Minimax Theory and its Applications
Volume 7 (2022), No. 2, 365–380

A Computer Assisted Proof of the Symmetries of Least Energy Nodal Solutions on Squares

Ariel Salort
Departamento de Matemática, FCEyN – Universidad de Buenos Aires
and: IMAS - CONICET, Buenos Aires, Argentina
asalort@dm.uba.ar

Christophe Troestler
Départment de Mathématique, Université de Mons, Belgium
christophe.troestler@umons.ac.be

Received: February 7, 2022
Accepted: March 7, 2022

Using a Lyapunov-Schmidt reduction on an asymptotic Nehari manifold and verified computations, we prove that the least energy nodal solutions to Lane-Emden equation $-\Delta u = |u|^{p-2}u$ with zero Dirichlet boundary conditions on a square are odd with respect to one diagonal and even with respect to the other one when p is close to 2.

Keywords: Least energy sign changing solutions, symmetries, interval arithmetic, verified computation.

2020 Mathematics Subject Classification: Primary 35J20; secondary 35B06, 65G40.

1. Introduction

Given a connected domain $\Omega \subset \mathbb{R}^N$, $N \geq 2$ and $2 < p < 2^* := \frac{2N}{N-2}$ ($+\infty$ if $N = 2$), we consider the so-called Lane-Emden problem with Dirichlet boundary conditions

$$\begin{cases} -\Delta u = |u|^{p-2}u & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{1}$$

In this article we aim at symmetry properties of solutions of (1) for values of p close to 2. The literature on this subject is vast. In the early '70s Ambrosetti and Rabinowitz [2] generalized the work of Nehari [14] to prove that (1) has a positive ground state solution which, in [9], was shown to inherit all symmetries of convex domains. These solutions are unique when p is close to 2 and "usually" non-degenerate as established by Dancer [6].

Later on, in the '90s, solutions with minimum energy among sign-changing ones (or *least energy nodal solutions* hereafter referred to as *l.e.n.s.*) were proved to exists by Castro, Cossio and Neuberger [5] by minimization on a "nodal Nehari set." A natural question is whether l.e.n.s. inherit (some) symmetries of the domain.

The authors are partially supported by the project "Existence and asymptotic behavior of solutions to systems of semilinear elliptic partial differential equations" (T.1110.14) of the *Fonds de la Recherche Fondamentale Collective*, Belgium.

ISSN 2199-1413 (printed), ISSN 2199-1421 (electronic) / \$ 2.50

There are several results in this direction. Aftalion and Pacella [1] proved that, on a ball, a l.e.n.s. cannot be radial. Bartsch, Weth and Willem [3] established that, on radial domains, l.e.n.s. are foliated Schwarz symmetric and, in particular, even w.r.t. $N-1$ orthogonal hyperplanes. In [4], symmetries on more general domains were studied, as well as the asymptotic behavior of solutions when $p \to 2$. A similar asymptotic analysis was performed by Grossi who showed that, from any eigenvalue of multiplicity m of the linear problem $-\Delta u = \lambda u$ with zero Dirichlet boundary conditions, m branches of solutions emanate [10].

It is worth to mention that the arguments leading to uniqueness and the symmetry properties of l.e.n.s. of (1) for values of p close to 2 strongly depend on the simplicity of λ_2. Indeed, when λ_2 is simple, uniqueness of the limit solution follows by an application the implicit function theorem. In this case, for p close to 2, l.e.n.s. are unique (up their sign) and possess the same symmetries as eigenfunctions corresponding to λ_2. In particular, on a rectangle, the nodal line is the small median [4].

However, the approach is different when λ_2 has multiplicity. When Ω is a ball, as all eigenfunctions for λ_2 can be obtained from a one-dimensional subspace by rotations, one is still able to apply the implicit function theorem and deduce that l.e.n.s. are odd w.r.t. an hyperplane passing through the center of the ball and invariant by rotations around the axis orthogonal to that hyperplane [4, Theorem 3]. In particular, their nodal surface is a $N-1$-dimensional ball. For general domains, the implicit function theorem can no longer be applied and symmetries of l.e.n.s. for values of p near 2 are less understood. Even for simple domains such as planar squares, many questions remain open. Particularly, numerical simulations carried out in [4] support the following assertion:

Conjecture 1.1. *If $\Omega \subset \mathbb{R}^2$ is a square and p is close to 2, l.e.n.s. are even with respect to a diagonal and odd with respect to the orthogonal diagonal. In particular, their nodal line is a diagonal.*

Despite the fact that everything is explicit, a “paper and pencil proof” of that fact seems not to be clear. The main aim of this article is to prove this conjecture by using a computer-assisted proof approach.

The manuscript is organized as follows. In section 2 we introduce some preliminary results. In Section 3, we explain the main ingredients of our arguments. Section 4 is devoted to the description and implementation our computer-assisted proof and in Section 5 we put it all together to establish the symmetry property.

2. Preliminary results

In this section we introduce some notations and preliminary results. We first recall a variational formulation for sign-changing solutions of (1) as well as the limit equation they fulfill as p approaches 2. Finally we present some abstract results on symmetries of solutions for values of p close to 2.

2.1. Variational formulation

In order to study solutions of (1), and to avoid further normalizations, we consider the following problem

$$(\mathcal{P}_p) \qquad \begin{cases} -\Delta u = \lambda_2 |u|^{p-2} u & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases}$$

where λ_2 is the second eigenvalue of $-\Delta$ with zero Dirichlet boundary conditions. It is straightforward that symmetry properties of solutions to (1) and (2.1) are the same since u solves (1) if and only if $\lambda_2^{1/(2-p)}u$ solves (2.1). Weak solutions to (2.1) are critical points of the functional $\mathcal{J}_p : H_0^1(\Omega) \to \mathbb{R}$ defined by

$$\mathcal{J}_p(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 \,\mathrm{d}x - \frac{\lambda_2}{p}\int_\Omega |u|^p \,\mathrm{d}x.$$

This functional has Fréchet derivative given by

$$\forall v \in H_0^1(\Omega), \qquad \mathcal{J}_p'(u)[v] = \int_\Omega \nabla u \cdot \nabla v \,\mathrm{d}x - \lambda_2 \int_\Omega |u|^{p-2}uv \,\mathrm{d}x.$$

We remark that standard regularity theory arguments imply that weak solutions to (2.1) belong to $C_0^2(\Omega) \cap C(\overline{\Omega})$ and hence classical solutions.

Clearly the zero function solves (2.1). All non-zero critical points of $\mathcal{J}_p$ belong to the Nehari manifold

$$\mathcal{N}_p := \left\{u \in H_0^1(\Omega) \setminus \{0\} : \mathcal{J}_p'(u)[u] = 0\right\}$$

and all sign-changing solutions belong to the nodal Nehari set

$$\mathcal{M}_p := \left\{u \in H_0^1(\Omega) : u^\pm \in \mathcal{N}_p\right\}$$

where as usual $u^\pm := \max\{\pm u, 0\}$. Given a function $u \in H_0^1(\Omega)$ with $u^\pm \neq 0$, $u \in \mathcal{M}_p$ if and only if $\int_\Omega |\nabla u^\pm|^2 \,\mathrm{d}x = \lambda_2 \int_\Omega |u^\pm|^p \,\mathrm{d}x$.

It can be shown [3, 5] that the minimum of $\mathcal{J}_p$ on $\mathcal{M}_p$ is achieved and that all minimizers are nodal solutions of (2.1) – referred to as *least energy nodal solutions* – with exactly two nodal domains.

