Optimal Control results for the sine–Gordon equation with and without damping

Pitágoras P. de Carvalho, Felipe W. Chaves-Silva, Gilcenio R. Sousa-Neto, Franciane B. Vieira

Abstract In this note, we study an optimal control problem for the sine-Gordon equation, with emphasis on the effect of the damping term on control performance. First, a theoretical framework ensuring the well-posedness of the problem and the consistency of the control strategy is established. Next, a numerical approach combining finite element and finite difference methods is employed, together with a Picard iterative scheme to ensure convergence. The interplay between control and damping is then analyzed through computational simulations.

1 Introduction

Natural systems often involve many degrees of freedom, and their dynamics can usually be described by Partial Differential Equations (PDEs). In this setting, it is natural to consider how such systems can be influenced by suitable control actions. In Control Theory, the objective is to design controls that steer the system toward a desired state within a finite time (see [14, 11, 20, 16]). To achieve this goal, a widely used approach is to formulate the problem as the minimization of a cost functional that measures the distance between the system state and a prescribed target, while penalizing the control effort. This formulation provides a rigorous framework for both theoretical analysis and numerical approximation.

In this note, we study an optimal control problem for the sine-Gordon equation (sGE), which is a nonlinear hyperbolic Partial Differential Equation arising in several areas of mathematical physics. One remarkable feature of this equation is

Pitágoras Pinheiro de Carvalho ✉
Universidade Estadual do Piauí, Teresina-Piauí-Brazil, e-mail: pitagorascarvalho@gmail.com

Felipe Wallison Chaves-Silva
Universidade Federal da Paraíba, João Pessoa-Paraíba-Brazil, e-mail: fchaves@mat.ufpb.br

Gilcenio Rodrigues de Sousa Neto
Universidade Federal do Piauí, Teresina-Piauí-Brazil, e-mail: gilceniorodrigues@ufpi.edu.br

Franciane de Brito Vieira
Universidade Federal do Piauí, Teresina-Piauí-Brazil, e-mail: franciane@ufpi.edu.br

the existence of localized and stable wave solutions that preserve their shape during propagation and interact elastically upon collision (see [8, 15, 17]). Beyond its mathematical interest, the sine-Gordon equation appears in a variety of physical contexts, including magnetic crystals, superconductivity, condensed matter physics, field theory, and nonlinear optics.

In order to formulate the problem, let y_0 and y_1 denote the initial data, f the forcing term, and v the control. We denote by $y = y(x,t;v)$ the corresponding controlled state, which satisfies

$$\begin{cases} y_{tt} + \beta y_t - \Delta y + \sin(y) = f + v\,\mathbb{1}_{\mathcal{O}} & \text{in } \; Q := \Omega \times (0,T), \\ \dfrac{\partial y}{\partial n} = 0 & \text{on } \Sigma := \partial\Omega \times (0,T), \\ y(x,0) = y_0(x), \quad y_t(x,0) = y_1(x) & \text{in } \; \Omega. \end{cases} \tag{1}$$

Here, $\Omega \subset \mathbb{R}^N$ $(N = 1,2,3)$ is a regular, connected domain, and $\mathcal{O} \subset \Omega$ denotes the region where the control acts. We also introduce $\mathcal{O}_d \subset \Omega$, the target region where the state is required to remain close to a prescribed profile y_d. Finally, $T > 0$ denotes the final time and $\beta \geq 0$ is a parameter.

In system (1), the damping term βy_t plays an important role in the dynamics of the system (see, for instance, [2, 5]). When $\beta = 0$, the equation is undamped and exhibits persistent oscillations; for $\beta > 0$, energy is dissipated, and the convergence to equilibrium becomes faster as β increases.

The optimal control problem for system (1) consists in finding $v \in L^2(\mathcal{O}\times(0,T))$ that minimizes a quadratic cost functional, defined by

$$J(v) = \frac{1}{2}\iint_{\mathcal{O}_d\times(0,T)} |y(x,t;v) - y_d(x,t)|^2\, dxdt + \frac{\mu}{2}\iint_{\mathcal{O}\times(0,T)} |v(x,t)|^2\, dxdt, \tag{2}$$

where $\mu > 0$ is a regularization parameter (see [16, Chapter 18]). The corresponding minimization problem can then be expressed as

$$J(v) = \min_{\hat{v}\in L^2(\mathcal{O}\times(0,T))} J(\hat{v}). \tag{3}$$

In this formulation, the first term in J aims to keep the controlled state as close as possible to the desired profile within the target region, while the second term penalizes the magnitude of the applied control.

