + \alpha\psi_2(x,\lambda)$ is also solution of equation (2.1) and satisfies the boundary condition $\psi'(0,\lambda) - \alpha\psi(0,\lambda) = 0$. For $\alpha = 0$, we have $\psi(x,\lambda) = \psi_1(x,\lambda)$ and for $\alpha = \infty$, we put $\psi(x,\lambda) = \psi_2(x,\lambda)$. We assume that for $\lambda \in \mathbb{C}$ and $x \geq 0$, we have

$$\psi(x,\lambda) = \cos\sqrt{\lambda}x + \int_0^x K(x,t)\cos\sqrt{\lambda}t\,dt, \tag{2.5}$$

where K is a function to be determined, subject to the condition of having partial derivatives of order one and order two continuous in the set of real pairs (x,t) such that : $0 \leq t \leq x$. In other words, we look for $\psi(x,.)$ as a perturbation of the function $x \mapsto \psi(x,\lambda) = \cos\sqrt{\lambda}x$ and precisely, as a transform $(I+K)\psi_1(x,.)$ where K is a Volterra operator in $[0,+\infty[$. We will look for the conditions that $K(x,t)$ must satisfy for the function (2.5) to be a solution of the differential equation (2.1). From equation (2.5), we get

$$\frac{\partial^2\psi}{\partial x^2}(x,\lambda) = -\lambda\cos\sqrt{\lambda}x + \frac{dK(x,x)}{dx}\cos\sqrt{\lambda}x - \sqrt{\lambda}K(x,x)\sin\sqrt{\lambda}x + \left.\frac{\partial K(x,t)}{\partial x}\right|_{t=x}\cos\sqrt{\lambda}x + \int_0^x \frac{\partial^2 K(x,t)}{\partial x^2}\cos\sqrt{\lambda}t\,dt. \tag{2.6}$$

Let's calculate the expression $\lambda\int_0^x K(x,t)\cos\sqrt{\lambda}t\,dt$, by doing two integrations in parts, we get

$$\lambda\int_0^x K(x,t)\cos\sqrt{\lambda}t\,dt = \sqrt{\lambda}K(x,x)\sin\sqrt{\lambda}x + \left.\frac{\partial K(x,t)}{\partial t}\right|_{t=x}\cos\sqrt{\lambda}x - \left.\frac{\partial K(x,t)}{\partial t}\right|_{t=0} - \int_0^x \frac{\partial^2 K(x,t)}{\partial t^2}\cos\sqrt{\lambda}t\,dt. \tag{2.7}$$

To calculate expression (2.1), substitute (2.6) and (2.7),

$$\begin{aligned} 0 &= \psi'' + (\lambda - u(x))\psi \\ &= \frac{dK(x,x)}{dx}\cos\sqrt{\lambda}x + \left(\frac{\partial K(x,t)}{\partial t} + \frac{\partial K(x,t)}{\partial x}\right)_{x=t}\cos\sqrt{\lambda}x - \left.\frac{\partial K(x,t)}{\partial t}\right|_{t=0} - u(x)\cos\sqrt{\lambda}x \\ &+ \int_0^x \left(\frac{\partial^2 K(x,t)}{\partial x^2} - \frac{\partial^2 K(x,t)}{\partial t^2} - u(x)K(x,t)\right)\cos\sqrt{\lambda}t\,dt. \end{aligned}$$

We have

$$\frac{\partial^2 K(x,t)}{\partial x^2} - u(x)K(x,t) = \frac{\partial^2 K(x,t)}{\partial t^2}, \tag{2.8}$$

with the boundary conditions

$$\left.\frac{\partial K(x,t)}{\partial t}\right|_{t=0} = 0, \tag{2.9}$$

$$\frac{dK(x,x)}{dx} = \frac{1}{2}u(x). \tag{2.10}$$

For the initial conditions, we have $\psi(0,\lambda) = 1$ and $\psi'(0,\lambda) = K(0,0)$. As $\psi'(0,\lambda) - \alpha\psi(0,\lambda) = 0$, then $K(0,0) = \alpha$. Therefore,

$$K(x,x) = \alpha + \frac{1}{2}\int_0^x u(t)dt. \tag{2.11}$$

If $u(x)$ has a continuous derivative, then there exists a unique solution of (2.8), satisfying conditions (2.9) and (2.11). Hence, there exists a satisfying function $K(x,t)$ (2.5). Let's solve equation (2.5) as an equation of Volterra, we get

$$\cos\sqrt{\lambda}x = \psi(x,\lambda) - \int_0^x K_1(x,t)\psi(t,\lambda)dt, \tag{2.12}$$

and in the same way as before, we show that $K_1(x,t)$ is solution of the equation

$$\frac{\partial^2 K_1(x,t)}{\partial x^2} = \frac{\partial^2 K_1(x,t)}{\partial t^2} - u(t)K_1(x,t),$$

with the conditions $\left(\frac{\partial K_1}{\partial t} - \alpha K_1\right)_{t=0} = 0$, $K_1(x,x) = \alpha + \frac{1}{2}\int_0^x u(t)dt$.

For the case $\alpha = \infty$, we look for $\psi(x,\lambda)$ as a perturbation of the function $x \longmapsto \psi(x,\lambda) = \frac{1}{\sqrt{\lambda}}\sin\sqrt{\lambda}x$ (see expression (2.4)) or what is equivalent as a transform $(I+K)\psi_1(x,.)$ where K is a Volterra operator in $[0,+\infty[$. In other words, we set $\lambda \in \mathbb{C}$ and $x \geq 0$,

$$\psi(x,\lambda) = \frac{\sin\sqrt{\lambda}x}{\sqrt{\lambda}} + \int_0^x L(x,t)\frac{\sin\sqrt{\lambda}x}{\sqrt{\lambda}t}dt, \tag{2.13}$$

where L is a function to be determined, subject to the condition of having partial derivatives of order one and order two continuous in the set of real pairs (x,t) such that : $0 \leq t \leq x$. By reasoning as before, we obtain the relation

$$\frac{\partial^2 L(x,t)}{\partial x^2} - u(x)L(x,t) = \frac{\partial^2 L(x,t)}{\partial t^2},$$

with the conditions $L(x,x) = 0$, $L(x,x) = \frac{1}{2}\int_0^x u(t)dt$. By solving equation (2.13), we obtain

$$\frac{\sin\sqrt{\lambda}x}{\sqrt{\lambda}} = \psi(x,\lambda) + \int_0^x L_1(x,t)\psi(t,\lambda)dt. \tag{2.14}$$

The functions $L(x,t)$ and $L_1(x,t)$ have the same properties as the functions $K(x,t)$ and $K_1(x,t)$ previously obtained.

Recall that for every function $f \in L^2(\mathbb{R})$, we have the Parseval identity $\int_0^{\infty} f^2(x)dx = \int_{-\infty}^{\infty} F^2(\lambda).d\rho(\lambda)$ where $\mathscr{F}\{f(x)\} \equiv F(\lambda) = \int_0^{\infty} f(x)\psi(x,\lambda)dx$ is the Fourier transform of $f(x)$ and $\rho(\lambda)$ a monotone function, bounded on any finite interval. The sequence of functions $F_n(\lambda) = \int_0^n f(x)\psi(x,\lambda)dx$, converges in quadratic mean (with respect to the spectral measure $\rho(\lambda)$) to $F(\lambda)$, i.e., $\lim_{n\to\infty}\int_{-\infty}^{\infty}(F(\lambda)-F_n(\lambda))^2 d\rho(\lambda) = 0$. We choose $\rho(\lambda)$ in the following form : $\rho(\lambda) = \frac{2}{\pi}\sqrt{\lambda} + \sigma(\lambda)$ if $\lambda > 0$ and $\rho(\lambda) = \sigma(\lambda)$ if $\lambda < 0$, where $\sigma(\lambda)$ is a measure with compact support satisfying the condition : $\int_{-\infty}^{\infty}|\lambda|.|d\sigma(\lambda)| < +\infty$. For $0 < b < y < a < x$, the functions $\int_a^x \psi(t,\lambda)dt$ and $\int_b^y \cos\sqrt{\lambda}tdt$ are orthogonal with respect to $\rho(\lambda)$. In other words, we have the orthogonality relation :

$$I \equiv \int_{-\infty}^{\infty}\left(\int_a^x \psi(t,\lambda)dt\right)\left(\int_b^y \cos\sqrt{\lambda}tdt\right)d\rho(\lambda) = 0.$$

Indeed, by integrating equation (2.12) from b to y, we obtain

$$\begin{aligned}\int_b^y \cos\sqrt{\lambda}tdt &= \int_b^y \psi(t,\lambda)dt - \int_b^y dt\int_0^t K_1(t,s)\psi(s,\lambda)ds,\\ &= \int_b^y \psi(t,\lambda)dt - \int_0^b \psi(s,\lambda)ds\int_b^y K_1(t,s)dt - \int_b^y \psi(s,\lambda)dt\int_s^y K_1(t,s)dt.\end{aligned}$$

By definition, this function is expressed using the transform (in $\psi(t,\lambda)$) of a null function outside the interval $]b,y[$. Since $]b,y[\cap]a,x[=\emptyset$, we deduce from Parseval's equality that we have $I = 0$.

To obtain the Gelfand-Levitan integral equation [4, 5], we proceed as follows: according to equation (2.5), we have

$$\begin{aligned}\int_a^x \psi(t,\lambda)dt &= \int_a^x \cos\sqrt{\lambda}tdt + \int_a^x dt\int_0^t K(t,s)\cos\sqrt{\lambda}sds,\\ &= \int_a^x \cos\sqrt{\lambda}tdt + \int_0^a \cos\sqrt{\lambda}sds\int_a^x K(t,s)dt + \int_a^x \cos\sqrt{\lambda}sds\int_s^x K(t,s)dt,\end{aligned}$$

by virtue of Lebesgue-Fubini's theorem. Therefore,

$$\begin{aligned}I &= \int_{-\infty}^{\infty}\left(\int_a^x \cos\sqrt{\lambda}tdt\right)\left(\int_b^y \cos\sqrt{\lambda}tdt\right)d\rho(\lambda)\\ &+\int_{-\infty}^{\infty}\left(\int_0^a \cos\sqrt{\lambda}sds\int_a^x K(t,s)dt + \int_a^x \cos\sqrt{\lambda}sds\int_s^x K(t,s)dt\right)\times\left(\int_b^y \cos\sqrt{\lambda}tdt\right)d\rho(\lambda) = 0.\end{aligned}$$

This expression can be written using the definition of $\rho(\lambda)$, in the form

$$\begin{aligned}I &= \int_{-\infty}^{\infty}\left(\int_a^x \cos\sqrt{\lambda}tdt\right)\left(\int_b^y \cos\sqrt{\lambda}tdt\right)d\sigma(\lambda)\\ &+\int_{-\infty}^{\infty}\left(\int_0^a \cos\sqrt{\lambda}sds\int_a^x K(t,s)dt + \int_a^x \cos\sqrt{\lambda}sds\int_s^x K(t,s)dt\right)\times\left(\int_b^y \cos\sqrt{\lambda}tdt\right)d\sigma(\lambda)\\ &+\frac{2}{\pi}\int_{-\infty}^{\infty}\left(\int_a^x \cos\sqrt{\lambda}tdt\right)\left(\int_b^y \cos\sqrt{\lambda}tdt\right)d\sigma(\lambda)\\ &+\frac{2}{\pi}\int_{-\infty}^{\infty}\left(\int_0^a \cos\sqrt{\lambda}sds\int_a^x K(t,s)dt + \int_a^x \cos\sqrt{\lambda}sds\int_s^x K(t,s)dt\right)\times\left(\int_b^y \cos\sqrt{\lambda}tdt\right)d\sigma(\lambda) = 0.\end{aligned}$$

Since $b < y < a < x$, then given the Parseval identity, the third term is equal to zero while the fourth is equal to

$$\int_b^y\left(\int_0^a \cos\sqrt{\lambda}sds\int_a^x K(t,s)dt + \int_a^x \cos\sqrt{\lambda}sds\int_s^x K(t,s)dt\right)ds = \int_b^y ds\int_a^x K(t,s)dt.$$

Therefore,

$$\begin{aligned} I &= \int_{-\infty}^{\infty} \frac{(\sin\sqrt{\lambda}x - \sin\sqrt{\lambda}a)(\sin\sqrt{\lambda}y - \sin\sqrt{\lambda}b)}{\lambda} d\sigma(s) \\ &+ \int_{-\infty}^{\infty} \left(\int_0^a \cos\sqrt{\lambda}sds \int_a^x K(t,s)dt + \int_a^x \cos\sqrt{\lambda}sds \int_s^x K(t,s)dt \right) \times \left(\int_b^y \cos\sqrt{\lambda}sds \right) d\sigma(\lambda) \\ &+ \int_b^y ds \int_a^x K(t,s)dt = 0. \end{aligned}$$

By setting

$$F(x,y) \equiv \int_{-\infty}^{\infty} \frac{\sin\sqrt{\lambda}x \sin\sqrt{\lambda}y}{\lambda} d\sigma(\lambda),$$

and

$$G(x,s) \equiv \begin{cases} \int_a^x K(t,s)dt, & 0 \le s \le a \\ \int_s^x K(t,s)dt, & a \le s \le x \\ 0, & s > x \end{cases}$$

the equation above becomes

$$F(x,y) - F(x,b) - F(a,y) + F(a,b) + \int_b^y ds \int_a^x K(t,s)dt + \int_{-\infty}^{\infty} \left(\int_0^x G(x,s)\cos\sqrt{\lambda}sds \right) \left(\int_b^y \cos\sqrt{\lambda}sds \right) d\sigma(\lambda) = 0.$$

This last equation can still be written, doing an integration by parts and noticing that $G(x,x) = 0$,

$$F(x,y) - F(x,b) - F(a,y) + F(a,b) + \int_b^y ds \int_a^x K(t,s)dt + \int_{-\infty}^{\infty} \left(\int_0^x \frac{\partial G(x,s)}{\partial s} \frac{\sin\sqrt{\lambda}s}{\sqrt{\lambda}} ds \right) \left(\frac{\sin\sqrt{\lambda}y - \sin\sqrt{\lambda}b}{\sqrt{\lambda}} \right) d\sigma(\lambda) = 0. \tag{2.15}$$

But

$$\begin{aligned} &\int_{-\infty}^{\infty} \left(\int_0^x \frac{\partial G(x,s)}{\partial s} \frac{\sin\sqrt{\lambda}s}{\sqrt{\lambda}} ds \right) \left(\frac{\sin\sqrt{\lambda}y - \sin\sqrt{\lambda}b}{\sqrt{\lambda}} \right) d\sigma(\lambda), \\ &= \int_0^x \frac{\partial G(x,s)}{\partial s} \left(\int_{-\infty}^{\infty} \left(\frac{\sin\sqrt{\lambda}s \sin\sqrt{\lambda}y - \sin\sqrt{\lambda}s \sin\sqrt{\lambda}b}{\lambda} \right) d\sigma(\lambda) \right) ds, \\ &= \int_0^x \frac{\partial G(x,s)}{\partial s} \left(F(s,y) - F(s,b) \right) ds, \\ &= -\int_0^x G(x,s) \left(\frac{\partial F(s,y)}{\partial s} - \frac{\partial F(s,b)}{\partial s} \right) ds, \\ &= -\int_0^a \left(\frac{\partial F(s,y)}{\partial s} - \frac{\partial F(s,b)}{\partial s} \right) ds \left(\int_a^x K(t,s)dt \right) - \int_a^x \left(\frac{\partial F(s,y)}{\partial s} - \frac{\partial F(s,b)}{\partial s} \right) ds \left(\int_s^x K(t,s)dt \right), \\ &= \int_a^x dt \int_0^t \left(\frac{\partial F(s,y)}{\partial s} - \frac{\partial F(s,b)}{\partial s} \right) ds, \end{aligned}$$

so equation (2.15) becomes

$$F(x,y) - F(x,b) - F(a,y) + F(a,b) + \int_b^y ds \int_a^x K(t,s)dt + \int_a^x dt \int_0^t \left(\frac{\partial F(s,y)}{\partial s} - \frac{\partial F(s,b)}{\partial s} \right) ds = 0.$$

Deriving this expression with respect to y and then with respect to x (the support of the measure σ is compact), we obtain

$$\frac{\partial^2 F}{\partial x \partial y} + \int_0^x K(x,s) \frac{\partial^2 F(s,y)}{\partial s \partial y} + K(x,y) = 0.$$

By setting $f(x,y) \equiv \dfrac{\partial^2 F}{\partial x \partial y}$, we finally obtain the Gelfand-Levitan integral equation for the function $x \longmapsto K(x,y)$ valid for $0 < y < x$,

$$f(x,y) + K(x,y) + \int_0^x K(x,s) f(s,y) ds = 0, \qquad y \leq x. \tag{2.16}$$

For the case $\alpha = \infty$, i.e., $\psi(x,\lambda) = \psi_2(x,\lambda)$, just integrate the two members of equation (2.14) from 0 to x and use a similar reasoning. Under the continuity assumption of K, equation (2.16) must be checked for $x = 0$ and $x = y$. Note also that if we set x in the previous equation, then we will obtain the so called Fredholm's linear integral equation. We can prove that, conversely, equation (2.16) admits a single continuous solution in the set of pairs of real numbers such that : $0 \leq t \leq x$. We will not look for the solution at this level, it will be done later (in the next section) when we treat the Korteweg-de-Vries equation.

3. KdV equation and the inverse diffusion method

Let us first examine some particular solutions of the equation of KdV (1.1), of the kind of progressive waves $u(x,t) = s(x - ct)$, where c is the phase velocity. By replacing this expression in (1.1), we obtain $-c\dfrac{\partial s}{\partial x} - 6s\dfrac{\partial s}{\partial x} + \dfrac{\partial^3 s}{\partial x^3} = 0$. By integrating this equation with respect to x and imposing the boundary condition that s and its derivatives decrease for $|x| \longrightarrow \infty$, we get $-cs - 3s^2 + \dfrac{\partial^2 s}{\partial x^2} = 0$, hence $-cs - 2s^3 + \left(\dfrac{\partial s}{\partial x}\right)^2 = 0$, and the exact expression of the solution s requires the use of elliptic functions. Suppose that $\dfrac{\partial s}{\partial x}(0) = 0$, in which case the solution of this last equation is $s(x - ct) = -\dfrac{c}{2}\,\mathrm{sech}^2 \dfrac{\sqrt{c}}{2}(x - ct)$, where sech denotes the hyperbolic secant, i.e., $\dfrac{1}{\cosh}$. Therefore $u(x,0) = u_0\,\mathrm{sech}^2 \dfrac{x}{l}$, where $u_0 \equiv -\dfrac{c}{2}$ et $l^2 \equiv \dfrac{4}{c}$. This expression shows that u removes infinitely long time in the position $u \simeq 0$, then it reaches the value u_0, is reflected on this point and returns again in the position of $u \simeq 0$. This solution is called soliton. To obtain this solution, we can use the so-called Bäcklund transformations for the Korteweg de Vries equation.

When solitons collide, dimensions and speeds of solutions do not change after collision. This phenomenon has suggested the idea of conservation laws. And indeed, Kruskal, Zabusky, Lax, Gardner, Green and Miura [6, 7, 8, 9, 10, 11, 12, 13, 14] have been able to find a whole series of first integrals for the KdV equation. These integrals are of the form $\int P_n\left(u, ..., u^{(n)}\right) dx$, where P_n is a polynomial. Indeed, the conservation equations that can be deduced from the KdV equation take the following general form : $\dfrac{\partial P_n}{\partial t} + \dfrac{\partial Q_n}{\partial x} = 0$, where P_n and Q_n form a series of functions of which here are the first three:

(i) The KdV equation can be written in the form

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left(-3u^2 + \frac{\partial^2 u}{\partial x^2}\right) = 0 \Longrightarrow P_1 = u, \quad Q_1 = -3u^2 + \frac{\partial^2 u}{\partial x^2}.$$

(ii) Multiply the KdV equation by u, this gives

$$u\frac{\partial u}{\partial t} - 6u^2\frac{\partial u}{\partial x} + u\frac{\partial^3 u}{\partial x^3} = 0, \qquad \frac{\partial}{\partial t}\left(\frac{u^2}{2}\right) + \frac{\partial}{\partial x}\left(-2u^3 + u\frac{\partial^2 u}{\partial x^2} - \frac{1}{2}\left(\frac{\partial u}{\partial x}\right)^2\right) = 0.$$

Hence,

$$P_2 = \frac{u^2}{2}, \qquad Q_2 = -2u^3 + u\frac{\partial^2 u}{\partial x^2} - \frac{1}{2}\left(\frac{\partial u}{\partial x}\right)^2.$$

(iii) We have

$$\left(3u^2 - \frac{\partial^2 u}{\partial x^2}\right)\left(\frac{\partial u}{\partial t} - 6u\frac{\partial u}{\partial x} + \frac{\partial^3 u}{\partial x^3}\right) = 0,$$

$$\left(3u^2\frac{\partial u}{\partial t} + \frac{\partial u}{\partial x}\frac{\partial^2 u}{\partial x \partial t}\right) + \left(-18u^3\frac{\partial u}{\partial x} + 3u^2\frac{\partial^3 u}{\partial t^3} + 6u\frac{\partial u}{\partial x}\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial x^2}\frac{\partial^3 u}{\partial x^3} - \frac{\partial^2 u}{\partial x^2}\frac{\partial u}{\partial t} - \frac{\partial u}{\partial x}\frac{\partial^2 u}{\partial x \partial t}\right) = 0.$$

Therefore,

$$\frac{\partial}{\partial t}\left(u^3 + \frac{1}{2}\left(\frac{\partial u}{\partial x}\right)^2\right) + \frac{\partial}{\partial t}\left(-\frac{9}{2}u^4 + 3u^2\frac{\partial^2 u}{\partial x^2} - \frac{1}{2}\left(\frac{\partial^2 u}{\partial x^2}\right)^2 - \frac{\partial u}{\partial x}\frac{\partial u}{\partial t}\right) = 0.$$

Consequently,

$$P_3 = u^3 + \frac{1}{2}\left(\frac{\partial u}{\partial x}\right)^2, \quad Q_3 = -\frac{9}{2}u^4 + 3u^2\frac{\partial^2 u}{\partial x^2} - \frac{1}{2}\left(\frac{\partial^2 u}{\partial x^2}\right)^2 - \frac{\partial u}{\partial x}\frac{\partial u}{\partial t}.$$

If u vanishes for $x \to \infty$, we get $\frac{\partial}{\partial t}\int P_n dx = 0$, then $\int P_n dx$ are first integrals of the KdV equation. Let $u(x,t) = \frac{\partial y}{\partial x}(x,t)$ and suppose that $\frac{\partial y}{\partial t}, \frac{\partial y}{\partial x}, \frac{\partial^3 y}{\partial t^3}$ decay when $|x| \to \infty$. The KdV equation, is written

$$\frac{\partial y}{\partial t} - 3\left(\frac{\partial y}{\partial x}\right)^2 + \frac{\partial^3 y}{\partial t^3} = 0.$$

Hence,

$$\frac{\partial}{\partial t}\int_{-\infty}^{\infty} y(x,t)dx = 3\int_{-\infty}^{\infty}\left(\frac{\partial y}{\partial x}\right)^2(x,t)dx = 3\int_{-\infty}^{\infty} u^2(x,t)dx = \text{constant}.$$

Since $u = \frac{\partial y}{\partial x}$, we have also

$$\frac{\partial}{\partial t}\int_{-\infty}^{\infty} y(x,t)dx = \frac{\partial}{\partial t}\int_{-\infty}^{\infty}\int_{-\infty}^{x} u(z,t)dzdx = x\frac{\partial}{\partial t}\int_{-\infty}^{x} u(z,t)dz\Big|_{-\infty}^{\infty} - \frac{\partial}{\partial t}\int_{-\infty}^{\infty} u(z,t)dx = -\frac{\partial}{\partial t}\int_{-\infty}^{\infty} xu(x,t)dx,$$

because by hypothesis u^2 and $\frac{\partial^2 u}{\partial x^2}$ tend to 0 when $|x| \to \infty$. Comparing the two expressions obtained, we obtain a new first integral

$$\frac{\partial}{\partial t}\int_{-\infty}^{\infty} xu(x,t)dx = \text{constante}.$$

Lax [7] showed that the equation of KdV is equivalent to the equation : $\frac{dA}{dt} = [A,B]$, where

$$A = -\frac{\partial^2}{\partial x^2} + u(x,t), \qquad B = 4\frac{\partial^3}{\partial x^3} - 3\left(u\frac{\partial}{\partial x} + \frac{\partial u}{\partial x}\right).$$

We deduce that the spectrum of A is conserved : if A is a symmetric operator ($A^\top = A$) and T an orthogonal transformation ($T^\top = T^{-1}$), then the spectrum of $T^{-1}AT$ coincides with that of A. The appearance of an infinite series of first integrals is easily explained by the Lax equation.
The Sturm-Liouville equation $A\psi = \lambda\psi$, where λ is a real parameter, can be written in the form

$$\frac{\partial^2 \psi}{\partial x^2} + (\lambda - u(x,t))\psi = 0. \tag{3.1}$$

This equation reminds us the stationary Schrödinger equation. We will see that the complete solution of the KdV equation is closely related to the solution of this equation. We will look at solutions for which u decreases fast enough for $x \longrightarrow \pm\infty$. There are other interesting conditions to know: the case where $u(x,t)$ tends to different constants for $|x| \longrightarrow \infty$ and the one where $u(x,t)$ is periodic in x. So consider equation (3.1) where $u(x,t)$ is the solution of the KdV equation (1.1). It is assumed that after a certain time equation (3.1) has N bound states with energy $\lambda_n = -k_n^2$, $n = 1,2,...,N$ and continuous states with for energy $\lambda = k^2$. We draw u from equation (3.1) and replace it in equation (1.1). After a long calculation, after multiplying by ψ^2, we get the expression

$$\frac{\partial \lambda}{\partial t}.\psi^2 + \frac{\partial}{\partial x}\left(\psi\frac{\partial \Upsilon}{\partial x} - \frac{\partial \psi}{\partial x}\Upsilon\right) = 0, \tag{3.2}$$

where $\Upsilon \equiv \frac{\partial \psi}{\partial t} + \frac{\partial^3 \psi}{\partial x^3} - 3(u+\lambda)\frac{\partial \psi}{\partial x}$.

Theorem 3.1. *a) For the study of the discrete part of the spectrum* $\lambda_n(t) = -k_n^2(t)$*, we consider* ψ_n *(measurable and square integrable function) and we show that if* ψ_n *and* $\dfrac{\partial \psi_n}{\partial x}$ *tend to zeros when* $|x|$ *goes to infinity, then* $\lambda_n(t) =$ *constant and the solution of equation* (3.1) *is given by* $\psi_n(t) = c_n(0)e^{k_n(x-4k_n^2 t)}$*, where* $c_n(0)$ *is determined by the initial condition* $u(x,0) = u_0(x)$ *of the KdV equation.*

b) For the study of the continuous part of the spectrum $\lambda(t) = k^2(t)$*, we assume that a stationary plane wave propagates from* $x = -\infty$ *and meets a potential* $u(x,t)$ *with a transmission coefficient* T *and a reflection coefficient* R*. In this case equation* (3.1) *admits a solution* ψ *such that :*

$$\psi = \begin{cases} T(k,t)e^{ikx}, & x \to +\infty \text{ (i.e., to the right of the potential barrier)} \\ e^{ikx} + R(k,t)e^{-ikx}, & x \to + -\infty \text{ (i.e., to the left of the potential barrier)} \end{cases}$$

where $|R|^2 + |T|^2 = 1$*. If* $u \simeq 0$ *for* $|x| \to \infty$*, then we have* $T(k,t) = T(k,0)$ *and* $R(k,t) = R(k,0)e^{-8ik^3 t}$ *where* $R(k,0)$ *and* $T(k,0)$ *are determined by the initial condition* $u(x,0) = u_0(x)$ *of the KdV equation.*

Proof. a) Just integrate equation (3.2), this gives

$$\frac{\partial \lambda_n}{\partial t}.\int_{-\infty}^{\infty} \psi_n^2 dx + \psi_n \frac{\partial \Upsilon}{\partial x} - \frac{\partial \psi_n}{\partial x}\Upsilon = 0.$$

By hypothesis, $\psi_n \in L^2$ and ψ_n, $\dfrac{\partial \psi_n}{\partial x}$ tend to zeros when $|x|$ goes to infinity, so $\psi_n \dfrac{\partial \Upsilon}{\partial x} - \dfrac{\partial \psi_n}{\partial x}\Upsilon$ tends to 0 for $|x| \to \infty$ and we deduce that $\lambda_n(t) =$ constant. Now, since $\dfrac{\partial \lambda}{\partial t} = 0$, then equation (3.2) becomes $\dfrac{\partial}{\partial x}\left(\psi \dfrac{\partial \Upsilon}{\partial x} - \dfrac{\partial \psi}{\partial x}\Upsilon\right) = 0$. Let's integrate this expression twice, $\dfrac{\left(\psi \frac{\partial \Upsilon}{\partial x} - \frac{\partial \psi}{\partial x}\Upsilon\right)}{\psi^2} = \dfrac{A}{\psi^2}$, i.e., $\left(\dfrac{\Upsilon}{\psi}\right)' = \dfrac{A}{\psi^2}$, hence, $\Upsilon = \psi \displaystyle\int \dfrac{A(t)}{\psi^2} dx + B(t)\psi$, where $A(t)$ and $B(t)$ are integration constants. So we have

$$\frac{\partial \psi_n}{\partial t} + \frac{\partial^3 \psi_n}{\partial x^3} - 3(u + \lambda_n)\frac{\partial \psi_n}{\partial x} = \psi_n \int \frac{A_n}{\psi_n^2} dx + B_n \psi_n. \tag{3.3}$$

Note that $A_n(t) = 0$ because ψ_n satisfies (3.3) and decreases to zeros for $t \to -\infty$. Let's consider $u \cong 0$ for $x \to -\infty$ because otherwise ψ_n would not have the decay assumption. Multiply (3.3) by ψ_n and integrate

$$\int_{-\infty}^{\infty} \psi_n \frac{\partial \psi_n}{\partial t} dx + \int_{-\infty}^{\infty} \left(\psi_n \frac{\partial^3 \psi_n}{\partial x^3} - 3\lambda_n \psi_n \frac{\partial \psi_n}{\partial x}\right) dx = B_n \int_{-\infty}^{\infty} \psi_n^2 dx.$$

This expression can be written, by adding and subtracting $\dfrac{\partial \psi_n}{\partial x}\dfrac{\partial^2 \psi_n}{\partial x^2}$,

$$\int_{-\infty}^{\infty} \frac{1}{2}\frac{\partial \psi_n^2}{\partial t} dx + \int_{-\infty}^{\infty} \frac{\partial}{\partial x}\left(\psi_n \frac{\partial^2 \psi_n}{\partial x^2} - \frac{3}{2}\lambda_n \psi_n^2 - \frac{1}{2}\left(\frac{\partial \psi_n}{\partial x}\right)^2\right) dx = B_n \int_{-\infty}^{\infty} \psi_n dx.$$

We have $B_n(t) = 0$ because $\psi_n \in L^2$ and decreases to zeros when $x \to -\infty$. Since $u \cong 0$ for $x \to -\infty$, then from equation (3.2), it comes $\psi_n(x,t) = c_n(t)e^{k_n x}$, $x \to -\infty$. By replacing the latter in equation (3.3), we obtain $\left(\dfrac{\partial c_n}{\partial t} + 4c_n k_n^3\right)e^{k_n x} = 0$, hence $c_n(t) = c_n(0)e^{-4k_n^3 t}$. Consequently, $\psi_n(x,t) = c_n(0)e^{k_n(x-4k_n^2 t)}$.

b) Choose $\lambda =$ constant since the spectrum for $\lambda > 0$ is continuous. So equation (3.3) remains valid,

$$\frac{\partial \psi}{\partial t} + \frac{\partial^3 \psi}{\partial x^3} - 3(u + \lambda)\frac{\partial \psi}{\partial x} = \psi \int \frac{A}{\psi^2} dx + B\psi. \tag{3.4}$$

For $u \cong 0$, when $x \to +\infty$, we replace $\psi = T(k,t)e^{ikx}$, $\lambda = k^2$ in equation (3.4) and we get $\dfrac{\partial T}{\partial t} - 4ik^3 T = \dfrac{A}{T}\displaystyle\int e^{-2ikx} dx + BT$. For this equation to preserve meaning when $x \to +\infty$, we must have $A = 0$, hence

$$\frac{\partial T}{\partial t} - (4ik^3 + B)T = 0. \tag{3.5}$$

Similarly, for $u \cong 0$, when $x \to -\infty$, we replace $\psi = e^{ikx} + R(k,t)e^{-ikx}$, $\lambda = k^2$ in equation (3.4) and we get

$$\left(\frac{\partial R}{\partial t} + 4ik^3R - BR\right)e^{-ikx} - (4ik^3 + B)e^{ikx} = A(e^{ikx} + Re^{-ikx})\int \frac{dx}{e^{2ikx} + R^2e^{-2ikx} + 2R}.$$

For $x \to +\infty$, the equation above preserves a sense if $A = 0$ and is written

$$\left(\frac{\partial R}{\partial t} + 4ik^3R - BR\right)e^{-ikx} - (4ik^3 + B)e^{ikx} = 0.$$

For $4ik^3 + B = 0$, equation (3.5) implies that $T(k,t) = T(k,0)$ while the condition $\frac{\partial R}{\partial t} + 4ik^3R - BR = 0$, gives us $R(k,t) = R(k,0)e^{-8ik^3t}$. □

The knowledge of $c_n(t)$, $k_n(t)$, $n = 1, 2, ..., N$ and $R(k,t)$ allows to express $u(x,t)$ for any time; it is the problem of the inverse diffusion. The latter is reduced to the solution $K(x,y;t)$ (to simplify the notations, the reader can obviously use $K(x,y)$ instead of $K(x,y;t)$), of the Gelfand-Levitan linear integral equation :

$$K(x,y;t) + I(x+y,t) + \int_{-\infty}^{x} I(y+z,t)K(x,z;t)dz = 0, \quad y \le x \tag{3.6}$$

where

$$I(x+y,t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} R(k,t)e^{-ik(x+y)}dk + \sum_{n=1}^{N} c_n^2(t)e^{k_n(t)(x+y)}.$$

The solution $u(x,t)$ of the KdV equation is then given (see (2.10)) by

$$u(x,t) = 2\frac{d}{dx}K(x,x;t). \tag{3.7}$$

The nonlinear KdV equation is transformed into the linear Gelfand-Levitan equation. The initial problem is thus completely solved. This method presents two major simplifications. First, in the analytical approach of the solution of the KdV equation, it suffices at each stage to solve only linear equations. Then t only appears parametrically and more than for all t the Gelfand-Levitan equation seems superficially to be an integral equation of two variables, actually x intervenes as a parameter and so we have to do to a family of integral equations for the functions $K(x,y)$ of a single variable y. Before dealing with the general case, i.e., the case of distinct N solitons, let us return first to the case of a soliton and therefore consider the solution $u(x,t) = -\frac{c}{2}\,\text{sech}^2\frac{\sqrt{c}}{2}(x - ct)$ of the KdV equation obtained previously with the following initial condition : $u(x,0) = -2\,\text{sech}^2 x$, where by convention we put $c = 4$. The Schrödinger equation (3.1) is written

$$\frac{\partial^2\psi}{\partial x^2} + (2\,\text{sech}^2 x + \lambda)\psi = 0. \tag{3.8}$$

To study equation (3.8), one poses

$$\psi = A\,\text{sech}^\alpha x.w(x), \tag{3.9}$$

where A is an arbitrary amplitude, $\alpha^2 = -\lambda$ and w satisfies the equation

$$\frac{\partial^2 w}{\partial x^2} - 2\alpha\tanh x\frac{\partial w}{\partial x} + (2 + \alpha - \alpha^2)\,\text{sech}^2 x.w = 0.$$

By doing the substitution $u = \frac{1}{2}(1 - \tanh x)$, the last equation comes down to a hypergeometric differential equation or Gaussian equation :

$$u(1-u)\frac{\partial^2 w}{\partial u^2} + (c - (a+b+1)u)\frac{\partial w}{\partial u} - abw = 0,$$

where a,b,c denote constants and are equal to $a=2+\alpha$, $b=-1+\alpha$, $c=1+\alpha$. This equation presents three regular singular points : $u=0$, $u=1$, $u=\infty$. The solution of this equation for $u=0$ is

$$w \equiv F(a,b,c,u) = 1+\frac{ab}{c}.\frac{u}{1!}+\frac{a(a+1)b(b+1)}{c(c+1)}.\frac{u^2}{2!}+\frac{a(a+1)...(a+n-1)b(b+1)...(b+n-1)}{c(c+1)...(c+n-1)}.\frac{u^n}{n!}+\cdots \tag{3.10}$$

For $x\to\infty$ (i.e., when $u\to 0$), we have $w\to 1$. According to (3.9), we have $\psi = A2^{\alpha}(e^x+e^{-x})^{-\alpha}.w(x)$ and this one tends to $Ae^{2\alpha}e^{-\alpha x}$ when $x\to\infty$. To represent a plane wave Ae^{ikx} going to $+\infty$, we will put $\alpha=-ik$. The asymptotic form of the wave function for $x\to-\infty$ ($u\to 1$) is obtained by transforming the hypergeometric function using the well-known functional relation :

$$F(a,b,c,u)=\frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}F(a,b,a+b-c+1,1-u)+(1-u)^{c-a-b}\frac{\Gamma(c)\Gamma(a+b-c)}{\Gamma(a)\Gamma(b)}F(c-a,c-b,c-a-b+1,1-u),$$

where $\Gamma(z)=\int_0^{\infty}e^{-t}e^{z-1}dt$, Re $z>0$, is the Euler Gamma function. Taking into account (3.10) and the expression above, the relation (3.9) becomes

$$\psi = A\ \text{sech}^{\alpha}x\left[\frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}\left(1+\frac{ab}{a+b-c+1}(1-u)+\cdots\right)\right.$$
$$\left.+\,(1-u)^{c-a-b}\frac{\Gamma(c)\Gamma(a+b-c)}{\Gamma(a)\Gamma(b)}\left(1+\frac{(c-a)(c-b)}{c-a-b+1}(1-u)+\cdots\right)\right].$$

When $u\to 1$ ($x\to-\infty$), we have $(1-u)^{c-a-b}\to e^{-2\alpha x}$ and since $\alpha=-ik$, then

$$\psi \longrightarrow Ae^{\alpha}\frac{\Gamma(c)\Gamma(a+b-c)}{\Gamma(a)\Gamma(b)}\left(e^{ikx}+\frac{\Gamma(c-a-b)\Gamma(a)\Gamma(b)}{\Gamma(c-a)\Gamma(c-b)\Gamma(a+b-c)}\right).$$

This expression combined with the fact that ψ tends to $Ae^{2\alpha}e^{-\alpha}$, $x\to\infty$, give us the transmission coefficient T and the reflection coefficient R,

$$T=\frac{\Gamma(a)\Gamma(b)}{\Gamma(c)\Gamma(a+b-c)},\qquad R=\frac{\Gamma(c-a-b)\Gamma(a)\Gamma(b)}{\Gamma(c-a)\Gamma(c-b)\Gamma(a+b-c)}.$$

We have $k_1=1$, $c(0)=\sqrt{2}$, $R(k,0)=0$. For an individual soliton, equation (1.1) has a precise solution. It turns out that the soliton of amplitude u_0 has only one discrete level with eigenvalue $\lambda=\frac{u_0}{2}$, while the next level corresponds to the point $\lambda=0$ (with the respective eigenfunction $\psi=\tanh x$) and already belongs to the continuous spectrum. The Gelfand-Levitan equation (3.6) where $I(\mu,t)=c_1^2(t)e^{k_1\mu}=c_1^2(0)e^{-8k_1t}e^{k_1t}=2e^{-8t+\mu}$ is written

$$K(x,y;t)+2e^{-8t+x+y}+2e^{-8t+y}\int_{-\infty}^{x}e^zK(x,z;t)dz=0.$$

By putting $K(x,y,t)=f(x)e^y$, we obtain $f(x)+2e^{-8t+x}+e^{-8t+2x}f(x)=0$, hence, $f(x)=-2\frac{e^{-x}}{1+e^{8t-2x}}$. Therefore, the solution (3.7) of the KdV equation in the case of a solitary wave is

$$u(x,t)=2\frac{d}{dx}K(x,x,t)=-\frac{2}{\cosh^2(x-4t)}=-2\ \text{sech}^2(x-4t).$$

We will now look at the case of N-solitons through the procedure suggested by Gardner, Green, Kruskal, Miura [8] and to use the results of [15]. In order to solve the Gelfand-Levitan equation (3.6), where $R(k,t)=0$, one poses

$$K(x,y)=\sum_{n=1}^{N}w_n(x,t)e^{k_ny}, \tag{3.11}$$

where w_n are functions to be determined. By replacing this expression in the Gelfand-Levitan equation, we obtain the following system of linear algebraic equations for w_n, $n=1,...,N$:

$$\begin{cases} w_1(x,t)+c_1^2(t)e^{k_1x}+\sum_{m=1}^{N}c_1^2(t)\frac{e^{(k_1+k_m)x}}{k_1+k_m}w_m(x,t)=0,\\ \vdots\\ w_N(x,t)+c_N^2(t)e^{k_Nx}+\sum_{m=1}^{N}c_N^2(t)\frac{e^{(k_N+k_m)x}}{k_N+k_m}w_m(x,t)=0.\end{cases}$$

Define the following notations : $A = \left(c_n^2(t)e^{(k_n+k_m)x}\right)$, $W = (w_1 \cdots w_N)^\top$, $G = \left(c_1^2(t)e^{k_1 x} \cdots c_N^2(t)e^{k_N x}\right)^\top$,

$$P \equiv (P_{nm}) = \left(\delta_{nm} + c_n^2(t)\frac{e^{(k_n+k_m)x}}{k_n+k_m}\right) = I + A, \tag{3.12}$$

where I is the unit matrix. With these notations, the system above is written $PW = -G$, and it is easy to show that it has a unique solution. From equation (3.11), we draw $K(x,x) = h^\top w = -h^\top P^{-1}G$, $h \equiv \left(e^{k_1 x} \cdots e^{k_N x}\right)^\top$. Or

$$\frac{d}{dx}P_{nm} = c_m^2 e^{k_n x}.e^{k_m x}, \qquad \det P = \sum_{n=1}^{N}\left(\delta_{nm} + c_n^2(t)\frac{e^{(k_n+k_m)x}}{k_n+k_m}\right)\alpha_{nm},$$

and $P^{-1} = \dfrac{\alpha_{nm}}{\det P}$, where α_{nm} is the cofactor of P, so

$$K(x,x) = -\sum_{n,m}\frac{\alpha_{nm}}{\det P}\frac{d}{dx}P_{nm} = -\frac{1}{\det P}\frac{d}{dx}(\det P) = -\frac{d}{dx}\ln\det P,$$

and according to (3.7), we have $u = 2\dfrac{d}{dx}K(x,x) = -2\dfrac{d^2}{dx^2}\ln\det P$. Therefore,

Theorem 3.2. *The solution of the KdV equation is given by the function*

$$u = -2\frac{d^2}{dx^2}\ln\det P,$$

where P is defined by (3.12) *and whose $c_n(t) = c_n(0)e^{-4k_n^3 t}$, with $k_n > 0$ distinct.*

The function obtained in the this theorem is negative for all x, continuous and behaves like the exponential when $|x| \to \infty$. To get an idea of the behavior of solitons and in particular their asymptotic behavior, suppose that $k_1 < k_2 < ... < k_{N-1} < k_N$. We will need the following result :

$$\Delta \equiv \begin{vmatrix} \frac{1}{a_1-b_1} & \frac{1}{a_1-b_2} & \cdots & \frac{1}{a_1-b_n} \\ \frac{1}{a_2-b_1} & \frac{1}{a_2-b_2} & \cdots & \frac{1}{a_2-b_n} \\ \vdots & & \ddots & \\ \frac{1}{a_n-b_1} & \frac{1}{a_n-b_2} & \cdots & \frac{1}{a_n-b_n} \end{vmatrix} = (-1)^{\frac{n(n-1)}{2}}\frac{\prod_{j<k}(a_j-a_k)\prod_{j<k}(b_j-b_k)}{\prod_{j,k}(a_j-b_k)}.$$

Consider the following determinant :

$$\Delta = \begin{vmatrix} 1+c_1^2 b_{11} & c_1^2 b_{12} & c_1^2 b_{13} \\ c_2^2 b_{21} & 1+c_2^2 b_{22} & c_2^2 b_{23} \\ c_3^2 b_{31} & c_3^2 b_{32} & 1+c_3^2 b_{33} \end{vmatrix},$$

where $c_1 c_2 c_3 \neq 0$. If we divide the 1^{st} line by c_1, the 2^{nd} line by c_2 and the 3^{nd} line by c_3, we will have

$$\Delta = c_1 c_2 c_3 \begin{vmatrix} \frac{1+c_1^2 b_{11}}{c_1} & c_1 b_{12} & c_1 b_{13} \\ c_2 b_{21} & \frac{1+c_2^2 b_{22}}{c_2} & c_2 b_{23} \\ c_3 b_{31} & c_3 b_{32} & \frac{1+c_3^2 b_{33}}{c_3} \end{vmatrix}.$$

Multiply the 1^{st} column by c_1, the 2^{nd} column by c_2 and the 3^{nd} column by c_3, we get

$$\Delta = \begin{vmatrix} 1+c_1^2 b_{11} & c_1 c_2 b_{12} & c_1 c_3 b_{13} \\ c_2 c_1 b_{21} & 1+c_2^2 b_{22} & c_2 c_3 b_{23} \\ c_3 c_1 b_{31} & c_3 c_2 b_{32} & 1+c_3^2 b_{33} \end{vmatrix}.$$

So for the determinant of order N, just use the same procedure, i.e., by dividing the j^{nd} row by c_j and multiply the j^{nd} column by c_j

Theorem 3.3. *The explicit solution of N-solitons of the KdV equation is given by*

$$u(x,t) = \begin{cases} -2\sum_{n=1}^{N} k_n^2 \, sech^2(k_n\xi_n + \delta_n^+), & t \to +\infty \\ -2\sum_{n=1}^{N} k_n^2 \, sech^2(k_n\xi_n + \delta_n^-), & t \to -\infty \end{cases}$$

where the phase changes are given by

$$\delta_n^+ \equiv \frac{1}{2}\ln\frac{c_n^2}{2k_n}\left(\prod_{j=1}^{n-1}\frac{k_j-k_n}{k_j+k_n}\right)^2, \qquad \delta_n^- \equiv \frac{1}{2}\ln\frac{c_n^2}{2k_n}\left(\prod_{j=n+1}^{N}\frac{k_j-k_n}{k_j+k_n}\right)^2.$$

Proof. The determinant of the matrix P is written explicitly in the form

$$\det P = \begin{vmatrix} 1+\frac{c_1^2(t)}{2k_1}e^{2k_1x} & \frac{c_1^2(t)}{k_1+k_2}e^{(k_1+k_2)x} & \dots & \frac{c_1^2(t)}{k_1+k_j}e^{(k_1+k_j)x} & \dots & \frac{c_1^2(t)}{k_1+k_N}e^{(k_1+k_N)x} \\ \frac{c_2^2(t)}{k_2+k_1}e^{(k_2+k_1)x} & 1+\frac{c_2^2(t)}{2k_2}e^{2k_2x} & & \frac{c_2^2(t)}{k_2+k_j}e^{(k_2+k_j)x} & \dots & \frac{c_2^2(t)}{k_2+k_N}e^{(k_2+k_N)x} \\ \vdots & & & \ddots & & \\ \frac{c_N^2(t)}{k_N+k_1}e^{(k_N+k_1)x} & \frac{c_N^2(t)}{k_N+k_2}e^{(k_N+k_2)x} & \dots & \frac{c_N^2(t)}{k_N+k_j}e^{(k_N+k_j)x} & & 1+\frac{c_N^2(t)}{2k_N}e^{2k_Nx} \end{vmatrix}$$

Applying the previous remark to the determinant $\det P$ above, we obtain

$$\det P = \begin{vmatrix} 1+\frac{c_1^2(t)}{2k_1}e^{2k_1x} & \frac{c_1(t)c_2(t)}{k_1+k_2}e^{(k_1+k_2)x} & \dots & \frac{c_1(t)c_N(t)}{k_1+k_N}e^{(k_1+k_N)x} \\ \frac{c_2(t)c_1(t)}{k_2+k_1}e^{(k_2+k_1)x} & 1+\frac{c_2^2(t)}{2k_2}e^{2k_2x} & \dots & \frac{c_2(t)c_N(t)}{k_2+k_N}e^{(k_2+k_N)x} \\ \vdots & & \ddots & \\ \frac{c_N(t)c_1(t)}{k_N+k_1}e^{(k_N+k_1)x} & \frac{c_N(t)c_2(t)}{k_N+k_2}e^{(k_N+k_2)x} & \dots & 1+\frac{c_N^2(t)}{2k_N}e^{2k_Nx} \end{vmatrix}$$

Since $c_j(t) = c_j(0)e^{-4k_j^3t}$, then

$$\det P = \begin{vmatrix} 1+\frac{c_1^2(0)}{2k_1}e^{2k_1\xi_1} & \frac{c_1(0)c_2(0)}{k_1+k_2}e^{k_1\xi_1+k_2\xi_2} & \dots & \frac{c_1(0)c_N(0)}{k_1+k_N}e^{k_1\xi_1+k_N\xi_n} \\ \frac{c_2(0)c_1(0)}{k_2+k_1}e^{k_2\xi_2+k_1\xi_1} & 1+\frac{c_2^2(0)}{2k_2}e^{2k_2\xi_2} & \dots & \frac{c_2(0)c_N(0)}{k_2+k_N}e^{k_2\xi_2+k_N\xi_N} \\ \vdots & & \ddots & \\ \frac{c_N(0)c_1(0)}{k_N+k_1}e^{k_N\xi_N+k_1\xi_1} & \frac{c_N(0)c_2(0)}{k_N+k_2}e^{k_N\xi_N+k_2\xi_2} & \dots & 1+\frac{c_N^2(0)}{2k_N}e^{2k_N\xi_N} \end{vmatrix}$$

where $\xi_n \equiv x - 4k_j^2 t$, $1 \le j \le N$. To get an idea of the behavior of solitons and in particular their asymptotic behavior, suppose that $k_1 < k_2 < \dots < k_{N-1} < k_N$. For $t \gg 0$ let's write ξ_j in the form $\xi_j \equiv \xi_n - \varepsilon_{jn}t$, $1 \le j \le N$ with $\varepsilon_{jn} \equiv 4k_j^2 - 4k_n^2$ and $c_j(0) \equiv c_j$. Note that $\varepsilon_{jn} < 0$ if $1 \le j < n$, $\varepsilon_{nn} = 0$, $\varepsilon_{jn} > 0$ if $n < j \le N$, and $\varepsilon_{jn} = -\varepsilon_{nj}$. We have also $\varepsilon_{nm} > \varepsilon_{(n-1)m} > \dots > \varepsilon_{(m+1)n} > 0$ if $n > m$, and $\varepsilon_{nm} < \varepsilon_{n(m-1)} < \dots < \varepsilon_{n(m+1)} < 0$ if $n < m$. Replace these expressions in the determinant above and approximate the elements of the diagonal (for $j < n$) :

$$1 + \frac{c_j^2}{2k_j}e^{2k_j(\xi_n - \varepsilon_{jn}t)} \cong \frac{c_j^2}{2k_j}e^{2k_j(\xi_n - \varepsilon_{jn}t)}, \quad j < n, \quad t \to \infty$$