2.2. The asymptotic equation

Let us analyze the behavior of l.e.n.s. solutions of (2.1) as $p \to 2$. To start, any family $(u_p)_{p>2}$ of l.e.n.s. of (2.1) can be proved to be bounded in $H_0^1(\Omega)$ and bounded away from 0 (see [4, Lemma 4.1 and Lemma 4.4]). This implies that any weak accumulation point u_* of $(u_p)_{p>2}$ verifies a limit equation [4, Theorem 4.5]: if $u_{p_n} \rightharpoonup u_*$ weakly in $H_0^1(\Omega)$ as $p_n \to 2$, then $u_{p_n} \to u_*$ in $H_0^1(\Omega) \setminus \{0\}$ and u_* fulfills

$$\begin{cases} -\Delta u_* = \lambda_2 u_* & \text{in } \Omega, \\ u_* = 0 & \text{on } \partial\Omega. \end{cases} \tag{2}$$

Thus $u_* \in E_2 \setminus \{0\}$ where E_2 denotes the eigenspace related to the second eigenvalue λ_2. Moreover, u_* minimizes the *asymptotic functional* $\mathcal{J}_*$ over the *asymptotic Nehari manifold* $\mathcal{N}_*$, where

$$\mathcal{J}_* : E_2 \to \mathbb{R} : u \mapsto \mathcal{J}_*(u) = \frac{\lambda_2}{4}\int_\Omega u^2 - u^2 \log u^2 \,\mathrm{d}x, \tag{3}$$

$$\mathcal{N}_* = \{u \in E_2 \setminus \{0\} : \mathcal{J}_*'(u)[u] = 0\},$$

and $t^2 \log t^2$ is extended continuously by 0 at $t = 0$.

In particular u_* is a critical point of $\mathcal{J}_*$. It is easy to see that

$$\forall v \in E_2, \qquad \mathcal{J}_*'(u)[v] = -\lambda_2 \int_\Omega vu \log|u| \,\mathrm{d}x. \tag{4}$$

Any nontrivial critical point of $\mathcal{J}_*$ belongs to $\mathcal{N}_*$. This manifold $\mathcal{N}_*$ is compact and such that $u \in \mathcal{N}_*$ if and only if $u \neq 0$ and

$$\int_\Omega u^2 \log|u| \,\mathrm{d}x = 0 \quad \text{or equivalently iff} \quad \mathcal{J}_*(u) = \frac{\lambda_2}{4} \int_\Omega u^2 \,\mathrm{d}x.$$

Let us now show that u_* satisfies an equivalent minimization problem. This problem will be the one used in our computer assisted proof. First, as for the classical Nehari manifold, one has that

$$\inf_{u \in \mathcal{N}_*} \mathcal{J}_*(u) = \inf_{u \in \mathbb{S}} \max_{t>0} \mathcal{J}_*(tu), \tag{5}$$

where $\mathbb{S}$ is the L^2-unit sphere of E_2. Indeed, observe that, for any $u \in E_2 \setminus \{0\}$, the ray $\{tu : t > 0\}$ intersects $\mathcal{N}_*$ at a single point $t_u^* u$ maximizing $\mathcal{J}_*$. The value of $t_u^* \in \,]0, +\infty[$ is easily computed:

$$t_u^* = \exp\left(- \frac{\int_\Omega u^2 \log|u| \,\mathrm{d}x}{\int_\Omega u^2 \,\mathrm{d}x} \right). \tag{6}$$

In view of the above expression, given $u \in \mathbb{S}$, we get

$$\mathcal{J}_*(t_u^* u) = \frac{\lambda_2}{4} \int_\Omega (t_u^* u)^2 \,\mathrm{d}x = \frac{\lambda_2}{4} \,\mathrm{e}^{-2 \int_\Omega u^2 \log|u| \,\mathrm{d}x}$$

and consequently, since the map $t \mapsto e^{2t}$ is increasing, (5) is equivalent to

$$\text{minimize } \mathcal{S}_*(u) := -\int_\Omega u^2 \log|u| \,\mathrm{d}x \text{ on the } \mathrm{L}^2\text{-sphere } \mathbb{S}. \tag{7}$$

Note that for any $u \in \mathbb{S}$ and $r > 0$,

$$\mathcal{S}_*(ru) = r^2 \mathcal{S}_*(u) - r^2 \log r \tag{8}$$

so, instead of (7), one can equivalently minimize $\mathcal{S}_*$ on the sphere $r\mathbb{S}$ of radius r.

The function $\mathcal{S}_*$ can be seen as the *entropy* associated to the density $|u|^2$ and accumulation points u_* of the minimal energy nodal solutions are multiples of the eigenfunctions with minimal entropy. Now that the limit functions u_* of l.e.n.s. sequences are characterized, our following concern is to understand how their symmetries yield symmetries of the u_p for p close to 2.

2.3. Symmetries

As already noted, when studying symmetries there are two possibles scenarios depending of the simplicity of λ_2.

When $\dim E_2 = 1$, using the *implicit function theorem*, it follows that (2.1) has a unique solution (up its sign) for $p \approx 2$. Therefore [4, Theorem 2], there exists p^*

sufficiently close to 2 such that, for any reflection R such that $R(\Omega) = \Omega$, a function $u_2 \in E_2 \setminus \{0\}$ is even or odd w.r.t. R if and only if u_p possesses the same invariance for $p \in \,]2, p^*]$. For instance, when Ω is a rectangle, since λ_2 is simple, l.e.n.s. to (2.1), with $p \approx 2$, are odd w.r.t. the small median of the rectangle (so their nodal line is that median) and even w.r.t. the large median.

When λ_2 is not simple, with the exception of radial domains, the questions of uniqueness and symmetries become more delicate. Symmetries of l.e.n.s. to (2.1) for p close to 2 were addressed in [4, Theorem 3.6] by the following abstract result. Consider a family of groups $(G_\alpha)_{\alpha \in E_2}$ acting on $H_0^1(\Omega)$ in a such way that, for every $\alpha \in E_2$, $g \in G_\alpha$, p close to 2, and $u \in H_0^1(\Omega)$, the following holds:

$$\text{(i) } g(E_2) = E_2, \quad \text{(ii) } g(E_2^\perp) = E_2^\perp, \quad \text{(iii) } g\alpha = \alpha, \quad \text{(iv) } \mathcal{J}_p(gu) = \mathcal{J}_p(u).$$

Then, for all $M > 0$, if p is close to 2, any l.e.n.s. $u_p \in \{u \in B_M : P_{E_2}(u) \notin B_{1/M}\}$ of (2.1) is invariant under the isotropy group $G_{\alpha_p} = \{g \in G : g\alpha_p = \alpha_p\}$ of $\alpha_p = P_{E_2} u_p$, the orthogonal projection of u_p on E_2. In particular, if G_α describes the symmetries (or antisymmetries) of α, for p close to 2, u_p possesses the same symmetries as its orthogonal projection α_p. For instance, when Ω is a square and p is close to 2, l.e.n.s. of (2.1) are odd with respect to the center of the square.

Moreover, as we anticipated, in [4, Conjecture 5.4] numerical computations suggest that u_* is odd w.r.t. a diagonal and even w.r.t. the other diagonal. It seems natural to think that these symmetries extend to u_p for p close to 2, whence conjecture 1.1. This is the content of our main result.

Theorem 2.1. *If $\Omega \subset \mathbb{R}^2$ is a square and p is close to 2, l.e.n.s. are even with respect to a diagonal and odd with respect to the orthogonal diagonal.*

Unfortunately the above mentioned abstract result is not powerful enough for this purpose. Indeed, the computer assisted proof below will characterize the symmetries of $u_* \in \mathcal{N}_*$ but that does not readily implies that $P_{E_2} u_p = u_*$, so it is not clear that u_p enjoys the same symmetries as u_*. To do that, the implicit function theorem needs to be replaced by a Lyapunov-Schmidt reduction. Such approach was done by Grossi and we extract here (with our notations) the part of his results that is useful for our purposes [10].

