Physica D 456 (2023) 133909

Contents lists available at ScienceDirect

Physica D

journal homepage: www.elsevier.com/locate/physd

# Collisionless shock region of the KdV equation and an entry in Gradshteyn and Ryzhik

Tewodros Amdeberhan [a], Victor H. Moll [a,*], John Lopez Santander [a], Ken McLaughlin [a], Christoph Koutschan [b]

[a] Department of Mathematics, Tulane University, New Orleans, LA 70118, United States of America
[b] RICAM, Austrian Academy of Sciences, Austria

ARTICLE INFO

*Article history:*
Received 31 March 2023
Received in revised form 5 August 2023
Accepted 4 September 2023
Available online 9 September 2023

To the memory of Hermann Flaschka

*Keywords:*
Asymptotics
Korteweg–de Vries
Elliptic integrals
Automatic proofs

ABSTRACT

The long-time behavior of solutions to the initial value problem for the Korteweg–de Vries equation on the whole line, with general initial conditions has been described uniformly using five different asymptotic forms. Four of these asymptotic forms were expected: the *quiescent behavior* (for $|x|$ very large), a *soliton region* (in which the solution behaves as a collection of isolated solitary waves), a *self-similar region* (in which the solution is described via a Painlevé transcendent), and a *similarity region* (where the solution behaves as a simple trigonometric function of the quantities $t$ and $x/t$). A fifth asymptotic form, lying between the self-similar (Painlevé) and the similarity one, has been described in terms of classical elliptic functions. An integral of elliptic type, giving an explicit representation of the phase, has appeared in this context. The same integral has appeared in the table of integrals by Gradshteyn and Ryzhik. Our goal here is to confirm the validity of this entry.



## 1. Introduction

The table of integrals created by Gradshteyn and Ryzhik [1] contains a large variety of entries where the answers are expressed in terms of the complete elliptic integral

$$\mathbf{K}(k) = \int_0^1 \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}.$$

The goal of this paper is to present a proof of entry 4.242.4 in [1]:

$$\begin{aligned} I(a,b) &:= \int_0^b \frac{\ln x\, dx}{\sqrt{(a^2-x^2)(b^2-x^2)}} \\ &= \frac{1}{2a}\left[\mathbf{K}\left(\frac{b}{a}\right)\ln(ab) - \frac{\pi}{2}\mathbf{K}\left(\frac{\sqrt{a^2-b^2}}{a}\right)\right] \end{aligned} \tag{1.1}$$

with $0 < b < a$. Note that the right-hand side equals $\frac{1}{2a}\left[\mathbf{K}(k)\ln(ab) - \frac{\pi}{2}\mathbf{K}(k')\right]$, with modulus $k = b/a$ and complementary modulus $k' = \sqrt{1-k^2}$.

Some integrals of this general type have been considered in [2] as part of a series of papers dedicated to establishing all entries in [1] starting with [3] and currently at [4]. Entry 4.242.1

$$\int_0^\infty \frac{\ln x\, dx}{\sqrt{(a^2+x^2)(b^2+x^2)}} = \frac{1}{2a}\mathbf{K}\left(\frac{\sqrt{a^2-b^2}}{a}\right)\ln(ab)$$

has been proved in [2]. The crucial point in the proof comes down to the identity

$$\sum_{\ell=0}^{j-1} \frac{a_\ell}{j-\ell} = 4a_j \sum_{\ell=0}^{j-1} \frac{1}{2\ell+1} \quad \text{where} \quad a_\ell = \frac{\left(\frac{1}{2}\right)_\ell^2}{\ell!^2}. \tag{1.2}$$

Two proofs of (1.2) were presented: one involves the manipulation of a balanced ${}_4F_3$ hypergeometric series and the other is an automatic proof based on the techniques developed in [5]. It is our aim to proceed in a similar manner to verify (1.1).

Section 3 presents an expression for $I(a,b)$ in terms of the power series

$$F(x) = \sum_{n=0}^\infty (-1)^n \binom{2n}{n}^2 H_n x^n, \tag{1.3}$$

where $H_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$ is the harmonic number. The series in (1.3) converges for $|x| < \frac{1}{16}$, it is not a hypergeometric function and it makes its appearance in Lemma 3.1. A remarkable fourth order differential equation for $F$, presented in Section 6, is then used to supply an automated proof of (1.1).

* Corresponding author.
*E-mail addresses:* tamdeber@tulane.edu (T. Amdeberhan), vhm@tulane.edu (V.H. Moll), jlopez12@tulane.edu (J.L. Santander), kmclaughlin@tulane.edu (K. McLaughlin), christoph@koutschan.de (C. Koutschan).

https://doi.org/10.1016/j.physd.2023.133909

## 2. Background on the integral

The Korteweg–de Vries equation (KdV)

$$u_t - 6uu_x + u_{xxx} = 0$$

originated in the 19th century as a description of the evolution of long waves in shallow water such as a canal [6,7]. Much later (in the late 1960s) an incredible connection was discovered between this equation and the scattering and inverse scattering theory of the 1-dimensional Schrödinger operator [8]

$$L = -\frac{d^2}{dx^2} + u\ . \tag{2.1}$$

For potentials $u(x)$ that are rapidly decaying as $|x| \to \infty$, the scattering data for $L$ consists of $(i)$ a reflection coefficient $r(z)$ describing the energy that is reflected back when an incoming (quantum) wave with velocity $z$ interacts with the medium represented by $u$, and $(ii)$ a finite number of eigenvalues and associated normalization constants (each eigenvalue corresponds to a bound state for the operator $L$).

The amazing discovery in [8] is that if the potential $u$ in (2.1) evolves according to the KdV equation, so that now $u$ depends on time $u = u(x,t)$ and as does the operator $L = L(t)$, this nonlinear evolution has two remarkable properties:

- The eigenvalues of the operator $L(t)$ are *constant in time* and the associated normalization constants evolve in a simple manner.
- The reflection coefficient $r = r(z,t)$ evolves explicitly in $t$:

  $$r(z,t) = r(z,0)e^{8iz^3t}$$

  where $r(z,0) = r_0(z)$ is the reflection coefficient corresponding to the initial potential $u(x,0) = u_0(x)$.

