The Sigmoid Neural Network Activation Function and its Connections to Airy's and the Nield-Kuznetsov Functions

M.H. HAMDAN
Department of Mathematics and Statistics
University of New Brunswick
100 Tucker Park Road, Saint John, New Brunswick, E2L 4L5
CANADA

D.C. ROACH
Department of Engineering
University of New Brunswick
100 Tucker Park Road, Saint John, New Brunswick, E2L 4L5
CANADA

Abstract: - Analysis and solution of Airy's inhomogeneous equation, when its forcing function is the sigmoid neural network activation function, are provided in this work. Relationship between the Nield-Kuznetsov, the Scorer, the sigmoid, the polylogarithm and Airy's functions are established. Solutions to initial and boundary value problems, when the sigmoid function is involved, are obtained. Computations were carried out using *Wolfram Alpha*.

Key-Words: - Sigmoid Logistic Function; Airy's Equation; Nield-Kuznetsov Functions.

Received: May 17, 2021. Revised: February 26, 2022. Accepted: March 21, 2022. Published: April 27, 2022.

1 Introduction

A function that continues to receive considerable attention in neural networks and deep learning, as it serves as an activation function, is the sigmoid logistic function, (*cf.* [1-6] and the references therein). Its characteristic S-shaped curve, [3], maps the number line onto a finite-length subinterval, such as (0,1), and its most common form is given by:

$$S(x) = \frac{e^x}{1+e^x} = 1 - S(-x) \tag{1}$$

Recently, this smoothly-increasing function has made it to the porous media literature, where it has been used in modelling variations in permeability across a porous layer, [5].

The *S-shaped* graph of the sigmoid function makes it appealing in the study of transition layer, and Roach and Hamdan, [5], provided a modification of $S(x)$ and used it in the modelling of Poiseuille flow through a Brinkman porous layer of variable permeability, wherein they created a continuously varying permeability between relatively constant permeability regions. An advantage of the Roach and Hamdan approach is that it treats the flow domain as one region with variable permeability, while replicating flow in layered media.

Depending on the choice of permeability variations in the transition layer, Brinkman's equation can sometimes be reduced to Airy's inhomogeneous ordinary differential equation (ODE), [7]. Airy's ODE, [8-10], has received considerable attention in the literature, and general approaches to solutions of Airy's inhomogeenoeus ODE have been introduced, (*cf.* [11-19] and the references therein). In addition to its importance in mathematical physics, solutions to the inhomogeneous Airy's ODE when its forcing function is a general function of the independent variable, give rise to new functions that are important in the advancement of our mathematical library of functions.

Of particular interest to the current work is the solution to Airy's inhomogeneous ODE when its forcing function is the sigmoid logistic function, $S(x)$, defined in (1), above. The objective is to find the general solution to the resulting inhomogeneous ODE, then use the general solution to obtain solutions to initial and boundary value problems. It will be shown that this choice of forcing function leads to the establishment of connections between the sigmoid function, Airy's functions, the Nield-Kuznetsov functions, Scorer functions, and polylogarithmic functions.

2 Properties of the Sigmoid Logistic Function

Some properties of the sigmoid function, $S(x)$, are listed in what follows, (some of these properties can be found in von Seggern, [3], and Weisstein, [6]). In what follows, "prime" notation denotes ordinary differentiation with respect to the argument.