(we can do it because for $j < n$, we have $\varepsilon_{jn} < 0$ and $1 + e^x \cong e^x$ for $x \to \infty$). Then, we put in factor the following common expressions : $e^{2k_1(\xi_n - \varepsilon_{1n}t)}$, $e^{2k_2(\xi_n - \varepsilon_{2n}t)}$,...,$e^{2k_{n-1}(\xi_n - \varepsilon_{(n-1)n}t)}$. By turning t to infinity, we have (since $\varepsilon_{jn} > 0$ for $n \le j \le N$) the following situation :

$$\det P = C \begin{vmatrix} \frac{c_1^2}{2k_1} & \frac{c_1c_2}{k_1+k_2} & \cdots & \frac{c_1c_{n-1}}{k_1+k_{n-1}} & \frac{c_1c_n}{k_1+k_n}e^{k_n\xi_n} & 0 & \cdots & 0 \\ \frac{c_2c_1}{k_2+k_1} & \frac{c_2^2}{2k_2} & \cdots & \frac{c_2c_{n-1}}{k_2+k_{n-1}} & \frac{c_2c_n}{k_2+k_n}e^{k_n\xi_n} & 0 & \cdots & 0 \\ \vdots & & \ddots & & \vdots & \vdots & \cdots & \vdots \\ \frac{c_{n-1}c_1}{k_{n-1}+k_1} & \frac{c_{n-1}c_2}{k_{n-1}+k_2} & \cdots & \frac{c_{n-1}^2}{2k_{n-1}} & \frac{c_{n-1}c_n}{k_{n-1}+k_n}e^{k_n\xi_n} & 0 & \cdots & 0 \\ \frac{c_nc_1}{k_n+k_1}e^{k_n\xi_n} & \frac{c_nc_2}{k_n+k_2}e^{k_n\xi_n} & \cdots & \frac{c_nc_{n-1}}{k_n+k_{n-1}}e^{k_n\xi_n} & 1+\frac{c_n^2}{2k_n}e^{2k_n\xi_n} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{vmatrix}$$

where $C \equiv \prod_{j=1}^{n-1} e^{2k_j(\xi_n - \varepsilon_{jn})}$. Obviously, we have

$$\det P = C \begin{vmatrix} \frac{c_1^2}{2k_1} & \frac{c_1c_2}{k_1+k_2} & \cdots & \frac{c_1c_{n-1}}{k_1+k_{n-1}} & \frac{c_1c_n}{k_1+k_n}e^{k_n\xi_n} \\ \frac{c_2c_1}{k_2+k_1} & \frac{c_2^2}{2k_2} & \cdots & \frac{c_2c_{n-1}}{k_2+k_{n-1}} & \frac{c_2c_n}{k_2+k_n}e^{k_n\xi_n} \\ \vdots & & \ddots & & \vdots \\ \frac{c_{n-1}c_1}{k_{n-1}+k_1} & \frac{c_{n-1}c_2}{k_{n-1}+k_2} & \cdots & \frac{c_{n-1}^2}{2k_{n-1}} & \frac{c_{n-1}c_n}{k_{n-1}+k_n}e^{k_n\xi_n} \\ \frac{c_nc_1}{k_n+k_1}e^{k_n\xi_n} & \frac{c_nc_2}{k_n+k_2}e^{k_n\xi_n} & \cdots & \frac{c_nc_{n-1}}{k_n+k_{n-1}}e^{k_n\xi_n} & 1+\frac{c_n^2}{2k_n}e^{2k_n\xi_n} \end{vmatrix}.$$

This determinant is still written in the form

$$\det P = C\prod_{l=1}^{n-1} c_l^2 \begin{vmatrix} \frac{1}{2k_1} & \frac{1}{k_1+k_2} & \cdots & \frac{1}{k_1+k_{n-1}} \\ \frac{1}{k_2+k_1} & \frac{1}{2k_2} & \cdots & \frac{1}{k_2+k_{n-1}} \\ \vdots & & \ddots & \vdots \\ \frac{1}{k_{n-1}+k_1} & \frac{1}{k_{n-1}+k_2} & \cdots & \frac{1}{2k_{n-1}} \end{vmatrix} + C\prod_{l=1}^{n} c_l^2 \begin{vmatrix} \frac{1}{2k_1} & \frac{1}{k_1+k_2} & \cdots & \frac{1}{k_1+k_{n-1}} & \frac{1}{k_1+k_n} \\ \frac{1}{k_2+k_1} & \frac{1}{2k_2} & \cdots & \frac{1}{k_2+k_{n-1}} & \frac{1}{k_2+k_n} \\ \vdots & & \ddots & & \vdots \\ \frac{1}{k_{n-1}+k_1} & \frac{1}{k_{n-1}+k_2} & \cdots & \frac{1}{2k_{n-1}} & \frac{1}{k_{n-1}+k_n} \\ \frac{1}{k_n+k_1} & \frac{1}{k_n+k_2} & \cdots & \frac{1}{k_n+k_{n-1}} & \frac{1}{2k_n} \end{vmatrix},$$

for $n \geq 2$, while for $n=1$, it equals to $1+\frac{c_1^2}{2k_1}e^{2k_1\xi_1}$. Using the previous lemma, we get for $t \gg 0$,

$$\det P = \prod_{i=1}^{n-1} e^{2k_i(\xi_n - \varepsilon_{in} t)} \left(\prod_{j=1}^{n-1} c_j^2 \frac{(\prod_{i<j}(k_i-k_j))^2}{\prod_{i,j}(k_i+k_j)} + \prod_{j=1}^{n-1} c_j^2 \frac{(\prod_{i<j}(k_i-k_j))^2}{\prod_{i,j}(k_i+k_j)} e^{2k_n\xi_n} \right).$$

By replacing this expression in the solution obtained in the last theorem, we obtain the explicit solution of N-solitons :

$$u(x,t) = -2\sum_{n=1}^{N} k_n^2 \operatorname{sech}^2 \left(k_n\xi_n + \frac{1}{2}\ln \frac{c_n^2}{2k_n} \left(\prod_{j=1}^{n-1} \frac{k_j-k_n}{k_j+k_n} \right)^2 \right), \quad t \to +\infty.$$

Similarly, it is shown that for $t \ll 0$

$$u(x,t) = -2\sum_{n=1}^{N} k_n^2 \operatorname{sech}^2 \left(k_n\xi_n + \frac{1}{2}\ln \frac{c_n^2}{2k_n} \left(\prod_{j=n+1}^{N} \frac{k_j-k_n}{k_j+k_n} \right)^2 \right), \quad t \to -\infty.$$

This completes the demonstration. □

This result can be interpreted as follows: for example for $t \to \infty$, we have

$$\lim_{t\to\infty} u(x,t) = \lim_{t\to\infty} u(x-ct) = \begin{cases} -2k_n^2 \operatorname{sech}^2\left(k_n(x-4k_n^2t)+\delta_n^+\right) & \text{if } c = 4k_n^2 \\ 0 & \text{if } c \neq 4k_n^2 \end{cases}$$

This is the form of a solitary wave of amplitude $2k_n^2$, propagating on the right with a constant velocity equal to $4k_n^2$. The solution of the KdV equation actually splits into N-solitons at the limit for $|t| \to \infty$. This indicates that each soliton preserves its shape after collisions. These are analyzed by the phase changes δ_n^+ and δ_n^-. The relative phase change is determined by

$$\delta_n^+ - \delta_n^- = \frac{1}{2}\ln\frac{c_n^2}{2k_n}\left(\prod_{j=1}^{n-1}\frac{k_j-k_n}{k_j+k_n}\right)^2 - \frac{1}{2}\ln\frac{c_n^2}{2k_n}\left(\prod_{j=n+1}^{N}\frac{k_j-k_n}{k_j+k_n}\right)^2 = \sum_{j=1}^{n-1}\ln\frac{k_j-k_n}{k_j+k_n} - \sum_{j=n+1}^{N}\ln\frac{k_j-k_n}{k_j+k_n},$$

and it is expressed in terms of k_j $(1 \le j \le N)$. Since k_j are invariant with respect to time, then the $\delta_n^+ - \delta_n^-$ are also invariant. Recall that we assumed $k_1 < k_2 < \ldots < k_N$, then

$$\delta_1^+ - \delta_1^- = -\sum_{j=2}^{N}\ln\frac{k_j-k_1}{k_j+k_1} > 0, \qquad \delta_N^+ - \delta_N^- = \sum_{j=1}^{N-1}\ln\frac{k_N-k_1}{k_N+k_1} < 0.$$

In addition, it is easy to show that $\sum_{n=1}^{N}\delta_n^+ = \sum_{n=1}^{N}\delta_n^-$.

We could not finish this section without indicating some results related to the KdV equation. The KdV equation (1.1) is written in the form

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(3u^2 - \frac{\partial^2 u}{\partial x^2}\right) = \frac{\partial}{\partial x}\frac{\delta H}{\delta u},$$

where

$$H = \int_{-\infty}^{\infty} P_3 dx = \int_{-\infty}^{\infty}\left(u^3 + \frac{1}{2}\left(\frac{\partial u}{\partial x}\right)^2\right)dx,$$

is the first integral (Hamiltonian) obtained previously and $\frac{\delta H}{\delta u(x)}$ denotes the gradient (Fréchet derivative) of the function H. This equation forms an infinite dimensional Hamiltonian system, completely integrable and the Hamiltonian structure is defined by the Poisson bracket : $\{F,H\} = \int \frac{\delta F}{\delta u(x)}\frac{\partial}{\partial x}\frac{\delta H}{\delta u(x)}dx$. We check that the latter satisfies the Jacobi identity. We will discuss further (in the following sections) the problem of studying the KdV equation via symplectic structures on operator algebra, the relation with the KP hierarchy [16], the Sato theory [17] τ functions and the work of Jimbo-Miwa-Kashiwara [18, 19]. The study of the periodic problem for the KdV equation allowed some authors to discover an interesting class of completely integrable systems. The obtained solutions are endowed with remarkable properties : they define functions $u(x)$ for which equation (2.1) with periodic coefficients has a finite number of zones of parametric resonance on the axis λ. The spectrum of the Schrödinger operator is invariant by the Hamiltonian flow defined by the KdV equation. And as we have already pointed out, this spectrum provides an infinity first integrals or invariants. The isospectral sets related to invariant manifolds defined by putting these invariants equal to generic constants are compact, connected, and infinite-dimensional tori. Each of these isospectral sets is isomorphic to the real part of a Jacobi variety associated with a hyperelliptic curve of finite or infinite genus. The periods of this torus can be expressed using hyperelliptic integrals; in short, the explicit linearization of the flow of the KdV equation is made on this Jacobian variety using the Abel application, the Jacobi inversion problem and the theta functions. For other interesting integrable systems that will not be discussed here, the solutions blow up after a finite time as Laurent series depending on many parameters (see for example [20, 21]).

4. Pseudo-differential operators

Let L be a pseudo-differential operator with holomorphic coefficients. The set of these operators form a Lie algebra that we note $\mathscr{A}$. The algebra $\mathscr{A}$ decomposes in two sub-algebras $\mathscr{A}_+$ and $\mathscr{A}_-$: $\mathscr{A} = \mathscr{A}_+ \oplus \mathscr{A}_-$, where $\mathscr{A}_+$ is the algebra of differential operators of the form $\zeta = \sum_{k\ge 0} u_k(x)\partial^k$, finite sum, $\partial = \frac{\partial}{\partial x}$, and $\mathscr{A}_-$ is the algebra of strictly pseudo-differential operators of the form

$$\eta = \sum_{k>0} u_{-k}(x)\partial^{-k} = \partial^{-1}v_0 + \partial^{-2}v_1 + \cdots, \quad \partial = \frac{\partial}{\partial x},$$

The algebra $\mathscr{A}$ is an associative algebra for the product of two pseudo-differential operators L and L',

$$L.L' = \sum_{k=0}^{\infty} \frac{1}{k!} {:}\partial_{\partial}^{k}(L).\partial_{x}^{k}(L'){:},$$

where $\partial = \dfrac{\partial}{\partial x}$ and the symbol **::** denotes the normal order, i.e., it means that the derivatives always appear on the right independently of the commutation relations.

For $m,n \in \mathbb{N}^*$, we have for all functions u, v,

$$u\partial^m.v\partial^n = \sum_{k=0}^{m} \frac{m!}{k!(m-k)!} uv^{(k)}\partial^{m+n-k} = \sum_{k=0}^{m} \frac{1}{k!} {:}\partial_{\partial}^{k}(u\partial^m).\partial_{x}^{k}(v\partial^n){:} \tag{4.1}$$

and

$$\partial^{-1}u = u\partial^{-1} - u'\partial^{-2} + u''\partial^{-3} + \cdots = \sum_{k=0}^{\infty} {:}\partial_{\partial}^{k}(\partial^{-1}).\partial_{x}^{k}(u){:} \tag{4.2}$$

where ∂^{-1} is a formal inverse of ∂, i.e., $\partial^{-1}.\partial = \partial.\partial^{-1} = 1$.

We define a coupling between $\mathscr{A}_+$ and $\mathscr{A}_-$ as follows : Let $\operatorname{Res}(\zeta\eta)$ be the coefficient of ∂^{-1} in $\zeta\eta$. We have

$$\langle \zeta, \eta \rangle = \left\langle \sum_{k\geq 0} u_k\partial^k, \sum_{k>0} u_{-k}\partial^{-k} \right\rangle = \left\langle u_0\partial^0 + u_1\partial^1 + \cdots, \partial^{-1}v_0 + \partial^{-2}v_1 + \cdots \right\rangle$$

i.e., $\langle \zeta, \eta \rangle = \int_{-\infty}^{\infty} \operatorname{Res}(\zeta\eta)dx = \int_{-\infty}^{\infty} \sum_{k\geq 0} u_k v_k dx$. Therefore, the Volterra group $(I + \mathscr{A}_-)$ acts on $\mathscr{A}_-$ by the adjoint action and on $\mathscr{A}_+$ by the coadjoint action. Let $\zeta \in \mathscr{A}_+$ and $\eta_k \in \mathscr{A}_-$. We obtain from [22],

$$\begin{aligned}
\langle ad^*_{\eta_1}(\zeta), \eta_2 \rangle = \langle \zeta, ad_{\eta_1}(\eta_2) \rangle &= \langle \zeta, [\eta_1, \eta_2] \rangle, \\
&= \int \left(\partial^{-1} - \text{term of } (\zeta\eta_1\eta_2 - \zeta\eta_2\eta_1)\right) dx, \\
&= \int \left(\partial^{-1} - \text{term of } \ (\zeta\eta_1 - \eta_1\zeta)_+ \eta_2\right) dx, \\
&= \langle [\zeta, \eta_1]_+, \eta_2 \rangle.
\end{aligned}$$

So the set $\mathscr{O}^*_{\mathscr{A}_+}(L)$ of the differential operators of the form

$$L = \partial^N + \sum_{k=0}^{N-2} u_k(x)\partial^k, \quad N \text{ fixed}, \tag{4.3}$$

is a coadjoint orbit in $\mathscr{A}_+$.

Let f be a function of class $\mathscr{C}^\infty$ in x and dependent on a finite number of derivatives $u_k^{(l)}$ of the coefficients u_k of L. Let

$$\nabla H(L) = \sum_{k=0}^{N-1} \partial^{-k-1} \sum_{l} (-1)^l \left(\frac{d}{dx}\right)^l \frac{\partial f}{\partial p_k^{(l)}} = \sum_{k=0}^{N-1} \partial^{-k-1} \frac{\delta H}{\delta u_k},$$

be the gradient of the functional, $H(L) = \int_{-\infty}^{\infty} f(x, ..., u_k^{(l)}, ...)dx$ defined on $\mathscr{A}_+$ and such that :

$$dH = \int_{-\infty}^{\infty} \frac{\delta H}{\delta u_k} du_k = \left\langle \sum_{k=0}^{N} du_k.\partial^k, \nabla H \right\rangle = \langle dL, \nabla H \rangle,$$

where $dL = \sum_{k=0}^{N} du_k.\partial^k$. We recall that the scalar product between two pseudo-differential operators L and L' is defined by $\langle L, L' \rangle = -\int_{-\infty}^{\infty} (LL')_- dx = \int_{-\infty}^{\infty} (L'L)_- dx)$. According to the Adler-Kostant-Symes [20, 22, 23, 24, 25, 26, 27, 28, 29], the Hamiltonian vector fields on the coadjoint orbit $\mathscr{O}^*_{\mathscr{A}_+}$, define commutative flows and are given by

$$\frac{dL}{dt} = ad^*_{\nabla H(L)}(L) = [L, \nabla H(L)]_+, \tag{4.4}$$

where $H(L)$ is the Hamiltonian on $\mathscr{A}_+$. The operator L does not contain the coefficient u_{N-1}. Since the vector field (4.4) applied to the operator L (4.3) imposes the condition Res $[L, \nabla H(L)] = 0$, we can replace the gradient $\dfrac{\delta H}{\delta p_{N-1}}$ by any expression satisfying this condition. A first Poisson bracket is given by

$$\{H,F\}_1 = \langle L, [\nabla F, \nabla H] \rangle = \int \text{Res}\,(\nabla H[L, \nabla F]_+)\,dx = \int \text{Res}\,(\nabla H[L, \nabla F])\,dx = \int \text{Res}\,([\nabla H, L]\nabla F)\,dx. \tag{4.5}$$

Consider the Hamiltonians $H_{k+N} = \dfrac{N}{k+N}\displaystyle\int \left(\text{Rés}\, L^{\frac{k+N}{N}}\right) dx$, $k \in \mathbb{N}^*$. We have $\nabla H^{(L)}_{k+N} = \left(L^{\frac{k}{N}}\right)_-$, and the vector fields (4.4) applied to these Hamiltonians, provide the integrable equations; N-reduction of Gel'fand Dickey equations of KP hierarchy (see below for definition) :

$$\frac{dL}{dt} = [L, \nabla H_{k+N}(L)]_+ = -\left[(L^{\frac{k}{N}})_-, L\right]_+ = \left[(L^{\frac{k}{N}})_+, L\right]. \tag{4.6}$$

Note that since $\left[(L^{\frac{k}{N}})_+, L\right]_+ = \left[L^{\frac{k}{N}} - (L^{\frac{k}{N}})_-, L\right]_+ = -\left[(L^{\frac{k}{N}})_-, L\right] \in \mathscr{A}^-$, then equations (4.6) determine an infinite number of commutative vector fields (see below) on $\mathscr{A}^+ + \mathscr{A}^-$.

We will now study [22, 30, 31, 32, 33] the existence of a second symplectic structure. Let $\widetilde{L} = L + z$ where L is a differential operator of order n. We have

$$\frac{dL}{dt} = \left(\widetilde{L}\nabla H\right)_+ \widetilde{L} - \widetilde{L}\left(\nabla H \widetilde{L}\right)_+. \tag{4.7}$$

Note that (4.7) is a Hamiltonian vector field generalizing (4.4). Indeed, let $J : \mathscr{A}_- / \mathscr{A}_{-\infty,N-1} \longrightarrow \mathscr{D}_{0,N-1}$, be the function defined by

$$J(\zeta) = \left(\widetilde{L}\zeta\right)_+ \widetilde{L} - \widetilde{L}\left(\zeta\widetilde{L}\right)_+ = -\left(\widetilde{L}\zeta\right)_- \widetilde{L} + \widetilde{L}\left(\zeta\widetilde{L}\right)_-, \quad \zeta \in \mathscr{A}_- / \mathscr{A}_{-\infty,N-1}.$$

Hence, $\dfrac{dL}{dt} = \partial_{J(\zeta)}(L) \equiv \left(\widetilde{L}\zeta\right)_+ \widetilde{L} - \widetilde{L}\left(\zeta\widetilde{L}\right)_+$, which shows that it is indeed a vector field on the differential operators L of order n. Similarly, we have $\dfrac{dL}{dt} = -\left(\widetilde{L}\nabla H\right)_- \widetilde{L} + \widetilde{L}\left(\nabla H\widetilde{L}\right)_-$, and the same conclusion remains valid. We also have the relation $\dfrac{dL}{dt} = (L\nabla H)_+ L - L(\nabla HL)_+ + z[\nabla H, L]_+$, which shows that this vector field is an interpolation between (4.4) for $z = \infty$ and a new vector field for $z = 0$. Consider the 2-differential form

$$\omega\left(\partial_{J(\zeta)}, \partial_{J(\eta)}\right) = \langle J(\zeta), \eta \rangle = \int \text{Res}\ (J(\zeta)\eta)dx.$$

This form is closed ($d\omega = 0$) and is antisymmetric $\langle J(\zeta), \eta \rangle = -\langle \zeta, J(\eta) \rangle$, and furthermore $\left[\partial_{J(\zeta)}, \partial_{J(\eta)}\right] = \partial_{J(\xi)}$, where

$$\xi = \left(-\zeta\left(\widetilde{L}\eta\right)_+ + \left(\zeta\widetilde{L}\right)_- \eta\right)_- \left(-\eta\left(\widetilde{L}\zeta\right)_+ + \left(\eta\widetilde{L}\right)_- \zeta\right)_- + \partial_{J(\zeta)}\eta - \partial_{J(\eta)}\zeta.$$

The functional algebra on the operator space of the form (4.3) for this symplectic form is the so called $\mathscr{W}$ algebra.

Theorem 4.1. *The Hamiltonians $H_k, H_{k+N}, H_{k+2N}, \ldots$, defined in* (4.6) *are all in involution for the bracket* (4.5)*.*

Proof. Indeed, let $J = J_1$ if $z = \infty$ and $J = J_2$ if $z = 0$, where the Poisson brackets $\{.,.\}_1$, $\{.,.\}_2$ are given by $\{H_j, H_k\}_1 = \int \text{Res}\ (\nabla H_j J_1(\nabla H_k))$ and

$$\{H_j, H_k\}_2 = \langle \nabla H, J_2(\nabla F) \rangle = \int \text{Res}\ (\nabla H((L\nabla F)_+ L - L(\nabla FL)_+))dx = \int \text{Res}\ (L\nabla H(L\nabla F)_+ - \nabla HL(\nabla FL)_+)dx.$$

We deduce from the relation

$$\left(L\left(L^{\frac{r}{n}-1}\right)_-\right)_+ L - L\left(\left(L^{\frac{r}{n}-1}\right)_- L\right)_+ + \left[\left(L^{\frac{r}{n}}\right)_-, L\right]_+ = 0,$$

the expression $\{H_j, H_k\}_1 = \int \text{Res } (\nabla H_j J_2 \nabla H_{k-N})$. Since the form ω is anti-symmetric, we have

$$\{H_j, H_k\}_1 = -\int \text{Res } (\nabla H_{k-N} J_2(\nabla H_j)) = -\int \text{Res } (\nabla H_{k-N} J_1(\nabla H_{j+N})),$$

$$\{H_j, H_k\}_1 = -\int \text{Res } (\nabla H_{j+N} J_1(\nabla H_{k-N})),$$

$$\{H_j, H_k\}_1 = \left\{H_{j+N}, H_{k-N}\right\}_1 = \{H_j, H_k\}_1 = \left\{H_{j+\alpha N}, H_{k-\alpha N}\right\}_1,$$

for α large enough with $J_1(\nabla H_{k-\alpha N}) = 0$, i.e., $H_{k-\alpha N}$ is trivial for α large enough and we get $\{H_j, H_k\}_1 = 0$, so H_j, H_k are in involution. □

5. KdV equation, Heisenberg and Virasoro algebras

Theorem 5.1. *a) The operator $L = \partial^2 + q$, corresponding to the case $N = 2$ with $q \equiv u_0$, is related to the KdV equation and the Poisson bracket is provided in this case by $\{q(x), q(y)\}_1 = \frac{d}{dx}\delta(x-y)$.*

b) By replacing in a), $q(x)$ by the Fourier series

$$q(x) = \alpha \sum_{n=-\infty}^{\infty} e^{-inx}\varphi_n + \beta, \quad -i\alpha^{-2} = 1, \tag{5.1}$$

where $(\varphi_k)_{k\in\mathbb{Z}}$ are new coordinates (Fourier coefficients), one obtains the Heisenberg algebra and the Poisson bracket is provided by $\{\varphi_n, \varphi_m\}_1 = n\delta_{m+n,0}$.

c) In the case $N = 2$ one obtains the Virasoro algebra and its structure is given by

$$\{\varphi_m, \varphi_n\}_2 = (m-n)\varphi_{m+n} + \frac{c}{12}(m^3 - m)\delta_{m+n,0}.$$

Proof. a) Indeed, since

$$\nabla H(L) = \partial^{-1}\frac{\delta H}{\delta q} + \partial^{-2}\frac{1}{2}\left(\frac{\delta H}{\delta q}\right)',$$

then the vector fields applied to the Hamiltonian $H = \int \left(q^3 - \frac{1}{2}q'^2\right) dx$, provide the KdV equation

$$\frac{dq}{dt} = \frac{dL}{dt} = \frac{1}{2}[L, \nabla H]_+ = \frac{d}{dx}\frac{\delta H}{\delta q} = \frac{dq}{dt} = \frac{\partial^3 q}{\partial x^3} + 6q\frac{\partial q}{\partial x}. \tag{5.2}$$

The Poisson bracket is in this case is $\{H, F\}_1 = \int \frac{\delta H}{\delta q}\frac{d}{dx}\frac{\delta F}{\delta q}$, and therefore, $\{q(x), q(y)\}_1 = \frac{d}{dx}\delta(x-y)$.

b) Let $\mathcal{M}$ be the set of matrices (a_{kl}), $(k, l \in \mathbb{Z})$, with complex coefficients and

$$\mathcal{N} = \{(a_{kl}) \in \mathcal{M} : \text{there is at least } r \text{ such that } a_{kl} = 0 \text{ for } |k-l| > r\},$$

the $\mathbb{C}$-algebra, i.e., the set of infinite matrices with support in a band around the diagonal. The product of two matrices belonging respectively to $\mathcal{N}$ and $\mathcal{M}$ is defined in the usual way. Note that $\mathcal{N}$ is a Lie algebra and $\mathcal{M}$ is a $\mathcal{N}$-module. Their extensions $\widetilde{\mathcal{N}}$ and $\widetilde{\mathcal{M}}$ are defined by

$$0 \longrightarrow \mathbb{C}c \longrightarrow \widetilde{\mathcal{N}} \longrightarrow \mathcal{N} \longrightarrow 0, \qquad 0 \longrightarrow \mathbb{C}c \longrightarrow \widetilde{\mathcal{M}} \longrightarrow \mathcal{M} \longrightarrow 0,$$

with $\widetilde{\mathcal{N}} = \mathcal{N} \oplus \mathbb{C}c$, $\widetilde{\mathcal{M}} = \mathcal{M} \oplus \mathbb{C}c$, where c is a central element, i.e., we have $[c, A] = [c, B] = 0$, $\forall A \in \widetilde{\mathcal{N}}$, $\forall B \in \widetilde{\mathcal{M}}$. We notice $e_{i,j} = (\delta_{ki}.\delta_{lj})_{kl}$ the elementary matrices, i.e., the matrices whose coefficients are all zero except the one of the line i and the column j which is equal to 1. Since a Jacobi matrix has no trace, then we consider the matrix $A[J, B]$ where $A \in \mathcal{N}$, $B \in \mathcal{M}$ and J is the matrix defined by $J = \sum_{i\in\mathbb{Z}} \varepsilon(i) e_{i,i}$, where $\varepsilon(i) = +1$ if $i < 0$ and -1 if $i \geq 0$. The elements of the matrix $A[J, B]$ are null except for a finite number, so it is a finite matrix and we define the cocycle of $A \in \mathcal{N}$ and $B \in \mathcal{M}$ using the formula

$$\rho(A, B) = \frac{1}{2}Tr(A[J, B]) = \frac{1}{2}\sum_{i,j}(\varepsilon(i) - \varepsilon(j))a_{ij}b_{ji}.$$

Therefore, the bracket $\widetilde{[,]}$ of $A \in \mathscr{N}$ and $B \in \mathscr{M}$ is defined by

$$\widetilde{[A,B]} = [A+\alpha c, B+\beta c] = [A,B] + \rho(A,B)c.$$

We note that the $\widetilde{\mathscr{N}}$ algebra is a non-trivial central extension of N while the subalgebra $\widetilde{\mathscr{M}_f} = \mathscr{M}_f \oplus \mathbb{C}c$, is a trivial central extension of

$$\mathscr{M}_f = \{(a_{ij}) \in \mathscr{M} : (i,j) \longmapsto (a_{ij}) \text{ with finished support}\}.$$

Let's put $E_i = \sum_{n\in\mathbb{Z}} e_{n,n+i}$, where $e_{i,i} = (\delta_{ki}.\delta_{ij})_{kl}$ are the elementary matrices defined above. The subspace $E = \bigoplus_{i\in\mathbb{Z}} \mathbb{C}E_i$, is a commutative subalgebra of $\mathscr{N}$. The subalgebra of $\mathscr{N}$ defined by setting $\widetilde{E} = E \oplus \mathbb{C}c$ is called Heisenberg subalgebra. We have

$$\widetilde{[E_i, E_j]} = i\delta_{i,-j}c. \tag{5.3}$$

We now reconsider the previous example and replace $q(x)$ with the Fourier series (5.1). Let H be a functional of q. Its Fréchet derivative in terms of the coordinates φ_k is written

$$\frac{\delta H}{\delta q} = \sum_{k=-\infty}^{\infty} \frac{\delta H}{\delta \varphi_k}.\frac{\partial \varphi_k}{\partial q} = \alpha^{-1} \sum_{k=-\infty}^{\infty} \frac{\delta H}{\delta \varphi_k} e^{ikx}. \tag{5.4}$$

We substitute (5.3) and (5.1) in equation (5.2) and we specify the Fourier coefficients; we get the relation $\alpha \frac{\partial \varphi_n}{\partial t} = -i\alpha^{-1} n \frac{\partial H}{\partial \varphi_n}$. Moreover, since the symplectic structure is given by the matrix of the Poisson brackets, we also have

$$\frac{\partial \varphi_n}{\partial t} = \sum_{m=-\infty}^{\infty} \{\varphi_n, \varphi_m\}_1 \frac{\partial H}{\partial \varphi_m}.$$

Therefore, $\{\varphi_n, \varphi_m\}_1 = -i\alpha^{-2} n \delta_{m+n,0}$. By putting $-i\alpha^{-2} = 1$, we obtain the Heisenberg algebra (where $\{,\}$ plays the role here of the bracket $\widetilde{[,]}$ (5.3) above).

c) Let $\mathrm{Diff}(S^1)$ be the group of diffeomorphisms of the unit circle : $S^1 = \{z \in \mathbb{C} : |z| = 1\}$. Let

$$F = \left\{ f(z)\frac{d}{dz} : f(z) \in \mathbb{C}[z, \frac{1}{z}] \right\},$$

be the set of vector fields (Laurent's polynomials). Note that F can be seen as the tangent space $\mathrm{Diff}(S^1)$ at its unit point, so F is a Lie algebra with respect to the bracket $[,]$. By setting $\varphi_m = -z^{m+1}\frac{d}{dz}$, we obtain

$$[\varphi_m, \varphi_n] = \left((n+1)z^{m+n+1} - (m+1)^{m+n+1}\right)\frac{d}{dz} = -(m+n)z^{m+n+1}\frac{d}{dz},$$

i.e., $[\varphi_m, \varphi_n] = (m-n)\varphi_{m+n}$. We show that $H^2(F,\mathbb{C}) \cong \mathbb{C}$ and $\rho(\varphi_m, \varphi_n) = \frac{1}{12}(m^3 - m)\delta_{m,-n}$. The vector space $F \oplus \mathbb{C}c$ is called Virasoro algebra, it is a central extension of the algebra of complex vector fields on the circle. The bracket is given by the formula

$$[\varphi_m, \varphi_n] = (m-n)\varphi_{m+n} + \frac{c}{12}(m^3 - m)\delta_{m,-n}. \tag{5.5}$$

Let us now consider the example of the KdV equation. We have $N = 2$ and

$$\frac{dq}{dt} = \frac{dL}{dt} = (L\nabla H)_+ - L(\nabla HL)_+ = \left(\partial^3 + 2(\partial q + q\partial)\right)\frac{\delta H}{\delta q}.$$

The (Poisson) bracket is written in this case

$$\{H, F\}_2 = \int \frac{\delta H}{\delta q}\left(\partial^3 + 2(\partial q + q\partial)\right)\frac{\delta F}{\delta q},$$

and we have $\{q(x),q(y)\}_2 = \left(\partial^3+2(\partial q+q\partial)\right)\delta(x-y)$. By reasoning as before (while taking into account the Fréchet derivative (5.4)), we obtain

$$\alpha\frac{\partial\varphi_m}{\partial t} = i\sum_n (n-m)\varphi_{m+n}\frac{\delta H}{\delta\varphi_n}+\frac{i}{2\alpha}(m^3-4\beta m)\frac{\delta H}{\delta\varphi_{-m}},$$

where $(\varphi_k)_{k\in\mathbb{Z}}$ are the Fourier coefficients of q. By setting $4\beta=1$, $\alpha=\frac{6i}{c}$ and taking into account the Fourier series (5.1), we obtain

$$\frac{\partial}{\partial t}\begin{pmatrix}\vdots\\ \varphi_m\\ \vdots\end{pmatrix} = \begin{pmatrix} & & -\text{nth column} & & \text{nth column}\\ & & \downarrow & & \downarrow\\ \text{mth line} & \longrightarrow & \frac{c}{12}(m^3-m) & \dots & (m-n)\varphi_{m+n}\end{pmatrix}\begin{pmatrix}\vdots\\ \frac{\delta H}{\delta\varphi_m}\\ \vdots\end{pmatrix}.$$

Consequently, we have $\{\varphi_m,\varphi_n\}_2=(m-n)\varphi_{m+n}+\frac{c}{12}(m^3-m)\delta_{m+n,0}$, i.e., the Virasoro structure [19] (where $\{,\}_2$ plays the role here of the bracket $[,]$ (5.5) above). This establishes the theorem. □

For $N=3$, $u\equiv u_2$, $v\equiv u_3$, $L=\partial^3+u\partial+v$, and $L^{\frac{2}{3}}=\partial^3+\frac{2}{3}u$, the flow (5.2) takes the form

$$\frac{\partial u}{\partial t_2}=-\frac{\partial^2 u}{\partial x^2}+2\frac{\partial v}{\partial x},\qquad \frac{\partial v}{\partial t_2}=\frac{\partial^2 v}{\partial x^2}-\frac{2}{3}\frac{\partial^3 u}{\partial x^3}-\frac{2}{3}u\frac{\partial u}{\partial x}.$$

Eliminating v from these equations yields the Boussinesq equation

$$3\left(\frac{\partial u}{\partial t_2}\right)^2+\frac{\partial^2}{\partial x^2}\left(\frac{\partial^2 u}{\partial x^2}+2u^2\right)=0.$$

6. KP hierarchy and vertex operators

Consider the pseudo-differential operator of infinite order

$$L=\partial+u_1\partial^{-1}+u_2\partial^{-2}+\dots,\qquad \partial\equiv\frac{\partial}{\partial x} \tag{6.1}$$

where $u_1,u_2,\dots$ are functions of class $\mathscr{C}^\infty$ depending on an infinity of independent variables $x\equiv t_1,t_2,\dots$. The compound operator L^n is calculated according to the rules (4.1) and (4.2). We obtain

$$L^n=\partial^n+p_{n,2}\partial^{n-2}+\dots+p_{n,n}+p_{n,n+1}\partial^{-1}+\dots=\partial^n+\sum_{j=2}^{n}p_{n,j}\partial^{n-j}+\sum_{j=1}^{\infty}p_{n,n+j}\partial^{-j},$$

where $p_{n,j}$ are polynomials in u_j and their derivatives in relation to x. The differential part L_+^n of L^n being equal to $L_+^n=\partial^n+\sum_{j=2}^{n}p_{n,j}\partial^{n-j}$, we have

$$L_+^1=\partial,\qquad L_+^2=\partial^2+2u_2,\qquad L_+^3=\partial^3+3u_2\partial+3(u_3+\partial u_2),\dots \tag{6.2}$$

The dependency between the functions $u_1,u_2,\dots$ and the variables $x=t_1,t_2,\dots$ is provided by the following system of partial differential equations :

$$\frac{\partial L}{\partial t_n}=[L_+^n,L],\qquad n\in\mathbb{N}^* \tag{6.3}$$

The set of these equations is called Kadomtsev-Petviashvili hierarchy (abbreviated KP hierarchy). It is a hierarchy of isospectral deformations of the pseudo-differential operator (6.1). We prove (see [34]) the following result :

Theorem 6.1. *There is an equivalence between (6.3) and the equations*

$$\frac{\partial}{\partial t_n}L_+^m - \frac{\partial}{\partial t_m}L_+^n = [L_+^n, L_+^m], \tag{6.4}$$

as well as their dual forms

$$\frac{\partial}{\partial t_n}L_-^m - \frac{\partial}{\partial t_m}L_-^n = -[L_-^n, L_-^m], \tag{6.5}$$

where $L_-^n = L^n - L_+^n$. Equations (6.3) determine an infinite number of commutative vector fields on algebra $\mathscr{A} = \mathscr{A}_+ \oplus \mathscr{A}_-$.

Proof. Note that since $L^n = L_+^n + L_-^n$, then $\dfrac{\partial L}{\partial t_n} = [L_+^n, L] = -[L_-^n, L] \in \mathscr{A}_-$. Equation (6.3) defines an infinite number of vector fields on $\mathscr{A}$. Since $\dfrac{\partial}{\partial t_n}$ and $[L_+^n, .]$ are derivations, then

$$\begin{aligned}
\frac{\partial L^m}{\partial t_n} &= [L_+^n, L_+^m] + [L_+^n, L_-^m],\\
&= -[L_-^n, L_+^m] - [L_-^n, L_-^m],\\
&= \frac{1}{2}\left([L_+^n, L_+^m] - [L_-^n, L_+^m]\right) + \frac{1}{2}\left(-[L_-^n, L_-^m] + [L_+^n, L_-^m]\right),\\
&= \frac{1}{2}\left([L_+^n, L_+^m] - [L_-^n, L_-^m]\right) + \frac{1}{2}\left([L_+^m, L_-^n] - [L_-^m, L_+^n]\right).
\end{aligned}$$

Similarly, we have (just swap n and m)

$$\frac{\partial L^n}{\partial t_m} = \frac{1}{2}\left([L_+^m, L_+^n] - [L_-^m, L_-^n]\right) + \frac{1}{2}\left([L_+^n, L_-^m] - [L_-^n, L_+^m]\right).$$

Hence, $\dfrac{\partial L^m}{\partial t_n} - \dfrac{\partial L^n}{\partial t_m} = [L_+^n, L_+^m] - [L_-^n, L_-^m]$. Or

$$\frac{\partial L^m}{\partial t_n} - \frac{\partial L^n}{\partial t_m} = \frac{\partial}{\partial t_n}L_+^m + \frac{\partial}{\partial t_n}L_-^m - \frac{\partial}{\partial t_m}L_+^n - \frac{\partial}{\partial t_m}L_-^n = \frac{\partial}{\partial t_n}L_+^m - \frac{\partial}{\partial t_m}L_+^n + \frac{\partial}{\partial t_n}L_-^m - \frac{\partial}{\partial t_m}L_-^n,$$

then

$$\frac{\partial}{\partial t_n}L_+^m - \frac{\partial}{\partial t_m}L_+^n - [L_+^n, L_+^m] = -\frac{\partial}{\partial t_n}L_-^m + \frac{\partial}{\partial t_m}L_-^n - [L_-^n, L_-^m].$$

Since the expression on the left belongs to $\mathscr{A}_+$ and the one on the right belongs to $\mathscr{A}_-$, then the result comes from the decomposition $\mathscr{A} = \mathscr{A}_+ \oplus \mathscr{A}_-$ since obviously $\mathscr{A}_+ \cap \mathscr{A}_- = \emptyset$. To show that the vector fields defined by these equations commute, we put $X(L) = [L_+^m, L]$ and $Y(L) = [L_+^n, L]$. Hence,

$$\begin{aligned}
[X,Y](L) &= (XY - YX)(L),\\
&= X\left([L_+^n, L]\right) - Y\left([L_+^m, L]\right),\\
&= \left[X(L_+^n) - Y(L_+^m), L\right] + \left[L_+^n, X(L)\right] - \left[L_+^m, Y(L)\right],\\
&= \left[X(L_+^n) - Y(L_+^m), L\right] + \left[L_+^n, [L_+^m, L]\right] - \left[L_+^m, [L_+^n, L]\right],\\
&= \left[X(L_+^n) - Y(L_+^m) - [L_+^m, L_+^n], L\right],
\end{aligned}$$

according to Jacobi's identity and taking into account (6.4), we deduce that the vector fields in question commute. □

By specifying the quantifiers of ∂^k in (6.4), one obtains an infinity of nonlinear partial differential equations [5] forming the Kadomtsev-Petviashvili hierarchy. These equations connect infinitely many functions u_j to infinitely many variables t_j. For example, for $m = 2$, $n = 3$, relations (6.4) and (6.2) determine two expressions based on u_2 and u_3. After eliminating u_3, we immediately obtain the Kadomtsev-Petviashvili equation (KP equation) :

$$3\frac{\partial^2 u_2}{\partial t_2^2} - \frac{\partial}{\partial t_1}\left(4\frac{\partial u_2}{\partial t_3} - 12u_2\frac{\partial u_2}{\partial t_1} - \frac{\partial^3 u_2}{\partial t_1^3}\right) = 0. \tag{6.6}$$

We can obtain particular solutions of this equation by solving the equations :

$$\frac{\partial u_2}{\partial t_2}=0, \qquad 4\frac{\partial u_2}{\partial t_3}-12u_2\frac{\partial u_2}{\partial t_1}-\frac{\partial^3 u_2}{\partial t_1^3}=0.$$

The second equation is precisely the KdV equation. The KP equation is therefore a generalization of the KdV equation, to which it is reduced when $\frac{\partial u_2}{\partial t_2}=0$.

Equations (6.3) and (6.4) (see also (6.5)) imply the existence of the following pseudo-differential operator of degree 0 (wave operator) $W\in\mathscr{I}+\mathscr{A}_-$:

$$W=1+w_1(t)\partial^{-1}+w_2(t)\partial^{-2}+\cdots \tag{6.7}$$

with $t=(t_1,t_2,...)\in\mathbb{C}^\infty$. The inverse W^{-1} of W is also a pseudo-differential operator of the form

$$W^{-1}=1+v_1(t)\partial^{-1}+v_2(t)\partial^{-2}+\cdots,$$

and can be calculated term by term. Indeed, by definition, we have $WW^{-1}=1$. So, using the fact that

$$\partial^m u=\sum_{k=0}^{\infty}\frac{m!}{k!(m-k)!}(\partial^k u\partial^u)\partial^{m-k}u, \qquad \partial^m\partial^n=\partial^{m+n},$$

as well as the formulas described above, we specify the quantifiers of ∂^{-1}, ∂^{-2},... in the equation $WW^{-1}=1$ and we determine relations between w_m et v_m. We obtain finally for W^{-1} the following expression :

$$W^{-1}=1-w_1\partial^{-1}+(-w_2+w_1^2)\partial^{-2}+(w_3+2w_1w_2-w_1\partial w_1-w_1^3)\partial^{-3}+\cdots$$

In terms of W, the operator L (6.1) can be written in the form

$$L=W.\partial.W^{-1}. \tag{6.8}$$

According to (6.1) and ((6.7), we deduce the relations :

$$u_2=\partial w_2, \qquad u_3=-\partial w_2-w_1\partial w_1, \qquad u_4=-\partial w_3+w_1\partial w_2+(\partial w_1)w_2-w_1^2\partial w_1-(\partial w_1)^2.$$

We have the following result :

Theorem 6.2. *Equations (6.3) or what amounts to the same equations (6.4)) are equivalent to the existence of the wave operator W (6.7) such that the system of differential equations*

$$LW=W\partial, \tag{6.9}$$

$$\frac{\partial W}{\partial t_n}=-L_-^nW, \tag{6.10}$$

has a solution (which can be inductively obtained).

Theorem 6.3. *a) Let $\xi(t,z)=\sum_{j=1}^{\infty}t_jz^j$, $z\in\mathbb{C}$ be the phase function with $\partial^m\xi(t,z)=z^m$ and $\partial^m e^{\xi(t,z)}=z^m e^{\xi(t,z)}$. There is an equivalence between (6.6), (6.10) and the following problem : there is a wave function ψ (Baker-Akhiezer function)*

$$\Psi(t,z)=\left(1+w_1(t)z^{-1}+w_2(t)z^{-2}+\cdots\right)e^{\xi(t,z)}=We^{\xi(t,z)}, \quad z\in\mathbb{C} \tag{6.11}$$

where W is identified as (6.7) and such that :

$$L\Psi=z\Psi, \qquad \frac{\partial\Psi}{\partial t_n}=L_+^n\Psi. \tag{6.12}$$

b) Introduce the conjugation $\partial^*=-\partial$ and let

$$L^*=1+(-\partial)^{-1}u_1+(-\partial)^{-2}u_2+\cdots, \quad W^*=1+(-\partial)^{-1}w_1+(-\partial)^{-2}w_2+\cdots$$

be the adjoints of L and W such that : $L^*=-(W^*)^{-1}.\partial.W^*$. The adjoint wave function

$$\Psi^*(t,z)=(W^*(t,\partial))^{-1}e^{-\xi(t,z)},$$

satisfies the following relations : $L^*\Psi^*=z\Psi^*$, $\frac{\partial\Psi^*}{\partial t_n}=-(L_+^n)^*\Psi^*$.

Proof. a) Indeed, we have from (6.11),

$$\begin{aligned}\frac{\partial\Psi}{\partial t_n} &= \frac{\partial W}{\partial t_n}e^{\xi(t,z)}+Wz^ne^{\xi(t,z)},\\ &= -L_-^nWe^{\xi(t,z)}+z^nWe^{\xi(t,z)}, \quad \text{according to (6.8)}\\ &= -L_-^n\Psi+z^n\Psi, \quad \text{according to (6.9)}\\ &= -L_-^n\Psi+L^n\Psi, \quad \text{according to (6.10)}\\ &= L_+^n\Psi.\end{aligned}$$

In other words, Ψ satisfies (6.11) and (6.12) is equivalent to the fact that W satisfies (6.7) and (6.10).

b) Just reason as in the proof of a). □

Therefore, the knowledge of Ψ implies the knowledge of W and also of W^* and L. Define the following residues : $\underset{z}{\text{Res}}\sum a_kz^k = a_{-1}$, $\underset{\partial}{\text{Res}}\sum a_k\partial^k = a_{-1}$ and consider the following result [33],

Theorem 6.4. *Let P and Q be two pseudo-differential operators. So*

$$\underset{z}{Res}((Pe^{xz}).(Qe^{-xz})) = \underset{\partial}{Res}\, PQ^*,$$

where Q^* is the adjoint of Q.

Proof. Indeed, we have

$$\underset{z}{\text{Res}}((Pe^{xz}).(Qe^{-xz})) = \underset{z}{\text{Res}}\left(\sum p_kz^k\sum q_l(-z)^l\right) = \sum_{k+l=-1}(-1)^lp_kq_l,$$

and

$$\underset{\partial}{\text{Res}}\, PQ^* = \underset{\partial}{\text{Res}}\sum_{kl}p_k\partial^k(-\partial)^lq_l = \sum_{k+l=-1}(-1)^lp_kq_l,$$

hence the result. □

Moreover, we have [33, 34] :

$$\begin{aligned}\underset{z}{\text{Res}}\,(\partial^k\Psi).\Psi^* &= \underset{z}{\text{Res}}\left(\partial^kWe^{\xi(t,z)}\right)(W^*)^{-1}e^{-\xi(t,z)},\\ &= \underset{z}{\text{Res}}\left(\partial^kWe^{xz}\right)(W^*)^{-1}e^{-xz}, \quad x\equiv t-1,\\ &= \underset{\partial}{\text{Res}}\,\partial^kW.W^{-1},\\ &= \underset{\partial}{\text{Res}}\,\partial^k,\\ &= 0.\end{aligned}$$

This bilinear identity can be written in the following symbolic form :

$$\underset{z=\infty}{\text{Res}}\left(\Psi(t,z).\Psi^*(t',z)\right) = 0,$$

for all t and t'. Therefore, we have

Theorem 6.5. *$\Psi(t,z)$ is a wave function for the KP hierarchy if and only if the residue identity is satisfied :*

$$\underset{z=\infty}{Res}\left(\Psi(t,z).\Psi^*(t',z)\right) = 0 \quad \forall t,t' \tag{6.13}$$

or what amounts to the same if and only if

$$\frac{1}{2\pi\sqrt{-1}}\int_\gamma\Psi(t,z).\Psi^*(t',z)dz = 0, \tag{6.14}$$

with γ a closed path around $z=\infty$ (such that : $\int_\gamma\frac{dz}{2\pi\sqrt{-1}} = 1$).

Recall that a $\tau(t)$ function is defined by the Fay differential identity (see next theorem) :

$$\{\tau(t-[y_1]),\tau(t-[y_2])\}+(y_1^{-1}-y_2^{-1})(\tau(t-[y_1])\tau(t-[y_2])-\tau(t)\tau(t-[y_1]-[y_2]))=0,$$

où $y_1,y_2\in\mathbb{C}^*$ and $\{u,v\}$ is the Wronskian $u'v-uv'$.