Theorem 2.2. (Lyapunov-Schmidt reduction) *Assume $\mathcal{J}_* \in C^2(E_2 \setminus \{0\}; \mathbb{R})$ and let $u_* \in E_2 \setminus \{0\}$ be a non-degenerate critical point of $\mathcal{J}_*$. There exists a neighborhood U_* of u_* in $H_0^1(\Omega)$ and a function $\gamma : [2, 2+\varepsilon[\,\to H_0^1(\Omega)$, $\varepsilon > 0$, such that $\gamma(2) = u_*$ and*

$$\forall p \in \,]2, 2+\varepsilon[, \ \forall u \in U_*, \qquad u \text{ solves } (\mathcal{P}_p) \iff u = \gamma(p).$$

Proposition 2.3. *Let $(u_p)_{p>2}$ be a family of l.e.n.s. converging to $u_* \in E_2 \setminus \{0\}$. Assume $\mathcal{J}_* \in C^2(E_2 \setminus \{0\}; \mathbb{R})$ and u_* is a non-degenerate critical point of $\mathcal{J}_*$. If $gu_* = g_*$ (where $gu(x) := u(g^{-1}x)$ or $gu(x) := -u(g^{-1}x)$) for some $g \in O(N)$ such that $g\Omega = \Omega$ and $\mathcal{J}_p(gu) = \mathcal{J}_p(u)$ for all $u \in H_0^1(\Omega)$ and p, then, for p close to 2, $gu_p = u_p$.*

Proof. If the conclusion does not hold, there exists a sequence (u_{p_n}) of l.e.n.s. to $(\mathcal{P}_{p_n})$ such that $\forall n$, $gu_{p_n} \neq u_{p_n}$. Going if necessary to a subsequence $u_{p_n} \rightharpoonup u_* \in E_2$ and thus $u_{p_n} \to u_*$ [4, Theorem 4.5].

In particular, for n large enough, $(p_n, u_{p_n}) \in [2, 2+\varepsilon[\times U_*$ and so $u_{p_n} = \gamma(p_n)$. But gu_{p_n} is also a solution to $(\mathcal{P}_{p_n})$ and $gu_{p_n} \to gu_* = u_*$. Therefore, for n large enough, $gu_{p_n} = \gamma(p_n) = u_{p_n}$ which is a contradiction. □

Remark 2.4. The regularity assumption will be shown to hold for the square in Section 4.5.

3. Reduction to a one-dimensional functional

We sketch our strategy to prove Theorem 2.1. Let $\Omega \subset \mathbb{R}^2$ be a square. Without loss of generality we can consider $\Omega =]0,1[^2$. In this case, a basis of eigenfunctions of E_2 is given by

$$\phi_1(x,y) = \sin(\pi x)\sin(2\pi y), \qquad \phi_2(x,y) = \sin(2\pi x)\sin(\pi y). \tag{9}$$

Moreover, it easy to check that these functions are orthogonal in $\mathrm{L}^2(\Omega)$, which allow us to parametrize $\frac{1}{2}\mathbb{S}$ as

$$\begin{aligned} \tfrac{1}{2}\mathbb{S} &= \bigl\{u \in E_2 : \|u\|_{\mathrm{L}^2(\Omega)} = \tfrac{1}{2}\bigr\} \\ &= \bigl\{u_\theta(x,y) := \phi_1(x,y)\cos\theta - \phi_2(x,y)\sin\theta \ : \ \theta \in [0, 2\pi[\bigr\}. \end{aligned} \tag{10}$$

Hence, thanks to (8), instead of (7) we can restrict ourselves to study the minimization problem

$$\min_{\theta \in [0,2\pi]} \mathcal{S}_*(u_\theta). \tag{11}$$

Since $u_{\theta+\pi/2}(x,y) = u_\theta(y,-x)$, $u_{\pi/4-\theta}(x,y) = u_{\pi/4+\theta}(y,x)$ and $\mathcal{S}_*$ is invariant under the group of symmetries of the square (the dihedral group of order 8), we have that the function $g(\theta) : [0, \frac{\pi}{4}] \to \mathbb{R}$ given by

$$g(\theta) := \mathcal{S}_*(u_\theta) = -\int_\Omega f\bigl(u_\theta(x,y)\bigr)\,\mathrm{d}x\,\mathrm{d}y, \quad \text{with } f(t) = \begin{cases} t^2 \log|t| & \text{if } t \neq 0, \\ 0 & \text{if } t = 0, \end{cases} \tag{12}$$

is $\frac{\pi}{2}$-periodic and $g(\frac{\pi}{4} - \theta) = g(\frac{\pi}{4} + \theta)$ (see Fig. 3.1).

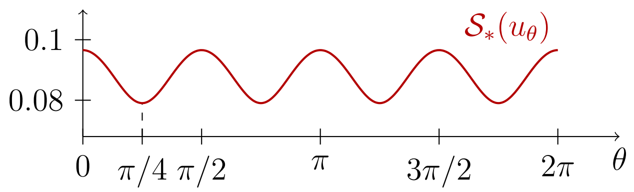

Figure 3.1: Graph of $\theta \mapsto \mathcal{S}_*(u_\theta)$.

Observe that $u_{\pi/4} \in \frac{1}{2}\mathbb{S}$ is even with respect to a diagonal and odd with respect to the second one. Because of the symmetries of g, it is clear that $\theta = \pi/4$ is a critical point of g. If we prove that the minimal energy in (11) is achieved at $\theta = \pi/4$ and that the corresponding critical point of $\mathcal{J}_*$ is non-degenerate, Theorem 2.3 will assert that l.e.n.s. to (2.1) would enjoy the symmetries of Conjecture 1.1.

4. Ingredients of the computer-assisted proof

In this section we prove that the function $\theta \mapsto g(\theta)$ reaches its minimum at $\theta = \frac{\pi}{4}$. Throughout this section Ω stands for the square $]0,1[^2$.

The arguments in the proof can be sketched roughly as follows. We design an adaptive integration method `integ` which allows to integrate a function of two variables on Ω. By applying it to $(x,y) \mapsto f(u_\theta(x,y))$, we obtain a function `entropy` which estimates g. With these tools, our task is reduced to discard subintervals of $[0, \frac{\pi}{4}]$ which are guaranteed not to contain the minimum of g. This is performed by the function `exclude`. The outcome is a small interval around $\pi/4$ in which the minimum must lie. Finally, in that interval, we analyze the concavity of g in order to guarantee that it contains a unique critical point. Consequently $\pi/4$ is the only critical point in that interval and is therefore the minimum.

When our scheme is implemented in a computer, one has to be conscious of the existence of several sources of error involved in the computations: numerical integration methods carry their corresponding error, rounding errors come from computing with floating point numbers, etc. Hence, we have to track the precision of our computations for our conclusions to be valid in spite of these various errors. Since in general real numbers are not *machine numbers*, to guarantee rigorous results when performing computations, instead of considering *values*, we will consider *intervals* containing the true values. In the next subsection, for the reader's convenience, we recall some basic facts about computing with floating numbers and interval arithmetic.

