
Department of Applied Mathematics
Faculty of EEMCS



University of Twente
The Netherlands

P.O. Box 217
7500 AE Enschede
The Netherlands

Phone: +31-53-4893400

Fax: +31-53-4893114

Email: memo@math.utwente.nl
www.math.utwente.nl/publications

Memorandum No. 1721

**Discontinuous Galerkin method for
linear free-surface gravity waves**

J.J.W. VAN DER VEGT AND S.K. TOMAR

April, 2004

ISSN 0169-2690

Discontinuous Galerkin Method for Linear Free-Surface Gravity Waves

J.J.W. van der Vegt and S.K. Tomar

Department of Applied Mathematics,
University of Twente, P.O. Box 217,
7500AE, Enschede, Netherlands.

E-mails: {j.j.w.vandervegt, s.k.tomar}@math.utwente.nl

Abstract

In this paper, we discuss a discontinuous Galerkin finite element method for linear free surface gravity waves. We prove that the algorithm is unconditionally stable and does not require additional smoothing or artificial viscosity terms in the free surface boundary condition to prevent numerical instabilities on a non-uniform mesh. A detailed error analysis of the full time-dependent algorithm is given, showing that the error in the wave height and velocity potential in the L^2 -norm is in both cases of optimal order and proportional to $O(\Delta t^2 + h^{p+1})$, without the need for a separate velocity reconstruction, with p the polynomial order, h the mesh size and Δt the time step. The error analysis is confirmed with numerical simulations. In addition, a Fourier analysis of the fully discrete scheme is conducted which shows the dependence of the frequency error and wave dissipation on the time step and mesh size. The algebraic equations for the DG discretization are derived in a way suitable for an unstructured mesh and result in a symmetric positive definite linear system. The algorithm is demonstrated on a number of model problems, including a wave maker, for discretizations with accuracy ranging from second to fourth order.

Key words: discontinuous Galerkin method, gravity waves, a-priori error analysis, elliptic partial differential equations

AMS Subject Classification: 35J05, 65N12, 65N15, 65N30, 76B07, 76B15, 76M10

1 Introduction

For many applications with free surface gravity waves it is sufficient to consider the fluid as incompressible, inviscid and irrotational. This greatly simplifies the problem, resulting in a Laplace equation with a non-linear free surface boundary condition, but the numerical solution of these equations still requires a delicate balance between the stability and accuracy of the discretization. Otherwise, spurious waves and serious errors in the wave propagation will occur. The theoretical analysis of free surface gravity waves is fairly complete, for overviews see [17, 18, 23], and many different numerical schemes have been developed over the past decades. Both for linear and nonlinear free surface gravity waves the vast majority of numerical techniques have been boundary element methods, because they offer the promise of a significant reduction in computational complexity by transforming the problem into an integral equation using a Greens function technique. For problems requiring higher order accuracy, complex geometries and, in particular, for non-linear free surface boundary conditions, the use of finite element methods can, however, be computationally more efficient, especially when combined with efficient iterative solution techniques to solve the resulting system of algebraic equations, see for instance [4, 8, 24]. More information about the asymptotic computational complexity of boundary and finite element methods can be found in [8].

Finite element methods also offer the possibility to apply hp -adaptive techniques, in which the mesh is locally refined (h -adaptation) or the order of the polynomial basis functions is locally adjusted (p -adaptation). Due to their local, element wise discretization discontinuous Galerkin (DG) finite element methods are well suited for this technique and combine excellent accuracy and parallel performance with great flexibility to adapt the mesh using hp -adaptation. Discontinuous Galerkin methods for elliptic and parabolic pde's have been around already for quite some time and were known initially as interior penalty methods. Recently, they have seen a revival due to their potential for hp -adaptive techniques. An excellent overview and detailed analysis of DG-methods for elliptic partial differential equations can be found in [2]. Further theoretical results on elliptic equations are discussed in [3, 6, 7, 9, 10, 14].

Despite their great potential, the application of discontinuous Galerkin finite element methods to free surface gravity waves has been very limited and mainly restricted to the shallow water equations [15, 16]. As a first step in the development of these algorithms we provide in this paper a complete analysis of a discontinuous Galerkin finite element method for linear free

surface gravity waves. This will provide essential information necessary for the extension to non-linear free surface wave problems. After some preliminaries, we start with the discussion of an efficient DG discretization for linear free surface gravity waves. This requires the solution of the Laplace equation, for which we have chosen the approach from Brezzi et al. [6], since this method results in the most compact stencil with the smallest number of non-zero entries, and uses data only from nearest neighboring elements. The differences in complexity and performance with the local discontinuous Galerkin method proposed by Cockburn and Shu [9, 10, 12] are, however, small.