Our objective is to carry out numerical simulations for the optimal control problem associated with the sine-Gordon equation (1), focusing on the behavior of the cost functional J at the optimal control. We also investigate the combined influence of the regularization parameter $\mu > 0$ and the damping coefficient $\beta \geq 0$ on the behavior of this functional.

The numerical analysis follows the methodology of R. Glowinski et al. [12], who studied the control of various PDEs, and P. P. de Carvalho et al. [9], who specifically focused on wave-type equations. Our problem, however, introduces two key differences: new boundary conditions and the inclusion of a damping term. To the best of our knowledge, these aspects have not yet been addressed in the numerical literature for the sine-Gordon equation.

2 Theoretical framework

To study the optimal control problem introduced in Section 1, the first step is to ensure the existence and uniqueness of solutions of the system (1). It is well-known that for any data

$$(f, y_0, y_1, y_d) \in \mathcal{S}_d := L^2(Q) \times H^1(\Omega) \times L^2(\Omega) \times L^2(O_d \times (0,T)), \tag{4}$$

and any control $v \in L^2(O \times (0,T))$, there exists a unique solution

$$y = y(v) \in C([0,T]; H^1(\Omega)) \cap C^1([0,T]; L^2(\Omega))$$

of the sine-Gordon equation (1) (see, e.g., [19, Chapter IV, Section 2]). Consequently, the cost functional J defined in (2) is well-defined.

For each such solution y of (1), we introduce the associated (adjoint) system

$$\begin{cases} \phi_{tt} - \beta\phi_t - \Delta\phi + \cos(y)\phi = (y - y_d)\mathbb{1}_{O_d} & \text{in } Q, \\ \dfrac{\partial \phi}{\partial n} = 0 & \text{on } \Sigma, \\ \phi(x,T) = \phi_t(x,T) = 0 & \text{in } \Omega. \end{cases} \tag{5}$$

A straightforward computation shows that the Gâteaux derivative of J satisfies

$$J'(v) \equiv 0 \quad \Leftrightarrow \quad v = -\frac{1}{\mu}\phi(y(v)).$$

Hence, the solution $\phi = \phi(y(v))$ of (5) characterizes the minimizers of (3), as these correspond to the zeros of J'. This not only clarifies the theoretical structure of the optimal control problem but also provides a practical starting point for iterative algorithms: the minimizers can be obtained as fixed points of the operator

$$\Lambda : v \mapsto -\frac{1}{\mu}\phi(y(v)) .$$

The following result establishes the existence and uniqueness of the optimal control.

Theorem 1 *Let $(f, y_0, y_1, y_d) \in \mathcal{S}_d$. There exists*[1] $\mu_0 > 0$*, such that if $\mu > \mu_0$, then the following statements are true:*

1. There exists unique $v \in L^2(O \times (0,T))$ such that $J'(v) \equiv 0$. Moreover,

[1] Actually, the constant μ_0 depends on the data (f, y_0, y_1, y_d), the parameter β, the time T, and the domain Ω.

$$v = -\frac{1}{\mu}\phi \mathbb{1}_{O}, \tag{6}$$

where (y, ϕ) *is the unique solution of the independent system*

$$\begin{cases} y_{tt} + \beta y_t - \Delta y + \sin(y) = f - \dfrac{1}{\mu}\phi \mathbb{1}_{O} & in \ \ Q, \\ \phi_{tt} - \beta \phi_t - \Delta \phi + \cos(y)\phi = (y - y_d)\mathbb{1}_{O_d} & in \ \ Q, \\ \dfrac{\partial \phi}{\partial n} = \dfrac{\partial y}{\partial n} = 0 & on \ \Sigma, \\ y(x,0) - y_0(x) = y_t(x,0) - y_1(x) = \phi(x,T) = \phi_t(x,T) = 0 & in \ \ \Omega. \end{cases} \tag{7}$$