4.1. Interval arithmetic

When implementations are made in digital electronic devices, due to the limited space and speed, only a finite subset of real numbers is available. For the present considerations, we will use the set of double precision floating point numbers that we will denote $\mathbb{F}$. Elementary arithmetic operations ($+$, $-$, $\cdot$, $/$) do not necessarily return an element in $\mathbb{F}$ even when their operands are in $\mathbb{F}$. The standard IEEE-754 mandates that their implementation returns a value in $\mathbb{F}$ closest to the exact result. This yields small rounding errors which may propagate badly along the computations. A classical example [11, 16] is to evaluate

$$f(x,y) = 333.75y^6 + x^2(11x^2y^2 - y^6 - 121y^4 - 2) + 5.5y^8$$

at $(x,y) = (77617, 33096)$. Despite the fact that all coefficients of f and x, y belong to $\mathbb{F}$, when all computations are performed with *double precision*, the result is $-1.180... \cdot 10^{21}$ while the correct value is -2. Rounding errors may also affect the constants in the program: for example $0.1 \notin \mathbb{F}$ (because it has an infinite binary expansion) and $\pi \notin \mathbb{F}$ are rounded in computer memory. Errors also come from the necessary approximation (think of the truncation of an infinite sum) performed by the algorithm computing functions such as sin, cos, log,...

In order to account for all these errors along computations performed on a computer, values $a \in \mathbb{R}$ will be represented by an interval

$$[a] = [\underline{a}, \overline{a}] = \{x \in \mathbb{R} : \overline{a} \le x \le \underline{a}\}$$

(where $[a]$ must read as a single symbol) such that $a \in [a]$. The set of intervals with endpoints in $\mathbb{R}$ will be denoted $\mathbb{IR}$. The *width* of $[a]$ is $\overline{a} - \underline{a}$ and denoted width($[a]$). We also define $\operatorname{low}([\underline{a}, \overline{a}]) := \underline{a}$ and $\operatorname{high}([\underline{a}, \overline{a}]) := \overline{a}$. Vectors $a \in \mathbb{R}^N$ will be represented as $[a] = \big([a_1], \dots, [a_N]\big) \in (\mathbb{IR})^N$ standing for the *box* $[a_1] \times \cdots \times [a_N]$.

Elementary arithmetic operations such as $(+,\ -,\ \cdot,\ /)$ and in general any function $f : \mathbb{R}^N \to \mathbb{R}^M$ must be extended to an interval function $[f] : (\mathbb{IR})^N \to (\mathbb{IR})^M$ so that the following *containment property* holds:

$$\forall [a] \in \operatorname{Dom}[f], \qquad f([a]) = \big\{ f(x) : x \in [a] \cap \operatorname{Dom} f \big\} \subseteq [f]\big([a]\big).$$

This means that all exact values $f(x)$ for x ranging in the interval $[a]$ are contained in the interval returned by the function $[f]$ on the operand $[a]$. In this case, we will call $[f]$ an *interval extension* of f. Composition of interval extensions are again interval extensions.

When implemented on a computer, interval extensions of $f : \mathbb{R}^N \to \mathbb{R}^M$ will naturally be functions $[f]$ defined on (a subset of) $(\mathbb{IF})^N$ and returning values in $(\mathbb{IF})^M$ where $\mathbb{IF}$ denotes the set of intervals with endpoints in $\mathbb{F}$. Rounding errors of arithmetic operators will be handled by using *directed rounding modes* of the processor: for example, the interval extension of the addition if $[a] \in \mathbb{IF}$ and $[b] \in \mathbb{IF}$ will be given by $\big[{\downarrow}(\underline{a} + \underline{b}), {\uparrow}(\overline{a} + \overline{b})\big]$ where ${\downarrow}(\dots)$ (resp. ${\uparrow}(\dots)$) means that the outcome of operations within the braces if rounded towards $-\infty$ (resp. $+\infty$). For approximations made to evaluate functions, a theoretical error bound must be derived in terms of computable quantities which will then be used to provide an (over)estimation of the error on the computer.

We refer the reader interested in more details about interval analysis to [12, 13].

4.2. Interval extension of $\mathbf{u^2 \log|u|}$

One of the basic functions to implement for our task is $u \mapsto u^2 \log|u|$. The naive extension is $[u] \mapsto [u]^2 \cdot [\log]\big|[u]\big|$ but it is unsuitable. Indeed if $0 \in [u]$, the logarithm interval extension will return an unbounded interval (of the form $[-\infty, a]$, which is possible because $\pm\infty \in \mathbb{F}$) and so will subsequent computations. Thus a more precise interval extension of this function is necessary. Because this function is even, one can suppose w.l.o.g. that $[u] \subseteq [0, +\infty]$. In order to derive interval bounds, one has to distinguish sub-intervals based on the monotonicity of $u \mapsto u^2 \log u$ (which is decreasing on $[0, \mathrm{e}^{-1/2}]$ and increasing after) and the sign of $\log u$. Thus one has first to compute an interval estimate $[\underline{m}, \overline{m}] = [\exp]\big([-0.5, -0.5]\big)$ of the location of the minimum. Then, for an interval of the form $[0, \overline{u}]$ with $0 < \overline{u} < \underline{m}$, the algorithm will return the interval $[{\downarrow}({\uparrow}(\overline{u}^2) \cdot \log \overline{u}), 0]$ (one has to round up the square because $\log \overline{u} < 0$). A similar analysis is done for other intervals $[u]$. This more clever interval extension returns tight bounds for the function and, in particular, returns bounded intervals even when $0 \in [u]$ (see Fig. 4.1).

4.3. Integration

In order to compute $\mathcal{S}_*$, an integration on the square Ω must be performed. There is a vast literature about integration methods, both for the one-variable and multivariate cases, including for interval arithmetic (see for example [8, 15, 15] and the references therein).

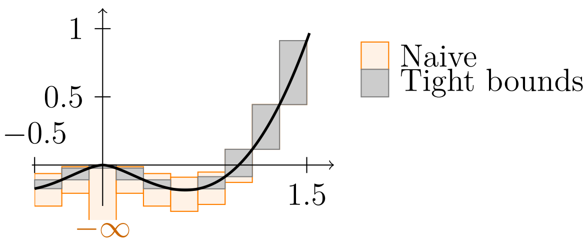

Figure 4.1: Evaluation of $u \mapsto u^2 \log|u|$.

For accurate evaluation of integrals in high dimensions, it is recommended to avoid tensor products of one-dimensional rules (because they require a number of function evaluations that is exponential in the dimension) and instead turn to rules such as the one devised by Smolyak [17]. In this paper however, the dimension is small and the fact that error bounds of tensor product formula do not depend on mixed derivatives is interesting. Our integrand is also not smooth everywhere.

In this subsection, we briefly describe the adaptive verified integration scheme that we are using. Given an interval extension $[f]$ of a function $f : \Omega \to \mathbb{R}$, we want to return an interval $[\mathcal{I}_\Omega]$ such that[1]

$$[\mathcal{I}_\Omega] \ni \mathcal{I}_\Omega := \frac{1}{|\Omega|} \int_\Omega f(x, y)\, \mathrm{d}x\, \mathrm{d}y.$$

To compute $[\mathcal{I}_\Omega]$, a *rule* that evaluates $[f]$ at various points and estimates the error in terms of the variation (on Ω) of f and its derivatives is used. Such a rule returns an interval $[\mathcal{I}^0_\Omega]$. If we are happy with the width of $[\mathcal{I}^0_\Omega]$, we take $[\mathcal{I}_\Omega] = [\mathcal{I}^0_\Omega]$ and stop. If not, we split Ω into four squares of equal sizes $\Omega = \Omega_1 \cup \Omega_2 \cup \Omega_3 \cup \Omega_4$ and recursively apply this procedure to have $[\mathcal{I}_{\Omega_i}]$, $i = 1, \dots, 4$. Then $[\mathcal{I}_\Omega]$ is defined by

$$[\mathcal{I}_\Omega] := \frac{1}{4} \sum_{i=1}^{4} [\mathcal{I}_{\Omega_i}] \ni \sum_{i=1}^{4} \frac{|\Omega_i|}{|\Omega|} \mathcal{I}_{\Omega_i} = \mathcal{I}_\Omega. \tag{13}$$

Since (13) implies $\operatorname{width}([\mathcal{I}_\Omega]) \le \max_{i=1,\dots,4} \operatorname{width}([\mathcal{I}_{\Omega_i}])$, a natural stopping criteria for the recursion is that $\operatorname{width}([\mathcal{I}_{\Omega_i}]) \le \varepsilon$ where ε is a desired tolerance.