Theorem 6.6. *Let's put* $[s]=\left(s,\frac{s^2}{2},\frac{s^3}{3},\ldots\right)$. *The τ function satisfies the following identities :*
(i) Fay identity :

$$\begin{aligned}\mathscr{F}(t,y_0,y_1,y_2,y_3) &\equiv (y_0-y_1)(y_2-y_3)\tau(t+[y_0]+[y_1])\tau(t+[y_2]+[y_3])\\ &\quad+(y_0-y_2)(y_3-y_1)\tau(t+[y_0]+[y_2])\tau(t+[y_2]+[y_1])\\ &\quad+(y_0-y_3)(y_1-y_2)\tau(t+[y_0]+[y_3])\tau(t+[y_1]+[y_2])\\ &= 0.\end{aligned}$$

(ii) Fay differential identity :

$$\{\tau(t-[y_1]),\tau(t-[y_2])\}+(y_1^{-1}-y_2^{-1})(\tau(t-[y_1])\tau(t-[y_2])-\tau(t)\tau(t-[y_1]-[y_2]))=0,$$

where $y_1,y_2\in\mathbb{C}^*$ and $\{u,v\}$ is the Wronskian $u'v-uv'$. This identity can still be written in the form

$$\partial^{-1}\psi(t,\lambda)\psi^*(t,\mu)=\frac{1}{\mu-\lambda}\frac{\tau\left(t-[\lambda^{-1}]+[\mu^{-1}]\right)}{\tau(t)}e^{\sum_{j=1}^{\infty}t_j(\mu^j-\lambda^j)}.$$

The following equation $\dot{\tau}=X(t,\lambda,\mu)\tau$, determines a vector field on the infinite dimension manifold of the τ functions where $X(t,\lambda,\mu)$ is the vertex operator (of Date-Jimbo-Kashiwara-Miwa) for the KP equation .

Proof. According to Sato theory [35, 36], the functions Ψ and Ψ^* can be expressed in terms of a tau function as follows :

$$\Psi(t,z)=We^{\xi(t,z)}=\frac{\tau(t-[z^{-1}])}{\tau(t)}e^{\xi(t,z)},$$

$$\Psi^*(t,z)=(W^*)^{-1}e^{-\xi(t,z)}=\frac{\tau(t+[z^{-1}])}{\tau(t)}e^{-\xi(t,z)}.$$

By replacing these expressions in the residue formula (6.13) or (6.14), we obtain a bilinear relation for the τ functions. Indeed, the equation (6.14) is written

$$\int_\gamma e^{\xi(t-t',z)}\tau(t-[z^{-1}])\tau(t'-[z^{-1}])dz=0.$$

'Using the following change : $t\leftarrow t+s$ and $t'\leftarrow t+s$, we obtain

$$\int_\gamma e^{\xi(-2s,z)}\tau(t-s-[z^{-1}])\tau(t+s+[z^{-1}])dz=0.$$

Using again the transformation

$$s\leftarrow t+\frac{1}{2}([y_0]+[y_1]+[y_2]+[y_3]),\qquad t\leftarrow\frac{1}{2}([y_0]-[y_1]-[y_2]-[y_3]),$$

and taking into account that $e^{\sum_1^\infty(ab^{-1})^j.j^{-1}}=1-ab^{-1}$, we obtain via the residue theorem

$$\begin{aligned}0 &= \int_\gamma\frac{1-zy_0}{\prod_{j=1}^3(1-zy_j)}\tau(t-s-[z^{-1}])\tau(t+s+[z^{-1}])dz,\\ &= 2\pi\sqrt{-1}\sum\operatorname*{Res}_{y_1^{-1},y_2^{-1},y_3^{-1}}\left(\frac{1-zy_0}{\prod_{j=1}^3(1-zy_j)}\tau(t-s-[z^{-1}])\tau(t+s+[z^{-1}])\right),\\ &= \frac{2\pi\sqrt{-1}}{(y_1-y_2)(y_2-y_3)(y_3-y_1)}\mathscr{F}(t,y_0,y_1,y_2,y_3),\end{aligned}$$

where

$$\begin{aligned}\mathscr{F}(t,y_0,y_1,y_2,y_3) \equiv\ & (y_0-y_1)(y_2-y_3)\tau(t+[y_0]+[y_1])\tau(t+[y_2]+[y_3])\\ &+(y_0-y_2)(y_3-y_1)\tau(t+[y_0]+[y_2])\tau(t+[y_2]+[y_1])\\ &+(y_0-y_3)(y_1-y_2)\tau(t+[y_0]+[y_3])\tau(t+[y_1]+[y_2]).\end{aligned}$$

The relation $\mathscr{F}(t,y_0,y_1,y_2,y_3)=0$ is the Fay identity. By making the transformation in the expression $(y_1y_2)^{-1}\frac{\partial \mathscr{F}}{\partial y_0}|_{y_0=y_3=0}$ and replacing t by $t-[y_1]-[y_2]$, we obtain the Fay differential identity which allows to define the τ functions :

$$\{\tau(t-[y_1]),\tau(t-[y_2])\}+(y_1^{-1}-y_2^{-1})\left(\tau(t-[y_1])\tau(t-[y_2])-\tau(t)\tau(t-[y_1]-[y_2])\right)=0,$$

where $y_1,y_2\in\mathbb{C}^*$ and $\{u,v\}$ is the Wronskian $u'v-uv'$. Consider the Fay differential identity above and replace t with $t+[y_1]$. We obtain

$$\{\tau(t),\tau(t+[y_1]-[y_2])\}+(y_1^{-1}-y_2^{-1})\left(\tau(t)\tau(t+[y_1]-[y_2])-\tau(t)\tau(t-[y_2])\right)=0.$$

By putting $\lambda=y_1^{-1}$, $\mu=y_2^{-1}$, we obtain after having multiplied the expression obtained by $\frac{1}{\tau(t)}e^{\sum_1^\infty t_j(\mu^j-\lambda^j)}$, the following formula :

$$\frac{\tau(t+[\lambda^{-1}])}{\tau(t)}e^{-\sum t_j\lambda^j}\frac{\tau(t-[\mu^{-1}])}{\tau(t)}e^{\sum t_j\mu^j}=\frac{1}{\mu-\lambda}\frac{\partial}{\partial x}\left(e^{\sum t_j(\mu^j-\lambda^j)}\frac{\tau(t+[\lambda^{-1}]-[\mu^{-1}])}{\tau(t)}\right).$$

Let

$$X(t,\lambda,\mu)=\frac{1}{\mu-\lambda}e^{\sum_1^\infty t_j(\mu^j-\lambda^j)}e^{\sum_1^\infty j^{-1}(\lambda^{-j}-\mu^{-j})\frac{\partial}{\partial t_j}},\quad \lambda\neq\mu,$$

be the vertex operator (of Date-Jimbo-Kashiwara-Miwa) for the KP equation, then $X(t,\lambda,\mu)\tau$ et $\tau+X(t,\lambda,\mu)\tau$ are also τ functions. Therefore, $\dot{\tau}=X(t,\lambda,\mu)\tau$ determines a vector field on the infinite dimension manifold of functions τ. We deduce, according to [33], that $\partial^{-1}\left(\Psi^*(t,\lambda)\Psi(t,\mu)\right)=\frac{1}{\tau(t)}X(t,\lambda,\mu)\tau(t)$. □

Let $\Delta(s_1,...,s_n)=\prod_{1\le j<i\le n}(s_i^{-1}-s_j^{-1})$, be the Vandermonde determinant. Fay identities (theorem 6.6) are generalized as follows. The τ function satisfies identities :

$$\tau\left(t-\sum_{j=1}^n[y_j]\right)\Delta(y_1,...,y_n)\left(\left(t-\sum_{j=1}^n[y_j]\right)\Delta(x_1,...,x_n)\right)^{n-1}$$

$$=\det\left(\left(t-\sum_{j=1}^n[x_k]+[x_j]-[y_l]\right)\Delta(x_1,...,x_{j+1},y_j,x_{j+1},...,x_n)\right)_{1\le j,l\le n},$$

and

$$\{\psi(t,y_1^{-1},...,\psi(t,y_n^{-1}\}=e^{\sum_{j=1}^\infty t_j(y_1^{-j}+\cdots+y_n^{-j})}\frac{\tau(t-[y_l]-\cdots-[y_n]}{\tau(t)}\Delta(y_1,...,y_n),$$

where $\{u_1,...,u_n\}$ is the Wronskian $\det\left(\left(\frac{\partial}{\partial x}\right)^{j-1}u_j\right)_{1\le i,j\le n}$.

We will see that τ functions characterize the KP hierarchy. Let $s_j(t)$ denote the elementary Schur polynomials, i.e., polynomials such that :

$$e^{\xi(t,z)}=e^{\sum_{j=1}^\infty t_jz_j}=\sum_{j=1}^\infty s_j(t)z_j=1+t_1z+\left(\frac{1}{2}t_1^2+t_2\right)z^2+\left(\frac{1}{6}t_1^3+t_1t_2+t_3\right)z^3+\cdots$$

with $s_j(t) = \frac{t_1^j}{j!} + \cdots + t_n$. By setting $\widetilde{\partial} = \left(\frac{\partial}{\partial t_1}, \frac{1}{2}\frac{\partial}{\partial t_2}, \frac{1}{3}\frac{\partial}{\partial t_3}, \ldots\right)$, we obtain

$$\Psi(t,z) = \frac{\tau\left(t_1 - z^{-1}, t_2 - \frac{z^{-2}}{2}, t_3 - \frac{z^{-3}}{3}, \ldots\right)}{\tau(t)} e^{\xi(t,z)} = \sum_{j=0}^{\infty} \frac{s_j(-\widetilde{\partial})\tau(t)}{\tau(t)} \partial^{-j} e^{\xi(t,z)} = W(t) e^{\xi(t,z)}, =$$

where $W(t) = \sum_{j=0}^{\infty} \frac{s_j(-\widetilde{\partial})\tau(t)}{\tau(t)} \partial^{-j}$, is the wave operator (6.7). Similarly, we have

$$W^{-1} = \sum_{j=0}^{\infty} \partial^{-j} \frac{s_j(\widetilde{\partial})\tau(t)}{\tau(t)}. \tag{6.15}$$

It follows from (6.8) that $L^n = W.\partial^n.W^{-1}$ and L^n is expressed in terms of the τ function, $L^n = \sum_{i,j=0}^{\infty} \frac{s_i(-\widetilde{\partial})\tau}{\tau} \partial^{n-i-j} \frac{s_j(\widetilde{\partial})}{\tau}\tau$. By developing this expression, we get

$$L^n = \partial^n + n(\log\tau)''\partial^{n-2} + \cdots + \sum_{i+j=n+1} \frac{s_i(\widetilde{\partial})\tau s_j(-\widetilde{\partial})\tau}{\tau^2} + \cdots$$

The formula (6.10) is written taking into account this last expression of L^n and the relation (6.15) as follows :

$$\frac{\partial}{\partial t_n}\left(1 - \frac{\tau'}{\tau}\partial^{-1} + \cdots\right) = \left(-\sum_{i+j=n+1} \frac{s_i(-\widetilde{\partial})\tau s_j(-\widetilde{\partial})\tau}{\tau^2}\partial^{-1} + \cdots\right)\left(1 - \frac{\tau'}{\tau}\partial^{-1} + \cdots\right).$$

Using the Hirota symbol[1], we have

$$\sum_{\substack{i+j=n+1\\ i,j\geq 0}} \left(s_i(\widetilde{\partial})\tau\right)\left(s_j(-\widetilde{\partial})\tau\right) = s_{n+1}(\widetilde{\partial})\tau.\tau,$$

and we obtain

$$\tau^2 \frac{\partial^2}{\partial t_n \partial t_1}\log\tau - \sum_{\substack{i+j=n+1\\ i,j\geq 0}} s_i(\widetilde{\partial})\tau s_j(-\widetilde{\partial})\tau = 0, \quad n \in \mathbb{N}^*.$$

These relations are called Hirota bilinear equations. They show that all the functions, $j \geq 2$, can be expressed in terms of the τ function. For example,

$$u_2 = \frac{\partial^2}{\partial t_1^2}\log\tau, \qquad u_3 = \frac{1}{2}\left(\frac{\partial^3}{\partial t_1^3} + \frac{\partial}{\partial t_1}\frac{\partial}{\partial t_3}\right)\log\tau, \qquad u_4 = \frac{1}{6}\left(\frac{\partial^4}{\partial t_1^4} - 3\frac{\partial^2}{\partial t_1^2}\frac{\partial}{\partial t_2} + 2\frac{\partial}{\partial t_1}\frac{\partial}{\partial t_2}\right)\log\tau - \left(\frac{\partial^2}{\partial t_1^2}\log\tau\right), \ldots$$

In particular, these equations provide the KP equation in the following bilinear form :

$$\frac{1}{12}\tau\left(\frac{\partial^4\tau}{\partial t_1^4} - 4\frac{\partial^2\tau}{\partial t_1\partial t_3} + 3\frac{\partial^2\tau}{\partial t_2^2}\right) - \frac{1}{3}\frac{\partial\tau}{\partial t_1}\left(\frac{\partial^3\tau}{\partial t_1^3} - \frac{\partial\tau}{\partial t_3}\right) + \frac{1}{4}\left(\frac{\partial^2\tau}{\partial t_1^2} + \frac{\partial\tau}{\partial t_2}\right)\left(\frac{\partial^2\tau}{\partial t_1^2} - \frac{\partial\tau}{\partial t_2}\right) = 0.$$

Therefore, we have

Theorem 6.7. *The τ functions characterize the KP hierarchy.*

The equations of soliton theory play an important role in the characterization of Jacobian varieties. Let

$$\mathscr{H}_g = \{Z \in M_g(\mathbb{C}) : \Omega = \Omega^\top, \ \mathrm{I}\Omega > 0\},$$

[1] $p(\partial_t)f(t).g(t) \equiv p\left(\frac{\partial}{\partial s_1}, \frac{\partial}{\partial s_2}, \ldots\right) f(t+s)g(t-s)\Big|_{s=0}$ where p is any polynomial, $f(t)$ and $g(t)$ are two differentiable functions.

be the Siegel half-space, $\Lambda = \mathbb{Z}^g \oplus Z\mathbb{Z}^g$ a lattice in $\mathbb{C}^g$ and $T = \mathbb{C}^g/\Lambda$ a principally polarised Abelian variety. We show [24, 37] that the following three conditions are equivalent : (i) There are vector fields v_1, v_2, v_3 on $\mathbb{C}^g$ and a quadratic form $q(t) = \sum_{k,l=1}^{3} q_{kl}(t) t_k t_l$, such that : for all $z \in \mathbb{C}^g$, the function $\tau(t) = e^{q(t)}\theta\left(\sum_{k=1}^{3} t_k v_k + z\right)$, satisfies the KP equation. The theta divisor does not contain an Abelian subvariety of T for which the vector v_1 is tangent. (ii) T is isomorphic to the Jacobian variety of a reduced non-singular complete curve of genus g. (iii) There is a matrix $\mathscr{V} = (v_1, v_2, \ldots)$ of order $g \times \infty$, $v_k \in \mathbb{C}^g$, of rank g and a quadratic form $Q(t) = \sum_{k,l=1}^{\infty} q_{kl}(t) t_k t_l$, such that: for all $z \in \mathbb{C}^g$, $\widetilde{\tau}(t) = e^{Q(t)}\theta\left(\mathscr{V}t + z\right)$, is a τ function for KP hierarchy.

References

[1] D. J. Korteweg, G. de Vries, *On the change of form of long waves advancing in a rectangular canal and on a new type of long stationary waves*, Phil. Mag., **39** (1895), 422-443.

[2] J. Boussinesq, *Essai sur la théorie des eaux courantes*, Mémoires présentés par divers savants, Acad. des Sci. Inst. Nat. France, **XXIII** (1877), 1-680.

[3] J. Scott-Russell, *Report on Waves*. Report to the 14th meeting British Association for Advancement of Science, (1844), 311-390.

[4] I. M. Gel'fand, B. M. Levitan, *On the determination of a differential equation from its spectral function*, Amer. Math. Soc. Transl., **2** (1) (1955), 253-304.

[5] N. Levinson, *The inverse Sturm-Liouville problem*, Mat. Tidsskr. B., (1949) 25-30.

[6] M. D. Kruskal, R. M. Miura, C. S. Gardner, N. J. Zabusky, *Korteweg-de Vries equation and generalizations V. Uniqueness and nonexistence of polynomial consevation laws*, J. Math. Phys., **11** (1970), 952-960.

[7] P. Lax, *Integrals of nonlinear equations of evolution and solitary waves*, Comm. Pure Appl. Math., **21** (1968), 467-490.

[8] C. S. Gardner, J. M. Greene, M. D. Kruskal, R. M. Miura, *Method for solving the Korteweg-de Vries equation*, Phys. Rev. Lett., **19** (1967), 1095-1097.

[9] C. S. Gardner, *Korteweg-de Vries equation and generalizations IV, The Korteweg-de Vries equation as a Hamiltonian system*, J. Math. Phys., **12** (1971), 1548-1551.

[10] C. S. Gardner, J. M. Greene, M. D. Kruskal, R. M. Miura, *Korteweg-de Vries equation and generalizations VI, Methods for exact solutions*, Comm. Pure Appl. Math., **27** (1974), 97-133.

[11] R. M. Miura, C. S. Gardner, M. D. Kruskal, *Korteweg-de Vries equation and generalizations II. Existence of conservation laws and constants of motion*, J. Math. Phys., **9** (1968), 1204-1209.

[12] P. Lax, *Periodic solutions of the KdV equation*, Comm. Pure Appl. Math., **28** (1975), 141-188.

[13] P. Lax, R. Phillips, *Scattering theory for automorphic functions*. Annals Math. Studies, Princeton, 1976.

[14] A. Lesfari, *Etude des équations stationnaire de Schrödinger, intégrale de Gelfand-Levitan et de Korteweg-de-Vries. Solitons et méthode de la diffusion inverse*, Aequat. Math., **85** (2013), 243-272.

[15] M. Wadati, M. Toda, *The exact N-soliton solution of the Korteweg-de Vries equation*, J. Phys. Soc. Japan, **32** (1972), 1403-1411.

[16] B. B. Kadomtsev, V I. and Petviashvili, *On the Stability of Solitary Waves in Weakly Dispersing Media*, Sov. Phys. Dokl., **15** (6) (1970), 539-541.

[17] M. Sato, *Soliton equations and the universal Grassmann manifold*, Math. Lect. Note Ser. 18. Sophia University, Tokyo, 1984.

[18] P. van Moerbeke, *Integrable foundations of string theory in Lectures on Integrable Systems*, (Sophia-Antipolis, 1991), World Sci., River Edge, N.J., (1994), 163-267.

[19] J.-L. Gervais, *Infinite family of polynomial functions of the Virasoro generators with vanishing Poisson bracket*, Phys. Lett. B, **160** (4-5) (1985), 277-278.

[20] M. Adler, P. van Moerbeke, P. Vanhaecke, *Algebraic integrability, Painlevé geometry and Lie algebras*, A series of modern surveys in mathematics, Volume 47, Springer-Verlag, 2004.

[21] A. Lesfari, *Abelian surfaces and Kowalewski's top*, Ann. Scient. École Norm. Sup., Paris, 4^e série, **.21** (1988), 193-223.

[22] M. Adler, *On a trace functional for formal pseudo differential operators and the symplectic structure of the Korteweg-de Vries type equations*, Invent. Math., **50**(3) (1979), 219-248.

[23] B. Kostant, *The solution to a generalized Toda lattice and representation theory*, Adv. Math., **34** (3) (1979), 195-338.

[24] W. Symes, *Systems of Toda type, inverse spectral problems and representation theory*, Invent. Math., **59** (1980), 13-53.

[25] A. I. Belokolos, V. Z. Bobenko, V.Z., Enol'skii, A. R. Its, V. B. Matveev, *Algebro-Geometric approach to nonlinear integrable equations*, Springer-Verlag, 1994.

[26] A. Lesfari, *Integrables Hamiltonian systems and the isospectral deformation method*, Int. J. of Appl. Math. And Mech., **3** (4) (2007), 35-55.

[27] A. Lesfari, *Théorie spectrale et problèmes non-linéaires*, Surv. Math. Appl., **5** (2010), 151-190.

[28] A. Lesfari, *Algèbres de Lie affines et opérateurs pseudo-différentiels d'ordre infini*, Math. Rep., **14** (64),1 (2012), 43-69 (2012).

[29] A. Lesfari, *Introduction à la géométrie algébrique complexe*, Éditions Hermann, Paris, 2015.

[30] I. M. Gel'fand, L. Dickey, *Family of Hamiltonian structures connected with integrable nonlinear differential equations*, Funct. Anal. Appl., **2** (1968), 92-93.

[31] L. Dickey, *Soliton equations and integrable systems*, World Scientific, 1991.

[32] L. Dickey, *Additional symmetries of KP, Grassmannian and the string equation*, Mod. Phys. Lett., A **8** (1993), 1259-1272.

[33] L. Dickey, *Lectures on classical W-algebras*, Acta Appl. Math., **47** (1997), 243-321.

[34] E. Date, M. Jimbo, M. Kashiwara, T. Miwa, *Transformation groups for soliton equations*, Proc. RIMS Symp. Nonlinear integrable systems (Kyoto, 1981). Classical and quantum theory, Singapore, World Scientific (1983), 39-119.

[35] M. Sato, *The KP hierarchy and infinite-dimensional Grassmann manifolds*, Proc. of Sympos. Pure Math., **49** (1989), 51-66.

[36] M. Sato, Y. Sato, *Soliton equations as dynamical systems on infinite dimensional Grassmann manifolds*, Lect. Notes in Num. Appl. Anal., **5** (1982), 259-271.

[37] T. Shiota, *Characterization of jacobian varieties in terms of soliton equations*, Invent. Math., **83** (1986), 333-382.

Communications in Advanced Mathematical Sciences
Vol. II, No. 2, 105-113, 2019
Research Article
e-ISSN:2651-4001
DOI: 10.33434/cams.486401

Construction of Exact Solutions to Partial Differential Equations with CRE Method

Filiz Taşcan[1]*, Arzu Akbulut[2]

Abstract
In this article, the consistent Riccati expansion (CRE) method is presented for constructing new exact solutions of (1+1) dimensional nonlinear dispersive modified Benjamin Bona Mahony (DMBBM) and mKdV-Burgers equations. The exact solutions obtained are composed of hyperbolic and exponential functions. The outcomes obtained confirm that the proposed method is an efficient technique for analytic treatment of a wide variety of nonlinear partial differential equations.

Keywords: Partial differential equations, Exact solution, The consistent Riccati expansion.
2010 AMS: Primary 35Q53, 35C07, Secondary 83C15

[1] *Eskisehir Osmangazi University, Art-Science Faculty, Department of Mathematics-Computer, Eskisehir-TURKEY, ORCID: 0000-0003-2697-5271*
[2] *Eskisehir Osmangazi University, Art-Science Faculty, Department of Mathematics-Computer, Eskisehir-TURKEY, ORCID: 0000-0003-2448-2481*
***Corresponding author**: ftascan@ogu.edu.tr
Received: 21 November 2018, **Accepted:** 15 February 2019, **Available online:** 27 June 2019

1. Introduction

Nonlinear evolution equations (NLEEs) in mathematical physics play a vital role in different fields, such as fluid mechanics, plasma physics, optical fibers, solid state physics, chemical kinematic, chemical physics and geochemistry. Since obtaining exact solutions of NLEEs come into prominence, there become significant improvements in this domain[1]. Many effective and powerful methods have been established and improved, such as modified simple equation method [2], symmetry reduction method[3], trial equation method [4], the (G'/G)-expansion method [5], sub equation method [6], $\exp(-\Phi(\xi))$ method[7], functional variable method[8], first integral method[9], modified exp-function method [10] and so on.

The aim of this paper is search new solutions of (1+1) dimensional nonlinear dispersive modified Benjamin Bona Mahony (DMBBM) equation and modified Korteweg-de Vries (mKdV)-Burgers equation with consistent Riccati expansion (CRE) method. In section 2, we give the definition of the method. In section 3, there found solutions of the given equations. In section 4, conslusions are given.

2. Consistent Riccati expansion (CRE) method

Lets assume that we have a nonlinear differential equation, remark in the independent variables x and t and dependent variable u, given by

$$F\left(u, u_x, u_t, u_{xx}, u_{tt}, \ldots\right) = 0, \tag{2.1}$$

where F is a polynomial of $u(x,t)$ and its various partial derivatives including the highest order derivatives and nonlinear terms.

According to the algoritm, we can seek for the solutions of Eq. (2.1) in the form

$$u=\sum_{i=0}^{n}u_i\left(x,t\right)R^i\left(w\right), \tag{2.2}$$

where u_i $(i=0,...,n)$ are functions to be detected later and the positive integer n can be detected by using homogeneous balance method. Here $R(w)$ is a solution of the Riccati equation

$$R_w=a_0+a_1R+a_2R^2 \tag{2.3}$$

where a_0,a_1,a_2 are parameters to be determined and w is an undetermined function of x and t.

The positive integer n can be detected by considering the homogeneous balance between the highest order derivative term with the highest order nonlinear term appearing in Eq. (2.1). Then by sett ing Eq. (2.2) along with Eq. (2.3) into Eq. (2.1) and equating the coefficients of all powers of $R\left(w\right)$ to zero yields a set of algebraic equations for unknowns u_i,a_0,a_1 and a_2[11, 12].

3. Exercises

In this part, we have dealed with two partial differential equations as an application of the CRE method.

3.1 (1+1) dimensional nonlinear dispersive modified Benjamin Bona Mahony (DMBBM) equation

Firstly, we look at the (1+1) dimensional nonlinear dispersive modified Benjamin Bona Mahony (DMBBM) equation [13]

$$u_t+u_x-\alpha u^2u_x+u_{xxx}=0, \tag{3.1}$$

where where α is a nonzero constant. This equation was first derived to describe an approximation for surface long waves in nonlinear dispersive media. It can also characterize the hydro magnetic waves in cold plasma, acoustic waves in inharmonic crystals and acoustic gravity waves in compressible fluids [14].

Here, it is clear from the homogoneous balance principle that the balancing number is 1. From here, we infer from that the exact solution of Eq. (3.1) is

$$u(x,t)=u_0\left(x,t\right)+u_1\left(x,t\right)R\left(w\left(x,t\right)\right) \tag{3.2}$$

where $u_0\left(x,t\right)$ and $u_1\left(x,t\right)$ are functions to be determined later. Setting Eq. (3.2) and its derivatives with the condition Eq. (2.3) into Eq. (3.1) and gathering all terms with the same power of $R\left(w\right),(i=0,1,...,4)$, we obtain the following system

$$R^4\left(w\right)\quad:\quad 6u_1w_x^3a_2^3-\alpha u_1^3w_xa_2=0, \tag{3.3}$$

$$\begin{aligned}R^3\left(w\right)\quad:\quad & -2\alpha u_0u_1^2w_xa_2+6\left(u_1\right)_xw_x^2a_2^2+12u_1w_x^3a_1a_2^2+6u_1w_xw_{xx}a_2^2\\ & -\alpha u_1^2\left(u_1\right)_x-\alpha u_1^3w_xa_1=0,\end{aligned} \tag{3.4}$$

$$\begin{aligned}R^2\left(w\right)\quad:\quad & 9\left(u_1\right)_xw_x^2a_1a_2+8u_1w_x^3a_0a_2^2-\alpha u_0^2u_1w_xa_2+9u_1w_xw_{xx}a_1a_2\\ & -\alpha u_1^2\left(u_0\right)_x+u_1w_xa_2-2\alpha u_0u_1^2w_xa_1+u_1w_ta_2-\alpha u_1^3w_xa_0\\ & +7u_1w_x^3a_1^2a_2+3\left(u_1\right)_{xx}w_xa_2-2\alpha u_0u_1\left(u_1\right)_x\\ & +3\left(u_1\right)_xw_{xx}a_2+u_1w_{xxx}a_2=0,\end{aligned} \tag{3.5}$$

$$\begin{aligned}R^1\left(w\right)\quad:\quad & \left(u_1\right)_{xxx}+u_1w_ta_1+3\left(u_1\right)_xw_{xx}a_1+3u_1w_xw_{xx}a_1^2\\ & +\left(u_1\right)_t-\alpha u_0^2u_1w_xa_1+8u_1w_x^3a_1a_2a_0\\ & -2\alpha u_0u_1^2w_xa_0-2\alpha u_0u_1\left(u_0\right)_x+u_1w_x^3a_1^3+\left(u_1\right)_x\\ & +6\left(u_1\right)_xw_x^2a_2a_0+6u_1w_xw_{xx}a_2a_0+u_1w_{xxx}a_1\\ & +3\left(u_1\right)_{xx}w_xa_1+3\left(u_1\right)_xw_x^2a_1^2+u_1w_xa_1-\alpha u_0^2\left(u_1\right)_x=0,\end{aligned} \tag{3.6}$$

$$\begin{aligned}R^0\left(w\right)\quad:\quad & 3\left(u_1\right)_xw_x^2a_1a_0+u_1w_{xxx}a_0+3u_1w_xw_{xx}a_1a_0\\ & +\left(u_0\right)_x+\left(u_0\right)_t+\left(u_0\right)_{xxx}+3\left(u_1\right)_{xx}w_xa_0+u_1w_ta_0\\ & +3\left(u_1\right)_xw_{xx}a_0-\alpha u_0^2\left(u_0\right)_x+u_1w_x^3a_1^2a_0\\ & +u_1w_xa_0+2u_1w_x^3a_2a_0^2-\alpha u_0^2u_1w_xa_0=0.\end{aligned} \tag{3.7}$$

From the Eq. (3.3), we get

$$u_1(x,t) = \sqrt{6}\sqrt{\frac{1}{\alpha}}a_2 w_x. \tag{3.8}$$

If we substitute Eq. (3.8) in Eq. (3.4), we obtain

$$u_0(x,t) = \frac{\sqrt{6}\sqrt{\frac{1}{\alpha}}w_{xx}}{2w_x} + \sqrt{6}\sqrt{\frac{1}{\alpha}}a_1 w_x - \frac{1}{2}w_x\alpha\sqrt{6}\sqrt{\left(\frac{1}{\alpha}\right)^3}a_1. \tag{3.9}$$

When we substitute Eq. (3.8) and Eq. (3.9) in Eq. (3.5), we get following partial differential equation

$$w_t w_x = -\frac{4a_2 w_x^4 a_0 - w_x^4 a_1^2 - 3w_{xx}^2 + 2w_x w_{xxx} + 2w_x^2}{2}. \tag{3.10}$$

If we use Eq. (3.10) in Eq. (3.6) and Eq. (3.7), these Eqs. are equal to zero.
If w is a solution of Eq. (3.10), then

$$u = \frac{\sqrt{6}\sqrt{\frac{1}{\alpha}}w_{xx}}{2w_x} + \sqrt{6}\sqrt{\frac{1}{\alpha}}a_1 w_x - \frac{1}{2}w_x\alpha\sqrt{6}\sqrt{\left(\frac{1}{\alpha}\right)^3}a_1 + \sqrt{6}\sqrt{\frac{1}{\alpha}}a_2 w_x R \tag{3.11}$$

is a solution of the DMBBM equation with $R \equiv R(w)$ being a solution of the Riccati equation (2.3).

We suppose that $w(x,t)$ be of the form

$$w(x,t) = a\cosh(kx+lt+\xi) + b\sinh(kx+lt+\xi) + r \tag{3.12}$$

where a,b,k,l and r are constants to be determined later and ξ is an arbitrary constant. Setting Eq. (3.12) into Eq. (3.10) , we obtain the following equations

$$-\frac{k\left(16a_2a_0a^3k^3b - 4a_1^2ak^3b^3 - 4a_1^2a^3k^3b + 16a_2a_0ak^3b^3\right)}{2} = 0,$$

$$-\frac{k\left(-16a_2a_0a^3k^3b + 4alb + 4a_1^2a^3k^3b - 2ak^3b + 4akb\right)}{2} = 0,$$

$$-\frac{k\left(-a_1^2a^4k^3 + 24a_2a_0a^2k^3b^2 - a_1^2b^4k^3\right)}{2}$$

$$-\frac{k\left(-6a_1^2a^2k^3b^2 + 4a_2a_0b^4k^3 + 4a_2a_0a^4k^3\right)}{2} = 0,$$

$$-\frac{k\left(2a_1^2a^4k^3 + 6a_1^2a^2k^3b^2 - 8a_2a_0a^4k^3 + 2a^2k + 2b^2k\right)}{2}$$

$$-\frac{k\left(-b^2k^3 - a^2k^3 + 2a^2l - 24a_2a_0a^2k^3b^2 + 2b^2l\right)}{2} = 0,$$

$$-\frac{k\left(-a_1^2a^4k^3 - 2a^2k - 2a^2l - 2a^2k^3 + 4a_2a_0a^4k^3 + 3b^2k^3\right)}{2} = 0$$

Solving above system, we get the following two solutions.

State 1:

$$\begin{aligned} &a = b,\ a_0 = \tfrac{a_1^2}{4a_2},\ a_1 = a_1,\ a_2 = a_2,\ b = b, \\ &k = k,\ \xi = \xi,\ l = \tfrac{k\left(k^2-2\right)}{2},\ r = r. \end{aligned} \tag{3.13}$$

State 2:

$$\begin{aligned} &a = -b,\ a_0 = \tfrac{a_1^2}{4a_2},\ a_1 = a_1,\ a_2 = a_2,\ b = b, \\ &k = k,\ \xi = \xi,\ l = \tfrac{k\left(k^2-2\right)}{2},\ r = r. \end{aligned} \tag{3.14}$$

Combining Eq. (3.11), Eq. (3.12) with Eq. (3.13) and Eq. (3.14), two families of exact explicit solutions to the DMBBM equation are obtained

$$\begin{aligned} u(x,t) &= \frac{1}{2}\sqrt{6}\sqrt{\frac{1}{\alpha}}k\left(a_1 b\cosh(\beta)+1+a_1 b\sinh(\beta)\right) \\ &+\sqrt{6}\sqrt{\frac{1}{\alpha}}a_2\left(bk\sinh(\beta)+bk\cosh(\beta)\right) \\ &\times R\left(b\cosh(\beta)+b\sinh(\beta)+r\right) \end{aligned}$$

and

$$\begin{aligned} u(x,t) &= \frac{1}{2}\sqrt{6}\sqrt{\frac{1}{\alpha}}k\left(a_1 b\cosh(\beta)-1-a_1 b\sinh(\beta)\right) \\ &+\sqrt{6}\sqrt{\frac{1}{\alpha}}a_2\left(-bk\sinh(\beta)+bk\cosh(\beta)\right) \\ &\times R\left(-b\cosh(\beta)+b\sinh(\beta)+r\right). \end{aligned}$$

where $\beta = kx - kt + \frac{k^3 t}{2} + \xi$.

We suppose that $w(x,t)$ be of the form

$$w(x,t) = A\exp\left(k_1 x + l_1 t + \xi_1\right) + B\exp\left(k_2 x + l_2 t + \xi_2\right) + C \tag{3.15}$$

where A, B, C, k_i and l_i are constants to be determined later and ξ_i are an arbitrary constant. Setting Eq. (3.15) into Eq. (3.10) , we get the following system

$$\begin{aligned} \frac{a_1^2 B^4 k_2^4}{2} - 2a_2 a_0 B^4 k_2^4 &= 0, \\ -8a_2 a_0 A k_1 B^3 k_2^3 + 2a_1^2 A k_1 B^3 k_2^3 &= 0, \\ -12a_2 a_0 A^2 k_1^2 B^2 k_2^2 + 3a_1^2 A^2 k_1^2 B^2 k_2^2 &= 0, \\ -B^2 k_2^2 + \tfrac{1}{2} B^2 k_2^4 - B^2 l_2 k_2 &= 0, \\ -8a_2 a_0 A^3 k_1^3 B k_2 + 2a_1^2 A^3 k_1^3 B k_2 &= 0, \\ -A l_1 B k_2 - A k_1^3 B k_2 - B l_2 A k_1 - B k_2^3 A k_1 + 3A k_1^2 B k_2^2 - 2A k_1 B k_2 &= 0, \\ \frac{a_1^2 A^4 k_1^4}{2} - 2a_2 a_0 A^4 k_1^4 &= 0, \\ \frac{A^2 k_1^4}{2} - A^2 k_1^2 - A^2 l_1 k_1 &= 0. \end{aligned}$$

Solving above system, one gets the following set of solution.

$$\begin{aligned} &A = A,\ B = B,\ C = C,\ a_0 = \tfrac{a_1^2}{4a_2},\ a_1 = a_1,\ a_2 = a_2,\ k_1 = k_2, \\ &k_2 = k_2,\ l_1 = -k_2 + \tfrac{k_2^3}{2},\ l_2 = -k_2 + \tfrac{k_2^3}{2},\ \xi_1 = \xi_1,\ \xi_2 = \xi_2 \end{aligned} \tag{3.16}$$

Combining Eq. (3.11), Eq. (3.15) with Eq. (3.16), exact explicit solution is obtained

$$\begin{aligned} u(x,t) &= \frac{k_2\sqrt{6}}{2}\sqrt{\frac{1}{\alpha}}\left(Aa_1\exp(\beta+\xi_1)+Ba_1\exp(\beta+\xi_2)+1\right) \\ &+a_2\sqrt{6}\sqrt{\frac{1}{\alpha}}\left(Ak_2\exp(\beta+\xi_1)+Bk_2\exp(\beta+\xi_2)\right) \\ &\times R\left(A\exp(\beta+\xi_1)+B\exp(\beta+\xi_2)+C\right) \end{aligned}$$

where $\beta = k_2 x+\left(-k_2+\frac{k_2^3}{2}\right)t$.

3.2 Modified Korteweg-de Vries (mKdV)-Burgers equation

mKdV-Burgers equation is given by [15]

$$u_t+qu^2u_x+ru_{xx}-su_{xxx}=0 \tag{3.17}$$

where $q,\ r$ and s are arbitrary constants. According to the homogeneous balance method, we get the balancing number as $n=1$. From here, we infered that the exact solution of Eq. (3.17) is

$$u(x,t)=u_0(x,t)+u_1(x,t)R(w(x,t)) \tag{3.18}$$

where $u_0(x,t)$ and $u_1(x,t)$ are functions to be detected later. Setting Eq. (3.18) and its derivatives with the condition Eq. (2.3) into Eq. (3.17) and picking all terms with the same power of $R(w),(i=0,1,...,4)$, we have the following system

$$R^4(w)\ :\ qu_1^3w_xa_2-6su_1w_x^3a_2^3=0, \tag{3.19}$$

$$\begin{aligned} R^3(w)\ :\ & 2qu_0u_1^2w_xa_2-6s(u_1)_xw_x^2a_2^2+2ru_1w_x^2a_2^2-12su_1w_x^3a_1a_2^2 \\ & +qu_1^2(u_1)_x+qu_1^3w_xa_1-6su_1w_xw_{xx}a_2^2=0, \end{aligned} \tag{3.20}$$

$$\begin{aligned} R^2(w)\ :\ & qu_1^2(u_0)_x+2r(u_1)_xw_xa_2-su_1w_{xxx}a_2 \\ & +u_1w_ta_2-3s(u_1)_xw_{xx}a_2-8su_1w_x^3a_0a_2^2 \\ & +3ru_1w_x^2a_1a_2+2qu_0u_1^2w_xa_1-3s(u_1)_{xx}w_xa_2 \\ & +2qu_0u_1(u_1)_x-9s(u_1)_xw_x^2a_1a_2-9su_1w_xw_{xx}a_1a_2 \\ & +qu_1^3w_xa_0+qu_0^2u_1w_xa_2-7su_1w_x^3a_1^2a_2+ru_1w_{xx}a_2=0, \end{aligned} \tag{3.21}$$

$$\begin{aligned} R^1(w)\ :\ & qu_0^2(u_1)_x-6su_1w_xw_{xx}a_2a_0+2qu_0u_1^2w_xa_0+r(u_1)_{xx} \\ & +ru_1w_{xx}a_1-su_1w_{xxx}a_1+(u_1)_t-3s(u_1)_{xx}w_xa_1 \\ & -3s(u_1)_xw_x^2a_1^2-su_1w_x^3a_1^3+ru_1w_x^2a_1^2-3s(u_1)_xw_{xx}a_1 \\ & +qu_0^2u_1w_xa_1-s(u_1)_{xxx}+u_1w_ta_1-3su_1w_xw_{xx}a_1^2 \\ & -8su_1w_x^3a_1a_2a_0+2r(u_1)_xw_xa_1+2ru_1w_x^2a_2a_0 \\ & +2qu_0u_1(u_0)_x-6s(u_1)_xw_x^2a_2a_0=0, \end{aligned} \tag{3.22}$$

$$\begin{aligned} R^0(w)\ :\ & u_1w_ta_0-3s(u_1)_xw_x^2a_1a_0+2r(u_1)_xw_xa_0+ru_1w_x^2a_1a_0 \\ & +ru_1w_{xx}a_0-3s(u_1)_{xx}w_xa_0+qu_0^2u_1w_xa_0-su_1w_{xxx}a_0 \\ & +r(u_0)_{xx}-2su_1w_x^3a_2a_0^2-3s(u_1)_xw_{xx}a_0+qu_0^2(u_0)_x \\ & -3su_1w_xw_{xx}a_1a_0+(u_0)_t-s(u_0)_{xxx}-su_1w_x^3a_1^2a_0=0. \end{aligned} \tag{3.23}$$

From the Eq. (3.19), we get

$$u_1(x,t)=\frac{\sqrt{6}\sqrt{s}a_2w_x}{\sqrt{q}}. \tag{3.24}$$

If we substitute Eq. (3.24) in Eq. (3.20), we obtain

$$u_0(x,t) = \frac{\sqrt{6}\left(3sw_{xx} - rw_x + 3sw_x^2 a_1\right)}{6\sqrt{s}\sqrt{q}w_x}. \tag{3.25}$$

When we substitute Eq. (3.24) and Eq. (3.25) in Eq. (3.21), we get following partial differential equation

$$w_t w_x = -3sw_{xx}^2 + 2sa_2 a_0 w_x^4 - \frac{sw_x^4 a_1^2}{2} + sw_{xxx}w_x - \frac{w_x^2 r^2}{6s}. \tag{3.26}$$

If we use Eq. (3.26) in Eq. (3.22), this Eq. is equal to zero. If we use Eq. (3.26) in Eq. (3.23), we obtain

$$\frac{r\sqrt{6s}\left(4w_{xx}a_2 w_x^4 a_0 + 3w_{xx}^3 + w_{xxxx}w_x^2 - 4w_{xx}w_{xxx}w_x - w_x^4 a_1^2 w_{xx}\right)}{2\sqrt{q}w_x^3} = 0. \tag{3.27}$$

If w is a solution of Eqs. (3.26) and (3.27), then

$$u = \frac{\sqrt{6}\left(3sw_{xx} - rw_x + 3sw_x^2 a_1\right)}{6\sqrt{s}\sqrt{q}w_x} + \frac{\sqrt{6}\sqrt{s}a_2 w_x}{\sqrt{q}}R \tag{3.28}$$

is a solution of the Eq. (3.17) with $R \equiv R(w)$ being a solution of the Riccati equation (2.3).

We suppose that $w(x,t)$ be of the form

$$w(x,t) = a\cosh(kx + lt + \xi) + b\sinh(kx + lt + \xi) + r \tag{3.29}$$

where a, b, k, l and r are constants to be determined later and ξ is an arbitrary constant. Setting Eq. (3.29) into Eqs. (3.26) and

(3.27), we obtain the following equations

$$sa^2k^4 - a^2kl - \frac{3sb^2k^4}{2} - \frac{r^2a^2k^2}{6s} = 0,\ 2sa_2a_0a^4k^4 - \frac{sa_1^2a^4k^4}{2} = 0,$$

$$-3sa_1^2a^2k^4b^2 + 12sa_2a_0a^2k^4b^2 = 0,\ 8sa_2a_0a^3k^4b - 2sa_1^2a^3k^4b = 0,$$

$$-\frac{sa_1^2b^4k^4}{2} + 2sa_2a_0b^4k^4 = 0,\ -sak^4b - 2akbl - \frac{r^2ak^2b}{3s} = 0,$$

$$\frac{2\sqrt{6}r\sqrt{s}bk^6a_2a^4a_0}{\sqrt{q}} - \frac{\sqrt{6}r\sqrt{s}a^4k^6ba_1^2}{2\sqrt{q}} = 0,\ -\frac{r^2b^2k^2}{6s} - b^2kl - \frac{3sa^2k^4}{2} + sb^2k^4 = 0,$$

$$-\frac{\sqrt{6}r\sqrt{s}a^5k^6a_1^2}{2\sqrt{q}} - \frac{2\sqrt{6}r\sqrt{s}a^3k^6b^2a_1^2}{\sqrt{q}} + 2a_0\left(\frac{4\sqrt{6}r\sqrt{s}a^3k^6b^2a_2}{\sqrt{q}} + \frac{\sqrt{6}r\sqrt{s}a^5k^6a_2}{\sqrt{q}}\right) = 0,$$

$$-\frac{2\sqrt{6}r\sqrt{s}a^4k^6ba_1^2}{\sqrt{q}} - \frac{3\sqrt{6}r\sqrt{s}a^2k^6b^3a_1^2}{\sqrt{q}} + 2a_0\left(\frac{6\sqrt{6}r\sqrt{s}a^2k^6b^3a_2}{\sqrt{q}} + \frac{4\sqrt{6}r\sqrt{s}a^4k^6a_2b}{\sqrt{q}}\right) = 0,$$

$$-2sa_1^2ak^4b^3 + 8sa_2a_0ak^4b^3 = 0,\ -\frac{3\sqrt{6}r\sqrt{s}bk^6a^2}{2\sqrt{q}} + \frac{3\sqrt{6}r\sqrt{s}b^3k^6}{2\sqrt{q}} = 0,$$

$$-\frac{3\sqrt{6}r\sqrt{s}a^3k^6b^2a_1^2}{\sqrt{q}} - \frac{2\sqrt{6}r\sqrt{s}ab^4k^6a_1^2}{\sqrt{q}} + 2a_0\left(\frac{4\sqrt{6}r\sqrt{s}ak^6b^4a_2}{\sqrt{q}} + \frac{6\sqrt{6}r\sqrt{s}a^3k^6a_2b^2}{\sqrt{q}}\right) = 0,$$

$$\frac{7\sqrt{6}r\sqrt{s}ak^6b^2}{2\sqrt{q}} + \frac{\sqrt{6}r\sqrt{s}a^3k^6}{2\sqrt{q}} - 2ak\left(\frac{\sqrt{6}r\sqrt{s}k^5b^2}{\sqrt{q}} + \frac{\sqrt{6}r\sqrt{s}a^2k^5}{\sqrt{q}}\right) = 0,$$

$$-\frac{2\sqrt{6}r\sqrt{s}a^2k^6b^3a_1^2}{\sqrt{q}} - \frac{\sqrt{6}r\sqrt{s}ab^5k^6a_1^2}{2\sqrt{q}} + 2a_0\left(\frac{\sqrt{6}r\sqrt{s}b^5k^6a_2}{\sqrt{q}} + \frac{4\sqrt{6}r\sqrt{s}a^2k^6a_2b^3}{\sqrt{q}}\right) = 0,$$

$$\frac{7\sqrt{6}r\sqrt{s}bk^6a^2}{2\sqrt{q}} + \frac{\sqrt{6}r\sqrt{s}b^3k^6}{2\sqrt{q}} - 2bk\left(\frac{\sqrt{6}r\sqrt{s}k^5b^2}{\sqrt{q}} + \frac{\sqrt{6}r\sqrt{s}a^2k^5}{\sqrt{q}}\right) = 0,$$

$$\frac{3\sqrt{6}r\sqrt{s}a^3k^6}{2\sqrt{q}} - \frac{3\sqrt{6}r\sqrt{s}ak^6b^2}{2\sqrt{q}} = 0,\ \frac{2\sqrt{6}r\sqrt{s}ak^6b^4a_2a_0}{\sqrt{q}} - \frac{\sqrt{6}r\sqrt{s}ab^4k^6a_1^2}{2\sqrt{q}} = 0,$$

Solving above system, we get the following two solutions.

State 1:

$$\begin{aligned} &a = b,\ a_0 = \tfrac{a_1^2}{4a_2},\ a_1 = a_1,\ a_2 = a_2,\ b = b,\\ &k = k,\ \xi = \xi,\ l = -\tfrac{k\left(3s^2k^2+r^2\right)}{6s},\ r = r. \end{aligned} \tag{3.30}$$

State 2:

$$\begin{aligned} &a = -b,\ a_0 = \tfrac{a_1^2}{4a_2},\ a_1 = a_1,\ a_2 = a_2,\ b = b,\\ &k = k,\ \xi = \xi,\ l = -\tfrac{k\left(3s^2k^2+r^2\right)}{6s},\ r = r. \end{aligned} \tag{3.31}$$

Combining Eq. (3.28), Eq. (3.29) with Eq. (3.30) and Eq. (3.31), two families of exact explicit solutions to the mKdV-Burgers equation are obtained

$$u(x,t) = \frac{\sqrt{6}b^3k^3(\cosh(\beta)-\sinh(\beta))^3(3sk-r+3sbka_1\cosh(\beta)-3sbka_1\sinh(\beta))}{6\sqrt{s}\sqrt{q}(-bk\sinh(\alpha)+bk\cosh(\alpha))^3} + \frac{\sqrt{6s}}{\sqrt{q}}(-bk\sinh(\alpha)+bk\cosh(\alpha))a_2R(b\cosh(\alpha)-b\sinh(\alpha)+r)$$

and

$$u(x,t) = \frac{\sqrt{6}b^3k^3(\cosh(\beta)-\sinh(\beta))^3(-3sk-r+3sbka_1\cosh(\beta)+3sbka_1\sinh(\beta))}{6\sqrt{s}\sqrt{q}(bk\sinh(\alpha)+bk\cosh(\alpha))^3} + \frac{\sqrt{6s}}{\sqrt{q}}(bk\sinh(\alpha)+bk\cosh(\alpha))a_2R(-b\cosh(\alpha)-b\sinh(\alpha)+r)$$

where $\alpha = -kx + \frac{k(3s^2k^2+r^2)t}{6s} - \xi$, $\beta = \frac{-6kxs+3k^3ts^2+ktr^2-6s\xi}{6s}$.