2. *The function* v *above is the unique solution for* (3)*, that is, the optimal control.*

A proof of Theorem 1 can be achieved as follows. First, we show that, for μ large enough, the operator Λ satisfies the hypothesis of Schauder's fixed-point theorem. This implies the existence of a fixed point $v \in L^2(O \times (0,T))$, which, by (6), satisfies $J'(v) \equiv 0$. Since the solution (y, ϕ) of system (7) is unique, we see that there is a unique v such that $J'(v) \equiv 0$.

To prove the second part of Theorem 1, we need to ensure that v is indeed a strict global minimum of J. Nevertheless, since we are working with a nonlinear equation, J may not be convex, and it is not obvious at first sight that v is a minimum. One way to prove this is to show that the second derivative of J in v is coercive (see, for instance, [6, Theorem 7.3-2]), i.e. there exists a constant $\alpha > 0$ such that

$$J''(v)(\xi,\xi) \geq \alpha \|\xi\|^2_{L^2(O\times(0,T))}, \qquad \forall \xi \in L^2(O \times (0,T)). \tag{8}$$

The estimate in (8) can be proved for μ sufficiently large by arguing, for instance, as Araruna et al. in [1].

In fact, the coercivity in (8) will imply that v is a local minimizer of J. However, since J satisfies the assumptions of [4, Corollary 3.23], it follows that it attains its minimum and, therefore, as J has a unique critical point, v must be the strict global minimizer of J.

3 Finite Dimensional Problem and Algorithmic

To implement numerical and computational approaches for the sine-Gordon equation, we first introduce a finite-dimensional approximation of (7), and then construct temporal and spatial discretizations, leading to the fully discrete formulation presented below.

For this part, we assume that $\Omega \subset \mathbb{R}^2$ is a bounded polygonal domain. Moreover, we suppose that the subsets O_d and O are polygonal and are unions of elements of the triangulation, as illustrated in Fig. 1. For simplicity, we introduce the notations for the control space

$$\mathcal{U} := L^2(0,T; L^2(O)),$$

and the state space

$$\mathcal{W} := \{w \in C([0,T]; H^1(\Omega));\ w_t \in C([0,T]; L^2(\Omega))\}.$$

Time discretization

For time discretization, let $\Delta t = T/M$, where M is a large positive integer, and define the uniform partition of $[0,T]$ by

$$t^m = m\Delta t, \qquad m = 0, 1, \dots, M,$$

so that $0 = t^0 < t^1 < \cdots < t^M = T$. We approximate the control space, in time, by

$$\mathcal{U}^{\Delta t} := \left(L^2(O)\right)^M,$$

whose elements are interpreted as piecewise constant, in time, controls.

Spatial Discretization

For spatial discretization, we consider $\mathcal{T}_h$ be a structured, uniform and shape-regular triangulation of $\Omega \subset \mathbb{R}^2$, constructed from a uniform partition of Ω into squares of side length h, each square being subdivided along one diagonal into two congruent right triangles. For each $h > 0$, we define the conforming finite element space to $\mathcal{W}$ using

$$\mathcal{W}_h := \left\{z \in C^0(\overline{\Omega}) : z|_K \in \mathbb{P}_1(K)\ ,\ \forall K \in \mathcal{T}_h\right\} \subset H^1(\Omega),$$

where $\mathbb{P}_1(K)$ denotes the space of polynomial functions of degree at most one and each $K \in \mathcal{T}_h$ represents a triangular element of the mesh. We denote by N_h the dimension of $\mathcal{W}_h$, which coincides with the number of vertices of $\mathcal{T}_h$. Similarly, the spatial approximate control space to $\mathcal{U}$, is given by

$$\mathcal{U}_h := \left\{z \in C^0(\overline{O}) : z|_K \in \mathbb{P}_1(K)\ ,\ \forall K \subset O\right\}.$$