4.3.1. Basic integration rule

The simplest way to estimate $\mathcal{I}_\Omega$ is to use the following simple fact (related to the mean value theorem for integrals): if f is integrable and $\forall \xi \in \Omega$, $\underline{m} \le f(\xi) \le \overline{m}$, then

$$\frac{1}{|\Omega|} \int_\Omega f(\xi)\, \mathrm{d}\xi \in [\underline{m}, \overline{m}].$$

If $\Omega \subseteq [x] \times [y]$, this is in particular true for $[\underline{m}, \overline{m}] = [f]([x], [y])$. Thus, one takes

$$[\mathcal{I}^0_\Omega] := [f]\big([x], [y]\big).$$

The associated adaptive procedure does not perform very well however. Indeed, we expect the width of $[f]\big([x], [y]\big)$ to be of size $\|\nabla f\|_{\mathrm{L}^\infty(\Omega)} \operatorname{diam} \Omega$ where $\operatorname{diam} \Omega$ is

[1]We choose to compute the integral mean because the weights in the sum approximating the integral are independent of the size of Ω.

the diameter of Ω. Thus, if h denotes the diameter of the small squares obtained at depth d of the recursion, for $f \in C^1(\overline{\Omega})$, $\operatorname{width}([\mathcal{I}_\Omega]) = O(h)$. In terms of the number n of points at which the function needs to be evaluated and in terms of the depth d of the recursion, this gives:

$$\operatorname{width}([\mathcal{I}_\Omega]) = O(h) = O\Big(\frac{1}{\sqrt{n}}\Big) = O\Big(\frac{1}{2^d}\Big). \tag{14}$$

This is quite slow. For example, integrating the analytic function ϕ_1 (defined in (9)) using a recursion depth of 14 must perform at least 268 468 225 function evaluations for a final precision of about 10^{-3}. This is pictured in Fig. 4.2. Clearly, this is not practical as the procedure to determine a small neighborhood of the minumum requires many evaluations of $\mathcal{S}_*$ with a good precision.

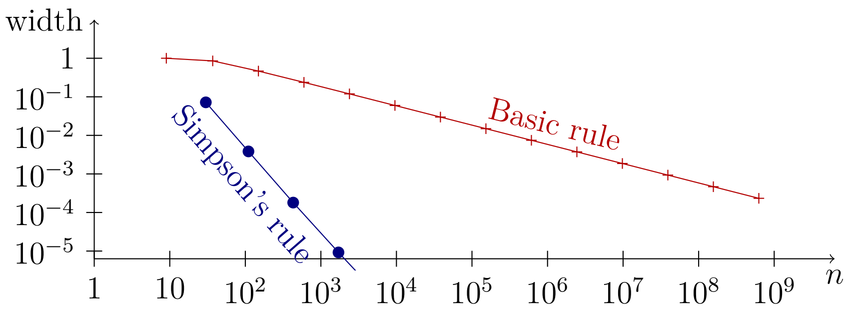

Figure 4.2: Errors of adaptive integration rules.

4.3.2. Simpson's rule

To locate the minimum of (12) reasonably fast, we evaluate $\mathcal{S}_*(u_\theta)$ at $\theta = \pi/4$ with a precision of 10^{-9}. A higher order integration rule is therefore desirable. The one we use is the tensor product of two Simpson's rules. Recall that the error of the one-dimensional Simpson's rule applied to a function $f \in C^4([a, a+h]; \mathbb{R})$ is given by [7, p. 288]:

$$\frac{1}{h}\int_a^{a+h} f(x)\,\mathrm{d}x - \frac{1}{6}\big(f(a) + 4f(a+\tfrac{1}{2}h) + f(a+h)\big) = -\frac{1}{2880}h^4 f^{(4)}(\xi)$$

for some $\xi \in \,]a, a+h[$. This yields a two-dimensional rule on a rectangle P:

$$\operatorname{Simpson}(f) = \frac{1}{36}\big(f_{00} + f_{20} + f_{02} + f_{22} + 4(f_{10} + f_{01} + f_{21} + f_{12}) + 16 f_{11}\big)$$

where $f_{ij} := f(a_1 + ih_1/2, a_2 + jh_2/2)$, and the following error bound:

$$\left|\frac{1}{|P|}\int_P f - \operatorname{Simpson}(f)\right| \le \frac{1}{2880}\big(h_1^4\|\partial_x^4 f\|_\infty + h_2^4\|\partial_y^4 f\|_\infty\big). \tag{15}$$

From this, an interval version is easily derived with convergence speed

$$\operatorname{width}[\mathcal{I}_P] = O(h^4) = O\big(\tfrac{1}{n^2}\big) = O\big(\tfrac{1}{16^d}\big).$$

4.3.3. Integrating a non-smooth function

For the purpose of integrating $u_\theta^2 \log|u_\theta|$, Simpson's rule cannot be used over the whole Ω. Indeed, whenever $u_\theta = 0$, the fourth derivatives of the function are unbounded. Thus, the following algorithm is used to estimate $\frac{1}{|\Omega|}\int_\Omega f(x,y)\,\mathrm{d}(x,y)$.

`mean_integ`$([f], [\partial_x^4 f], [\partial_y^4 f], [x_{\text{left}}], [x_{\text{right}}], [y_{\text{left}}], [y_{\text{right}}], d, \text{tol})$

- Let $[x] := \big[\text{low}([x_{\text{left}}]), \text{high}([x_{\text{right}}])\big]$ and $[y] := \big[\text{low}([y_{\text{left}}]), \text{high}([y_{\text{right}}])\big]$.
- For $i \in \{x, y\}$, compute $[\underline{d}_i, \overline{d}_i] := [\partial_i^4 f]\big([x],[y]\big)$ and bound the derivatives on the box $[x]\times[y]$ with $\|\partial_i^4 f\|_\infty \le \max\{-\underline{d}_i, \overline{d}_i\}$. Then compute $\overline{e}$ using the right hand side of (15), rounding all operations towards $+\infty$.
- Let $[x_{\text{mid}}]$ (resp. $[y_{\text{mid}}]$) be the middle points of $[x_{\text{left}}]$ and $[x_{\text{right}}]$ (resp. $[y_{\text{left}}]$ and $[y_{\text{right}}]$).
- Let $[\mathcal{I}^0]$ be defined by $\big(\text{Simpson}(f) + [-\overline{e}, \overline{e}]\big) \cap [f]\big([x],[y]\big)$.
- If $d \le 0 \vee \overline{e} \le \frac{1}{2}\texttt{tol}$, then return $[\mathcal{I}^0]$;
  else, call recursively `mean_integ` on the four sub-squares (see Fig. 4.3) with $d := d - 1$. This will return $[\mathcal{I}_1^0], \dots, [\mathcal{I}_4^0]$. Finally, in accordance with (13), return $\big([\mathcal{I}_1^0] + \cdots + [\mathcal{I}_4^0]\big)/4$.