We suppose that $w(x,t)$ be of the form

$$w(x,t) = A + B\exp(k_1x + l_1t + \xi_1) \tag{3.32}$$

where A, B, k_1 and l_1 are constants to be determined later and ξ_1 are an arbitrary constant. Setting Eq. (3.32) into Eqs. (3.26) and (3.27), we get the following system

$$\frac{2\sqrt{6}r\sqrt{s}B^2k_1^3a_2a_0}{\sqrt{q}} - \frac{\sqrt{6}r\sqrt{s}B^2k_1^3a_1^2}{2\sqrt{q}} = 0,$$

$$2sB^3k_1^3a_2a_0 - \frac{sB^3k_1^3a_1^2}{2} = 0,$$

$$-\frac{sBk_1^3}{2} - \frac{Bk_1r^2}{6s} - Bl_1 = 0,$$

Solving above system, one gets the following set of solution

$$\begin{array}{l} A = A,\ B = B,\ a_0 = \frac{a_1^2}{4a_2},\ a_1 = a_1,\ a_2 = a_2, \\ k_1 = k_1, l_1 = -\frac{k_1(3s^2k_1^2+r^2)}{6s},\ \xi_1 = \xi_1. \end{array} \tag{3.33}$$

Combining Eq. (3.28), Eq. (3.29) with Eq. (3.33), exact explicit solution is obtained

$$u(x,t) = \frac{\sqrt{6}\exp(\alpha)\left(3k_1s-r+3sBk_1a_1\exp\left(\frac{\alpha}{3}\right)\right)}{6\sqrt{sq}(\exp(\beta))^3} + \frac{\sqrt{6}\sqrt{s}Bk_1a_2\exp(\beta)R(A+B\exp(\beta))}{\sqrt{q}}$$

where $\alpha = -\frac{-6k_1xs+3k_1^3ts^2+k_1tr^2-6\xi_1s}{2s}, \beta = k_1x - \frac{k_1\left(3s^2k_1^2+r^2\right)t}{6s} + \xi_1$

4. Conclusions

In this paper, by introducing CRE method we apply to DMBBM and mKdV-Burgers equations. We had exact explicit solutions of given equations with the help of Riccati equation. The obtained exact solutions are consist of hyperbolic and exponential functions. We checked all solutions of given equations by the Maple.

It is also shown that the CRE method can be performed to other kinds of integrable systems and can be obtained other kind of solutions.

References

[1] M. Kaplan, A. Akbulut, A. Bekir, *Exact travelling wave solutions of the nonlinear evolution equations by auxiliary equation method*, Z. Naturforsch A, **70** (2015), 969–974.

[2] A. Bekir, A. Akbulut, M. Kaplan, *Exact solutions of nonlinear evolution equations by using modified simple equation method*, Int. J. Nonlinear Sci., **19** (2015), 159-164.

[3] F. Taşcan, A. Yakut, *Conservation laws and exact solutions with symmetry reduction of nonlinear reaction diffusion equations*, Int. J. Nonlinear Sci. Numer. Simul., **16** (2015), 191–196.

[4] M. Ekici, M. Mirzazadeh, Q. Zhou, S. P. Moshokoa, A. Biswas, M. Belic, *Solitons in optical metamaterials with fractional temporal evolution*, Optik, **127** (2016), 10879-10897.

[5] M. Mirzazadeh, M. Eslami, D. Milovic, A. Biswas, *Topological solitons of resonant nonlinear Schödinger'sequation with dual-power law nonlinearity by* (G'/G)*-expansion technique*, Optik, **125** (19), (2014) 5480-5489.

[6] Q. Feng, F. Meng, *Explicit solutions for space-time fractional partial differential equations in mathematical physics by a new generalized fractional Jacobi elliptic equation-based sub-equation method*, Optik, **127** (2016), 7450-7458.

[7] A. Akbulut, M. Kaplan, F. Taşcan, *The investigation of exact solutions of nonlinear partial differential equations by using exp*$(-\Phi(\xi))$ *method*, Optik, **132** (2017), 382-387.

[8] A. Biswas, M. Mirzazadeh, M. Eslami, D. Milovic, M. Belic, *Solitons in optical metamaterials by functional variable method and first integral approach*, Frequenz, **68** (11-12) (2014), 525-530.

[9] B. Lu, *The first integral method for some time fractional differential equations*, J. Math. Anal. Appl., **395** (2012), 684-693.

[10] H.M. Başkonuş, H. Bulut , *Analytical studies on the (1+1)-dimensional nonlinear Dispersive Modified Benjamin-Bona-Mahony equation defined by seismic sea waves*, Waves Random Complex Media, doi:10.1080/17455030.1062577, 2015.

[11] J. Xiang-Li, L. Sen-Yue, *CRE method for solving mKdV equation and new interactions between solitons and cnoidal periodic waves*, Commun. Theor. Phys., **63** (2015), 7-9.

[12] M. Chen, H. Hu, H. Zhu, *Consistent Riccati expansion and exact solutions of the Kuramoto-Sivashinsky equation*, Appl. Math. Lett., **49** (2015), 147-151.

[13] K. Khan, M.A. Akbar, S.M. Raynaul Islam , *Exact solutions for (1 + 1)-dimensional nonlinear dispersive modified Benjamin-Bona-Mahony equation and coupled Klein-Gordon equations*, SpringerPlus, **3** (2014), 8 pages.

[14] E.M.E. Zayed, S. Al-Joudi, *Applications of an extended (G'/G)-expansion method to find exact solutions of nonlinear PDEs in mathematical physics*, Math. Probl. Eng., Article ID 768573, doi:10.1155/2010/768573, (2010), 19 pages.

[15] M. Wang, *Exact solutions for a compound KdV-Burgers equation*, Phys. Lett. A., **213** (1996), 279-287.

Communications in Advanced Mathematical Sciences
Vol. II, No. 2, 114-120, 2019
Research Article
e-ISSN:2651-4001
DOI: 10.33434/cams.505485

The Graceful Coalescence of Alpha Cycles

Christian Barrientos[1]*, Sarah Minion[2]

Abstract
The standard coalescence of two graphs is extended, allowing to identify two isomorphic subgraphs instead of a single vertex. It is proven here that any succesive coalescence of cycles of size n, where n is divisible by four, results in an α-graph, that is, the most restrictive kind of graceful graph, when the subgraphs identified are paths of sizes not exceeding $\frac{n}{2}$. Using the coalescence and another similar technique, it is proven that some subdivisions of the ladder $L_n = P_2 \times P_n$ also admit an α-labeling, extending and generalizing the existing results for this type of subdivided graphs.

Keywords: Coalescence, α-labeling, Graceful labeling, Ladder.
2010 AMS: Primary 05C78, Secondary 05C76

[1] *Department of Mathematics, Valencia College, Orlando, Florida, USA, ORCID: 0000-0003-2838-8687*
[2] *Department of Mathematics, Valencia College, Orlando, Florida, USA, ORCID: 0000-0002-8523-3369*
***Corresponding author**: chr_barrientos@yahoo.com
Received: 30 December 2018, **Accepted:** 21 February 2019, **Available online:** 27 June 2019

1. Introduction

A *difference vertex labeling* of a graph G of size n is an injective mapping f from $V(G)$ into a set M of nonnegative integers, such that every edge uv of G has assigned a *weight* defined by $|f(u) - f(v)|$. All labelings considered in this work are difference vertex labelings. A labeling is called *graceful* when $M = \{0, 1, \ldots, n\}$ and the induced weights are $1, 2, \ldots, n$. If G admits such a labeling, it is called a *graceful graph*. Let G be a bipartite graph where $\{A, B\}$ is the natural bipartition of $V(G)$, we refer to A and B as *stable* sets of $V(G)$. A *bipartite labeling* of G is an injection $f : V(G) \to \{0, 1, \ldots, t\}$ for which there exists an integer λ, named the *boundary value* of f, such that $f(u) \leq \lambda < f(v)$ for every $(u, v) \in A \times B$, that induces n different weights. This is an extension of the definition given by Rosa and Širáň [1]. From the definition we conclude that the labels assigned by f on the vertices of A and B are in the interval $[0, \lambda]$ and $[\lambda + 1, t]$, respectively. When $t = n$, the function f is an *α-labeling* and G is an *α-graph*.

Let $f : V(G) \to \{0, 1, \ldots, t\}$ be a labeling of a graph G of size $n \leq t$. The labeling $g : V(G) \to \{c, c+1, \ldots, c+t\}$, defined for every $v \in V(G)$ and $c \in \mathbb{Z}$ as $g(v) = c + f(v)$, is the *shifting* of f in c units. Note that this labeling preserves the weights induced by f.

If f is bipartite with boundary value λ, the labeling $h : V(G) \to \{0, 1, \ldots, t+d-1\}$, defined for every $v \in V(G)$ and $d \in \mathbb{Z}^+$ as $h(v) = f(v)$ when $f(v) \leq \lambda$ and $h(v) = f(v) + d - 1$ when $f(v) > \lambda$, is the *bipartite d-labeling* of G obtained from f. This labeling uses labels from $[1, \lambda] \cup [\lambda + d, t + d - 1]$. In other terms, this labeling shifts the weights induced by f in $d - 1$ units. Thus, if f is an α-labeling of G and d is a positive constant, then h is a d-graceful labeling. This concept was introduced, independently by Maheo and Thuillier [2] and Slater [3] in 1982.

In the following sections we study α-labelings of the coalescence of α-cycles. Suppose that G_1 and G_2 are two graphs such that H is an induced subgraph of both of them. The *H-coalescence*, or simply *coalescence*, of G_1 and G_2, denoted by $G_1 \cdot G_2$, is the graph obtained by identifying the copy of H in G_1 with the copy of H in G_2. Assuming that for $i = 1, 2$, the graph G_i has order n_i and size m_i, and H has order p and size q, then $G_1 \cdot G_2$ has order n_1+n_2-p and size m_1+m_2-q. In Figure 1.1 we show an example of this operation where G_1 and G_2 are isomorphic and $H \cong C_4$, which is represented in the picture with green edges.

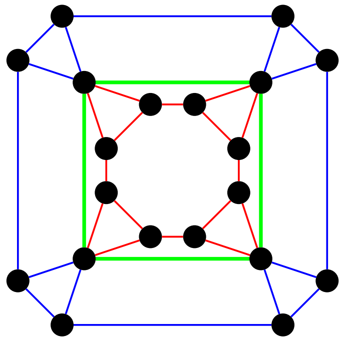

Figure 1.1. The C_4-coalescence of two isomorphic graphs of order 12 and size 20

In [4], Barrientos proved that if $H \cong P_1$ and G_1 and G_2 are α-graphs, there is a coalescence (also named one-point union or vertex amalgamation) of them is an α-graph. In [5], Barrientos and Minion showed that if $H \cong P_2$ and G_1 and G_2 are α-graphs, the coalescence (or edge amalgamation) of them is an α-graph if the edge of minimum weight in G_1 is identified with the edge of maximum weight of G_2. In this article, we extend the idea of the edge amalgamation, presented in [5], considering H to be a path. All graphs considered here are finite with no loops or multiple edges. We use the notation and terminology used in [6] and [7].

2. Preliminary results

In his seminal paper, Rosa [8] showed that when $n \equiv 0 (\text{mod } 4)$, there exists an α-labeling of the cycle C_n. We present here two labelings of C_n that are going to be used in the proof of the main result of the next section.

Suppose that $V(C_n) = \{v_1, v_2, \dots, v_n\}$ is the vertex set of C_n and its edge set is $\{v_i v_{i+1} : 1 \leq i \leq n\}$ where the addition is taken modulo n. The labelings f and g given below are two well-known α-labelings of C_n. The interested reader can easily verify this statement.

$$f(v_i) = \begin{cases} \frac{i-1}{2} & \text{if } i \text{ is odd and } 1 \leq i \leq \frac{n}{2}-1, \\ \frac{i+1}{2} & \text{if } i \text{ is odd and } \frac{n}{2}+1 \leq i \leq n-1, \\ n+1-\frac{i}{2} & \text{if } i \text{ is even.} \end{cases}$$

$$g(v_i) = \begin{cases} \frac{i-1}{2} & \text{if } i \text{ is even and } 2 \leq i \leq \frac{n}{2}-1, \\ \frac{i}{2} & \text{if } i \text{ is even and } \frac{n}{2}+2 \leq i \leq n, \\ n-\frac{i-1}{2} & \text{if } i \text{ is odd.} \end{cases}$$

In Figure 2.1 we show two examples for each of these labelings. The graphs on the first row are labeled using the function f, while g is used to label the graphs on the second row. The arrow inside the cycle shows the orientation of the vertices within each graph.

3. Graceful coalescence

Let $C_{n_1}, C_{n_2}, \dots, C_{n_k}$ be a collection of cycles, where the vertex set is $V(C_{n_j}) = \{v_{1,j}, v_{2,j}, \dots, v_{n_j,j}\}$ and the edge set is $E(C_{n_j}) = \{v_{1,j}v_{2,j}, v_{2,j}v_{3,j}, \dots, v_{n_k-1,j}v_{n_k,j}, v_{n_k,j}v_{1,j}\}$ for each $j \in \{1,2,\dots,k\}$. For every $j \in \{1,2,\dots,k-1\}$, select a positive integer t_j such that $2t_j \leq \min\{n_j, n_{j+1}\}$. A graph G is a *coalescence* of the cycles $C_{n_1}, C_{n_2}, \dots, C_{n_k}$ when the vertices $v_{1,j+1}, v_{2,j+1}, \dots, v_{t_j,j+1}$ of $C_{n_{j+1}}$ (together with the induced edges) are identified with the vertices $v_{n_j-t_j+1,j}, v_{n_j-t_j+2,j}, \dots v_{n_j,j}$ of C_{n_j}, respectively. Note that G is a graph of order

$$\sum_{j=1}^{k} n_j - \sum_{j=1}^{k-1} t_j = n_k + \sum_{j=1}^{k-1} (n_j - t_j)$$

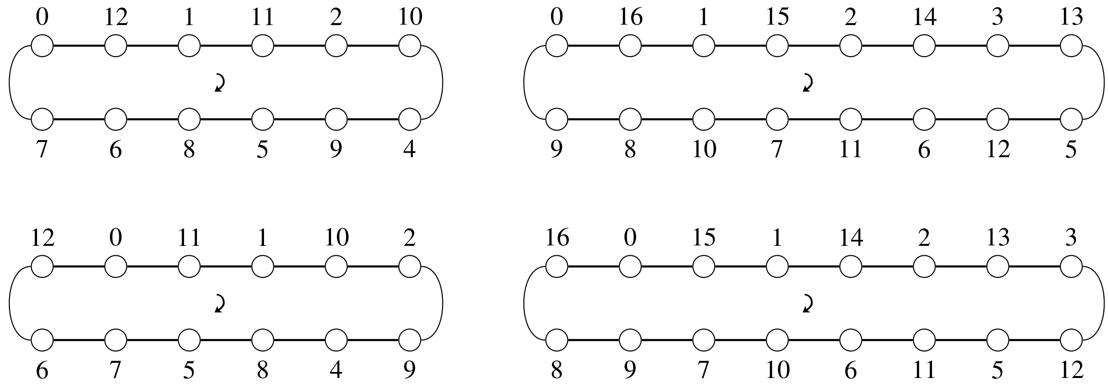

Figure 2.1. Two α-labelings of C_{12}, and C_{16}

and size

$$\sum_{j=1}^{k} n_j - \sum_{j=1}^{k-1}(t_j - 1) = n_k + \sum_{j=1}^{k-1}(n_j - t_j + 1)$$

In Figure 3.1 we show an example of this construction where $n_1 = 8$, $n_2 = 12$, $n_3 = 12$, $n_4 = 8$, and $n_5 = 4$, and $t_1 = 3$, $t_2 = 4$, $t_3 = 4$, and $t_4 = 2$. The numbers inside each cycle correspond to their respective vertices within that cycle.

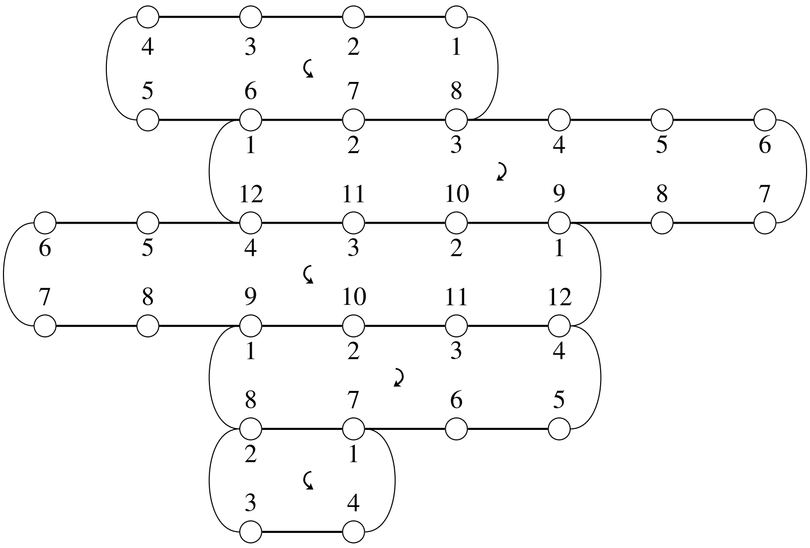

Figure 3.1. A coalescence of C_8, C_{12}, C_{12}, C_8 and C_4

We claim that when each $n_j \equiv 0 (\text{mod } 4)$, the coalescence G of the cycles $C_{n_1}, C_{n_2}, \ldots, C_{n_k}$, determined by $t_1, t_2, \ldots, t_{k-1}$, is an α-graph. Within the proof of this theorem we use the labelings f and g of C_n given in Section 2.

Theorem 3.1. *Let G be the coalescence of the cycles $C_{n_1}, C_{n_2}, \ldots, C_{n_k}$ determined by the integers $t_1, t_2, \ldots, t_{k-1}$, where $2t_j \le min\{n_j, n_{j+1}\}$. If for every $j \in \{1, 2, \ldots, k\}$, $n_j \equiv 0(mod\ 4)\}$, then G is an α-graph.*

Proof. Let G be the coalescence of the cycles $C_{n_1}, C_{n_2}, \ldots, C_{n_k}$ determined by the integers $t_1, t_2, \ldots, t_{k-1}$, where, for every $j \in \{1, 2, \ldots, k\}$, $2t_j \le min\{n_j, n_{j+1}\}$ and $n_j \equiv 0 (\text{mod } 4)\}$. Thus, every C_{n_j} admits an α-labeling.

To start, we label the vertices of C_{n_1} using the labeling f. For each $j \in \{2, 3, \ldots, k\}$, the selection of the initial labeling used on the vertices of C_{n_j} depends on the labeling used on $C_{n_{j-1}}$, according to the following criteria:

- If t_{j-1} is even, both $C_{n_{j-1}}$ and C_{n_j} have the same type of labeling.
- If t_{j-1} is odd, $C_{n_{j-1}}$ and C_{n_j} have different types of labelings.

Now that every cycle C_{n_j} has been α-labeled, we proceed to modify these initial labelings to obtain the desired α-labeling of G.

Recall that for every $j \in \{1,2,\ldots,k\}$, the size of the coalescence of the cycles $C_{n_1}, C_{n_2}, \ldots, C_{n_k}$ determined by the integers $t_1, t_2, \ldots, t_{k-1}$, is

$$\sum_{i=j}^{k} n_i - \sum_{i=j}^{k-1}(t_i - 1) = n_k + \sum_{i=j}^{k-1}(n_i - t_i + 1),$$

where the term $\sum_{i=j}^{k-1}(t_i - 1)$ is the number of edges shared by C_{n_j} and $C_{n_{j+1}}$.

The α-labeling of C_{n_j} is transformed into a d_j-graceful labeling (the intermediate labeling), where

$$d_j = \left(1 + n_k + \sum_{i=j}^{k-1}(n_i - t_i + 1)\right) - n_j.$$

In this way, the weights on the edges of C_{n_j} form the interval

$$I_j = \left[n_k + n_k + \sum_{i=j}^{k-1}(n_i - t_i + 1) - (n_j - 1), n_k + \sum_{i=j}^{k-1}(n_i - t_i + 1)\right].$$

Since $\min\{I_j : 1 \leq j \leq k\} = 1$ and $\max\{I_j : 1 \leq j \leq k\} = n_k + \sum_{i=1}^{k-1}(n_i - t_i + 1)$, that is, the size of G, we get

$$\bigcup_{j=1}^{k} I_j = [1, n_k + \sum_{i=1}^{k-1}(n_i - t_i + 1)].$$

Now, we need to shift these labelings to perform the coalescence of the labeled cycles. The labels assigned to the vertices of C_{n_1} constitute the final labeling of this cycle. For every $j \in \{2,3,\ldots,k\}$, the final labeling of C_{n_j} is obtained recursively in the following manner:

Assume that the labeling of $C_{n_{j-1}}$ is its final labeling. Let L_{j-1} be the set of the labels assigned to the vertices shared by $C_{n_{j-1}}$ and C_{n_j}. The final labeling of C_{n_j} is obtained by adding the constant $\min L_{j-1}$ to every label of C_{n_j}. Thus, the only overlapping of vertex labels between $C_{n_{j-1}}$ and C_{n_j}occurs on the vertices used to produce the coalescence.

Once this process has been completed, we have a bipartite labeling of G where the induced weights are $1, 2, \ldots, n_k + \sum_{j=1}^{k-1}(n_j - t_j + 1)$, with no label repeated.

Therefore, G is an α-graph. □

In Figure 3.2 we show the final α-labeling of the coalescence of the cycles C_{16}, C_{12}, C_8, C_{12}, C_8, determined by the integers $t_1 = 5$, $t_2 = 3$, $t_3 = 4$, $t_4 = 3$. The starting α-labelings of the cycles are: (0, 16, 1, 15, 2, 14, 3, 13, 5, 12, 6, 11, 7, 10, 8, 9), (12, 0, 11, 1, 10, 2, 9, 4, 8, 5, 7, 6), (0, 8, 1, 7, 3, 6, 4, 5), (0, 12, 1, 11, 2, 10, 4, 9, 5, 8, 6, 7), (8, 0, 7, 1, 6, 3, 5, 4). The intermediate d-graceful labelings are: (0, 45, 1, 44, 2, 43, 3, 42, 5, 41, 6, **40, 7, 39, 8, 38**), (**33, 0, 32, 1, 31**, 2, 30, 4, 29, **5, 28, 6**), (**0, 23, 1**, 22, **3, 21, 4, 20**), (**0, 18, 1, 17**, 2, 16, 4, 15, 5, **14, 6, 13**), (**8, 0, 7**, 1, 6, 3, 5, 4). The shifting constants are 7, 12, 15, and 21, respectively. The highlighted numbers correspond to the vertices that are going to be identified to produce the graph G.

Suppose that we want to form all nonisomorphic coalescence graphs constructed with k copies of C_n, where $n \equiv 0 \pmod 4$. How many of these graphs exist? Since two consecutive copies of C_n shared at most $\frac{n-2}{2}$ edges, any graph obtained by the coalescence of these cycles can be described by a sequence (or string) of numbers from $\{1, 2, \ldots, \frac{n-2}{2}\}$. Thus, counting nonismorphic coalescence graphs is equivalent to count nonoriented strings with $k-1$ beads of $\frac{n-2}{2}$ or fewer colors. This number is known and can be found in OEIS A277504 [9]. In the following table we show the first values for $n \in \{4, 8, 12, 16, 20\}$ and $k \in \{1, 2, \ldots, 11\}$.

In Figure 3.3 we show the α-labelings of the six graphs obtained using three copies of C_8.

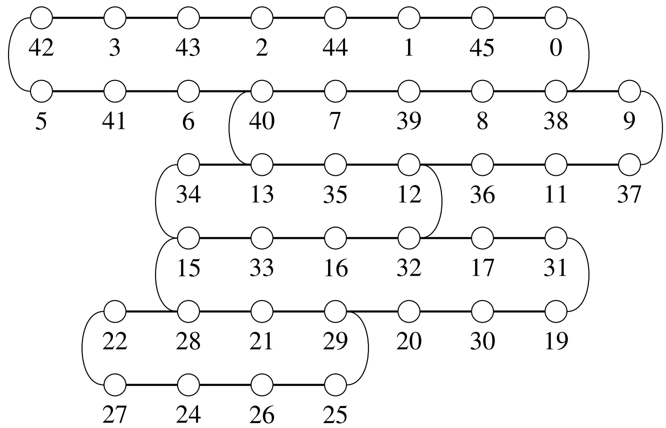

Figure 3.2. α-labeling of a coalescence of cycles

$k\backslash n$	4	8	12	16	20
1	1	1	1	1	1
2	1	3	5	7	9
3	1	6	15	28	45
4	1	18	75	196	405
5	1	45	325	1225	3321
6	1	135	1625	8575	29889
7	1	378	7875	58996	266085
8	1	1134	39375	412972	2394765
9	1	3321	195625	2883601	21526641
10	1	9963	978125	20185207	193739769
11	1	29643	4884375	141246028	1743421725

Table 1. Number of nonisomorphic coalescence graphs formed with k copies of C_8

4. Graceful subdivision of ladders

In this section we present two graceful labelings of subdivisions of ladders; the first result is a corollary of Theorem 3.1, the second one is a new construction. The *ladder* L_n is the result of the Cartesian product of the paths P_2 and P_n. The edges of P_2 within $L_n = P_2 \times P_n$ are called the *steps* of L_n. This type of graph can be seen as the coalescence of $n-1$ copies of C_4, therefore, L_n is an α-graph. The fact that L_n is graceful was proven first by Acharya and Gill [10].

In a graph G, a *subdivision* of an edge uv is the operation of replacing uv with a path u, w, v through a new vertex w. If the edhe uv is replaced with the path $u, w_1, w_2, \ldots, w_t, v$, we say that uv has been *subdivided an even* (resp. *odd*) *number of times* when t is even (resp. odd). Kathiresan [11] has shown that graphs obtained from ladders by subdividing each step exactly once are graceful.

If every step of L_n is subdivided an even number of times, then two consecutive subdivided steps, together with the two edges connecting them, form a cycle of size divisible by 4. Using Theorem 3.1, we can prove that this type of subdivided ladder is a graceful graph; in fact, it is an α-graph.

Corollary 4.1. *If every step of the ladder $L_n = P_2 \times P_n$ is subdivided an even number of times, the resulting graph is an α-graph.*

In Figure 4.1 we show, together with the original labeling of L_5, two examples of this subdivided ladder.

Unfortunately, the argument used in this corollary does not work when the edges are subdivided an odd number of times. So it is an open problem to find an α-labeling or a graceful labeling of these subdivided ladders.

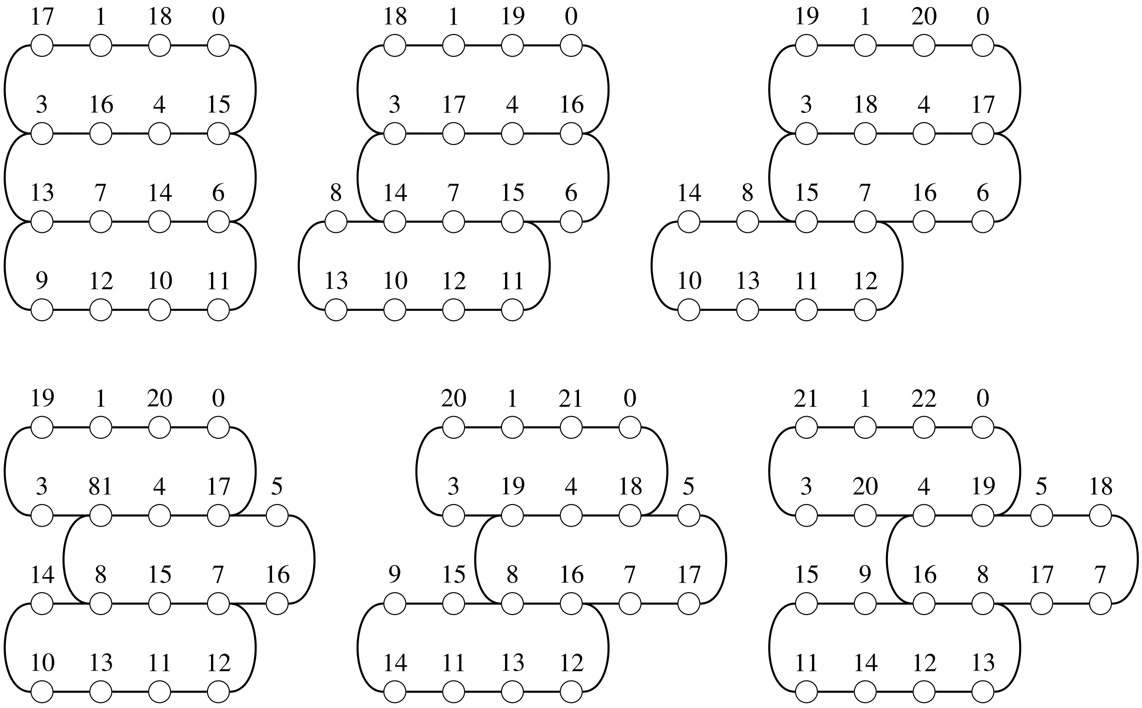

Figure 3.3. α-labelings of all the coalescences of three copies of C_8

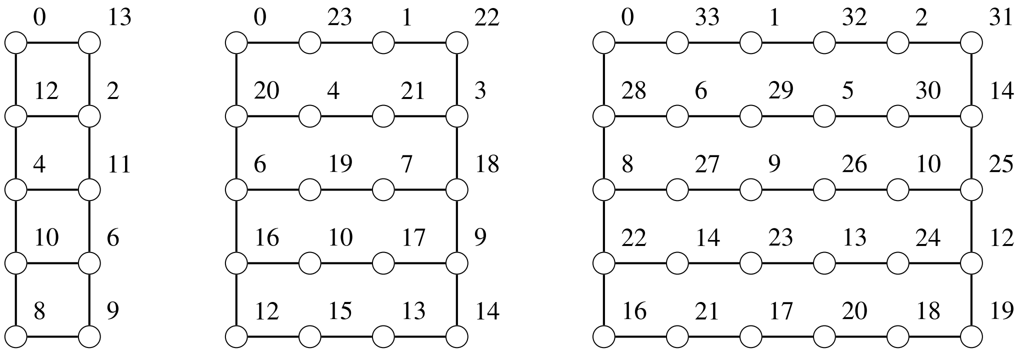

Figure 4.1. α-labeling of the ladder L_5 and some of its even subdivisions

Now we turn our attention to the graph obtained by subdividing every edge of L_n exactly once. We claim that this graph admits an α-labeling. Even when the resulting graph can be seen as the coalescence of $n-1$ copies of C_8, we use here a different construction based on the facts that C_8 and P_5 are α-graphs. The basic labelings of these graphs are given below:

- For C_8, the consecutive labels are: 5, 0, 8, 1, 7, 3, 6, 4.
- For P_5, the consecutive labels are: 2, 1, 3, 0, 4.

Suppose that G_n denotes the graph obtained by subdividing once all the edges of L_n. When $n = 2$, $G_2 \cong C_8$; we use on G_2 the α-labeling given above. To obtain an α-labeling of G_3 we transform the α-labeling of G_2 by shifting its weights in such a way that the new largest label is 14 = 8 + 6, that is, the size of G_2 plus 6. The labeling of P_5 is shifted $\lambda_2 + 2$ units, where λ_2 is the boundary value of the α-labeling of G_2. The vertices $\lambda_2 - 1$ and λ_2 in G_2 are connected to the vertices to the vertices $\lambda_2 + 4$ and $\lambda_2 + 6$ in P_5; thus, the new edges have weights 5 and 6, respectively. The resulting graph is G_3 with an α-labeling. We continue this process until G_n has been labeled. In this way we have proved the following theorem.

Theorem 4.2. *The graph G, obtained by subdividing every edge of the ladder L_n exactly once, is an α-graph.*

In Figure 4.2 we show the first four cases of this construction.

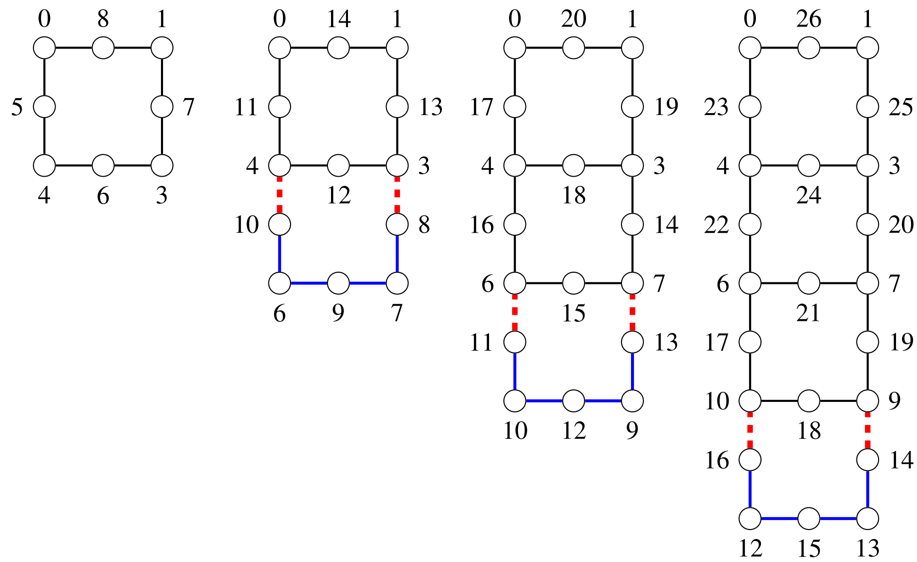

Figure 4.2. α-labelings of subdivided ladders

References

[1] A. Rosa, J. Širáň, *Bipartite labelings of trees and the gracesize*, J. Graph Theory, **19** (1995), 201-215.

[2] M. Maheo, H. Thuillier, *On d-graceful graphs*, Ars Combin., **13** (1982), 181-192.

[3] P. J. Slater, *On k-graceful graphs*, Proceedings of the 13th Southeastern International Conference on Combinatorics, Graph Theory and Computing, (1982), 53-57.

[4] C. Barrientos, *Graceful labelings of chain and corona graphs*, Bull. Inst. Combin. Appl., **34** (2002), 17-26.

[5] C. Barrientos, S. Minion, *Alpha labelings of snake polyominoes and hexagonal chains*, Bull. Inst. Combin. Appl., **74** (2015), 73-83.

[6] G. Chartrand, L. Lesniak, Graphs & Digraphs, 2nd ed. Wadsworth & Brooks/Cole, 1986.

[7] J. A. Gallian, *A dynamic survey of graph labeling*, Electronic J. Combin., **21**(#DS6), 2018.

[8] A. Rosa, *On certain valuations of the vertices of a graph*, Theory of Graphs (International Symposium, Rome, July 1966), Gordon and Breach, N. Y. and Dunod Paris (1967), 349-355.

[9] J.-F. Sloane, https://oeis.org/A277504, 2016.

[10] B. D. Acharya, M. K. Gill, *On the index of gracefulness of a graph and the gracefulness of two-dimensional square lattice graphs*, Indian J. Math., **23** (1981), 81-94.

[11] K. Kathiresan, *Subvisions of ladders are graceful*, Indian J. Pure Appl. Math., **23** (1992), 21-23.

Communications in Advanced Mathematical Sciences
Vol. II, No. 2, 121-128, 2019
Research Article
e-ISSN:2651-4001
DOI: 10.33434/cams.507917

Extending the Applicability of a Newton-Kurchatov-Type Method for Solving Non-Differentiable Equations in Banach Spaces

Ioannis K. Argyros[1]*, Santhosh George[2]

Abstract
We provide a new local convergence analysis of a Newton-Kurchatov-like method to solve non-differentiable equations in Banach spaces. Our result improve the earlier works in literature. The examples were used to test our hypotheses.

Keywords: Banach spaces, Local convergence, Newton-Kurchatov-type method, Non-differentiable equations
2010 AMS: Primary 65G99, 65H10, 65J15, 49M15.

[1] *Department of Mathematical Sciences, Cameron University, Lawton, OK 73505, USA, iargyros@cameron.edu, ORCID:0000-0002-9189-9298*
[2] *Department of Mathematical and Computational Sciences, National Institute of Technology Karnataka, India-575 025, sgeorge@nitk.ac.in, ORCID: 0000-0002-3530-5539*
***Corresponding author**: sgeorge@nitk.ac.in

Received: 4 January 2019, **Accepted:** 28 March 2019, **Available online:** 27 June 2019

1. Introduction

In this work, we solve

$$H(x) = 0, \tag{1.1}$$

to find a solution $x^* \in \Omega$, where $H : \Omega \subseteq X \longrightarrow Y$ and X, Y stand for Banach spaces. Iterative methods are mostly used to solve (1.1), since solutions in closed form are hard to find. If H is a differentiable operator, Newton's method is the most used method to solve the equation of (1.1), which is given by [1,2]

$$x_{n+1} = x_n - H'(x_n)^{-1}H(x_n), \text{ for all } n = 0,1,2,\ldots\, x_0 \in \Omega. \tag{1.2}$$

If H is not differentiable, Remember that an operator $[x,y;H] \in L(X,Y)$ is called a divided difference of order one for the operator H on the points x and y $(x \neq y)$ if the following equality holds:

$$[x,y;H] = H(x) - H(y).$$

Replacing the Fréchet derivative H' by divided differences of the operator H in Newton's method (1.2) at different points, we can define two iterative methods as follows: one is the secant method is given by [3,4]

$$x_{n+1} = x_n - [x_{n-1},x_n;H]^{-1}H(x_n) \quad n \geq 0,\; x_0, x_{-1} \in \Omega,$$

and the other is Kurchatov's method given by [5]

$$x_{n+1} = x_n - [x_{n-1}, 2x_n - x_{n-1}; H]^{-1} H(x_n) \quad n \geq 0,\ x_0, x_{-1} \in \Omega.$$

Note that, Kurchatov's method is as simple as Newton's method and has the same rate of convergence as Newton's method. This means it has the higher rate of convergence than the Secant method. A lot of study about the convergence of Kurchatov's method have been given, see [5–10].

We split it as

$$H(x) = F(x) + G(x),$$

where $F : \Omega \longrightarrow Y$ and $G : \Omega \longrightarrow Y$. F is differentiable and G is continuous but non-differentiable. Then, we use the following Newton-Kurchatov-type method given by

$$x_{n+1} = x_n - (F'(x_n) + [x_{n-1}, 2x_n - x_{n-1}; G])^{-1} H(x_n) \quad n \geq 0,\ x_0, x_{-1} \in \Omega \tag{1.3}$$

to solve (1.1). Recently, M. A. Hernández and M. J. Rubio [8] gave an analysis of method (1.3). Cases where method (1.3) is efficient for solving systems and also arguments about its efficiency were also given in [8]. A novel idea of [8] is that the usual condition of $H'(x^\star)$ is reduced to a new type condition, which means that H can be a non-differentiable operator. We give a more precise local analysis for (1.3) than [8]. Advantages of our local convergence analysis over the work, in [8] :

(a) Larger radius of convergence lending to wider choice of initial guesses,

(b) More precise estimates on the distances $\|x_{n+1} - x^\star\|$. Hence fewer iterates are need to obtain a desired error tolerance.

(c) At least as precise information on the uniqueness ball of the solution.

These advantages are obtained under the same computational cost, since in practice the new majorizing functions are special cases of the majorizing functions in [8].

The paper is organized as follows: Section 2 contains the local convergence analysis of method (1.3). The numerical examples including favorable comparisons with earlier study [8] are presented in the concluding Section 3.

Throughout the paper we denote $B(x,\rho) = \{y \in X : \|y - x\| < \rho\}$ and $\overline{B}(x,\rho) = \{y \in X : \|y - x\| \leq \rho\}$.

2. Local convergence analysis

From now on by differentiable operator, we mean differentiable in the sense of Fréchet. We shall use condition (C) to show the local convergence analysis of the Kurchatov-type method (1.3):

(C_1) $F : \Omega \subset X \to Y$ is continuously differentiable operator.

(C_2) $G : \Omega \subset X \to Y$ is continuous and a divided difference of order one $[\cdot, \cdot; G] : \Omega \times \Omega \to L(X,Y)$, exists.

(C_3) There exist $x^\star \in \Omega$ and $\bar{x} \in \Omega$ with $\|\bar{x} - x^\star\| = \delta > 0$ such that $H(x^\star) = 0$ and $M = F'(x^\star) + [x^\star, \bar{x}; G]$ is invertible.

(C_4) $\|M^{-1}\left(F'(x) - F'(x^\star)\right)\| \leq v_0(\|x - x^\star\|)$ for some function $v_0 : [0,+\infty) \to [0,+\infty)$ continuous, nondecreasing with $v_0(0) = 0$ and each $x \in \Omega$.

(C_5) $\|M^{-1}\left([y, 2x - y; G] - [x^\star, \bar{x}; G]\right)\| \leq w_0\left(\|y - x^\star\|, \|2x - y - \bar{x}\|\right)$ for some function $w_0 : [0,+\infty) \times [0,+\infty) \to [0,+\infty)$ continuous, nondecreasing for each x, y, $2x - y \in \Omega$.

(C_6) Equation $v_0(t) + w_0(t, \delta + t) = 1$ has a minimal positive solution $\overline{r_0}$. Pick $r_0 \in (0, \overline{r_0}]$. Define

$$p_0 := v_0(r_0) + w_0(r_0, \delta + r_0) < 1.$$

Let $\Omega_0 = \Omega \cap B(x^\star, r_0)$.

(C_7) $\|M^{-1}\left(F'(y)-F'(x)\right)\| \le v(\|y-x\|)$ for function $v:[0,r_0)\to[0,+\infty)$ continuous, nondecreasing with $v(0)=0$ and all $x,\, y\in\Omega_0$.

(C_8) $\|M^{-1}\left([y,\,2x-y;\,G]-[x^\star,\,x;\,G]\right)\| \le w\big(\|y-x^\star\|,\|x-y\|\big)$ for some function $w:[0,r_0)\times[0,r_0)\to[0,+\infty)$ continuous, nondecreasing for each $x,\,y,\,2x-y\in\Omega_0$.
Let $p=\int_0^1 v(\theta\, r_0)\,d\theta + w(r_0,2r_0)$ and $q=\frac{p}{1-p_0}$.

(C_9) Let $\bar r\ge 0$, there exists minimal $r\in(0,r_0)$ solving the equation

$$\bar r = t\,[1-\frac{2(\int_0^1 v(\theta\, t)d\theta + w(t,2t))}{1-\big(v_0(t)+w_0(t,\delta+t)+2\big(\int_0^1 v(\theta\, t)d\theta+w(t,2t)\big)\big)}].$$

Notice that $r>\bar r$.

(C_{10}) $B(x^\star,r)\subseteq\Omega$.

(C_{11})

$$\int_0^1 v_0(\theta\, r^\star)\,d\theta + w_0(0,\delta+r^\star)<1$$

for some $r^*\ge r$.

First, we need a perturbation result.

Lemma 2.1. *Assume* $(C_1)-(C_6)$. *Then, operator* $F'(x)+[y,\,2x-y;\,G]$ *is invertible for all* $x,y,2x-y\in B(x^\star,r_0)$ *with* $x\ne y$ *and*

$$\|\big(F'(x)+[y,2x-y;G]\big)^{-1}M\| \le \frac{1}{1-p_0}.$$

Proof. Operator $[y,2x-y;G]$ is well defined, since $y\ne 2x-y$. Using $(C_3)-(C_5)$, we have in turn that

$$\begin{aligned}
&\|M^{-1}\big(M-F'(x)-[y,2x-y;G]\big)\|\\
&\le \|M^{-1}\big(F'(x^\star)-F'(x)\big)\| + \|M^{-1}\big([x^\star,\bar x;G]-[y,2x-y;G]\big)\|\\
&\le v_0(\|x^\star-x\|)+w_0\big(\|x^\star-y\|,\|\bar x-(2x-y)\|\big)\\
&\le v_0(r_0)+w_0(r_0,\delta+r_0)=p_0<1.
\end{aligned}$$

The result follows from the preceding estimate and the Banach lemma on invertible operators [1,2]. ☐

Secondly, we establish the sequence $\{x_n\}$ generated by the Newton-Kurchatov-type method (1.3) is well defined.

Lemma 2.2. *Suppose the conditions* $(C_1)-(C_8)$ *hold,* $x_{n-1},x_{n-2},2x_{n-1}-x_{n-2}\in B(x^\star,r_0)$ *and* $x_{n-1}\ne x_{n-2}$, *then* x_n *is well defined and*

$$\|x_n-x^\star\|\le q\,\|x_{n-1}-x^\star\|.$$

Proof. We shall use the notation

$$M_{n-1}=F'(x_{n-1})+[x_{n-2},\,2x_{n-1}-x_{n-2};G].$$

Newton-Kurchatov-type method (1.3) gives

$$\begin{aligned}
x_n-x^\star &= x_{n-1}-x^\star-M_{n-1}^{-1}H(x_{n-1})\\
&= M_{n-1}^{-1}\left(M_{n-1}(x_{n-1}-x^\star)-H(x_{n-1})\right)\\
&= M_{n-1}^{-1}\Big(\big(F'(x_{n-1})(x_{n-1}-x^\star)-F(x_{n-1})+F(x^\star)\big)\\
&\quad +G(x^\star)-G(x_{n-1})+[x_{n-2},2x_{n-1}-x_{n-2};G]\,(x_{n-1}-x^\star)\Big)\\
&= -[M_{n-1}^{-1}M]\,[M^{-1}\int_0^1\big(F'(x_{n-1}+\theta\,(x^\star-x_{n-1}))-F'(x_{n-1})\big)\,(x_{n-1}-x^\star)\,d\theta]\\
&\quad +[M_{n-1}^{-1}M]\,[M^{-1}\big([x_{n-2},2x_{n-1}-x_{n-2};G]-[x^\star,x_{n-1};G]\big)\,(x_{n-1}-x^\star)].
\end{aligned}$$

Using Lemma 2.1, (C_7), (C_8) and the triangle inequality in the preceding identity,

$$\begin{aligned}\|x_n - x^\star\| &\leq \|M_{n-1}^{-1} M\| \Big(\int_0^1 v(\theta\, \|x^\star - x_{n-1}\|)\, d\theta \\ &\quad + w\big(\|x_{n-2} - x^\star\|, \|x_{n-1} - x_{n-2}\|\big)\Big)\, \|x_{n-1} - x^\star\| \\ &\leq \frac{1}{1-p_0} \Big(\int_0^1 v(\theta\, r_0)\, d\theta + w(r_0, 2r_0)\Big)\, \|x_{n-1} - x^\star\| \\ &= \frac{p}{1-p_0}\, \|x_{n-1} - x^\star\| = q\, \|x_{n-1} - x^\star\|.\end{aligned}$$

□

Let $\bar{r} = \|x_0 - x^\star\|$. As in [7], [11], we must somehow drop $2x_{n-1} - x_{n-2} \in B(x^\star, r_0)$, if $x_{n-1}, x_{n-2} \in B(x^\star, r_0)$. Suppose that $x_1, x_2, \cdots, x_{n-1} \in B(x^\star, r_0)$, then

$$\begin{aligned}\|2x_{n-1} - x_{n-2} - x^\star\| &\leq \|x_{n-1} - x_{n-2}\| + \|x_{n-1} - x^\star\| \\ &\leq 2\, \|x_{n-1} - x^\star\| + \|x_{n-2} - x^\star\| \\ &\leq (2\,q + 1)\, \|x_{n-2} - x^\star\|\end{aligned}$$

and

$$\|x_{n-2} - x^\star\| \leq q^{n-2}\, \|x_0 - x^\star\|.$$

Then, if $q < 1$, we have

$$\|x_{n-2} - x^\star\| < \|x_0 - x^\star\| = \bar{r}$$

and

$$\|2x_{n-1} - x_{n-2} - x^\star\| < (2q+1)\, \bar{r}.$$

Clearly, if $p_0 + p < 1$, then $q < 1$. To show $2x_{n-1} - x_{n-2} \in B(x^\star, r_0)$, it suffices to have $(2q+1)\,\bar{r} = r$ leading to the condition (C_9).

Theorem 2.3. *Assume (C) with $p_0 + p < 1$. Then, sequence $\{x_n\}$ generated by the Kurchatov-type method (1.3) for $x_0 \in B(x^\star, r) - \{x^\star\}$ and $x_{-1} \in B(x_0, r - \bar{r})$ with $x_{-1} \neq x_0$ and $\bar{r} = \|x_0 - x^\star\|$ exists in $B(x^\star, r)$, stayes in $B(x^\star, r)$ for all $n = 0, 1, 2, \cdots$ and $\lim_{n \longrightarrow \infty} x_n = x^\star$.*

Proof. Notice

$$\|x_{-1} - x^\star\| \leq \|x_{-1} - x_0\| + \|x_0 - x^\star\| \leq r - \bar{r} + \bar{r} = r$$

and

$$\|2x_0 - x_{-1} - x^\star\| \leq \|x_{-1} - x_0\| + \|x_0 - x^\star\| \leq r - \bar{r} + \bar{r} = r,$$

so $x_{-1}, 2x_0 - x_{-1} \in B(x^\star, r)$ and $2x_0 - x_{-1} \neq x_1$. By Lemma 2.1, x_1 exists and by Lemma 2.2

$$\|x_1 - x^\star\| \leq q\, \|x_0 - x^\star\| < \|x_0 - x^\star\| = \bar{r} < r,$$

so, $x_1 \in B(x^\star, r)$ and $x_1 \neq x_0$. Analogously,

$$\begin{aligned}\|2x_1 - x_0 - x^\star\| &\leq \|x_1 - x_0\| + \|x_1 - x^\star\| \\ &\leq 2\|x_1 - x^\star\| + \|x_0 - x^\star\| < (2q+1)\, \|x_0 - x^\star\| < r,\end{aligned}$$

so $2x_1 - x_0 \in B(x^\star, r)$. Assume for $k \geq 2$, if $x_{k-1}, x_{k-2} \in B(x^\star, r)$ for $x_{k-1} \neq x_{k-2}$, then $2x_{k-1} - x_{k-2} \in B(x^\star, r)$ and hence M_{k-1}^{-1} is well defined. Then, x_k is well defined and from Lemma 2.2, we get that

$$\|x_k - x^\star\| \leq q\|x_{k-1} - x^\star\|.$$

Suppose using mathematical induction that the preceding two inequalities hold for $k = 2,3,\cdots,m$, we shall show that they hold for $k = m+1$. If $x_k, x_{k-1} \in B(x^\star, r)$ with $x_k \neq x_{k-1}$, we get that

$$\begin{aligned}\|2x_k - x_{k-1} - x^\star\| &\le \|x_k - x_{k-1}\| + \|x_k - x^\star\| \\ &\le 2\|x_k - x^\star\| + \|x_{k-1} - x^\star\| < (2q+1)\|x_{k-1} - x^\star\| \\ &\le (2q+1)\, q^{k-1}\|x_0 - x^\star\| < (2q+1)\|x_0 - x^\star\| < r,\end{aligned}$$

so $2x_k - x_{k-1} \in B(x^\star, r)$ and $2x_k - x_{k-1} \neq x_{k-1}$. That is by Lemma 2.1 M_k^{-1} exists and x_{k+1} is well defined. Moreover, by Lemma 2.2,

$$\|x_{k-1} - x^\star\| \le q\|x_k - x^\star\|$$

which completes the induction. That is $\{x_k\} \subseteq B(x^\star, r)$ and

$$\|x_k - x^\star\| \le q^k\|x_0 - x^\star\|,$$

from which we deduce that $\lim_{k\to\infty} x_k = x^\star$. □

Next, a uniqueness result is given.