These remarkable properties followed hard upon the heels of the earlier discovery of Zabusky and Kruskal [9]. Previously these authors had made the observation that the KdV equation possesses not only the classical solitary wave solution, but also other special solutions that spread out at long times like a collection of separated individual solitary waves (both at large negative as well as large positive times), but interact with each other at intermediate times that seem to defy their interpretation as individual solitary waves. Shortly after Gardner, Green, Kruskal, and Miura made the connection to the Schrödinger operator, Lax [10] introduced the framework that bears his name. The pair of operators discovered in [8] represented the first example of a *Lax pair,* which brought forth the so-called inverse scattering method for the analysis of certain special nonlinear partial differential equations.

From these origins there emerged numerous collections of nonlinear partial differential equations and associated Lax pairs of operators. Primary amongst these was the discovery of Zakharov and Shabat [11] that the nonlinear Schrödinger equation falls into the framework, and the discovery independently by Flaschka [12,13] and Manakov [14,15] that the Toda lattice does, too. In each case, the scattering data for one of the operators evolves simply in the time variable. The inverse scattering theory was then exploited, or developed, to yield a solution procedure for these new integrable nonlinear partial differential equations. As it turns out, the solution procedure is extremely powerful. Using it, researchers discovered remarkable phenomena in the behavior of these nonlinear equations which turns out to be ubiquitous even outside the class of integrable equations.

A basic example of this type of discovery is the complete understanding of the long-time behavior of solutions to these equations *from general initial conditions*. This was first established for the nonlinear Schrödinger equation [16]. It was presumed for some time that the calculations for the KdV equation would be entirely similar, but a curious technical obstacle to the direct application of the method reared its head, and led to a new asymptotic phenomenon discovered in the behavior of general solutions of the KdV equation, the so-called *collisionless shock region* for the KdV equation [17].

For generic initial conditions, the behavior of the solution of the KdV equation is described uniformly using five different asymptotic forms each occurring in a different spatial region. Four of these asymptotic forms were expected: the *quiescent behavior* (for $|x|$ very large), a *soliton region* (in which the solution behaves as a collection of isolated solitary waves), a *self-similar region* (in which the solution is described via a Painlevé transcendent), and a *similarity region* (where the solution behaves as a simple trigonometric function of the quantities $t$ and $x/t$).

The fifth region, it turns out, lies in between the self-similar (Painlevé) and the similarity one. It emerges because of the surprisingly benign fact that for generic potentials (i.e., generic initial data for the KdV equation), the reflection coefficient takes on a specific extreme value at $z = 0$: $r(0) = -1$. The reflection coefficient satisfies $|r(z)| \leq 1$, and for all values of $z$ *other than* 0, this inequality is *strict*, but not so at $z = 0$. In the calculations, $|x|, t \to \infty$ at different rates depending on which asymptotic region one is studying, and the quantity $\log\left(1 - \left|r\left(\frac{x}{12t}\right)\right|^2\right)$ is a key ingredient in the asymptotic description. The fact that this quantity can diverge, because $r(0) = -1$, is the source of the fifth region.

In this new region, the behavior of the solution is described in terms of a Jacobi cnoidal function [18]. In explicit form, the asymptotic form of the solution in this region is

$$u \sim \left(\frac{-2x}{3t}\right)\left(a(\alpha) + b(\alpha)\mathrm{cn}^2\left(2\mathbf{K}(\nu)\theta + \theta_0;\, \nu(Z)\right)\right)\ , \tag{2.2}$$

where $\mathrm{cn}(\cdot;\nu)$ represents the Jacobi elliptic function [18], the quantity $\alpha$ is determined to depend on a slow variable $Z$, $\theta$ is a fast variable, and $K(\alpha)$ is the complete elliptic integral of the first kind. The quantity $\theta_0$ is a phase which was undetermined in the original work of Albowitz and Segur.

This work led to foundational questions regarding how to provide a rigorous proof of the asymptotic formulae in [17] and, more generally, of the long-time analysis of integrable nonlinear partial differential equations. Such problems remain outside of the reach of any classical methods. The Riemann–Hilbert machinery, developed by Deift and Zhou [19–22] for integrable problems, was used by Deift, Venakides, and Zhou [23] to analyze the collisionless shock region.

In the scattering and inverse scattering theory applied to the modified KdV equation in [20], and to the KdV equation in [23], the solution to the KdV equation was characterized through the solution of a vector-valued Riemann–Hilbert problem. In order to carry out the long-time analysis of this Riemann–Hilbert problem, and extract the long-time behavior of the solution to the partial differential equation, the authors invented explicit transformations relating one Riemann–Hilbert problem to another, each transformation in turn simplifying the nature of the subsequent one, until arriving at a final one for which an existence and uniqueness theorem could be established. Unraveling the sequence of transformations, precise analytical descriptions of the behavior of the solution to the partial differential equation can be extracted.

In each different asymptotic region, the sequence of transformations is different, and tailored to extract from the original Riemann–Hilbert problem the dominant contributing elements to the eventual asymptotic form of the solution. In fact, for the first four regions, the sequence of transformations showed that the dominant contribution comes from a finite number (usually one or two) of isolated points called stationary phase points in the spectral plane.

However, the Riemann–Hilbert analysis for the new collisionless shock region presented a leap in complexity. Indeed, the dominant contribution arose from an evolving (finite) collection of intervals in the spectral plane. Specifically, for $(x,t)$ in the collisionless shock region, four real endpoints emerged: $\pm a(x,t)$, $\pm b(x,t)$, with $0 < a(x,t) < b(x,t) < \sqrt{2}$, and $a^2+b^2 = 2$. These define intervals $(-b,-a)$ and $(a,b)$. The authors used the Riemann surface $\mathbb{X}$ associated to the function $f(z)=(z+b)^{1/2}(z+a)^{1/2}(z-a)^{1/2}(z-b)^{1/2}$, as a fundamental ingredient in their analysis. The genus of $\mathbb{X}$ is 1 and the Jacobi elliptic function appearing in (2.2) is constructed using the Abel map and the periods associated to this surface.

