A variational characterization of 2-soliton profiles for the KdV equation

John P. Albert

Department of Mathematics, University of Oklahoma
Norman, OK 73019
jalbert@ou.edu

Nghiem V. Nguyen

Department of Mathematics and Statistics, Utah State University
Logan, UT 84322
nghiem.nguyen@usu.edu

January 27, 2021

Abstract

We use profile decomposition to characterize 2-soliton solutions of the KdV equation as global minimizers to a constrained variational problem involving three of the polynomial conservation laws for the KdV equation.

1 Introduction

The variational properties of multi-soliton solutions of the Kortweg-de Vries (KdV) equation have played a central role in the study of this equation since shortly after the discovery of its remarkable properties in the 1960s. An early milestone was the paper [12] of Lax, in which it is pointed out that multi-soliton profiles are critical points of constrained variational problems in which the constraint functionals and the objective functional are conserved under the flow defined by the KdV equation. This suggested that it might be possible to establish stability properties of multi-soliton solutions by using the conserved quantities as Lyapunov-type functionals.

Benjamin [4] (see also Bona [5]) took a first step in this direction by showing that solitary-wave profiles are local minimizers in $H^1(\mathbb{R})$ for the conserved functional E_3, subject to the constraint that E_2 be held constant (see Section 2 for the definition of E_j), and deducing the orbital stability of solitary-wave solutions as a consequence. Later, the work of Cazenave and Lions (see for example, [7], and the expository treatment in [1]) established that solitary-wave profiles are actually global minimizers of this variational problem in a strong sense: every minimizing sequence for the variational problem has a subsequence which, after appropriate translations, converges in $H^1(\mathbb{R})$ to a solitary-wave profile. Orbital stability of solitary waves is an immediate consequence.

Maddocks and Sachs [18] generalized the theory of Benjamin and Bona to obtain a stability result for multi-soliton solutions of KdV. A key step in their proof was to show that the profiles of N-soliton solutions are local minimizers in $H^N(\mathbb{R})$ of the conserved functional E_{N+2}, when the functionals E_2, E_3, $\dots$, E_{N+1} are held constant. Their proof, like that of Benjamin and Bona for single solitons, did not yield information about global minimizers of the variational problem.

In this paper, we consider the special case $N = 2$ of the variational problem considered in [18]: that is, the problem of minimizing E_4 when E_2 and E_3 are held constant. In our main result, Theorem 2.6 below, we show that indeed 2-soliton solutions represent the global minimizers for this variational problem.

An easy consequence of Theorem 2.6 is a stability result for 2-soliton solutions in $H^2(\mathbb{R})$, stated below as Corollary 2.7. Of course, this is only a special case of the stability result of [18], which was asserted for N-solitons for general N, not just for $N = 2$. Moreover, in recent years a number of papers have appeared on the topic of stability of multi-solitons which have improved on the result of [18]. In particular, Killip and Visan [11] have proved a stability result for N-soliton solutions of KdV which is in some sense optimal: it asserts stability in H^{-1}, or more generally in any space H^s with $s \geq -1$. Instead of the variational characterization of multi-solitons used here or in [18], they use a different variational characterization, which is motivated by the inverse scattering theory for KdV, yet which is well-adapted to potentials in low-regularity Sobolev spaces where the classical inverse scattering theory does not apply. We also note the recent work of Le Coz and Wang [15], who by building on and elucidating the work of [18] obtain a stability result for N-soliton solutions of the modified KdV equation. Their methods should be transferrable to other integrable equations as well; and in particular it would be worth using them to revisit the stability theory for KdV multisolitons.

We feel that the result and proof of Theorem 2.6 are interesting in their own right, apart from the consequences for stability theory. The result settles, at least in the case $N = 2$, the question of whether multisolitons are actually global minimizers of a natural variational problem for KdV, expressed in terms of polynomial conservation laws which can be viewed as action variables in a formulation of KdV as an infinite-dimensional Hamiltonian system. It can thus be viewed as a step towards obtaining an analogue for KdV on the real line of the elegant theory produced for the periodic KdV equation by Lax [13] and Novikov [21]. The proof has the advantage of simplicity: it shows that the result is a straightforward consequence of the profile decomposition, a general phenomenon unconnected with the KdV equation or its structure, once one shows that 2-soliton profiles are minimizers for the constrained variational problem when consideration is restricted to the set of multi-soliton profiles. In other words, the fact that 2-soliton profiles are global minimizers in H^2 can be shown to follow from the profile decomposition, together with the fact that they are minimizers within the set of all multi-soliton profiles.

An important caveat, however, is that the argument can only proceed because of the uniqueness result for 2-solitons stated as Theorem 2.2 below; and such uniqueness results can be very difficult to prove in other settings. In fact, one of the main reasons we have restricted ourselves to 2-solitons in the present paper is that an analogue of Theorem 2.2 is not yet available for general N-solitons (see [2] for a discussion of what remains to be shown).

The plan of the remainder of the paper is as follows. In Section 2 we review some basic properties of N-soliton solutions and polynomial conservation laws for the KdV equation, state our main result Theorem 2.6, and sketch its proof. In Section 3, we review the profile decomposition, following [19]. In Section 4, we analyze a finite-dimensional minimization problem which arises from restricting the admissible functions in (2.6) and (2.7) to N-soliton profiles. Section 5 contains the proof of our main result, Theorem 2.6, and concludes with a proof of Corollary 2.7.

Notation.

If E is a measurable subset of $\mathbb{R}$ and $1 \leq p < \infty$, we define $L^p(E)$ to be the space of Lebesgue measurable real-valued functions u on E such that $\|u\|_{L^p(E)} = \left(\int_E |u|^p \, dx\right)^{1/p}$ is finite. In the case when $E = \mathbb{R}$, we sometimes denote $L^p(\mathbb{R})$ by simply L^p, and denote the norm of u in $L^2(\mathbb{R})$ by $\|u\|_{L^2}$.

When E is an open set in $\mathbb{R}$, for $l \in \mathbb{N}$, we define the L^2-based Sobolev space $H^l = H^l(E)$ to be the closure of the space $C^\infty(E)$ of all infinitely smooth real-valued functions on E with respect to the norm

$$\|u\|_{H^l(E)} = \left(\sum_{i=0}^{l} \int_E \left(\frac{d^i u}{dx^i} \right)^2 \, dx \right)^{1/2}.$$

Note that $H^0(E) = L^2(E)$. In the case when $E = \mathbb{R}$, we sometimes denote $H^l(\mathbb{R})$ by simply H^l, and denote the norm of u in $H^l(\mathbb{R})$ by $\|u\|_{H^l}$.

For $x \in \mathbb{R}$ and $r > 0$ we denote by $B(x, r)$ the open ball in $\mathbb{R}$ centered at x with radius r, or in other words the interval $(x - r, x + r)$. Also, for any subset E of $\mathbb{R}$, we denote by χ_E the characteristic function of E, so that $\chi_E(x) = 1$ for $x \in E$ and $\chi_E(x) = 0$ for $x \notin E$.

2 Statement of main result

We begin by reviewing the definition and some basic properties of N-soliton solutions of the Korteweg-de Vries (KdV) equation, and the associated sequence of polynomial conservation laws. For more details and further references, the reader is referred to, for example, [8] and [9]; the early papers [12, 13, 14] of Lax are also very readable.

Suppose $N \in \mathbb{N}$, $0 < C_1 < \cdots < C_N$ and $(\gamma_1, \ldots, \gamma_N) \in \mathbb{R}^N$. An *$N$-soliton profile function* is a function of the form

$$\psi_{C_1,\ldots,C_N;\gamma_1,\ldots,\gamma_N}(x) = 3(D'/D)'$$

where D is defined as the $N \times N$ determinant of Wronskian form,

$$D = D(y_1, \ldots, y_N) = \begin{vmatrix} y_1 & \ldots & y_N \\ y_1' & \ldots & y_N' \\ \ldots & \ldots & \ldots \\ y_1^{(N-1)} & \ldots & y_N^{(N-1)} \end{vmatrix},$$

with

$$y_j(x) = e^{\sqrt{C_j}(x-\gamma_j)} + (-1)^{j-1} e^{-\sqrt{C_j}(x-\gamma_j)}, \quad j = 1, \ldots, N.$$

From N-soliton profile functions, we can construct *N-soliton solutions* of the KdV equation,

$$u_t + uu_x + u_{xxx} = 0, \tag{2.1}$$

simply by defining

$$u(x,t) = \psi_{C_1,\ldots,C_N;\gamma_1(t),\ldots,\gamma_N(t)}(x)$$

where for $j = 1, \ldots, N$,

$$\gamma_j(t) = a_j + C_j t,$$

and $(a_1, \ldots, a_N) \in \mathbb{R}^N$ is arbitrary.

In particular, a single-soliton profile is obtained by taking $N = 1$, in which case we have, for $C > 0$ and $\gamma \in \mathbb{R}$,

$$\psi_{C,\gamma}(x) = 3(D'/D)'$$

where $D = e^{\sqrt{C}(x-\gamma)} + e^{-\sqrt{C}(x-\gamma)}$. In other words,

$$\psi_{C,\gamma}(x) = \frac{3C}{\cosh^2(\sqrt{C}(x-\gamma))}.$$

Then a single-soliton solution of the KdV equation is obtained by taking

$$u(x,t) = \psi_{C,\gamma(t)}(x)$$

where $\gamma(t) = a + Ct$ and $a \in \mathbb{R}$ is arbitrary.

If the constants $\gamma_1, \ldots, \gamma_N$ are widely separated, then the profile function $\psi_{C_1,\ldots,C_N;\gamma_1,\ldots,\gamma_N}$ closely resembles a sum of single-soliton profiles $\sum_{i=1}^N \psi_{C_i,\gamma_i}$. A particular instance of this well-known fact that we will use below is the following:

Lemma 2.1. *Suppose $0 < C_1 < C_2$; $\gamma_1, \gamma_2 \in \mathbb{R}$; and $\{x_n^1\}$ and $\{x_n^1\}$ are sequences such that*

$$\lim_{n\to\infty} |x_n^1 - x_n^2| = \infty.$$

Then

$$\lim_{n\to\infty} \left\| \psi_{C_1,\gamma_1+x_n^1} + \psi_{C_2,\gamma_2+x_n^2} - \psi_{C_1,C_2,\gamma_1+x_n^1,\gamma_2+x_n^2} \right\|_{H^2(\mathbb{R})} = 0.$$

Proof. This follows immediately from Lemma 3.6 of [3]. □

The variational problem we are concerned with here has, as objective and constraint functionals, polynomial conservation laws for the KdV equation. These are functionals of the form

$$E_j(u) = \int_{\mathbb{R}} P\left(u, u_x, u_{xx}, \dots, \frac{\partial^{j-2}u}{\partial x^{j-2}}\right)\, dx$$

where $j \geq 2$ and P is a polynomial in its arguments. They are conservation laws in the sense that, if $u(x,t)$ is a solution of (2.1), then (at least formally),

$$\frac{d}{dt}\left[E_j(u(x,t))\right] = 0$$

for all $t \in \mathbb{R}$. As shown for example in [14], the KdV equation has an infinite family of such conserved functionals, the first three of which are given by:

$$\begin{aligned}
E_2(u) &= \int_{\mathbb{R}} \frac{1}{2}u^2\, dx,\\
E_3(u) &= \int_{\mathbb{R}} \left(\frac{1}{2}u_x^2 - \frac{1}{6}u^3\right)\, dx,\\
E_4(u) &= \int_{\mathbb{R}} \left(\frac{1}{2}u_{xx}^2 - \frac{5}{6}uu_x^2 + \frac{5}{32}u^4\right)\, dx.
\end{aligned}$$

Sobolev embedding theorems imply that for each $N \geq 0$, E_{N+2} defines a continuous functional on the Sobolev space $H^N = H^N(\mathbb{R})$ of real-valued functions whose derivatives up to order N are in $L^2(\mathbb{R})$. For $u \in H^N$, we denote by $\nabla E_k(u)$ the Fréchet derivative of E_k at u, which coincides with the Gateaux derivative of E_k and is therefore defined as a linear functional on H^N by

$$\nabla E_k(u)[v] = \lim_{\epsilon \to 0} \frac{E_k(u + \epsilon v) - E(u)}{\epsilon},$$

and identified with a function $\nabla E_k(u)(x)$ in the usual way, so that $\nabla E_k(u)[v] = \int_{\mathbb{R}} \nabla E_k(u)(x) \cdot v(x)\, dx$ for $v \in H^N$. In particular, the Fréchet derivatives of the first few functionals are given by

$$\begin{aligned}
\nabla E_2(u) &= u,\\
\nabla E_3(u) &= -u_{xx} - \frac{1}{2}u,\\
\nabla E_4(u) &= u_{xxxx} + \frac{5}{3}u_{xx} + \frac{5}{6}(u_x)^2 + \frac{5}{8}u^3.
\end{aligned}$$

For fixed $0 < C_1 < \cdots < C_N$, define

$$S = S(C_1, \dots, C_N) = \left\{\psi_{C_1,\dots,C_N;\gamma_1,\dots,\gamma_N}(x) : (\gamma_1, \dots, \gamma_N) \in \mathbb{R}^N\right\}. \tag{2.2}$$

It is well-known (for a proof, see for example Theorem 3.8 of [2]) that there exist constants $\lambda_2, \ldots \lambda_{N+1}$ such that each $\psi \in S$ satisfies the ordinary differential equation

$$\nabla E_{N+2}(\psi) = \sum_{k=2}^{N+1} \lambda_k \nabla E_k(\psi). \tag{2.3}$$

Thus N-soliton profiles are (non-isolated) critical points for constrained variational problems involving the functionals E_j. The family of ordinary differential equations (2.3) is collectively known as the stationary KdV hierarchy. Due to work of Novikov, Its/Matveev, and Gelfand/Dickey in the 1970's (see [2] for references), it is known that each equation in the hierarchy has the structure of a completely integrable Hamiltonian system, and indeed can be explicitly solved by integration.

In the case $N = 2$, equation (2.3) takes the form

$$\nabla E_4(\psi) = \lambda_2 \nabla E_2(\psi) + \lambda_3 \nabla E_3(\psi), \tag{2.4}$$

or

$$\psi'''' + \frac{5}{3}\psi\psi'' + \frac{5}{6}(\psi')^2 + \frac{5}{8}\psi^3 = \lambda_2\psi - \lambda_3(\psi'' + \frac{1}{2}\psi^2).$$

A fact which is crucial for the proof of our main result below is that there is no choice of the numbers λ_2, λ_3 for which (2.4) has any solutions in $H^2(\mathbb{R})$ besides 1-soliton and 2-soliton profiles:

Theorem 2.2 ([2])**.** *Suppose $\psi \in H^2$ is a solution of* (2.4)*, in the sense of distributions, for some $\lambda_2, \lambda_3 \in \mathbb{R}$. Then ψ is either a 1-soliton or a 2-soliton profile for the KdV equation.*

The proof given in [2] for Theorem 2.2 relies on the fact that, as mentioned above, (2.4) can be explicitly integrated.

We will also make crucial use of the fact that, for all $k \geq 2$, E_k is constant on S, with its value on S given by

$$E_k(C_1, \ldots, C_N) := (-1)^k \frac{36}{2k-1} \sum_{j=1}^{N} C_j^{(2k-1)/2}. \tag{2.5}$$

To prove (2.5), one first shows that it is valid in the case of a single-soliton profile, when $N = 1$ (cf. equation (3.18) of [18]). For a general N-soliton profile $\psi(x) = \psi_{C_1,\ldots,C_N;\gamma_1,\ldots,\gamma_N}(x)$, one then proves (2.5) by considering a solution $u(x,t)$ of (2.1) with $u(x,0) = \psi(x)$. Since $u(x,t)$ resolves into widely separated single-soliton profiles as $t \to \infty$, it follows that $\lim_{t\to\infty} E_k(u(x,t)) = E_k(C_1, \ldots, C_N)$. But since E_k is a conserved functional for KdV, it then follows that $E_k(\psi) = E_k(C_1, \ldots, C_N)$ as well.

We now consider the constrained variational problem of minimizing the functional E_4 over $H^2(\mathbb{R})$, subject to the constraints that E_2 and E_3 be held constant. This is the same variational problem considered, in the case $N = 2$, in the stability theory for N-solitons presented in [18]. Whereas it was shown in [18] that 2-soliton profiles are local minimizers of the variational problem, we will show in Theorem 2.6 that they are global minimizers (and in fact, by the uniqueness result Theorem 2.2, every global minimizer is a 2-soliton profile). Moreover, every minimizing sequence for the problem must converge strongly in $H^2(\mathbb{R})$ to the set of minimizers.

We begin by introducing some notation concerning the variational problem.

Proposition 2.3. *Suppose $(a,b) \in \mathbb{R}^2$. Then there exists a nonzero function g in $H^2(\mathbb{R})$ such that $E_2(g) = a$ and $E_3(g) = b$ if and only if $(a,b) \in \Sigma$, where*

$$\Sigma = \left\{ (a,b) \in \mathbb{R}^2 : a > 0 \text{ and } b \geq -ma^{5/3} \right\},$$

and

$$m = \frac{36}{5}\left(\frac{1}{12}\right)^{5/3}.$$

Proof. From Cazenave and Lions' variational characterization of solitary waves (see for example Theorem 2.9 and Proposition 2.11 of [1]), we know that for all $a > 0$ and all $g \in H^1(\mathbb{R})$ such that $E_2(g) = a$, we have $E_3(g) \geq -ma^{5/3}$; and the minimum value $E_3(g) = -ma^{5/3}$ is attained when (and only when) $g = \psi_{C,\gamma}$, where $C = (a/12)^{2/3}$ and $\gamma \in \mathbb{R}$ is arbitrary. It follows that if g is a nonzero function in $H^2(\mathbb{R})$ with $E_2(g) = a$ and $E_3(g) = b$, then $(a,b) \in \Sigma$.

Conversely, suppose $(a,b) \in \Sigma$. For $\alpha > 0$, define $g_\alpha(x) = \sqrt{\alpha}\psi_{C,\gamma}(\alpha x)$, where C and γ are as above. Then $E_2(g_\alpha) = a$ for every $\alpha > 0$, $E_3(g_\alpha) = -ma^{5/3}$ when $\alpha = 1$, and $E_3(g_\alpha) \to +\infty$ as $\alpha \to \infty$. Therefore, since $b \geq -ma^{5/3}$, the intermediate value theorem implies the existence of some $\alpha \in [1,\infty)$ for which $E_3(g_\alpha) = b$. Thus by taking $g = g_\alpha$, we can satisfy $g \in H^2(\mathbb{R})$, $E_2(g) = a$, and $E_3(g) = b$. □

For $(a,b) \in \Sigma$, define $\Lambda(a,b) \subseteq \mathbb{R}$ by

$$\Lambda(a,b) = \{ r \in \mathbb{R} : \text{for some } g \in H^2(\mathbb{R}),\ E_2(g) = a,\ E_3(g) = b, \text{ and } E_4(b) = r \}. \tag{2.6}$$

By the definition of Σ, the set on the right-hand side is nonempty, and we can therefore define

$$J(a,b) = \inf \Lambda(a,b). \tag{2.7}$$

(Notice that we do not exclude here the possibility that $J(a,b) = -\infty$. However, as shown below at the beginning of Section 5, in fact $J(a,b) > -\infty$ for all $(a,b) \in \Sigma$.)