Property 1: Domain of $S(x)$ is the set of real numbers, $-\infty < x < +\infty$ and its range is the interval $(0,1)$. Graph of *S(x)* is shown in **Fig. 1** for $-5 \le x \le 5$. This graph was obtained using *Wolfram Alpha.*

Property 2: $S(x)$ has horizontal asymptotes at $x = \mp\infty$ with

$$\lim_{x\to+\infty} S(x) = 1 \text{ and } \lim_{x\to-\infty} S(x) = 0 \quad (2)$$

Property 3: The first two derivatives of $S(x)$ are given by

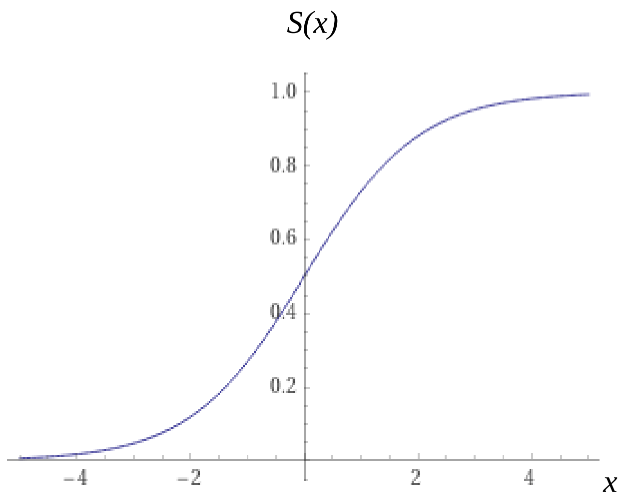

Fig. 1 Graph of *S(x)* for $-5 \le x \le 5$

$$S'(x) = \frac{e^x}{(1+e^x)^2} \quad (3)$$

$$S''(x) = \frac{e^x(1-e^x)}{(1+e^x)^3} \quad (4)$$

Property 4: Values at zero of $S(x)$ and its first two derivatives are $S(0) = \frac{1}{2}$; $S'(0) = \frac{1}{4}$; $S''(0) = 0$.

Property 5: Properties 3 and 4 imply that $S(x)$ is increasing on its domain and has a point of inflection at $x = 0$.

Property 6: $S(x)$ represents solution to Bernoulli's ODE of the form

$$y' - y = g(y) \quad (5)$$

where $g(y) = -y^2$, with initial value $y(0) = \frac{1}{2}$.

Property 7: Indefinite integral of $S(x)$ is given by

$$\int S(x)dx = \log(1 + e^x) + C \quad (6)$$

where C is a constant and *log* stands for the natural logarithm.

Property 8: Definite integral of $S(x)$ over the interval [0,1] is given by

$$\int_0^1 S(x)dx = \log(1+e) - \log 2 \quad (7)$$

Property 9: Higher derivatives of $S(x)$ can be written as polynomials in $S(x)$. This might be of convenience in obtaining higher deivatives of $S(x)$.

$$S'(x) = S(x) - S^2(x) \quad (8)$$

$$S''(x) = S(x) - 3S^2(x) + 2S^3(x) \quad (9)$$

Equation (28) suggests that $S(x)$ is related to the solution of an ODE of the form:

$$y'' = y - 3y^2 + 2y^3 \quad (10)$$

Property 10: Higher derivatives of $S(x)$ can be written in terms of the first derivative, $S'(x)$. This might also be of convenience in obtaining higher derivatives of $S(x)$. With the first derivative given by (8), the second derivative can be written as:

$$S''(x) = S'(x)[1 - 2S(x)] \quad (11)$$

Equation (11) suggests that $S(x)$ is related to the solution of an ODE of the form:

$$y'' - (1-2y)y' = 0 \quad (12)$$

Property 11: $S(x)$ possesses the following Maclaurin series representation, [6]:

$$S(x) = \sum_{n=0}^{\infty} \frac{(-1)^n E_n(0)}{2\, n!} x^n = \frac{1}{2} + \frac{x}{4} - \frac{x^3}{48} + \frac{x^5}{480} - \cdots \quad (13)$$

where $E_n(x)$ is an Euler polynomial.