The integral (1.1) appears in [23], when the authors established an explicit representation of the phase $\theta_0$ (for arbitrary initial conditions):

$$\theta_0 = \mathbf{K}(\alpha) - \int_1^{\sqrt{b/a}} \left(\left(w^2-1\right)\left(1-(a/b)^2w^2\right)\right)^{-1/2} dw - \frac{1}{2\pi b}\int_{-a}^{a} \frac{\log\left(2\gamma w^2\right)}{\left(\left(w^2-a^2\right)\left(w^2-b^2\right)\right)^{1/2}}dw.$$

The explicit determination of the asymptotic form (2.2), including the phase $\theta_0$ as well as the size of the error term, highlights the reach of integrability.

But the occurrence of the integral (1.1) is more than just as a part of the answer. In order to understand the overlap between the collisionless shock region and its two neighboring regions, one must study the behavior of the Riemann surface $\mathbb{X}$, the Jacobi elliptic function, and all internal parameters, in two singular limits: in one of them the branch point $a$ converges to 0, and in the other one the branch points $a$ and $b$ merge together. Having an explicit representation of this integral in terms of well-known special functions assists greatly in matching the asymptotic form of the KdV solution in the transition regions. There lies the value of (1.1).

## 3. An analytic representation of the integral $I(a,b)$

This section presents an analytic expression for the integral $I(a,b)$ in (1.1) in terms of the function $F$ from (1.3). The change of variables $t = x/b$ yields

$$I(a,b) = \frac{\ln b}{a}\int_0^1 \frac{dt}{\sqrt{(1-t^2)(1-c^2t^2)}} + \frac{1}{a}\int_0^1 \frac{\ln t\, dt}{\sqrt{(1-t^2)(1-c^2t^2)}}$$

with $c = b/a$. Observe that $0 < c < 1$. The first integral is $\mathbf{K}(c)$ and so

$$I(a,b) = \frac{\ln b}{a}\mathbf{K}(c) + \frac{1}{a}\mathbf{J}(c), \tag{3.1}$$

where

$$\mathbf{J}(c) = \int_0^1 \frac{\ln x\, dx}{\sqrt{(1-x^2)(1-c^2x^2)}}. \tag{3.2}$$

**Lemma 3.1.** *The integral* $\mathbf{J}(c)$ *in* (3.2) *is given by*

$$\mathbf{J}(c) = -\ln 2\,\mathbf{K}(c) - \frac{\pi}{4\sqrt{1-c^2}}F\left(\frac{c^2}{16(1-c^2)}\right).$$

**Proof.** The change of variables $s = t^2$ yields

$$\mathbf{J}(c) = \frac{1}{4}\int_0^1 \frac{\ln s\, ds}{\sqrt{s(1-s)(1-c^2s)}}.$$

To evaluate $\mathbf{J}(c)$ consider the function

$$\mathbf{A}(c,r) = \int_0^1 s^r s^{-1/2}(1-c^2s)^{-1/2}(1-s)^{-1/2}\, ds$$

and observe that

$$\mathbf{J}(c) = \frac{1}{4}\frac{d}{dr}\mathbf{A}(c,r)\Big|_{r=0}.$$

The integral representation ([1, 9.111]) of the hypergeometric function

$${}_2F_1\left(\begin{matrix}\alpha & \beta\\ & \gamma\end{matrix}\Big|z\right) = \frac{1}{B(\beta,\gamma-\beta)}\int_0^1 t^{\beta-1}(1-t)^{\gamma-\beta-1}(1-tz)^{-\alpha}\, dt$$

implies

$$\mathbf{A}(c,r) = B\left(r+\frac{1}{2},\frac{1}{2}\right){}_2F_1\left(\begin{matrix}\frac{1}{2} & r+\frac{1}{2}\\ & r+1\end{matrix}\Big|c^2\right).$$

The relation (see [1, 9.131.1])

$${}_2F_1\left(\begin{matrix}\alpha & \beta\\ & \gamma\end{matrix}\Big|z\right) = (1-z)^{-\alpha}{}_2F_1\left(\begin{matrix}\alpha & \gamma-\beta\\ & \gamma\end{matrix}\Big|\frac{z}{z-1}\right)$$

yields

$$\mathbf{A}(c,r) = (1-c^2)^{-1/2}B\left(r+\frac{1}{2},\frac{1}{2}\right){}_2F_1\left(\begin{matrix}\frac{1}{2} & \frac{1}{2}\\ r+1 & \end{matrix}\Big|\frac{c^2}{c^2-1}\right).$$

Write $\mathbf{A}(c,r) = (1-c^2)^{-1/2}A_1(r)C_1(c,r)$ where

$$A_1(r) = B\left(r+\tfrac{1}{2},\tfrac{1}{2}\right) \quad\text{and}\quad C_1(c,r) = {}_2F_1\left(\begin{matrix}\frac{1}{2} & \frac{1}{2}\\ r+1 & \end{matrix}\Big|\frac{c^2}{c^2-1}\right),$$

so that

$$\mathbf{J}(c) = \frac{1}{4}\frac{d}{dr}\mathbf{A}(c,r)\Big|_{r=0} = \frac{1}{4\sqrt{1-c^2}} \times \left[A_1(0)C_1'(c,0) + A_1'(0)C_1(c,0)\right],$$

where $C_1'$ is the derivative with respect to $r$. Each of these four terms is evaluated individually.