Definition 2.4. *Suppose $(a,b) \in \Sigma$. We say that a function $\phi \in H^2(\mathbb{R})$ is a* minimizer *for $J(a,b)$ if $E_2(\phi) = a$, $E_3(\phi) = b$, and $E_4(\phi) = J(a,b)$. We say that a sequence $\{\phi_n\}$ of functions in $H^2(\mathbb{R})$ is a* minimizing sequence *for $J(a,b)$ if $\lim_{n\to\infty} E_2(\phi_n) = a$, $\lim_{n\to\infty} E_3(\phi_n) = b$, and $\lim_{n\to\infty} E_4(\phi_n) = J(a,b)$.*

From the uniqueness property stated in Theorem 2.2, it follows that if minimizers for $J(a,b)$ exist, then they must of necessity be 1-soliton or 2-soliton profiles:

Proposition 2.5. *Suppose a and b are real numbers and $u \in H^2$ is a minimizer for $J(a,b)$. Then u must be either a 1-soliton profile or a 2-soliton profile. In other words we have $u(x) = \psi_{D_1,D_2;\gamma_1,\gamma_2}(x)$ for some real numbers D_1, D_2, γ_1, γ_2, with $0 \leq D_1 < D_2$.*

Proof. According to Theorem 2 on page 188 of [16], if u is a regular point of the constraint functionals E_3 and E_4, meaning that the Fréchet derivatives $\nabla E_3(u)$ and $\nabla E_4(u)$ are linearly independent, then there must exist real numbers λ_2 and λ_3 such that the equation $\nabla E_4(u) = \lambda_2 \nabla E_3(u) + \lambda_3 \nabla E_2(u)$ holds, at least in the sense of distributions. In this case, by Theorem 2.2, u is either a 1-soliton or a 2-soliton profile. On the other hand, if u is not a regular point of the constrained functionals, then u satisfies the equation $\nabla E_3(u) = \lambda \nabla E_2(u)$ for some $\lambda \in \mathbb{R}$, and it is an elementary exercise (see for example Theorem 4.2 of [2]) to show that the only possible solutions of this equation in H^2 are 1-soliton profiles. □

However, the preceding result of course leaves open the question of whether any 1-soliton or 2-soliton profiles are in fact minimizers for $J(a,b)$. Our main result determines the set of values of (a,b) within Σ for which minimizers for $J(a,b)$ exist; and for such values of (a,b) determines the value of $J(a,b)$, describes all the minimizers for $J(a,b)$, and describes the behavior of minimizing sequences for $J(a,b)$. If $S \subseteq H^2(\mathbb{R})$, we say that a sequence $\{\phi_n\}$ converges to S in $H^2(\mathbb{R})$ norm if

$$d(\phi_n, S) = \inf_{\psi \in S} \|\phi_n - \psi\|_{H^2(\mathbb{R})} \to 0 \quad \text{as } n \to \infty,$$

or, equivalently, if there exists a sequence $\{\psi_n\}$ of elements of S such that

$$\lim_{n\to\infty} \|\phi_n - \psi_n\|_{H^2(\mathbb{R})} = 0.$$

Theorem 2.6. *Suppose $(a,b) \in \Sigma$; that is, $a > 0$ and $b \geq -ma^{5/3}$, where $m = \frac{36}{5}\left(\frac{1}{12}\right)^{5/3}$.*

1. *If*
$$b = -ma^{5/3}, \tag{2.8}$$
then every minimizing sequence for $J(a,b)$ converges to $S(C)$ in $H^2(\mathbb{R})$ norm, where $S(C)$ is as defined in (2.2), with $C = (a/12)^{2/3} = (-5b/36)^{2/5}$. Every element of $S(C)$ is a minimizer for $J(a,b)$, and $J(a,b) = E_4(C)$.

2. *If*
$$\frac{-ma^{5/3}}{2^{2/3}} > b > -ma^{5/3}, \tag{2.9}$$

then every minimizing sequence for $J(a,b)$ converges to $S(C_1,C_2)$ in $H^2(\mathbb{R})$ norm, where (C_1,C_2) is the unique pair of numbers such that $0 < C_1 < C_2$, $E_2(C_1,C_2) = a$, and $E_3(C_1,C_2) = b$. Every element of $S(C_1,C_2)$ is a minimizer for $J(a,b)$, and $J(a,b) = E_4(C_1,C_2)$.

3. *If*

$$b \geq -\frac{ma^{5/3}}{2^{2/3}}, \tag{2.10}$$

then there do not exist any minimizers for $J(a,b)$ in H^2.

We remark that the method used below to analyze the behavior of minimizing sequences under the assumptions (2.8) or (2.9) should also be applicable in case (2.10) holds, and suggests that in the latter case, minimizing sequences $\{\phi_n\}$ should, as $n \to \infty$, come to resemble superpositions of widely separated single-soliton profiles. Thus, for example, if $b = -\frac{ma^{5/3}}{2^{2/3}}$ we expect that if $\{\phi_n\}$ is a minimizing sequence for $J(a,b)$, then there will exist a number $C > 0$ and sequences $\{\gamma_{1n}\}$ and $\{\gamma_{2n}\}$ with $\lim_{n\to\infty} |\gamma_{1n} - \gamma_{2n}| = \infty$ such that

$$\lim_{n\to\infty} \|\phi_n - (\psi_{C,\gamma_{1n}} + \psi_{C,\gamma_{2n}})\|_{H^2} = 0.$$

Similarly, if $b > -\frac{ma^{5/3}}{2^{2/3}}$, we expect that the functions in a typical minimizing sequence for $J(a,b)$ would resemble a superposition of three or more single-soliton profiles, two of which have equal amplitudes and whose distance from each other increases to infinity as $n \to \infty$. We do not pursue this topic further here, however.

We also remark that the method of proof of Theorem 2.6 should apply as well to the variational problems satisfied by N-soliton profiles for general $N \in \mathbb{N}$. One important obstacle we have encountered, however, is that of proving an analogue of Theorem 2.2: i.e., of showing that for all possible choices of the numbers $\lambda_2, \ldots, \lambda_{N+1}$, the Euler-Lagrange equation

$$\nabla E_{N+2}(\psi) = \lambda_2 \nabla E_2(\psi) + \cdots + \lambda_{N+1} \nabla E_{N+1}(\psi) \tag{2.11}$$

has no solutions in H^N besides N-soliton profiles. As noted in [2], the explicit integration of equation (2.11) can be carried out for general N just as it can for $N = 2$, whenever the value of $(\lambda_2, \ldots, \lambda_{N+1})$ corresponds to that of an N-soliton profile; but technical difficulties arise in proving that solutions corresponding to other values of $(\lambda_2, \ldots, \lambda_{N+1})$ are singular.

An immediate consequence of Theorem 2.6 is a stability result for 2-soliton solutions of the KdV equation. This recovers (by a different proof) the special case $N = 2$ of the general result for N-soliton solutions given by Maddocks and Sachs in [18].

Corollary 2.7. *Suppose $0 < C_1 < C_2$. Then every minimizing sequence $\{\phi_n\}$ for $J(E_2(C_1, C_2), E_3(C_1, C_2))$ converges strongly to $S(C_1, C_2)$ in $H^2(\mathbb{R})$. Moreover, $S = S(C_1, C_2)$ is stable, in the sense that for every $\epsilon > 0$ there exists $\delta > 0$ such that if $u_0 \in H^2(\mathbb{R})$ and $d(u_0, S) < \delta$, then $d(u(\cdot, t), S) < \epsilon$ for all $t > 0$.*

Remark. It follows from this stability result that there are C^1 functions $\gamma_1(t)$ and $\gamma_2(t)$ defined for $t \geq 0$ such that

$$\|u(\cdot, t) - \psi_{C_1, C_2; \gamma_1(t), \gamma_2(t)}\|_{H^2(\mathbb{R})} < \epsilon$$

and $|\gamma_1'(t) - C_1| < \epsilon$, $|\gamma_2'(t) - C_2| < \epsilon$ for all $t > 0$. See [3].

We conclude this section by sketching the idea of the proof of Theorem 2.6 which is given in the succeeding sections.

Suppose $(a, b) \in \Sigma$ with $-ma^{5/3} \leq b < \dfrac{-ma^{5/3}}{2^{2/3}}$, and suppose $\{\phi_n\}$ is a minimizing sequence for $J(a, b)$. We can apply the profile decomposition, in the form of Corollary 3.4, to the sequence $\{\rho_n\}$ defined by

$$\rho_n = |\phi_n|^2 + |\phi_n'|^2 + |\phi_n''|^2.$$

It follows that we can decompose ϕ_n as

$$\phi_n = \sum_{i=1}^{n} v_n^i + w_n,$$

where the sequence $\{w_n\}$ is vanishing in the sense of Definition 3.1, and for each $i \in \mathbb{N}$ the sequence $\{v_n^i\}_{n \in \mathbb{N}}$ concentrates around some sequence $\{x_n^i\}_{n \in \mathbb{N}}$, in the sense of Definition 3.2. From the concentration property of $\{v_n^i\}_{n \in \mathbb{N}}$ and the fact that $\{\phi_n\}$ is a minimizing sequence, it follows that the sequence $\{v_n^i\}_{n \in \mathbb{N}}$ can be suitably translated so that it converges, weakly in $H^2(\mathbb{R})$ and strongly in $H^1(\mathbb{R})$, to a minimizer g_i for the variational problem

$$E_4(g_i) = \inf \left\{ E_4(\phi) : \phi \in H^2(\mathbb{R}), E_2(\phi) = a_i, E_3(\phi) = b_i \right\}$$

for some real numbers a_i, b_i. As a critical point of this constrained variational problem, $\psi = g_i$ must satisfy the Euler-Lagrange equation (2.4). From Theorem 2.2 it then follows that for each i we have $g_i = \psi_{D_{1i}, D_{2i}}$ for some numbers D_{1i} and D_{2i} with $0 \leq D_{1i} < D_{2i}$.

In parts 1 and 2 of Theorem 2.6, our assumption on (a, b) implies that there exist numbers C_1 and C_2 with $0 \leq C_1 < C_2$ such that $E_2(\psi_{C_1, C_2}) = a$ and $E_3(\psi_{C_1, C_2}) = b$. Therefore, by definition of $J(a, b)$, we have that

$$J(a, b) \leq E_4(\psi_{C_1, C_2}) = (36/7) \left(C_1^{7/2} + C_2^{7/2} \right).$$

From the profile decomposition and the fact that $\{\phi_n\}$ is a minimizing sequence, we can obtain that

$$\sum_{i=1}^{\infty} E_2(g_i) = 36 \sum_{i=1}^{\infty} \left(D_{1i}^3 + D_{2i}^3\right) \leq \lim_{n\to\infty} E_2(\phi_n) = 36 \left(C_1^3 + C_2^3\right)$$
$$\sum_{i=1}^{\infty} E_3(g_i) = \frac{-36}{5} \sum_{i=1}^{\infty} \left(D_{1i}^5 + D_{2i}^5\right) \leq \lim_{n\to\infty} E_3(\phi_n) = \frac{-36}{5} \left(C_1^5 + C_2^5\right)$$
$$\sum_{i=1}^{\infty} E_4(g_i) = \frac{36}{7} \sum_{i=1}^{\infty} \left(D_{1i}^7 + D_{2i}^7\right) \leq \lim_{n\to\infty} E_4(\phi_n) \leq \frac{36}{7} \left(C_1^7 + C_2^7\right).$$

Permuting the terms of the sequence $(D_{11}^{1/2}, D_{21}^{1/2}, D_{21}^{1/2}, D_{22}^{1/2}, D_{31}^{1/2}, D_{32}^{1/2}, \dots)$ so that they form a decreasing sequence $(x_1, x_2, x_3, \dots)$; and defining $y_1 = C_2^{1/2}$ and $y_2 = C_1^{1/2}$, we thus have that

$$\sum_{i=1}^{\infty} x_i^3 \leq y_1^3 + y_2^3$$
$$\sum_{i=1}^{\infty} x_i^5 \geq y_1^5 + y_2^5$$
$$\sum_{i=1}^{\infty} x_i^7 \leq y_1^7 + y_2^7.$$

We analyze this system of inequalities in Section 4, where we show (cf. Lemma 4.10) that it can only be satisfied if $x_1 = C_2$, $x_2 = C_1$, and $x_i = 0$ for all $i \geq 3$. This can be interpreted as saying that, among the N-soliton profiles of the KdV equation, the only ones which could possibly solve the variational problem are 1-soliton profiles (in the case when $C_1 = 0$) and 2-soliton profiles (in the case when $C_1 > 0$).

This information, combined with the control on the functions $\{v_n^i\}$ and w_n afforded by the profile decomposition, is enough to allow us to deduce that the functions in the minimizing sequence $\{\phi_n\}$ are either of the form

$$\phi_n(x) = \psi_{C_1,C_2}(x + x_n) + r_n(x)$$

for some sequence $\{x_n\}$ of real numbers, where $r_n \to 0$ in $H^2(\mathbb{R})$; or of the form

$$\phi_n(x) = \phi_{C_1}(x + x_n^1) + \phi_{C_2}(x + x_n^2) + r_n(x)$$

for some pair of sequences $\{x_n^1\}$ and $\{x_n^2\}$ of real numbers with $|x_n^1 - x_n^2| \to \infty$, where again $r_n \to 0$ in $H^2(\mathbb{R})$. In either case, this shows that the set of minimizing functions for the variational problem consists of the set $S(C_1, C_2)$, and that $\{\phi_n\}$ converges to $S(C_1, C_2)$ in $H^2(\mathbb{R})$ norm.

Part 3 of Theorem 2.6 will follow from a simpler argument: under the given assumptions on (a, b), no 1-soliton profile or 2-soliton profile ψ can exist satisfying the constraints $E_2(\psi) = a$ and $E_3(\psi) = b$. But any minimizer for the variational problem must satisfy the associated Euler-Lagrange equation (2.4), and therefore by Theorem 2.2 must be either a 1-soliton profile or a 2-soliton profile. Hence no minimizers can exist.

3 Profile decomposition

The idea of the proof of Theorem 2.6 is to use an elaboration of the method of concentration compactness, known as *profile decomposition*, which details the ways in which a sequence of measures of bounded total mass can lose compactness.

The technique of profile decomposition dates back to [10] and in some form even earlier (see for example [6]). We will use a version which is due to Mariş [19]. Actually, although the result of [19] is valid for arbitrary bounded sequences of Borel measures on any metric space, for simplicity of notation we here restrict consideration to bounded sequences of nonnegative functions in $L^1(\mathbb{R})$.

Definition 3.1. *We say that a sequence $\{f_n\}$ of nonnegative functions in $L^1(\mathbb{R})$ is* vanishing *if for every $r > 0$, we have*

$$\lim_{n\to\infty} \sup_{y\in\mathbb{R}} \int_{B(y,r)} f_n = 0.$$

Definition 3.2. *If $\{f_n\}$ is a sequence of nonnegative functions in $L^1(\mathbb{R})$ and $\{x_n\}$ is a sequence of real numbers, we say that $\{f_n\}$* concentrates around $\{x_n\}$ *if for every $\epsilon > 0$, there exists $r_\epsilon > 0$ such that*

$$\int_{\mathbb{R}\setminus B(x_n,r_\epsilon)} f_n < \epsilon$$

for every $n \in \mathbb{N}$.

Theorem 3.3 ([19])**.** *Suppose $\{\rho_n\}$ is a sequence of nonnegative functions which is bounded in $L^1(\mathbb{R})$. Then either $\{\rho_n\}$ is vanishing, or there exists a subsequence of $\{\rho_n\}$ (which we continue to denote by $\{\rho_n\}$), which satisfies one of the following two properties: either*

(1) there exist $k \in \mathbb{N}$ and for each $i \in \{1,\dots,k\}$ a number $m_i > 0$ and sequence of balls $\{B(x_n^i, r_n^i)\}_{n\in\mathbb{N}}$ in $\mathbb{R}$ with $\lim\limits_{n\to\infty} r_n^i = \infty$, such that

(a) $B(x_n^i, r_n^i) \cap B(x_n^j, r_n^j) = \emptyset$ for all $n \in \mathbb{N}$ and all $i, j \in \{1,\dots,k\}$ with $i \neq j$,

(b) for each $i \in \{1,\dots,k\}$, $\lim\limits_{n\to\infty} \int_{B(x_n^i, r_n^i/2)} \rho_n = m_i$,

(c) for each $i \in \{1,\dots,k\}$, $\lim\limits_{n\to\infty} \int_{B(x_n^i, r_n^i)\setminus B(x_n^i, r_n^i/2)} \rho_n = 0$,

(d) for each $i \in \{1,\dots,k\}$, the sequence $\{\rho_n \chi_{B(x_n^i, r_n^i)}\}_{n\in\mathbb{N}}$ concentrates around $\{x_n^i\}_{n\in\mathbb{N}}$,

(e) the sequence $\{\rho_n \chi_{\mathbb{R}\setminus \cup_{i=1}^k B(x_n^i, r_n^i)}\}_{n\in\mathbb{N}}$ is vanishing;

or

(2) *for each* $i \in \mathbb{N}$ *there is a number* $m_i > 0$ *and a sequence of balls* $\{B(x_n^i, r_n^i)\}_{n=i,i+1,i+2,\ldots}$ *in* $\mathbb{R}$, *with* $\lim_{n\to\infty} r_n^i = \infty$, *such that*

   (a) $B(x_n^i, r_n^i) \cap B(x_n^j, r_n^j) = \emptyset$ *for all* $i, j \in \mathbb{N}$ *with* $i \neq j$, *and all* $n \in \mathbb{N}$ *with* $n \geq i$ *and* $n \geq j$,

   (b) *for each* $i \in \mathbb{N}$, $\displaystyle\lim_{n\to\infty} \int_{B(x_n^i, r_n^i/2)} \rho_n = m_i$,

   (c) *for each* $i \in \mathbb{N}$, $\displaystyle\sum_{n=i}^{\infty} \int_{B(x_n^i, r_n^i)\setminus B(x_n^i, r_n^i/2)} \rho_n \leq \frac{1}{2^i}$,

   (d) *for each* $i \in \mathbb{N}$, *the sequence* $\{\rho_n \chi_{B(x_n^i, r_n^i)}\}_{n\geq i}$ *concentrates around* $\{x_n^i\}_{n\geq i}$,

   (e) *the sequence* $\{\rho_n \chi_{\mathbb{R}\setminus \cup_{i=1}^n B(x_n^i, r_n^i)}\}_{n\in\mathbb{N}}$ *is vanishing, and*

   (f) *if for each* $N \in \mathbb{N}$ *and each* $n \geq N$, *we define* $g_n^N = \rho_n \chi_{\mathbb{R}\setminus \cup_{i=1}^N B(x_n^i, r_n^i)}$, *and define the increasing function* $q_n^N(r)$ *for* $r > 0$ *by*

$$q_n^N(r) = \sup_{y\in\mathbb{R}} \int_{B(y,r)} g_n^N,$$

   then

$$\lim_{N\to\infty} \left(\lim_{r\to\infty} \left(\limsup_{n\to\infty} q_n^N(r) \right) \right) = 0. \tag{3.1}$$

One can view (3.1) as saying that although, for any given value of N, the sequence $\{g_n^N\}_{n\in\mathbb{N}}$ is not necessarily vanishing, it does come closer, in some sense, to being a vanishing sequence as $N \to \infty$.

To shorten our proof of Theorem 2.6, we observe that the two cases in Theorem 3.3 can be combined into one, if we drop the requirement that $m_i > 0$ for each i:

Corollary 3.4. *Suppose* $\{\rho_n\}$ *is a sequence of nonnegative functions which is bounded in* $L^1(\mathbb{R})$, *and suppose* $\{\rho_n\}$ *is not vanishing. Then there exists a subsequence of* $\{\rho_n\}$ *(which we continue to denote by* $\{\rho_n\}$*), a sequence* $\{m_i\}_{i\in\mathbb{N}}$ *of nonnegative numbers, and for each* $i \in \mathbb{N}$ *a sequence of balls* $\{B(x_n^i, r_n^i)\}_{n\in\mathbb{N}}$ *in* $\mathbb{R}$, *with* $\lim_{n\to\infty} r_n^i = \infty$, *such that*

(a) $B(x_n^i, r_n^i) \cap B(x_n^j, r_n^j) = \emptyset$ *for all* $i, j \in \mathbb{N}$ *with* $i \neq j$, *and all* $n \in \mathbb{N}$,

(b) *for each* $i \in \mathbb{N}$, $\displaystyle\lim_{n\to\infty} \int_{B(x_n^i, r_n^i/2)} \rho_n = m_i$,

(c) *for each* $i \in \mathbb{N}$, $\displaystyle\sum_{n=i}^{\infty} \int_{B(x_n^i, r_n^i)\setminus B(x_n^i, r_n^i/2)} \rho_n \leq \frac{1}{2^i}$,

(d) *for each* $i \in \mathbb{N}$, *the sequence* $\{\rho_n \chi_{B(x_n^i, r_n^i)}\}_{n\in\mathbb{N}}$ *concentrates around* $\{x_n^i\}_{n\in\mathbb{N}}$,

(e) *the sequence* $\{\rho_n\chi_{\mathbb{R}\setminus\cup_{i=1}^n B(x_n^i,r_n^i)}\}_{n\in\mathbb{N}}$ *is vanishing, and*

(f) *if for each* $N\in\mathbb{N}$ *and each* $n\in\mathbb{N}$*, we define* $g_n^N=\rho_n\chi_{\mathbb{R}\setminus\cup_{i=1}^N B(x_n^i,r_n^i)}$*, and define the increasing function* $q_n^N(r)$ *for* $r>0$ *by*

$$q_n^N(r)=\sup_{y\in\mathbb{R}}\int_{B(y,r)} g_n^N,$$

then

$$\lim_{N\to\infty}\left(\lim_{r\to\infty}\left(\limsup_{n\to\infty} q_n^N(r)\right)\right)=0.$$

Proof. To obtain Corollary 3.4 from Theorem 3.3, we observe that if (2) holds in Theorem 3.3, then the statements in Corollary 3.4 will also hold if we simply define $B(x_n^i,r_n^i)=\emptyset$ when $i<n$. So we need only consider the case when (1) holds in Theorem 3.3.