Proposition 2.4. *Assume conditions (C). Then, $x^\star$ is the only solution of equation $H(x) = 0$ in $\Omega_1 = \Omega \cap \overline{B}(x^\star, r^\star)$.*

Proof. Let $y^\star \in \Omega_1$ with $H(y^\star) = 0$. Define operator $T = \int_0^1 F'(x^\star + \theta(y^\star - x^\star))\, d\theta + [x^\star, y^\star; G]$. Then, using (C_4), (C_5) and (C_{11}).

$$\begin{aligned}\|M^{-1}(T - M)\| &\le \|\int_0^1 M^{-1}\Big(F'(x^\star + \theta(y^\star - x^\star))\Big)\, d\theta\| \\ &\quad + \|M^{-1}([x^\star, y^\star; G] - [x^\star, \bar{x}; G])\| \\ &\le \int_0^1 v_0(\theta\|x^\star - y^\star\|)\, d\theta + w_0(\|x^\star - x^\star\|, \|y^\star - \bar{x}\|) \\ &\le \int_0^1 v_0(\theta r^\star)\, d\theta + w_0(0, r^\star + \delta) < 1,\end{aligned}$$

so, T^{-1} exists.
But from

$$0 = H(x^\star) - H(y^\star) = T(x^\star - y^\star),$$

we conclude that $y^\star = x^\star$. □

Remark 2.5. (a) *We can set $\bar{x} = x_0$. In this case $\delta = \bar{r}$.*

(b) *If $\Omega = X$, condition $2x - y \in \Omega$ is automatically satisfied. To relax this condition, let*

$$p_1 = v_0(r_0) + w_0(r_0, \delta + 3r_0).$$

Then, we use the condition $p_1 + p < 1$, instead of using (C_9) to calculate r, or the equation

$$v_0(t) + + w_0(t, \delta + 3t) + 2\Big(\int_0^1 v(\theta t)\, dt + w(t, 2t)\Big) = 1. \tag{2.1}$$

Note that in this case $q_1 = \frac{p}{1-p_1} < 1$. Hence, we arrived at:

Proposition 2.6. *Assume conditions $(C_1) - (C_8)$, and*

$(C_9)'$ *There exists a solution $R \in (0, r_0)$ of equation (2.1)*

$(C_{10})'$ *For* $R_1 < R$, $B(x^\star, R_1) \subset \Omega$

$(C_{11})'$ $2x - y \in \Omega$ *for all* $x, y \in B(x^\star, R_1)$ *with* $x \neq y$

(C_{12})

$$\int_0^1 v_0(\theta R^\star)\, d\theta + w_0(0, \delta + R^\star) < 1$$

for some $R^\star \geq R_1$. *Then, sequence* $\{x_n\}$ *generated for* $x_0 \in B(x^\star) - x^\star$ *and* $x_1 \in B(x_0, R_1 - \bar{r})$ *with* $x_{-1} \neq x_0$ *by the Newton-Kurchatov-type method (1.3) exists in* $B(x^\star, R_1)$, *stays in* $B(x^\star, R_1)$ *for all* $n = 0, 1, \cdots$ *and* $\lim_{n \longrightarrow \infty} x_n = x^\star$, *which is the only solution of equation* $H(x) = 0$ *in* $\Omega_2 = \Omega \cap \overline{B}(x^\star, R_1)$.

Remark 2.7. *Clearly condition* $(C_{11})'$ *can be exchanged by*

$$(C_{10})'' B(x^\star, 3R) \subseteq \Omega,$$

since if $x, y \in B(x^\star, R) \Rightarrow \|x^\star - (2x - y)\| \leq 2\|x^\star - x\| + \|x^\star - y\| < 3R \Rightarrow 2x - y \in B(x^\star, 3R)$ *(see also [5–9] and the numerical examples).*

Remark 2.8. *The results in this study improve the corresponding ones in [8]. Indeed, we have the following advantages:*

(1) Affine invariant results are given here which are more advantageous than non affine results given in [8].
(2) The following conditions have been used in [8]

(h_7) $\|F'(y) - F'(x)\| \leq \bar{v}(\|y - x\|)$ for all $x, y \in \Omega$,

(h_8) $\|[x, y; G] - [u, v; G]\| \leq \overline{w}(\|x - u\|, \|y - v\|)$ for all $x, y, u, v \in \Omega$,

$(h_6)'$ $\overline{p} = \gamma\Big(\int_0^1 \bar{v}(\theta\rho)\, d\theta + \overline{w}(\rho, 2\rho)\Big)$, $\overline{q} = \frac{\overline{p}}{1 - \overline{p_0}}$, $\|M^{-1}\| \leq \gamma$,

$$\overline{p_0} = \gamma(\bar{v}(\rho) + w(\rho, \rho + \delta)) < 1,$$

$$\bar{r} = t\,[1 - \frac{2\gamma\Big(\int_0^1 \bar{v}(\theta t)\, d\theta + \overline{w}(t, 2t)\Big)}{1 - \gamma\Big(\bar{v}(t) + \overline{w}(t\gamma, \delta + t) + 2\gamma\Big(\int_0^1 \bar{v}(\theta t)\, d\theta + \overline{w}(t, 2t)\Big)\Big)}]$$

and

(h_9) $B(x^\star, \rho) \subseteq \Omega$,

(h_{10}) There exists $\rho^\star \geq \rho$ such that

$$\gamma\Big(\int_0^1 \bar{v}(\theta\rho^\star)\, d\theta + \overline{w}(0, \delta + \rho^\star)\Big) < 1.$$

However, we have that

$$v_0(t) \leq \gamma\bar{v}(t), v(t) \leq \gamma\bar{v}(t)$$

$$w_0(s, t) \leq w(s, t) \leq \gamma\overline{w}(s, t)$$

$$q \leq \overline{q}$$

$$\rho \leq r$$

and

$$\rho^\star \leq r^\star$$

which lead to the improvements listed in the introduction. It is worth noticing that improvements are given using the same computational cost, because in practice the computation of functions $\bar{v}$, $\overline{w}$ needs the computation of the functions v_0, v, w_0, w as special cases.

Remark 2.9. *Let us see the radii for Newton's method (1.2), i.e., when the $\bar{v}$ and the the $\bar{v}_0$ functions are choosen by $\bar{v}(t) = \mu t$ and $\bar{v}_0(t) = \lambda t$, $G = 0, w = w_0 = 0$ and $x^\star = \bar{x}$ (i.e, $\delta = 0$). The radius ρ given in [8] is*

$$\rho = \frac{2}{3\gamma\mu}.$$

The radius ρ coincides with radius given independently by Rheinboldt [12] and Traub [13]. This value improves the radius

$$\rho_0 = \frac{1}{2\gamma\mu},$$

given also by Dennis and Schnabel [12, 13]. Our radius of convergence r is given by

$$r = \frac{2}{(2\lambda + \mu_0)\gamma}.$$

Then, we have that

$$\rho_0 \le \rho \le r. \tag{2.2}$$

The right hand side inequality in (2.2) can be strict (see (c4), (h7) and the numerical examples).

3. Numerical examples

Choose the divided difference $[x, y; F] = \int_0^1 F'(y + \theta(x - y))d\theta$.

Example 3.1. **Case 1** *Newton's method. Let F, G be defined on $\Omega = [-1,1] \times [-1,1] \times [-1,1]$ by*

$$F(x,y,z) = (e^x - 1, \frac{(e-1)y^2}{2} + y, z)^T, \quad \text{and} \quad G = 0. \tag{3.1}$$

*Choose $\lambda = e - 1, \mu_0 = e^{\frac{1}{e-1}}$, $\mu = e$ for $x^\star = (0,0,0)^T$ and $\gamma = 1$
we have*

$$\rho_0 = 0.1839 < \rho = 0.2453 < r = 0.3827.$$

Newton's method is very efficient. In general, if the method is inefficient, then we use a better method. The new error bounds are also better, since

$$\|x_{n+1} - x^\star\| \le \frac{\mu_0 \|x_n - x^\star\|^2}{2(1 - \lambda \|x_n - x^\star\|)} \quad n = 1, 2, \dots$$

and

$$\|x_1 - x^\star\| \le \frac{\lambda \|x_0 - x^\star\|^2}{2(1 - \lambda \|x_n - x^\star\|)}$$

but λ the old ones are given by

$$\|x_{n+1} - x^\star\| \le \frac{\mu \|x_n - x^\star\|^2}{2(1 - \mu \|x_n - x^\star\|)}, n = 0, 1, 2, 3, \cdots.$$

The old uniqueness ball is $B(x^\star, \frac{2}{e})$. The new uniqueness ball is $B(x^\star, \frac{2}{e-1})$ is better, since

$$B(x^\star, \frac{2}{e}) \subseteq B(x^\star, \frac{2}{e-1}).$$

Case 2 *Newton-Kurchatov-type method. Let F be given as in (3.1) and define $G(x) = |x|$. We have for $\bar{x} = (0.01, 0.01, 0.01)^T, \gamma = \frac{1}{2}, \delta = 0.01$, $v_0(t) = (e-1)t$, $v(t) = e^{\frac{1}{e-1}}t, w_0(s,t) = w(s,t) = 1, \bar{r} = \delta$ and $v_0(t) < v(t)$. Then*

$$r_{old} = 0.4905 < r = 0.7654.$$

References

[1] I. K. Argyros, *Computational theory of iterative methods, Series: Studies in Computational Mathematics 15*, C.K. Chui and L. Wuytack (editors), Elservier Publ. Co. New York, USA, 2007.

[2] J.M. Ortega, W.C. Rheinbolt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[3] I. K. Argyros, *On the Secant method*, Publ. Math. Debrecen, **43** (1993), 223-238.

[4] F. A. Potra, V. Pták, *Nondiscrete Induction and Iterative Methods*, Pitman Publishing Limited, London, 1984.

[5] V. A. Kurchatov, *On the method of linear interpolation for the solution of functional equations*, (Russion) Dolk. Akad. Nauk SSSR, 1998 (1971) 524-526, translation in Soviet Math. Dolk., 12 (1971) 835-838.

[6] I. K. Argyros, *On the two point Newton-like methods of convergent R-order two*, Int. J. Comput. Math., **82** (2005), 219-233.

[7] I. K. Argyros, *A Kantorovich-type analysis for a fast iterative method for solving nonlinear equations*, J. Math. Anal. Appl. **332** (2007), 97-108.

[8] M. A. Hernández, M. J. Rubio, *On the local convergence of a Newton-Kurchatov-type method for non-differentiable operators*, Appl. Math. Comput., **304** (2017), 1-9.

[9] S. Shakhno, *On the Secant method under generalized Lipschitz conditions for the divided operator*, PAMM-Proc. Appl. Math. Mech., **7** (2007), 2060083-2060084.

[10] A. Cordero, F. Soleymani, J. R. Torregrosa, F. K. Haghani, *A family of Kurchatov-type methods and its stability*, Appl. Math. Comput., **294** (2017), 264-279.

[11] I. K. Argyros, *On a quadratically convergent iterative method using divided differences of order one*, J. Korean. Math. S. M. E. Ser. B, **14** (3) (2007), 203-221.

[12] W. C. Rheinboldt, *An adaptive continuation process for solving systems of nonlinear equations*, Banach Center Publ., **3** (1977), 129-142.

[13] J. F. Traub, *Iterative Methods for the Solution of Equations*, Prentice Hall Englewood Cliffs, New Jersey, USA, 1994.

Communications in Advanced Mathematical Sciences
Vol. II, No. 2, 129-134, 2019
Research Article
e-ISSN:2651-4001
DOI: 10.33434/cams.508305

Stability Conditions for Perturbed Semigroups on a Hilbert Space via Commutators

Michael Gil' [1]*

Abstract
Let A and B be linear operators on a Hilbert space. Let A and $A+B$ generate C_0-semigroups e^{tA} and $e^{t(A+B)}$, respectively, and e^{tA} be exponentially stable. We establish exponential stability conditions for $e^{t(A+B)}$ in terms of the commutator $AB-BA$, assuming that it has a bounded extension. Besides, B can be unbounded.

Keywords: Commutator, Hilbert space, Semigroups, Stability
2010 AMS: Primary 47D06, Secondary 47D60

[1] *Department of Mathematics, Ben Gurion University of the Negev, Beer-Sheva 84105, Israel, ORCID: orcid.org/0000-0002-6404-9618*
***Corresponding author**: gilmi@bezeqint.net
Received: 4 January 2019, **Accepted:** 24 April 2019, **Available online:** 27 June 2019

1. Statement of the main result

Let $\mathscr{H}$ be a Hilbert space with a scalar product $\langle .,. \rangle$, the norm $\|.\| = \sqrt{\langle ,. \rangle}$ and unit operator I. For a linear operator C, Dom (C) is the domain, C^* is the adjoint operator, $\sigma(C)$ is the spectrum. If C is a bounded operator, then $\|C\|$ is its operator norm.

Throughout this paper A and B are linear operators on $\mathscr{H}$ with $\text{Dom}(B) \supseteq \text{Dom}(A)$. In addition, A and $A+B$ generate C_0-semigroups e^{At} and $e^{t(A+B)}$, respectively.

We consider the following problem: let e^{At} be exponentially stable, i.e.

$$\|e^{At}\| \leq ce^{-\nu t} \quad (t \geq 0; c = const \geq 1, \nu = const > 0).$$

What are the conditions that provide the exponential stability of $e^{t(A+B)}$? The literature on the stability of semigroups is very rich. The classical results are presented in the books [1, 2], about the recent investigations for instance see [3] -[6], [7, 8, 9, 10]. In particular, in [7] the author investigates the uniform, strong, weak and almost weak stabilities of multiplication semigroups on Banach space valued L^p-spaces. In the paper [9] Lyapunov based proofs are presented for the well-known Arendt-Batty-Lyubich-Vu Theorem for strongly continuous and discrete semigroups. In [10] the authors obtain continuous-time and discrete-time Lyapunov operator inequalities for the exponential stability of strongly continuous, one-parameter semigroups acting on Banach spaces. Thus they extend the classic result of Datko from Hilbert spaces to Banach spaces. Recall also that various conditions, under which the perturbed operator generates a C_0-semigroup can be found for instance in [11, Chapter III]. For example, if B is A-compact and the semigroups generated by A is analytic, then by Corollary III.2.17 from [11, p. 180] $A+B$ generates an analytic semigroup. Certainly, we could not survey the whole subject here and refer the reader to the above listed publications and references given therein.

To the best of our knowledge, the exponential stability conditions for the perturbed semigroup in terms of the commutator $[A,B] = AB-BA$ have not been investigated in the available literature. In the paper [12] in the case of a Banach space, an estimate has been established for the L^1-norm of a semigroup generated by $A+B$, provided that both $[A,B]$ and B are bounded. The aim of this paper is to establish exponential stability conditions for $e^{t(A+B)}$ in terms of $[A,B]$, assuming that

$$B\mathrm{Dom}(A^2) \subseteq \mathrm{Dom}(A) \tag{1.1}$$

and

$$[A,B] \text{ has a bounded extension.} \tag{1.2}$$

Besides, B can be unbounded. Since A generates a C_0-semigroup, $\mathrm{Dom}(A^2)$ is dense in $\mathscr{H}$, cf. [13, Theorem I.2.3]. So the operators AB and BA are defined on $\mathrm{Dom}(A^2)$. Thus (1.2) means that $[A,B]$ is defined and uniformly bounded on $\mathrm{Dom}(A^2)$, and therefore admits the extension to the whole space as a bounded operator. Our approach in the present paper is considerably different from the one in [12]. In addition, we considerably generalize the main result from [14].

Introduce the operator

$$W := \int_0^\infty e^{A^*t} e^{At} dt.$$

This integral converges in the operator norm, since e^{At} is exponentially stable, and

$$\|W\| \le \int_0^\infty \|e^{At}\|^2 dt \le c^2 \int_0^\infty e^{-2\nu t} dt = \frac{c^2}{2\nu}. \tag{1.3}$$

The integral

$$\zeta(A) := 2\int_0^\infty \|e^{At}\| \int_0^t \|e^{sA}\| \|e^{(t-s)A}\| ds\, dt$$

also converges, and

$$\zeta(A) \le 2c^3 \int_0^\infty e^{-\nu t} \int_0^t e^{-\nu s} e^{-\nu(t-s)} ds\, dt = 2c^3 \int_0^\infty e^{-2\nu t} t dt = \frac{c^3}{2\nu^2}. \tag{1.4}$$

Finally assume that

$$\Lambda(B) := \sup_{h\in \mathrm{Dom}(B); \|h\|=1} \Re\langle Bh, h\rangle < \infty$$

and put

$$\psi(W,B) := \begin{cases} 2\Lambda(B)\|W\| & \text{if } \Lambda(B) > 0, \\ 0 & \text{if } \Lambda(B) \le 0. \end{cases}$$

Now we are in a position to formulate the main result of the paper.

Theorem 1.1. *Let conditions (1.1) and (1.2) hold, and e^{At} be exponentially stable. If, in addition, $\Lambda(B) < \infty$ and*

$$\psi(W,B) + \|[A,B]\|\zeta(A) < 1, \tag{1.5}$$

then $e^{t(A+B)}$ is also exponentially stable.

This theorem is proved in the next section. It is sharp. Indeed, let A and B be commuting normal operators, with $\alpha(A) := \sup \Re\sigma(A) < 0$. $\Lambda(B) = \alpha(B) > 0$. Then $\|e^{At}\| = e^{\alpha(A)t}$ $(t \ge 0)$, and by (1.3) $\|W\| \le \frac{1}{2|\alpha(A)|}$. Consequently,

$$\psi(W,B) = \frac{\alpha(B)}{|\alpha(A)|}$$

By Theorem 1.1 $e^{t(A+B)}$ is stable if $\alpha(B) < |\alpha(A)|$. But $\|e^{t(A+B)}\| = e^{(\alpha(A)+\alpha(B))t}$ $(t \ge 0)$. Therefore, in the considered case $e^{t(A+B)}$ is stable, provided $\alpha(A) + \alpha(B) < 0$. So Theorem 1.1 is really sharp.

2. Proof of Theorem 1.1

Lemma 2.1. *Let A generate a C_0-semigroup e^{At} on $\mathscr{H}$, and conditions (1.1) and (1.2) hold. Then the operator $[e^{At},B] := e^{At}B - Be^{At}$ is bounded. Moreover,*

$$[e^{At},B] = \int_0^t e^{sA}[A,B]e^{(t-s)A}ds \;\; (t \geq 0). \tag{2.1}$$

Proof. For any $x \in \mathrm{Dom}(A^2)$, we have $e^{sA}x \in \mathrm{Dom}(A^2)$ and $Ae^{sA}x \in \mathrm{Dom}(A) \subseteq \mathrm{Dom}(B)$. So $BAe^{sA}x \in \mathscr{H}$. In addition, according to (1.1) , $ABe^{sA}x \in \mathscr{H}$. Thus,

$$e^{A(t-s)}(AB - BA)e^{sA}x \in \mathscr{H} \;\; (x \in \mathrm{Dom}(A^2)).$$

But

$$e^{A(t-s)}(AB - BA)e^{sA}x = -\frac{\partial}{\partial s}e^{A(t-s)}Bx - B\frac{\partial}{\partial s}e^{sA}x = -\frac{\partial}{\partial s}e^{A(t-s)}Be^{sA}x.$$

Integrating this equality, we get

$$\begin{aligned}\int_0^t e^{A(t-s)}(AB - BA)e^{sA}xds &= -\int_0^t \frac{\partial}{\partial s}e^{A(t-s)}Be^{sA}xds = -e^{A(t-s)}Be^{sA}x|_0^t \\ &= (e^{tA}B - Be^{tA})x.\end{aligned}$$

Thus

$$[e^{At},B]x = \int_0^t e^{A(t-s)}[A,B]e^{sA}xds.$$

Since $[A,B]$ is bounded, we can extend $[e^{At},B]$ to the whole space. This proves the required relation (2.1). ☐

Proof of Theorem 1.1: Since e^{At} is exponentially stable, W is a unique solution of the Lyapunov equation

$$WA + (WA)^* = -I. \tag{2.2}$$

Equation (2.2) is understood in the sense

$$\langle Az_1, Wz_2\rangle + \langle Wz_1, Az_2\rangle = -\langle z_1, z_2\rangle \;\; (z_1, z_2 \in \mathrm{Dom}(A)). \tag{2.3}$$

Besides, $W : \mathrm{Dom}\,(A) \to \mathrm{Dom}\,(A^*)$, cf. [1, p. 252, Section 5.3].

For all $h \in \mathrm{Dom}(A)$ with $\|h\| = 1$, by (2.3) we can write

$$\begin{aligned}\langle (A+B)h, Wh\rangle + \langle Wh, (A+B)h\rangle &= \langle Ah, Wh\rangle + \langle Wh, Ah\rangle + \langle Bh, Wh\rangle + \langle Wh, Bh\rangle \\ &= -1 + \langle Bh, Wh\rangle + \langle Wh, Bh\rangle = -1 + \langle Bh, \int_0^\infty e^{A^*t}e^{At}dt\;\; h\rangle + \langle \int_0^\infty e^{A^*t}e^{At}dt\;\; h, Bh\rangle \\ &= -1 + \int_0^\infty (\langle e^{At}Bh, e^{At}h\rangle + \langle e^{At}h, e^{At}Bh\rangle)dt = -1 + \int_0^\infty (\langle Be^{At}h, e^{At}h\rangle + \langle e^{At}h, Be^{At}h\rangle)dt \\ &+ \int_0^\infty (\langle [e^{At},B]h, e^{At}h\rangle + \langle e^{At}h, [e^{At},B]h\rangle)dt.\end{aligned}$$

Thus,

$$\langle (A+B)h, Wh\rangle + \langle Wh, (A+B)h\rangle = -1 + J_1(h) + J_2(h), \tag{2.4}$$

where

$$J_1(h) = \int_0^\infty (\langle Be^{At}h, e^{At}h\rangle + \langle e^{At}h, Be^{At}h\rangle)dt$$

and

$$J_2(h) = \int_0^\infty (\langle [e^{At},B]h, e^{At}h\rangle + \langle e^{At}h, [e^{At},B]h\rangle)dt.$$

Since

$$\langle Be^{At}h, e^{At}h\rangle + \langle e^{At}h, Be^{At}h\rangle = 2\Re\langle Be^{At}h, e^{At}h\rangle \le 2\Lambda(B)\langle e^{At}h, e^{At}h\rangle,$$

we have

$$J_1(h) \le 2\Lambda(B)\int_0^\infty \langle e^{At}h, e^{At}h\rangle dt = 2\Lambda(B)\langle Wh, h\rangle. \tag{2.5}$$

If $\Lambda(B) > 0$, then $J_1(h) \le 2\Lambda(B)\|W\|$. If $\Lambda(B) < 0$, then $J_1(h) \le 0$. So $J_1(h) \le \psi(W,B)$. In addition, by Lemma 2.1

$$|J_2(h)| \le 2\int_0^\infty \|e^{At}\|\|[e^{At},B]\|dt \le 2\int_0^\infty \|e^{At}\|\|[A,B]\|\int_0^t \|e^{sA}\|\|e^{(t-s)A}\|ds\,dt = \|[A,B]\|\zeta(A). \tag{2.6}$$

Consequently, due to (1.5), for all $h \in \text{Dom}(A), \|h\| = 1$,

$$\langle (A+B)h, Wh\rangle + \langle Wh, (A+B)h\rangle = -1 + J_1(h) + J_2(h) \le -(1 - \psi(W,B) - \|[A,B]\|\zeta(A)) < 0.$$

Now the required result is due to the generalized Lyapunov theorem [2, Theorem 7.1].

3. Example

Let $\mathscr{H} = L^2(0,1)$, where $L^2(0,1)$ is the space of square-integrable functions defined on $[0,1]$ with the traditional scalar product. Let $a(x)$ be a complex valued function having a bounded measurable derivative, b be a real constant,

$$(Af)(x) = \frac{d^2 f(x)}{dx^2} + a(x)f(x) \text{ and } (Bf)(x) = bf'(x) \ \ (0 < x < 1, f \in \text{Dom}\,(A))$$

with

$$\text{Dom}\,(A) = \{h \in L^2(0,1) : h'' \in L^2(0,1), h(0) = h(1) = 0\}.$$

Then the commutator is defined by $([A,B]f)(x) = -ba'(x)f(x)$ and $\|[A,B]\| = |b|\sup_x |a'(x)|$. Clearly $A+B$ and A generate C_0-semigroups. Assume that $q := \max_x \Re a(x) < \pi^2$. Since the largest eigenvalue of the operator defined on Dom(A) by d^2/dx^2 is $-\pi^2$, we easily obtain

$$\sup_{h\in \text{Dom}\, A; \|h\|=1} \Re\langle Ah, h\rangle \le q - \pi^2 < 0.$$

So A is dissipative and therefore,

$$\|e^{At}\| \le \exp\left[-t(\pi^2 - q)\right] \ \ (t \ge 0).$$

Hence, by (1.4)

$$\zeta(A) \le \frac{1}{2(\pi^2 - q)^2}.$$

Since $(f', f) = -(f, f')$ $(f \in \text{Dom}\,(A))$, we have $\Lambda(B) = 0$ and consequently, $\psi(W,B) = 0$. Thus, due to Theorem 1.1 the semigroup generated by the operator $\tilde{A} = A + B$ defined by

$$(\tilde{A}f)(x) = \frac{d^2 f(x)}{dx^2} + bf'(x) + a(x)f(x) \ \ (0 < x < 1, f \in \text{Dom}\,(A))$$

is exponentially stable, provided

$$|b|\sup_x |a'(x)| < 2(\pi^2 - q)^2.$$

4. A particular case

In this section we refine Theorem 1.1, assuming that

$$\lambda(A) := \inf_{h\in \mathrm{Dom}(A);\|h\|=1} \Re\langle Ah,h\rangle > -\infty. \tag{4.1}$$

For example, let

$$(Af)(x) = \frac{df(x)}{dx} + a(x)f(x) \quad (0<x<1; f\in L^2(0,1))$$

with

$$\mathrm{Dom}\,(A) = \{h\in L^2(0,1) : h' \in L^2(0,1), h(0) = h(1)\}$$

and a complex bounded measurable function $a(x)$ with $\sup_x \Re\, a(x) < 0$. Simple calculations show that in this case $\lambda(A) = \inf_x \Re\, a(x) > -\infty$.

Furthermore, put

$$\hat{\psi}(W,B) := \begin{cases} 2\Lambda(B)\|W\| & \text{if } \Lambda(B)\geq 0,\\ \frac{\Lambda(B)}{|\lambda(A)|} & \text{if } \Lambda(B) < 0.\end{cases}$$

Theorem 4.1. *Let conditions (1.1), (1.2) and (4.1) hold, and e^{At} be exponentially stable. If, in addition, $\lambda(B) < \infty$ and*

$$\hat{\psi}(W,B) + \|[A,B]\|\zeta(A) < 1,$$

then $e^{t(A+B)}$ is also exponentially stable.

Proof. Define $J_1(h)$ and $J_2(h)$ $(h\in\mathrm{Dom}(A), \|h\| = 1)$ as in Section 2. Under condition $\Lambda(B) < 0$ we have

$$J_1(h) = 2\int_0^\infty \Re\langle Be^{At}h, e^{At}h\rangle dt \leq 2\Lambda(B)\int_0^\infty \langle e^{At}h, e^{At}h\rangle dt < 0.$$

Put $y(t) = e^{At}h$ $(h\in\mathrm{Dom}(A))$. Then

$$\frac{d}{dt}\langle y(t),y(t)\rangle = 2\Re\langle \frac{d}{dt}y, y(t)\rangle = 2\Re\langle Ay, y(t)\rangle \geq 2\lambda(A)\langle y(t),y(t)\rangle.$$

Solving this inequality, we get

$$\|e^{At}h\| \geq e^{\lambda(A)t}\|h\|.$$

Since A generates a stable semigroup $\lambda(A) < 0$. Consequently,

$$\int_0^\infty \langle e^{At}h, e^{At}h\rangle dt \geq \int_0^\infty e^{2\lambda(A)}dt\langle h,h\rangle = \frac{1}{2|\lambda(A)|}\langle h,h\rangle.$$

If $\Lambda(B)\geq 0$, then according to (2.5) $|J_1(h)| \leq 2\Lambda(B)\|W\|$. Thus $|J_1(h)| \leq \hat{\psi}(W,B)$. Taking into account (2.4) and (2.6) , under condition (4.1) for all $h\in\mathrm{Dom}(A), \|h\| = 1$, we obtain

$$\langle (A+B)h, Wh\rangle + \langle Wh, (A+B)h\rangle = -1 + J_1(h) + J_2(h) \leq -(1-\hat{\psi}(W,B) - \|[A,B]\|\zeta(A)) < 0.$$

Now the required result is due to the above mentioned generalized Lyapunov theorem [2, Theorem 7.1]. □

Acknowledgement

I am very grateful to the referee of this article for his (her) helpful remarks.

References

[1] R. Curtain, H. Zwart, *Introduction to Infinite-Dimensional Systems Theory*, Springer, New York, 1995.

[2] T. Eisner, *Stability of Operators and Operator Semigroups*, Operator Theory: Advances and Applications, Vol. 209, Birkhåuser Verlag, Basel, 2010.

[3] A. Batkai, K.-J. Engel, J. Pruus, R. Schnaubelt, *Polynomial stability of operator semigroups*, Math. Nachr., **279** (13-14) (2006), 1425-1440.

[4] C.J.K. Batty, T. Duyckaerts, Non-uniform stability for bounded semi-groups on Banach spaces, J. Evol. Equ., **8** (2008), 765-780.

[5] C. Buse, A. Khan, G. Rahmat, O. Saierli, *Weak real integral characterizations for exponential stability of semigroups in reflexive spaces*, Semigr. Forum, **88** (2014), 195-204.

[6] C. Buse, C. Niculescu, *A condition of uniform exponential stability for semigroups*, Math. Inequal. Appl. **11(3)** (2008), 529-536.

[7] R. Heymann, *Eigenvalues and stability properties of multiplication operators and multiplication semigroups*, Math. Nachr., **287**(5-6) (2014), 574-584.

[8] L. Maniar, S.Nafiri, *Approximation and uniform polynomial stability of C_0-semigroups*, ESAIM Control Optim. Calc. Var., **22**(1) (2016), 208-235.

[9] L. Paunonen, H. Zwart, *A Lyapunov approach to strong stability of semigroups*, Syst. & Control Let., **62** (2013), 673-678.

[10] C. Preda, P. Preda, *Lyapunov operator inequalities for exponential stability of Banach space semigroups of operators*, Appl. Math. Let., **25** (2012), 401-403.

[11] K.-J. Engel, R. Nagel, *One-Parameter Semigroups for Linear Evolution Equations*, Springer, New York, 2000.

[12] M.I. Gil', *Semigroups of sums of two operators with small commutators*, Semigroup Forum, **98**(1) (2019), 22-30.

[13] S.G. Krein, *Linear Equations in a Banach Space*, Amer. Math. Soc., Providence, R.I, 1971.

[14] M.I. Gil', *Stability of sums of operators*, Ann. Univ. Ferrara, **62** (2016), 61-70.

Communications in Advanced Mathematical Sciences
Vol. II, No. 2, 135-153, 2019
Research Article
e-ISSN:2651-4001
DOI: 10.33434/cams.512796

Standard and Corrected Numerical Differentiation Formulae

François Dubeau*

Abstract
Standard numerical differentiation rules that might be established by the method of undetermined coefficients are revisited. Best truncation error bounds are established by a direct method and by the method of integration by parts "backwards". A new method to increase the order of the truncation error using a primitive is presented. This approach leads to corrected numerical differentiation rules. Differentiation formulae and numerical tests are presented.

Keywords: Absolutely continuous function, Method of undetermined coefficients, Numerical differentiation rules, Peano kernel, Taylor's expansion.
2010 AMS: Primary 65D25, Secondary 26A24

Départment de Mathématiques, Faculté des sciences, Université de Sherbrooke, Sherbrooke (Qc), Canada, ORCID: 0000000229563208
***Corresponding author**: francois.dubeau@usherbrooke.ca
Received: 14 January 2019, **Accepted:** 12 April 2019, **Available online:** 27 June 2019

1. Introduction

Many textbooks of numerical analysis present the method of undetermined coefficients to find and approximation of the integral or the derivative of a given function, for example [1, 2, 3, 4] and many others. The method of undetermined coefficients used to estimate the k-th derivative $f^{(k)}(0)$ of a given function $f(x)$ consists in finding a $(n+1)$-dimensional weight vector $\vec{a} = (a_0, \ldots, a_n)$ associated to a given $(n+1)$-dimensional vector of distinct coordinates (or nodes) $\vec{x} = (x_0, \ldots, x_n)$ with $n \geq k$ and $|x_i| \leq 1$ for all i, such that the quantity $D^{(k)}(f;h)$ given by the formula

$$f^{(k)}(0) \approx \frac{1}{h^k} \sum_{i=0}^{n} a_i f(hx_i).$$

The method of undetermined coefficients is based on the requirement that the truncation error

$$R_{D^{(k)}}(f;h) = h^k f^{(k)}(0) - \sum_{i=0}^{n} a_i f(hx_i), \tag{1.1}$$

be such that

$$R_{D^{(k)}}(f;h) = o(h^{r(n)}), \tag{1.2}$$

where $r(n) \geq n$ depends on the regularity of $f(x)$.

It is possible to increase the order of the truncation error term. Indeed, if a primitive $F(x)$ is available, that is to say $F'(x) = f(x)$, we can add an expression to the preceding approximation as follows

$$f^{(k)}(0) \approx \frac{1}{h^k}\left(\sum_{i=0}^{n} a_i f(hx_i) + c\left[\frac{1}{h}\sum_{j=0}^{m}\beta_j F(h\xi_j) - \sum_{i=0}^{n} b_i f(hx_i)\right]\right),$$

where $(n+1)$-dimensional weight vector $\vec{b} = (b_0,\dots,b_n)$, the two $(m+1)$-dimensional vectors of weight $\vec{\beta} = (\beta_0,\dots,\beta_m)$ and distinct coordinates (or nodes) $\vec{\xi} = (\xi_0,\dots,\xi_m)$, and c are chosen in such a way that the truncation error given by

$$R^c_{D^{(k)}}(f;h) = h^k f^{(k)}(0) - \sum_{i=0}^{n} a_i f(hx_i) - c\left[\frac{1}{h}\sum_{j=0}^{m}\beta_j F(h\xi_j) - \sum_{i=0}^{n} b_i f(hx_i)\right], \tag{1.3}$$

is such that

$$R^c_{D^{(k)}}(f;h) = o(h^{r^c(n)}),$$

with $r^c(n) > r(n)$.

The plan of the paper is the following. In the next section, we present preliminaries about polynomials, Vandermonde matrix, and Taylor's expansions. Section 3 presents the standard approach for obtaining differentiation rules using the method of undetermined coefficients. We establish optimal truncation error bounds by a direct approach and by the method of integration by parts "backward". Total error bound composed of the truncation term and of the roundoff error term is given. In Section 4, we present a method to improve the error by adding information coming from a primitive. Examples of formula are given in Section 5 and numerical tests are included in Section 6.

We will use $f^{(l)}(x)$ for the l-th derivative of $f(x)$ for $l = 0,1,2,\dots$, and $f^{(0)}(x) = f(x)$. Let $1 \le p \le \infty$, if $f(x)$ is defined on a set E, $\|f\|_{p,E}$ will be its p-norm on E, and if $\vec{v}$ is a vector in $\mathbb{R}^n$, its p-norm will be $\|\vec{v}\|_p$.

2. Preliminaries

2.1 Small o and big O notations

Let $f(x)$ be a function such that $\lim_{x\to\alpha} f(x) = 0$. We say that $g(x)$ is a small o of $f(x)$ around α, and write $g(x) = o(f(x))$, if for any $\varepsilon > 0$ there exists a $\delta_\varepsilon > 0$ such that

$$|g(x)| \le \varepsilon\,|f(x)|\,.$$

holds for $0 < |x-\alpha| < \delta_\varepsilon$. We say that $g(x)$ is a big O of $f(x)$ around α, and write $g(x) = O(f(x))$, if there exist a constant C and a $\delta > 0$ such that

$$|g(x)| \le C\,|f(x)|\,.$$

holds for $0 < |x-\alpha| < \delta$.

Lemma 2.1. *Let us assume that the real number $r > 0$ and $n = \lfloor r \rfloor \ge 0$. Let $\pi_m(x)$ be a polynomial of degree m such that*

$$\pi_m(x) = o(|x-\alpha|^r)\,.$$

Then,

$$\pi_m(x) = \begin{cases} (x-\alpha)^n \pi_{m-n}(x) & \text{if} \quad m > r, \\ 0 & \text{if} \quad m \le r, \end{cases}$$

where $\pi_{m-n}(x)$ is a polynomial of degree $m-n$. □

2.2 Vandermonde matrix and Lagrange interpolation polynomials

Let $\vec{x} = (x_0,\dots,x_n)$ bet a $n+1$-vector of distincts real (or complex) numbers and its associated Vandermonde matrix $V(\vec{x})$,

$$V(\vec{x}) = \begin{bmatrix} 1 & \dots & 1 \\ x_0 & & x_n \\ \vdots & & \vdots \\ x_0^n & \dots & x_n^n \end{bmatrix}.$$

Let $\vec{e}_l$ be the $(n+1)$-column vector, the transpose of $(\delta_{l,0},\cdots,\delta_{l,j}\cdots,\delta_{l,n})$, where

$$\delta_{l,j}=\begin{cases}1 & \text{if} \quad j=l,\\ 0 & \text{if} \quad j\neq l,\end{cases}$$

for $0\leq l,j\leq n$.

Lemma 2.2. *[5] The Vandermonde matrix $V(\vec{x})$ is invertible and the l-th column of $V^{-1}(\vec{x})$ is*

$$V^{-1}(\vec{x})\vec{e}_l=\frac{1}{l!}\begin{bmatrix} w_{n,0}^{(l)}(0)\\ w_{n,1}^{(l)}(0)\\ \vdots \\ w_{n,n}^{(l)}(0)\end{bmatrix},$$

for $l=0,\dots,n$, where $\{w_{n,j}(x)\}_{j=0}^{n}$ is the Lagrange's basis of the space of polynomial of degree at most n. □

2.3 Taylor's expansion

Let $I_h=[-h,h]$, $I_h^+=[0,h]$ and $I_h^-=[-h,0]$. For $h=1$ we will simply use $I=[-1,1]$, $I^+=[0,1]$ and $I^-=[-1,0]$. Let p and q be two real numbers such that $1\leq p,q\leq\infty$ and $\frac{1}{p}+\frac{1}{q}=1$. Let $C^l(I_h)$ be the set of continuously differentiable functions up to order l on I_h, and the set of absolutely continuous function on I_h be defined by

$$AC^{l+1,p}(I_h)=\left\{f\in C^l(I_h)\,\middle|\,\begin{array}{l}(a)\ f^{(l+1)}\in L^p(I_h),\ \ and\\ (b)\ f^{(l)}(s)=f^{(l)}(r)+\int_r^s f^{(l+1)}(\xi)d\xi,\ \forall r,s\in I_h\end{array}\right\}.$$

Taylor's expansion of $f(x)\in AC^{l+1,p}(I_h)$ around $x=0$ of order $l+1$ is

$$f(x)=\sum_{j=0}^{l}\frac{f^{(j)}(0)}{j!}x^j+\int_{-h}^{h}f^{(l+1)}(y)K_{T,l}(x,y;h)dy.$$

where $K_{T,l}(x,y;h)$ is the kernel

$$K_{T,l}(x,y;h)=\frac{1}{l!}\left[(x-y)_+^l\mathbf{1}_{I_h^+}(y)+(-1)^{l+1}(y-x)_+^l\mathbf{1}_{I_h^-}(y)\right],$$

for any x, y in I_h [6, 7]. This kernel is a piecewise polynomial function of degree l. In this expression, if E is a set, then

$$\mathbf{1}_E(y)=\begin{cases}1 & \text{if} \quad y\in E,\\ 0 & \text{if} \quad y\notin E.\end{cases}$$

Also for any $l\geq 0$, $(\eta)_+^l$ is defined

$$(\eta)_+^l=\eta^l\mathbf{1}_{(0,+\infty)}(\eta).$$

If we set $x=h\xi$, and $y=h\eta$, then the kernel becomes

$$K_{T,l}(x,y;h)=K_{T,l}(h\xi,h\eta;h)=h^lK_{T,l}(\xi,\eta;1),$$

for any ξ, η in I.

3. Standard numerical differentiation rules

3.1 Existence: method of undetermined coefficients

Let us observe that $R_{D^{(k)}}(f;h)$ is linear expression with respect to $f(x)$ and also if $f(x)$ is a polynomial of degree $\leq m$ with respect to x, then $R_{D^{(k)}}(f;h)$ is a polynomial of degree $\leq m$ with respect to h. The condition (1.2), combined to Lemma 2.1, implies that $R_{D^{(k)}}(f;h) = 0$ for any polynomial $f(x)$ of degree $\leq n$. So using the standard basis $\{x^l\}_{l=0}^n$, we have to solve the linear system

$$\sum_{i=0}^{n} a_i x_i^l = \frac{D^{(k)}(x^l;h)}{h^l} = k!\delta_{k,l} \quad \text{for} \quad l = 0,\ldots,n, \tag{3.1}$$

for which the solution is

$$\vec{a} = k! V^{-1}(\vec{x})\vec{e}_k = \begin{bmatrix} w_{n,0}^{(k)}(0) \\ w_{n,1}^{(k)}(0) \\ \vdots \\ w_{n,n}^{(k)}(0) \end{bmatrix}.$$

We obtain the method

$$f^{(k)}(0) \approx \frac{\sum_{i=0}^{n} a_i f(hx_i)}{h^k}.$$

It might happen that $R_{D^{(k)}}(f;h) = 0$ for some polynomials of degree $k > n$. Let us define the degree of accuracy (or precision) k_a of the approximation process (1.1) to be the largest integer $k_a \geq n$ such that $R_{D^{(k)}}(f;h) = 0$ holds for any polynomial $f(x)$ of degree $l \leq k_a$.

3.2 Truncation error

Two different approaches can be used to establish the best bounds for the truncation error in terms of the regularity of $f(x)$. The first approach will be called the standard direct approach, while the second way is the integration by parts ”backwards”. This second approach presented in [8, 9, 10], usually presented for numerical integration [11], can be used in general when we consider the method of undetermined coefficients [12]. Let us note some bounds were already presented for specific formulae elsewhere, for example in [13].

3.2.1 Direct approach

For any integer l such that $k \leq l \leq k_a$, let $f(x) \in AC^{l+1,p}(I_h)$. Since the process is exact for polynomials of degree $\leq l$, using a Taylor's expansion of order $l+1$, the truncation error is

$$R_{D^{(k)}}(f;h) = \int_{-h}^{h} f^{(l+1)}(y) K_{D^{(k)},l}(y;h)dy,$$

where $K_{D^{(k)},l}(y;h)$ is the Peano kernel associated to the process given by

$$\begin{aligned} K_{D^{(k)},l}(y;h) &= R_{D^{(k)}}(K_{T,l}(\cdot,y;h);h) \\ &= h^k K_{T,l-k}(0,y;h) - \sum_{i=0}^{n} a_i K_{T,l}(hx_i,y;h). \end{aligned}$$

It follows that

$$\left|R_{D^{(k)}}(f;h)\right| \leq \left\|f^{(l+1)}\right\|_{p,I_h} \left\|K_{D^{(k)},l}(\cdot;h)\right\|_{q,I_h}.$$

Moreover, $K_{D^{(k)},l}(y;h) = K_{D^{(k)},l}(h\eta;h) = h^l K_{D^{(k)},l}(\eta;1)$, then

$$\left\|K_{D^{(k)},l}(\cdot;h)\right\|_{q,I_h} = h^{l+1-\frac{1}{p}} \left\|K_{D^{(k)},l}(\cdot;1)\right\|_{q,I}.$$

So

$$\left|R_{D^{(k)}}(f;h)\right| \leq h^{l+1-\frac{1}{p}} C_{k;l,p} \left\|f^{(l+1)}\right\|_{p,I_h}, \tag{3.2}$$

where

$$C_{k;l,p} = \left\|K_{D^{(k)},l}(\cdot;1)\right\|_{q,I}, \tag{3.3}$$

does not depend on h. Since because

$$\lim_{h\to 0}\left\|f^{(l+1)}\right\|_{p,I_h}=\begin{cases}0 & \text{for} \quad 1\le p<\infty,\\ C & \text{for} \quad p=\infty,\end{cases}$$

we have

$$R_{D^{(k)}}(f;h)=\begin{cases}o\left(h^{l+1-\frac{1}{p}}\right) & \text{for} \quad 1\le p<\infty,\\ O\left(h^{l+1}\right) & \text{for} \quad p=\infty.\end{cases}$$

Since an $o\left(h^{l+1-\frac{1}{p}}\right)$ and an $O\left(h^{l+1}\right)$ are $o\left(h^{l}\right)$, it means that $R_{D^{(k)}}(f;h)=o\left(h^{l}\right)$. In summary we have proved the following theorem which presents necessary and sufficient conditions to obtain the desired error order.

Theorem 3.1. *For any l such that $n\le l\le k_a$, a necessary and sufficient condition to have $R_{D^{(k)}}(f;h)=o(h^l)$ for any $f\in AC^{l+1,p}(I_h)$ is that $R_{D^{(k)}}(f;h)=0$ for any polynomial $f(x)$ of degree $\le l$.* □

Theorem 3.2. *If $R_{D^{(k)}}(f;h)=0$ for any polynomial of degree $\le k_a$, then (3.2) and (3.3) hold for any $f\in AC^{l+1,p}(I_h)$ for $k\le l\le k_a$.* □

Remark 3.3. *The bounds given by (3.2) and (3.3) are the best one as it has been shown in the general case of the method of undetermined coefficient [12].*

Remark 3.4. *Let us specify the kernels $K_{D^{(k)},l}(y;h)$ and $K_{D^{(k)},l}(\eta;1)$.*

$$\begin{aligned}K_{D^{(k)},l}(y;h) &= h^k K_{T,l-k}(x,y;h)-\sum_{i=0}^{n}a_iK_{T,l}(hx_i,y;h)\\ &= \frac{h^k}{(l-k)!}\left[(0-y)_+^{l-k}\boldsymbol{1}_{I_h^+}(y)+(-1)^{l-k+1}(y-0)_+^{l-k}\boldsymbol{1}_{I_h^-}(y)\right]\\ &\quad -\frac{1}{l!}\sum_{i=0}^{n}a_i\left[(hx_i-y)_+^l\boldsymbol{1}_{I_h^+}(y)+(-1)^{l+1}(y-hx_i)_+^l\boldsymbol{1}_{I_h^-}(y)\right],\end{aligned}$$

and

$$K_{D^{(k)},l}(\eta;1)=-\frac{1}{l!}\sum_{i=0}^{n}a_i\left[(x_i-\eta)_+^l\boldsymbol{1}_{I^+}(\eta)+(-1)^{l+1}(\eta-x_i)_+^l\boldsymbol{1}_{I^-}(\eta)\right].$$

3.2.2 Integration by parts ”backwards”

The method of integration by parts ”backwards” is based on the Taylor’s expansion of

$$W_{D^{(k)}}(f;h)=hR_{D^{(k)}}(f;h).$$

We suppose that $f\in AC^{l+1,p}(I_h)$ for $k\le l\le k_a$, and we proceed as follows. We have

$$W_{D^{(k)}}(f;h)=hR_{D^{(k)}}(h;f)=h^{k+1}f^{(k)}(0)-h\sum_{i=0}^{n}a_if(hx_i),$$

so $W_{D^{(k)}}(f;0)=0$. For $k\le l\le k_a$ and $1\le j\le l-1$

$$W_{D^{(k)}}^{(j)}(f;h)=(k+1)_jh^{k+1-j}f^{(k)}(0)-j\sum_{i=0}^{n}a_ix_i^{j-1}f^{(j-1)}(hx_i)-h\sum_{i=0}^{n}a_ix_i^jf^{(j)}(hx_i),$$

where for two non negative integers k and j

$$(k)_j=\begin{cases}\frac{k!}{(k-j)!} & \text{for} \quad j=0,\dots,k,\\ 0 & \text{for} \quad j>k.\end{cases}$$

Then

$$\lim_{h\to 0} W^{(j)}_{D^{(k)}}(f;h) = (k+1)_j \delta_{j,k+1} f^{(k)}(0) - j f^{(j-1)}(0)\sum_{i=0}^{n} a_i x_i^{j-1} = 0.$$

Also for $j = l$

$$W^{(l)}_{D^{(k)}}(f;h) = (k+1)_l h^{k+1-l} f^{(k)}(0) - l\sum_{i=0}^{n} a_i x_i^{l-1} f^{(l-1)}(hx_i) - h\sum_{i=0}^{n} a_i x_i^{l} f^{(l)}(hx_i),$$

and using Taylor's expansions of order 2 for $f^{(l-1)}(x)$ which is in $AC^{2,p}(I_h)$ and of order 1 for $f^{(l)}(x)$ which is in $AC^{1,p}(I_h)$, we obtain

$$W^{(l)}_{D^{(k)}}(f;h) = (k+1)_l h^{k+1-l} f^{(k)}(0) - l f^{(l-1)}(0)\sum_{i=0}^{n} a_i x_i^{l-1} - h(l+1) f^{(l)}(0)\sum_{i=0}^{n} a_i x_i^{l} + \int_{-h}^{h} f^{(l+1)}(y) K_{W,l}(y;h)dy.$$

Now, considering (3.1) to simplify, we obtain

$$W^{(l)}_{D^{(k)}}(f;h) = \int_{-h}^{h} f^{(l+1)}(y) K_{W,l}(y;h)dy,$$

where

$$K_{W,l}(y;h) = -l\sum_{i=0}^{n} a_i x_i^{l-1} K_{T,1}(hx_i,y;h) - h\sum_{i=0}^{n} a_i x_i^{l} K_{T,0}(hx_i,y;h).$$

Let us remark that $K_{W,l}(y;h) = hK_{W,l}(\eta;1)$ for $y = h\eta$. So the Taylor's expansion of order l for $W_n(f;h)$ leads to

$$\begin{aligned}
W_{D^{(k)}}(f;h) &= \int_0^h W^{(l)}_{D^{(k)}}(f;z)\frac{(h-z)^{l-1}}{(l-1)!}dz \\
&= \int_0^h \left[\int_{-z}^{z} f^{(l+1)}(y) K_{W,l}(y;z)dy\right]\frac{(h-z)^{l-1}}{(l-1)!}dz \\
&= \int_{-h}^{h} f^{(l+1)}(y)\left[\int_{|y|}^{h} K_{W,l}(y;z)\frac{(h-z)^{l-1}}{(l-1)!}dz\right]dy \\
&= \int_{-h}^{h} f^{(l+1)}(y)\widehat{K}_{D^{(k)},l}(y;h)dy,
\end{aligned}$$

for

$$\widehat{K}_{D^{(k)},l}(y;h) = \int_{|y|}^{h} K_{W,l}(y;z)\frac{(h-z)^{l-1}}{(l-1)!}dz.$$

As indicated in Remark 3.3, we can obtain the best bound from this expression. So we get the following result.