3 Solution to Airy's ODE with Sigmoid Forcing Function

Airy's inhomogeneous ODE with $S(x)$ as its forcing function takes the form:

$$y'' - xy = S(x) \quad (14)$$

The general solution to (14) can be expressed in the form

$$y = c_1 A_i(x) + c_2 B_i(x) + y_p \quad (15)$$

where $A_i(x)$ and $B_i(x)$ are the linearly independent Airy's functions of the first and second kind, whose non-zero Wronskian is given by, [9,10]:

$$A_i(x)B'_i(x) - B_i(x)A'_i(x) = \frac{1}{\pi} \quad (16)$$

and y_p is the particular solution, expressible as, [11]:

$$y_p = \pi\left\{B_i(x)\int_0^x S(t)A_i(t)\,dt - A_i(x)\int_0^x S(t)B_i(t)\,dt\right\} \quad (17)$$

The integrals on the right of (17) are evaluated using integration by parts, with the help of (6), to yield:

$$\int_0^x S(t)A_i(t)\,dt = A_i(x)\log(1+e^x) - A'_i(x)\int_0^x \log(1+e^t)dt \quad (18)$$

$$\int_0^x S(t)B_i(t)\,dt = B_i(x)\log(1+e^x) - B'_i(x)\int_0^x \log(1+e^t)dt \quad (19)$$

Using (18) and (19) in (17), yields

$$y_p = \pi[A_i(x)B'_i(x) - B_i(x)A'_i(x)]\int_0^x \log(1+e^t)dt \quad (20)$$

Using (16), equation (20) can be written as:

$$y_p = \int_0^x \log(1+e^t)dt = Li_2(-1) - Li_2(-e^x) \quad (21)$$

where

$$Li_2(x) = \sum_{n=1}^{\infty} \frac{x^n}{n^2} \quad (21)$$

is the dilogarithm function, [20,21].

Using the value $Li_2(-1) = -\frac{\pi^2}{12}$ in (21), and subsequent use in (15), the general solution to (14) can be written as:

$$y = c_1 A_i(x) + c_2 B_i(x) - Li_2(-e^x) - \frac{\pi^2}{12} \quad (22)$$

The above analysis furnishes the proof to the following Theorem.

Theorem 1.

The particular solution to Airy's inhomogeneous ODE with the sigmoid logistic function as its forcing function is given by

$$y_p = -Li_2(-e^x) - \frac{\pi}{12} \quad ...(i)$$

and its general solution is given by

$$y = c_1 A_i(x) + c_2 B_i(x) - Li_2(-e^x) - \frac{\pi^2}{12} \quad ...(ii)$$

4 Connections with the Scorer Functions and the Nield-Kuznetsov Functions

Based on the analysis provided in Hamdan and Kamel [11], the particular solution (17) can also be expressed in the form

$$y_p = \pi K_i(x) - \pi S(x) N_i(x) \tag{23}$$

where the integral function, $K_i(x)$, is the Nield-Kuznetsov function of the second kind that is defined by the following equivalent forms, [11]:

$$K_i(x) = A_i(x)\int_0^x \left\{\int_0^t B_i(\tau)d\tau\right\} S'(t)dt - B_i(x)\int_0^x \left\{\int_0^t A_i(\tau)d\tau\right\} S'(t) \tag{24}$$

$$K_i(x) = f(x)N_i(x) - \left\{A_i(x)\int_0^x S(t)B_i(t)\,dt - B_i(x)\int_0^x S(t)A_i(t)\,dt\right\} \tag{25}$$

and the integral function $N_i(x)$ is the Nield-Kuznetsov function of the first kind that is defined by, [7,11]:

$$N_i(x) = A_i(x)\int_0^x B_i(t)dt - B_i(x)\int_0^x A_i(t)dt \tag{26}$$

Hamdan and Kamel, [11], showed that

$$N_i(x) = \tfrac{2}{3}G_i(x) - \tfrac{1}{3}H_i(x) \tag{27}$$

where $H_i(x)$ and $G_i(x)$ are the Scorer functions, [22], that represent particular solutions to the Airy's inhomogeneous equation when the forcing functions are $\mp\frac{1}{\pi}$.