Define $E_n=\mathbb{R}\setminus\cup_{i=1}^k B(x_n^i,r_n^i)$ for $n\in\mathbb{N}$. Since $\{\rho_n\chi_{E_n}\}_{n\in\mathbb{N}}$ is vanishing, then for each fixed $j\in\mathbb{N}$,

$$\lim_{n\to\infty}\sup_{y\in\mathbb{R}}\int_{B(y,j)}\rho_n\chi_{E_n}=0.$$

Therefore we can define a sequence $n_1<n_2<n_3<\dots$ such that

$$\sup_{y\in\mathbb{R}}\int_{B(y,j)}\rho_{n_j}\chi_{E_{n_j}}\le\frac{1}{2^{j+1}}$$

for all $j\in\mathbb{N}$. If we now pass to the subsequence $\{\rho_{n_j}\}_{j\in\mathbb{N}}$, continuing to denote this subsequence by $\{\rho_n\}_{n\in\mathbb{N}}$, we have that

$$\sup_{y\in\mathbb{R}}\int_{B(y,n)}\rho_n\chi_{E_n}\le\frac{1}{2^{n+1}} \tag{3.2}$$

for all $n\in\mathbb{N}$. Also, because of part (1)(c) of Theorem 3.3, by passing to a further subsequence we can guarantee that

$$\int_{B(x_n^i,r_n^i)\setminus B(x_n^i,r_n^i/2)}\rho_n\le\frac{1}{2^{n+1}} \tag{3.3}$$

holds for all $i\in\{1,2,\dots,k\}$ and all $n\in\mathbb{N}$ as well.

For each $i\ge k+1$, we set $m_i=0$, and for all $n\in\mathbb{N}$ we define $r_n^i=n$ if $n\ge i$ and $r_n^i=0$ if $n<i$. For each fixed $n\in\mathbb{N}$, we define a sequence $\{x_n^j\}$ inductively for all $i\ge k+1$ by choosing x_n^i to be any real number such that $B(x_n^i,r_n^i)$ is disjoint from $\cup_{j=1}^{i-1}B(x_n^j,r_n^j)$. Then we have that $\lim_{n\to\infty} r_n^i=\infty$ for each $i\in\mathbb{N}$, and part (a) of the Corollary holds.

For all $i\ge k+1$ and for all $n\in\mathbb{N}$, since $B(x_n^i,r_n^i)\subset E_n$, it follows from (3.2) that

$$\int_{B(x_n^i,r_n^i)}\rho_n\le\frac{1}{2^{n+1}}. \tag{3.4}$$

This implies that parts (b) and (c) of the Corollary hold for all $i \geq k+1$. For $1 \leq i \leq k$ we already know that part (b) holds, and part (c) follows from (3.3).

To prove part (d), we fix i such that $i \geq k+1$, and observe that by (3.4), for every $\epsilon > 0$ we can find $N \in \mathbb{N}$ such that $\int_{\mathbb{R}} \rho_n \chi_{B(x_n^i, r_n^i)} < \epsilon$ for all $n > N$. Also, for each $n \in \{1, \dots, N\}$, since $\rho_n \chi_{B(x_n^i, r_n^i)} \in L^1(\mathbb{R})$, we can find $r_{\epsilon,n} > 0$ such that

$$\int_{\mathbb{R}\setminus B(x_n^i, r_{\epsilon,n})} \rho_n \chi_{B(x_n^i, r_n^i)} < \epsilon.$$

Then if we set $r_\epsilon = \max\{r_{\epsilon,1}, \dots, r_{\epsilon,N}\}$, we have that

$$\int_{\mathbb{R}\setminus B(x_n^i, r_\epsilon)} \rho_n \chi_{B(x_n^i, r_n^i)} < \epsilon$$

for all $n \in \mathbb{N}$. This proves part (d) for all i such that $i \geq k+1$, and we already know that part (d) holds for $1 \leq i \leq k$.

Finally, part (e) of the Corollary follows immediately from part (1)(e) of Theorem 3.3, as does part (f) of the Corollary; since (1)(e) of Theorem 3.3 implies that $\lim_{n\to\infty} q_n^N(r) = 0$ for every $r > 0$ and every $N \geq k$. ☐

We record the following important feature of vanishing sequences.

Lemma 3.5. *Suppose $2 < p \leq \infty$. Then there exists a constant $C_p > 0$ such that for all $u \in H^1(\mathbb{R})$,*

$$\|u\|_{L^p(\mathbb{R})} \leq C_p \left(\sup_{y\in\mathbb{R}} \int_{B(y,1)} |u'|^2 + |u|^2 \; dx \right)^{\frac{1}{2}-\frac{1}{p}} \|u\|_{H^1(\mathbb{R})}^{\frac{2}{p}} \tag{3.5}$$

Proof. This lemma is standard; a proof can be found, for example, in [19]. ☐

Corollary 3.6. *Suppose $\{g_n\}$ is a bounded sequence in $H^1(\mathbb{R})$. If $\{|g_n|^2 + |g_n'|^2\}$ is vanishing, then $\lim_{n\to\infty} \|g_n\|_{L^p} = 0$ for all $p > 2$.*

4 Minimizers among the set of N-solitons

Lemma 4.1. *Suppose $A, B > 0$ and $k \in \mathbb{N}$. If the system of equations*

$$\begin{aligned} \sum_{i=1}^k x_i^3 &= A^3 \\ \sum_{i=1}^k x_i^5 &= B^5 \end{aligned} \tag{4.1}$$

has a solution $(x_1, \dots, x_k)$ with $x_k \geq 0$ for $i = 1, \dots, k$; then

$$\left(\frac{1}{k}\right)^{2/15} \leq \frac{B}{A} \leq 1. \tag{4.2}$$

Proof. Suppose the system (4.1) has a solution $(x_1, \dots, x_k)$ with $x_i \geq 0$ for $i = 1, \dots, k$. Defining $p_i = x_i^3/A^3$ for each i, we have that $0 \leq p_i \leq 1$, so $p_i^{5/3} \leq p_i$. Therefore

$$(B/A)^5 = \sum_{i=1}^{k} p_i^{5/3} \leq \sum_{i=1}^{k} p_i = 1,$$

which implies that $B/A \leq 1$. Also, by Hölder's inequality we have

$$A^3 = \sum_{i=1}^{k} x_i^3 \leq \left(\sum_{i=1}^{k} x_i^5\right)^{3/5} \left(\sum_{i=1}^{k} 1\right)^{2/5} = B^3 k^{2/5},$$

which implies that $(1/k)^{2/15} \leq B/A$. ☐

Lemma 4.2. *Suppose $A, B > 0$ and consider the systems*

$$\begin{aligned} y_1^3 + y_2^3 &= A^3 \\ y_1^5 + y_2^5 &= B^5 \end{aligned} \tag{4.3}$$

and

$$\begin{aligned} 2y_1^3 + y_2^3 &= A^3 \\ 2y_1^5 + y_2^5 &= B^5 \end{aligned} \tag{4.4}$$

for (y_1, y_2) in the first quadrant $U = \{(y_1, y_2) \in \mathbb{R}^2 : y_1 \geq 0, y_2 \geq 0\}$.

1. *If $B/A = 1$, then (4.3) has exactly two solutions in U, given by $(0, A)$ and $(A, 0)$, and and (4.4) has exactly one solution in U, given by $(0, A)$.*

2. *If $(1/2)^{2/15} < B/A < 1$, then (4.3) has exactly two solutions in U, which are of the form (α, β) and (β, α),where $0 < \alpha < \beta$; and (4.4) has exactly one solution in U, which is of the form (γ, δ) where $0 < \gamma < \delta$.*

3. *If $B/A = (1/2)^{2/15}$, then (4.3) has exactly one solution in U, which is given by $(A/2^{1/3}, A/2^{1/3})$; and (4.4) has exactly two solutions in U: one given by $(A/2^{1/3}, 0)$, and one of the form (γ, δ) where $0 < \gamma < \delta$.*

4. *If $(1/3)^{2/15} < B/A < (1/2)^{2/15}$, then (4.3) has no solutions in U, and (4.4) has exactly two solutions in U, which are of the form (γ_1, δ_1) and (γ_2, δ_2), where $0 < \gamma_1 < \delta_1$ and $0 < \delta_2 < \gamma_2$.*

5. *If $B/A = (1/3)^{2/15}$, then (4.3) has no solutions in U, and (4.4) has exactly one solution in U, which is given by $(A/3^{1/3}, A/3^{1/3})$.*

Proof. Suppose (y_1, y_2) solves (4.3) and $y_1, y_2 > 0$. Letting $\theta = y_2/y_1$, we obtain that $y_1 = A/(1 + \theta^3)^{1/3}$ and $g(\theta) = (B/A)^{15}$, where

$$g(t) := \frac{(1 + t^5)^3}{(1 + t^3)^5}. \tag{4.5}$$

Conversely, for each choice of $\theta > 0$ such that $g(\theta) = (B/A)^{15}$, we have a solution (y_1, y_2) of (4.3) given by the positive numbers $y_1 = A/(1+\theta^3)^{1/3}$ and $y_2 = \theta y_1$. Therefore, for a given choice of B/A, the number of solutions (y_1, y_2) of (4.3) with $y_1 > 0$ and $y_2 > 0$ is equal to the number of solutions $\theta > 0$ to the equation $g(\theta) = (B/A)^{15}$.

We have that $g(0) = 1$, $g(1) = 1/4$, $\lim_{t\to\infty} g(t) = 1$, and

$$g'(t) = \frac{15(1+t^5)^2(t^4-t^2)}{(1+t^3)^6},$$

so $g(t)$ is monotone decreasing for $0 \le t \le 1$ and monotone increasing for $1 \le t < \infty$. Therefore the equation $g(\theta) = (B/A)^{15}$ has exactly two solutions when $B/A \in ((1/2)^{2/15}, 1)$, and has exactly one solution when $B/A = (1/2)^{2/15}$. The assertions of the lemma concerning (4.3) then follow.

On the other hand, when $y_1 > 0$, we have that (y_1, y_2) solves (4.4) if and only if $y_1 = A/(2+\theta^3)^{1/3}$, $y_2 = \theta y_1$, and $h(\theta) = (B/A)^{15}$, where

$$h(t) := \frac{(2+t^5)^3}{(2+t^3)^5}.$$

We have $h(0) = 1/4$, $h(1) = 1/9$ and $\lim_{t\to\infty} h(t) = 1$, and

$$h'(t) = \frac{30(2+t^5)^2(t^4-t^2)}{(2+t^3)^6},$$

so that $h(t)$ is monotone decreasing for $0 \le t \le 1$ and monotone increasing for $1 \le t < \infty$. When $B/A = (1/3)^{3/15}$, the equation $h(\theta) = (B/A)^{15}$ has exactly one solution, namely $\theta = 1$. When $B/A \in ((1/3)^{2/15}, (1/2)^{2/15})$, the equation $h(\theta) = (B/A)^{15}$ has exactly two solutions, one of which is greater than one and one of which is less than one. When $B/A = (1/2)^{2/15}$ there are again exactly two solutions, one of which is $\theta = 0$ and the other of which is a value $\theta > 1$. Finally, when $B/A \in ((1/2)^{2/15}, 1)$, there is exactly one solution to $h(\theta) = (B/A)^{15}$, and it satisfies $\theta > 1$. These statements imply the assertions of the lemma concerning (4.4). □

Definition 4.3. *Let*

$$D = \{(A,B) \in \mathbb{R}^2 : A > 0, B > 0, (1/2)^{2/15} \le B/A \le 1\}.$$

For each $(A,B) \in D$, *we define*

$$m(A,B) = y_1^7 + y_2^7,$$

where (y_1, y_2) *is the unique solution to* (4.3) *satisfying* $0 \le y_1 \le y_2$ *and* $y_2 > 0$, *guaranteed by Lemma 4.2.*

Lemma 4.4.

1. *For all* $(A, B) \in D$*, and for every* $\lambda > 0$*, we have*

$$m(\lambda A, \lambda B) = \lambda^7 m(A, B). \tag{4.6}$$

2. *The function* $m(A, B)$ *is continuous on the set* D*.*

Proof. The homogeneity property (4.6) of $m(A, B)$ is an easy consequence of the definition of $m(A, B)$. To see that $m(A, B)$ is continuous on D, observe first that for given $(A, B) \in D$, the numbers y_1 and y_2 given in Definition 4.3 are given by $y_2 = A/(\tilde{\theta}^3 + 1)^{1/3}$ and $y_1 = \tilde{\theta} y_2$, where $\tilde{\theta} = \tilde{\theta}(A, B)$ is the unique solution in $[0, 1]$ of the equation $g(\tilde{\theta}) = (B/A)^{15}$, and g is the function defined in (4.5). Since g is continuous and monotone decreasing on $[0, 1]$, and $g([0, 1]) = [1/4, 1]$, then the inverse map $h : [1/4, 1] \to [0, 1]$ defined by $h(g(t)) = t$ is also continuous. Therefore $\tilde{\theta}(A, B) = h((B/A)^{15})$ is continuous on D, so y_2 and hence also y_1 depend continuously on (A, B). So $m(A, B) = y_1^7 + y_2^7$ is continuous on D as well. □

Lemma 4.5. *Suppose* $y_1 \geq y_2 \geq 0$ *and* $z_1 \geq z_2 \geq 0$*, and*

$$\begin{aligned} z_1^3 + z_2^3 &\leq y_1^3 + y_2^3 \\ z_1^5 + z_2^5 &\geq y_1^5 + y_2^5. \end{aligned} \tag{4.7}$$

Then

$$z_1^7 + z_2^7 \geq y_1^7 + y_2^7. \tag{4.8}$$

Equality holds in (4.8) *only if* $y_1 = z_1$ *and* $y_2 = z_2$*.*

Proof. We may assume $y_1 > 0$ and $z_1 > 0$, or otherwise there is nothing to prove. Let $A^3 = y_1^3 + y_2^3$, $B^5 = y_1^5 + y_2^5$, and $C^7 = y_1^7 + y_2^7$. By Lemmas 4.1 and 4.2, we have $(1/2)^{2/15} \leq B/A \leq 1$. Define $\theta = y_2/y_1 \in [0, 1]$ and $\tilde{\theta} = z_2/z_1 \in [0, 1]$, and define the function g as in (4.5). Then from the definitions of A and B we deduce that $g(\theta) = (B/A)^{15} \in [1/4, 1]$; and from (4.7) we have that

$$\frac{B}{(1 + \tilde{\theta}^5)^{1/5}} \leq z_1 \leq \frac{A}{(1 + \tilde{\theta}^3)^{1/3}}, \tag{4.9}$$

which implies that $g(\tilde{\theta}) \geq (B/A)^{15} = g(\theta)$. Since, as shown in the proof of Lemma 4.2, g is monotone decreasing on $[0, 1]$, it follows that $\tilde{\theta} \leq \theta$. Now define

$$k(t) := \frac{(1 + \theta^7)^5}{(1 + \theta^5)^7}.$$

Then as in the proof of Lemma 4.2, an elementary computation (whose details we omit) shows that $k(t)$ is, like $g(t)$, strictly decreasing on $[0, 1]$ and strictly increasing on $[1, \infty)$. Therefore $k(\tilde{\theta}) \geq k(\theta) = (C/B)^{35}$, and so

$$\frac{C}{(1 + \tilde{\theta}^7)^{1/7}} \leq \frac{B}{(1 + \tilde{\theta}^5)^{1/5}}. \tag{4.10}$$

Taken with (4.9), this implies that

$$\frac{C}{(1+\tilde{\theta}^7)^{1/7}} \le z_1, \tag{4.11}$$

which yields (4.8).

If equality holds in (4.8), then equality also holds in (4.11), so from (4.9) and (4.10) we have that equality holds in (4.10). Therefore $k(\tilde{\theta}) = k(\theta)$. Since k is strictly decreasing on $[0,1]$, this implies that $\tilde{\theta} = \theta$, and hence $g(\tilde{\theta}) = g(\theta)$ and so $\tilde{B}/\tilde{A} = B/A$. But from (4.7) we have that $\tilde{A} \le A$ and $\tilde{B} \ge B$, so it follows that $\tilde{A} = A$ and $\tilde{B} = B$. Hence z_1 and z_2 satisfy the same equation (4.3) as y_1 and y_2, so by Lemma 4.2, we must have $z_1 = y_1$ and $z_2 = y_2$. ☐

Lemma 4.6. *Suppose $A, B > 0$ and $(1/3)^{2/15} < B/A < 1$. For $x = (x_1, x_2, x_3) \in \mathbb{R}^3$, define*

$$\begin{aligned} g_1(x) &= x_1^3 + x_2^3 + x_3^3 \\ g_2(x) &= x_1^5 + x_2^5 + x_3^5 \\ f(x) &= x_1^7 + x_2^7 + x_3^7, \end{aligned}$$

and define

$$\begin{aligned} \Gamma &= \left\{x \in \mathbb{R}^3 : g_1(x) = A^3 \text{ and } g_2(x) = B^5\right\} \\ \Omega &= \left\{x \in \mathbb{R}^3 : x_1 > 0,\, x_2 > 0, \text{ and } x_3 > 0\right\}. \end{aligned} \tag{4.12}$$

Then $\Gamma \cap \Omega$ is nonempty, and is a smooth one-dimensional submanifold of $\mathbb{R}^3$.

If we assume further that $B/A \ge (1/2)^{2/15}$, then $\Gamma \cap \Omega$ must consist of three nonempty connected components Γ_1, Γ_2, and Γ_3. For each $i = 1, 2, 3$, let $\overline{\Gamma}_i$ denote the closure of Γ_i, and $\partial\Gamma_i$ the boundary of Γ_i, in the topology of $\mathbb{R}^3$. Then the restriction of f to $\overline{\Gamma}_i$ takes its maximum value at a single point in Γ_i, and takes its minimum value f_{min} on $\partial\Gamma_i$. For each $(x_1, x_2, x_3) \in \Gamma \cap \Omega$, we have $f(x_1, x_2, x_3) > f_{min}$.

Proof. Since $(1/3)^{2/15} < B/A < 1$, then by Lemma 4.2, there exists a solution $(y_1, y_2) = (\gamma, \delta)$ to (4.4) with $0 < \gamma < \delta$. Setting $x_1 = x_2 = \gamma$ and $x_3 = \delta$ then defines a point $Q = (\gamma, \gamma, \delta) \in \Gamma \cap \Omega$, and shows that $\Gamma \cap \Omega$ is nonempty. The fact that $\Gamma \cap \Omega$ is a smooth one-dimensional submanifold of $\mathbb{R}^3$ follows from the implicit function theorem (see, for example, Theorem 1.38 of [22]) and the fact that the gradients $\nabla g_1(x)$ and $\nabla g_2(x)$ are linearly independent at all $x \in \Gamma \cap \Omega$. Indeed, if for some $c_1, c_2 \in \mathbb{R}$, with c_1, c_2 not both zero, we have $c_1 \nabla g_1(x) + c_2 \nabla g_2(x) = 0$, then it follows easily that $x_1 = x_2 = x_3$. But then $g_1(x) = A^3$ and $g_2(x) = B^5$ imply that $B/A = (1/3)^{2/15}$, contradicting our assumption about B/A.