Theorem 3.5. *Let $h > 0$, the kernels $K_{D^{(k)},l}(y;h)$ and $\widehat{K}_{D^{(k)},l}(y;h)$ are such that*

$$hK_{D^{(k)},l}(\cdot;h) = \widehat{K}_{D^{(k)},l}(\cdot;h)$$

almost everywhere.

As a consequence both methods lead to the same best error bounds.

Remark 3.6. *The kernel is*

$$\begin{aligned}
K_{W,l}(y;h) &= -\left[l\sum_{i=0}^{n} a_i x_i^{l-1}(hx_i-y)_+ + h\sum_{i=0}^{n} a_i x_i^{l}(hx_i-y)_+^0\right]\mathbf{1}_{I_h^+}(y) \\
&\quad -\left[l\sum_{i=0}^{n} a_i x_i^{l-1}(y-hx_i)_+ - h\sum_{i=0}^{n} a_i x_i^{l}(y-hx_i)_+^0\right]\mathbf{1}_{I_h^-}(y)
\end{aligned}$$

or, after the substitution $y = h\eta$ and simplification,

$$\begin{aligned}
K_{W,l}(\eta;1) &= -\left[l\sum_{i=0}^{n} a_i x_i^{l-1}(x_i-\eta)_+ + \sum_{i=0}^{n} a_i x_i^{l}(x_i-\eta)_+^0\right]\mathbf{1}_{I^+}(\eta) \\
&\quad -\left[l\sum_{i=0}^{n} a_i x_i^{l-1}(\eta-x_i)_+ - \sum_{i=0}^{n} a_i x_i^{l}(\eta-x_i)_+^0\right]\mathbf{1}_{I^-}(\eta).
\end{aligned}$$

3.3 Total error

In effective numerical computation, the quantity $f^{(k)}(0)$ is approximated by

$$\frac{1}{h^k}\sum_{i=0}^{n} a_i \tilde{f}(hx_i),$$

which uses $\tilde{f}(hx_i)$ instead of $f(hx_i)$, so introducing roundoff error $e_i = f(hx_i) - \tilde{f}(hx_i)$. The total error $E(f;h)$ is decomposed in two types of error: the truncation error $R_{D^{(k)}}(f;h)$ and the roundoff error $S(f;h)$. Hence

$$\begin{aligned} E(f;h) &= f^{(k)}(0) - \frac{1}{h^k}\sum_{i=0}^{n} a_i \tilde{f}(hx_i) \\ &= \left[f^{(k)}(0) - \frac{1}{h^k}\sum_{i=0}^{n} a_i f(hx_i)\right] + \frac{1}{h^k}\sum_{i=0}^{n} a_i \left(f(hx_i) - \tilde{f}(hx_i)\right) \\ &= \frac{1}{h^k}\left[R_{D^{(k)}}(f;h) + S(f;h)\right], \end{aligned}$$

so

$$|E(f;h)| \le \frac{1}{h^k}\left[\left|R_{D^{(k)}}(f;h)\right| + |S(f;h)|\right].$$

For the truncation error

$$\left|R_{D^{(k)}}(f;h)\right| \le h^{l+1-\frac{1}{p}} C_{k;l,p} \left\|f^{(l+1)}\right\|_{p,I_h},$$

and for the roundoff error we have

$$S(f;h) = \sum_{i=0}^{n} a_i \left(f(hx_i) - \tilde{f}(hx_i)\right) = \sum_{i=0}^{n} a_i e_i.$$

So

$$|S(f;h)| = \left|\sum_{i=0}^{n} a_i e_i\right| \le \sum_{i=0}^{n} |a_i|\,|\varepsilon_i| \le \|\vec{a}\|_q \|\vec{e}\|_p,$$

where $\|\vec{a}\|_q$ is independant of h.

Consequently we have

$$|E_n(f;h)| \le h^{l+1-\frac{1}{p}-k} C_{k;l,p} \left\|f^{(l+1)}\right\|_{p,I_h} + \frac{\|\vec{a}\|_q \|\vec{e}\|_p}{h^k},$$

for $l = k, \ldots, k_a$. This expression shows that the derivation process ($k > 0$) is numerically unstable. See also [14] for more precision on stability of such processes.

4. Corrected numerical differentiation rules

4.1 The idea

In this section we suggest a way to improve the order of the truncation error term when we have a primitive $F(x)$ of $f(x)$, which means that $F'(x) = f(x)$. Suppose that

$$h^k f^{(k)}(0) = \sum_{i=0}^{n} a_i f(hx_i) + o(h^{k_a}),$$

so the degree of accuracy of the process is k_a, and the truncation error is

$$R_{D^{(k)}}(f;h) = h^k f^{(k)}(0) - \sum_{i=0}^{n} a_i f(hx_i) = o(h^{k_a}).$$

Suppose now that using the same $(n+1)$-dimensional vector of distinct coordinates (or nodes) $\vec{x} = (x_0, \ldots, x_n)$ we can find a $(n+1)$-dimensional weight vector $\vec{b} = (b_0, \ldots, b_n)$ such that we can determine an expression of the form

$$\frac{1}{h}\sum_{j=0}^{m} \beta_j F(h\xi_j) = \sum_{i=0}^{n} b_i f(hx_i) + o(h^{k_a}),$$

for two $(m+1)$-dimensional vectors $\vec{\beta} = (\beta_0, \ldots, \beta_m)$ and $\vec{\xi} = (\xi_0, \ldots, \xi_m)$. Its truncation error is

$$R_P(f;h) = \frac{1}{h}\sum_{j=0}^{m} \beta_j F(h\xi_j) - \sum_{i=0}^{n} b_i f(hx_i),$$

and

$$R_P(f;h) = o(h^{k_a}).$$

We can combine the two truncation error terms as follows

$$R^c_{D^{(k)}}(f;h) = R_{D^{(k)}}(f;h) - cR_P(f;h)$$

to get (1.3), and this expression is at least of degree of accuracy k_a. Since the error terms are both $o(h^{k_a})$, this expression is also exact for polynomials of degree $\leq k_a$. We can select the parameter c such that $R^c_{D^{(k)}}(f;h)$ will be also exact for polynomials of degree k_a+1. Indeed, if

$$c = \frac{R_{D^{(k)}}(x^{k_a+1};h)}{R_P(x^{k_a+1};h)},$$

then $R^c_{D^{(k)}}(f;h)$ will be also exact for polynomials of degree k_a+1, so its degree of accuracy will be at least k_a+1.

We will have

$$\begin{aligned} h^k f^{(k)}(0) &= \sum_{i=0}^{n} a_i f(hx_i) + c\left[\frac{1}{h}\sum_{j=0}^{m} \beta_j F(h\xi_j) - \sum_{i=0}^{n} b_i f(hx_i)\right] + o(h^{k_a+1}) \\ &= \sum_{i=0}^{n} [a_i - cb_i] f(hx_i) + \frac{c}{h}\sum_{j=0}^{m} \beta_j F(h\xi_j) + o(h^{k_a+1}), \end{aligned}$$

or

$$f^{(k)}(0) = \frac{1}{h^k}\sum_{i=0}^{n} [a_i - cb_i] f(hx_i) + \frac{c}{h^{k+1}}\sum_{j=0}^{m} \beta_j F(h\xi_j) + o(h^{k_a+1-k}),$$

which will be exact for polynomials of degree up to k_a+1, and we have increased the order of the error term.

4.2 Existence

The vectors $\vec{b}$, $\vec{\beta}$, and $\vec{\xi}$ of

$$\frac{1}{h}\sum_{j=0}^{m} \beta_j F(h\xi_j) = \sum_{i=0}^{n} b_i f(hx_i)$$

can be determined using the method of undetermined coefficients. It is required that

$$\sum_{j=0}^{m} \beta_j = 0,$$

and

$$\sum_{j=0}^{m} \beta_j \frac{\xi_j^{l+1}}{l+1} = \sum_{i=0}^{n} b_i x_i^l$$

for $l = 0, \ldots, n$. We also need that

$$\sum_{j=0}^{m} \beta_j \frac{\xi_j^{l+1}}{l+1} = \sum_{i=0}^{n} b_i x_i^l.$$

for $l = n+1, \ldots, k_a$. We will not present a complete analysis of this problem here. Examples of solutions of these equations are given in the last section of this paper.

4.3 Total error

In effective numerical computation, with these corrected rules the quantity $f^{(k)}(0)$ is approximated by

$$f^{(k)}(0) \approx \frac{1}{h^k}\sum_{i=0}^{n}[a_i - cb_i]\,\tilde{f}(hx_i) + \frac{c}{h^{k+1}}\sum_{j=0}^{m}\beta_j\tilde{F}(h\xi_j),$$

which uses $\tilde{f}(hx_i)$ and $\tilde{F}(h\xi_j)$ instead of $f(hx_i)$ and $F(\xi_j)$. So roundoff errors are introduced as $\mathrm{e}_i = f(hx_i) - \tilde{f}(hx_i)$ and $E_j = f(hx_j) - \tilde{f}(hx_j)$. The total error $E^c(f;h)$ is decomposed in two types of error: the truncation error $R^c_{D^{(k)}}(f;h)$ and the roundoff error $S^c(f;h)$. Hence

$$E^c(f;h) = \frac{1}{h^k}\left[R^c_{D^{(k)}}(f;h) + S^c(f;h)\right],$$

where

$$R^c_{D^{(k)}}(f;h) = h^k f^{(k)}(0) - \sum_{i=0}^{n}[a_i - cb_i]\,f(hx_i) + \frac{c}{h^1}\sum_{j=0}^{m}\beta_j F(h\xi_j),$$

and

$$S^c(f;h) = \sum_{i=0}^{n}[a_i - cb_i]\left(f(hx_i) - \tilde{f}(hx_i)\right) + \frac{c}{h^1}\sum_{j=0}^{m}\beta_j\left(F(h\xi_j) - \tilde{F}(h\xi_j)\right).$$

For the truncation error, if we proceed as we did in the preceding section, we could establish the bound

$$\left|R^c_{D^{(k)}}(f;h)\right| \le h^{l+1-\frac{1}{p}} C^c_{k;l,p}\left\|f^{(l+1)}\right\|_{p,I_h}.$$

For the roundoff error we have

$$\begin{aligned} S^c(f;h) &= \sum_{i=0}^{n}[a_i - cb_i]\left(f(hx_i) - \tilde{f}(hx_i)\right) + \frac{c}{h}\sum_{j=0}^{m}\beta_j\left(F(h\xi_j) - \tilde{F}(\xi_j)\right) \\ &= \sum_{i=0}^{n}[a_i - cb_i]\,\mathrm{e}_i + \frac{c}{h}\sum_{j=0}^{m}\beta_j E_i, \end{aligned}$$

so

$$\begin{aligned} |S^c(f;h)| &\le \sum_{i=0}^{n}|a_i - cb_i|\,|\mathrm{e}_i| + \frac{c}{h}\sum_{j=0}^{m}|\beta_j|\,|E_i| \\ &\le \left\|\vec{a} - c\vec{b}\right\|_q \|\vec{\mathrm{e}}\|_p + \frac{c}{h}\left\|\vec{\beta}\,\|_q\|\vec{E}\right\|_p \end{aligned}$$

where $\left\|\vec{a} - c\vec{b}\right\|_q$ and $\left\|\vec{\beta}\right\|_q$ are independant of h.

Consequently we have

$$|E^c(f;h)| \le h^{l+1-\frac{1}{p}-k} C^c_{k;l,p}\left\|f^{(l+1)}\right\|_{p,I_h} + \frac{1}{h^k}\left[\left\|\vec{a} - c\vec{b}\right\|_q \|\vec{\mathrm{e}}\|_p + \frac{c}{h}\left\|\vec{\beta}\,\|_q\|\vec{E}\right\|_p\right],$$

not only for $l = k, \ldots, k_a$ but also for $l = k_a + 1$. Obviously, this process is intereating for regular functions $f(x) \in AC^{k_a+2,p}(I_h)$ for which we know a primitive.

5. Examples of formula

Example 5.1. *First derivative: the 2-points symmetric formula is*

$$f^{(1)}(0) = \frac{1}{2h}[f(h) - f(-h)] + \mathrm{o}(h),$$

which is exact for polynomials of degree up to 2*. The truncation error we consider is*

$$R_{D^{(1)}}(f;h) = hf^{(1)}(0) - \frac{1}{2}[f(h) - f(-h)] = o(h^2).$$

For $f(x) = x^l$*,* $R_{D^{(1)}}(x^l;h) = 0$ *for* $l = 0,1,2$*, and for* $l \geq 3$ *we have*

$$R_{D^{(1)}}(x^l;h) = -h^l \frac{\left[1-(-1)^l\right]}{2} = \begin{cases} 0 & \text{for even} \quad l, \\ -h^l & \text{for odd} \quad l. \end{cases} \tag{5.1}$$

The corresponding expression involving the primitive $F(x)$ *is*

$$\frac{1}{h}[F(h) - 2F(0) + F(-h)] = \frac{1}{2}[f(h) - f(-h)] + o(h^2).$$

with its truncation error

$$R_P(f;h) = \frac{1}{h}[F(h) - 2F(0) + F(-h)] - \frac{1}{2}[f(h) - f(-h)] = o(h^2).$$

For $f(x) = x^l$*,* $R_P(x^l;h) = 0$ *for* $l = 0,1,2$*, and for* $l \geq 3$ *we have*

$$R_P(x^l;h) = h^l\left[1-(-1)^l\right]\left[\frac{1}{l+1} - \frac{1}{2}\right] = \begin{cases} 0 & \text{for even} \quad l, \\ 2h^l\left[\frac{1}{l+1} - \frac{1}{2}\right] & \text{for odd} \quad l. \end{cases} \tag{5.2}$$

Then we choose

$$c = \frac{R_{D^{(1)}}(x^3;h)}{R_P(x^3;h)} = \frac{R_{D^{(1)}}(x^3;1)}{R_P(x^3;1)} = \frac{-1}{-1/2} = 2.$$

The resulting formula will be exact not only for polynomials of degree 3 *but also for polynomial of degree* 4*, since (5.1) and (5.2) hold. We obtain*

$$hf^{(1)}(0) = -\frac{1}{2}[f(h) - f(-h)] + \frac{2}{h}[F(h) - 2F(0) + F(-h)] + o(h^4)$$

or

$$f^{(1)}(0) = -\frac{1}{2h}[f(h) - f(-h)] + \frac{2}{h^2}[F(h) - 2F(0) + F(-h)] + o(h^3).$$

Example 5.2. *First derivative: the one-sided formula is*

$$f^{(1)}(0) = \frac{1}{h}[f(h) - f(0)] + o(1).$$

Its corresponding truncation error is

$$R_{D^{(1)}}(f;h) = hf^{(1)}(0) - [f(h) - f(0)] = o(h).$$

For $f(x) = x^l$*,* $R_{D^{(1)}}(x^l;h) = 0$ *for* $l = 0,1$*, and for* $l \geq 2$ *we have*

$$R_{D^{(1)}}(x^l;h) = -h^l$$

We consider

$$\frac{1}{h}[F(h) - F(0)] = \frac{1}{2}[f(h) + f(0)] + o(h).$$

with its truncation error

$$R_P(f;h) = \frac{1}{h}[F(h) - F(0)] - \frac{1}{2}[f(h) + f(0)] = o(h).$$

For $f(x) = x^l$, $R_P(x^l;h) = 0$ *for* $l = 0,1$, *and for* $l \geq 2$ *we have*

$$R_P(x^l;h) = h^l \left[\frac{1}{l+1} - \frac{1}{2}\right].$$

Then we choose

$$c = \frac{R_{D^{(1)}}(x^2;h)}{R_P(x^2;h)} = \frac{R_{D^{(1)}}(x^2;1)}{R_P(x^2;1)} = \frac{-1}{-1/6} = 6.$$

The resulting formula will be exact for polynomial of degree 2*, and we obtain*

$$hf^{(1)}(0) = -2\left[f(h)+2f(0)\right] + \frac{6}{h}\left[F(h)-F(0)\right] + \mathrm{o}(h^2)$$

or

$$f^{(1)}(0) = -\frac{2}{h}\left[f(h)+2f(0)\right] + \frac{2}{h^2}\left[F(h)-F(0)\right] + \mathrm{o}(h).$$

Example 5.3. *Second derivative: the* 3*-points symmetric formula is*

$$f^{(2)}(0) = \frac{1}{h^2}\left[f(h)-2f(0)+f(-h)\right] + \mathrm{o}(h),$$

and its truncation error

$$R_{D^{(2)}}(f;h) = h^2 f^{(2)}(0) - \left[f(h)-2f(0)+f(-h)\right] = \mathrm{o}(h^3).$$

For $f(x) = x^l$, $R_{D^{(2)}}(x^l;h) = 0$ *for* $l = 0,1,2,3$, *and for* $l \geq 4$ *we have*

$$R_{D^{(2)}}(x^l;h) = -h^l\left[1+(-1)^l\right] = \begin{cases} 0 & \text{for odd} \quad l, \\ -2h^l & \text{for even} \quad l. \end{cases} \tag{5.3}$$

We consider

$$\frac{1}{h}\left[F(h)-F(-h)\right] = \left[\frac{1}{3}f(h)+\frac{4}{3}f(0)+\frac{1}{3}f(-h)\right] + \mathrm{o}(h^3),$$

with its truncation error

$$R_P(f;h) = \frac{1}{h}\left[F(h)-F(-h)\right] - \left[\frac{1}{3}f(h)+\frac{4}{3}f(0)+\frac{1}{3}f(-h)\right] = \mathrm{o}(h^3).$$

For $f(x) = x^l$, $R_P(x^l;h) = 0$ *for* $l = 0,1,2,3$, *and for* $l \geq 4$ *we have*

$$R_P(x^l;h) = h^l\left[1+(-1)^l\right]\left[\frac{1}{l+1}-\frac{1}{3}\right] = \begin{cases} 0 & \text{for odd} \quad l, \\ 2h^l\left[\frac{1}{l+1}-\frac{1}{3}\right] & \text{for even} \quad l. \end{cases} \tag{5.4}$$

Then we choose

$$c = \frac{R_{D^{(2)}}(x^4;h)}{R_P(x^4;h)} = \frac{R_{D^{(2)}}(x^4;1)}{R_P(x^4;1)} = \frac{-2}{-4/15} = \frac{15}{2}.$$

The resulting formula is not only exact for polynomials of degree 4 *but also for polynomials of degree* 5*, since (5.3) and (5.4) hold. We obtain*

$$h^2 f^{(2)}(0) = -\frac{3}{2}\left[f(h)+8f(0)+f(-h)\right] + \frac{15}{2h}\left[F(h)-F(-h)\right] + \mathrm{o}(h^5)$$

or

$$f^{(2)}(0) = -\frac{3}{2h^2}\left[f(h)+8f(0)+f(-h)\right] + \frac{15}{2h^3}\left[F(h)-F(-h)\right] + \mathrm{o}(h^3).$$

Example 5.4. *Second derivative: the* 4*-points symmetric formula is*

$$f^{(2)}(0) = \frac{9}{8h^2}\left[f(h) - f\left(\frac{h}{3}\right) - f\left(-\frac{h}{3}\right) + f(-h)\right] + o(h).$$

Its truncation error is

$$R_{D^{(2)}}(f;h) = h^2 f^{(2)}(0) - \frac{9}{8}\left[f(h) - f\left(\frac{h}{3}\right) - f\left(-\frac{h}{3}\right) + f(-h)\right] = o(h^3).$$

For $f(x) = x^l$, $R_{D^{(2)}}(x^l;h) = 0$ *for* $l = 0,1,2,3,$ *and for* $l \geq 4$ *we have*

$$R_{D^{(2)}}(x^l;h) = -\frac{9}{8}h^l\left[1+(-1)^l\right]\left[1-\frac{1}{3^l}\right] = \begin{cases} 0 & \textit{for odd} \quad l, \\ -\frac{9}{4}h^l\left[1-\frac{1}{3^{l-1}}\right] & \textit{for even} \quad l. \end{cases} \tag{5.5}$$

We consider

$$\frac{1}{h}\left[F(h) - F(-h)\right] = \frac{1}{4}\left[f(h) + 3f\left(\frac{h}{3}\right) + 3f\left(-\frac{h}{3}\right) + f(-h)\right] + o(h^3).$$

with its truncation error

$$R_P(f;h) = \frac{1}{h}\left[F(h) - F(-h)\right] - \frac{1}{4}\left[f(h) + 3f\left(\frac{h}{3}\right) + 3f\left(-\frac{h}{3}\right) + f(-h)\right] = o(h^3).$$

For $f(x) = x^l$, $R_P(x^l;h) = 0$ *for* $l = 0,1,2,3,$ *and for* $l \geq 4$ *we have*

$$R_P(x^l;h) = h^l\left[1+(-1)^l\right]\left[\frac{1}{l+1} - \frac{1}{4}\left(1+\frac{1}{3^{l-1}}\right)\right] = \begin{cases} 0 & \textit{for odd} \quad l, \\ 2h^l\left[\frac{1}{l+1} - \frac{1}{4}\left(1+\frac{1}{3^{l-1}}\right)\right] & \textit{for even} \quad l. \end{cases} \tag{5.6}$$

Then we choose

$$c = \frac{R_{D^{(2)}}(x^4;h)}{R_P(x^4;h)} = \frac{R_{D^{(2)}}(x^4;1)}{R_P(x^4;1)} = \frac{-20/9}{-16/135} = \frac{75}{4}.$$

The resulting formula will be exact not only for polynomials of degree 4 *but also for polynomials of degree* 5 *since (5.5) and (5.6) hold, and we obtain*

$$h^2 f^{(2)}(0) = -\frac{1}{16}\left[57f(h) + 243f\left(\frac{h}{3}\right) + 243f\left(-\frac{h}{3}\right) + 57f(-h)\right] + \frac{75}{4h}\left[F(h) - F(-h)\right] + o(h^5)$$

or

$$f^{(2)}(0) = -\frac{1}{16h^2}\left[57f(h) + 243f\left(\frac{h}{3}\right) + 243f\left(-\frac{h}{3}\right) + 57f(-h)\right] + \frac{75}{4h^3}\left[F(h) - F(-h)\right] + o(h^3).$$

Example 5.5. *Third derivative: the* 4*-point symmetric formula is*

$$f^{(3)}(0) = \frac{27}{8h^3}\left[f(h) - 3f\left(\frac{h}{3}\right) + 3f\left(-\frac{h}{3}\right) - f(-h)\right] + o(h).$$

Its truncation error is

$$R_{D^{(3)}}(f;h) = h^3 f^{(3)}(0) - \frac{27}{8}\left[f(h) - 3f\left(\frac{h}{3}\right) + 3f\left(-\frac{h}{3}\right) - f(-h)\right] = o(h^4).$$

For $f(x) = x^l$, $R_{D^{(3)}}(x^l;h) = 0$ *for* $l = 0,1,2,3,4,$ *and for* $l \geq 5$ *we have*

$$R_{D^{(3)}}(x^l;h) = -\frac{27}{8}h^l\left[1-(-1)^l\right]\left[1-\frac{1}{3^{l-1}}\right] = \begin{cases} 0 & \textit{for even} \quad l, \\ -\frac{27}{4}h^l\left[1-\frac{1}{3^{l-1}}\right] & \textit{for odd} \quad l. \end{cases} \tag{5.7}$$

We consider

$$\frac{1}{h}\left[F(-h)-2F(0)+F(h)\right]=\frac{1}{32}\left[7f(h)+27f\left(\frac{h}{3}\right)-27f\left(-\frac{h}{3}\right)-7f(-h)\right]+o(h^4).$$

with its truncation error

$$R_P(f;h)=\frac{1}{h}\left[F(h)-2F(0)-F(-h)\right]-\frac{1}{32}\left[7f(h)+27f\left(\frac{h}{3}\right)-27f\left(-\frac{h}{3}\right)-7f(-h)\right]=o(h^4).$$

For $f(x)=x^l$, $R_P(x^l;h)=0$ *for* $l=0,1,2,3,4$, *and for* $l\geq 5$ *we have*

$$R_P(x^l;h)=h^l\left[1-(-1)^l\right]\left[\frac{1}{l+1}-\frac{1}{32}\left(7+\frac{1}{3^{l-3}}\right)\right]=\begin{cases}0 & \text{for even} \quad l,\\ 2h^l\left[\frac{1}{l+1}-\frac{1}{32}\left(7+\frac{1}{3^{l-3}}\right)\right] & \text{for odd} \quad l.\end{cases} \tag{5.8}$$

Then we choose

$$c=\frac{R_{D^{(3)}}(x^5;h)}{R_P(x^5;h)}=\frac{R_{D^{(3)}}(x^5;1)}{R_3^P(x^5;1)}=\frac{-20/3}{-1/9}=60.$$

The resulting formula will also be exact for polynomials of degree 4 *since (5.7) and (5.8) hold, and we obtain*

$$h^3f^{(3)}(0)=-\frac{1}{4}\left[39f(h)+243f\left(\frac{h}{3}\right)-243f\left(-\frac{h}{3}\right)-39f(-h)\right]+\frac{60}{h}\left[F(h)-2F(0)+F(-h)\right]+o(h^6)$$

or

$$f^{(3)}(0)=-\frac{1}{4h^3}\left[39f(h)+243f\left(\frac{h}{3}\right)-243f\left(-\frac{h}{3}\right)-39f(-h)\right]+\frac{60}{h^4}\left[F(h)-2F(0)+F(-h)\right]+o(h^3).$$

Example 5.6. *Fourth derivative: the* 5*-points symmetric formula is*

$$f^{(4)}(0)=\frac{16}{h^4}\left[f(h)-4f\left(\frac{h}{2}\right)+6f(0)-4f\left(-\frac{h}{2}\right)+f(-h)\right]+o(h).$$

Its truncation error is

$$R_{D^{(4)}}(f;h)=h^4f^{(4)}(0)-16\left[f(h)-4f\left(\frac{h}{2}\right)+6f(0)-4f\left(-\frac{h}{2}\right)+f(-h)\right]=o(h^5).$$

For $f(x)=x^l$, $R_{D^{(4)}}(x^l;h)=0$ *for* $l=0,1,2,3,4,5$, *and for* $l\geq 6$ *we have*

$$R_{D^{(4)}}(x^l;h)=-16h^l\left[1+(-1)^l\right]\left[1-\frac{1}{2^{l-2}}\right]=\begin{cases}0 & \text{for odd} \quad l,\\ -32h^l\left[1-\frac{1}{2^{l-2}}\right] & \text{for even} \quad l.\end{cases} \tag{5.9}$$

We consider

$$\frac{1}{h}\left[F(h)-F(-h)\right]=\frac{1}{45}\left[7f(h)+32f\left(\frac{h}{2}\right)+12f(0)+32f\left(-\frac{h}{2}\right)+7f(-h)\right]+o(h^5).$$

with its truncation error

$$R_P(f;h)=\frac{1}{h}\left[F(h)-F(-h)\right]-\frac{1}{45}\left[7f(h)+32f\left(\frac{h}{2}\right)+12f(0)+32f\left(-\frac{h}{2}\right)+7f(-h)\right]=o(h^5).$$

For $f(x)=x^l$, $R_P(x^l;h)=0$ *for* $l=0,1,2,3,4,5$, *and for* $l\geq 6$ *we have*

$$R_P(x^l;h)=h^l\left[1+(-1)^l\right]\left[\frac{1}{l+1}-\frac{1}{45}\left(7+\frac{1}{2^{l-5}}\right)\right]=\begin{cases}0 & \text{for odd} \quad l,\\ 2h^l\left[\frac{1}{l+1}-\frac{1}{45}\left(7+\frac{1}{2^{l-5}}\right)\right] & \text{for even} \quad l.\end{cases} \tag{5.10}$$

Then we choose

$$c = \frac{R_{D^{(4)}}(x^6;h)}{R_P(x^6;h)} = \frac{R_{D^{(4)}}(x^6;1)}{R_P(x^6;1)} = \frac{-30}{-1/21} = 630.$$

The resulting formula will also be exact not only for polynomials of degree 6 *but also for polynomials of degree* 7 *since (5.9) and (5.10) hold. We obtain*

$$h^4 f^{(4)}(0) = -\left[82f(h) + 512f\left(\frac{h}{2}\right) + 72f(0) + 512f\left(-\frac{h}{2}\right) + 82f(-h)\right] + \frac{630}{h}\left[F(h) - F(-h)\right] + o(h^7)$$

or

$$f^{(4)}(0) = -\frac{1}{h^4}\left[82f(h) + 512f\left(\frac{h}{2}\right) + 72f(0) + 512f\left(-\frac{h}{2}\right) + 82f(-h)\right] + \frac{630}{h^5}\left[F(h) - F(-h)\right] + o(h^3)$$

6. Numerical examples

To illustrate the results, we will apply the formulae to the functions given in Table 1. To get the exponent L of expression of the form $o(h^L)$, we compute the absolute error which is of the form $O(h^{L+1})$ for a regular enough function, which is the case for the chosen functions in Table 1. In the formula, we will replace $o(h^L)$ by $O(h^{L+1})$. So the order $L+1$ is estimated by the expression

$$\log_2\left(\frac{\text{absolute error for } h}{\text{absolute error for } h/2}\right) \approx \log_2\left(\frac{O(h^{L+1})}{O((h/2)^{L+1})}\right) \approx \log_2\left(2^{L+1}\right) = L+1$$

The approximations of $L+1$ are given in the last column of the tables below. Obviously, the derivative can be estimated at any value a not only at 0 as expressed in the formula. We reconsider the 6 examples of the preceding section, and we numerically observe the predicted order $L+1$ of $O(h^{L+1})$.

Example 6.1. *First derivative: the* 2*-points symmetric formula is*

$$f^{(1)}(0) = \frac{1}{2h}\left[f(h) - f(-h)\right] + O(h^2),$$

and the corrected formula is

$$f^{(1)}(0) = -\frac{1}{2h}\left[f(h) - f(-h)\right] + \frac{2}{h^2}\left[F(h) - 2F(0) + F(-h)\right] + O(h^4).$$

The numerical results are given in Table 2, which indicates the order of the method.

Example 6.2. *First derivative: the one-sided formula is*

$$f^{(1)}(0) = \frac{1}{h}\left[f(h) - f(0)\right] + O(h),$$

and the corrected formula is

$$f^{(1)}(0) = -\frac{2}{h}\left[f(h) + 2f(0)\right] + \frac{2}{h^2}\left[F(h) - F(0)\right] + O(h^2).$$

The numerical results are given in Table 3, which indicates the order of the method.

Example 6.3. *Second derivative: the* 3*-points symmetric formula is*

$$f^{(2)}(0) = \frac{1}{h^2}\left[f(h) - 2f(0) + f(-h)\right] + O(h^2),$$

and the corrected formula is

$$f^{(2)}(0) = -\frac{3}{2h^2}\left[f(h) + 8f(0) + f(-h)\right] + \frac{15}{2h^3}\left[F(h) - F(-h)\right] + O(h^4).$$

The numerical results are given in Table 4, which indicates the order of the method.

Example 6.4. *Second derivative: the* 4*-points symmetric formula is*

$$f^{(2)}(0) = \frac{9}{8h^2}\left[f(h) - f\left(\frac{h}{3}\right) - f\left(-\frac{h}{3}\right) + f(-h)\right] + O(h^2).$$

and

$$f^{(2)}(0) = -\frac{1}{16h^2}\left[57f(h) + 243f\left(\frac{h}{3}\right) + 243f\left(-\frac{h}{3}\right) + 57f(-h)\right] + \frac{75}{4h^3}\left[F(h) - F(-h)\right] + O(h^4).$$

The numerical results are given in Table 5, which indicates the order of the method.

Example 6.5. *Third derivative: the* 4*-point symmetric formula is*

$$f^{(3)}(0) = \frac{27}{8h^3}\left[f(h) - 3f\left(\frac{h}{3}\right) + 3f\left(-\frac{h}{3}\right) - f(-h)\right] + O(h^2).$$

and

$$f^{(3)}(0) = -\frac{1}{4h^3}\left[39f(h) + 243f\left(\frac{h}{3}\right) - 243f\left(-\frac{h}{3}\right) - 39f(-h)\right] + \frac{60}{h^4}\left[F(h) - 2F(0) + F(-h)\right] + O(h^4).$$

The numerical results are given in Table 6, which indicates the order of the method.

Example 6.6. *Fourth derivative: the* 5*-points symmetric formula is*

$$f^{(4)}(0) = \frac{16}{h^4}\left[f(h) - 4f\left(\frac{h}{2}\right) + 6f(0) - 4f\left(-\frac{h}{2}\right) + f(-h)\right] + O(h^2).$$

and

$$f^{(4)}(0) = -\frac{1}{h^4}\left[82f(h) + 512f\left(\frac{h}{2}\right) + 72f(0) + 512f\left(-\frac{h}{2}\right) + 82f(-h)\right] + \frac{630}{h^5}\left[F(h) - F(-h)\right] + O(h^4)$$

The numerical results are given in Table 7, which indicates the order of the method.

7. Conclusion

In this paper, we have presented a complete analysis of the standard numerical differentiation formulae for which we have established, using two different methods, the best error bounds depending on the regularity of absolutely continuous functions. Moreover we have presented a way to improve the order of those formulae by adding information coming from a primitive of the function. Obviously, this process is possible if we can get values of the primitive, directly of by an indirect method.

Acknowledgement

This work has been financially supported by an individual discovery grant from NSERC (Natural Sciences and Engineering Research Council of Canada).

References

[1] P. J. Davis, *Interpolation and Approximation*, Dover, N.Y., 1975.

[2] A. S. Househoulder, *Principles of Numerical Analysis*, McGraw Hill, Columbus, N.Y., 1953.

[3] D. Kincaid, W. Cheney, *Numerical Analysis*, Brooks/Cole Pub. Co., Cal., 1991.

[4] A. Ralston, *A First Course in Numerical Analysis*, McGraw-Hill, N.Y., 1965.

[5] N. Macon, A. Spitzbart, *Inverses of Vandermonde matrices*, Amer. Math. Monthly, **65**(2) (1958), 95-100. http://dx.doi.org/10.2307/2308881

[6] E. Asplund, L. Bungart, *A First Course in Integration*, Holt, Rinehart and Winston, N.Y., 1966.

[7] L. L. Schumaker, *Spline Functions Basic Theory*, Wiley, N.Y., 1981.

[8] G. Peano, *Resto nelle formule di quadratura espresso con un integrale definito*, Atti. Accad. Naz. Lincei, Rend. Cl. Sci. Fis. Mat. Nat., Serie 5, **22**(I) (1913), 562-569.

[9] R. von Mises, *Über allgemeine quadraturformeln*, J. Reine Angew. Math., **174** (1935), 56-67; reprinted in Selected Papers of Richard von Mises, Vol. 1, 559-574, American Mathematical Society, Providence, R.I., 1963.

[10] A. Ghizzetti, A. Ossicini, *Quadrature Formulae*, Academic Press, N.Y., 1970.

[11] F. Dubeau, *Revisited optimal error bounds for interpolatory integration rules*, Adv. Numer. Anal., **2016** (2016), Article ID 3170595, 8 pages, http://dx.doi.org/10.1155/2016/3170595.

[12] F. Dubeau, *The method of undetermined coefficients: general approach and optimal error bounds*, J. Math. Anal., **5**(4) (2014), 1-11.

[13] J. S. C. Prentice, *Truncation and roundoff errors in three-point approximations of first and second derivatives*, Appl. Math. Comput., **217** (2011), 4576-4581. http://dx.doi.org/10.1016/j.amc.2010.11.008

[14] N. J. Higham, *Accuracy and Stability of Numerical Algorithms*, 2nd Ed., SIAM, Philadelphia, PN, 2002. http://dx.doi.org/10.1137/1.9780898718027

Test functions		
$f(x)$	$\frac{1}{1+x^2}$	$\tan(x)$
$f^{(1)}(x)$	$-2\frac{x}{(1+x^2)^2}$	$1+\tan^2(x)$
$f^{(2)}(x)$	$-2\frac{(1-3x^2)}{(1+x^2)^3}$	$2\tan(x)(1+\tan^2(x))$
$f^{(3)}(x)$	$24x\frac{(1-x^2)}{(1+x^2)^4}$	$2(1+\tan^2(x))(1+3\tan^2(x))$
$f^{(4)}(x)$	$24\frac{(1-10x^2+5x^4)}{(1+x^2)^5}$	$8\tan(x)(1+\tan^2(x))(2+3\tan^2(x))$
$F(x)$	$\arctan(x)$	$-\ln\|\cos(x)\|$

Table 1. Test functions for numerical differentiation.

	Symmetric first derivative rule					
	Standard rule			Corrected rule		
h	computed derivative	absolute error	estimated order	computed derivative	absolute error	estimated order
	Estimation of $f^{(1)}(2) = -0.16$ for $f(x) = 1/(1+x^2)$					
1.0000	-2.00000000e-01	4.00000000e-02		-1.59707000e-01	2.93000415e-04	
0.5000	-1.69761273e-01	9.76127321e-03	2.03	-1.59948828e-01	5.11719430e-05	2.52
0.2500	-1.62410785e-01	2.41078509e-03	2.02	-1.59996441e-01	3.55871874e-06	3.85
0.1250	-1.60600684e-01	6.00684200e-04	2.00	-1.59999772e-01	2.27514621e-07	3.98
0.0625	-1.60150043e-01	1.50042917e-04	2.00	-1.59999986e-01	1.42973008e-08	3.99
0.0313	-1.60037503e-01	3.75026847e-05	2.00	-1.59999999e-01	8.95188673e-10	4.00
0.0156	-1.60009375e-01	9.37516783e-06	2.00	-1.60000000e-01	5.67215996e-11	3.98
	Estimation of $f^{(1)}(\pi/4) = 2$ for $f(x) = \tan(x)$					
1.0000	-2.18503986e+00	4.18503986e+00		3.93847408e+00	6.57541629e+00	
0.5000	3.11481545e+00	1.11481545e+00	1.91	1.81019631e+00	1.89803686e-01	5.11
0.2500	2.18520996e+00	1.85209959e-01	2.59	1.99348573e+00	6.51426518e-03	4.86
0.1250	2.04273537e+00	4.27353698e-02	2.12	1.99963919e+00	3.60810092e-04	4.17
0.0625	2.01048219e+00	1.04821852e-02	2.03	1.99997809e+00	2.19088367e-05	4.04
0.0313	2.00260824e+00	2.60824212e-03	2.01	1.99999864e+00	1.35955952e-06	4.01
0.0156	2.00065130e+00	6.51296080e-04	2.00	1.99999992e+00	8.48203887e-08	4.00

Table 2. Estimation of $f^{(1)}(x)$ using a symmetric rule.

	Unilateral first derivative rule					
	Standard rule			Corrected rule		
h	computed derivative	absolute error	estimated order	computed derivative	absolute error	estimated order
	Estimation of $f^{(1)}(2) = -0.16$ for $f(x) = 1/(1+x^2)$					
1.0000	-1.00000000e-01	6.00000000e-02		-1.48617672e-01	1.13823276e-02	
0.5000	-1.24137931e-01	3.58620690e-02	0.74	-1.56334573e-01	3.66542739e-03	1.63
0.2500	-1.40206186e-01	1.97938144e-02	0.86	-1.58952804e-01	1.04719561e-03	1.81
0.1250	-1.49575071e-01	1.04249292e-02	0.93	-1.59719803e-01	2.80196710e-04	1.90
0.0625	-1.54646840e-01	5.35315985e-03	0.96	-1.59927519e-01	7.24808122e-05	1.95
0.0313	-1.57287102e-01	2.71289769e-03	0.98	-1.59981568e-01	1.84323851e-05	1.98
0.0156	-1.58634325e-01	1.36567488e-03	0.99	-1.59995352e-01	4.64763034e-06	1.99
	Estimation of $f^{(1)}(\pi/4) = 2$ for $f(x) = \tan(x)$					
1.0000	-5.58803782e+00	7.58803782e+00		1.23765843e+01	2.15169529e+01	
0.5000	4.81644688e+00	2.81644688e+00	1.43	4.68917926e-01	1.53108207e+00	3.81
0.2500	2.74318567e+00	7.43185669e-01	1.92	1.84910116e+00	1.50898839e-01	3.34
0.1250	2.29941556e+00	2.99415562e-01	1.31	1.97268790e+00	2.73120951e-02	2.47
0.0625	2.13630119e+00	1.36301191e-01	1.14	1.99406905e+00	5.93095126e-03	2.20
0.0313	2.06521013e+00	6.52101292e-02	1.06	1.99861227e+00	1.38772563e-03	2.10
0.0156	2.03191402e+00	3.19140168e-02	1.03	1.99966405e+00	3.35953571e-04	2.05

Table 3. Estimation of $f^{(1)}(x)$ using a non symmetric rule.

	Second derivative rule					
	Standard rule			Corrected rule		
h	computed derivative	absolute error	estimated order	computed derivative	absolute error	estimated order
	Estimation of $f^{(2)}(2) = 0.176$ for $f(x) = 1/(1+x^2)$					
1.0000	2.00000000e-01	2.40000000e-02		1.77357068e-01	1.35706751e-03	
0.5000	1.82493369e-01	6.49336870e-03	1.89	1.76033533e-01	3.35330618e-05	5.34
0.2500	1.77636796e-01	1.63679619e-03	1.99	1.76001446e-01	1.44601079e-06	4.54
0.1250	1.76409814e-01	4.09814051e-04	2.00	1.76000081e-01	8.08135942e-08	4.16
0.0625	1.76102489e-01	1.02488599e-04	2.00	1.76000005e-01	4.90153890e-09	4.04
0.0313	1.76025624e-01	2.56242908e-05	2.00	1.76000000e-01	2.88695456e-10	4.09
0.0156	1.76006406e-01	6.40620560e-06	2.00	1.76000000e-01	1.94850247e-11	3.89
	Estimation of $f^{(2)}(\pi/4) = 4$ for $f(x) = \tan(x)$					
1.0000	-6.80599592e+00	1.08059959e+01		6.63488721e+00	2.37088144e+01	
0.5000	6.80652574e+00	2.80652574e+00	1.95	3.36168164e+00	6.38318359e-01	5.21
0.2500	4.46380567e+00	4.63805674e-01	2.60	3.97858106e+00	2.14189379e-02	4.90
0.1250	4.10688307e+00	1.06883070e-01	2.12	3.99881917e+00	1.18083079e-03	4.18
0.0625	4.02620819e+00	2.62081947e-02	2.03	3.99992838e+00	7.16213649e-05	4.04
0.0313	4.00652078e+00	6.52077514e-03	2.01	3.99999556e+00	4.44325997e-06	4.01
0.0156	4.00162825e+00	1.62825080e-03	2.00	3.99999972e+00	2.77308572e-07	4.00

Table 4. Estimation of $f^{(2)}(x)$ using a 3-points symmetric rule.

	Second derivative rule					
	Standard rule			Corrected rule		
h	computed derivative	absolute error	estimated order	computed derivative	absolute error	estimated order
	Estimation of $f^{(2)}(2) = 0.176$ for $f(x) = 1/(1+x^2)$					
1.0000	2.02636917e-01	2.66369168e-02		1.78991046e-01	2.99104605e-03	
0.5000	1.83214004e-01	7.21400359e-03	1.88	1.76072351e-01	7.23506264e-05	5.37
0.2500	1.77818622e-01	1.81862247e-03	1.99	1.76003075e-01	3.07542484e-06	4.56
0.1250	1.76455347e-01	4.55346643e-04	2.00	1.76000171e-01	1.70955263e-07	4.17
0.0625	1.76113876e-01	1.13876080e-04	2.00	1.76000010e-01	1.03522543e-08	4.05
0.0313	1.76028471e-01	2.84714255e-05	2.00	1.76000001e-01	6.04290118e-10	4.10
	Estimation of $f^{(2)}(\pi/4) = 4$ for $f(x) = \tan(x)$					
1.0000	-8.26975149e+00	1.22697515e+01		7.03551181e+00	5.89830241e+01	
0.5000	7.13309802e+00	3.13309802e+00	1.97	2.60314120e+00	1.39685880e+00	5.40
0.2500	4.51592823e+00	5.15928230e-01	2.60	3.95439864e+00	4.56013565e-02	4.94
0.1250	4.11879260e+00	1.18792597e-01	2.12	3.99750194e+00	2.49806479e-03	4.19
0.0625	4.02912227e+00	2.91222737e-02	2.03	3.99984872e+00	1.51279188e-04	4.05
0.0313	4.00724543e+00	7.24543362e-03	2.01	3.99999062e+00	9.38143057e-06	4.01

Table 5. Estimation of $f^{(2)}(x)$ using a 4-points symmetric rule.

	Third derivative rule					
	Standard rule			Corrected rule		
h	computed derivative	absolute error	estimated order	computed derivative	absolute error	estimated order
	Estimation of $f^{(3)}(2) = -2.30400000e-01$ for $f(x) = 1/(1+x^2)$					
1.0000	-2.40973631e-01	1.05736308e-02		-2.37051773e-01	6.65177254e-03	
0.5000	-2.34696213e-01	4.29621292e-03	1.30	-2.30683859e-01	2.83859197e-04	4.55
0.2500	-2.31550154e-01	1.15015437e-03	1.90	-2.30415897e-01	1.58974761e-05	4.16
0.1250	-2.30691910e-01	2.91909546e-04	1.98	-2.30400966e-01	9.65795639e-07	4.04
0.0625	-2.30473245e-01	7.32447324e-05	1.99	-2.30400059e-01	5.94454010e-08	4.02
	Estimation of $f^{(3)}(\pi/4) = 16$ for $f(x) = \tan(x)$					
1.0000	-3.06825876e+01	4.66825876e+01		-3.90107225e-01	1.89206795e+02	
0.5000	2.80069421e+01	1.20069421e+01	1.96	1.14873564e+01	4.51264360e+00	5.39
0.2500	1.79802007e+01	1.98020069e+00	2.60	1.58535764e+01	1.46423628e-01	4.95
0.1250	1.64561085e+01	4.56108524e-01	2.11	1.59919858e+01	8.01418527e-03	4.19
0.0625	1.61118262e+01	1.11826164e-01	2.02	1.59995148e+01	4.85235909e-04	4.05

Table 6. Estimation of $f^{(3)}(x)$ using a 4-points symmetric rule.

	Fourth derivative rule					
	Standard rule			Corrected rule		
h	computed derivative	absolute error	estimated order	computed derivative	absolute error	estimated order
	Estimation of $f^{(4)}(2) = 3.14880000e-01$ for $f(x) = 1/(1+x^2)$					
1.0000	2.80106101e-01	3.47738992e-02		3.38842477e-01	2.39624769e-02	
0.5000	3.10820640e-01	4.05935957e-03	3.10	3.16060349e-01	1.18034893e-03	4.34
0.2500	3.14107429e-01	7.72571418e-04	2.39	3.14948783e-01	6.87831106e-05	4.10
0.1250	3.14701263e-01	1.78737171e-04	2.11	3.14884220e-01	4.22036637e-06	4.03
0.0625	3.14836205e-01	4.37952495e-05	2.03	3.14880209e-01	2.09043622e-07	4.34
	Estimation of $f^{(4)}(\pi/4) = 80$ for $f(x) = \tan(x)$					
1.0000	-2.17800347e+02	2.97800347e+02		-6.13368590e+02	2.09714234e+03	
0.5000	1.49934084e+02	6.99340843e+01	2.09	4.20050565e+01	3.79949435e+01	5.79
0.2500	9.13721867e+01	1.13721867e+01	2.62	7.88356148e+01	1.16438517e+00	5.03
0.1250	8.26110723e+01	2.61107229e+00	2.12	7.99370476e+01	6.29523769e-02	4.21
0.0625	8.06396706e+01	6.39670610e-01	2.03	7.99961997e+01	3.80027294e-03	4.05

Table 7. Estimation of $f^{(4)}(x)$ using a 5-points symmetric rule.

Communications in Advanced Mathematical Sciences
Vol. II, No. 2, 154-160, 2019
Research Article
e-ISSN:2651-4001
DOI: 10.33434/cams.553548

On $\mathscr{I}_2$-Cauchy Double Sequences in Fuzzy Normed Spaces

Erdinç Dündar [1]*, Muhammed Recai Türkmen[2]

Abstract
In this paper, we investigate relationship between $\mathscr{I}_2$-convergence and $\mathscr{I}_2$-Cauchy double sequences in fuzzy normed spaces. After, we introduce the concepts of $\mathscr{I}_2^*$-Cauchy double sequences and study relationships between $\mathscr{I}_2$-Cauchy and $\mathscr{I}_2^*$-Cauchy double sequences in fuzzy normed spaces.