Note that, in the above algorithm, the parameter `tol` is used as an indication of when to stop the recursion, the final accuracy being actually given by the width of the returned interval. Fig. 4.4 depicts how this algorithm subdivides the square $[0,1]^2$ to compute $\int_{[0,1]^2} u_\theta^2 \log|u_\theta|$ with $\theta = \pi/4$ to achieve a tolerance of 0.05 with depth $d = 5$. A sub-square is colored whenever the higher order formula was used on it.

The function `integ(`$[f], [\partial_x^4 f], [\partial_y^4 f], d,$ `tol)` mentioned at the beginning of section 4 is simply defined as `mean_integ(`$[f], [\partial_x^4 f], [\partial_y^4 f]$, [0,0], [1,1], [0,0], [1,1], d, `tol)`.

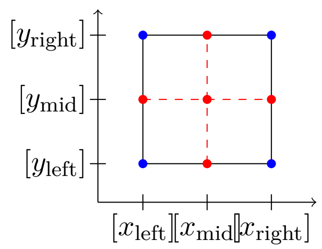

Figure 4.3: Square subdivisions.

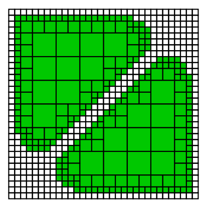

Figure 4.4: Using the appropriate integration rule with Simpson's rule.

Now that verified integration is available, it is easy to define `entropy(`$[\theta], d,$ `tol)` that provides an enclosure $[g]([\theta])$ for $g([\theta])$. Indeed, given an interval $[\theta]$, it suffices to use `integ` with $[f]$ being the lexical closure $[F]([\theta],\cdot,\cdot)$ where $[F] : \mathbb{IR}\times\mathbb{IR}\times\mathbb{IR} \to \mathbb{IR}$ is an interval extension of the map $F : (\theta, x, y) \mapsto -(u_\theta(x,y))^2 \cdot \log|u_\theta(x,y)|$. The routine `integ` also requires fourth order derivatives $[\partial_x^4 f]$ and $[\partial_y^4 f]$. Given that the map F is explicit, we compute $\partial_x^4 F$ and $\partial_y^4 F$ by hand[2] and use these expressions to derive interval extensions $[\partial_i^4 F] : \mathbb{IR}\times\mathbb{IR}\times\mathbb{IR} \to \mathbb{IR}$, $i \in \{x, y\}$ of these functions. As for $[f]$, we pass the lexical closures $[\partial_i^4 F]([\theta],\cdot,\cdot)$ to `integ`.

[2]We did not use automatic differentiation as the functions are simple enough and manual computations enable us to better factorize common sub-expressions.

4.4. Locating the minimum

As we described before, our task is reduced to minimize the entropy function $\mathcal{S}_*(u_\theta)$ on the sphere $\mathbb{S}$. Due to symmetry considerations, it is enough to prove that the function $g(\theta) = -\int_\Omega u_\theta^2 \log|u_\theta| \,\mathrm{d}x\,\mathrm{d}y$ defined on $[0, \frac{\pi}{4}]$ attains its minimum at $\theta = \frac{\pi}{4}$, where u_θ is defined in (10). The symmetries also imply that $\frac{\pi}{4}$ is a critical point of g.

In this section, we determine a "small" interval around $\pi/4$ in which the minimum is guaranteed to lie. To that effect, we use the interval extension $[g] : \mathbb{IF} \to \mathbb{IF}$ of g which was constructed in the previous section. We first compute an interval $[\underline{m}, \overline{m}] := [g]\big([\pi]/4\big) \in \mathbb{IF}$ to have bounds on the value $g(\pi/4)$, where $[\pi] \in \mathbb{IF}$ denotes a small interval containing π. Then, starting from the left of $[0, \pi/4]$, we try to remove intervals in which we are sure the minimum does not lie. More precisely, an interval $[x]$ is discarded if the lower bound of $[g]\big([x]\big)$ (which contains all $g(\xi)$ for $\xi \in [x]$) is greater than $\overline{m}$. This simple idea leads to the following algorithm.

$\texttt{exclude}([g], \underline{x}, \overline{x}, \texttt{step}, n)$
  $\texttt{if}\ n \le 0\ \texttt{then return}\ [\underline{x}, \overline{x}]$
  $\texttt{else}$
    $\texttt{let}\ x_{\texttt{next}} = \min\{\underline{x} + \texttt{step}, \overline{x}\}$
    $\texttt{let}\ [\underline{y}, \overline{y}] = [g]\big([\underline{x}, \overline{x}]\big)$
    $\texttt{if}\ \underline{y} > \overline{m}\ \texttt{then return exclude}([g], x_{\texttt{next}}, \overline{x}, \texttt{step}, n-1)$
    $\texttt{else return exclude}([g], \underline{x}, \overline{x}, \frac{1}{2}\texttt{step}, n-1)$

Remark that n is the number of iterations to be performed. In practice, the depth (and hence the precision) with which the integral in $[g]$ is computed is gradually increased as $\underline{x}$ gets closer to $\pi/4$. This function is called with $\underline{x} = 0$ and $\overline{x} = \mathrm{high}([\pi]/4)$ to obtain the location of the minimum. This is illustrated in Fig. 4.5 for $n = 80$ where the gray rectangles are $[x] \times [g]\big([x]\big)$ with $[x]$ excluded.

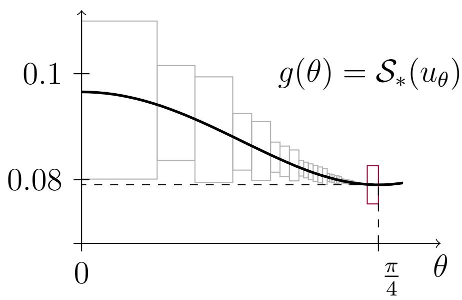

Figure 4.5: The exclusion procedure for locating the minimum of $g(\theta)$.

4.5. Analysis of concavity

The final step is to analyze the concavity of g in the interval I returned by `exlude` to guarantee that $\theta = \frac{\pi}{4}$ is the unique critical point in I, hence the only minimum.

Proposition 4.1. *The functional $\mathcal{J}_* : E_2 \to \mathbb{R}$ defined by* (3) *is of class C^2 on $E_2 \setminus \{0\}$ and its second derivative $\partial_u^2 \mathcal{J}_*$ at $u \in E_2 \setminus \{0\}$ is the bilinear map given by*

$$\partial_u^2 \mathcal{J}_*(u)[u_1, u_2] = -\lambda_2 \int_\Omega u_1 u_2 (1 + \log|u|) \,\mathrm{d}x\,\mathrm{d}y. \qquad (16)$$

The function $g : \mathbb{R} \to \mathbb{R}$ defined by (12) *with u_θ being given in* (10) *is C^2 and*

$$g''(\theta) = -2\left(\frac{1}{4} + g(\theta) + \int_\Omega |u'_\theta|^2 \log|u_\theta| \,\mathrm{d}x\,\mathrm{d}y\right) \qquad \text{where } ' = \frac{\partial}{\partial\theta}. \tag{17}$$