`First term`: $A_1(0)$. The beta function is given by

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)},$$

so that

$$A_1(0) = \frac{\Gamma(r+\frac{1}{2})\Gamma(\frac{1}{2})}{\Gamma(r+1)}\Big|_{r=0} = \pi.$$

`Second term`: $C_1(c,0)$. The value is given by

$$C_1(c,0) = {}_2F_1\left(\begin{matrix}\frac{1}{2} & \frac{1}{2}\\ 1 & \end{matrix}\Big|\frac{c^2}{c^2-1}\right).$$

On the other hand, this hypergeometric value corresponds to the complete elliptic integral (see [1, 8.113.1])

$$\mathbf{K}(k) = \frac{\pi}{2}{}_2F_1\left(\begin{matrix}\frac{1}{2} & \frac{1}{2}\\ 1 & \end{matrix}\Big|k^2\right)$$

so that

$$C_1(c,0) = \frac{2}{\pi}\mathbf{K}\left(\sqrt{\frac{c^2}{c^2-1}}\right).$$

Observe that $c^2 = b^2/a^2 < 1$, therefore the argument of $\mathbf{K}$ above is purely imaginary. The transformation rule (of the imaginary modulus) reads

$$\mathbf{K}(it) = \frac{1}{\sqrt{1+t^2}}\mathbf{K}\left(\frac{t}{\sqrt{1+t^2}}\right). \tag{3.3}$$

(see [24, Page 82]), which gives

$$C_1(c,0) = \frac{2}{\pi}\sqrt{1-c^2}\,\mathbf{K}(c).$$

`Third term`: $A_1'(0)$. The function $A_1(r)$ is given by

$$A_1(r) = B(r+\tfrac{1}{2},\tfrac{1}{2}) = \Gamma(\tfrac{1}{2})\frac{\Gamma(r+\frac{1}{2})}{\Gamma(r+1)}.$$

Differentiation at $r=0$ yields

$$A_1'(0) = \frac{\Gamma(\frac{1}{2})}{\Gamma^2(1)}\left[\Gamma'(\tfrac{1}{2})\Gamma(1) - \Gamma'(1)\Gamma(\tfrac{1}{2})\right].$$

The digamma function $\psi$, defined by $\Gamma'(x) = \psi(x)\Gamma(x)$, shows that

$$A_1'(0) = \frac{\Gamma(\frac{1}{2})}{\Gamma^2(1)}\left[\Gamma(\tfrac{1}{2})\psi(\tfrac{1}{2})\Gamma(1) - \Gamma(1)\psi(1)\Gamma(\tfrac{1}{2})\right],$$

and the special values $\Gamma(1)=1$, $\Gamma(\frac{1}{2})=\sqrt{\pi}$, $\psi(1)=-\gamma$, $\psi(\frac{1}{2}) = -\gamma - 2\ln 2$ appearing in [1, 8.338.1, 8.338.2, 8.366.1 and 8.366.2] generate

$$A_1'(0) = -2\pi\ln 2.$$

`Fourth term`: $C_1'(c,0)$. Start with

$$C_1(c,r) = {}_2F_1\left(\begin{matrix}\frac{1}{2} & \frac{1}{2}\\ r+1 & \end{matrix}\middle|\frac{c^2}{c^2-1}\right)$$

and since $0 < c^2 = b^2/a^2 < 1$, it is convenient to introduce the parameter

$$t = \frac{c^2}{1-c^2},$$

so that $t > 0$ and

$$C_1(c,r) = {}_2F_1\left(\begin{matrix}\frac{1}{2} & \frac{1}{2}\\ r+1 & \end{matrix}\middle|-t\right) = \sum_{\ell=0}^{\infty}(-1)^{\ell}\frac{(\frac{1}{2})_{\ell}^2}{(r+1)_{\ell}}\frac{t^{\ell}}{\ell!}.$$

The only term that needs to be differentiated is the Pochhammer symbol $(r+1)_{\ell}$. To accomplish this, proceed in the manner

$$\frac{d}{dr}\frac{1}{(r+1)_{\ell}} = -\frac{1}{(r+1)_{\ell}^2}\frac{d}{dr}(r+1)_{\ell},$$

and then expanding $(r+1)_{\ell} = (r+1)(r+2)\cdots(r+\ell)$ leads to

$$\frac{d}{dr}(r+1)_{\ell} = (r+1)_{\ell}\sum_{j=1}^{\ell}\frac{1}{r+j}.$$

Evaluating at $r=0$ results in

$$\frac{d}{dr}\frac{1}{(r+1)_{\ell}}\Big|_{r=0} = -\frac{1}{(r+1)_{\ell}}\sum_{j=1}^{\ell}\frac{1}{r+j}\Big|_{r=0} = -\frac{1}{\ell!}H_{\ell}$$

where $H_{\ell}$ is the harmonic number $H_{\ell} = 1 + \frac{1}{2} + \cdots + \frac{1}{\ell}$. The `fourth term` becomes

$$C_1'(c,0) = -\sum_{\ell=0}^{\infty}(-1)^{\ell}\binom{2\ell}{\ell}^2\left(\frac{c^2}{16(1-c^2)}\right)^{\ell}H_{\ell},$$

using the identity

$$\binom{\frac{1}{2}}{n} = \frac{(-1)^{n-1}}{2^{2n}(2n-1)}\binom{2n}{n} \tag{3.4}$$

to simplify the result. The proof is complete. □

The integral $I(a,b)$ is now expressed in terms of the function $F$ defined in (1.3).

**Theorem 3.2.** *Preserving the notations from Lemma 3.1, we have*

$$I(a,b) = \frac{1}{a}\ln\left(\frac{b}{2}\right)\mathbf{K}\left(\frac{b}{a}\right) - \frac{\pi}{4\sqrt{a^2-b^2}}F\left(\frac{b^2}{16(a^2-b^2)}\right).$$

## 4. An analytic proof of the main identity

The previous section has given an expression for

$$I(a,b) = \int_0^b \frac{\ln x\,dx}{\sqrt{(a^2-x^2)(b^2-x^2)}}$$

in terms of the function $F$ defined in (1.3). This section delivers a direct analytic proof of the main identity displayed in (1.1).