Now, equations (22) to (27) render the following relationships between $K_i(x)$, $N_i(x)$, $S(x)$, $G_i(x)$, $H_i(x)$, $A_i(x)$, $B_i(x)$ and $Li_2(x)$:

$$K_i(x) = S(x)N_i(x) - \tfrac{1}{\pi}Li_2(-e^x) - \tfrac{\pi^2}{12} \tag{28}$$

$$K_i(x) = S(x)\left\{A_i(x)\int_0^x B_i(t)dt - B_i(x)\int_0^x A_i(t)dt\right\} - Li_2(-e^x) - \tfrac{\pi^2}{12} \tag{29}$$

Using (13), equation (41) can be written in terms of the Scorer functions, as:

$$K_i(x) = S(x)\left\{\tfrac{2}{3}G_i(x) - \tfrac{1}{3}H_i(x)\right\} - \tfrac{1}{\pi}Li_2(-e^x) - \tfrac{\pi^2}{12} \tag{30}$$

The above analysis furnishes the proof to the following Theorem.

Theorem 2.

The Nield-Kuznetsov functions of the first and second kinds are related to dilogarithm, sigmoid, Airy's, and Scorer's functions through the equations

$$K_i(x) = S(x)N_i(x) - \frac{1}{\pi}Li_2(-e^x) - \frac{\pi^2}{12} \quad ...(iii)$$

$$K_i(x) = S(x)\left\{A_i(x)\int_0^x B_i(t)dt - B_i(x)\int_0^x A_i(t)dt\right\} - Li_2(-e^x) - \frac{\pi^2}{12} \quad ...(iv)$$

$$K_i(x) = S(x)\left\{\frac{2}{3}G_i(x) - \frac{1}{3}H_i(x)\right\} - \frac{1}{\pi}Li_2(-e^x) - \frac{\pi^2}{12} \quad ...(v)$$

5 Solution to Initial Value Problem

Consider the initial value problem composed of solving equation (14) subject to the initial conditions

$$y(0) = \alpha \tag{31}$$

$$y'(0) = \beta \tag{32}$$

where α and β are known constants.

From (15) and (21), the general solution is written as:

$$y = c_1 A_i(x) + c_2 B_i(x) + \int_0^x \log(1+e^t)dt \qquad (33)$$

Evaluating (33) at $x = 0$, and using (31) yields

$$\alpha = c_1 A_i(0) + c_2 B_i(0) \qquad (34)$$

Differentiating (33) once results in:

$$y' = c_1 A'_i(x) + c_2 B'_i(x) + \log(1+e^x) \qquad (35)$$

Evaluating (35) at $x = 0$, and using (32) yields

$$\beta = c_1 A'_i(0) + c_2 B'_i(0) + \log 2 \qquad (36)$$

Equations (34) and (36) provide the following solution for c_1 and c_2:

$$c_1 = \frac{\alpha[1-\pi A'_i(0)B_i(0)]}{A_i(0)} - [\pi\{\beta - \log 2\}B_i(0)] \qquad (37)$$

$$c_2 = \pi[\{\beta - \log 2\}A_i(0) - \alpha A'_i(0)] \qquad (38)$$

where

$$A_i(0) = \frac{1}{3^{\frac{2}{3}}\Gamma(\frac{2}{3})} \qquad (39)$$

$$B_i(0) = \frac{\sqrt{3}}{3^{\frac{2}{3}}\Gamma(\frac{2}{3})} = \sqrt{3}\,A_i(0) \qquad (40)$$

$$A'_i(0) = \frac{-1}{3^{\frac{1}{3}}\Gamma(\frac{1}{3})} \qquad (41)$$

$$B'_i(0) = \frac{\sqrt{3}}{3^{\frac{1}{3}}\Gamma(\frac{1}{3})} = -\sqrt{3}A'_i(0) \qquad (42)$$

wherein $\Gamma(.)$ is the Gamma function.