Proof. 1. The first derivative of $\partial_u \mathcal{J}_*$ exists and is given by (4). It is sufficient to show that the map $(t,\theta) \mapsto \partial_u \mathcal{J}_*(tu_\theta)$ is C^1. Indeed, first notice that $\mathcal{S} :]0,+\infty[\times \mathbb{R} \to E_2 \setminus \{0\} : (t,\theta) \mapsto tu_\theta$ is surjective. Second, one readily shows that, for all (t,θ), the differential $\partial_{(t,\theta)}\mathcal{S}(t,\theta)$ is an one-to-one linear map so the inverse function theorem says that $\mathcal{S}$ is locally invertible as a C^1-map which proves the differentiability of $\partial_u \mathcal{J}_*$ on $E_2 \setminus \{0\}$. Since $\partial_u \mathcal{J}_*(tu)[v] = t\,\partial_u \mathcal{J}_*(u)[v] - \lambda_2 t \log t \int uv$, the derivative w.r.t. t clearly exists and is continuous. It remains to show that the derivative w.r.t. θ exists, is given by

$$\partial_\theta\big(\partial_u \mathcal{J}_*(tu_\theta)\big)[v] = -\lambda_2 \int_\Omega v\, tu'_\theta(1 + \log|tu_\theta|) \,\mathrm{d}x\,\mathrm{d}y \tag{18}$$

and is continuous. Given the above formula for $\partial_u \mathcal{J}_*(tu)$, it is enough to prove (18) for $t = 1$. From the form of both partial derivatives as well as $\partial_{(t,\theta)}\mathcal{S}$, one readily deduces (16).

2. Remark that the function $f : \mathbb{R} \to \mathbb{R}$ defined in (12) is C^1 and u_θ and u'_θ are bounded functions of (x,y). Thus Lebesgue dominated convergence Theorem implies that g is continuously differentiable and

$$g'(\theta) = -\int_\Omega \partial_t f(u_\theta) u'_\theta \,\mathrm{d}x\,\mathrm{d}y = -\int_\Omega u'_\theta u_\theta (1 + 2\log|u_\theta|) \,\mathrm{d}x\,\mathrm{d}y.$$

A direct computation shows that $\int_\Omega u'_\theta u_\theta = 0$, so g' may be rewritten as

$$g'(\theta) = -2\int_\Omega u'_\theta u_\theta \log|u_\theta| \,\mathrm{d}x\,\mathrm{d}y. \tag{19}$$

Notice this is essentially $\partial_u \mathcal{J}_*(u_\theta)[u'_\theta]$ so, if we establish (18), (17) will easily follow by the chain rule and the facts: $u''_\theta = -u_\theta$ and $\|u'_\theta\|_{L^2(\Omega)} = 1/2$.

3. We now establish (18) for $t = 1$. Using the mean value Theorem, the differential quotient may be written as

$$\frac{\partial_u \mathcal{J}_*(u_{\theta+\varepsilon})[v] - \partial_u \mathcal{J}_*(u_\theta)[v]}{\varepsilon} = -\lambda_2 \int_\Omega v u'_{\theta_\varepsilon}(1 + \log|u_{\theta_\varepsilon}|) \,\mathrm{d}x\,\mathrm{d}y$$

for some $\theta_\varepsilon \in]\theta, \theta+\varepsilon[$, so the theorem will be proved if we show that the map

$$\theta \mapsto \int_\Omega \varphi(\theta,x,y) \,\mathrm{d}x\,\mathrm{d}y, \quad \text{where } \varphi(\theta,x,y) := v\, u'_\theta\,(1 + \log|u_\theta|), \tag{20}$$

is well defined and continuous for all $v \in E_2$. To prove this, the Lebesgue dominated convergence Theorem cannot be applied directly because, as θ varies, the nodal lines $\{(x,y) : u_\theta(x,y) = 0\}$ cover a non-zero measure set making any dominating function non-integrable.

We will show the continuity of (20) for $\theta \in [0, \pi/4]$ as, thanks to symmetries (see p. 370), the remaining range of θ is very similar. First note that u_θ may be written as $u_\theta(x,y) = z(x,y)\, w_\theta(x,y)$ with

$$z(x,y) := 2\sin(\pi x)\sin(\pi y) \quad \text{and} \quad w_\theta(x,y) := \cos\theta\cos(\pi y) - \sin\theta\cos(\pi x).$$

The nodal line of u_θ is the subset of Ω where w_θ vanishes. For $\theta \in [0, \pi/4]$, it is the graph of a function $r_\theta : [0,1] \to [0,1] : x \mapsto r_\theta(x)$ defined implicitly by

$$\forall x \in [0,1], \qquad \cos\theta \cos(\pi r_\theta(x)) - \sin\theta \cos(\pi x) = 0. \tag{21}$$

Let us split Ω as the union of Ω^+ and Ω^- where

$$\Omega^+ := \{(x,y) : x \in]0,1[\text{ and } 0 < y \le r_\theta(x)\},$$
$$\Omega^- := \{(x,y) : x \in]0,1[\text{ and } r_\theta(x) < y < 1\}.$$

Thanks to the oddness of u_θ and u'_θ w.r.t. to the center of Ω, we only have to consider Ω^+, the case of Ω^- being similar. One has

$$\int_{\Omega^+} \varphi(\theta,x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_0^1 \int_0^{r_\theta(x)} \varphi(\theta,x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_0^1 \int_0^1 \varphi(\theta,x,sr_\theta(x)) r_\theta(x)\,\mathrm{d}s\,\mathrm{d}x.$$

If we find an integrable function (independent of θ) dominating $\varphi(\theta,x,sr_\theta(x))r_\theta(x)$, the proof will be done. Note that, because $\|u'_\theta\|_\infty \le 2$,

$$\bigl|\varphi(\theta,x,sr_\theta(x))r_\theta(x)\bigr| \le 2\|v\|_\infty\bigl(1 + \bigl|\log|u_\theta(x,sr_\theta(x))|\bigr|\bigr)$$

so it suffices to find an integrable function dominating $\bigl|\log|u_\theta(x,sr_\theta(x))|\bigr|$. Now, since $u_\theta \ge 0$ on Ω^+ and $u_\theta \le 2$, it remains to find a lower bound of u_θ whose logarithm is integrable. Observe that $\log u_\theta = \log z + \log w_\theta$. Since $z(x,y) \ge 2x(1-x)y(1-y)$, $\log z$ is integrable on Ω, whence on Ω^+. For w_θ, using (21), one has

$$\begin{aligned}
w_\theta(x,sr_\theta(x)) &= \cos(\theta)\bigl(\cos(\pi s r_\theta(x)) - \cos(\pi r_\theta(x))\bigr) \\
&= \cos(\theta)\pi r_\theta(x) \int_s^1 \sin(\pi\sigma r_\theta(x))\,\mathrm{d}\sigma \\
&\ge \cos(\theta)\pi r_\theta(x) \int_s^1 \pi\,\sigma r_\theta(x)(1-\sigma r_\theta(x))\,\mathrm{d}\sigma && (\text{as } \sigma r_\theta \in [0,1]) \\
&= \cos(\theta)\pi^2 r_\theta^2(x)(1-s)\Bigl(\frac{1+s}{2} - \frac{1+s+s^2}{3} r_\theta(x)\Bigr) \\
&\ge \cos(\theta)\pi^2 r_\theta^2(x)(1-s)\Bigl(\frac{1+s}{2} - \frac{1+s+s^2}{3}\Bigr) && (\text{as } r_\theta \le 1) \\
&\ge \frac{1}{6}\cos(\theta)\pi^2 r_\theta^2(x)(1-s)^2 && (\text{for } s \in [0,1]) \\
&\ge \frac{\sqrt{2}}{12}\pi^2 \min\{x^2, \tfrac{1}{4}\}\,(1-s)^2 && (\text{as } \theta \in [0, \tfrac{\pi}{4}]).
\end{aligned}$$

The logarithm of this last bound is easily seen to be an integrable function of the argument $(x,s) \in]0,1[^2$, thereby concluding the proof. □

In view of (17), to check the concavity of g, the following condition must be verified:

$$g''(\theta) > 0 \Leftrightarrow h(\theta) > \tfrac{1}{4} + g(\theta), \qquad \text{where } h(\theta) := -\int_\Omega |u'_\theta|^2 \log|u_\theta|\,\mathrm{d}x\,\mathrm{d}y. \tag{22}$$

Because the integrand of h is singular whenever $u_\theta = 0$, when evaluated with interval arithmetic, [log] will return intervals of the form $[-\infty, \overline{y}] \in \mathbb{IF}$ which will contaminate the sum computing the integral. The important remark however is that the singularity helps h to be large and so does not need to be controlled. Interval extensions $[h]$ of h will return intervals of the form $[\underline{y}, +\infty]$ and, to check condition (22), one will verify that $\underline{y} > \mathrm{high}\bigl(1 + [g]([\theta])\bigr)$.