**Theorem 4.1.** *The following identity holds:*

$$I(a,b) = \frac{1}{2a}\left[\mathbf{K}\left(\frac{b}{a}\right)\ln(ab) - \frac{\pi}{2}\mathbf{K}\left(\frac{\sqrt{a^2-b^2}}{a}\right)\right]. \tag{4.1}$$

**Proof.** The evaluation (4.1) has been reduced, in (3.1), to

$$I(a,b) = \frac{\ln b}{a}\mathbf{K}(c) + \frac{1}{a}\mathbf{J}(c),$$

with $c = b/a$ and where, after an elementary change of variables,

$$\mathbf{J}(c) = \frac{1}{4}\int_0^1 \frac{\ln s\,ds}{\sqrt{s(1-s)(1-c^2s)}}. \tag{4.2}$$

So, the proof in this section amounts to a direct computation of $\mathbf{J}(c)$.

Start by transforming the interval of integration to a half-line of the new variable $x$, defined via $s = 3/(3x+c^2+1)$. Then (4.2) becomes

$$\begin{aligned}\mathbf{J}(c) &= -\frac{1}{2}\int_{\frac{1}{3}(2-c^2)}^{\infty}\frac{\log(x+\frac{1}{3}(c^2+1))\,dx}{\sqrt{4(x+\frac{c^2+1}{3})(x+\frac{c^2-2}{3})(x+\frac{1-2c^2}{3})}} \\ &= -\frac{1}{2}\int_{\frac{1}{3}(2-c^2)}^{\infty}\frac{\log(x+\frac{1}{2}(c^2+1))\,dx}{\sqrt{4x^3-g_2x-g_3}},\end{aligned} \tag{4.3}$$

with

$$g_2 = \tfrac{4}{3}(c^4-c^2+1) \quad\text{and}\quad g_3 = \tfrac{4}{27}(1+c^2)(1-2c^2)(2-c^2).$$

The discriminant of the cubic is $g_2^3 - 27g_3^2 = 16c^4(c+1)^2(c-1)^2 \neq 0$ for $0 < c < 1$. The roots of the cubic are real and are given by

$$e_1 = \tfrac{1}{3}(2-c^2) > e_2 = \tfrac{1}{3}(2c^2-1) > e_3 = -\tfrac{1}{3}(c^2+1).$$

Consider the curve $\mathcal{E}$ defined by the equation

$$y^2 = 4x^3 - g_2x - g_3.$$

It is well-known that (the projectivation of) $\mathcal{E}$ is a torus $\mathbb{C}/(\mathbb{Z}2\omega_1 \oplus \mathbb{Z}2\omega_2)$, with an associated Weierstrass $\wp$-function satisfying

$$\begin{aligned}(\wp'(z))^2 &= 4\wp(z)^3 - g_2\wp(z) - g_3 \\ &= 4(\wp(z)-e_1)(\wp(z)-e_2)(\wp(z)-e_3).\end{aligned}$$

The parameters $\omega_1$, $\omega_2$ being the half-periods for $\wp(z)$, satisfy

$$\wp(\omega_1) = e_1, \quad \wp(\omega_2) = e_3 \quad\text{and}\quad \wp(\omega_3) = e_2,$$

with $\omega_3 = \frac{1}{2}(\omega_1+\omega_2)$. The periods can be written, using $P(x) = 4x^3 - g_2x - g_3$, as

$$\frac{\omega_1}{2} = \int_{\infty}^{e_1}\frac{dt}{\sqrt{P(t)}} \quad\text{and}\quad \frac{\omega_2}{2} = \int_{e_1}^{e_2}\frac{dt}{\sqrt{P(t)}}$$

so that $\omega_1$ is real and $\omega_2$ is purely imaginary. See [18] for details.

The integral in (4.3) now becomes

$$\mathbf{J}(c) = -\frac{1}{2}\int_{e_1}^{\infty} \frac{\log(x-e_3)\,dx}{\sqrt{4(x-e_1)(x-e_2)(x-e_3)}}.$$

Substitute $x = \wp(z)$ with $\wp'(z) = -\sqrt{4(\wp(z)-e_1)(\wp(z)-e_2)(\wp(z)-e_3)}$ (the negative sign is due to the fact that since $\wp$ is real and decreasing in the interval of integration). Since $\wp$ is an even function, it follows that

$$\begin{aligned}\mathbf{J}(c) &= -\frac{1}{2}\int_0^{\omega_1} \log(\wp(z)-\wp(\omega_2))\,dz \\ &= -\frac{1}{4}\int_{-\omega_1}^{\omega_1} \log(\wp(z)-\wp(\omega_2))\,dz. \qquad (4.4)\end{aligned}$$

At this point, introduce the Weierstrass zeta-function $\zeta(z;g_2,g_3)$ ([18, page 467]) defined by

$$\frac{d}{dz}\zeta(z;g_2,g_3) = -\wp(z;g_2,g_3)$$

and the normalization $\lim_{z\to 0}[\zeta(z;g_2,g_3)-1/z] = 0$. The Weierstrass sigma-function ([18, page 469]) is then defined by

$$\frac{d}{dz}\log\sigma(z;g_2,g_3) = \zeta(z;g_2,g_3)$$

with the corresponding normalization $\lim_{z\to 0}\frac{\sigma(z)}{z} = 1$. The parameters $g_2$ and $g_3$ are suppressed and we simply write $\zeta(z)$ and $\sigma(z)$. The function $\sigma$ is odd, `quasi-periodic` [18, page 470] and satisfies the relations

$$\sigma(z+2\omega_j) = -e^{2\eta_j(z+\omega_j)}\sigma(z), \qquad (4.5)$$

with $\eta_j = \zeta(\omega_j)$.