Now suppose we are at a point $x_0 = (x_{10}, x_{20}, x_{30}) \in \Gamma \cap \Omega$ where $x_{20} \ne x_{30}$. (Note that the point Q defined above is such a point.) Then from the implicit function theorem it follows that there exists a neighborhood I of x_0 in $\mathbb{R}$ such that for all $t \in I$, there are unique numbers $x_2(t)$ and $x_3(t)$ so that

$(t, x_2(t), x_3(t)) \in \Gamma \cap \Omega$. Moreover, $x_2(t)$ and $x_3(t)$ are smooth functions of $t \in I$, with

$$\begin{aligned} \frac{dx_3}{dt} &= \frac{t^2(x_2^2 - t^2)}{x_3^2(x_3^2 - x_2^2)} \\ \frac{dx_2}{dt} &= \frac{t^2(t^2 - x_3^2)}{x_2^2(x_3^2 - x_2^2)} \end{aligned} \tag{4.13}$$

on I.

In particular, this analysis when applied to the point Q shows that that there are functions $x_2(t)$ and $x_3(t)$ defined for t in a neighborhood I of γ such that $(t, x_2(t), x_3(t)) \in \Gamma \cap \Omega$ for all $t \in I$, and equations (4.13) hold on I. From (4.13) we have that $\frac{dx_2}{dt} < 0$ at $t = \gamma$, so there exists an $\epsilon > 0$ such that $0 < x_2(t) < t < x_3(t)$ for all t such that $\gamma < t < \gamma + \epsilon$.

Assume now further that $B/A \geq (1/2)^{2/15}$. Then by Lemma 4.2, the point (γ, δ) defined above is the only solution (y_1, y_2) to (4.4) with $y_1 > 0$ and $y_2 > 0$. Let S be the set of all $t_0 > \gamma$ such that there exist smooth functions $x_2(t)$, $x_3(t)$ defined for all $t \in (\gamma, t_0)$ such that $(t, x_2(t), x_3(t)) \in \Gamma \cap \Omega$ and

$$0 < x_2(t) < t < x_3(t)$$

for all $t \in (\gamma, t_0)$. Then S is nonempty and bounded, since $\epsilon \in S$ and $t_0 \leq A$ for all $t_0 \in S$. Therefore S has a finite supremum, which we denote by t_m. Equations (4.13) imply that $\frac{dx_3}{dt} \leq 0$ and $\frac{dx_2}{dt} \leq 0$ for all $t \in [\gamma, t_m)$, so $x_2(t)$ and $x_3(t)$ have limits as t approaches t_m from the left; we denote these limits by $x_2(t_m)$ and $x_3(t_m)$ respectively.

We have that $0 \leq x_2(t_m) \leq t_m \leq x_3(t_m)$. It cannot be the case that $0 < x_2(t_m) < t_m < x_3(t_m)$, for then an application of the implicit function theorem would allow us to extend $x_2(t)$ and $x_3(t)$ to an open interval containing $t = t_m$, contradicting the maximality of t_m. Since $x_2(t)$ is nonincreasing on $[\gamma, t_m)$ and $x_2(\gamma) = \gamma$, we have $x_2(t_m) < t_m$. Also, we cannot have $0 < x_2(t_m) < t_m = x_3(t_m)$, for then setting $\tilde{\gamma} = t_m = x_3(t_m)$ and $\tilde{\delta} = x_2(t_m)$ would produce a solution $(y_1, y_2) = (\tilde{\gamma}, \tilde{\delta})$ of (4.4) with $0 < \tilde{\delta} < \tilde{\gamma}$, which is distinct from (γ, δ) and therefore contradicts the uniqueness of positive solutions to (4.4). Finally, if $0 < x_2(t_m) = t_m = x_3(t_m)$, we would obtain a point in $\Gamma \cap \Omega$ where $x_1 = x_2 = x_3$, which we have already seen is impossible. We have thus ruled out all the possibilities in which $x_2(t_m) > 0$, so it follows that $x_2(t_m) = 0$.

We have shown that $\Gamma \cap \Omega$ contains the smooth arc

$$\{(t, x_2(t), x_3(t)) : \gamma \leq t \leq t_m\}$$

whose endpoints are Q and $P_1 = (t_m, 0, x_3(t_m)) \in \partial\Omega$. By symmetry, $\Gamma \cap \Omega$ also contains a smooth arc whose endpoints are Q and $P_2 = (0, t_m, x_3(t_m)) \in \partial\Omega$. The interior of the union of these two arcs is a connected component Γ_1 of $\Gamma \cap \Omega$.

We now consider the problem of maximizing or minimizing $f(x)$ subject to the constraint $x \in \overline{\Gamma}_1$. If the maximum or minimum occurs at an interior point $x = (x_1, x_2, x_3) \in \Gamma_1$, then x must be a critical point of the constrained variational problem, in the sense that

$$\nabla f(x) = \lambda_1 \nabla g_1(x) + \lambda_2 \nabla g_2(x)$$

for some $\lambda_1, \lambda_2 \in \mathbb{R}$. This implies that

$$7x_i^6 = 3\lambda_1 x_i^2 + 5\lambda_2 x_i^4$$

for $i = 1, 2, 3$. Letting $z_i = x_i^2$, we have for all $i, j \in \{1, 2, 3\}$ that

$$\frac{5\lambda_2}{7}(z_i - z_j) = z_i^2 - z_j^2.$$

Therefore either $z_i = z_j$ or $z_i + z + j = 5\lambda_2/7$. It follows that the set $\{z_1, z_2, z_3\}$ cannot consist of three distinct numbers: if, for example, $z_1 \neq z_3$ and $z_2 \neq z_3$, then we must have that $z_1 + z_3 = z_2 + z_3 = 5\lambda_2/7$, so $z_1 = z_2$. It follows then from Lemma 4.2 that the only possible critical points of f on $\Gamma \cap \Omega$ are $Q = (\gamma, \gamma, \delta)$, (γ, δ, γ), and (γ, δ, γ). But since $\delta > \gamma$, and $x_3 > x_2$ at all points on Γ_1 except Q, we conclude that Q is the only critical point of f on Γ_1.

We have now shown that either f takes its maximum value over $\overline{\Gamma}_1$ at Q and its minimum value at P_1 and P_2, or f takes its minimum value over $\overline{\Gamma}_1$ at Q and its maximum value at P_1 and P_2. To decide between these two alternatives, it suffices to determine whether the restriction of f to Γ_1 has a local maximum or a local minimum at Q. For this purpose, we use the second derivative test for constrained extrema, as expounded for example in [20].

Consider the Lagrangian $L(x)$ defined by

$$L(x) = f(x) - \lambda_1(g_1(x) - A^3) - \lambda_2(g_2(x) - B^5),$$

and form the "augmented Hessian", a 5×5 matrix $\mathbf{H}$ defined by

$$\mathbf{H} = \begin{bmatrix} \mathbf{0} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix},$$

where $\mathbf{B}$ is the 2×3 matrix given by

$$\mathbf{B} = \begin{bmatrix} -(g_1)_{x_1} & -(g_1)_{x_2} & -(g_1)_{x_3} \\ -(g_2)_{x_1} & -(g_2)_{x_2} & -(g_2)_{x_3} \end{bmatrix},$$

$\mathbf{C} = \mathbf{B}^T$ is the transpose of $\mathbf{B}$, and $\mathbf{D}$ is the 3×3 Hessian of L, given by

$$\mathbf{D}_{ij} = L_{x_i x_j} \quad \text{for } i, j \in \{1, 2, 3\}.$$

Here and in what follows we use $\mathbf{0}$ to denote matrices of various sizes (in this case, a 2×2 matrix) with all zero entries.

We want to compute the determinant $\det \mathbf{H}$ of $\mathbf{H}$ at $x = Q$. Calculations show that at $x = Q$, we have

$$\mathbf{B} = \begin{bmatrix} -3\gamma^2 & -3\gamma^2 & -3\delta^2 \\ -5\gamma^4 & -5\gamma^4 & -5\delta^4 \end{bmatrix}$$

and

$$\mathbf{D} = \begin{bmatrix} -14\gamma^3(\gamma^2 + \delta^2) & 0 & 0 \\ 0 & -14\gamma^3(\gamma^2 + \delta^2) & 0 \\ 0 & 0 & -14\delta^3(\gamma^2 + \delta^2) \end{bmatrix};$$

from which one finds that

$$\det(\mathbf{0} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}) = \frac{-450\gamma\delta}{196}$$

and

$$\det \mathbf{D} = -14^3(\gamma^2 - \delta^2)^3\delta^3\gamma^6.$$

Let $\mathbf{I}_2$ be the 2×2 identity matrix, and $\mathbf{I}_3$ the 3×3 identity matrix. Then the matrix

$$\begin{bmatrix} \mathbf{I}_2 & \mathbf{0} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_3 \end{bmatrix}$$

has determinant equal to one, and so we can write

$$\det \mathbf{H} = \det \begin{bmatrix} \mathbf{0} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{I}_2 & \mathbf{0} \\ -\mathbf{D}^{-1}\mathbf{C} & \mathbf{I}_3 \end{bmatrix} = \det \begin{bmatrix} \mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C} & \mathbf{B} \\ \mathbf{0} & \mathbf{D} \end{bmatrix} =$$
$$= \det(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}) \det \mathbf{D} = 14 \cdot 450\gamma^7\delta^4(\gamma^2 - \delta^2)^3.$$

Since $\gamma < \delta$, we have shown that $\det \mathbf{H} < 0$ at $x = Q$. It is easy to check that $\mathbf{B}$ has full rank at $x = Q$. Therefore, according to Theorem 36 on p. 58 of [20], we have that $\mathbf{v}^T\mathbf{D}\mathbf{v} < 0$ for all nonzero column vectors $\mathbf{v} \in \mathbb{R}^3$ satisfying $\mathbf{B}\mathbf{v} = \mathbf{0}$. In other words, the Hessian $\mathbf{D}$ of L is negative definite in all directions $\mathbf{v}$ which are tangent to both the surfaces $\{x : g_1(x) = A^3\}$ and $\{x : g_2(x) = B^5\}$ at Q. From a classical result in the calculus of variations (see for example page 334 of [17]), it follows that $f(x)$ has a local maximum at Q subject to the restriction $x \in \Gamma_1$.

We have now proved that the restriction of f takes its maximum over $\overline{\Gamma}_1$ at $Q = (\gamma, \gamma, \delta) \in \Gamma_1$ and its minimum value at the endpoints $P_1 = (t_m, 0, x_3(t_m))$ and $P_2 = (0, t_m, x_3(t_m))$ of $\overline{\Gamma}_1$. Let us define $f_{\max} = f(Q)$ and $f_{\min} = f(P_1) = f(P_2)$. Since the restriction of f to Γ_1 has no critical points in $\Gamma_1 \backslash Q$, we must have $f(x) > f_{\min}$ for all $x \in \Gamma_1$.

By symmetry, it follows that $\Gamma \cap \Omega$ also contains a component Γ_2 which includes the point (γ, δ, γ) and whose closure has endpoints $(0, x_3(t_m), t_m)$ and $(t_m, x_3(t_m), 0)$; and a component Γ_3 which includes the point (δ, γ, γ) and whose closure has endpoints $(x_3(t_m), 0, t_m)$ and $(x_3(t_m), t_m, 0)$. Furthermore, we know that the maximum value of f on $\overline{\Gamma}_2$ is attained at (γ, δ, γ), and is equal to $f_{\max}$; the minimum value of f on $\overline{\Gamma}_2$ is attained at the boundary points of $\overline{\Gamma}_2$, and is equal to $f_{\min}$; and $f(x) > f_{\min}$ for all $x \in \Gamma_2$. Similar statements hold for $\overline{\Gamma}_3$.

To complete the proof of the Lemma, it remains only to show that $\Gamma \cap \Omega$ contains no other components besides Γ_1, Γ_2, and Γ_3. To prove this, assume $Q_0 = (x_{10}, x_{20}, x_{30}) \in \Gamma \cap \Omega$; we wish to show that $Q_0 \in \Gamma_i$ for some $i \in \{1, 2, 3\}$. We know x_{10}, x_{20}, and x_{30} cannot all be equal (for this would imply $B/A = (1/3)^{2/15}$); and if any two of x_{10}, x_{20}, x_{30} are equal, then by Lemma 4.2, Q_0 must be one of the points (γ, γ, δ), (γ, δ, γ), or (δ, γ, γ), and therefore lies in one of the Γ_i. We may therefore assume without loss of generality that $x_{10} < x_{20} < x_{30}$. Then the analysis above shows that there exists some $\epsilon > 0$ and a smooth curve $x(t) = (t, x_2(t), x_3(t))$ mapping $I = (x_{10} - \epsilon, x_{10} + \epsilon)$ into

$\Gamma \cap \Omega$, such that $x(0) = Q_0$, and satisfying $t < x_2(t) < x_3(t)$ on I. Moreover, $x(t)$ satisfies equations (4.13), which imply that $\frac{dx_2}{dt} < 0$ and $\frac{dx_3}{dt} > 0$ on I.

Now let S be the set of all $t_0 > x_{10}$ such that there exist smooth functions $x_2(t)$, $x_3(t)$ defined for all $t \in (x_{10}, t_0)$ such that $(t, x_2(t), x_3(t)) \in \Gamma \cap \Omega$ and

$$0 < t < x_2(t) < x_3(t)$$

for all $t \in (x_{10}, t_0)$. Again we define $t_m = \sup S$ and let $x_2(t_m)$ denote the limit of $x_2(t)$ as t approaches t_m from the left. The implicit function theorem and the maximality of t_m imply that we must have $x_2(t_m) = t_m$. But this then implies that the point $(t_m, x_2(t_m), x_3(t_m)) = (\gamma, \gamma, \delta) \in \Gamma_1$. Therefore the set $S_1 = \{t \in [x_{01}, t_m] : (t, x_2(t), x_3(t)) \in \Gamma_1\}$ is non-empty. The uniqueness assertion in the implicit function theorem tells us that for every $t \in [x_{01}, t_m]$, the equations $g_1(x) = A^3$ and $g_2(x) = B^5$ determine x_2 and x_3 uniquely as functions of x_1 in some open neighborhood of $(t, x_2(t), x_3(t))$. Therefore S_1 is open. On the other hand, S_1 is clearly closed, by the continuity of $x_2(t)$ and $x_3(t)$ and the fact that Γ_1 is a closed subset of $\mathbb{R}^3$. So we must have $S = [x_{01}, t_m]$, and therefore $Q_0 \in \Gamma_1$. □

Remark: In the case $(1/3)^{2/15} < B/A < (1/2)^{2/15}$, a similar analysis shows that $\Gamma \cap \Omega$ is homeomorphic to a circle, and contains all six of the points $P_1(\gamma_1, \gamma_1, \delta_1)$, $P_2(\gamma_1, \delta_1, \gamma_1)$, $P_3(\delta_1, \gamma_1, \delta_1)$, $P_4(\gamma_2, \gamma_2, \delta_2)$, $P_5(\gamma_2, \delta_2, \gamma_2)$, and $P_6(\delta_2, \gamma_2, \delta_2)$, where (γ_1, δ_1) and (γ_2, δ_2) are as described in part 4 of Lemma 4.2. Moreover, points P_1, P_2, and P_3 are local maxima for the restriction of f to $\Gamma \cap \Omega$; while points P_4, P_5, and P_6 are local minima. However, we will not need these facts in what follows.

Lemma 4.7. *Suppose $x_1, \ldots, x_n$ are numbers such that $x_1 \geq x_2 \geq \cdots \geq x_n \geq 0$, with $x_1 > 0$, and for each $m \in \{1, \ldots, n\}$ define*

$$A_m = \left(\sum_{i=1}^{m} x_i^3 \right)^{1/3}$$
$$B_m = \left(\sum_{i=1}^{m} x_i^5 \right)^{1/5}.$$

Then for each $m \in \{2, \ldots, n\}$,

$$\frac{B_{m-1}}{A_{m-1}} \geq \frac{B_m}{A_m}, \tag{4.14}$$

and the inequality is strict if $x_m > 0$.

Proof. The statement is obvious if $x_m = 0$, so we may assume $x_m > 0$. Let $f(x) = \dfrac{(B_m^5 - x^5)^3}{(A_m^3 - x^3)^5}$. Then $f(x_m) = (B_{m-1}/A_{m-1})^{15}$ and $f(0) = (B_m/A_m)^{15}$, so it suffices to show that $f(x_m) > f(0)$. Now

$$f'(x) = \frac{15x^2(B_m^5 - A_m^3 x^2)(B_m^5 - x^5)^2}{(A_m^3 - x^3)^6}.$$

So $f'(x) > 0$ for $0 \leq x < x_0$, where $x_0 = \sqrt{B_m^5/A_m^3}$. But since $x_m \leq x_i$ for all $i \in \{1, \ldots, m\}$, we have

$$A_m^3 x_m^2 = \left(\sum_{i=1}^m x_i^3\right) x_m^2 \leq \sum_{i=1}^m x_i^5 = B_m^5,$$

so $x_m \leq x_0$. Therefore $f(x_n) > f(0)$, as desired. □

Lemma 4.8. *Let $A, B > 0$ be such that $(1/2)^{2/15} \leq B/A \leq 1$, and let $n \in \mathbb{N}$, with $n \geq 3$. Suppose $x_1, \ldots, x_n$ are numbers such that $x_1 \geq \cdots \geq x_n \geq 0$, with $x_3 > 0$, and*

$$\begin{aligned} \sum_{i=1}^n x_i^3 &= A^3 \\ \sum_{i=1}^n x_i^5 &= B^5. \end{aligned} \tag{4.15}$$

Then

$$\sum_{i=1}^n x_i^7 \geq m(A, B) + E, \tag{4.16}$$

where $E = E(x_1, x_2, x_3)$ is defined by

$$E(x_1, x_2, x_3) := x_1^7 + x_2^7 + x_3^7 - m((x_1^3 + x_2^3 + x_3^3)^{1/3}, (x_1^5 + x_2^5 + x_3^5)^{1/5}). \tag{4.17}$$

In particular,

$$E(x_1, x_2, x_3) > 0. \tag{4.18}$$

Proof. Let $\tilde{A} = (x_1^3 + x_2^3 + x_3^3)^{1/3}$ and $\tilde{B} = (x_1^5 + x_2^5 + x_3^5)^{1/5}$. If we define Γ and Ω as in (4.12) with A replaced by $\tilde{A}$ and B replaced by $\tilde{B}$, then since $x_1 \geq x_2 \geq x_3 > 0$, the point $x = (x_1, x_2, x_3)$ lies in $\Gamma \cap \Omega$. The inequality (4.18) thus follows from Lemma 4.6.

To prove (4.16), we use induction on n. When $n = 3$, the result is trivial. Suppose $n \geq 4$ and assume the statement of the lemma is true for $n - 1$; we wish to prove it for n.