4.6. Non-degeneracy

Proposition 4.2. *If θ_* is a non-degenerate critical point of g, then $u_* = t^*_{u_{\theta_*}} u_{\theta_*}$, where $t^*_u > 0$ is defined by* (6)*, is a non-degenerate critical point of $\mathcal{J}_*$.*

Proof. Let θ_* be a critical point of g. On one hand, by definition (6) of $t^*_{u_{\theta_*}}$, one has $\partial_u \mathcal{J}_*(u_*)[u_{\theta_*}] = 0$. On the other hand, (19) implies that $\partial_u \mathcal{J}_*(u_{\theta_*})[u'_{\theta_*}] = 0$. Since $\partial_u \mathcal{J}_*(u_{\theta_*})$ vanishes in two orthogonal directions, u_* is a critical point of $\mathcal{J}_*$.

In view of Proposition 4.1,

$$\partial^2_u \mathcal{J}_*(u_*)[u_*, v] = \partial^2_u \mathcal{J}_*(u_*)[u_*, v] - \partial_u \mathcal{J}_*(u_*)[v] = -\lambda_2 \int_\Omega u_* v.$$

In particular, $\partial^2_u \mathcal{J}_*(u_*)[u_*, u_*] = -\lambda_2 \int_\Omega u_*^2 < 0$ and $\partial^2_u \mathcal{J}_*(u_*)[u_*, u'_{\theta_*}] = 0$ (recalling that $\int u_{\theta_*} u'_{\theta_*} = 0$). Using again the definition (6) of $t^*_{u_{\theta_*}}$, one obtains

$$\partial^2_u \mathcal{J}_*(u_*)[u'_{\theta_*}, u'_{\theta_*}] = \frac{\lambda_2}{2} g''(\theta_*) \neq 0.$$

As a consequence, the bilinear form $\partial^2_u \mathcal{J}_*(u_*)$ is non-degenerate. □

Consequently, our main result is a consequence of the following proposition.

Proposition 4.3. *Let $\Omega =]0,1[^2$. Then the minimization problem* (11) *is solved by $\theta = \pi/4$ and the minimum is non-degenerate. Therefore the minimum of $\mathcal{J}_*$ on $\mathcal{N}_*$ is achieved by a multiple of $u_{\pi/4}$ and it is a non-degenerate critical point.*

5. Proof: putting it all together

In this section, we finally run the machinery described above on a computer and give the numerical results. We will sometimes write intervals as, for example, $[1.74_{67}, 1.74_{81}]$ instead of $[1.7467, 1.7481]$ to make easier to spot the common digits.

First, to locate an interval $[\theta_{\min}] \ni \frac{\pi}{4}$ where the entropy function (12) attains its minimum, an small interval $[\pi]/4$ containing $\pi/4$ is determined and $[g]\big([\pi]/4\big)$ is estimated with good precision by calling `entropy(`$[\pi]/4$`, depth, tol)` with `depth` $=$ 13 and `tol` $= 10^{-9}$. The result is

$$[g]\big([\pi]/4\big) = [0.07905032_{01}, 0.07905032_{53}].$$

Then `exclude` with $n = 100$ is run on $[g]$ and outputs the desired interval $[\theta_{\min}]$:

$$[\theta_{\min}] = [0.7_{51367}, 0.7_{85398}].$$

Finally, in order to verify that $\frac{\pi}{4}$ is the unique critical point in $[\theta_{\min}]$, we analyze the second derivative of (12) in $[\theta_{\min}]$ by checking condition (22). Note that, if this condition is fulfilled, our task is done. We evaluate $[h]$ by using `integ` with a depth $d = 8$ and tolerance `tol` $= 10^{-3}$ and get:

$$[h]\big([\theta_{\min}]\big) = [0.35179120, +\infty].$$

Computing the right hand side of (22) yields:

$$\tfrac{1}{4} + [g]\big([\theta_{\min}]\big) = [0.324879, 0.333196].$$

Clearly, these results show that condition (22) is satisfied, thereby proving Theorem 2.1.

References

[1] A. Aftalion, F. Pacella: *Qualitative properties of nodal solutions of semilinear elliptic equations in radially symmetric domains*, C. R. Math. 339/5 (2004) 339–344.

[2] A. Ambrosetti, P. H. Rabinowitz: *Dual variational methods in critical point theory and applications*, J. Functional Analysis 14/4 (1973) 349–381.

[3] T. Bartsch, T. Weth, M. Willem: *Partial symmetry of least energy nodal solutions to some variational problems*, J. d'Analyse Math. 96/1 (2005) 1–18.

[4] D. Bonheure, V. Bouchez, C. Grumiau, J. Van Schaftingen: *Asymptotics and symmetries of least energy nodal solutions of Lane-Emden problems with slow growth*, Comm. Contemp. Math. 10/4 (2008) 609–631.

[5] A. Castro, J. Cossio, J. M. Neuberger: *A sign-changing solution for a superlinear Dirichlet problem*, Rocky Mountain J. Math. 27/4 (1997) 1041–1053.

[6] E. N. Dancer: *Real analyticity and non-degeneracy*, Math. Ann. 325/2 (2003) 369–392.

[7] P. J. Davis, P. Rabinowitz: *Methods of Numerical Integration*, 2nd ed., Computer Science and Applied Mathematics, Academic Press, Orlando (1984).

[8] S. Galdino: *Interval integration revisited*, Open J. Appl. Sciences 2 (2012) 108–111.

[9] B. Gidas, W. M. Ni, L. Nirenberg: *Symmetry and related properties via the maximum principle*, Comm. Math. Phys. 68/3 (1979) 209–243.

[10] M. Grossi: *On the shape of solutions of an asymptotically linear problem*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 8/3 (2009) 429–449.

[11] E. Loh, G. W. Walster: *Rump's example revisited*, Reliab. Comput. 8/3 (2002) 245–248.

[12] G. Mayer: *Interval Analysis and Automatic Result Verification*, De Gruyter Studies in Mathematics 65, De Gruyter, Berlin (2017).

[13] R. E. Moore, R. Baker Kearfott, M. J. Cloud: *Introduction to Interval Analysis*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, (2009).

[14] Z. Nehari: *On a class of nonlinear integral equations*, Math. Zeitschrift 72 (1959/60) 175–183.

[15] K. Petras: *Principles of verified numerical integration*, J. Comput. Appl. Math. 199/2 (2007) 317–328.

[16] S. M. Rump: *Algorithms for verified inclusions: Theory and practice*, in: *Reliability in Computing*, Perspect. Comput. 19 (1988) 109–126.

[17] S. A. Smolyak: *Quadrature and interpolation formulas for tensor products of certain classes of functions*, Soviet Math. Dokl. 4 (1963) 240–243.