Next, we analyse the convergence and stability properties of the numerical discretization in detail and give an a-priori error analysis of the free surface wave height and potential. The application of finite element methods has been hampered by the occurrence of saw-tooth waves resulting from a dependence of the stability on the smoothness of the mesh, for an overview see [19]. In order to overcome this problem additional smoothing has been suggested [24], viscosity terms were added to the free surface equation [19], or a finite difference reconstruction of the free surface has been applied [22], but none of these approaches are very appealing. The smoothing and viscosity approaches unnecessarily damp the wave amplitude. The finite difference approach is not attractive in general geometries and is not easily combined with *hp*-adaptation.

We could identify as a significant source of the numerical instabilities the weak coupling between the Laplace equation describing the velocity potential and the free surface condition, which is generally updated after the Laplace equation is solved. We will demonstrate that these numerical instabilities can be removed by introducing the free surface boundary condition directly into the weak formulation for the velocity potential. This results in a fully coupled set of equations with a symmetric positive definite matrix. The numerical scheme is unconditionally stable and has a very small numerical dissipation, which is important to simulate water waves for a long period of time. A nice feature of the proposed DG finite element discretization is also that all boundary conditions are imposed weakly. This alleviates the problems at the free surface - wave maker intersection which occur with standard node based finite element methods, where it is unclear which boundary condition to impose at the intersection point.

The organization of this paper is as follows. In Section 2, we present the equations governing linear free-surface gravity waves. We define the tessellation, function spaces and traces in Section 3. Next, we derive the discontinuous Galerkin discretization in Section 4. In Section 5, we analyse

the stability and convergence of the discretization and provide an a-priori error analysis. In Section 6 we derive the algebraic equations for the discontinuous Galerkin discretization. The effects of the DG method on the dispersion and dissipation of free surface waves are analyzed in Section 7. We conclude in Section 8 with a numerical study of some model problems with discretizations ranging from second to fourth order accuracy.

2 Equations governing linear free-surface gravity waves

Let us assume the fluid to be incompressible and inviscid, with the velocity field irrotational. The flow domain $\Omega \subset \mathbb{R}^3$ is bounded by a free surface Γ_S , a solid boundary Γ_N , and periodic boundaries Γ_P , such that $\bar{\Gamma}_S \cup \bar{\Gamma}_N \cup \bar{\Gamma}_P = \partial\Omega$. We assume that Γ_S and Γ_N have a non-zero measure and the boundaries Γ_S , Γ_N , and Γ_P are non-overlapping and Lipschitz continuous. The free surface Γ_S is defined as $\Gamma_S := \{(x, y) \in \mathcal{S} \mid z = 0\}$, with $\mathcal{S} \subset \mathbb{R}^2$ and x, y, z the coordinates in a standard Cartesian coordinate system. The flow domain Ω is defined as $\Omega := \{(x, y, z) \in \mathbb{R}^3 \mid -h_b(x, y) < z < 0, \forall (x, y) \in \mathcal{S}\}$, with $h_b : \mathcal{S} \rightarrow \mathbb{R}^+ \cup \{0\}$ representing the bottom topography. We also assume that the wave height $\zeta : \mathcal{S} \times [t_0, t_n] \rightarrow \mathbb{R}$ is a single valued function of its arguments. Let $\phi : \Omega \times [t_0, t_n] \rightarrow \mathbb{R}$ denote the velocity potential and $u = \nabla\phi$ the fluid velocity. We make the various parameters dimensionless by redefining them as:

$$\phi \rightarrow H\sqrt{Hg_c}\phi, \quad (x, y, z) \rightarrow H(x, y, z), \quad t \rightarrow \sqrt{\frac{H}{g_c}}t, \quad \text{and} \quad u \rightarrow u\sqrt{Hg_c},$$

where g_c is the gravitational constant, H the average depth, and t denotes time. The following equations then govern the flow field of linear free-surface gravity waves:

$$-\Delta\phi = 0 \quad \text{in } \Omega, \quad n \cdot \nabla\phi = g_N \quad \text{at } \Gamma_N, \quad (1)$$