The following solution to the initial value problem is thus obtained:

$$y = \left\{\frac{\alpha[1-\pi A'_i(0)B_i(0)]}{A_i(0)} - [\pi\{\beta - \log 2\}B_i(0)]\right\} A_i(x) + \{\pi[\{\beta - \log 2\}A_i(0) - \alpha A'_i(0)]\}B_i(x) - Li_2(-e^x) - \frac{\pi^2}{12} \qquad (43)$$

For the sake of illustration, consider the case of $\alpha = 0$ and $\beta = 1$ in the initial conditions (31) and (32). Equations (37) and (38) render the following values for the arbitrary constants:

$$c_1 = -0.592793305 \qquad (44)$$

$$c_2 = 0.3422493741 \qquad (45)$$

Solution (43) then gives:

$$y = -0.592793305A_i(x) + 0.3422493741\, B_i(x) - Li_2(-e^x) - \frac{\pi^2}{12} \qquad (46)$$

Graph of this solution is shown in **Fig. 2**, below.

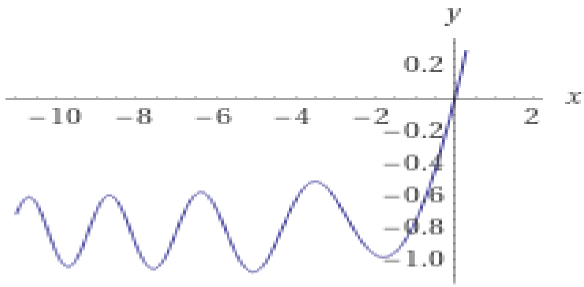

Fig. 2. Solution to the Initial Value Problem $\alpha = 0$ and $\beta = 1$

Computation and graphing of (46) was carried out on *Wolfram Alpha*.

6 Solution to Boundary Value Problem

Consider the two-point boundary value problem composed of solving equation (14) subject to the following conditions on interval $[a, b]$:

$$y(a) = \alpha \qquad (47)$$
$$y(b) = \beta \qquad (48)$$

where α and β are known constants.

Using conditions (47) and (48) in the general solution (2) yields:

$$c_1 A_i(a) + c_2 B_i(a) = \alpha + \frac{\pi^2}{12} + Li_2(-e^a) \qquad (49)$$

$$c_1 A_i(b) + c_2 B_i(b) = \beta + \frac{\pi^2}{12} + Li_2(-e^b) \qquad (50)$$

Solutions to (49) and (50) are given by

$$c_1 = \frac{\alpha + \frac{\pi^2}{12} + Li_2(-e^a)}{A_i(a)} - \frac{\left[\beta + \frac{\pi^2}{12} + Li_2(-e^b)\right]B_i(a) - \left\{\alpha + \frac{\pi^2}{12} + Li_2(-e^a)\right\}\frac{A_i(b)B_i(a)}{A_i(a)}}{[A_i(a)B_i(b) - A_i(b)B_i(a)]} \qquad (51)$$

$$c_2 = \frac{\left[\beta+\frac{\pi^2}{12}+Li_2(-e^b)\right]A_i(a)-\left\{\alpha+\frac{\pi^2}{12}+Li_2(-e^a)\right\}A_i(b)}{A_i(a)B_i(b)-A_i(b)B_i(a)} \quad (52)$$

Using (51) and (52) in (22) gives the following solution to the posed boundary value problem:

$$y = \left\{\frac{\alpha+\frac{\pi^2}{12}+Li_2(-e^a)}{A_i(a)} - \frac{\left[\beta+\frac{\pi^2}{12}+Li_2(-e^b)\right]B_i(a)-\left\{\alpha+\frac{\pi^2}{12}+Li_2(-e^a)\right\}\frac{A_i(b)B_i(a)}{A_i(a)}}{[A_i(a)B_i(b)-A_i(b)B_i(a)]}\right\} * A_i(x) + \left\{\frac{\left[\beta+\frac{\pi^2}{12}+Li_2(-e^b)\right]A_i(a)-\left\{\alpha+\frac{\pi^2}{12}+Li_2(-e^a)\right\}A_i(b)}{A_i(a)B_i(b)-A_i(b)B_i(a)}\right\}B_i(x) - Li_2(-e^x) - \frac{\pi^2}{12} \quad (53)$$