Suppose that $x_1 \geq \cdots \geq x_n \geq 0$, with $x_3 > 0$, and that (4.15) holds. If $x_n = 0$, then we are done by the inductive hypothesis, so we may assume $x_n > 0$. Let

$$\begin{aligned} A_{n-1} &= \frac{(A^3 - x_n^3)^{1/3}}{x_n} \\ B_{n-1} &= \frac{(B^5 - x_n^5)^{1/5}}{x_n}, \end{aligned}$$

and define $y_i = x_i/x_n$ for $1 \le i \le n-1$. Then $y_1 \ge \cdots \ge y_{n-1}$, and

$$\sum_{i=1}^{n-1} y_i^3 = A_{n-1}^3$$
$$\sum_{i=1}^{n-1} y_i^5 = B_{n-1}^5.$$

From Lemma 4.7 it follows that $B_{n-1}/A_{n-1} > B/A$, and from Lemma 4.1 we have that $B_{n-1}/A_{n-1} \le 1$. Hence $(1/2)^{2/15} \le B_{n-1}/A_{n-1} \le 1$, and we may therefore apply the inductive hypothesis to the numbers $y_1 \ge y_2 \ge \cdots \ge y_{n-1} \ge 0$. There results the inequality

$$\sum_{i=1}^{n-1} y_i^7 \ge m(A_{n-1}, B_{n-1}) + E_1, \tag{4.19}$$

where

$$E_1 = y_1^7 + y_2^7 + y_3^7 - m((y_1^3 + y_2^3 + y_3^3)^{1/3}, (y_1^5 + y_2^5 + y_3^5)^{1/5}).$$

From (4.6), however, it follows that $E_1 = E/x_n^7$, so multiplying (4.19) by x_n^7, we conclude that

$$\sum_{i=1}^{n-1} x_i^7 \ge x_n^7 m(A_{n-1}, B_{n-1}) + E. \tag{4.20}$$

From Lemma 4.2 we have that there exist w_1, w_2 with $0 \le w_1 < w_2$ such that

$$w_1^3 + w_2^3 = A_{n-1}^3$$
$$w_1^5 + w_2^5 = B_{n-1}^5.$$

By definition of the function m, we have

$$w_1^7 + w_2^7 = m(A_{n-1}, B_{n-1}). \tag{4.21}$$

Letting $z_1 = x_n w_1$, $z_2 = x_n w_2$, and $z_3 = x_n$, we see that

$$z_1^3 + z_2^3 + z_3^3 = A^3$$
$$z_1^5 + z_2^5 + z_3^5 = B^5.$$

Therefore (z_1, z_2, z_3) is in the closure of the set $\Gamma \cap \Omega$ defined in Lemma 4.6. From Lemma 4.2 we see that the boundary of $\Gamma \cap \Omega$ consists exactly of the six points $(0, \alpha, \beta)$, $(0, \beta, \alpha)$, $(\alpha, 0, \beta)$, $(\beta, 0, \alpha)$, $(\alpha, \beta, 0)$, and $(\beta, \alpha, 0)$. At each of these boundary points, the function f defined in Lemma 4.6 takes the same value $\alpha^7 + \beta^7$, which by definition is equal to $m(A, B)$. Hence, by Lemma 4.6, we have that

$$f(z_1, z_2, z_3) \ge m(A, B).$$

Since $f(z_1, z_2, z_3) = x_n^7(1 + w_1^7 + w_2^7)$, using (4.21) we deduce that

$$x_n^7 + x_n^7 m(A_{n-1}, B_{n-1}) \geq m(A, B).$$

Therefore, by (4.20),

$$\sum_{i=1}^{n} x_i^7 = x_n^7 + \sum_{i=1}^{n-1} x_i^7 \geq x_n^7 + x_n^7 m(A_{n-1}, B_{n-1}) + E \geq m(A, B) + E,$$

as was desired. □

We are now ready for the main results of this section.

Lemma 4.9. *Suppose $y_1 \geq y_2 \geq 0$ and $y_1 > 0$. Let $n \in \mathbb{N}$, and suppose $x_1, \ldots, x_n$ are numbers such that $x_1 \geq \cdots \geq x_n \geq 0$, and*

$$\begin{aligned} \sum_{i=1}^{n} x_i^3 &\leq y_1^3 + y_2^3 \\ \sum_{i=1}^{n} x_i^5 &\geq y_1^5 + y_2^5. \end{aligned} \tag{4.22}$$

1. *If $n \geq 2$ and $y_2 = 0$, then $x_2 = 0$ and $x_1 = y_1$.*
2. *If $n \geq 3$ and $x_3 > 0$, then*

$$\sum_{i=1}^{n} x_i^7 \geq y_1^7 + y_2^7 + E(x_1, x_2, x_3), \tag{4.23}$$

where $E(x_1, x_2, x_3) > 0$ is as in (4.17) *and* (4.18).

Proof. Define $A = \left(y_1^3 + y_2^3\right)^{1/3}$, $B = \left(y_1^5 + y_2^5\right)^{1/5}$, $A_n = \left(\sum_{i=1}^n x_i^3\right)^{1/3}$, and $B_n = \left(\sum_{i=1}^n x_i^5\right)^{1/5}$. From Lemma 4.1, Lemma 4.7, and (4.22), we have that

$$1 = \frac{B_1}{A_1} \geq \frac{B_2}{A_2} \geq \frac{B_3}{A_3} \geq \cdots \geq \frac{B_n}{A_n} \geq \frac{B}{A} \geq \left(\frac{1}{2}\right)^{2/15}. \tag{4.24}$$

To prove part 1 of the Lemma, we simply observe that if $y_2 = 0$ then $B/A = 1$, and if $x_2 > 0$ then $B_1/A_1 > B_2/A_2$ by Lemma 4.7. Thus (4.24) immediately gives a contradiction. So if $y_2 = 0$, we must have $x_2 = 0$ and hence $x_1 = y_1$.

To prove part 2 of the Lemma, we first observe that from (4.24), Lemma 4.2, and the definition of the function m, we obtain that there exist z_1 and z_2 with $0 \leq z_1 \leq z_2$ and $z_2 > 0$ such that

$$\begin{aligned} z_1^3 + z_2^3 &= A_n^3 \\ z_1^5 + z_2^5 &= B_n^5 \end{aligned} \tag{4.25}$$

and $z_1^7 + z_2^7 = m(A_n, B_n)$.

Now if $n \geq 3$ and $x_3 > 0$, then from Lemma 4.8 it follows that

$$\sum_{i=1}^{n} x_i^7 \geq m(A_n, B_n) + E(x_1, x_2, x_3) = z_1^7 + z_2^7 + E(x_1, x_2, x_3). \tag{4.26}$$

But since

$$z_1^3 + z_2^3 \leq A^3$$
$$z_1^5 + z_2^5 \geq B^5,$$

it follows from Lemma 4.5 that $z_1^7 + z_2^7 \geq y_1^7 + y_2^7$. This, combined with (4.26), gives (4.23). □

Note that an interesting, and immediate, consequence of Lemma 4.9 is the following: among the set of all N-soliton profiles for the KdV equation, the ones which minimize E_4 subject to the constraints that E_3 and E_2 be held constant are precisely the 1-soliton and 2-soliton profiles.

Lemma 4.10. *Suppose $y_1 \geq y_2 \geq 0$ and $y_1 > 0$. Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence such that $x_1 \geq x_2 \geq x_3 \geq \cdots \geq 0$, and*

$$\begin{aligned}\sum_{i=1}^{\infty} x_i^3 &\leq y_1^3 + y_2^3 \\ \sum_{i=1}^{\infty} x_i^5 &\geq y_1^5 + y_2^5.\end{aligned} \tag{4.27}$$

1. *If $y_2 = 0$, then $x_2 = 0$ and $x_1 = y_1$.*
2. *If $x_3 > 0$, then*
$$\sum_{i=1}^{\infty} x_i^7 \geq y_1^7 + y_2^7 + E(x_1, x_2, x_3), \tag{4.28}$$
where $E(x_1, x_2, x_3) > 0$ is as in (4.17) and (4.18).

Proof. Define $A = \left(y_1^3 + y_2^3\right)^{1/3}$, $B = \left(y_1^5 + y_2^5\right)^{1/5}$, $A_\infty = \left(\sum_{i=1}^{\infty} x_i^3\right)^{1/3}$, and $B_\infty = \left(\sum_{i=1}^{\infty} x_i^5\right)^{1/5}$; and for $n \in \mathbb{N}$ define $A_n = \left(\sum_{i=1}^{n} x_i^3\right)^{1/3}$ and $B_n = \left(\sum_{i=1}^{n} x_i^5\right)^{1/5}$. Thus $\lim_{n\to\infty} A_n = A$ and $\lim_{n\to\infty} B_n = B$. From Lemmas 4.1 and 4.7 and our assumptions, we have that

$$1 \geq \frac{B_1}{A_1} \geq \frac{B_2}{A_2} \geq \cdots \geq \frac{B_n}{A_n} \geq \cdots \geq \frac{B_\infty}{A_\infty} \geq \frac{B}{A} \geq \left(\frac{1}{2}\right)^{2/15}.$$

Part 1 of the Lemma is now proved by the same argument as part 1 of Lemma (4.9).

To prove part 2 of the Lemma, we suppose $n \geq 3$ and $x_3 > 0$, and consider first the case when the first inequality in (4.27) is strict: that is, when $\sum_{i=1}^{\infty} x_i^3 <$

$y_1^3+y_2^3$. For each $i \in \mathbb{N}$ and $n \in \mathbb{N}$, define $\alpha_n = B_n/B_\infty$ and $x_{in} = x_i/\alpha_n$. Then $\lim_{n\to\infty} \alpha_n = 1$, and for each $n \in \mathbb{N}$ we have

$$\sum_{i=1}^{n} x_{in}^5 = \sum_{i=1}^{\infty} x_i^5 \geq y_1^5 + y_2^5.$$

Also,

$$\lim_{n\to\infty} \sum_{i=1}^{n} x_{in}^3 = \sum_{i=1}^{\infty} x_i^3 < y_1^3 + y_2^3,$$

so by choosing n sufficiently large we have $\sum_{i=1}^{n} x_{in}^3 \leq y_1^3 + y_2^3$. We can thus apply Lemma 4.9 to $x_{1n} \geq \cdots \geq x_{nn} \geq 0$ for all sufficiently large $n \in \mathbb{N}$, and obtain that

$$\sum_{i=1}^{n} x_{in}^7 \geq y_1^7 + y_2^7 + E(x_{1n}, x_{2n}, x_{3n}),$$

or

$$\frac{1}{\alpha_n^7} \sum_{i=1}^{n} x_i^7 \geq y_1^7 + y_2^7 + \frac{1}{\alpha_n^7} E(x_1, x_2, x_3).$$

Taking the limit as $n \to \infty$ then gives the desired result (4.28).

Next, consider the case when $\sum_{i=1}^{\infty} x_i^3 > y_1^3 + y_2^3$. Then for sufficiently large n, (4.22) holds, so by Lemma 4.9 we conclude that (4.23) holds, which immediately implies (4.28).

It remains then only to consider the case when $A_\infty = A$ and $B_\infty = B$. In this case we argue as follows. For each $n \in \mathbb{N}$, since $B_n/A_n \geq (1/2)^{2/15}$, by Lemma 4.2 we can choose $z_{1n} \geq z_{2n} \geq 0$ such that $A_n = z_{1n}^3 + z_{2n}^3$ and $B_n = z_{1n}^5 + z_{2n}^5$; we then have that $m(A_n, B_n) = z_{1n}^7 + z_{2n}^7$. From Lemma 4.9, we have that

$$\sum_{i=1}^{n} x_i^7 \geq m(A_n, B_n) + E(x_1, x_2, x_3). \tag{4.29}$$

But, by Lemma 4.6, we have $\lim\limits_{n\to\infty} m(A_n, B_n) = m(A_\infty, B_\infty) = m(A, B)$. Taking the limit on both sides of (4.29) as $n \to \infty$ then gives the desired result. □

5 Proof of Theorem 2.6

We first prove part 3 of Theorem 2.6, which is an easy consequence of the results of the preceding sections. Suppose $(a, b) \in \Sigma$, and suppose that (2.10) holds. Assume for contradiction that there exists a minimizer $u \in H^2(\mathbb{R})$ for $J(a, b)$. Then by Proposition 2.5, there must exist real numbers $D_1, D_2, \gamma_1, \gamma_2$ with $0 \leq D_1 < D_2$ such that $u = \psi_{D_1,D_2;\gamma_1,\gamma_2}$. Since $E_2(\psi) = a$ and $E_3(\psi) = b$, it follows from (2.5) that

$$\begin{aligned} 12\left(D_1^{3/2} + D_2^{3/2}\right) &= a \\ -\frac{36}{5}\left(D_1^{5/2} + D_2^{5/2}\right) &= b. \end{aligned}$$

Hence the equations (4.1) hold with $A = (a/12)^{1/3}$, $B = (-5b/36)^{1/5}$, $k = 2$, $x_1 = D_1^{1/2}$, and $x_2 = D_2^{1/2}$. Therefore, by Lemma 4.1, we must have that $B/A \geq (1/2)^{2/15}$. Further, we cannot have that $B/A = (1/2)^{2/15}$, for by part 3 of Lemma 4.2, this would imply that $x_1 = x_2$, contradicting the fact that $D_1 < D_2$. Hence $B/A > (1/2)^{2/15}$. But this means that

$$b < -\frac{ma^{5/3}}{2^{2/3}},$$

which contradicts our assumption (2.10). This then completes the proof of part 3 of the Theorem.

Turning to the proof of parts 1 and 2 of Theorem 2.6, we now suppose that $(a, b) \in \Sigma$ and either (2.8) or (2.9) holds. In particular we must have that $b < 0$.

Let $\{\phi_n\}$ be any minimizing sequence for $J(a, b)$, so that $\lim_{n\to\infty} E_2(\phi_n) = a$, $\lim_{n\to\infty} E_3(\phi_n) = b$, and $\lim_{n\to\infty} E_4(\phi_n) = J(a, b)$. (Note that minimizing sequences always exist. For example, from the definition of $J(a, b)$ it follows that we can choose $\{r_n\}$ to be any sequence in $\Lambda(a, b)$ such that $r_n \to J(a, b)$, and then take $\{\phi_n\}$ such that $E_2(\phi_n) = a$, $E_3(\phi_n) = b$, and $E_4(\phi_n) = r_n$ for each $n \in \mathbb{N}$.)

Since $\{E_2(\phi_n)\}$ converges, then $\{\phi_n\}$ is bounded in L^2. Also, since by Sobolev embedding and interpolation we have

$$\begin{aligned}\int_{\mathbb{R}} (\phi_n')^2\ dx &= 2E_3(\phi_n) + \frac{1}{3}\int_{\mathbb{R}} u^3\\ &\leq 2E_3(\phi_n) + C\|\phi_n\|_{H^{1/6}}^3 \leq 2E_3(\phi_n) + C\|\phi_n\|_{L^2}^{5/2}\|\phi_n\|_{H^1}^{1/2},\end{aligned}$$

it follows that

$$\|\phi_n\|_{H^1}^2 \leq C(1 + \|\phi_n\|_{H^1}^{1/2}),$$

which implies that $\{\phi_n\}$ is bounded in H^1. Finally, we have

$$\int_{\mathbb{R}} (\phi_n'')^2\ dx = 2E_4(\phi_n) + \int_{\mathbb{R}} \left(\frac{5}{3}uu_x^2 - \frac{5}{16}u^4\right)\ dx, \tag{5.1}$$

and since $\{\phi_n\}$ is bounded in H^1, it follows from Sobolev inequalities that the integral on the right is bounded. Since $\{E_4(\phi_n)\}$ is bounded above and $\{\int_{\mathbb{R}} (\phi_n'')^2\ dx\}$ is bounded below, it follows from (5.1) that both these sequences are in fact bounded. Therefore $\{\phi_n\}$ is bounded in H^2, and $J(a, b) > -\infty$.

Define

$$\rho_n := \phi_n^2 + (\phi_n')^2 + (\phi_n'')^2.$$

Since $\{\phi_n\}$ is bounded in H^2, then $\{\rho_n\}$ is bounded in L^1, and we can apply Corollary 3.4 to $\{\rho_n\}$.

We observe that $\{\rho_n\}$ is not a vanishing sequence in the sense of Definition 3.1. Indeed, if $\{\rho_n\}$ did vanish, then it would follow from Lemma 3.5 that $\lim_{n\to\infty} \|\phi_n\|_{L^3} = 0$, which in turn implies that

$$\liminf_{n\to\infty} E_3(\phi_n) = \frac{1}{2}\liminf_{n\to\infty}\int_{\mathbb{R}} (\phi_n')^2 \geq 0.$$

But on the other hand,

$$\lim_{n\to\infty} E_3(\phi_n) = b < 0,$$

giving a contradiction.

Since $\{\rho_n\}$ does not vanish, then from Corollary 3.4 we obtain a sequence of balls $\{B(x_n^i, r_n^i)\}_{n\in\mathbb{N}}$ for each $i \in \mathbb{N}$, satisfying properties (a) to (f).

Define η to be a smooth function on $\mathbb{R}$ such that $\eta(x) = 1$ for $|x| \leq 1/2$ and $\eta(x) = 0$ for $|x| \geq 1$; and for $R > 0$, define $\eta_R(x) = \eta(x/R)$. For each $i \in \mathbb{N}$ and $n \in \mathbb{N}$, define

$$v_n^i(x) = \phi_n(x)\eta_{r_n^i}(x - x_n^i). \tag{5.2}$$

We then have the decomposition

$$\phi_n = \sum_{i=1}^{n} v_n^i + w_n, \tag{5.3}$$

where for each $n \in \mathbb{N}$ we define

$$w_n(x) = \phi_n(x)\widetilde{\eta}_n(x),$$

with

$$\widetilde{\eta}_n(x) = 1 - \sum_{i=1}^{n} \eta_{r_n^i}(x - x_n^i).$$

For all $n \in \mathbb{N}$ and $i \in \mathbb{N}$, define

$$\begin{aligned} A_n^i &= B(x_n^i, r_n^i)\backslash B(x_n^i, r_n^i/2) \\ Z_n^i &= \mathbb{R}\backslash B(x_n^i, r_n^i/2) \\ W_n &= \mathbb{R}\backslash \cup_{i=1}^n B(x_n^i, r_n^i). \end{aligned}$$

Then

$$\begin{aligned} &\operatorname{supp} v_n^i \subseteq B(x_n^i, r_n^i) \\ &\operatorname{supp} w_n \subseteq W_n \cup \left(\cup_{i=1}^n A_n^i\right) \\ &\operatorname{supp} w_n' \subseteq \cup_{i=1}^n A_n^i. \end{aligned} \tag{5.4}$$

We will need some preliminary results on the behavior of the decomposition (5.3), which we state in the next few lemmas.

Lemma 5.1. *We have*

$$\lim_{n\to\infty} \sum_{i=1}^{n} \int_{A_n^i} \rho_n = 0. \tag{5.5}$$

Proof. For given $\epsilon > 0$, choose $N_1 \in \mathbb{N}$ so that $\displaystyle\sum_{i=N_1}^{\infty} \frac{1}{2^i} < \frac{\epsilon}{2}$. If $n \geq N_1$ then we can use part (c) of Corollary 3.4 to write

$$\sum_{i=N_1}^{n} \int_{A_n^i} \rho_n \leq \sum_{i=N_1}^{n} \sum_{j=i}^{\infty} \int_{A_j^i} \rho_j \leq \sum_{i=N_1}^{n} \frac{1}{2^i} < \frac{\epsilon}{2}.$$

But we also have from part (c) of Corollary 3.4 that for each fixed $i \in \mathbb{N}$, $\lim_{n\to\infty} \int_{A_n^i} \rho_n = 0$. Therefore, once N_1 has been chosen, we can find $N_2 \in \mathbb{N}$ so that if $n \geq N_2$ then $\sum_{i=1}^{N_1-1} \int_{A_n^i} \rho_n < \frac{\epsilon}{2}$. It follows that for $n \geq \max(N_1, N_2)$ we have $\sum_{i=1}^{n} \int_{A_n^i} \rho_n < \epsilon$, as desired. □

Lemma 5.2. *Suppose $B > 0$, and suppose $\{\phi_n\}$ is any sequence of functions in $H^2(\mathbb{R})$ satisfying $\|\phi_n\|_{H^1} \leq B$ for all $n \in \mathbb{N}$. Let v_n^i and w_n be defined for $n \in \mathbb{N}$ and $i \in \mathbb{N}$ by (5.2) and (5.3). Then there exists a constant $C > 0$ depending only on η and B (and in particular, not on k or n) such that for all $n \in \mathbb{N}$, and for $m = 2, 3, 4$,*

$$\left| E_m(\phi_n) - \sum_{i=1}^{n} E_m(v_n^i) - E_m(w_n) \right| \leq C \sum_{i=1}^{n} \int_{A_n^i} \rho_n. \tag{5.6}$$

Proof. We substitute (5.3) into $E_2(\phi_n) = \frac{1}{2}\int_{\mathbb{R}} \phi_n^2$, and expand, expressing $E_2(\phi_n)$ as a sum of integrals. Since v_n^i and v_n^j have disjoint supports for $i \neq j$, all integrals in this expression whose integrands contain two factors of v_n^i with distinct values of i will vanish. We thus obtain

$$E_2(\phi_n) = \sum_{i=1}^{n} E_2(v_n^i) + E_2(w_n) + \sum_{i=1}^{n} \int_{\mathbb{R}} v_n^i w_n.$$

Since the intersection of the supports of v_n^i and w_n is contained in A_n^i, and $|v_n^i| \leq |\phi_n|$ and $|w_n| \leq |\phi_n|$ everywhere on $\mathbb{R}$, we have

$$\int_{\mathbb{R}} |v_n^i w_n| = \int_{A_n^i} |v_n^i w_n| \leq \int_{A_n^i} |\phi_n|^2 \leq \int_{A_n^i} \rho_n. \tag{5.7}$$

This then establishes (5.6) for $m = 2$.