Now use the identity [18, page 473]

$$\wp(u)-\wp(v) = -\frac{\sigma(u+v)\sigma(u-v)}{\sigma^2(u)\sigma^2(v)}, \qquad (4.6)$$

to write

$$\wp(z)-\wp(\omega_2) = -\frac{\sigma(z+\omega_2)\sigma(z-\omega_2)}{\sigma^2(z)\sigma^2(\omega_2)} = \frac{\sigma^2(z+\omega_2)e^{-2\eta_2 z}}{\sigma^2(z)\sigma^2(\omega_2)},$$

and convert (4.4) into

$$\begin{aligned}\mathbf{J}(c) &= -\frac{1}{2}\int_{-\omega_1}^{\omega_1} \log(\sigma(z+\omega_2))\,dz \\ &\quad + \frac{1}{2}\int_{-\omega_1}^{\omega_1} \log(\sigma(z))\,dz + \omega_1\log(\sigma(\omega_2)),\end{aligned}$$

since the integral arising from $e^{2\eta_2 z}$ vanishes. Define

$$L(\tau) := \int_{-\omega_1}^{\omega_1} \log\sigma(z+\tau)\,dz$$

to write

$$\mathbf{J}(c) = \omega_1\log\sigma(\omega_2) - \tfrac{1}{2}\left[L(\omega_2)-L(0)\right]. \qquad (4.7)$$

The evaluation of $L(\tau)$ begins with

$$\frac{d^3}{d\tau^3}L(\tau) = -\int \wp'(z+\tau)\,dz = -\wp(\tau+\omega_1)+\wp(\tau-\omega_1) = 0$$

since $2\omega_1$ is a period of $\wp$. It follows that

$$L(\tau) = C_0\tau^2 + C_1\tau + C_2,$$

for some constants $C_0,\ C_1,\ C_2$.

The value $C_0$ comes from $J''(\tau) = \zeta(\tau+\omega_1)-\zeta(\tau-\omega_1)$ being constant. Now $L''(\tau) \equiv L''(0) = 2\zeta(\omega_1)$, since $\zeta$ is an odd function. This gives $C_0 = \zeta(\omega_1) = \eta_1$. The computation of $C_1 = L'(0)$ uses the branch cut and gives

$$\begin{aligned}C_1 &= \lim_{\tau\to 0}\left(\log\sigma(\tau+\omega_1) - \log\sigma(\tau-\omega_1)\right) \\ &= \lim_{\tau\to 0}\left(\log\sigma(\tau+\omega_1) - \log\left[-\sigma(-\tau+\omega_1)\right]\right) \\ &= -\pi i,\end{aligned}$$

since $\sigma$ is odd. Thus $L(\tau) = \eta_1\tau^2 - \pi i\tau + C_2$. Legendre's identity $\eta_1\omega_2 - \eta_2\omega_1 = \frac{1}{2}\pi i$ [18, page 469] and (4.7) yield

$$\mathbf{J}(c) = \omega_1\log\sigma(\omega_2) - \tfrac{1}{2}\eta_2\omega_1\omega_2 + \tfrac{1}{4}\pi i\omega_2. \qquad (4.8)$$

The next step requires an identity for the $\sigma$-function:

**Lemma 4.2.** *The Weierstrass $\sigma$-function satisfies*

$$\sigma^2(\omega_1+\omega_2) = e^{2\eta_2\omega_1}\sigma^2(\omega_1)\sigma^2(\omega_2). \qquad (4.9)$$

**Proof.** Observe that (4.7) reveals

$$1 = e_1 - e_3 = \wp(\omega_1) - \wp(\omega_2) = e^{-2\eta_2\omega_1}\frac{\sigma^2(\omega_1+\omega_2)}{\sigma^2(\omega_1)\sigma^2(\omega_2)},$$

and this verifies the identity in (4.9). □

The next result generates an expression for $\log\sigma(\omega_2)$.

**Lemma 4.3.** *The identity*

$$c^2 = \wp(\omega_3) - \wp(\omega_2) = \frac{e^{2\eta_2\omega_2}}{\sigma^4(\omega_2)},$$

*holds. Therefore* $\log\sigma(\omega_2) = \frac{1}{2}\eta_2\omega_2 - \frac{1}{2}\log c$.

**Proof.** Use (4.6) combined with (4.5) to produce

$$\begin{aligned}c^2 &= e_2 - e_3 = \wp(\omega_3) - \wp(\omega_2) = -\frac{\sigma(\omega_1+2\omega_2)\sigma(\omega_1)}{\sigma^2(\omega_3)\sigma^2(\omega_2)} \\ &= \frac{\sigma^2(\omega_1)e^{2\eta_2\omega_3}}{\sigma^2(\omega_3)\sigma^2(\omega_2)},\end{aligned}$$

and the result follows from (4.9). □

Lemmas 4.2 and 4.3, Eq. (4.8) and the expressions for the half-periods [24, page 114]

$$\omega_1 = \sqrt{e_1-e_3}\,\mathbf{K}(c) \quad \text{and} \quad \omega_2 = i\sqrt{e_1-e_3}\,\mathbf{K}(\sqrt{1-c^2})$$

and the fact that $e_1 - e_3 = 1$, readily show that

$$\mathbf{J}(c) = -\frac{\pi}{4}\mathbf{K}(\sqrt{1-c^2}) - \frac{1}{2}\mathbf{K}(c)\log c. \qquad (4.10)$$

Replacing (4.10) back in (3.1) completes the proof. □

The computation of $I(a,b)$ presented above and Theorem 3.2 yield an expression for $F$ in terms of complete elliptic integrals.

**Corollary 4.4.** *The function $F$ defined in (1.3) is given by*

$$F(x) = \frac{1}{\pi\sqrt{1+16x}}\left[\ln\left(\frac{x}{1+16x}\right)\mathbf{K}\left(\sqrt{\frac{16x}{1+16x}}\right) + \pi\mathbf{K}\left(\frac{1}{\sqrt{1+16x}}\right)\right].$$

## 5. A new expression for the function *F*

The function $F$, defined in (1.3), appeared in the first computation of $I(a,b)$ given in Theorem 3.2. This section presents a different approach to this function and establishes a connection with the integral $\mathbf{J}(c)$ defined in (3.2).