For the sake of illustration, consider the values $a = \alpha = 1, b = \beta = 2$ in boundary conditions (47) and (48). Expressions (51) and (52) then render the following values for the arbitrary constants:

$$c_1 = 2.19841 \quad (54)$$

$$c_2 = -0.232932 \quad (55)$$

and solution (53) then takes the form:

$$y = 2.19841 * Ai(x) - 0.232932 * Bi(x) - Li_2(-e^x) - \frac{\pi^2}{12} \quad (56)$$

Graph of (56) is shown in **Fig. 3** below.

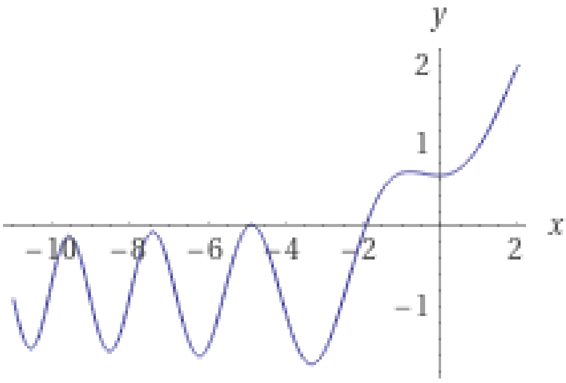

Fig. 3. Solution to the Initial Value Problem $a = \alpha = 1, b = \beta = 2$

Computation and graphing of (56) was carried out using *Wolfram Alpha*.

7 Conclusion

The problem of solving the inhomogeneous Airy's equation (14), when its forcing function is the sigmoid logistic function, was provided in this work. Relationships between the Nield-Kuznetsov functions of the first and second kinds, Airy's functions, the Scorer functions, the sigmoid logistic function and the dilogarithm function have been established and given in ***Theorems 1 and 2.***

General solution to the posed problem was cast in terms of Euler's dilogarithm function, as given in ***Theorem 1***. General formulations of an initial value problem and a two-point boundary value problem were given and solutions were obtained for particular values of the initial and boundary conditions. Solutions and graphs have been carried out using *Wolfram Alpha*.

References:

[1] Skapura, D.M., *Building Neural Networks*. ACM Press, New York, NY, 1996.

[2] Haykin, S., *Neural Networks: A Comprehensive Foundation*, 2/e. Prentice-Hall, Inc., 1999.

[3] von Seggern, D., *CRC Standard Curves and Surfaces with Mathematics, 2nd ed.* Boca Raton, FL: CRC Press, 2007.

[4] Wood, T., *Sigmoid Function*. Retrieved March 8, 2022. https://deepai.org/machine-learning-glossary-and-terms/sigmoid-function.

[5] Roach, D.C. and Hamdan, M.H., On the Sigmoid Function as a Variable Permeability Model for Brinkman Equation, *Trans. on Applied and Theoretical Mechanics*, *WSEAS*, Vol. 17, 2022, pp. 29-38.

[6] Weisstein, Eric W., *Sigmoid Function*. From *MathWorld*--A Wolfram Web Resource, 2022, https://mathworld.wolfram.com/SigmoidFunction.html

[7] Nield, D.A. and Kuznetsov, A.V., The effect of a transition layer between a fluid and a porous medium: shear flow in a channel, *Transp Porous Med*, Vol. 78, 2009, pp. 477-487.