with linear free surface boundary conditions at Γ_S :

$$\frac{\partial\phi}{\partial t} + \zeta = 0, \quad \text{and} \quad \frac{\partial\zeta}{\partial t} - n \cdot \nabla\phi = 0, \quad (2)$$

and periodic boundary conditions at Γ_P :

$$\phi(x + L_x, y + L_y, z, t) = \phi(x, y, z, t). \quad (3)$$

Here, $n \in \mathbb{R}^3$ is the unit outward normal vector to $\partial\Omega$, $g_N : \Gamma_N \rightarrow \mathbb{R}$ the prescribed normal velocity at Γ_N , and L_x, L_y the length of the periodic

domain in the x - and y -directions. For analysis purposes, and in order to obtain an efficient discretization technique, it is beneficial to transform (2) into a single equation for ϕ at Γ_S :

$$\frac{\partial^2 \phi}{\partial t^2} + n \cdot \nabla \phi = 0. \quad (4)$$

As initial conditions, we either start without any waves with $\phi(x, y, z, t_0) = \zeta(x, y, z, t_0) = 0$ and the waves are generated by the wave maker by specifying a periodic normal velocity at Γ_N ; or, we start with an analytic wave field in a periodic domain and ϕ, ζ are known at initial time.

Note, if we change $\phi \rightarrow \phi + c$, with $c \in \mathbb{R}$ an arbitrary constant, then the equations and boundary conditions (1), (3) and (4) remain unchanged. Hence, the potential ϕ is undetermined up to an arbitrary constant.

3 Tessellation, function spaces and traces

Let \mathcal{T}_h denote a tessellation of Ω with shape-regular elements K , with maximum diameter h . Conditions to ensure that elements are shape regular can be found in e.g. [5, 11]. For simplicity we assume that Ω is a polyhedral domain. Let Γ denote the union of the boundary faces of the elements $K \in \mathcal{T}_h$, i.e. $\Gamma = \bigcup_{K \in \mathcal{T}_h} \partial K$, and $\Gamma_0 = \Gamma \setminus \partial\Omega$. We denote the set of all faces in \mathcal{T}_h by $\{\mathcal{F}_h\}$, all internal faces by $\{\mathcal{F}_h^I\}$, faces on Γ_N by $\{\mathcal{F}_h^N\}$, faces on Γ_S by $\{\mathcal{F}_h^S\}$, and faces on $\partial\Omega$ by $\{\mathcal{F}_h^\partial\}$. Note, faces at Γ_P are considered internal faces, where the periodicity relation (3) is used to connect the external part of Γ_P to the interior of the domain Ω .

For a general domain $\Omega \subset \mathbb{R}^d$, with $d = \dim(\Omega)$, we use the standard definition for the Sobolev space $H^s(\Omega)$, $s \in \mathbb{R}$, for real valued functions, see [5]. For $s = 0$, the Sobolev space $H^s(\Omega)$ is denoted as $L^2(\Omega)$, with inner product and norm

$$(u, v)_\Omega = \int_\Omega u v dx, \quad \|v\|_{0, \Omega} = (v, v)_\Omega^{\frac{1}{2}},$$

and for $m \geq 0$, integer, we define the $H^m(\Omega)$ norm and semi-norm as:

$$\|v\|_{m, \Omega} := \left(\sum_{|\alpha| \leq m} \|D^\alpha v\|_{0, \Omega}^2 \right)^{\frac{1}{2}}, \quad |v|_{m, \Omega} := \left(\sum_{|\alpha|=m} \|D^\alpha v\|_{0, \Omega}^2 \right)^{\frac{1}{2}},$$

where $D^\alpha = (\partial/\partial x_1)^{\alpha_1} \dots (\partial/\partial x_d)^{\alpha_d}$ denotes an arbitrary derivative with multi-index symbol $\alpha = (\alpha_1, \dots, \alpha_d)$, $\alpha_i \in \mathbb{N} \cup \{0\}$ and $|\alpha| = \sum_{j=1}^d \alpha_j$. We also introduce the space V , which is defined as:

$$V := \left\{ v \in H^1(\Omega) \mid \int_\Omega v dx = 0 \right\}.$$

Now we introduce the finite element spaces associated with the tessellation \mathcal{T}_h . We denote with $H^l(\mathcal{T}_h)$, $l \in \mathbb{Z}$, the space of functions such that their restriction to each element $K \in \mathcal{T}_h$ belongs to the Sobolev space $H^l(K)$. Let $\mathcal{P}_p(K)$ be the space of polynomials of degree at most $p \geq 1$ on $K \in \mathcal{T}_h$. We define the spaces V_h^p and Σ_h^p as:

$$V_h^p := \left\{ v \in L^2(\Omega) \mid v|_K \in \mathcal{P}_p(K), \forall K \in \mathcal{T}_h \right\},$$

$$\Sigma_h^p := \left\{ \sigma \in [L^2(\Omega)]^d \mid \sigma|_K \in [\mathcal{P}_p(K)]^d, \forall K \in \mathcal{T}_h \right\},$$

with $d = \dim(\Omega)$. For consistency reasons, we need to assume that $\nabla V_h^p \subset \Sigma_h^p$.

The *traces* of functions in $H^l(\mathcal{T}_h)$ belong to $T(\Gamma) := \prod_{K \in \mathcal{T}_h} H^{l-\frac{1}{2}}(\partial K)$. Functions in $T(\Gamma)$ are double-valued on Γ_0 and single-valued on $\partial\Omega$. Next, we introduce some trace operators to manipulate the numerical fluxes in the discontinuous Galerkin formulation. For $v \in T(\Gamma)$, we define the *average* $\langle v \rangle$ and *jump* $[[v]]$ operators of v at an internal face $\mathcal{F} \in \mathcal{F}_h^I$ as follows:

$$\langle v \rangle := \frac{1}{2}(v_L + v_R), \quad [[v]] := v_L n_L + v_R n_R, \quad (5)$$

with $v_L := v|_{\partial K_L}$ and $v_R := v|_{\partial K_R}$, and K_L, K_R the elements connected to the face \mathcal{F}^I with unit outward normal vectors n_L and n_R , respectively. For $q \in [T(\Gamma)]^d$ we similarly define q_L and q_R and set:

$$\langle q \rangle = \frac{1}{2}(q_L + q_R), \quad [[q]] = q_L \cdot n_L + q_R \cdot n_R, \quad \text{at } \mathcal{F} \in \mathcal{F}_h^I. \quad (6)$$

For $\mathcal{F} \in \mathcal{F}_h^\partial$, the set of exterior boundary faces, each $v \in T(\Gamma)$ and $q \in [T(\Gamma)]^d$ has a uniquely defined restriction on \mathcal{F} ; and we define:

$$[[v]] = vn, \quad \langle q \rangle = q \quad \text{at } \mathcal{F} \in \mathcal{F}_h^\partial. \quad (7)$$

Since we do not require either of the quantities $\langle v \rangle$ or $[[q]]$ on boundary faces, we leave them undefined.

For the definition of the primal DG formulation we need to define the lifting operators: $\mathcal{L} : L^2(\Gamma_0) \rightarrow \Sigma_h^p$ and $\mathcal{R} : [L^2(\Gamma)]^d \rightarrow \Sigma_h^p$:

$$\int_{\Omega} \mathcal{L}(q) \cdot \sigma dx = \int_{\Gamma_0} q [[\sigma]] ds, \quad \forall \sigma \in \Sigma_h^p, \quad (8)$$

$$\int_{\Omega} \mathcal{R}(q) \cdot \sigma dx = \int_{\Gamma} q \cdot \langle \sigma \rangle ds, \quad \forall \sigma \in \Sigma_h^p, \quad (9)$$

and for a face $\mathcal{F} \in \mathcal{F}_h$ also the *local* lifting operator $\mathcal{R}_{\mathcal{F}} : [L^2(\mathcal{F})]^d \rightarrow \Sigma_h^p$:

$$\int_{\Omega} \mathcal{R}_{\mathcal{F}}(q) \cdot \sigma dx = \int_{\mathcal{F}} q \cdot \langle \sigma \rangle ds, \quad \forall \sigma \in \Sigma_h^p. \quad (10)$$

Note that $\mathcal{R}_{\mathcal{F}}(q)$ vanishes outside the elements connected to the face \mathcal{F} and that for a particular element $K \in \mathcal{T}_h$ we have the relation:

$$\mathcal{R}(q) = \sum_{\mathcal{F} \in \mathcal{F}_h} \mathcal{R}_{\mathcal{F}}(q), \quad \forall q \in [L^2(\Gamma)]^d. \quad (11)$$