Keywords: Double sequences, Fuzzy normed space, $\mathscr{I}_2$-Cauchy, $\mathscr{I}_2$-convergence.
2010 AMS: Primary 34C41, 40A35, Secondary 40G15

[1] *Department of Mathematics, Faculty of Science and Literature, Afyon Kocatepe University, 03200, Afyonkarahisar, Turkey, ORCID: 0000-*
[2] *Department of Mathematics, Faculty of Education, Afyon Kocatepe University, 03200, Afyonkarahisar, Turkey, ORCID: 0000-*
***Corresponding author**: edundar@aku.edu.tr
Received: 14 April 2019, **Accepted:** 20 May 2019, **Available online:** 27 June 2019

1. Introduction and background

Throughout the paper $\mathbb{N}$ and $\mathbb{R}$ denote the set of all positive integers and the set of all real numbers, respectively. The concept of convergence of a sequence of real numbers has been extended to statistical convergence independently by Fast [1] and Schoenberg [2]. A lot of developments have been made in this area after the various studies of researchers [3, 4]. The idea of $\mathscr{I}$-convergence was introduced by Kostyrko et al. [5] as a generalization of statistical convergence which is based on the structure of the ideal $\mathscr{I}$ of subset of the set of natural numbers $\mathbb{N}$. Das et al. [6] introduced the concept of $\mathscr{I}$-convergence of double sequences in a metric space and studied some properties of this convergence. A lot of developments have been made in this area after the works of [7, 8, 9, 10].

The concept of ordinary convergence of a sequence of fuzzy numbers was firstly introduced by Matloka [11] and proved some basic theorems for sequences of fuzzy numbers. Nanda [12] studied the sequences of fuzzy numbers and showed that the set of all convergent sequences of fuzzy numbers are a complete metric space. Şençimen and Pehlivan [13] introduced the notions of statistically convergent sequence and statistically Cauchy sequence in a fuzzy normed linear space. Hazarika [14] studied the concepts of $\mathscr{I}$-convergence, $\mathscr{I}^*$-convergence and $\mathscr{I}$-Cauchy sequence in a fuzzy normed linear space. Dündar and Talo [15, 16] introduced the concepts of $\mathscr{I}_2$-convergence and $\mathscr{I}_2$-Cauchy sequence for double sequences of fuzzy numbers and studied some properties and relations of them. Hazarika and Kumar [17] introduced the notion of $\mathscr{I}_2$-convergence and $\mathscr{I}_2$-Cauchy double sequences in a fuzzy normed linear space. Dündar and Türkmen [18] studied some properties of $\mathscr{I}_2$-convergence and $\mathscr{I}_2^*$-convergence of double sequences in fuzzy normed spaces. A lot of developments have been made in this area after the various studies of researchers [19, 20, 21, 22].

Now, we recall the concept of ideal, convergence, statistical convergence, ideal convergence of sequence, double sequence and fuzzy normed and some basic definitions (see [1, 3, 4, 13, 15, 20, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34]).

Fuzzy sets are considered with respect to a nonempty base set X of elements of interest. The essential idea is that each element $x \in X$ is assigned a membership grade $u(x)$ taking values in $[0,1]$, with $u(x) = 0$ corresponding to nonmembership,

$0 < u(x) < 1$ to partial membership, and $u(x) = 1$ to full membership. According to Zadeh [35], a fuzzy subset of X is a nonempty subset $\{(x,u(x)) : x \in X\}$ of $X \times [0,1]$ for some function $u : X \to [0,1]$. The function u itself is often used for the fuzzy set.

A fuzzy set u on $\mathbb{R}$ is called a fuzzy number if it has the following properties:
1. u is normal, that is, there exists an $x_0 \in \mathbb{R}$ such that $u(x_0) = 1$;
2. u is fuzzy convex, that is, for $x,y \in \mathbb{R}$ and $0 \le \lambda \le 1$, $u(\lambda x + (1-\lambda)y) \ge \min[u(x),u(y)]$;
3. u is upper semicontinuous;
4. $supp\,u = cl\{x \in \mathbb{R} : u(x) > 0\}$, or denoted by $[u]_0$, is compact.

Let $L(\mathbb{R})$ be set of all fuzzy numbers. If $u \in L(\mathbb{R})$ and $u(t) = 0$ for $t < 0$, then u is called a non-negative fuzzy number. We write $L^*(\mathbb{R})$ by the set of all non-negative fuzzy numbers. We can say that $u \in L^*(\mathbb{R})$ iff $u_\alpha^- \ge 0$ for each $\alpha \in [0,1]$. Clearly we have $\widetilde{0} \in L(\mathbb{R})$. For $u \in L(\mathbb{R})$, the α level set of u is defined by

$$[u]_\alpha = \begin{cases} \{x \in \mathbb{R} : u(x) \ge \alpha\}, & \text{if } \alpha \in (0,1] \\ supp\,u, & \text{if } \alpha = 0. \end{cases}$$

A partial order $\preceq$ on $L(\mathbb{R})$ is defined by $u \preceq v$ if $u_\alpha^- \le v_\alpha^-$ and $u_\alpha^+ \le v_\alpha^+$ for all $\alpha \in [0,1]$.

Arithmetic operation for $t \in \mathbb{R}$, $\oplus, \ominus, \odot$ and $\oslash$ on $L(\mathbb{R}) \times L(\mathbb{R})$ are defined by
$(u \oplus v)(t) = \sup_{s\in\mathbb{R}} \{u(s) \wedge v(t-s)\}$, $\quad (u \ominus v)(t) = \sup_{s\in\mathbb{R}} \{u(s) \wedge v(s-t)\}$,
$(u \odot v)(t) = \sup_{s\in\mathbb{R}, s\neq 0} \{u(s) \wedge v(t/s)\}$ and $(u \oslash v)(t) = \sup_{s\in\mathbb{R}} \{u(st) \wedge v(s)\}$.
For $k \in \mathbb{R}^+$, ku is defined as $ku(t) = u(t/k)$ and $0u(t) = \tilde{0}$, $t \in \mathbb{R}$.
Some arithmetic operations for $\alpha-$level sets are defined as follows:
$u, v \in L(\mathbb{R})$ and $[u]_\alpha = [u_\alpha^-, u_\alpha^+]$ and $[v]_\alpha = [v_\alpha^-, v_\alpha^+]$, $\alpha \in (0,1]$. Then,
$[u \oplus v]_\alpha = [u_\alpha^- + v_\alpha^-, u_\alpha^+ + v_\alpha^+]$, $[u \ominus v]_\alpha = [u_\alpha^- - v_\alpha^+, u_\alpha^+ - v_\alpha^-]$,
$[u \odot v]_\alpha = [u_\alpha^- . v_\alpha^-, u_\alpha^+ . v_\alpha^+]$ and $\left[\tilde{1} \oslash u\right]_\alpha = \left[\frac{1}{u_\alpha^+}, \frac{1}{u_\alpha^-}\right]$, $u_\alpha^- > 0$.
For $u, v \in L(\mathbb{R})$, the supremum metric on $L(\mathbb{R})$ defined as

$$D(u,v) = \sup_{0\le\alpha\le1} \max\left\{\left|u_\alpha^- - v_\alpha^-\right|, \left|u_\alpha^+ - v_\alpha^+\right|\right\}.$$

It is known that D is a metric on $L(\mathbb{R})$ and $(L(\mathbb{R}), D)$ is a complete metric space.

A sequence $x = (x_k)$ of fuzzy numbers is said to be convergent to the fuzzy number x_0, if for every $\varepsilon > 0$ there exists a positive integer k_0 such that $D(x_k, x_0) < \varepsilon$ for $k > k_0$ and a sequence $x = (x_k)$ of fuzzy numbers convergent to levelwise to x_0 if and only if $\lim_{k\to\infty} [x_k]_\alpha = [x_0]_\alpha^-$ and $\lim_{k\to\infty} [x_k]_\alpha = [x_0]_\alpha^+$, where $[x_k]_\alpha = \left[(x_k)_\alpha^-, (x_k)_\alpha^+\right]$ and $[x_0]_\alpha = \left[(x_0)_\alpha^-, (x_0)_\alpha^+\right]$, for every $\alpha \in (0,1)$.

Let X be a vector space over $\mathbb{R}$, $\|.\| : X \to L^*(\mathbb{R})$ and the mappings $L;R$ (respectively, left norm and right norm) : $[0,1] \times [0,1] \to [0,1]$ be symetric, nondecreasing in both arguments and satisfy $L(0,0) = 0$ and $R(1,1) = 1$.

The quadruple $(X, \|.\|, L, R)$ is called fuzzy normed linear space (briefly FNS) and $\|.\|$ a fuzzy norm if the following axioms are satisfied

1. $\|x\| = \widetilde{0}$ iff $x = 0$,
2. $\|rx\| = |r| \odot \|x\|$ for $x \in X$, $r \in \mathbb{R}$,
3. For all $x, y \in X$
   (a) $\|x+y\|(s+t) \ge L(\|x\|(s), \|y\|(t))$, whenever $s \le \|x\|_1^-$, $t \le \|y\|_1^-$ and $s+t \le \|x+y\|_1^-$,
   (b) $\|x+y\|(s+t) \le R(\|x\|(s), \|y\|(t))$, whenever $s \ge \|x\|_1^-$, $t \ge \|y\|_1^-$ and $s+t \ge \|x+y\|_1^-$.

Let $(X, \|.\|_C)$ be an ordinary normed linear space. Then, a fuzzy norm $\|.\|$ on X can be obtained by

$$\|x\|(t) = \begin{cases} 0, & \text{if } 0 \le t \le a\|x\|_C \text{ or } t \ge b\|x\|_C \\ \frac{t}{(1-a)\|x\|_C} - \frac{a}{1-a}, & \text{if } a\|x\|_C \le t \le \|x\|_C \\ \frac{-t}{(b-1)\|x\|_C} + \frac{b}{b-1}, & \text{if } \|x\|_C \le t \le b\|x\|_C \end{cases}$$

where $\|x\|_C$ is the ordinary norm of $x\ (\neq 0)$, $0 < a < 1$ and $1 < b < \infty$. For $x = 0$, define $\|x\| = \widetilde{0}$. Hence, $(X, \|.\|)$ is a fuzzy normed linear space.

Let us consider the topological structure of an $FNS\ (X, \|.\|)$. For any $\varepsilon > 0, \alpha \in [0,1]$ and $x \in X$, the $(\varepsilon, \alpha)-$ neighborhood of x is the set $\mathscr{N}_x(\varepsilon, \alpha) = \{y \in X : \|x-y\|_\alpha^+ < \varepsilon\}$.

Let $(X,\|.\|)$ be an FNS. A sequence $(x_n)_{n=1}^{\infty}$ in X is convergent to $x\in X$ with respect to the fuzzy norm on X and we denote by $x_n \overset{FN}{\to} x$, provided that $(D)-\lim_{n\to\infty}\|x_n-x\|=\widetilde{0}$; i.e., for every $\varepsilon>0$ there is an $N(\varepsilon)\in\mathbb{N}$ such that $D\left(\|x_n-x\|,\widetilde{0}\right)<\varepsilon$ for all $n\geq N(\varepsilon)$. This means that for every $\varepsilon>0$ there is an $N(\varepsilon)\in\mathbb{N}$ such that for all $n\geq N(\varepsilon)$, $\sup\limits_{\alpha\in[0,1]}\|x_n-x\|_{\alpha}^{+}=\|x_n-x\|_0^{+}<\varepsilon$.

Let $(X,\|.\|)$ be an FNS. Then a double sequence (x_{jk}) is said to be convergent to $x\in X$ with respect to the fuzzy norm on X if for every $\varepsilon>0$ there exist a number $N=N(\varepsilon)$ such that $D\left(\|x_{jk}-x\|,\widetilde{0}\right)<\varepsilon,$ for all $j,k\geq N.$

In this case, we write $x_{jk}\overset{FN}{\longrightarrow}x$. This means that, for every $\varepsilon>0$ there exist a number $N=N(\varepsilon)$ such that $\sup\limits_{\alpha\in[0,1]}\|x_{jk}-x\|_{\alpha}^{+}=\|x_{jk}-x\|_0^{+}<\varepsilon$, for all $j,k\geq N$. In terms of neighnorhoods, we have $x_{jk}\overset{FN}{\longrightarrow}x$ provided that for any $\varepsilon>0$, there exists a number $N=N(\varepsilon)$ such that $x_{jk}\in\mathscr{N}_x(\varepsilon,0)$, whenever $j,k\geq N$.

Let $X\neq\emptyset$. A class $\mathscr{I}$ of subsets of X is said to be an ideal in X provided:

(i) $\emptyset\in\mathscr{I}$, (ii) $A,B\in\mathscr{I}$ implies $A\cup B\in\mathscr{I}$, (iii) $A\in\mathscr{I}$, $B\subset A$ implies $B\in\mathscr{I}$.

$\mathscr{I}$ is called a nontrivial ideal if $X\notin\mathscr{I}$. A nontrivial ideal $\mathscr{I}$ in X is called admissible if $\{x\}\in\mathscr{I}$ for each $x\in X$.

A nontrivial ideal $\mathscr{I}_2$ of $\mathbb{N}\times\mathbb{N}$ is called strongly admissible if $\{i\}\times\mathbb{N}$ and $\mathbb{N}\times\{i\}$ belong to $\mathscr{I}_2$ for each $i\in\mathbb{N}$. It is evident that a strongly admissible ideal is also admissible. Throughout the paper we take $\mathscr{I}_2$ as a strongly admissible ideal in $\mathbb{N}\times\mathbb{N}$.

Let $\mathscr{I}_2^0=\{A\subset\mathbb{N}\times\mathbb{N}:(\exists m(A),(i,j)\geq m(A)\Rightarrow(i,j)\notin A)\}$. Then $\mathscr{I}_2^0$ is a nontrivial strongly admissible ideal and clearly an ideal $\mathscr{I}_2$ is strongly admissible if and only if $\mathscr{I}_2^0\subset\mathscr{I}_2$.

Let $X\neq\emptyset$. A non empty class $\mathscr{F}$ of subsets of X is said to be a filter in X provided:

(i) $\emptyset\notin\mathscr{F}$, (ii) $A,B\in\mathscr{F}$ implies $A\cap B\in\mathscr{F}$, (iii) $A\in\mathscr{F}$, $A\subset B$ implies $B\in\mathscr{F}$.

Let $\mathscr{I}$ is a nontrivial ideal in X, $X\neq\emptyset$, then the class $\mathscr{F}(\mathscr{I})=\{M\subset X:(\exists A\in\mathscr{I})(M=X\backslash A)\}$ is a filter on X, called the filter associated with $\mathscr{I}$.

Let (X,ρ) be a linear metric space and $\mathscr{I}_2\subset 2^{\mathbb{N}\times\mathbb{N}}$ be a strongly admissible ideal. A double sequence $x=(x_{mn})$ in X is said to be $\mathscr{I}_2$-convergent to $L\in X$, if for any $\varepsilon>0$ we have $A(\varepsilon)=\{(m,n)\in\mathbb{N}\times\mathbb{N}:\rho(x_{mn},L)\geq\varepsilon\}\in\mathscr{I}_2$ and we write $\mathscr{I}_2-\lim\limits_{m,n\to\infty}x_{mn}=L.$

Let $(X,\|.\|)$ be fuzzy normed space. A sequence $x=(x_m)_{m\in\mathbb{N}}$ in X is said to be $\mathscr{I}-$ convergent to $L\in X$ with respect to fuzzy norm on X if for each $\varepsilon>0$, the set $A(\varepsilon)=\left\{m\in\mathbb{N}:\|x_m-L\|_0^{+}\geq\varepsilon\right\}$ belongs to $\mathscr{I}$. In this case, we write $x_m\overset{F\mathscr{I}}{\longrightarrow}L$. The element L is called the $\mathscr{I}-$limit of (x_m) in X.

Let $(X,\|.\|)$ be a fuzzy normed space. A double sequence $x=(x_{mn})_{(m,n)\in\mathbb{N}\times\mathbb{N}}$ in X is said to be $\mathscr{I}_2-$ convergent to $L_1\in X$ with respect to fuzzy norm on X if for each $\varepsilon>0$, the set $A(\varepsilon)=\left\{(m,n)\in\mathbb{N}\times\mathbb{N}:\|x_{mn}-L_1\|_0^{+}\geq\varepsilon\right\}\in\mathscr{I}_2$. In this case, we write $x_{mn}\overset{F\mathscr{I}_2}{\longrightarrow}L_1$ or $x_{mn}\to L_1\,(F\mathscr{I}_2)$ or $F\mathscr{I}_2-\lim\limits_{m,n\to\infty}x_{mn}=L_1$. The element L_1 is called the $F\mathscr{I}_2-$limit of (x_{mn}) in X. In terms of neighborhoods, we have $x_{mn}\overset{F\mathscr{I}_2}{\longrightarrow}L_1$ provided that for each $\varepsilon>0$, $\{(m,n)\in\mathbb{N}\times\mathbb{N}:x_{mn}\notin\mathscr{N}_{L_1}(\varepsilon,0)\}\in\mathscr{I}_2$. A useful interpretation of the above definition is the following;

$$x_{mn}\overset{F\mathscr{I}_2}{\longrightarrow}L_1\Leftrightarrow F\mathscr{I}_2-\lim_{m,n\to\infty}\|x_{mn}-L_1\|_0^{+}=0.$$

Note that $F\mathscr{I}_2-\lim\limits_{m,n\to\infty}\|x_{mn}-L_1\|_0^{+}=0$ implies that

$$F\mathscr{I}_2-\lim\|x_{mn}-L_1\|_{\alpha}^{-}\quad=\quad F\mathscr{I}_2-\lim\|x_{mn}-L_1\|_{\alpha}^{+}=0,$$

for each $\alpha\in[0,1]$, since $0\leq\|x_{mn}-L_1\|_{\alpha}^{-}\leq\|x_{mn}-L_1\|_{\alpha}^{+}\leq\|x_{mn}-L_1\|_0^{+}$ holds for every $m,n\in\mathbb{N}$ and for each $\alpha\in[0,1]$.

Let $(X,\|.\|)$ be a fuzzy normed space. A double sequence $x=(x_{mn})$ in X is said to be $\mathscr{I}_2$-Cauchy (or $F\mathscr{I}_2$-Cauchy) double sequence with respect to the fuzzy norm on X if, for each $\varepsilon>0$, there exists integers $p=p(\varepsilon)$ and $q=q(\varepsilon)$ such that the set $\left\{(m,n)\in\mathbb{N}\times\mathbb{N}:\|x_{mn}-x_{pq}\|_0^{+}\geq\varepsilon\right\}$ belongs to $\mathscr{I}_2$.

We say that an admissible ideal $\mathscr{I}_2\subset 2^{\mathbb{N}\times\mathbb{N}}$ satisfies the property (AP2), if for every countable family of mutually disjoint sets $\{A_1,A_2,...\}$ belonging to $\mathscr{I}_2$, there exists a countable family of sets $\{B_1,B_2,...\}$ such that $A_j\cap B_j\in\mathscr{I}_2^0$, i.e., $A_j\cap B_j$ is included in the finite union of rows and columns in $\mathbb{N}\times\mathbb{N}$ for each $j\in\mathbb{N}$ and $B=\bigcup_{j=1}^{\infty}B_j\in\mathscr{I}_2$ (hence $B_j\in\mathscr{I}_2$ for each $j\in\mathbb{N}$).

Lemma 1.1. *([27], Theorem 3.3) Let $\{P_i\}_{i=1}^{\infty}$ be a countable collection of subsets of $\mathbb{N}\times\mathbb{N}$ such that $P_i\in\mathscr{F}(\mathscr{I}_2)$ for each i, where $\mathscr{F}(\mathscr{I}_2)$ is a filter associated with a strongly admissible ideal $\mathscr{I}_2$ with the property (AP2). Then there exists a set $P\subset\mathbb{N}\times\mathbb{N}$ such that $P\in\mathscr{F}(\mathscr{I}_2)$ and the set $P\backslash P_i$ is finite for all i.*

Lemma 1.2. *([17], Theorem 3.5) Let $(X,\|.\|)$ be fuzzy normed space and $\mathscr{I}_2$ be a admissible ideal. Then, every $\mathscr{I}_2$-convergent sequence is $\mathscr{I}_2$-Cauchy sequence.*

2. Main results

In this section, we investigate relationship between $\mathscr{I}_2$-convergence and $\mathscr{I}_2$-Cauchy double sequences in fuzzy normed spaces. After, we introduce the concepts of $\mathscr{I}_2^*$-Cauchy double sequences and study relationships between $\mathscr{I}_2$-Cauchy and $\mathscr{I}_2^*$-Cauchy double sequences in fuzzy normed spaces.

Theorem 2.1. *Let $(X,\|.\|)$ be a fuzzy normed space. Then, a double sequence (x_{mn}) is $F\mathscr{I}_2$-convergent if and only if it is $F\mathscr{I}_2$-Cauchy double sequence.*

Proof. Hazarika and Kumar proved that every $F\mathscr{I}_2$-convergent sequence is $F\mathscr{I}_2$-Cauchy sequence in Lemma 1.2.

Assume that (x_{mn}) is $F\mathscr{I}_2$-Cauchy double sequence. We prove that (x_{mn}) is $F\mathscr{I}_2$-convergent. To this effect, let (ε_{pq}) be a strictly decreasing sequence of numbers converging to zero. Since (x_{mn}) is $F\mathscr{I}_2-$Cauchy double sequence, there exist two strictly increasing sequences (k_p) and (l_q) of positive integers such that the set

$$A(\varepsilon_{pq})=\left\{(m,n)\in\mathbb{N}\times\mathbb{N}:\left\|x_{mn}-x_{k_pl_q}\right\|_0^+\geq\varepsilon_{pq}\right\}$$

belongs to $\mathscr{I}_2$, $(p,q\in\mathbb{N})$. This implies that

$$\emptyset\neq\left\{(m,n)\in\mathbb{N}\times\mathbb{N}:\left\|x_{mn}-x_{k_pl_q}\right\|_0^+<\varepsilon_{pq}\right\} \tag{2.1}$$

belongs to $\mathscr{F}(\mathscr{I}_2)$, $(p,q\in\mathbb{N})$. Let p,q,s,t be four positive integers such that $p\neq q$ and $s\neq t$. By (2.1), both the sets

$$D(\varepsilon_{pq})=\{(m,n)\in\mathbb{N}\times\mathbb{N}:\left\|x_{mn}-x_{k_pl_q}\right\|_0^+<\varepsilon_{pq}\}$$

and

$$C(\varepsilon_{st})=\{(m,n)\in\mathbb{N}\times\mathbb{N}:\|x_{mn}-x_{k_sl_t}\|_0^+<\varepsilon_{st}\}$$

are non empty sets in $\mathscr{F}(\mathscr{I}_2)$. Since $\mathscr{F}(\mathscr{I}_2)$ is a filter on $\mathbb{N}\times\mathbb{N}$, therefore $\emptyset\neq D(\varepsilon_{pq})\cap C(\varepsilon_{st})\in\mathscr{F}(\mathscr{I}_2)$. Thus, for each pair (p,q) and (s,t) of positive integers with $p\neq q$ and $s\neq t$, we can select a pair $\left(m_{(p,q),(s,t)},n_{(p,q),(s,t)}\right)\in\mathbb{N}\times\mathbb{N}$ such that

$$\left\|x_{m_{pqst}n_{pqst}}-x_{k_pl_q}\right\|_0^+<\varepsilon_{pq}\text{ and }\left\|x_{m_{pqst}n_{pqst}}-x_{k_sl_t}\right\|_0^+<\varepsilon_{st}.$$

It follows that

$$\begin{aligned}\left\|x_{k_pl_q}-x_{k_sl_t}\right\|_0^+ &\leq \left\|x_{m_{pqst}n_{pqst}}-x_{k_pl_q}\right\|_0^+ + \left\|x_{m_{pqst}n_{pqst}}-x_{k_sl_t}\right\|_0^+\\ &< \varepsilon_{pq}+\varepsilon_{st}\to 0,\text{ as } p,q,s,t\to\infty.\end{aligned}$$

This implies that $\left(x_{k_pl_q}\right)$ $(p,q\in\mathbb{N})$ is a Cauchy double sequence in fuzzy normed space, therefore it satisfies the Cauchy convergence criterion. Thus, the sequence $\left(x_{k_pl_q}\right)$ converges to a finite limit L_1 that is,

$$\lim_{p,q\to\infty}x_{k_pl_q}=L_1.$$

Also, we have $\varepsilon_{pq}\to 0$ as $p,q\to\infty$, so for each $\varepsilon>0$ we can choose the positive integers p_0,q_0 such that for $p\geq p_0$ and $q\geq q_0$,

$$\varepsilon_{p_0q_0}<\frac{\varepsilon}{2}\text{ and }\left\|x_{k_pl_q}-L_1\right\|_0^+<\frac{\varepsilon}{2}. \tag{2.2}$$

Now, we define the set

$$A(\varepsilon)=\left\{(m,n)\in\mathbb{N}\times\mathbb{N}:\|x_{mn}-L_1\|_0^+\geq\varepsilon\right\}.$$

We prove that $A(\varepsilon)\subset A\left(\varepsilon_{p_0q_0}\right)$. Let $(m,n)\in A(\varepsilon)$, then by second half of (2.2) we have

$$\begin{aligned}\varepsilon\leq\|x_{mn}-L_1\|_0^+ &\leq \left\|x_{mn}-x_{k_{p_0}l_{q_0}}\right\|_0^+ + \left\|x_{k_{p_0}l_{q_0}}-L_1\right\|_0^+\\ &\leq \left\|x_{mn}-x_{k_{p_0}l_{q_0}}\right\|_0^+ + \frac{\varepsilon}{2}.\end{aligned}$$

This implies that

$$\frac{\varepsilon}{2} \leq \left\| x_{mn} - x_{k_{p_0} l_{q_0}} \right\|_0^+$$

and therefore by first half of (2.2) we have

$$\varepsilon_{p_0 q_0} \leq \left\| x_{mn} - x_{k_{p_0} l_{q_0}} \right\|_0^+ .$$

This implies that $(m,n) \in A(\varepsilon_{p_0 q_0})$ and therefore $A(\varepsilon)$ is contained in $A(\varepsilon_{p_0 q_0})$. Since $A(\varepsilon_{p_0 q_0})$ belongs to $\mathscr{I}_2$ therefore, $A(\varepsilon)$ belongs to $\mathscr{I}_2$. This proves that (x_{mn}) is $F\mathscr{I}_2$-convergent to L_1. □

Definition 2.2. *Let $(X, \|.\|)$ be a fuzzy normed space. A double sequence $x = (x_{mn})$ in X is said to be $\mathscr{I}_2^*$-Cauchy (or $F\mathscr{I}_2^*$-Cauchy) double sequence with respect to fuzzy norm on X if, there exists a set $M \in \mathscr{F}(\mathscr{I}_2)$ (i.e., $H = \mathbb{N} \times \mathbb{N} \backslash M \in \mathscr{I}_2$) and $k_0 = k_0(\varepsilon)$ such that for every $\varepsilon > 0$ and for $(m,n),(s,t) \in M$, $\|x_{mn} - x_{st}\|_0^+ < \varepsilon$, whenever $m,n,s,t > k_0$. In this case, we write $\lim\limits_{m,n,s,t\to\infty} \|x_{mn} - x_{st}\|_0^+ = 0$.*

Theorem 2.3. *Let $\mathscr{I}_2$ be an admissible ideal of $\mathbb{N} \times \mathbb{N}$. If a double sequence (x_{mn}) in X is an $F\mathscr{I}_2^*$-Cauchy sequence, then it is $F\mathscr{I}_2$-Cauchy sequence.*

Proof. Suppose that (x_{mn}) is an $F\mathscr{I}_2^*$-Cauchy sequence. Then, there exists a set $M \in \mathscr{F}(\mathscr{I}_2)$ (i.e., $H = \mathbb{N} \times \mathbb{N} \backslash M \in \mathscr{I}_2$) and $k_0 = k_0(\varepsilon)$ such that for every $\varepsilon > 0$ and for $(m,n),(s,t) \in M$ $\|x_{mn} - x_{st}\|_0^+ < \varepsilon$, whenever $m,n,s,t \geq k_0$. Then,

$$\begin{aligned} A(\varepsilon) &= \{(m,n) \in \mathbb{N} \times \mathbb{N} : \|x_{mn} - x_{st}\|_0^+ \geq \varepsilon\} \\ &\subset H \cup [M \cap ((\{1,...,k_0\} \times \mathbb{N}) \cup (\mathbb{N} \times \{1,...,k_0\}))]. \end{aligned}$$

Since $\mathscr{I}_2$ be an admissible ideal, then

$$H \cup [M \cap ((\{1,...,k_0\} \times \mathbb{N}) \cup (\mathbb{N} \times \{1,...,k_0\}))] \in \mathscr{I}_2.$$

Therefore, we have $A(\varepsilon) \in \mathscr{I}_2$. This shows that (x_{mn}) is $F\mathscr{I}_2$-Cauchy sequence in X. □

Theorem 2.4. *Let $\mathscr{I}_2$ be an admissible ideal of $\mathbb{N} \times \mathbb{N}$ with the property $(AP2)$ and (x_{mn}) be a double sequence in X. Then, the concepts $\mathscr{I}_2-$Cauchy double sequence with respect to fuzzy norm on X and $\mathscr{I}_2^*$-Cauchy double sequence with respect to fuzzy norm on X coincide.*

Proof. If a double sequence is $F\mathscr{I}_2^*$-Cauchy, then it is $F\mathscr{I}_2$-Cauchy by Theorem 2.3, where $\mathscr{I}_2$ need not have the property $(AP2)$. Now, it is sufficient to prove that a double sequence (x_{mn}) in X is a $F\mathscr{I}_2^*$-Cauchy double sequence under assumption that it is an $F\mathscr{I}_2$-Cauchy double sequence. Let (x_{mn}) be an $F\mathscr{I}_2$-Cauchy double sequence in X. Then, there exists $s = s(\varepsilon), t = t(\varepsilon) \in \mathbb{N}$ such that for every $\varepsilon > 0$,

$$A(\varepsilon) = \left\{(m,n) \in \mathbb{N} \times \mathbb{N} : \|x_{mn} - x_{st}\|_0^+ \geq \varepsilon\right\} \in \mathscr{I}_2.$$

Let

$$P_i = \left\{(m,n) \in \mathbb{N} \times \mathbb{N} : \|x_{mn} - x_{s_i t_i}\|_0^+ < \frac{1}{i}\right\},$$

where $s_i = s(1 \backslash i)$, $(i \in \mathbb{N})$, $t_i = t(1 \backslash i)$. It is clear that $P_i \in \mathscr{F}(\mathscr{I}_2)$ for all $i \in \mathbb{N}$. Since $\mathscr{I}_2$ has the property $(AP2)$, then by Lemma 1.1 there exists a set $P \subset \mathbb{N} \times \mathbb{N}$ such that $P \in \mathscr{F}(\mathscr{I}_2)$ and $P \backslash P_i$ is finite for all $i \in \mathbb{N}$. Now we show that

$$\lim_{m,n,s,t\to\infty} \|x_{mn} - x_{st}\|_0^+ = 0,$$

for $(m,n),(s,t) \in P$. To prove this, let $\varepsilon > 0$ and $j \in \mathbb{N}$ such that $j > 2/\varepsilon$. If $(m,n),(s,t) \in P$ then $P \backslash P_i$ is a finite set, so there exists $N = N(j)$ such that $(m,n),(s,t) \in P_j$ for all $m,n,s,t > N(j)$. Therefore,

$$\|x_{mn} - x_{s_i t_i}\|_0^+ < \frac{1}{j} \text{ and } \|x_{st} - x_{s_i t_i}\|_0^+ < \frac{1}{j},$$

for all $m,n,s,t > N(j)$. Hence it follows that

$$\begin{aligned} \|x_{mn} - x_{st}\|_0^+ &\leq \|x_{mn} - x_{s_it_i}\|_0^+ + \|x_{st} - x_{s_it_i}\|_0^+ \\ &\leq \frac{1}{j} + \frac{1}{j} = \frac{2}{j} < \varepsilon, \end{aligned}$$

for all $m,n,s,t > N(j)$. Thus, for any $\varepsilon > 0$ there exists $N = N(\varepsilon)$ such that for $m,n,s,t > N(j)$ and $(m,n),(s,t) \in P$ we have

$$\|x_{mn} - x_{st}\|_0^+ < \varepsilon.$$

This shows that the double sequence (x_{mn}) in X is an $F\mathscr{I}_2^*$-Cauchy double sequence in fuzzy normed spaces. □

References

[1] H. Fast, *Sur la convergence statistique*, Colloq. Math., **2** (1951), 241–244.

[2] I.J. Schoenberg, *The integrability of certain functions and related summability methods*, Amer. Math. Monthly **66** (1959), 361–375.

[3] M. Mursaleen, O.H.H. Edely, *Statistical convergence of double sequences*, J. Math. Anal. Appl. **288** (2003), 223–231.

[4] D. Rath, B.C. Tripaty, *On statistically convergence and statistically Cauchy sequences*, Indian J. Pure Appl. Math. **25**(4) (1994), 381–386.

[5] P. Kostyrko, T. Šalát, W. Wilczyński, *$\mathscr{I}$-convergence*, Real Anal. Exchange, **26**(2) (2000), 669–686.

[6] P. Das, P. Kostyrko, W. Wilczyński, P. Malik, *$\mathscr{I}$ and $\mathscr{I}^*$-convergence of double sequences*, Math. Slovaca, **58**(5) (2008), 605–620.

[7] P. Das, P. Malik, *On extremal $\mathscr{I}$-limit points of double sequences*, Tatra Mt. Math. Publ., **40** (2008), 91–102.

[8] V. Kumar, *On $\mathscr{I}$ and $\mathscr{I}^*$-convergence of double sequences*, Math. Commun. **12** (2007), 171–181.

[9] T. Šalát, B.C. Tripaty, M. Ziman, *On $\mathscr{I}$-convergence field*, Ital. J. Pure Appl. Math. **17** (2005), 45–54.

[10] B. Tripathy, B.C. Tripathy, *On $\mathscr{I}$-convergent double sequences*, Soochow J. Math. **31** (2005), 549–560.

[11] M. Matloka, *Sequences of fuzzy numbers*, Busefal, **28** (1986), 28–37.

[12] S. Nanda, *On sequences of fuzzy numbers*, Fuzzy Sets Syst. **33** (1989), 123–126.

[13] C. Şençimen, S. Pehlivan, *Statistical convergence in fuzzy normed linear spaces*, Fuzzy Sets and Systems, **159** (2008), 361–370.

[14] B. Hazarika, *On ideal convergent sequences in fuzzy normed linear spaces*, Afrika Mat., **25**(4) (2013), 987–999.

[15] E. Dündar, Ö. Talo, *$\mathscr{I}_2$-convergence of double sequences of fuzzy numbers*, Iran. J. Fuzzy Syst., **10**(3) (2013), 37–50

[16] E. Dündar, Ö. Talo, Ö., *$\mathscr{I}_2$-Cauchy Double Sequences of Fuzzy Numbers*, Gen. Math. Notes, **16**(2) (2013), 103–114.

[17] B. Hazarika, V. Kumar, *Fuzzy real valued $\mathscr{I}$-convergent double sequences in fuzzy normed spaces*, J. Intell. Fuzzy Syst., **26** (2014), 2323–2332.

[18] E. Dündar, M. R. Türkmen, *On $\mathscr{I}_2$-Convergence and $\mathscr{I}_2^*$-Convergence of Double Sequences in Fuzzy Normed Spaces*, Konuralp J. Math., (in press).

[19] V. Kumar, K. Kumar, *On the ideal convergence of sequences of fuzzy numbers*, Inform. Sci. **178** (2008), 4670–4678.

[20] S.A. Mohiuddine, H. Şevli, M. Cancan, *Statistical convergence of double sequences in fuzzy normed spaces*, Filomat, **26**(4) (2012), 673–681.

[21] M. R. Türkmen, E. Dündar, *On Lacunary Statistical Convergence of Double Sequences and Some Properties in Fuzzy Normed Spaces*, J. Intell. Fuzzy Syst., **36**(1) (2019), 467-472.

[22] R. Saadati, *On the $\mathscr{I}$-fuzzy topological spaces*, Chaos, Solitons Fractals, **37** (2008), 1419–1426.

[23] T.Bag, S.K. Samanta, *Fixed point theorems in Felbin's type fuzzy normed linear spaces*, J. Fuzzy Math., **16**(1) (2008), 243–260.

[24] B. Bede, S.G. Gal, *Almost periodic fuzzy-number-valued functions*, Fuzzy Sets Syst., **147**(2004), 385–403.

[25] P. Diamond, P. Kloeden, *Metric Spaces of Fuzzy Sets-Theory and Applications*, World Scientific Publishing, Singapore, **(1994).**

[26] E. Dündar, B. Altay, *$\mathscr{I}_2$-convergence and $\mathscr{I}_2$-Cauchy of double sequences*, Acta Math. Sci., **34**(2) (2014), 343–353.

[27] E. Dündar, B. Altay, *On some properties of $\mathscr{I}_2$-convergence and $\mathscr{I}_2$-Cauchy of double sequences,* Gen. Math. Notes, **7**(1) (2011), 1–12.

[28] J.-X. Fang, H. Huang, *On the level convergence of a sequence of fuzzy numbers*, Fuzzy Sets and Systems, **147** (2004), 417–415.

[29] C. Felbin, *Finite-dimensional fuzzy normed linear space*, Fuzzy Sets and Systems, **48**(2) (1992), 239–248.

[30] Ö. Kişi, *$\mathscr{I}_2 - \lambda^-$-statistically convergence of double sequences in fuzzy normed spaces,* J. Intell. Fuzzy Syst., **36**(2) (2019), 1–12.

[31] M. Mizumoto, K. Tanaka, *Some properties of fuzzy numbers*, Advances in Fuzzy Set Theory and Applications, North-Holland (Amsterdam), **1979**, 153–164.

[32] A. Nabiev, S. Pehlivan, M. Gürdal, *On $\mathscr{I}$-Cauchy sequences*, Taiwanese J. Math. **11**(2) (2007) 569–5764.

[33] A. Pringsheim, *Zur theorie der zweifach unendlichen Zahlenfolgen*, Math. Ann. **53** (1900), 289–321.

[34] E. Savaş, M. Mursaleen, *On statistically convergent double sequences of fuzzy numbers*, Inform. Sci. **162** (2004), 183–192

[35] L.A. Zadeh, *Fuzzy sets*, Inform. Control **8**(1965), 338–353.

Communications in Advanced Mathematical Sciences
Vol. II, No. 2, 161-175, 2019
Research Article
e-ISSN:2651-4001
DOI: 10.33434/cams.556815

Complex Multivariate Montgomery Type Identity Leading to Complex Multivariate Ostrowski and Grüss Inequalities

George A. Anastassiou [1]*

Abstract
We give a general complex multivariate Montgomery type identity which is a representation formula for a complex multivariate function. Using it we produce general tight complex multivariate high order Ostrowski and Grüss type inequalities. The estimates involve L_p norms, any $1 \leq p \leq \infty$. We include also applications.

Keywords: Multivariate complex integral, Multivariate complex continuous functions, Multivariate complex analytic functions, Multivariate complex montgomery identity, Multivariate complex Ostrowski and Grüss inequalities.
2010 AMS: Primary 26D10, Secondery 26D15, 30A10.

[1] *Department of Mathematical Sciences University of Memphis Memphis, TN 38152, U.S.A.*
*__Corresponding author__: ganastss@memphis.edu
Received: 22 April 2019, **Accepted:** 14 June 2019, **Available online:** 27 June 2019

1. Introduction

Our motivation comes from the following results:

Theorem 1.1. *(A. Ostrowski, 1938 [1]). Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b) such that $f' : (a,b) \to \mathbb{R}$ is bounded on (a,b), i.e., $\|f'\|_\infty := \sup_{t\in(a,b)} |f'(t)| < \infty$. Then*

$$\left| \frac{1}{b-a} \int_a^b f(t)\,dt - f(x) \right| \leq \left[\frac{1}{4} + \left(\frac{x - \frac{a+b}{2}}{b-a} \right)^2 \right] \|f'\|_\infty (b-a),$$

for all $x \in [a,b]$ and the constant $\frac{1}{4}$ is the best possible.

Theorem 1.2. *(G. Grüss, 1934 [2]). Let $f, g : [a,b] \to \mathbb{R}$ be Lebesgue integrable functions, and $m, M, n, N \in \mathbb{R}$ such that: $-\infty < m \leq f \leq M < \infty$, $-\infty < n \leq g \leq N < \infty$, a.e. on $[a,b]$. Then*

$$\left| \frac{1}{b-a} \int_a^b f(t)\,g(t)\,dt - \left(\frac{1}{b-a} \int_a^b f(t)\,dt \right) \left(\frac{1}{b-a} \int_a^b g(t)\,dt \right) \right| \leq \frac{1}{4} (M-m)(N-n),$$

with the constant $\frac{1}{4}$ being the best possible.

Let $f \in C^1([a,b])$ and the kernel $p : [a,b]^2 \to \mathbb{R}$ be such that

$$p(x,t) := \begin{cases} t-a, & \text{if } t \in [a,x], \\ t-b, & \text{if } t \in (x,b]. \end{cases}$$

Then, we have the basic Montgomery integral identity [3, p. 565],

$$f(x) = \frac{1}{b-a}\int_a^b f(t)\,dt + \frac{1}{b-a}\int_a^b p(x,t)\,f'(t)\,dt, \ \ \forall\, x \in [a,b].$$

In order to describe complex extensions of Ostrowski and Grüss inequalities using the complex integral we need the following preparation.

Suppose γ is a smooth path parametrized by $z(t)$, $t \in [a,b]$ and f is a complex function which is continuous on γ. Put $z(a) = u$ and $z(b) = w$ with $u, w \in \mathbb{C}$. We define the integral of f on $\gamma_{u,w} = \gamma$ as

$$\int_\gamma f(z)\,dz = \int_{\gamma_{u,w}} f(z)\,dz := \int_a^b f(z(t))\,z'(t)\,dt.$$

We observe that the actual choice of parametrization of γ does not matter.

This definition immediately extends to paths that are piecewise smooth. Suppose γ is parametrized by $z(t)$, $t \in [a,b]$, which is differentiable on the intervals $[a,c]$ and $[c,b]$, then assuming that f is continuous on γ we define

$$\int_{\gamma_{u,w}} f(z)\,dz := \int_{\gamma_{u,v}} f(z)\,dz + \int_{\gamma_{v,w}} f(z)\,dz,$$

where $v := z(c)$. This can be extended for a finite number of intervals.

We also define the integral with respect to arc-length

$$\int_{\gamma_{u,w}} f(z)\,|dz| := \int_a^b f(z(t))\,\left|z'(t)\right| dt$$

and the length of the curve γ is then

$$l(\gamma) = \int_{\gamma_{u,w}} |dz| := \int_a^b \left|z'(t)\right| dt.$$

Let f and g be holomorphic in G, an open domain and suppose $\gamma \subset G$ is a piecewise smooth path from $z(a) = u$ to $z(b) = w$. Then we have the integration by parts formula

$$\int_{\gamma_{u,w}} f(z)\,g'(z)\,dz = f(w)\,g(w) - f(u)\,g(u) - \int_{\gamma_{u,w}} f'(z)\,g(z)\,dz.$$

We recall also the triangle inequality for the complex integral, namely

$$\left|\int_\gamma f(z)\,dz\right| \le \int_\gamma |f(z)|\,|dz| \le \|f\|_{\gamma,\infty}\, l(\gamma),$$

where $\|f\|_{\gamma,\infty} := \sup_{z\in\gamma} |f(z)|$.

We also define the p-norm with $p \ge 1$ by

$$\|f\|_{\gamma,p} := \left(\int_\gamma |f(z)|^p\,|dz|\right)^{\frac{1}{p}}.$$

For $p = 1$ we have

$$\|f\|_{\gamma,1} := \int_\gamma |f(z)|\,|dz|.$$

If $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, then by Hölder's inequality we have

$$\|f\|_{\gamma,1} \le [l(\gamma)]^{\frac{1}{q}}\,\|f\|_{\gamma,p}.$$

First, we mention a Complex extension of Ostrowski inequality to the complex integral by providing upper bounds for the quantity

$$\left| f(v)(w-u) - \int_\gamma f(z)\,dz \right|$$

under the assumption that γ is a smooth path parametrized by $z(t)$, $t \in [a,b]$, $u = z(a)$, $v = z(x)$ with $x \in (a,b)$ and $w = z(b)$ while f is holomorphic in G, an open domain and $\gamma \subset G$.

Secondly, we mention a Complex extension of Grüss inequality:

Suppose $\gamma \subset \mathbb{C}$ is a piecewise smooth path parametrized by $z(t)$, $t \in [a,b]$ from $z(a) = u$ to $z(b) = w$ with $w \neq u$. If f and g are continuous on γ, we consider the complex Čebyšev functional defined by

$$\mathscr{D}_\gamma(f,g) := \frac{1}{w-u}\int_\gamma f(z)\,g(z)\,dz - \frac{1}{w-u}\int_\gamma f(z)\,dz \frac{1}{w-u}\int_\gamma g(z)\,dz.$$

We display upper bounds to $\left|\mathscr{D}_\gamma(f,g)\right|$.

We have the following results for functions of a complex variable:

Theorem 1.3. *(S. Dragomir, 2019 [4]). Let f be holomorhic in G, an open domain and suppose $\gamma \subset G$ is a smooth path from $z(a) = u$ to $z(b) = w$. If $v = z(x)$ with $x \in (a,b)$, then $\gamma_{u,w} = \gamma_{u,v} \cup \gamma_{v,w}$,*

$$\left| f(v)(w-u) - \int_\gamma f(z)\,dz \right| \leq \left\|f'\right\|_{\gamma_{u,v};\infty} \int_{\gamma_{u,v}} |z-u|\,|dz| + \left\|f'\right\|_{\gamma_{v,w};\infty} \int_{\gamma_{u,w}} |z-w|\,|dz| \leq$$

$$\left[\int_{\gamma_{u,v}} |z-u|\,|dz| + \int_{\gamma_{u,w}} |z-w|\,|dz|\right] \left\|f'\right\|_{\gamma_{u,w};\infty},$$

and

$$\left| f(v)(w-u) - \int_\gamma f(z)\,dz \right| \leq \max_{z\in\gamma_{u,v}} |z-u| \left\|f'\right\|_{\gamma_{u,v};1} + \max_{z\in\gamma_{v,w}} |z-w| \left\|f'\right\|_{\gamma_{v,w};1} \leq$$

$$\max\left\{\max_{z\in\gamma_{u,v}} |z-u|, \max_{z\in\gamma_{v,w}} |z-w|\right\} \left\|f'\right\|_{\gamma_{u,w};1}.$$

If $p,q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\left| f(v)(w-u) - \int_\gamma f(z)\,dz \right| \leq \left(\int_{\gamma_{u,v}} |z-u|^q\,|dz|\right)^{\frac{1}{q}} \left\|f'\right\|_{\gamma_{u,v};p} + \left(\int_{\gamma_{u,w}} |z-w|^q\,|dz|\right)^{\frac{1}{q}} \left\|f'\right\|_{\gamma_{v,w};p} \leq$$

$$\left(\int_{\gamma_{u,v}} |z-u|^q\,|dz| + \int_{\gamma_{u,w}} |z-w|^q\,|dz|\right)^{\frac{1}{q}} \left\|f'\right\|_{\gamma_{u,w};p}.$$

Suppose $\gamma \subset \mathbb{C}$ is a piecewise smooth path parametrized by $z(t)$, $t \in [a,b]$ from $z(a) = u$ to $z(b) = w$. Now, for $\phi, \Phi \in \mathbb{C}$ define the set of complex-valued functions

$$\overline{\Delta}_\gamma(\phi,\Phi) := \left\{ f : \gamma \to \mathbb{C} \left| \left| f(z) - \frac{\phi+\Phi}{2} \right| \leq \frac{1}{2}|\Phi-\phi| \text{ for each } z \in \gamma \right.\right\}.$$

We have the following complex Grüss type inequalities:

Theorem 1.4. *(S. Dragomir, 2018 [5]). Suppose $\gamma \subset \mathbb{C}$ is a piecewise smooth path parametrized by $z(t)$, $t \in [a,b]$ from $z(a) = u$ to $z(b) = w$ with $w \neq u$. If f and g are continuous on γ and there exist $\phi, \Phi, \psi, \Psi \in \mathbb{C}$, $\phi \neq \Phi$, $\psi \neq \Psi$ such that $f \in \overline{\Delta}_\gamma(\phi,\Phi)$ and $g \in \overline{\Delta}_\gamma(\psi,\Psi)$ then*

$$\left|\mathscr{D}_\gamma(f,g)\right| \leq \frac{1}{4}|\Phi-\phi|\,|\Psi-\psi| \frac{l^2(\gamma)}{|w-u|^2}.$$

If the path γ is a segment $[u,w]$ connecting two distinct points u and w in $\mathbb{C}$ then we write $\int_\gamma f(z)\,dz$ as $\int_u^w f(z)\,dz$.

If f, g are continuous on $[u,w]$ and there exists $\phi,\Phi,\psi,\Psi\in\mathbb{C}$, $\phi\neq\Phi$, $\psi\neq\Psi$ such that $f\in\overline{\Delta}_{[u,w]}(\phi,\Phi)$ and $g\in\overline{\Delta}_{[u,w]}(\psi,\Psi)$ then

$$\left|\frac{1}{w-u}\int_u^w f(z)\,g(z)\,dz-\frac{1}{w-u}\int_u^w f(z)\,dz\frac{1}{w-u}\int_u^w g(z)\,dz\right|\leq\frac{1}{4}\left|\Phi-\phi\right|\left|\Psi-\psi\right|.$$

We will use the complex Montgomery identity which follows:

Theorem 1.5. *(S. Dragomir, 2018 [4]) Let f be holomorphic in G, an open domain and suppose $\gamma\subset G$ is a smooth path from $z(a)=u$ to $z(b)=w$. If $v=z(t)$ with $t\in[a,b]$, then $\gamma_{u,w}=\gamma_{u,v}\cup\gamma_{v,w}$, and*

$$f(v)=\frac{1}{w-u}\left[\int_\gamma f(z)\,dz+\int_{\gamma_{u,v}}(z-u)\,f'(z)\,dz+\int_{\gamma_{v,w}}(z-w)\,f'(z)\,dz\right].$$

Define

$$p(v,z):=\begin{cases} z-u, & \text{if } z\in\gamma_{u,v}\\ z-w, & \text{if } z\in\gamma_{v,w}.\end{cases}$$

Thus, it holds

$$f(v)=\frac{1}{w-u}\int_\gamma f(z)\,dz+\frac{1}{w-u}\int_\gamma p(v,z)\,f'(z)\,dz, \tag{1.1}$$

a form which we will use a lot in this article.