The fully discrete version of $\mathcal{W}$ and $\mathcal{U}$ are given respectively by

$$\mathcal{W}_h^{\Delta t} := (\mathcal{W}_h)^M \quad \text{and} \quad \mathcal{U}_h^{\Delta t} := (\mathcal{U}_h)^M.$$

Then, the finite dimensional version for the cost functional (2) is given by

$$J_h^{\Delta t}(v) := \frac{1}{2}\iint_{O_d\times(0,T)} |y_h^{\Delta t} - y_d|^2\, dxdt + \frac{\mu}{2}\iint_{O\times(0,T)} |v|^2\, dxdt, \tag{9}$$

and the corresponding optimal control is approximated by the solution of the finite-dimensional condition $\dfrac{\partial J_h^{\Delta t}}{\partial v}(v) = 0$.

Remark 1 A fundamental issue in numerical optimal control concerns whether one should first discretize and then optimize, or optimize first and then discretize. Here, we adopt the former approach. An alternative viewpoint is presented in [18], with related analysis and further developments presented in [12] and [16].

3.1 Picard Method

The Picard iteration method consists of evaluating the nonlinear coefficients using the previous iterate, thereby generating, at each step, a linear and well-posed problem. For a nonlinear operator $F(y)$, the following linearization is considered:

$$F\left(y_h^{n,m+1}\right) \approx F\left(y_h^{n,m}\right) + F'\left(y_h^{n,m}\right)\left(y_h^{n,m+1} - y_h^{n,m}\right), \tag{10}$$

where the derivative $F'(y_h^{n,m})$ is treated as a fixed coefficient.

Accordingly, our nonlinear term is given by $F(y) = \sin(y)$, whose derivative is $F'(y) = \cos(y)$. The iterative scheme is carried out until the relative increment convergence criteria are satisfied. These criteria are evaluated at each iteration and are explicitly described in Step 4 of the fully discretized variational algorithm.

3.2 Discretized Optimal Control

From Theorem 1, the optimal control associated with (1) is characterized by (6)-(7). Now, we present the full-discretized version of the variational formulation of problem (6)-(7), that describe a numerical scheme for computational simulations.

Variational Full Discretized Algorithm

1. **Step 1.** Choose $v_h^0 \in \mathcal{U}_h^{\Delta t}$ and fix an approximation $\left(y_h^0, y_h^1\right) \in \mathcal{W}_h^0 \times \mathcal{W}_h^0$ to (y_0, y_1), where
$$y_h^1 \equiv y_h^0(x) + \Delta t y_h^1(x)\,;$$
2. **Step 2.** For given $n \geq 0$, $v_h^n \in \mathcal{U}_h^{\Delta t}$ and $F(y_h^{m+1})$ rewritten as in (10), compute the approximate states $y_h^{n,0}, y_h^{n,1}, \ldots$ as follows:
$$\begin{cases} y_h^{n,0} = y_{0,h} \;\text{ and }\; y_h^{n,1} = y_h^{n,0} + \Delta t y_h^1\,; \\ \displaystyle\int_\Omega \left[\left(\frac{y_h^{n,m+1} - 2y_h^{n,m} + y_h^{n,m-1}}{\Delta t^2}\right)z + \beta\left(\frac{y_h^{n,m+1} - y_h^{n,m}}{\Delta t}\right)z + \nabla y_h^{n,m+1} \cdot \nabla z \right. \\ \qquad\qquad\qquad\qquad \left. + F(y_h^{n,m+1})z\right] dx, \\ \quad = \displaystyle\int_\Omega f(x, t^{m+1})z \;+\; v_h^{n,m+1} \mathbb{1}_O z\, dx \\ \forall z \in \mathcal{W}_h,\; y_h^{n,m+1} \in \mathcal{W}_h^{\Delta t},\; m = 1, \ldots, M-1, \end{cases}$$

   and the approximate adjoint states $\phi_h^{n,M}, \phi_h^{n,M-1}, \ldots$ by solving