**Lemma 5.1.** *Let $B_n$ be a sequence and define $A_n$ by $A_n = 4^{2n}(2n-1)B_n$. Then*

$$\sum_{n=0}^{\infty}(-1)^n B_n \binom{2n}{n}^2 H_n = \frac{2}{\pi}\int_0^1 \frac{\log(1-y^2)}{\sqrt{1-y^2}}\left(\sum_{n=0}^{\infty} A_n \binom{\frac{1}{2}}{n} y^{2n}\right) dy$$
$$-2\log 2\sum_{n=0}^{\infty}(-1)^n B_n \binom{2n}{n}^2; \tag{5.1}$$

*where $H_n$ are the harmonic numbers.*

**Proof.** Use the value

$$\int_0^1 \frac{y^{2n}\log(1-y^2)}{\sqrt{1-y^2}}\,dy = -\frac{\pi}{2^{2n+1}}\binom{2n}{n}\left[H_n + 2\log 2\right],$$

to compute the integral in (5.1), and (3.4) to simplify the result. □

The expression for $F$ is established next.

**Theorem 5.2.** *The function $F$ defined in (1.3) is given by*

$$F(x) = -\frac{4}{\pi\sqrt{1+16x}}\left[\mathbf{J}\left(\frac{16x}{1+16x}\right) + \log 2\,\mathbf{K}\left(\sqrt{\frac{16x}{1+16x}}\right)\right]. \tag{5.2}$$

**Proof.** Let $B_n = x^n$ in Lemma 5.1 and use the classical series

$$\sum_{n=0}^{\infty}\binom{2n}{n}u^n = (1-4u)^{-1/2} \quad\text{and}\quad \sum_{n=0}^{\infty}\binom{2n}{n}^2 u^n = \frac{2}{\pi}\mathbf{K}(\sqrt{16u})$$

to obtain

$$F(x) = -\frac{2}{\pi}\int_0^1 \frac{\log(1-y^2)\,dy}{\sqrt{1-y^2}\sqrt{1+16xy^2}}$$
$$-\frac{4\log 2}{\pi}\frac{1}{\sqrt{1+16x}}\mathbf{K}\left(\sqrt{\frac{16x}{1+16x}}\right), \tag{5.3}$$

using the transformation (3.3) for the complete elliptic integral. The change of variables $u = 1-y^2$ converts (5.3) into (5.2) and produces the result. □

## 6. An automatic proof the last identity for *F*

In this final section we show how the holonomic summation techniques [25] can be employed to prove the closed-form evaluation of $F(x)$ stated in Corollary 4.4, restated here for the convenience of the reader:

**Theorem 6.1.** *The function $F$ in (1.3) is given by*

$$F(x) = \frac{1}{\pi\sqrt{1+16x}}\left[\ln\left(\frac{x}{1+16x}\right)\mathbf{K}\left(\sqrt{\frac{16x}{1+16x}}\right)\right.$$
$$\left.+\pi\mathbf{K}\left(\frac{1}{\sqrt{1+16x}}\right)\right]. \tag{6.1}$$

**Proof.** Let $f_n(x)$ denote the expression inside the sum, i.e.,

$$f_n(x) = (-1)^n\binom{2n}{n}^2 H_n x^n.$$

Note that $f_n(x)$ is not hypergeometric in $n$, due to the presence of harmonic numbers $H_n$, and therefore the original Almkvist–Zeilberger algorithm [26] is not applicable. Instead it requires the corresponding generalization to arbitrary holonomic functions, as implemented in the `HolonomicFunctions` package [27], which delivers the following telescopic relation:

$$x^2(16x+1)^2 f_n^{(4)}(x) + 5x(32x+1)(16x+1)f_n^{(3)}(x)$$
$$+4(1568x^2+98x+1)f_n''(x) + 108(32x+1)f_n'(x) + 144f_n(x)$$
$$= g_n(x) - g_{n+1}(x), \tag{6.2}$$

where $f_n^{(i)}$ denotes the $i$th-derivative and

$$g_n(x) = \frac{n}{x^2}\Big(\big((n-1)n^2 + 4(2n+1)^3x\big)f_n(x) + n(n+1)^3 f_{n+1}(x)\Big).$$

The correctness of (6.2) can be established by routine calculations: divide both sides by $\binom{2n+1}{n+1}^2(-x)^n$ and observe that it reduces to

$$(n+2)H_{n+2} - (2n+3)H_{n+1} + (n+1)H_n = 0,$$

which is indeed a valid relation for harmonic numbers. This can be established by writing the sums for the harmonic numbers and splitting in the manner

$$H_{n+2} = H_n + \frac{1}{n+1} + \frac{1}{n+2} \text{ and } H_{n+1} = H_n + \frac{1}{n+1}. \tag{6.3}$$

Summing the right-hand side of (6.2) over $n = 0, 1, \ldots$ gives, for $|x| < \frac{1}{16}$,

$$g_0(x) - \lim_{n\to\infty} g_n(x) = 0,$$

while summing the left-hand side of (6.2) yields the desired fourth-order differential equation for $F(x)$:

$$x^2(16x+1)^2F^{(4)}(x) + 5x(32x+1)(16x+1)F^{(3)}(x)$$
$$+4(1568x^2+98x+1)F''(x) + 108(32x+1)F'(x) + 144F(x)$$
$$= 0. \tag{6.4}$$

To derive a differential equation for the right-hand side of (6.1), one transforms the standard differential equation for the complete elliptic integral (see [24, page 68]):

$$x(x^2-1)\mathbf{K}''(x) + (3x^2-1)\mathbf{K}'(x) + x\,\mathbf{K}(x) = 0$$

into

$$x(16x+1)^2y''(x) + (16x+1)^2y'(x) - 4y(x) = 0,$$

which is satisfied by both

$$y(x) = \mathbf{K}\left(\sqrt{\frac{16x}{1+16x}}\right) \quad\text{and}\quad y(x) = \mathbf{K}\left(\frac{1}{\sqrt{1+16x}}\right).$$

Combining it with the differential equation

$$x(16x+1)^2y''(x) + (48x+1)(16x+1)y'(x) + 8(24x+1)y(x) = 0,$$

satisfied by

$$y(x) = \frac{1}{\sqrt{1+16x}}\ln\left(\frac{x}{1+16x}\right),$$

yields exactly the same differential equation as in (6.4). These kinds of closure properties are executed algorithmically and automatically by the `Annihilator` command in [27].