Substituting (5.3) into the expression for $E_3(\phi_n)$, we obtain the estimate

$$\left| E_3(\phi_n) - \sum_{i=1}^{n} E_3(v_n^i) - E_3(w_n) \right| \leq$$
$$\sum_{i=1}^{n} \left(\int_{\mathbb{R}} |{v_n^i}'||w_n'| + \int_{\mathbb{R}} |v_n^i|^2 |w_n| + \int_{\mathbb{R}} |v_n^i||w_n|^2 \right).$$

Then we write

$$\begin{aligned}
\int_{\mathbb{R}} |{v_n^i}'||w_n'| &= \int_{A_n^i} |{v_n^i}'||w_n'| \leq C \int_{A_n^i} \left(|\phi_n|^2 + |\phi_n'|^2\right) \leq C \int_{A_n^i} \rho_n, \\
\int_{\mathbb{R}} |v_n^i|^2 |w_n| &\leq \|v_n\|_{L^\infty} \int_{A_n^i} |v_n^i w_n| \leq C \|\phi_n\|_{L^\infty} \int_{A_n^i} \rho_n \leq C \int_{A_n^i} \rho_n,
\end{aligned} \tag{5.8}$$

and similarly for $\int_{\mathbb{R}} |v_n^i||w_n|^2$. (Here and in what follows we use C to stand for various constants which depend only on η and B.) This establishes (5.6) for $m = 3$.

Finally, to prove (5.6) for $m = 4$, we substitute (5.3) into $E_4(\phi_n)$ and write

$$\left| E_4(\phi_n) - \sum_{i=1}^{n} E_4(v_n^i) - E_4(w_n) \right| \leq$$
$$C \int_{\mathbb{R}} \sum_{i=1}^{n} \left(|v_n^i{}'' w_n''| + |v_n^i v_n^i{}' w_n'| + |v_n^i||w_n'|^2 + |v_n^i{}'|^2 |w_n| + \sum_{\substack{s,t \geq 1 \\ s+t \leq 4}} |v_n^i|^s |w_n|^t \right). \tag{5.9}$$

All the terms on the right side of (5.9) can be estimated like the terms in the preceding paragraphs. For example, we have

$$\int_{\mathbb{R}} |v_n^i|^3 |w_n| \leq \|v_n\|_{L^\infty}^2 \int_{A_n^i} |v_n^i w_n| \leq C\|\phi_n\|_{L^\infty}^2 \int_{A_n^i} \rho_n \leq C\|\phi_n\|_{H^1}^2 \int_{A_n^i} \rho_n$$
$$\int_{\mathbb{R}} |v_n^i{}'|^2 |w_n| \leq \|w_n\|_{L^\infty} \int_{A_n^i} |v_n^i{}'|^2 \leq C\|\phi_n\|_{L^\infty} \int_{A_n^i} \rho_n \leq C\|\phi_n\|_{H^1} \int_{A_n^i} \rho_n.$$

Clearly, similar estimates hold for the remaining integrals; we omit the details. □

Lemma 5.3. *In addition to the assumptions of Lemma 5.2, assume that $f \in H^2$ with $\|f\|_{H^1} \leq B$. For each $n \in \mathbb{N}$ and each $i \in \{1, \ldots, n\}$, define*

$$\tilde{\phi}_n = f + \sum_{\substack{j=1 \\ j \neq i}}^{n} v_n^j + w_n.$$

Then there exists a constant $C > 0$, depending only on η and B, such that

$$\left| E_m(\tilde{\phi}_n) - E_m(f) - \sum_{\substack{j=1 \\ j \neq i}}^{n} E_m(v_n^j) - E_m(w_n) \right| \leq C \sum_{\substack{j=1 \\ j \neq i}}^{n} \int_{A_n^j} \rho_n + C\|f\|_{H^{m-2}(Z_n^i)}, \tag{5.10}$$

for $m = 2$, 3, and 4.

Proof. Proceeding as in the proof of Lemma 5.2, we write

$$\left| E_2(\tilde{\phi}_n) - E_2(f) - \sum_{\substack{j=1 \\ j \neq i}}^{n} E_2(v_n^j) - E_2(w_n) \right|$$
$$\leq C \int_{\mathbb{R}} |f w_n| + C \sum_{\substack{j=1 \\ j \neq i}}^{n} \left(\int_{\mathbb{R}} |f v_n^j| + \int_{\mathbb{R}} |v_n^j w_n| \right)$$

The sum $\sum_{\substack{j=1\\ j\neq i}}^{n} \int_{\mathbb{R}} |v_n^j w_n|$ is estimated as in (5.7). Also, since the support of w_n is contained in Z_n^i, and the same is true of the support of v_n^j whenever $j \neq i$, then

$$\int_{\mathbb{R}} |f w_n| \leq \|f\|_{L^2(Z_n^i)} \|w_n\|_{L^2} \leq C\|f\|_{L^2(Z_n^i)} \|\phi_n\|_{L^2} \leq C\|f\|_{L^2(Z_n^i)}.$$

Finally, since the supports of the functions $\{v_n^j\}_{j\in\mathbb{N}}$ are mutually disjoint, and for $j \neq i$ are all contained in Z_n^i, there exists C depending only on η and B (and not on k, n or i) such that

$$\sum_{\substack{j=1\\ j\neq i}}^{n} \int_{\mathbb{R}} |f v_n^j| \leq C \int_{Z_n^i} |f \phi_n| \leq C\|f\|_{L^2(Z_n^i)}.$$

This proves (5.10) for $m = 2$.

Similarly, (5.10) is proved for $m = 3$ by expanding $E_3(\phi_n)$ as in Lemma 5.2, then using estimates such as (5.8), together with the estimates

$$\int_{\mathbb{R}} |f'| \left(\sum_{j\neq i} |{v_n^j}'| + |w_n'| \right) \leq C\|f'\|_{L^2(Z_n^i)} \|\phi_n\|_{H^1} \leq C\|f\|_{H^1(Z_n^i)} \|\phi_n\|_{H^1},$$

$$\int_{\mathbb{R}} |f|^2 \left(\sum_{j\neq i} |v_n^j| + |w_n| \right) \leq \|f\|_{L^\infty} \|f\|_{L^2(Z_n^i)} \|\phi_n\|_{L^2} \leq C\|f\|_{L^2(Z_n^i)},$$

and a similar estimate for $\int_{\mathbb{R}} |f|(\sum_{j\neq i} |v_n^j|^2 + |w_n|^2)$.

Finally, (5.10) is proved for $m = 4$ by expanding $E_4(\phi_n)$ to obtain an expression similar to (5.9), but with additional terms on the right-hand side of the form

$$\begin{aligned} &\sum_{j\neq i} \int_{\mathbb{R}} \left(|f''||{v_n^j}''| + |f||{v_n^j}'|^2 + |f|^2|{v_n^j}'| + |f||f'||{v_n^j}'| \right) + \\ &+ \int_{\mathbb{R}} \left(|f''||w_n''| + |f||w_n'|^2 + |f|^2|w_n'| + |f||f'||w_n'| \right) \\ &+ \sum_{\substack{s,t\geq 1\\ s+t\leq 4}} \int_{\mathbb{R}} |f|^s \left(\sum_{j\neq i} |v_n^j| + |w_n| \right)^t . \end{aligned}$$

These can each be estimated by the terms on the right-hand side of (5.10). For example, we have

$$\int_{\mathbb{R}} |f|^3 |w_n| \leq \|f\|_{L^\infty} \|w_n\|_{L^\infty} \int_{Z_n^i} |f|^2 \leq C\|f\|_{H^1}^2 \|w_n\|_{H^1} \|f\|_{L^2(Z_n^i)}$$

and

$$\int_{\mathbb{R}} |f||f'||w_n'| \leq \|f\|_{L^\infty}\|w_n'\|_{L^2}\left(\int_{Z_n^i}|f'|^2\right)^{1/2} \leq C\|f\|_{H^1}\|w_n\|_{H^1}\|f\|_{H^1(Z_n^i)}.$$

The remaining terms are estimated similarly. We omit the details, which are straightforward. □

Lemma 5.4. *Suppose* $2 < p \leq \infty$. *Then* $\lim_{n\to\infty}\|w_n\|_{L^p(\mathbb{R})} = 0$.

Proof. For each $n \in \mathbb{N}$, we have from (5.4) that $w_n = w_n\left(\chi_{W_n} + \sum_{i=1}^n \chi_{A_n^i}\right)$ and $w_n' = w_n'\left(\chi_{W_n} + \sum_{i=1}^n \chi_{A_n^i}\right)$; and from the definition of w_n we see that there exists a constant $C > 0$ such that for every $n \in \mathbb{N}$ and $x \in \mathbb{R}$, $w_n^2(x) \leq C\phi_n^2(x) \leq C\rho_n(x)$, and $(w_n')^2(x) \leq C\left(\phi_n^2(x) + (\phi_n'(x))^2\right) \leq C\rho_n(x)$. Therefore, we can write

$$\begin{aligned}
&\sup_{y\in\mathbb{R}}\int_{B(y,1)}\left(w_n^2 + (w_n')^2\right) \\
&\qquad = \sup_{y\in\mathbb{R}}\left[\int_{B(y,1)}\left(w_n^2 + (w_n')^2\right)\chi_{W_n} + \sum_{i=1}^n\int_{B(y,1)}\left(w_n^2 + (w_n')^2\right)\chi_{A_n^i}\right] \\
&\qquad \leq C\sup_{y\in\mathbb{R}}\int_{B(y,1)}\rho_n\chi_{W_n} + C\sum_{i=1}^n\int_{A_n^i}\rho_n.
\end{aligned} \tag{5.11}$$

But

$$\lim_{n\to\infty}\sup_{y\in\mathbb{R}}\int_{B(y,1)}\rho_n\chi_{W_n} = 0, \tag{5.12}$$

by part (e) of Corollary 3.4. Combining Lemma 5.1, (5.11), and (5.12) gives

$$\lim_{n\to\infty}\sup_{y\in\mathbb{R}}\int_{B(y,1)}\left(w_n^2 + (w_n')^2\right)\,dx = 0.$$

Since $\{w_n\}$, like $\{\phi_n\}$, is a bounded sequence in $H^1(\mathbb{R})$, the proof is then completed by applying Lemma 3.5. □

Lemma 5.5. *We have*

$$\lim_{N\to\infty}\limsup_{n\to\infty}\sum_{i=N}^n\int_{\mathbb{R}}|v_n^i|^3 = 0 \tag{5.13}$$

and

$$\lim_{N\to\infty}\limsup_{n\to\infty}\sum_{i=N}^n\int_{\mathbb{R}}\left[\left|v_n^i\left({v_n^i}'\right)^2\right| + |v_n^i|^4\right] = 0. \tag{5.14}$$

Proof. For a fixed value of $n \in \mathbb{N}$, the supports of the functions $\{v_n^i\}_{i=1,\dots,n}$ are mutually disjoint. Therefore, if we define

$$f_{N,n} = \sum_{i=N}^{n} v_n^i,$$

we can write, for all N, n such that $N < n$,

$$\sum_{i=N}^{n} \int_{\mathbb{R}} \left|v_n^i\right|^3 = \int_{\mathbb{R}} |f_{N,n}|^3 .$$

Now

$$\begin{aligned}\int_{\mathbb{R}} |f_{N,n}|^3 &\leq C\|f_{N,n}\|^2_{H^1(\mathbb{R})} \left(\sup_{y\in\mathbb{R}} \int_{B(y,1)} \left(|f'_{N,n}|^2 + |f_{N,n}|^2\right)\right)^{1/2} \\ &\leq C \left(\sup_{y\in\mathbb{R}} \int_{B(y,1)} \rho_n \chi_{\mathbb{R}\setminus\cup_{i=1}^N B(x_n^i, r_n^i)}\right)^{1/2} = C\left(q_n^N(1)\right)^{1/2},\end{aligned} \tag{5.15}$$

where $q_n^N(r)$ is the function defined in part (f) of Corollary 3.4. (In obtaining (5.15), we used Lemma 3.5 along with the facts that the support of $f_{N,n}$ lies outside $\cup_{i=1}^N B(x_n^i, r_n^i)$, and that $|f_{N,n}|^2$ and $|f'_{N,n}|^2$ are majorized pointwise by $C\rho_n$, where C depends only on the cutoff function η.) Since $q_n^N(r)$ is an increasing function of r, it follows from part (f) of Corollary 3.4 that

$$\lim_{N\to\infty} \limsup_{n\to\infty} q_n^N(1) = 0. \tag{5.16}$$

Therefore (5.13) follows from (5.15).

Similarly, we can use Lemma 3.5 to write

$$\begin{aligned}\sum_{i=N}^{n} \int_{\mathbb{R}} |v_n^i|^4 = \int_{\mathbb{R}} (f_{N,n})^4 &\leq C\|f_{N,n}\|^2_{H^1(\mathbb{R})} \left(\sup_{y\in\mathbb{R}} \int_{B(y,1)} \left(|f'_{N,n}|^2 + |f_{N,n}|^2\right)\right) \\ &\leq Cq_n^N(1),\end{aligned} \tag{5.17}$$

and by Hölder's inequality, Sobolev embedding, and (5.15), we have

$$\begin{aligned}\sum_{i=N}^{n} \int_{\mathbb{R}} \left|v_n^i \left({v_n^i}'\right)^2\right| = \int_{\mathbb{R}} |f_{N,n}(f'_{N,n})^2| &\leq \left(\int_{\mathbb{R}} |f'_{N,n}|^{8/3}\right)^{3/4} \left(\int_{\mathbb{R}} |f_{N,n}|^3\right)^{1/3} \\ &\leq C\|f_{N,n}\|^{8/3}_{H^1(\mathbb{R})} \left(q_n^N(1)\right)^{1/6}.\end{aligned} \tag{5.18}$$

Estimates (5.17) and (5.18) together with (5.16) then imply (5.14). □

Fix $i \in \mathbb{N}$, and for $n \in \mathbb{N}$ define $\theta_n^i(x) = v_n^i(x + x_n^i)$. Since $\{\theta_n^i\}_{n\in\mathbb{N}}$ is a bounded sequence in $H^2(\mathbb{R})$, then by passing to a subsequence, we may assume that it converges weakly in $H^2(\mathbb{R})$ to some function $g_i \in H^2(\mathbb{R})$. By a diagonalization argument, and replacing $\{\phi_n\}_{n\in\mathbb{N}}$ by an appropriate subsequence, we may assume that for every $i \in \mathbb{N}$, the sequence $\{\theta_n^i\}_{n\in\mathbb{N}}$ converges weakly in $H^2(\mathbb{R})$ to g_i. (In what follows we will often replace sequences by subsequences without changing notation.) Further, by again passing to appropriate subsequences, we may assume that the sequences $\{E_2(\theta_n^i)\}_{n\in\mathbb{N}}$ and $\{E_3(\theta_n^i)\}_{n\in\mathbb{N}}$ converge. Define

$$\begin{aligned} a_i &= \lim_{n\to\infty} E_2(\theta_n^i) = \lim_{n\to\infty} E_2(v_n^i) \\ b_i &= \lim_{n\to\infty} E_3(\theta_n^i) = \lim_{n\to\infty} E_3(v_n^i). \end{aligned} \tag{5.19}$$

Lemma 5.6. *For each $i \in \mathbb{N}$, $\{\theta_n^i\}_{n\in\mathbb{N}}$ converges strongly to g_i in $H^1(\mathbb{R})$.*

Proof. First note that by part (d) of Corollary 3.4, for every $\epsilon > 0$ there exists $R_\epsilon > 0$ such that

$$\int_{B(x_n^i, r_n^i)\setminus B(x_n^i, R_\epsilon)} \rho_n < \epsilon$$

for all $n \in \mathbb{N}$. By taking R_ϵ larger if necessary, we may assume as well that

$$\int_{\mathbb{R}\setminus B(0,R_\epsilon)} \left((g_i'')^2 + (g_i')^2 + g_i^2\right) < \epsilon.$$

From the definition of θ_n^i, it is easy to see that there exists a constant C such that for all n,

$$\|\theta_n^i\|^2_{H^2(\mathbb{R}\setminus B(0,R_\epsilon))} \le C \int_{B(x_n^i, r_n^i)\setminus B(x_n^i, R_\epsilon)} \rho_n,$$

and therefore

$$\|\theta_n^i - g_i\|_{H^2(\mathbb{R}\setminus B(0,R_\epsilon))} \le \|\theta_n^i\|_{H^2(\mathbb{R}\setminus B(0,R_\epsilon))} + \|g_i\|_{H^2(\mathbb{R}\setminus B(0,R_\epsilon))} < 2\epsilon.$$

On the other hand, since the inclusion of $H^2(B(0,R_\epsilon))$ into $H^1(B(0,R_\epsilon))$ is compact, then $\{\theta_n^i\}_{n\in\mathbb{N}}$ has a subsequence $\{\theta_{n_k}^i\}_{k\in\mathbb{N}}$ that converges strongly to g_i in $H^1(B(0,R_\epsilon))$. Then for all sufficiently large k, $\|\theta_{n_k}^i - g_i\|_{H^1(B(0,R_\epsilon))} < \epsilon$, and therefore $\|\theta_{n_k}^i - g_i\|_{H^1(\mathbb{R})} < 3\epsilon$.