Representation formula (1.1) is the main inspiration to write this article.

We will use (1.1) to derive a multivariate Complex Montgomery type identity then based on it, we will produce Complex multivariate Ostrowski and Grüss type inequalities.

For the last we need:

Definition 1.6. *Here we extend the notion of line (curve) integral into multivariate case. Let γ_j, $j=1,\dots,m$, be a smooth path parametrized by $z_j(t_j)$, $t_j\in[a_j,b_j]$ and f is a complex valued function which is continuous on $\prod_{j=1}^m\gamma_j\subseteq\mathbb{C}^m$. Put $z_j(a_j)=u_j$ and $z_j(b_j)=w_j$, with $u_j,w_j\in\mathbb{C}$, $j=1,\dots,m$.*

We define the complex multivariate integral of f on $\prod_{j=1}^m\gamma_j:=\prod_{j=1}^m\gamma_{u_j,w_j}$ as

$$\int_{\gamma_1}\dots\int_{\gamma_m}f(z_1,\dots,z_m)\,dz_1\dots dz_m:=\int_{\prod_{j=1}^m\gamma_j}f(z_1,\dots,z_m)\,dz_1\dots dz_m:=$$

$$\int_{\gamma_{u_1,w_1}}\dots\int_{\gamma_{u_m,w_m}}f(z_1,\dots,z_m)\,dz_1\dots dz_m:=\int_{\prod_{j=1}^m\gamma_{u_j,w_j}}f(z_1,\dots,z_m)\,dz_1\dots dz_m:=$$

$$\int_{a_1}^{b_1}\int_{a_2}^{b_2}\dots\int_{a_m}^{b_m}f(z_1(t_1),\dots,z_m(t_m))\prod_{j=1}^m z_j'(t_j)\,dt_1\dots dt_m. \tag{1.2}$$

We make

Remark 1.7. *Clearly here $z_j\in C^1([a_j,b_j],\mathbb{C})$, $j=1,\dots,m$. The integrand in (1.2) is a continuous complex valued function over $\prod_{j=1}^m[a_j,b_j]$. Therefore $|f(z_1(t_1),\dots,z_m(t_m))|\left|\prod_{j=1}^m z_j'(t_j)\right|$ is also continuous but from $\prod_{j=1}^m[a_j,b_j]$ into $\mathbb{R}$, hence it is bounded. Consequently it holds*

$$\int_{\prod_{j=1}^m[a_j,b_j]}|f(z_1(t_1),\dots,z_m(t_m))|\prod_{j=1}^m\left|z_j'(t_j)\right|\prod_{j=1}^m dt_j<+\infty.$$

Therefore, by Fubini's theorem, the order integration in (1.2) is immaterial.

Clearly it holds

$$\left|\int_{a_1}^{b_1}\dots\int_{a_m}^{b_m} f\left(z_1\left(t_1\right),\dots,z_m\left(t_m\right)\right)\prod_{j=1}^{m} z_j'\left(t_j\right)dt_1\dots dt_m\right| \leq \int_{a_1}^{b_1}\dots\int_{a_m}^{b_m}\left|f\left(z_1\left(t_1\right),\dots,z_m\left(t_m\right)\right)\right|\prod_{j=1}^{m}\left|z_j'\left(t_j\right)\right|dt_1\dots dt_m. \tag{1.3}$$

We also define the integral with respect to arc-lengths

$$\int_{\prod_{j=1}^{m}\gamma_{u_j,w_j}} f\left(z_1,\dots,z_m\right)\left|dz_1\right|\left|dz_2\right|\dots\left|dz_m\right| := \int_{\prod_{j=1}^{m}\left[a_j,b_j\right]} f\left(z_1\left(t_1\right),\dots,z_m\left(t_m\right)\right)\prod_{j=1}^{m}\left|z_j'\left(t_j\right)\right|dt_1\dots dt_m. \tag{1.4}$$

It holds (by (1.3), (1.4))

$$\left|\int_{\prod_{j=1}^{m}\gamma_j} f\left(z_1,\dots,z_m\right)dz_1\dots dz_m\right| \leq \int_{\prod_{j=1}^{m}\gamma_{u_j,w_j}}\left|f\left(z_1,\dots,z_m\right)\right|\left|dz_1\right|\left|dz_2\right|\dots\left|dz_m\right| \leq \|f\|_{\prod_{j=1}^{m}\gamma_j,\infty}\prod_{j=1}^{m} l\left(\gamma_j\right),$$

where

$$\|f\|_{\prod_{j=1}^{m}\gamma_j,\infty} := \sup_{\left(z_1,\dots,z_m\right)\in\prod_{j=1}^{m}\gamma_j}\left|f\left(z_1,\dots,z_m\right)\right|,$$

and

$$l\left(\gamma_j\right) = \int_{\gamma_{u_j,w_j}}\left|dz_j\right| = \int_{a_j}^{b_j}\left|z_j'\left(t_j\right)\right|dt_j, \quad j=1,\dots,m.$$

We also define the p-norm with $p \geq 1$ by

$$\|f\|_{\prod_{j=1}^{m}\gamma_j,p} := \left(\int_{\prod_{j=1}^{m}\gamma_j}\left|f\left(z_1,\dots,z_m\right)\right|^p\left|dz_1\right|\left|dz_2\right|\dots\left|dz_m\right|\right)^{\frac{1}{p}}.$$

For $p=1$ we have

$$\|f\|_{\prod_{j=1}^{m}\gamma_j,1} := \int_{\prod_{j=1}^{m}\gamma_j}\left|f\left(z_1,\dots,z_m\right)\right|\left|dz_1\right|\left|dz_2\right|\dots\left|dz_m\right|.$$

If $p,q>1$ with $\frac{1}{p}+\frac{1}{q}=1$, then by Hölder's inequality we have

$$\|f\|_{\prod_{j=1}^{m}\gamma_j,1} \leq \left(\prod_{j=1}^{m} l\left(\gamma_j\right)\right)^{\frac{1}{q}}\|f\|_{\prod_{j=1}^{m}\gamma_j,p}.$$

2. Main results

We start by presenting a complex trivariate Montgomery type representation identity of complex functions:

Theorem 2.1. *Let $f:\prod_{j=1}^{3} D_j \subseteq \mathbb{C}^3 \to \mathbb{C}$ be a continuous function that is analytic per coordinate on the domain D_j, $j=1,2,3$, and $x=\left(x_1,x_2,x_3\right)\in\prod_{j=1}^{3} D_j$. For $j=1,2,3$, suppose $\gamma_j \subset D_j$ is a smooth path parametrized by $z_j\left(t_j\right)$, $t_j\in\left[a_j,b_j\right]$ with $z_j\left(a_j\right)=u_j$, $z_j\left(t_j\right)=x_j$ and $z_j\left(b_j\right)=w_j$, where $u_j,w_j\in D_j$, $u_j\neq w_j$. Assume also that all partial derivatives of f up to order three are continuous functions on $\prod_{j=1}^{3} D_j$.*

Here we define the kernels for $i = 1,2,3$, $p_i : \gamma_i^2 \to \mathbb{C}$

$$p_i(x_i, s_i) := \begin{cases} s_i - u_i, & \text{if } s_i \in \gamma_{u_i,x_i}, \\ s_i - w_i, & \text{if } s_i \in \gamma_{x_i,w_i}. \end{cases}$$

Then

$$f(x_1,x_2,x_3) = \frac{1}{\prod_{i=1}^{3}(w_i-u_i)} \left\{ \int_{\gamma_1}\int_{\gamma_2}\int_{\gamma_3} f(s_1,s_2,s_3)\,ds_3ds_2ds_1 + \sum_{j=1}^{3}\left(\int_{\gamma_1}\int_{\gamma_2}\int_{\gamma_3} p_j(x_j,s_j)\frac{\partial f(s_1,s_2,s_3)}{\partial s_j} ds_3ds_2ds_1 \right) \right.$$

$$\left. + \sum_{\substack{l=1 \\ j<k}}^{3} \left(\int_{\gamma_1}\int_{\gamma_2}\int_{\gamma_3} p_j(x_j,s_j)\,p_k(x_k,s_k)\frac{\partial^2 f(s_1,s_2,s_3)}{\partial s_k \partial s_j} ds_3ds_2ds_1 \right)(l) + \int_{\gamma_1}\int_{\gamma_2}\int_{\gamma_3} \left(\prod_{i=1}^{3} p_i(x_i,s_i) \right) \frac{\partial^3 f(s_1,s_2,s_3)}{\partial s_3 \partial s_2 \partial s_1} ds_3ds_2ds_1 \right\}. \tag{2.1}$$

Above l counts $(j,k) : j<k;\ j,k \in \{1,2,3\}$.

Proof. Here we apply (1.1) repeatedly.

First we see that

$$f(x_1,x_2,x_3) = A_0 + B_0,$$

where

$$A_0 := \frac{1}{w_1-u_1}\int_{\gamma_1} f(s_1,x_2,x_3)\,ds_1,$$

and

$$B_0 := \frac{1}{w_1-u_1}\int_{\gamma_1} p_1(x_1,s_1)\frac{\partial f(s_1,x_2,x_3)}{\partial s_1} ds_1.$$

Furthermore we have

$$f(s_1,x_2,x_3) = A_1 + B_1,$$

where

$$A_1 := \frac{1}{w_2-u_2}\int_{\gamma_2} f(s_1,s_2,x_3)\,ds_2,$$

and

$$B_1 := \frac{1}{w_2-u_2}\int_{\gamma_2} p_2(x_2,s_2)\frac{\partial f(s_1,s_2,x_3)}{\partial s_2} ds_2.$$

Also we find that

$$f(s_1,s_2,x_3) = \frac{1}{w_3-u_3}\int_{\gamma_3} f(s_1,s_2,s_3)\,ds_3 +$$

$$\frac{1}{w_3-u_3}\int_{\gamma_3} p_3(x_3,s_3)\frac{\partial f(s_1,s_2,s_3)}{\partial s_3} ds_3.$$

Next we put things together, and we derive

$$A_1 = \frac{1}{(w_2-u_2)(w_3-u_3)}\int_{\gamma_2}\int_{\gamma_3} f(s_1,s_2,s_3)\,ds_3ds_2 + \frac{1}{(w_2-u_2)(w_3-u_3)}\int_{\gamma_2}\int_{\gamma_3} p_3(x_3,s_3)\frac{\partial f(s_1,s_2,s_3)}{\partial s_3} ds_3ds_2.$$

And we get

$$A_0 = \frac{1}{\prod_{i=1}^{3}(w_i - u_i)} \int_{\gamma_1}\int_{\gamma_2}\int_{\gamma_3} f(s_1,s_2,s_3)\, ds_3 ds_2 ds_1 + \frac{1}{\prod_{i=1}^{3}(w_i - u_i)} \int_{\gamma_1}\int_{\gamma_2}\int_{\gamma_3} p_3(x_3,s_3) \frac{\partial f(s_1,s_2,s_3)}{\partial s_3} ds_3 ds_2 ds_1$$

$$+\frac{1}{(w_1-u_1)(w_2-u_2)} \int_{\gamma_1}\int_{\gamma_2} p_2(x_2,s_2) \frac{\partial f(s_1,s_2,x_3)}{\partial s_2} ds_2 ds_1.$$

Also we obtain

$$\frac{\partial f(s_1,s_2,x_3)}{\partial s_2} = \frac{1}{w_3-u_3}\int_{\gamma_3} \frac{\partial f(s_1,s_2,s_3)}{\partial s_2} ds_3 + \frac{1}{w_3-u_3}\int_{\gamma_3} p_3(x_3,s_3) \frac{\partial^2 f(s_1,s_2,s_3)}{\partial s_3 \partial s_2} ds_3.$$

Therefore we get

$$A_0 = \frac{1}{\prod_{i=1}^{3}(w_i - u_i)} \int_{\gamma_1}\int_{\gamma_2}\int_{\gamma_3} f(s_1,s_2,s_3)\, ds_3 ds_2 ds_1 + \frac{1}{\prod_{i=1}^{3}(w_i - u_i)} \int_{\gamma_1}\int_{\gamma_2}\int_{\gamma_3} p_3(x_3,s_3) \frac{\partial f(s_1,s_2,s_3)}{\partial s_3} ds_3 ds_2 ds_1$$

$$+\frac{1}{\prod_{i=1}^{3}(w_i - u_i)} \int_{\gamma_1}\int_{\gamma_2}\int_{\gamma_3} p_2(x_2,s_2) \frac{\partial f(s_1,s_2,s_3)}{\partial s_2} ds_3 ds_2 ds_1 + \frac{1}{\prod_{i=1}^{3}(w_i - u_i)} \int_{\gamma_1}\int_{\gamma_2}\int_{\gamma_3} p_2(x_2,s_2)\, p_3(x_3,s_3) \frac{\partial^2 f(s_1,s_2,s_3)}{\partial s_3 \partial s_2} ds_3 ds_2 ds_1.$$

Similarly we obtain that

$$B_0 = \frac{1}{\prod_{i=1}^{3}(w_i - u_i)} \int_{\gamma_1}\int_{\gamma_2}\int_{\gamma_3} p_1(x_1,s_1) \frac{\partial f(s_1,s_2,s_3)}{\partial s_1} ds_3 ds_2 ds_1 +$$

$$\frac{1}{\prod_{i=1}^{3}(w_i - u_i)} \int_{\gamma_1}\int_{\gamma_2}\int_{\gamma_3} p_1(x_1,s_1)\, p_3(x_3,s_3) \frac{\partial^2 f(s_1,s_2,s_3)}{\partial s_3 \partial s_1} ds_3 ds_2 ds_1 +$$

$$\frac{1}{\prod_{i=1}^{3}(w_i - u_i)} \int_{\gamma_1}\int_{\gamma_2}\int_{\gamma_3} p_1(x_1,s_1)\, p_2(x_2,s_2) \frac{\partial^2 f(s_1,s_2,s_3)}{\partial s_2 \partial s_1} ds_3 ds_2 ds_1 +$$

$$\frac{1}{\prod_{i=1}^{3}(w_i - u_i)} \int_{\gamma_1}\int_{\gamma_2}\int_{\gamma_3} p_1(x_1,s_1)\, p_2(x_2,s_2)\, p_3(x_3,s_3) \frac{\partial^3 f(s_1,s_2,s_3)}{\partial s_3 \partial s_2 \partial s_1} ds_3 ds_2 ds_1.$$

We have proved (2.1). □

Next comes the general complex multivariate Montgomery type representation identity of complex functions:

Theorem 2.2. *Let* $f : \prod_{j=1}^{m} D_j \subseteq \mathbb{C}^m \to \mathbb{C}$ *be a continuous function that is analytic per coordinate on the domain* D_j, $j = 1,...,m$, *and* $x = (x_1,...,x_m) \in \prod_{j=1}^{m} D_j$. *For* $j = 1,...,m$, *suppose* $\gamma_j \subset D_j$ *is a smooth path parametrized by* $z_j(t_j)$, $t_j \in [a_j, b_j]$ *with* $z_j(a_j) = u_j$, $z_j(t_j) = x_j$ *and* $z_j(b_j) = w_j$, *where* $u_j, w_j \in D_j$, $u_j \neq w_j$. *Assume also that all partial derivatives of* f *up to order* $m \in \mathbb{N}$ *are continuous functions on* $\prod_{j=1}^{m} D_j$.

We define the kernels $p_i : \gamma_i^2 \to \mathbb{C}$

$$p_i(x_i,s_i) := \begin{cases} s_i - u_i, & \text{if } s_i \in \gamma_{u_i,x_i}, \\ s_i - w_i, & \text{if } s_i \in \gamma_{x_i,w_i}, \end{cases}$$

for $i = 1,2,...,m$.

Then

$$f(x_1,x_2,...,x_m) = \frac{1}{\prod_{i=1}^{m}(w_i-u_i)}\left\{\int_{\prod_{i=1}^{m}\gamma_i} f(s_1,s_2,...,s_m)\,ds_m ds_{m-1}...ds_1 + \sum_{j=1}^{m}\left(\int_{\prod_{i=1}^{m}\gamma_i} p_j(x_j,s_j)\frac{\partial f(s_1,s_2,...,s_m)}{\partial s_j}ds_m...ds_1\right)+\right.$$

$$\sum_{\substack{l_1=1\\ j<k}}^{\binom{m}{2}}\left(\int_{\prod_{i=1}^{m}\gamma_i} p_j(x_j,s_j)\,p_k(x_k,s_k)\frac{\partial^2 f(s_1,s_2,...,s_m)}{\partial s_k\partial s_j}ds_m...ds_1\right)_{(l_1)}+$$

$$\sum_{\substack{l_2=1\\ j<k<r}}^{\binom{m}{3}}\left(\int_{\prod_{i=1}^{m}\gamma_i} p_j(x_j,s_j)\,p_k(x_k,s_k)\,p_r(x_r,s_r)\frac{\partial^3 f(s_1,...,s_m)}{\partial s_r\partial s_k\partial s_j}ds_m...ds_1\right)_{(l_2)}+...+$$

$$\sum_{l=1}^{\binom{m}{m-1}}\left(\int_{\prod_{i=1}^{m}\gamma_i} p_1(x_1,s_1)...\widehat{p_l(x_l,s_l)}...p_m(x_m,s_m)\frac{\partial^{m-1} f(s_1,,...,s_m)}{\partial s_m...\widehat{\partial s_l}...\partial s_1}ds_m...\widehat{ds_l}...ds_1\right)$$

$$\left.+\int_{\prod_{i=1}^{m}\gamma_i}\left(\prod_{i=1}^{m}p_i(x_i,s_i)\right)\frac{\partial^m f(s_1,,...,s_m)}{\partial s_m...\partial s_1}ds_m...ds_1\right\}. \tag{2.2}$$

Above l_1 counts $(j,k): j<k;\ j,k\in\{1,2,...,m\}$, also l_2 counts $(j,k,r): j<k<r;\ j,k,r\in\{1,2,...,m\}$, etc. Also $\widehat{p_l(x_l,s_l)}$ and $\widehat{\partial s_l}$ means that $p_l(x_l,s_l)$ and ∂s_l are missing, respectively.

Proof. Similar to Theorem 2.1. □

We make

Remark 2.3. *(on Theorems 2.1, 2.2)*
By (2.1) we get

$$E_f(x_1,x_2,x_3) := f(x_1,x_2,x_3) - \frac{1}{\prod_{i=1}^{3}(w_i-u_i)}\left\{\int_{\prod_{i=1}^{3}\gamma_i} f(s_1,s_2,s_3)\,ds_3ds_2ds_1\right.$$

$$\left.-\sum_{j=1}^{3}\left(\int_{\prod_{i=1}^{3}\gamma_i} p_j(x_j,s_j)\frac{\partial f(s_1,s_2,s_3)}{\partial s_j}ds_3ds_2ds_1\right) - \sum_{\substack{l=1\\ j<k}}^{3}\left(\int_{\prod_{i=1}^{3}\gamma_i} p_j(x_j,s_j)\,p_k(x_k,s_k)\frac{\partial^2 f(s_1,s_2,s_3)}{\partial s_k\partial s_j}ds_3ds_2ds_1\right)(l)\right\}$$

$$= \frac{1}{\prod_{i=1}^{3}(w_i-u_i)}\left(\int_{\prod_{i=1}^{3}\gamma_i}\left(\prod_{i=1}^{3}p_i(x_i,s_i)\right)\frac{\partial^3 f(s_1,s_2,s_3)}{\partial s_3\partial s_2\partial s_1}ds_1ds_2ds_3\right).$$

Above l counts $(j,k): j<k;\ j,k\in\{1,2,3\}$.
Similarly, by (2.2) we find

$$E_f(x_1,x_2,...,x_m) = f(x_1,x_2,...,x_m) -$$

$$\frac{1}{\prod_{i=1}^{m}(w_i-u_i)}\left\{\int_{\prod_{i=1}^{m}\gamma_i} f(s_1,...,s_m)\,ds_1...ds_m - \sum_{j=1}^{m}\left(\int_{\prod_{i=1}^{m}\gamma_i} p_j(x_j,s_j)\frac{\partial f(s_1,...,s_m)}{\partial s_j}ds_1...ds_m\right)-\right.$$

$$\sum_{\substack{l_1=1\\ j<k}}^{\binom{m}{2}} \left(\int_{\prod_{i=1}^{m} \gamma_i} p_j(x_j,s_j)\, p_k(x_k,s_k) \frac{\partial^2 f(s_1,\dots,s_m)}{\partial s_k \partial s_j} ds_1 \dots ds_m \right)_{(l_1)} -$$

$$\sum_{\substack{l_2=1\\ j<k<r}}^{\binom{m}{3}} \left(\int_{\prod_{i=1}^{m} \gamma_i} p_j(x_j,s_j)\, p_k(x_k,s_k)\, p_r(x_r,s_r) \frac{\partial^3 f(s_1,\dots,s_m)}{\partial s_r \partial s_k \partial s_j} ds_1 \dots ds_m \right)_{(l_2)} - \dots -$$

$$\left. \sum_{l=1}^{\binom{m}{m-1}} \left(\int_{\prod_{i=1}^{m} \gamma_i} p_1(x_1,s_1) \dots \widehat{p_l(x_l,s_l)} \dots p_m(x_m,s_m) \frac{\partial^{m-1} f(s_1,,\dots,s_m)}{\partial s_m \dots \widehat{\partial s_l} \dots \partial s_1} ds_1 \dots \widehat{ds_l} \dots ds_m \right) \right\}$$

$$= \frac{1}{\prod_{i=1}^{m}(w_i-u_i)} \left(\int_{\prod_{i=1}^{m} \gamma_i} \left(\prod_{i=1}^{m} p_i(x_i,s_i) \right) \frac{\partial^m f(s_1,,\dots,s_m)}{\partial s_m \dots \partial s_1} ds_1 \dots ds_m \right).$$

Above l_1 counts $(j,k) : j<k;\ j,k \in \{1,\dots,m\}$, l_2 counts $(j,k,r) : j<k<r;\ j,k,r \in \{1,2,\dots,m\}$, etc. Also $\widehat{p_l(x_l,s_l)}$ and $\widehat{\partial s_l}$ means that $p_l(x_l,s_l)$ and ∂s_l are missing, respectively.

Hence it holds

$$\left| E_f(x_1,x_2,x_3) \right| \le \frac{1}{\prod_{i=1}^{3} |w_i-u_i|} \times \left(\int_{\prod_{i=1}^{3} \gamma_i} \left(\prod_{i=1}^{3} |p_i(x_i,s_i)| \right) \left| \frac{\partial^3 f(s_1,s_2,s_3)}{\partial s_3 \partial s_2 \partial s_1} \right| |ds_1|\,|ds_2|\,|ds_3| \right), \tag{2.3}$$

and

$$\left| E_f(x_1,\dots,x_m) \right| \le \frac{1}{\prod_{i=1}^{m} |w_i-u_i|} \times \left(\int_{\prod_{i=1}^{m} \gamma_i} \left(\prod_{i=1}^{m} |p_i(x_i,s_i)| \right) \left| \frac{\partial^m f(s_1,,\dots,s_m)}{\partial s_m \dots \partial s_1} \right| |ds_1| \dots |ds_m| \right). \tag{2.4}$$

We give the following complex multivariate Ostrowski type inequalities:

Theorem 2.4. *All as in Theorem 2.1. Here $r_1,r_2,r_3,r_4>0 : \frac{1}{r_1}+\frac{1}{r_2}+\frac{1}{r_3}+\frac{1}{r_4}=1$. Then*

$$\left| E_f(x_1,x_2,x_3) \right| \le \frac{1}{\prod_{i=1}^{3} |w_i-u_i|} \times \min \left\{ \left(\prod_{i=1}^{3} \int_{\gamma_i} |p_i(x_i,s_i)|\,|ds_i| \right) \left\| \frac{\partial^3 f}{\partial s_3 \partial s_2 \partial s_1} \right\|_{\infty, \prod_{j=1}^{3} \gamma_j}, \right.$$

$$\left(\prod_{i=1}^{3} \|p_i(x_i,s_i)\|_{r_j,\gamma_j} \right) \left(\prod_{i=1}^{3} (l(\gamma_j))^{\frac{2}{r_j}} \right) \left\| \frac{\partial^3 f}{\partial s_3 \partial s_2 \partial s_1} \right\|_{r_4, \prod_{i=1}^{3} \gamma_j},$$

$$\left. \left(\sup_{(s_1,s_2,s_3) \in \prod_{j=1}^{3} \gamma_j} \left(\prod_{i=1}^{3} |p_i(x_i,s_i)| \right) \right) \left\| \frac{\partial^3 f}{\partial s_3 \partial s_2 \partial s_1} \right\|_{1, \prod_{j=1}^{3} \gamma_j} \right\},$$

$\forall\ (x_1,x_2,x_3) \in \prod_{j=1}^{3} \gamma_j$.

Proof. By (2.3) and generalized Hölder's inequality. □

Theorem 2.5. *All as in Theorem 2.2. Here* $r_1, r_2, ..., r_m, r_{m+1} > 0 : \sum_{i=1}^{m+1} \frac{1}{r_i} = 1$. *Then*

$$\left|E_f(x_1,...,x_m)\right| \leq \frac{1}{\prod_{i=1}^{m} |w_i - u_i|} \times \min\left\{ \left(\prod_{i=1}^{m} \int_{\gamma_i} |p_i(x_i,s_i)|\,|ds_i|\right) \left\|\frac{\partial^m f}{\partial s_m ... \partial s_1}\right\|_{\infty, \prod_{j=1}^{m}\gamma_j},\right.$$

$$\left(\prod_{i=1}^{m} \|p_i(x_i,s_i)\|_{r_j,\gamma_j}\right)\left(\prod_{i=1}^{m} (l(\gamma_j))^{\frac{m-1}{r_j}}\right)\left\|\frac{\partial^m f}{\partial s_m ... \partial s_1}\right\|_{r_{m+1}, \prod_{i=1}^{m}\gamma_j},$$

$$\left.\left(\sup_{(s_1,...,s_m)\in\prod_{j=1}^{m}\gamma_j}\left(\prod_{i=1}^{m}|p_i(x_i,s_i)|\right)\right)\left\|\frac{\partial^m f}{\partial s_m ... \partial s_1}\right\|_{1, \prod_{j=1}^{m}\gamma_j}\right\},$$

$\forall\, (x_1,...,x_m) \in \prod_{j=1}^{m} \gamma_j$.

Proof. By (2.4) and generalized Hölder's inequality. □

We make

Remark 2.6. *Working further on (2.1) we call:*

$$A_f^{(3)} := A_f^{(3)}(x_1,x_2,x_3) := \sum_{j=1}^{3}\left(\int_{\prod_{i=1}^{3}\gamma_i} p_j(x_j,s_j)\frac{\partial f(s_1,s_2,s_3)}{\partial s_j} ds_1 ds_2 ds_3\right)$$

$$+\sum_{\substack{l=1\\ j<k}}^{3}\left(\int_{\prod_{i=1}^{3}\gamma_i} p_j(x_j,s_j)\, p_k(x_k,s_k)\frac{\partial^2 f(s_1,s_2,s_3)}{\partial s_k \partial s_j} ds_1 ds_2 ds_3\right)(l),$$

and

$$B_f^{(3)} := B_f^{(3)}(x_1,x_2,x_3) := \int_{\prod_{i=1}^{3}\gamma_i}\left(\prod_{i=1}^{3} p_i(x_i,s_i)\right)\frac{\partial^3 f(s_1,s_2,s_3)}{\partial s_3 \partial s_2 \partial s_1} ds_1 ds_2 ds_3.$$

Set also

$$T_f^{(3)} := T_f^{(3)}(x_1,x_2,x_3) := A_f^{(3)} + B_f^{(3)}.$$

Thus, we have ($x = (x_1,x_2,x_3)$)

$$f(x) = f(x_1,x_2,x_3) = \frac{1}{\prod_{i=1}^{3}(w_i - u_i)}\int_{\prod_{i=1}^{3}\gamma_i} f(s_1,s_2,s_3)\, ds_1 ds_2 ds_3 + \frac{1}{\prod_{i=1}^{3}(w_i - u_i)}\left(A_f^{(3)} + B_f^{(3)}\right) =$$

$$\frac{1}{\prod_{i=1}^{3}(w_i - u_i)}\int_{\prod_{i=1}^{3}\gamma_i} f(s_1,s_2,s_3)\, ds_1 ds_2 ds_3 + \frac{1}{\prod_{i=1}^{3}(w_i - u_i)} T_f^{(3)}.$$

Working further on (2.2) we call:

$$A_f^{(m)} := A_f^{(m)}(x_1,...,x_m) := \sum_{j=1}^{m}\left(\int_{\prod_{i=1}^{m}\gamma_i} p_j(x_j,s_j)\frac{\partial f(s_1,...,s_m)}{\partial s_j} ds_1 ... ds_m\right) +$$

$$\sum_{\substack{l_1=1\\ j<k}}^{\binom{m}{2}} \left(\int_{\prod_{i=1}^{m}\gamma_i} p_j(x_j,s_j)\, p_k(x_k,s_k) \frac{\partial^2 f(s_1,\dots,s_m)}{\partial s_k \partial s_j} ds_1 \dots ds_m \right)_{(l_1)} +$$

$$\sum_{\substack{l_2=1\\ j<k<r}}^{\binom{m}{3}} \left(\int_{\prod_{i=1}^{m}\gamma_i} p_j(x_j,s_j)\, p_k(x_k,s_k)\, p_r(x_r,s_r) \frac{\partial^3 f(s_1,\dots,s_m)}{\partial s_r \partial s_k \partial s_j} ds_1 \dots ds_m \right)(l_2) + \dots +$$

$$\sum_{l=1}^{\binom{m}{m-1}} \left(\int_{\prod_{i=1}^{m}\gamma_i} p_1(x_1,s_1) \dots \widehat{p_l(x_l,s_l)} \dots p_m(x_m,s_m) \frac{\partial^{m-1} f(s_1,,\dots,s_m)}{\partial s_m \dots \widehat{\partial s_l} \dots \partial s_1} ds_1 \dots \widehat{ds_l} \dots ds_m \right),$$

and

$$B_f^{(m)} := B_f^{(m)}(x_1,\dots,x_m) := \int_{\prod_{i=1}^{m}\gamma_i} \left(\prod_{i=1}^{m} p_i(x_i,s_i) \right) \frac{\partial^m f(s_1,,\dots,s_m)}{\partial s_m \dots \partial s_1} ds_1 \dots ds_m.$$

Set also

$$T_f^{(m)} := T_f^{(m)}(x_1,\dots,x_m) := A_f^{(m)} + B_f^{(m)}.$$

Thus, we have ($x = (x_1,\dots,x_m)$)

$$f(x) = f(x_1,\dots,x_m) = \frac{1}{\prod_{i=1}^{m}(w_i-u_i)} \int_{\prod_{i=1}^{m}\gamma_i} f(s_1,\dots,s_m)\, ds_1 \dots ds_m + \frac{1}{\prod_{i=1}^{m}(w_i-u_i)} \left(A_f^{(m)} + B_f^{(m)} \right) =$$

$$\frac{1}{\prod_{i=1}^{m}(w_i-u_i)} \int_{\prod_{i=1}^{m}\gamma_i} f(s_1,\dots,s_m)\, ds_1 \dots ds_m + \frac{1}{\prod_{i=1}^{m}(w_i-u_i)} T_f^{(m)}. \tag{2.5}$$

Let function g as in Theorem 2.2. Then as in (2.5) we obtain

$$g(x) = g(x_1,\dots,x_m) = \frac{1}{\prod_{i=1}^{m}(w_i-u_i)} \int_{\prod_{i=1}^{m}\gamma_i} g(s_1,\dots,s_m)\, ds_1 \dots ds_m + \frac{1}{\prod_{i=1}^{m}(w_i-u_i)} \left(A_g^{(m)} + B_g^{(m)} \right) = \tag{2.6}$$

$$\frac{1}{\prod_{i=1}^{m}(w_i-u_i)} \int_{\prod_{i=1}^{m}\gamma_i} g(s_1,\dots,s_m)\, ds_1 \dots ds_m + \frac{1}{\prod_{i=1}^{m}(w_i-u_i)} T_g^{(m)}.$$

Above $A_g^{(m)}, B_g^{(m)}, T_g^{(m)}$ have the obvious meaning.

By (2.5) we get

$$f(x)\, g(x) = \frac{g(x)}{\prod_{i=1}^{m}(w_i-u_i)} \int_{\prod_{i=1}^{m}\gamma_i} f(s_1,\dots,s_m) \prod_{i=1}^{m} ds_i + \frac{g(x)}{\prod_{i=1}^{m}(w_i-u_i)} T_f^{(m)},$$

and by (2.6) we get

$$g(x)\, f(x) = \frac{f(x)}{\prod_{i=1}^{m}(w_i-u_i)} \int_{\prod_{i=1}^{m}\gamma_i} g(s_1,\dots,s_m) \prod_{i=1}^{m} ds_i + \frac{f(x)}{\prod_{i=1}^{m}(w_i-u_i)} T_g^{(m)}.$$

Consequently after integration we get:

(set $s := (s_1, ..., s_m)$)

$$\int_{\prod_{i=1}^m \gamma_i} f(s) g(s) \prod_{i=1}^m ds_i = \frac{\int_{\prod_{i=1}^m \gamma_i} g(s) \prod_{i=1}^m ds_i}{\prod_{i=1}^m (w_i - u_i)} \int_{\prod_{i=1}^m \gamma_i} f(s) \prod_{i=1}^m ds_i + \frac{1}{\prod_{i=1}^m (w_i - u_i)} \int_{\prod_{i=1}^m \gamma_i} g(s) T_f^{(m)}(s) \prod_{i=1}^m ds_i, \tag{2.7}$$

and

$$\int_{\prod_{i=1}^m \gamma_i} f(s) g(s) \prod_{i=1}^m ds_i = \frac{\int_{\prod_{i=1}^m \gamma_i} f(s) \prod_{i=1}^m ds_i}{\prod_{i=1}^m (w_i - u_i)} \int_{\prod_{i=1}^m \gamma_i} g(s) \prod_{i=1}^m ds_i + \frac{1}{\prod_{i=1}^m (w_i - u_i)} \int_{\prod_{i=1}^m \gamma_i} f(s) T_g^{(m)}(s) \prod_{i=1}^m ds_i. \tag{2.8}$$

By (2.7) and (2.8) we obtain

$$\int_{\prod_{i=1}^m \gamma_i} f(s) g(s) \prod_{i=1}^m ds_i - \frac{1}{\prod_{i=1}^m (w_i - u_i)} \left(\int_{\prod_{i=1}^m \gamma_i} f(s) \prod_{i=1}^m ds_i \right) \left(\int_{\prod_{i=1}^m \gamma_i} g(s) \prod_{i=1}^m ds_i \right) = \frac{1}{\prod_{i=1}^m (w_i - u_i)} \int_{\prod_{i=1}^m \gamma_i} f(s) T_g^{(m)}(s) \prod_{i=1}^m ds_i =$$

$$\frac{1}{\prod_{i=1}^m (w_i - u_i)} \int_{\prod_{i=1}^m \gamma_i} g(s) T_f^{(m)}(s) \prod_{i=1}^m ds_i.$$

We conclude that (set $d\overrightarrow{s} := \prod_{i=1}^m ds_i$)

$$\int_{\prod_{i=1}^m \gamma_i} f(s) g(s) d\overrightarrow{s} - \frac{1}{\prod_{i=1}^m (w_i - u_i)} \left(\int_{\prod_{i=1}^m \gamma_i} f(s) d\overrightarrow{s} \right) \left(\int_{\prod_{i=1}^m \gamma_i} g(s) d\overrightarrow{s} \right) = \frac{1}{2\left(\prod_{i=1}^m (w_i - u_i)\right)} \left[\int_{\prod_{i=1}^m \gamma_i} \left(f(s) T_g^{(m)}(s) + g(s) T_f^{(m)}(s) \right) d\overrightarrow{s} \right].$$

Therefore we have

$$\frac{1}{\prod_{i=1}^m (w_i - u_i)} \int_{\prod_{i=1}^m \gamma_i} f(s) g(s) d\overrightarrow{s} - \frac{1}{\prod_{i=1}^m (w_i - u_i)} \left(\int_{\prod_{i=1}^m \gamma_i} f(s) d\overrightarrow{s} \right) \frac{1}{\prod_{i=1}^m (w_i - u_i)} \left(\int_{\prod_{i=1}^m \gamma_i} g(s) d\overrightarrow{s} \right) = \frac{1}{2\left(\prod_{i=1}^m (w_i - u_i)\right)^2}$$

$$\left[\int_{\prod_{i=1}^m \gamma_i} \left\{ f(s) \left(A_g^{(m)}(s) + B_g^{(m)}(s) \right) + g(s) \left(A_f^{(m)}(s) + B_f^{(m)}(s) \right) \right\} d\overrightarrow{s} \right].$$

Hence it holds

$$\Delta(f,g) := \frac{1}{\prod_{i=1}^m (w_i - u_i)} \int_{\prod_{i=1}^m \gamma_i} f(s) g(s) d\overrightarrow{s} - \frac{1}{\prod_{i=1}^m (w_i - u_i)} \left(\int_{\prod_{i=1}^m \gamma_i} f(s) d\overrightarrow{s} \right) \frac{1}{\prod_{i=1}^m (w_i - u_i)} \left(\int_{\prod_{i=1}^m \gamma_i} g(s) d\overrightarrow{s} \right) -$$

$$\frac{1}{2\left(\prod_{i=1}^m (w_i - u_i)\right)^2} \left[\int_{\prod_{i=1}^m \gamma_i} \left\{ f(s) A_g^{(m)}(s) + g(s) A_f^{(m)}(s) \right\} d\overrightarrow{s} \right] = \frac{1}{2\left(\prod_{i=1}^m (w_i - u_i)\right)^2} \left[\int_{\prod_{i=1}^m \gamma_i} \left\{ f(s) B_g^{(m)}(s) + g(s) B_f^{(m)}(s) \right\} d\overrightarrow{s} \right].$$

Clearly we derive that ($\left|d\overrightarrow{s}\right| := \prod_{i=1}^m |ds_i|$)

$$|\Delta(f,g)| \leq \frac{1}{2\left(\prod_{i=1}^m |w_i - u_i|\right)^2} \left[\int_{\prod_{i=1}^m \gamma_i} \left\{ |f(s)| \left| B_g^{(m)}(s) \right| + |g(s)| \left| B_f^{(m)}(s) \right| \right\} \left| d\overrightarrow{s} \right| \right] = \tag{2.9}$$

$$\frac{1}{2\left(\prod_{i=1}^{m}|w_i-u_i|\right)^2}\left[\int_{\prod_{i=1}^{m}\gamma_i}|f(s)|\left|B_g^{(m)}(s)\right|\left|d\overrightarrow{s}\right|+\int_{\prod_{i=1}^{m}\gamma_i}|g(s)|\left|B_f^{(m)}(s)\right|\left|d\overrightarrow{s}\right|\right]\leq$$

$$\frac{1}{2\left(\prod_{i=1}^{m}|w_i-u_i|\right)^2}\left[\|f\|_{\infty,\prod_{i=1}^{m}\gamma_i}\left(\int_{\prod_{i=1}^{m}\gamma_i}\left|B_g^{(m)}(s)\right|\left|d\overrightarrow{s}\right|\right)+\|g\|_{\infty,\prod_{i=1}^{m}\gamma_i}\left(\int_{\prod_{i=1}^{m}\gamma_i}\left|B_f^{(m)}(s)\right|\left|d\overrightarrow{s}\right|\right)\right].$$

We have established the following complex multivariate Grüss type inequality:

Theorem 2.7. *Let f, g and all as in Theorem 2.2. Then*

$$\left|\frac{1}{\prod_{i=1}^{m}(w_i-u_i)}\int_{\prod_{i=1}^{m}\gamma_i}f(s)\,g(s)\,d\overrightarrow{s}-\frac{1}{\prod_{i=1}^{m}(w_i-u_i)}\left(\int_{\prod_{i=1}^{m}\gamma_i}f(s)\,d\overrightarrow{s}\right)\frac{1}{\prod_{i=1}^{m}(w_i-u_i)}\left(\int_{\prod_{i=1}^{m}\gamma_i}g(s)\,d\overrightarrow{s}_i\right)-\right.$$

$$\left.\frac{1}{2\left(\prod_{i=1}^{m}(w_i-u_i)\right)^2}\left[\int_{\prod_{i=1}^{m}\gamma_i}\left(f(s)A_g^{(m)}(s)+g(s)A_f^{(m)}(s)\right)d\overrightarrow{s}\right]\right|\leq$$

$$\frac{1}{2\left(\prod_{i=1}^{m}|w_i-u_i|\right)^2}\left[\|f\|_{\infty,\prod_{i=1}^{m}\gamma_i}\left(\int_{\prod_{i=1}^{m}\gamma_i}\left|B_g^{(m)}(s)\right|\left|d\overrightarrow{s}\right|\right)+\|g\|_{\infty,\prod_{i=1}^{m}\gamma_i}\left(\int_{\prod_{i=1}^{m}\gamma_i}\left|B_f^{(m)}(s)\right|\left|d\overrightarrow{s}\right|\right)\right].$$

The corresponding L_p Grüss inequality follows:

Theorem 2.8. *Let f, g and all as in Theorem 2.2 and $p, q > 1$ such that $\frac{1}{p}+\frac{1}{q}=1$. Then*

$$\left|\frac{1}{\prod_{i=1}^{m}(w_i-u_i)}\int_{\prod_{i=1}^{m}\gamma_i}f(s)\,g(s)\,d\overrightarrow{s}-\frac{1}{\prod_{i=1}^{m}(w_i-u_i)}\left(\int_{\prod_{i=1}^{m}\gamma_i}f(s)\,d\overrightarrow{s}\right)\frac{1}{\prod_{i=1}^{m}(w_i-u_i)}\left(\int_{\prod_{i=1}^{m}\gamma_i}g(s)\,d\overrightarrow{s}\right)-\right.$$

$$\left.\frac{1}{2\left(\prod_{i=1}^{m}(w_i-u_i)\right)^2}\left[\int_{\prod_{i=1}^{m}\gamma_i}\left(f(s)A_g^{(m)}(s)+g(s)A_f^{(m)}(s)\right)d\overrightarrow{s}\right]\right|\leq$$

$$\frac{1}{2\left(\prod_{i=1}^{m}|w_i-u_i|\right)^2}\left[\|f\|_{p,\prod_{i=1}^{m}\gamma_i}\left\|B_g^{(m)}\right\|_{q,\prod_{i=1}^{m}\gamma_i}+\|g\|_{p,\prod_{i=1}^{m}\gamma_i}\left\|B_f^{(m)}\right\|_{q,\prod_{i=1}^{m}\gamma_i}\right].$$

Proof. Use of (2.9) and Hölder inequality. ∎

The corrsponding L_1 Grüss inequality follows:

Theorem 2.9. *Let f, g and all as in Theorem 2.2. Then*

$$\left|\frac{1}{\prod_{i=1}^{m}(w_i-u_i)}\int_{\prod_{i=1}^{m}\gamma_i}f(s)\,g(s)\,d\overrightarrow{s}-\frac{1}{\prod_{i=1}^{m}(w_i-u_i)}\left(\int_{\prod_{i=1}^{m}\gamma_i}f(s)\,d\overrightarrow{s}\right)\frac{1}{\prod_{i=1}^{m}(w_i-u_i)}\left(\int_{\prod_{i=1}^{m}\gamma_i}g(s)\,d\overrightarrow{s}\right)-\right.$$

$$\left.\frac{1}{2\left(\prod_{i=1}^{m}(w_i-u_i)\right)^2}\left[\int_{\prod_{i=1}^{m}\gamma_i}\left(f(s)A_g^{(m)}(s)+g(s)A_f^{(m)}(s)\right)d\overrightarrow{s}\right]\right|\leq$$

$$\frac{1}{2\left(\prod_{i=1}^{m}|w_i-u_i|\right)^2}\left[\|f\|_{1,\prod_{i=1}^{m}\gamma_i}\left\|B_g^{(m)}\right\|_{\infty,\prod_{i=1}^{m}\gamma_i}+\|g\|_{1,\prod_{i=1}^{m}\gamma_i}\left\|B_f^{(m)}\right\|_{\infty,\prod_{i=1}^{m}\gamma_i}\right].$$

Proof. By (2.9) □

Corollary 2.10. *Let f,g and all as in Theorem 2.1. Then*

$$\left|\frac{1}{\prod_{i=1}^{3}(w_i-u_i)}\int_{\prod_{i=1}^{3}\gamma_i}f(s)g(s)d\overrightarrow{s}-\frac{1}{\prod_{i=1}^{3}(w_i-u_i)}\left(\int_{\prod_{i=1}^{3}\gamma_i}f(s)d\overrightarrow{s}\right)\frac{1}{\prod_{i=1}^{3}(w_i-u_i)}\left(\int_{\prod_{i=1}^{3}\gamma_i}g(s)d\overrightarrow{s}\right)-\right.$$

$$\left.\frac{1}{2\left(\prod_{i=1}^{3}(w_i-u_i)\right)^2}\left[\int_{\prod_{i=1}^{3}\gamma_i}\left(f(s)A_g^{(3)}(s)+g(s)A_f^{(3)}(s)\right)d\overrightarrow{s}\right]\right|\leq$$

$$\frac{1}{2\left(\prod_{i=1}^{m}|w_i-u_i|\right)^2}\left[\|f\|_{\infty,\prod_{i=1}^{3}\gamma_i}\left(\int_{\prod_{i=1}^{3}\gamma_i}\left|B_g^{(3)}(s)\right|\left|d\overrightarrow{s}\right|\right)+\|g\|_{\infty,\prod_{i=1}^{3}\gamma_i}\left(\int_{\prod_{i=1}^{3}\gamma_i}\left|B_f^{(3)}(s)\right|\left|d\overrightarrow{s}\right|\right)\right].$$

Proof. By Theorem 2.7 for $m=3$. □

Corollary 2.11. *Let f,g and all as in Theorem 2.1 and $p,q>1:\frac{1}{p}+\frac{1}{q}=1$. Then*

$$\left|\frac{1}{\prod_{i=1}^{3}(w_i-u_i)}\int_{\prod_{i=1}^{3}\gamma_i}f(s)g(s)d\overrightarrow{s}-\frac{1}{\prod_{i=1}^{3}(w_i-u_i)}\left(\int_{\prod_{i=1}^{3}\gamma_i}f(s)d\overrightarrow{s}\right)\frac{1}{\prod_{i=1}^{3}(w_i-u_i)}\left(\int_{\prod_{i=1}^{3}\gamma_i}g(s)d\overrightarrow{s}\right)-\right.$$

$$\left.\frac{1}{2\left(\prod_{i=1}^{3}(w_i-u_i)\right)^2}\left[\int_{\prod_{i=1}^{3}\gamma_i}\left(f(s)A_g^{(3)}(s)+g(s)A_f^{(3)}(s)\right)d\overrightarrow{s}\right]\right|\leq$$

$$\frac{1}{2\left(\prod_{i=1}^{3}|w_i-u_i|\right)^2}\left[\|f\|_{p,\prod_{i=1}^{3}\gamma_i}\left\|B_g^{(3)}\right\|_{q,\prod_{i=1}^{3}\gamma_i}+\|g\|_{p,\prod_{i=1}^{3}\gamma_i}\left\|B_f^{(3)}\right\|_{q,\prod_{i=1}^{3}\gamma_i}\right].$$

Proof. By Theorem 2.8 for $m=3$. □

Corollary 2.12. *Let f,g and all as in Theorem 2.1. Then*

$$\left|\frac{1}{\prod_{i=1}^{3}(w_i-u_i)}\int_{\prod_{i=1}^{3}\gamma_i}f(s)g(s)d\overrightarrow{s}-\frac{1}{\prod_{i=1}^{3}(w_i-u_i)}\left(\int_{\prod_{i=1}^{3}\gamma_i}f(s)d\overrightarrow{s}\right)\frac{1}{\prod_{i=1}^{3}(w_i-u_i)}\left(\int_{\prod_{i=1}^{3}\gamma_i}g(s)d\overrightarrow{s}\right)-\right.$$

$$\frac{1}{2\left(\prod_{i=1}^{3}(w_i-u_i)\right)^2}\left[\int_{\prod_{i=1}^{3}\gamma_i}\left(f(s)A_g^{(3)}(s)+g(s)A_f^{(3)}(s)\right)d\overrightarrow{s}\right]\Bigg| \leq$$

$$\frac{1}{2\left(\prod_{i=1}^{3}|w_i-u_i|\right)^2}\left[\|f\|_{1,\prod_{i=1}^{3}\gamma_i}\left\|B_g^{(3)}\right\|_{\infty,\prod_{i=1}^{3}\gamma_i}+\|g\|_{1,\prod_{i=1}^{3}\gamma_i}\left\|B_f^{(3)}\right\|_{\infty,\prod_{i=1}^{3}\gamma_i}\right].$$

Proof. By Theorem 2.9 for $m=3$. □

References

[1] A. Ostrowski, *Über die Absolutabweichung einer differentiebaren Funktion von ihrem Integral mittelwert*, Comment. Math. Helv. **10** (1938), 226-227.

[2] G. Grüss, *Über das Maximum des absoluten Betrages von* $\frac{1}{b-a}\int_a^b f(x)g(x)\,dx-\frac{1}{(b-a)^2}\int_a^b f(x)\,dx\int_a^b g(x)\,dx$, Math. Z., **39** (1935), 215-226.

[3] D. S. Mitrinović, J.E. Pečarić, A.M. Fink, *Inequalities for Functions and Their Integrals and Derivatives*,Kluwer Academic Publishers, Dordrecht, 1994.

[4] S. S. Dragomir, *An extension of Ostrowski's inequality to the Complex integral*, RGMIA Res. Rep. Call. **21** (2018), Art. 112, 17 pp, [http://rgmia.org/papers/v21/v21/v21a112.pdf].

[5] S. S. Dragomir, *On some Grüss type inequalities for the complex integral*, RGMIA Res. Rep. Call. **21** (2018), Art. 121, 12 pp, [http://rgmia.org/papers/v21/v21a121.pdf].