$$\begin{cases} \phi_h^{n,M} = 0, \quad \phi_h^{n,M-1} = 0, \\ \displaystyle\int_\Omega \Bigg[\left(\frac{\phi_h^{n,m+1} - 2\phi_h^{n,m} + \phi_h^{n,m-1}}{\Delta t^2} \right) z + \nabla \phi_h^{n,m-1} \cdot \nabla z - \beta \left(\frac{\phi_h^{n,m-1} - \phi_h^{n,m}}{\Delta t} \right) z \\ \qquad\qquad + \left(F'\left(y_h^{n,m}(x, t^{m-1}) \right) \phi_h^{n,m-1} \right) z \Bigg] \, dx \\ \quad = \displaystyle\int_{O_d} \left(y_h^{n,m}(x, t^{m-1}) - y_d(x, t^{m-1}) \right) z \, dx \\ \forall z \in \mathcal{W}_h, \ \phi_h^{n,m-1} \in \mathcal{W}_h^{\Delta t}, \quad m = M-1, \ldots, 1. \end{cases}$$

3. **Step 3.** Compute $v_h^{n,m+1} \in \mathcal{U}_h^{\Delta t}$, using

$$v_h^{n,m+1} = -\frac{1}{\mu} \phi_h^{n,m} \Big|_{O \times (0,T)}.$$

4. **Step 4.** Evaluate simultaneously the stop criteria, `ErrA` and `ErrB` using

$$\mathtt{ErrA} = \frac{\left\| y_h^{n,m+1} - y_h^{n,m} \right\|_{L^2(\Omega)}}{\left\| y_h^{n,m+1} \right\|_{L^2(\Omega)}} \quad \text{and} \quad \mathtt{ErrB} = \frac{\left\| v_h^{n,m+1} - v_h^{n,m} \right\|_{L^2(\Omega)}}{\left\| v_h^{n,m+1} \right\|_{L^2(\Omega)}}. \tag{11}$$

5. **Step 5.** Let $\varepsilon > 0$. If `ErrA`$\leq \varepsilon$ and `ErrB`$\leq \varepsilon$ occurs, break the n-iterate process and take

$$(y, v) := \left(y_h^{n,m+1}, \ v_h^{n,m+1} \right),$$

the last computed state. Otherwise, repeat the previous steps, with $n = n + 1$.

---

4 Numerical Simulations and Computational Analysis

In what follows, we present two-dimensional simulations using FreeFem++ [13], focused on the interaction of two orthogonal line solitons, for which theoretical and numerical reference solutions are available in the literature, such as [2], [3], [8], and [10], among others.

> Courant–Friedrichs–Lewy condition

We adopt the CFL (Courant-Friedrichs-Lewy) criterion (see [7], [12], and [16]), to guide our choice of spatial and temporal discretizations, ensuring the stability of our numerical schemes. In our setting, the CFL condition is expressed as

$$\Delta t \leq \frac{h}{\sqrt{2}}.$$

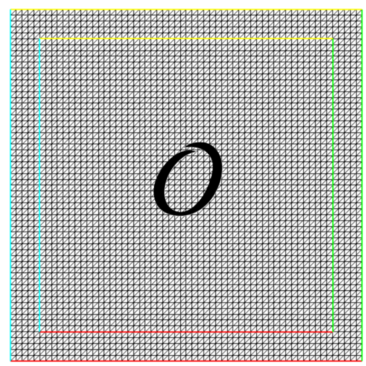
4.1 Optimal Control of Orthogonal Line Solitons

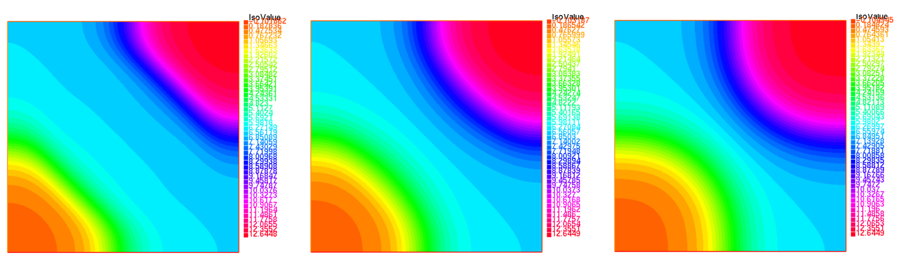
To preserve physical characteristics of the sGE in our simulations, we take $f \equiv 0$ and the initial conditions are chosen as in [2] or [3], that is

$$y_0 \equiv 4\,\mathrm{arctg}(\exp(x_1)) + 4\,\mathrm{arctg}(\exp(x_2)) \quad \text{and} \quad y_1 \equiv 0\,.$$