It follows from the preceding that for every $\epsilon > 0$, there exists a subsequence $\{\theta_{n_k}^i\}_{k\in\mathbb{N}}$ of $\{\theta_n^i\}_{n\in\mathbb{N}}$ such that $\|\theta_{n_k}^i - g_i\|_{H^1(\mathbb{R})} < \epsilon$ for all $k \in \mathbb{N}$. By now taking a sequence of values of ϵ tending to zero and using a diagonalization argument, we obtain a subsequence of $\{\theta_n^i\}_{n\in\mathbb{N}}$ which converges to g_i strongly in $H^1(\mathbb{R})$. Since the same argument shows that every subsequence of $\{\theta_n^i\}_{n\in\mathbb{N}}$ has a subsubsequence which converges to g_i in $H^1(\mathbb{R})$, it follows that $\{\theta_n^i\}_{n\in\mathbb{N}}$ itself converges to g_i in $H^1(\mathbb{R})$. ∎

Lemma 5.7. *For each $i \in \mathbb{N}$, we have*

$$\lim_{n\to\infty} \int_{\mathbb{R}} |v_n^i|^p = \int_{\mathbb{R}} |g_i|^p \tag{5.20}$$

for all p such that $2 \le p \le \infty$, and

$$\lim_{n\to\infty} \int_{\mathbb{R}} v_n^i \left({v_n^i}' \right)^2 = \int_{\mathbb{R}} g_i (g_i')^2. \tag{5.21}$$

In particular,

$$\begin{aligned} \lim_{n\to\infty} E_2(v_n^i) &= E_2(g_i) = a_i \\ \lim_{n\to\infty} E_3(v_n^i) &= E_3(g_i) = b_i. \end{aligned} \tag{5.22}$$

Proof. Equation (5.20) follows from Lemma 5.6 and the fact that, by standard Sobolev embedding theorems, L^p embeds continuously in $H^1(\mathbb{R})$ when $2 \le p \le \infty$. For (5.21), we can write

$$\int_{\mathbb{R}} v_n^i ({v_n^i}')^2 - \int_{\mathbb{R}} g_i (g_i')^2 = \int_{\mathbb{R}} (v_n^i - g_i) \left({v_n^i}' \right)^2 + \int_{\mathbb{R}} g_i ({v_n^i}' - g_i')({v_n^i}' + g_i').$$

Using Hölder's inequality and Sobolev embedding, we can majorize the integrals on the right-hand side by

$$\begin{aligned} &\|v_n^i - g_i\|_{L^\infty(\mathbb{R})} \|v_n^i\|_{H^1(\mathbb{R})}^2 + \|g_i\|_{L^\infty(\mathbb{R})} \|v_n^i - g_i\|_{H^1(\mathbb{R})} \left(\|v_n^i\|_{H^1(\mathbb{R})} + \|g^i\|_{H^1(\mathbb{R})} \right) \\ &\quad \le C \|v_n^i - g_i\|_{H^1(\mathbb{R})} \left(\|v_n^i\|_{H^1(\mathbb{R})}^2 + \|g^i\|_{H^1(\mathbb{R})}^2 \right). \end{aligned}$$

Since $g_i \in H^1(\mathbb{R})$, and $\{v_n^i\}_{n\in\mathbb{N}}$ converges to g_i in $H^1(\mathbb{R})$, the preceding expression has limit zero as $n \to \infty$, proving (5.21). Finally, (5.22) follows immediately from Lemma 5.6 and (5.20). □

Lemma 5.8. *For each $i \in \mathbb{N}$,*

$$E_4(g_i) \le \liminf_{n\to\infty} E_4(\theta_n^i) = \liminf_{n\to\infty} E_4(v_n^i). \tag{5.23}$$

Proof. Choose a subsequence $\{\theta_{n_k}^i\}_{k\in\mathbb{N}}$ of $\{\theta_n^i\}_{n\in\mathbb{N}}$ such that $\lim_{k\to\infty} E_4(\theta_{n_k}^i) = \liminf_{n\to\infty} E_4(\theta_n^i)$. By the weak compactness of bounded sets in Hilbert space, we can find a further subsequence, also denoted by $\{\theta_{n_k}^i\}$, which converges weakly in $H^2(\mathbb{R})$ to g_i. By the lower semicontinuity of the norm in Hilbert space, we have that

$$\|g_i\|_{H^2} \le \liminf_{k\to\infty} \|\theta_{n_k}^i\|_{H^2}$$

Since $\{\theta_{n_k}^i\}_{k\in\mathbb{N}}$ converges strongly in $H^1(\mathbb{R})$ to g_i, this implies that

$$\|g_i''\|_{L^2} \le \liminf_{k\to\infty} \|(\theta_{n_k}^i)''\|_{L^2}.$$

On the other hand, by Lemma 5.7, we have that

$$\int_{\mathbb{R}} \left(-\frac{5}{6} g_i g_{ix}^2 + \frac{5}{32} g_i^4 \right) = \lim_{k\to\infty} \int_{\mathbb{R}} \left(-\frac{5}{6} \theta_{n_k}^i (\theta_{n_k}^i)_x^2 + \frac{5}{32} (\theta_{n_k}^i)^4 \right).$$

Combining the last two statements, we obtain (5.23). □

For each $i \in \mathbb{N}$, if $g_i \equiv 0$, then obviously $a_i = b_i = 0$. If on the other hand g_i is not identically zero, then by Proposition 2.3, $(a_i, b_i) \in \Sigma$, and so $J(a_i, b_i)$ is well-defined by (2.7). In that case, we have:

Lemma 5.9. *For each $i \in \mathbb{N}$, if g_i is not identically zero, then g_i is a minimizer for $J(a_i, b_i)$.*

Proof. We prove the lemma by contradiction. If g_i is not a minimizer for $J(a_i, b_i)$, then there must exist a function $h \in H^2$ such that $E_2(h) = a_i$, $E_3(h) = b_i$, and $E_4(h) < E_4(g_i)$. Define, for $n \in \mathbb{N}$,

$$h_n(x) = h(x - x_n^i)$$

and

$$\tilde{\phi}_n = h_n + \sum_{\substack{j=1 \\ j \neq i}}^{n} v_n^j + w_n.$$

To obtain the desired contradiction, we will show that

$$\lim_{n\to\infty} E_2(\tilde{\phi}_n) = a$$
$$\lim_{n\to\infty} E_3(\tilde{\phi}_n) = b$$
$$\liminf_{n\to\infty} E_4(\tilde{\phi}_n) < J(a, b).$$

To begin with, use the triangle inequality to write

$$\left|E_2(\phi_n) - E_2(\tilde{\phi}_n)\right| \leq \left|E_2(\phi_n) - \sum_{j=1}^{n} E_2(v_n^j) - E_2(w_n)\right| +$$
$$+ \left|E_2(\tilde{\phi}_n) - E_2(h_n) - \sum_{\substack{j=1 \\ j \neq i}}^{n} E_2(v_n^j) - E_2(w_n)\right| + \left|E_2(v_n^i) - E_2(h_n)\right|.$$

We can use Lemma 5.2 and Lemma 5.3 with $f = h_n$ to estimate the first two terms on the right-hand side of the preceding inequality, and thus obtain that

$$\left|E_2(\phi_n) - E_2(\tilde{\phi}_n)\right| \leq |E_2(v_n^i) - E_2(h_n)| + C \sum_{j=1}^{n} \int_{A_n^j} \rho_n + C\|h_n\|_{L^2(Z_n^i)}.$$

Since

$$\|h_n\|_{L^2(Z_n^i)} = \left(\int_{\mathbf{R} \setminus B(0, r_n^i/2)} h^2(x)\, dx \right)^{1/2},$$

and $h \in L^2(\mathbb{R})$ and $\lim_{n\to\infty} r_n^i = \infty$, it follows that $\lim_{n\to\infty} \|h_n\|_{L^2(Z_n^i)} = 0$. Finally, we have that

$$\lim_{n\to\infty} (E_2(v_n^i) - E_2(h_n)) = \lim_{n\to\infty} (E_2(v_n^i) - E_2(h)) = \lim_{n\to\infty} (E_2(v_n^i) - a_i) = 0.$$

Combining these results, we obtain that $\lim_{n\to\infty}(E_2(\phi_n)-E_2(\tilde\phi_n))=0$, so $\lim_{n\to\infty}E_2(\tilde\phi_n)=a$.

Similar arguments apply to $E_3(\phi_n)-E_3(\tilde\phi_n)$ and $E_4(\phi_n)-E_4(\tilde\phi_n)$. From Lemmas 5.2 and 5.3 we obtain that

$$|E_3(\phi_n)-E_3(\tilde\phi_n)|\leq|E_3(v_n^i)-E_3(h_n)|+C\sum_{j=1}^n\int_{A_n^j}\rho_n+C\|h_n\|_{H^1(Z_n^i)} \quad (5.24)$$

and

$$E_4(\phi_n)-E_4(\tilde\phi_n)\geq E_4(v_n^i)-E_4(h_n)-C\sum_{j=1}^n\int_{A_n^j}\rho_n-C\|h_n\|_{H^2(Z_n^i)}. \quad (5.25)$$

The same considerations as in the preceding paragraph show that it follows from (5.24) that $\lim_{n\to\infty}E_3(\tilde\phi_n)=b$. Also, from (5.25) and (5.23) we obtain that

$$\liminf_{n\to\infty}\,[E_4(\phi_n)-E_4(\tilde\phi_n)]\geq\lim_{n\to\infty}\,[E_4(v_n^i)-E_4(h_n)]\geq E_4(g_i)-E_4(h)>0,$$

and hence

$$\limsup_{n\to\infty}E_4(\tilde\phi_n)<\lim_{n\to\infty}E_4(\phi_n)=J(a,b).$$

In particular, it follows that there exists some sufficiently large n for which $E_4(\tilde\phi_n)<J(a,b)$. But since $E_2(\tilde\phi_n)=a$ and $E_3(\tilde\phi_n)=b$, this contradicts the definition of $J(a,b)$. □

From Proposition 2.5 and Lemma 5.9, we conclude that for each $i\in\mathbb{N}$, there exist $D_{1i},D_{2i},\gamma_{1i},\gamma_{2i}\in\mathbb{R}$ with $0\leq D_{1i}\leq D_{2i}$ such that

$$g_i(x)=\psi_{D_{1i},D_{2i};\gamma_{1i},\gamma_{2i}}(x). \quad (5.26)$$

Here we follow the conventions that if $D_{1i}=0$, then $\psi_{D_{1i},D_{2i};\gamma_{1i},\gamma_{2i}}=\psi_{D_{2i};\gamma_{1i}}$; and if $D_{1i}=D_{2i}=0$, then $\psi_{D_{1i},D_{2i}}\equiv 0$. Also, in what follows we will occasionally omit the subscripts γ_{1i} and γ_{2i}, referring to g_i simply as $\psi_{D_{1i},D_{2i}}$.

Lemma 5.10. *For the numbers D_{1i} and D_{2i} defined for $i\in\mathbb{N}$ by* (5.26), *we have*

$$\begin{aligned}12\sum_{i=1}^\infty\left(D_{1i}^{3/2}+D_{2i}^{3/2}\right)&\leq a\\ \frac{36}{5}\sum_{i=1}^\infty\left(D_{1i}^{5/2}+D_{2i}^{5/2}\right)&\geq -b\\ \frac{36}{7}\sum_{i=1}^\infty\left(D_{1i}^{7/2}+D_{2i}^{7/2}\right)&\leq J(a,b).\end{aligned} \quad (5.27)$$

Proof. For $m = 2, 3, 4$, if we define ϵ_{mn} for $n \in \mathbb{N}$ by

$$\epsilon_{mn} = E_m(\phi_n) - \sum_{i=1}^{n} E_m(v_n^i) - E_m(w_n),$$

then we have from Lemmas 5.2 and 5.1 that $\lim_{n\to\infty} \epsilon_{mn} = 0$.

In case $m = 2$, we have $E_2(f) \geq 0$ for all $f \in L^2$. Therefore we have, for all $N, n \in \mathbb{N}$ such that $n > N$,

$$\sum_{i=1}^{N} E_2(v_n^i) = E_2(\phi_n) - \sum_{i=N+1}^{n} E_2(v_n^i) - E_2(w_n) - \epsilon_{2n} \leq E_2(\phi_n) - \epsilon_{2n}.$$

Holding N fixed and taking the limit on both sides as $n \to \infty$, and recalling (5.19), we obtain that

$$\sum_{i=1}^{N} E_2(g_i) = 12 \sum_{i=1}^{N} \left(D_{1i}^{3/2} + D_{2i}^{3/2} \right) \leq a.$$

Then taking the limit as $N \to \infty$ yields the first inequality in (5.27).

Next, we consider the case $m = 3$. We have, for all $N, n \in \mathbb{N}$ such that $n > N$,

$$\begin{aligned}
\sum_{i=1}^{N} E_3(v_n^i) &= E_3(\phi_n) - \sum_{i=N+1}^{n} E_3(v_n^i) - E_3(w_n) - \epsilon_{3n} \\
&= E_3(\phi_n) - \sum_{i=N+1}^{n} \int_{\mathbb{R}} \left(\frac{1}{2}({v_n^i}')^2 - \frac{1}{6}(v_n^i)^3 \right) - \int_{\mathbb{R}} \left(\frac{1}{2}({w_n}')^2 - \frac{1}{6} w_n^3 \right) - \epsilon_{3n} \\
&\leq E_3(\phi_n) + \frac{1}{6} \sum_{i=N+1}^{n} \int_{\mathbb{R}} (v_n^i)^3 + \frac{1}{6} \int_{\mathbb{R}} w_n^3 - \epsilon_{3n}.
\end{aligned} \tag{5.28}$$

Let $\epsilon > 0$ be given. From (5.13) it follows that there exists $N \in \mathbb{N}$ such that

$$\limsup_{n\to\infty} \sum_{i=N+1}^{n} \int_{\mathbb{R}} \left| v_n^i \right|^3 < \epsilon.$$

For this fixed value of N, by taking the limit as n goes to infinity of both sides of (5.28) and using Lemma 5.4, we obtain that

$$-\frac{36}{5} \sum_{i=1}^{N} \left(D_{1i}^{5/2} + D_{2i}^{5/2} \right) = \sum_{i=1}^{N} b_i \leq b + \frac{\epsilon}{6},$$

and hence

$$\frac{36}{5} \sum_{i=1}^{\infty} \left(D_{1i}^{5/2} + D_{2i}^{5/2} \right) \geq \frac{36}{5} \sum_{i=1}^{N} \left(D_{1i}^{5/2} + D_{2i}^{5/2} \right) \geq -b - \frac{\epsilon}{6}.$$

Since this inequality holds for all $\epsilon > 0$, we have proved the second inequality in (5.27).

In case $m = 4$, we have

$$\begin{aligned}\sum_{i=1}^{N} E_4(v_n^i) \le E_4(\phi_n) + \sum_{i=N+1}^{n} \int_{\mathbb{R}} \left[\frac{5}{6}\left|v_n^i({v_n^i}')^2\right| + \frac{5}{32}\left|v_n^i\right|^4\right] \\ + \frac{5}{6}\int_{\mathbb{R}} w_n(w_n')^2 - \frac{5}{32}\int_{\mathbb{R}} w_n^4 - \epsilon_{4n},\end{aligned} \tag{5.29}$$

for all $N, n \in \mathbb{N}$ such that $n > N$. Using Lemma 5.4 we see that $\lim_{n\to\infty} \int_{\mathbb{R}} w_n^4 = 0$ and

$$\lim_{n\to\infty} \left|\int_{\mathbb{R}} w_n(w_n')^2\right| \le \lim_{n\to\infty} \|w_n\|_{L^\infty} \|w_n\|_{H^1}^2 = 0.$$

For given $\epsilon > 0$, by (5.14), we can choose $N_0 \in \mathbb{N}$ such that for all $N \ge N_0$,

$$\limsup_{n\to\infty} \sum_{i=N+1}^{n} \int_{\mathbb{R}} \left[\frac{5}{6}\left|v_n^i({v_n^i}')^2\right| + \frac{5}{32}\left|v_n^i\right|^4\right] < \epsilon.$$

For each fixed value of $N \ge N_0$, taking the limit on both sides of (5.29) as n goes to infinity and using (5.23), we then obtain

$$\frac{36}{7}\sum_{i=1}^{N}\left(D_{1i}^{7/2} + D_{2i}^{7/2}\right) = \sum_{i=1}^{N} E_4(g_i) \le J(a,b) + \epsilon.$$

Since this is true for all $N \ge N_0$, it follows that

$$\frac{36}{7}\sum_{i=1}^{\infty}\left(D_{1i}^{7/2} + D_{2i}^{7/2}\right) \le J(a,b) + \epsilon,$$

and since $\epsilon > 0$ was arbitrary, this proves the final inequality in (5.27). □

By Lemma 5.10, only finitely many of the numbers D_{1i} and D_{2i} can be greater than any fixed positive number. Therefore it is possible to re-order the numbers in the sequence

$$\left(D_{11}^{1/2}, D_{21}^{1/2}, D_{12}^{1/2}, D_{22}^{1/2}, D_{13}^{1/2}, D_{23}^{1/2} \dots\right) \tag{5.30}$$

so that they form a non-increasing sequence, whose terms we denote by $\{x_n\}$, with $x_1 \ge x_2 \ge x_3 \ge \dots$.

Proof of part 1 of Theorem 2.6. Suppose that (2.8) holds. We let $C = (a/12)^{2/3} = (-5b/36)^{2/5} > 0$. For every $\gamma \in \mathbb{R}$, we have

$$\begin{aligned}E_2(\psi_{C,\gamma}) &= 12C^{3/2} = a, \\ E_3(\psi_{C,\gamma}) &= -(36/5)C^{5/2} = b.\end{aligned} \tag{5.31}$$

From the definition of $J(a,b)$ it therefore follows that

$$J(a,b) \leq E_4(\psi_{C,\gamma}) = E_4(C) = (36/7)C^{7/2}. \tag{5.32}$$

Let $y_1 = C$ and $y_2 = 0$. From (5.27) and (5.31), it follows that the inequalities (4.22) are satisfied by the numbers $\{x_n\}$ defined after (5.30). Therefore, by part 1 of Lemma 4.10, we must have $x_2 = 0$ and $x_1 = y_1$. Thus $g_1 = \psi_{C,\gamma}$ for some $\gamma \in \mathbb{R}$, and $g_i \equiv 0$ and $a_i = b_i = 0$ for all $i \geq 2$.

We therefore have that

$$\begin{aligned} \lim_{n\to\infty} E_2(\phi_n) &= a = 12C^{3/2} = E_2(g_1) \\ \lim_{n\to\infty} E_3(\phi_n) &= b = -\frac{36}{5}C^{5/2} = E_3(g_1). \end{aligned}$$

Also, from (5.27) we have that $(36/7)C^{7/2} \leq J(a,b)$, and combined with (5.32), this gives

$$\lim_{n\to\infty} E_4(\phi_n) = J(a,b) = \frac{36}{7}C^{7/2} = E_4(g_1). \tag{5.33}$$

In particular, we have now shown that g_1, and hence also every element of $S(C)$, is a minimizer for $J(a,b)$.

From Lemmas 5.1 and 5.2, we have that, for $m = 2, 3, 4$,

$$E_m(\phi_n) = E_m(v_n^1) + \sum_{i=2}^{n} E_m(v_n^i) + E_m(w_n) + \epsilon_n^m,$$

where $\lim_{n\to\infty} \epsilon_n^m = 0$. From Lemma 5.6 we conclude that

$$\begin{aligned} \lim_{n\to\infty} E_2(v_n^1) &= \lim_{n\to\infty} E_2(\theta_n^1) = E_2(g_1) = a = \lim_{n\to\infty} E_2(\phi_n) \\ \lim_{n\to\infty} E_3(v_n^1) &= \lim_{n\to\infty} E_3(\theta_n^1) = E_3(g_1) = b = \lim_{n\to\infty} E_3(\phi_n). \end{aligned}$$

Therefore for $m = 2$ and $m = 3$ we have

$$\lim_{n\to\infty} \left[E_m(w_n) + \sum_{i=2}^{n} E_m(v_n^i) \right] = 0. \tag{5.34}$$

When $m = 2$, (5.34) immediately implies that

$$\lim_{n\to\infty} \left[\|w_n\|_{L^2(\mathbb{R})} + \sum_{i=2}^{n} \|v_n^i\|_{L^2(\mathbb{R})} \right] = 0. \tag{5.35}$$

We claim that when $m = 3$, (5.34) implies that

$$\lim_{n\to\infty} \left[\|(w_n)'\|_{L^2(\mathbb{R})} + \sum_{i=2}^{n} \|(v_n^i)'\|_{L^2(\mathbb{R})} \right] = 0. \tag{5.36}$$

To prove this, since $\lim_{n\to\infty} \int_{\mathbb{R}} |w_n|^3 = 0$ by Lemma 5.4, it is enough to show that

$$\lim_{n\to\infty} \sum_{i=2}^{n} \int_{\mathbb{R}} |v_n^i|^3 = 0. \tag{5.37}$$

Let $\epsilon > 0$ be given. By (5.13), we can choose N_1 such that

$$\limsup_{n\to\infty} \sum_{i=N_1}^{n} \int_{\mathbb{R}} |v_n^i|^3 < \epsilon.$$

Therefore, there exists N_2 such that for all $n \geq N_2$,

$$\sum_{i=N_1}^{n} \int_{\mathbb{R}} |v_n^i|^3 < \epsilon.$$

For each fixed $i \geq 2$, since $g_i \equiv 0$, it follows from Lemma 5.6 that $\lim_{n\to\infty} \|v_n^i\|_{H^1(\mathbb{R})} = 0$, and hence by Sobolev embedding that $\lim_{n\to\infty} \|v_n^i\|_{L^p(\mathbb{R})} = 0$ for all $p \geq 2$. So there exists N_3 such that for all $n \geq N_3$,

$$\sum_{i=2}^{N_1-1} \int_{\mathbb{R}} |v_n^i|^3 < \epsilon.$$

Then for all $n \geq \max(N_2, N_3)$,

$$\sum_{i=2}^{n} \int_{\mathbb{R}} |v_n^i|^3 < 2\epsilon,$$

proving (5.37) and (5.36).

Define $\tilde{\phi}_n(x) := \phi_n(x + x_n^1)$ for $n \in \mathbb{N}$. From (5.3) we have

$$\tilde{\phi}_n = \theta_n^1 + \sum_{i=2}^{n} \tilde{v}_n^i + \tilde{w}_n$$

for all $n \in \mathbb{N}$, where $\tilde{v}_n^i(x) := v_n^i(x + x_n^1)$ and $\tilde{w}_n(x) := w_n(x + x_n^1)$. Therefore

$$\|\tilde{\phi}_n - g_1\|_{H^1(\mathbb{R})} \leq \|\theta_n^1 - g_1\|_{H^1(\mathbb{R})} + \sum_{i=2}^{n} \|\tilde{v}_n^i\|_{H^1(\mathbb{R})} + \|\tilde{w}_n\|_{H^1(\mathbb{R})},$$

and so from Lemma 5.6, (5.35), and (5.36), we conclude that $\tilde{\phi}_n(x)$ converges strongly to g_1 in $H^1(\mathbb{R})$.