[8] Airy, G.B., On the Intensity of Light in the Neighbourhood of a Caustic, *Trans. Cambridge Phil. Soc.*, Vol. 6, 1838, pp. 379-401.

[9] Abramowitz, M. and Stegun, I.A., *Handbook of Mathematical Functions*, Dover, New York, 1984.

[10] Vallée, O. and Soares, M., *Airy functions and applications to Physics*. World Scientific, London, 2004.

[11] Hamdan, M.H. and Kamel, M.T., On the Ni(x) Integral Function and its Application to the Airy's Non-homogeneous Equation, *Applied Math. Comput.* Vol. 21(17), 2011, pp. 7349-7360.

[12] Hamdan, M.H., Alzahrani, S.M., Abu Zaytoon, M.S. and Jayyousi Dajani, S., Inhomogeneous Airy's and Generalized Airy's Equations with Initial and Boundary Conditions, *Int. J. Circuits, Systems and Signal Processing*, Vol. 15, 2021, pp. 1486-1496.

[13] Hamdan, M.H., Jayyousi Dajani, S.and Abu Zaytoon, M,S., Nield-Kuznetsov Functions: Current Advances and New Results, *Int. J. Circuits, Systems and Signal Processing*, Vol. 15, 2021, pp. 1506-1520.

[14] Hamdan, M.H., Jayyousi Dajani, S., and Abu Zaytoon, M.S., Higher Derivatives and Polynomials of the Standard Nield-Kuznetsov Function of the First Kind, *Int. J. Circuits, Systems and Signal Processing*, Vol. 15, 2021, pp. 1737-1743.

[15] Miller, J. C. P. and Mursi, Z., Notes on the solution of the equation $y'' - xy = f(x)$, *Quarterly J. Mech. Appl. Math.*, Vol. 3, 1950, pp. 113-118.

[16] Alzahrani, S.M., Gadoura, I. and Hamdan, M.H., Ascending Series Solution to Airy's Inhomogeneous Boundary Value Problem, *Int. J. Open Problems Compt. Math.*, Vol. 9(1), 2016, pp. 1-11.

[17] S.M. Alzahrani, I. Gadoura, and M.H. Hamdan, Initial and boundary value problems involving the inhomogeneous Weber equation and the Nield-Kuznetsov parametric function, *The International Journal of Engineering and Science*, Vol.6 No.3, 2017, pp. 9-24.

[18] Abu Zaytoon, M. S., Alderson, T. L. and Hamdan, M. H., Flow through a Variable Permeability Brinkman Porous Core, *J. Appl. Mathematics and physics*, Vol. 4, 2016, pp. 766– 778.

[19] Jayyousi Dajani, S. and Hamdan, M.H., Airy's Inhomogeneous Equation with Special Forcing Function, *İSTANBUL International Modern Scientific Research Congress –II*, Istanbul, TURKEY, Proceedings, ISBN: 978-625-7898-59-1, IKSAD Publishing House, Dec. 23-25, 2021, pp. 1367-1375.

[20] L.C. Maximon, The Dilogarithm Function for Complex Argument, *Proc. R. Soc. Lond. A*, Vol. 459, 2003, pp. 2807–2819.

[21] Lewin, L., ed., Structural Properties of Polylogarithms, *Mathematical Surveys and Monographs*, Vol. 37, 1991, Providence, RI: Amer. Math. Soc. ISBN 978-0-8218-1634-9.

[22] Scorer, R.S., Numerical Evaluation of Integrals of the Form $I = \int_{x1}^{x2} f(x)e^{i\varphi(x)}dx$ and the Tabulation of the Function $\frac{1}{\pi}\int_0^\infty \sin\left(uz + \frac{1}{3}u^3\right) du$, *Quarterly J. Mech. Appl. Math.*, Vol. 3, 1950, pp. 107-112.

Contribution of individual authors

Both authors contributed to literature review, problem formulation and solution, and manuscript preparation.

Sources of funding

This work has not received any financial support from any source.