The parameter $y_{\mathcal{R}}$ will be used as the reference value for orthogonal line solitons at time $T = 4$, without the influence of control, as reported in [2] and [3], with the corresponding plots shown in Figure 2. The objective function is defined as $y_d(x_1, x_2) = C_d y_{\mathcal{R}}$, where $C_d = 0.9$. Thus, $C_d y_{\mathcal{R}}$ represents the desired percentage reduction in the wave magnitude evolution. The stop criterion adopted is $\varepsilon = 10^{-5}$.

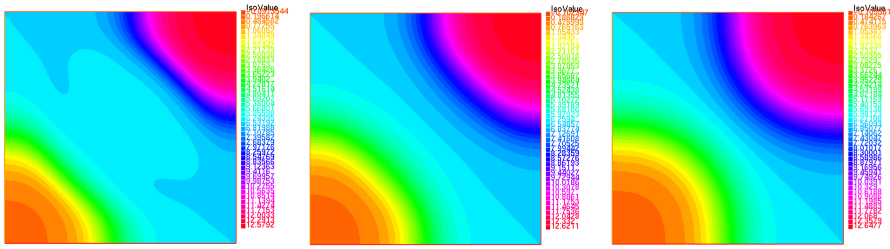
Fig. 1 We start our experiments with orthogonal line solitons in time interval $[0, 4]$. Here we consider the domains and subdomain given by $\Omega = [-6, 6] \times [-6, 6]$, $\mathcal{O} = [-5, 5] \times [-5, 5]$, $\mathcal{O}_d = [-6, 6] \times [-6, 6]$. The discretized parameters are fixed as: $\Delta t = 0.02$, and the structured triangular mesh was constructed with spacing $\Delta x = \Delta y \equiv h = 0.2$, that satisfy the CFL condition.





a) $\beta = 0$ b) $\beta = 0.5$ c) $\beta = 1.5$

Fig. 2: Reference to final state $y_R(x, 4)$.



a) $\beta \equiv 0$ b) $\beta \equiv 0.5$ c) $\beta \equiv 1.5$

Fig. 3: Controlled state $y(x, 4)$ for $C_d = 0.9$ in $\mu = 0.5$.

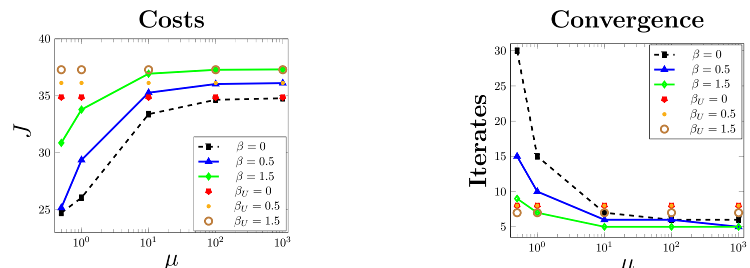


Fig. 4: Costs and Iterates in $\mu \in \{0.5, 1, 10, 100, 1000\}$. The parameter β_U represents the specific value of β corresponding to the uncontrolled system.

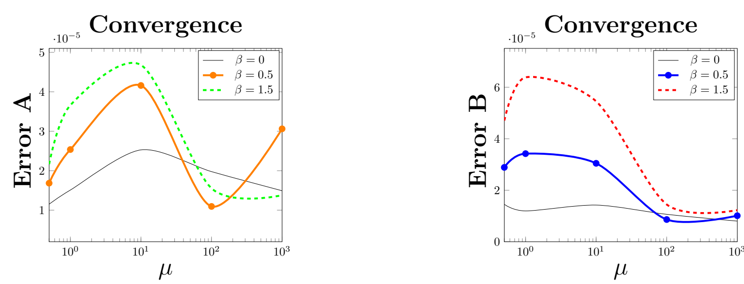


Fig. 5: Relative errors obtained according to the pre-established convergence criteria.