In particular, since $\{\tilde{\phi}_n\}$ is bounded in $H^2(\mathbb{R})$, it follows by the same arguments used to prove Lemma 5.7 that

$$\lim_{n\to\infty} \int_{\mathbb{R}} \tilde{\phi}_n (\tilde{\phi}_n')^2 = \int_{\mathbb{R}} g_1 (g_1')^2$$

and

$$\lim_{n\to\infty}\int_{\mathbb{R}}\tilde{\phi}_n^4=\int_{\mathbb{R}}g_1^4.$$

Therefore

$$\begin{aligned}\lim_{n\to\infty}\int_{\mathbb{R}}(\tilde{\phi}_n'')^2&=\lim_{n\to\infty}\left(E_4(\tilde{\phi}_n)+\frac{5}{3}\int_{\mathbb{R}}\tilde{\phi}_n(\tilde{\phi}_n')^2-\frac{5}{16}\int_{\mathbb{R}}\tilde{\phi}_n^4\right)\\&=E_4(g_1)+\frac{5}{3}\int_{\mathbb{R}}g_1(g_1')^2-\frac{5}{16}\int_{\mathbb{R}}g_1^4=\int_{\mathbb{R}}(g_1'')^2.\end{aligned}$$

Hence we have that

$$\lim_{n\to\infty}\|\tilde{\phi}_n\|_{H^2(\mathbb{R})}=\|g_1\|_{H^2(\mathbb{R})}. \tag{5.38}$$

But, from the weak compactness of the unit sphere in Hilbert space, we may assume by passing to a further subsequence that $\{\tilde{\phi}_n\}$ converges weakly in $H^2(\mathbb{R})$, and the limit must be g_1. From (5.38) it then follows that $\{\tilde{\phi}_n\}$ must converge strongly to g_1 in $H^2(\mathbb{R})$. This implies that

$$\lim_{n\to\infty}\|\phi_n-\psi_{C,\gamma+x_n^1}\|_{H^2(\mathbb{R})}=0,$$

which, since $\psi_{C,\gamma+x_n^1}\in S(C)$ for all $n\in\mathbb{N}$, shows that $\{\phi_n\}$ converges strongly to $S(C)$ in $H^2(\mathbb{R})$. This then completes the proof of part 1 of Theorem 2.6.

Proof of part 2 of Theorem 2.6. Assume that (2.9) holds. Applying part 2 of Lemma 4.2 with $A=(a/12)^{1/3}$ and $B=(-5b/36)^{1/5}$, we obtain that there exists a unique pair of numbers y_1 and y_2 such that $0<y_2<y_1$ and (4.3) holds. Define $C_1=y_2^2$ and $C_2=y_1^2$; then we have $0<C_1<C_2$ and

$$\begin{aligned}E_2(C_1,C_2)&=12\left(C_1^{3/2}+C_2^{3/2}\right)=a\\E_3(C_1,C_2)&=\frac{-36}{5}\left(C_1^{5/2}+C_2^{5/2}\right)=b.\end{aligned} \tag{5.39}$$

Therefore, for every pair $(\gamma_1,\gamma_2)\in\mathbb{R}^2$, we have $E_2(\psi_{C_1,C_2;\gamma_1,\gamma_2})=a$ and $E_3(\psi_{C_1,C_2;\gamma_1,\gamma_2})=b$; and hence from the definition of $J(a,b)$ we have that

$$E_4(\psi_{C_1,C_2;\gamma_1,\gamma_2})=E_4(C_1,C_2)=\frac{36}{7}\left(C_1^{7/2}+C_2^{7/2}\right)\geq J(a,b). \tag{5.40}$$

Let $\{x_n\}$ be the numbers defined after (5.30). From (5.27), (5.39), and (5.40) it follows that, for these numbers, the inequalities (4.27) hold, along with the inequality

$$\sum_{i=1}^{\infty}x_i^7\leq y_1^7+y_2^7.$$

Therefore, by part 2 of Lemma 4.10, we must have that $x_i=0$ for all $i\geq 3$, and so we conclude that

$$\begin{aligned}x_1^3+x_2^3&\leq y_1^3+y_2^3\\x_1^5+x_2^5&\geq y_1^5+y_2^5\\x_1^7+x_2^7&\leq y_1^7+y_2^7.\end{aligned}$$

It therefore follows from Lemma 4.5 that $x_1 = y_1$ and $x_2 = y_2$.

We thus see that (after relabelling the numbers D_{1i} and D_{2i} if necessary), we can reduce consideration to two possible cases: Case I in which

$$0 < D_{11} = C_1 < D_{21} = C_2,$$

and $D_{1i} = D_{2i} = 0$ for all $i \geq 2$, and Case II in which

$$0 = D_{11} < D_{21} = C_1, \quad 0 = D_{12} < D_{22} = C_2,$$

and $D_{1i} = D_{2i} = 0$ for all $i \geq 3$.

In Case I, we have that $g_1 = \psi_{C_1,C_2,\gamma_1,\gamma_2}$ for some $(\gamma_1, \gamma_2) \in \mathbb{R}^2$. Then

$$\lim_{n\to\infty} E_2(\phi_n) = a = E_2(g_1)$$
$$\lim_{n\to\infty} E_3(\phi_n) = b = E_3(g_1),$$

and from (5.27) and (5.40) we have that

$$\lim_{n\to\infty} E_4(\phi_n) = J(a,b) = E_4(g_1).$$

In particular, this implies that g_1, along with every other element of $S(C_1, C_2)$, is a minimizer for $J(a,b)$.

The same argument as in the paragraphs following equation (5.33) now shows that the translated sequence $\tilde{\phi}_n(x) = \phi_n(x + x_n^1)$ converges strongly in $H^2(\mathbb{R})$ to g_1. Hence

$$\lim_{n\to\infty} \|\phi_n - \psi_{C_1,C_2,\gamma_1+x_n^1,\gamma_2+x_n^1}\|_{H^2(\mathbb{R})} = 0,$$

which shows that $\{\phi_n\}$ converges strongly to $S(C_1, C_2)$ in $H^2(\mathbb{R})$. This completes the proof of part 2 of Theorem 2.6 in Case I.

We turn now to Case II. In this case, we have that $g_1 = \psi_{C_1,\gamma_1}$ and $g_2 = \psi_{C_2,\gamma_2}$ for some $(\gamma_1, \gamma_2) \in \mathbb{R}^2$; and $g_i \equiv 0$ for all $i \geq 3$. Then from (5.39) we have that

$$\lim_{n\to\infty} E_2(\phi_n) = a = E_2(g_1) + E_2(g_2)$$
$$\lim_{n\to\infty} E_3(\phi_n) = b = E_3(g_1) + E_3(g_2),$$

and from (5.27) and (5.40) we have that

$$\lim_{n\to\infty} E_4(\phi_n) = J(a,b) = E_4(g_1) + E_4(g_2). \tag{5.41}$$

In particular, this implies again that every element of $S(C_1, C_2)$ is a minimizer for $J(a,b)$. However, now it is no longer the case that one can translate the functions in the sequence $\{\phi_n\}$ to obtain a strongly convergent sequence in $H^2(\mathbb{R})$. Instead, we must modify the argument in the proof of part 1 of the Theorem, as follows.

Repeating the argument used above to obtain (5.34), we obtain in this case that

$$\lim_{n\to\infty} \left[E_m(w_n) + \sum_{i=3}^{n} E_m(v_n^i) \right] = 0 \tag{5.42}$$

for $m = 2$ and $m = 3$. When $m = 2$, (5.42) immediately implies that

$$\lim_{n\to\infty}\left[\|w_n\|_{L^2(\mathbb{R})} + \sum_{i=3}^{n}\|v_n^i\|_{L^2(\mathbb{R})}\right] = 0.$$

Also, by the same proof used above to prove (5.37), we have in this case that

$$\lim_{n\to\infty}\sum_{i=3}^{n}\int_{\mathbb{R}}|v_n^i|^3 = 0, \tag{5.43}$$

and together with Lemma 5.4 and (5.42) for $m = 3$, this implies that

$$\lim_{n\to\infty}\left[\|w_n\|_{H^1(\mathbb{R})} + \sum_{i=3}^{n}\|v_n^i\|_{H^1(\mathbb{R})}\right] = 0. \tag{5.44}$$

Since, by the Sobolev embedding theorem,

$$\int_{\mathbb{R}}\left|w_n(w_n')^2\right| \leq \|w_n\|_{L^\infty(\mathbb{R})}\int_{\mathbb{R}}(w_n')^2 \leq C\|w_n\|_{H^1(\mathbb{R})}^3,$$

it follows that

$$\lim_{n\to\infty}\left|E_4(w_n) - \frac{1}{2}\int_{\mathbb{R}}(w_n'')^2\right| = 0, \tag{5.45}$$

and hence

$$\liminf_{n\to\infty} E_4(w_n) \geq 0. \tag{5.46}$$

From (5.6) and (5.41), we have

$$\lim_{n\to\infty} E_4(\phi_n) = \lim_{n\to\infty}\left(\sum_{i=1}^{n}E_4(v_n^i) + E_4(w_n)\right) = E(g_1) + E(g_2).$$

Now observe that, by the same argument used to deduce (5.37) and (5.43) from (5.13), it follows from (5.14) that

$$\lim_{n\to\infty}\sum_{i=3}^{n}\int_{\mathbb{R}}\left[\left|v_n^i\left({v_n^i}'\right)^2\right| + |v_n^i|^4\right] = 0. \tag{5.47}$$

Therefore we can write

$$\lim_{n\to\infty} E_4(\phi_n) = \lim_{n\to\infty}\left[E_4(v_n^1) + E_4(v_n^2) + \frac{1}{2}\sum_{i=3}^{n}\int_{\mathbb{R}}\left({v_n^i}''\right)^2 + E_4(w_n)\right]. \tag{5.48}$$

For every sequence $\{n_k\}_{k\in\mathbb{N}}$ of integers approaching infinity, it follows from (5.47), (5.48), and Fatou's Lemma that

$$\begin{aligned}\lim_{k\to\infty} E_4(\phi_{n_k}) &\geq \liminf_{k\to\infty} E_4(v_{n_k}^1) + \liminf_{k\to\infty} E_4(v_{n_k}^2) + \frac{1}{2}\sum_{i=3}^{\infty}\liminf_{k\to\infty}\int_{\mathbb{R}}\left({v_{n_k}^i}''\right)^2 + \liminf_{k\to\infty} E_4(w_{n_k})\\ &= \liminf_{k\to\infty} E_4(v_{n_k}^1) + \liminf_{k\to\infty} E_4(v_{n_k}^2) + \sum_{i=3}^{\infty}\liminf_{k\to\infty} E_4(v_{n_k}^i) + \liminf_{k\to\infty} E_4(w_{n_k}).\end{aligned} \tag{5.49}$$

We claim now that

$$\lim_{n\to\infty} E_4(w_n) = 0 \tag{5.50}$$

and, for every $i \in \mathbb{N}$,

$$\lim_{n\to\infty} E_4(v_n^i) = E_4(g_i). \tag{5.51}$$

For otherwise, we would have either $\limsup_{n\to\infty} E_4(w_n) \geq \epsilon$ for some $\epsilon > 0$, or $\limsup_{n\to\infty} E_4(v_n^{i_0}) \geq E_4(g_{i_0}) + \epsilon$ for some $i_0 \in \mathbb{N}$ and some $\epsilon > 0$. In either case it would follow from Lemma 5.8 and (5.46) that there exists a sequence of integers $\{n_k\}_{k\in\mathbb{N}}$ approaching infinity for which (5.49) implies

$$\lim_{k\to\infty} E_4(\phi_{n_k}) \geq \sum_{i=1}^{\infty} E_4(g_i) + \epsilon.$$

But then from (5.41), since $g_i \equiv 0$ for $i \geq 3$, we obtain that

$$\lim_{k\to\infty} E_4(\phi_{n_k}) \geq \lim_{n\to\infty} E_4(\phi_n) + \epsilon.$$

This contradiction proves our claim.

From (5.45) and (5.50) we conclude that $\lim_{n\to\infty} \int_{\mathbb{R}} (w_n'')^2 = 0$, and together with (5.44) this gives that

$$\lim_{n\to\infty} \|w_n\|_{H^2(\mathbb{R})} = 0. \tag{5.52}$$

For all $i \in \mathbb{N}$, since $\{\theta_n^i\}_{n\in\mathbb{N}}$ converges to g_i strongly in $H^1(\mathbb{R})$ and weakly in $H^2(\mathbb{R})$, and since (5.51) implies that $\lim_{n\to\infty} E_4(\theta_n^i) = E_4(g_i)$ as well, it follows from the same argument used to prove (5.38) that

$$\lim_{n\to\infty} \|\theta_n^i - g_i\|_{H^2(\mathbb{R})} = 0.$$

Therefore, we have

$$\lim_{n\to\infty} \|v_n^i - \psi_{C_i,\gamma_i+x_n^i}\|_{H^2(\mathbb{R})} = 0 \quad \text{for } i = 1,2; \tag{5.53}$$

and

$$\lim_{n\to\infty} \|v_n^i\|_{H^2(\mathbb{R})} = 0 \quad \text{for } i \geq 3. \tag{5.54}$$

From Corollary 3.4, we see that $\lim_{n\to\infty} |x_n^1 - x_n^2| = \infty$, since $|x_n^1 - x_n^2| \geq r_n^1 + r_n^2$ and $\lim_{n\to\infty} r_n^1 = \lim_{n\to\infty} r_n^2 = \infty$. From (5.3) and the triangle inequality, we have

$$\begin{aligned}\|\phi_n - \psi_{C_1,C_2,\gamma_1+x_n^1,\gamma_2+x_n^2}\|_{H^2(\mathbb{R})} &\leq \|v_n^1 - \psi_{C_1,\gamma_1+x_n^1}\|_{H^2} + \|v_n^2 - \psi_{C_2,\gamma_2+x_n^2}\|_{H^2} \\ &+ \left\|\psi_{C_1,\gamma_1+x_n^1} + \psi_{C_2,\gamma_2+x_n^2} - \psi_{C_1,C_2,\gamma_1+x_n^1,\gamma_2+x_n^2}\right\|_{H^2} + \sum_{i=3}^{k} \|v_n^i\|_{H^2} + \|w_n\|_{H^2}.\end{aligned}$$

But by Lemma 2.1, (5.52), (5.53), and (5.54), all the terms on the right-hand side of the preceding inequality have limit zero as n goes to infinity. This then completes the proof of part 2 of Theorem 2.6.

Proof of Corollary 2.7. Corollary 2.7 follows from Theorem 2.6 by a standard argument, which we include here for the reader's convenience. Suppose C_1 and C_2 are given such that $0 < C_1 < C_2$, and define

$$\begin{aligned} a &= E_2(C_1, C_2) = 12(C_1^{3/2} + C_2^{3/2}) \\ b &= E_3(C_1, C_2) = \frac{-36}{5}(C_1^{5/2} + C_2^{5/2}). \end{aligned}$$

Then from Lemma 4.2 it follows that a and b satisfy (2.9), and so the assertion about convergence of minimizing sequences to $S(C_1, C_2)$ follows from Theorem 2.6.

To prove stability, we argue by contradiction: if S were not stable, then there would exist a sequence of initial data $\{u_{0n}\}_{n\in\mathbb{N}}$ in $H^2(\mathbb{R})$ such that $\lim_{n\to\infty} d(u_{0n}, S) = 0$ and a number $\epsilon > 0$ and a sequence of times $\{t_n\}_{n\in\mathbb{N}}$ such that the solutions $u_n(x,t)$ of KdV with initial data $u_n(\cdot, 0) = u_{0n}$ would satisfy

$$d(u(\cdot, t_n), S) \geq \epsilon \tag{5.55}$$

for all $n \in \mathbb{N}$. Let $\phi_n = u(\cdot, t_n)$ for $n \in \mathbb{N}$. Since E_2, E_3, and E_4 are continuous functionals on $H^2(\mathbb{R})$, and are conserved under the time evolution of the KdV equation, we have

$$\begin{aligned} \lim_{n\to\infty} E_2(\phi_n) &= \lim_{n\to\infty} E_2(u_{0n}) = E_2(C_1, C_2) = a \\ \lim_{n\to\infty} E_2(\phi_n) &= \lim_{n\to\infty} E_3(u_{0n}) = E_3(C_1, C_2) = b \\ \lim_{n\to\infty} E_4(\phi_n) &= \lim_{n\to\infty} E_4(u_{0n}) = E_4(C_1, C_2) = J(a, b). \end{aligned}$$

Hence $\{\phi_n\}$ is a minimizing sequence for $J(a, b)$, and so must converge strongly to S in $H^2(\mathbb{R})$. But this then contradicts (5.55).

6 Acknowledgements

The authors would like to thank Jerry Bona for introducing us to the problems considered in this paper, and more generally for his help and guidance over many years.

References

[1] J. Albert, Concentration compactness and the stability of solitary-wave solutions to nonlocal equations, in *Applied analysis (Baton Rouge, LA, 1996), Contemp. Math.*, **221**, Amer. Math. Soc., Providence, RI, 1999.

[2] J. Albert, A uniqueness result for 2-soliton solutions of the KdV equation, *Discrete and Continuous Dynamical Systems - A* **39** (2019), 3635–3670.

[3] J. P. Albert, J. L. Bona, and N. V. Nguyen, On the stability of KdV multi-solitons, *Differential Integral Equations* **20** (2007), 841–878.

[4] T. B. Benjamin, The stability of solitary waves, *Proc. Roy. Soc. London Ser. A* **328** (1972), 153–183.

[5] J. Bona, On the stability theory of solitary waves, *Proc. Roy. Soc. London Ser. A* **344**(1975), 363–374.

[6] H. Brezis and J.-M. Coron, Convergence of solutions of H-systems or how to blow bubbles, Arch. Rational Mech. Anal. **89** (1985), 21–56.

[7] T. Cazenave and P.-L. Lions, Orbital stability of standing waves for some nonlinear Schrödinger equations, *Comm. Math. Phys.* **85** (1982), 549–561.

[8] L. A. Dickey, *Soliton equations and Hamiltonian systems (1st edition)*, World Scientific, 1991.

[9] L. A. Dickey, *Soliton equations and Hamiltonian systems (2nd edition)*, World Scientific, 2003.

[10] P. Gérard, Description du défaut de compacité de l'injection de Sobolev, *ESAIM Control Optim. Calc. Var.* **3** (1998), 213–233.

[11] R. Killip and M. Visan, Orbital stability of KdV multisolitons in H^{-1}, `arXiv:2009.06746[math.AP]`

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

[13] P. D. Lax, Periodic solutions of the KdV equation, *Comm. Pure Appl. Math.* **2**8 (1975), 141–188.

[14] P. D. Lax, Almost periodic solutions of the KdV equation, *SIAM Rev.* **18** (1976), 351–375.

[15] S. Le Coz and Z. Wang, Stability of the multi-solitons of the modified Korteweg-de Vries equation, `arXiv:2010.00814[math.AP]`

[16] D. G. Luenberger, *Optimization by Vector Space Methods*, Wiley, 1969.

[17] D. G. Luenberger and Yinyu Ye, *Linear and Nonlinear Programming*, Springer, 2008 .

[18] J. Maddocks and R. Sachs, On the stability of KdV multi-solitons, *Comm. Pure Appl. Math.* **46** (1993), 867–901.

[19] M. Mariş, Profile decomposition for sequences of Borel measures, arXiv:1410.6125.

[20] Y. Murata, *Mathematics for Stability and Optimization of Economic Systems*, Academic Press, 1977.

[21] S. P. Novikov, A periodic problem for the Korteweg-de Vries equation. I. (Russian) *Funkcional. Anal. i Priložen.* **8** (1974), 54–66. English translation available at `http://www.mi-ras.ru/~snovikov/43.pdf`

[22] F. Warner, *Foundations of Differentiable Manifolds and Lie Groups*, Scott, Foresman, 1971.