Since both sides of (6.1) satisfy the same fourth-order ODE, it suffices to compare four initial values to establish equality. Using the Taylor series

$$\mathbf{K}(x) = \pi\left(\frac{1}{2} + \frac{1}{8}x^2 + \frac{9}{128}x^4 + \cdots\right)$$

one computes the series expansion

$$-4x + 54x^2 - \frac{2200}{3}x^3 + \frac{30625}{3}x^4 + \cdots$$

for the right-hand side of (6.1). Truncating the sum on the left-hand side produces exactly the same coefficients, thereby completing the proof. □

Replacing (6.1) in Theorem 3.2 gives a proof of the original identity (1.1). The same procedure applies to $I(a, b)$ directly, at least in principle. The result is a system of PDEs in $a$ and $b$, but it turns out that comparing the initial values is more delicate.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

No data was used for the research described in the article.

## References

[1] I.S. Gradshteyn, I.M. Ryzhik, Table of Integrals, Series, and Products, eighth ed., Academic Press, New York, 2015,
[2] S. Böettner, V. Moll, The integrals in Gradshteyn and Ryzhik. Part 16: Complete elliptic integrals, Scientia 20 (2011) 45–59.
[3] V. Moll, The integrals in Gradshteyn and Ryzhik. Part 1: A family of logarithmic integrals, Scientia 14 (2007) 1–6.
[4] D. Chen, T. Dunaisky, Victor H. Moll, A.R. McCurdy, C. Nguyen, V. Sharma, The integrals in Gradshteyn and Ryzhik. Part 32: Powers of trigonometric functions, Scientia 32 (2022) 71–98.
[5] M. Petkovšek, H . Wilf, D. Zeilberger, A=B. A.K. Peters, 1st. ed., 1996.
[6] J. Boussinesq, Théorie de l'intumescence liquid appelée onde solitaire ou de translation, se propageant dans un canal rectangulaire, C. R. Acad. Sci., Paris 72 (1871) 755–759.
[7] 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.
[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] N.J. Zabusky, M.D. Kruskal, Interaction of solitons in a collisionless plasma and the recurrence of initial states, Phys. Rev. Lett. 15 (1965) 240–243.
[10] P.D. Lax, Integrals of non-linear equations and solitary waves, Comm. Pure Appl. Math. 21 (1968) 467–490.
[11] V.E. Zakharov, A.B. Shabat, Interaction between solitons in a stable medium, Sov. Phys.-JETP 37 (1973) 823–828.
[12] H. Flaschka, On the Toda lattice II. inverse-scattering solution, Progr. Theoret. Phys 51 (1974) 703–716.
[13] H. Flaschka, The Toda lattice I. Existence of integrals, Phys. Rev. B 9 (1974) 1924–1925.
[14] S.V. Manakov, Complete integrability and stochasticity for discrete dynamical systems, J. Exp. Theor. Phys. 40 (1974) 269–274.
[15] S.V. Manakov, Integration of discrete dynamical systems, Zh. Eksp. Teor. Fiz. 67 (1974) 543.
[16] V.E. Zakharov, S.V. Manakov, Asymptotic behavior of nonlinear wave systems integrated by the inverse method, Sov. Phys.-JETP 44 (1976) 106–112.
[17] M.J. Ablowitz, H. Segur, Asymptotics solutions of the Korteweg-de Vries equation, Stud. Appl. Math. 57 (1977) 13–44.
[18] E.T. Whittaker, G.N. Watson, A Course of Modern Analysis, Cambridge University Press. Prepared for publication by Victor H. Moll, 2020.
[19] P.A. Deift, A.R. Its, X. Zhou, Long-time asymptotics for integrable nonlinear wave equations, in: Important Developments in Soliton Theory, Nonlinear Dynamics, Springer-Verlag, Berlin-Heidelberg, 1993, pp. 181–204.
[20] P.A. Deift, X. Zhou, A steepest descent method for oscillary Riemann-Hilbert problems, Bull. Amer. Math. Soc. 26 (1992) 119–123.
[21] P.A. Deift, X. Zhou, A steepest descent method for oscillary Riemann-Hilbert problems, asymptotics for the mKDV equation, Ann. Math. 137 (1993) 295–368.
[22] P.A. Deift, X. Zhou, Long-time behavior of the non-focusing nonlinear Schrödinger equation-A case study, in: Lecture in Math Sciences, vol. 5, University of Tokyo, 1994.
[23] P.A. Deift, S. Venakides, X. Zhou, The collisionless shock region for the long-time behavior of solution of the KdV equation, Comm. Pure Appl. Math. 47 (1994) 199–206.
[24] H.P. McKean, V. Moll, Elliptic Curves: Function Theory, Geometry, Arithmetic, Cambridge University Press, New York, 1997.
[25] C. Koutschan, Creative telescoping for holonomic functions, in: Computer Algebra in Quantum Field Theory: Integration, Summation and Special Functions, Texts & Monographs in Symbolic Computation, Springer, Wien, 2013, pp. 171–194.
[26] G. Almkvist, D. Zeilberger, The method of differentiating under the integral sign, J. Symbolic Comput. 10 (1990) 571–591.
[27] C. Koutschan, HolonomicFunctions (User's Guide), Technical Report 10-01, RISC Report Series, Johannes Kepler University Linz, 2010, http://www.risc.uni-linz.ac.at/research/combinat/software/HolonomicFunctions/.