5 Conclusion

The numerical simulations provide a clear picture of how the regularization parameter μ and the dissipation strength β jointly influence the optimal control problem. Figure 2 shows the reference values for the sGE problem (Orthogonal line solitons). Figure 3 presents the final states under the action of the control term for $\mu = 0.5$, where the control effects are most pronounced.

Figure 4 emphasizes the influence of μ on the cost functional J and on the number of iterations required to reach the stopping criteria. This allows a clear comparison between controlled and uncontrolled states and enables a detailed analysis of the combined effect of μ and β. The simulations indicate that the control intensity is higher for small values of μ and decreases as μ increases and as β grows.

Within the Picard method, the number of iterations required to satisfy the convergence criteria decreases as μ increases, while an increase in the damping term β reduces the influence of the control on the system. Figure 5 illustrates the behavior of the convergence criteria, showing that the stability of the method is preserved throughout the iterative process for different combinations of μ and β.

References

1. F. D. Araruna, E. Fernández-Cara, and M. C. Santos. Stackelberg–nash exact controllability for linear and semilinear parabolic equations. *ESAIM: Control, Optimisation and Calculus of Variations*, 21(3):835–856, (2015).
2. J. Argyris, M. Haase, and J. C. Heinrich. Finite element approximation to two-dimensional sine-gordon solitons. *Computer methods in applied mechanics and engineering*, 86(1):1–26, (1991).
3. A. G. Bratsos. The solution of the two-dimensional sine-gordon equation using the method of lines. *Journal of Computational and Applied Mathematics*, 206(1):251–277, (2007).
4. H. Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, 2011.
5. P. L. Christiansen and P. S. Lomdahl. Numerical solution of 2 + 1 dimensional sine-gordon solitons. *Physica 2D*, pages 482–494, (1981).
6. P. G. Ciarlet, B. Miara, and J. M. Thomas. *Introduction to numerical linear algebra and optimisation*. Cambridge university press, (1989).
7. R. Courant, K. Friedrichs, and H. Lewy. Über die partiellen differenzengleichungen der mathematischen physik. *Mathematische annalen*, 100(1):32–74, (1928).
8. J. Cuevas-Maraver, P. G. Kevrekidis, and F. Williams. The sine-gordon model and its applications: From pendula and josephson junctions to gravity and. *High-energy Physics*, 10:263, (2014).
9. P. P. de Carvalho, E. Fernández-Cara, and J. B. Limaco. On the computation of nash and pareto equilibria for some bi-objective control problems for the wave equation. *Advances in Computational Mathematics*, 46(5):1–30, (2020).
10. E. G. Drazin and R. S. Johnson. *Solitons: An Introduction*. Cambridge Univ. Press, (1989).
11. R. Glowinski. *Numerical Methods for Control Problems in Partial Differential Equations*. Springer, New York, (2015).
12. R. Glowinski, J.L. Lions, and J. He. *Exact and Approximate Controllability for Distributed Parameter Systems: A Numerical Approach*. Cambridge University Press, Cambridge, (2008).
13. F Hecht. New development in freefem++. *Journal of numerical mathematics*, 20(3-4):251–266, (2012).
14. J. L. Lions. *Contrôle Optimal de Systèmes Gouvernés par des Équations aux Dérivées Partielles*. Gauthier-Villars, Paris, (1971).
15. T. Mohaupt. Introduction to string theory. In *Quantum Gravity: From Theory to Experimental Search*, pages 173–251. Springer, (2003).
16. A. Quarteroni. *Numerical Models for Differential Problems*. Springer, Cham, (2016).
17. R Rajaraman. *Solitons and instantons. An introduction to solitons and instantons in quantum field theory*. North-Holland, (1987).
18. A. M. Ramos, R. Glowinski, and J. Periaux. Pointwise control of the burgers equation and related nash equilibria problems: A computational approach. *Journal of Optimization Theory and Applications*, 112:499–516, (2002).
19. R. Temam. *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*. Springer, (1997).
20. E. Zuazua. *Control and Stabilization of Partial Differential Equations*. Springer, Berlin, (2007).