We can now define the following mesh-dependent norms and semi-norms for $v \in H^m(\mathcal{T}_h)$, with $m \geq 0$, integer:

$$\begin{aligned} \|v\|_{m,h}^2 &:= \sum_{K \in \mathcal{T}_h} \|v\|_{m,K}^2, & |v|_{m,h}^2 &:= \sum_{K \in \mathcal{T}_h} |v|_{m,K}^2, \\ |v|_*^2 &:= \sum_{\mathcal{F} \in \mathcal{F}_h} \|R_{\mathcal{F}}([v])\|_{0,\Omega}^2, \\ |v|_{1,h,*}^2 &:= |v|_{1,h}^2 + |v|_*^2, \end{aligned} \quad (12)$$

and for $v \in V(h) := H^1(\mathcal{T}_h) + V \subset H^2(\mathcal{T}_h)$:

$$\|v\|^2 := |v|_{1,h}^2 + \sum_{K \in \mathcal{T}_h} h_K^2 |v|_{2,K}^2 + |v|_*^2. \quad (13)$$

Note, for $v \in V(h)$, both (12) and (13) define norms, not just semi-norms, also in the case $h = 0$. This follows directly from the discrete Poincaré-Friedrichs inequality, see [1], equation (2.2) (which can be straightforwardly extended to general convex elements with a Lipschitz continuous boundary), and estimate (4.5) in [2], which together imply:

$$\|v\|_{0,\Omega} \leq C(|v|_{1,h}^2 + |v|_*^2)^{\frac{1}{2}}, \quad \forall v \in V(h),$$

with C a positive constant independent of v . In addition, the $H^1(\Omega)$ semi-norm is a norm in V when Ω is a bounded Lipschitz domain due to the Poincaré-Friedrichs inequality (see [13], Page 127). Note, restricted to $v \in V(h)$ the norm $|v|_{1,h,*}$ is, for finite h , equivalent with the DG-norm $\|v\|$, which follows directly from a local inverse inequality (see [5], Section 4.5).

4 Discontinuous Galerkin formulation

In this section we summarize the derivation of the discontinuous Galerkin finite element discretization for linear free-surface gravity waves given by (1), (3) and (4). We will follow the approach from Brezzi et al. [6], which is analyzed in detail in [2]. First, we transform the Laplace equation into a first order system of equations:

$$u = \nabla \phi, \quad -\nabla \cdot u = 0, \quad \text{in } \Omega, \quad (14)$$

where u represents the fluid velocity and ϕ the potential. Multiplying (14) with arbitrary test functions $\sigma \in \Sigma_h^p$ and $v \in V_h^p$, integrating by parts over each element $K \in \mathcal{T}_h$ (twice for (15)), and adding over all elements in \mathcal{T}_h , we obtain for $u_h \in \Sigma_h^p$ and $\phi_h \in V_h^p$ the relation:

$$\begin{aligned} \int_{\Omega} u_h \cdot \sigma dx &= \sum_{K \in \mathcal{T}_h} \int_K \nabla_h \phi_h \cdot \sigma dx + \sum_{K \in \mathcal{T}_h} \int_{\partial K} (\hat{\phi}_K - \phi_h) \sigma \cdot n_K ds, \quad (15) \\ - \sum_{K \in \mathcal{T}_h} \int_K (\nabla_h \cdot u_h) v dx &= \sum_{K \in \mathcal{T}_h} \int_K u_h \cdot \nabla_h v dx - \sum_{K \in \mathcal{T}_h} \int_{\partial K} n_K \cdot \hat{u}_K v ds = 0, \end{aligned} \quad (16)$$

for $\forall \sigma \in \Sigma_h^p$ and $\forall v \in V_h^p$, where $\nabla_h = \nabla|_K$ for all $K \in \mathcal{T}_h$. Here, the numerical fluxes $\hat{u}(u)$ and $\hat{\sigma}(\nabla_h u, \sigma_h)$, defined as:

$$\hat{u} : H^1(\mathcal{T}_h) \rightarrow T(\Gamma), \quad \hat{\sigma} : H^2(\mathcal{T}_h) \times [H^1(\mathcal{T}_h)]^d \rightarrow [T(\Gamma)]^d,$$

are introduced to account for the multivalued trace at Γ . The different discontinuous Galerkin formulations proposed so far in the literature can all be derived by specifying these numerical fluxes. A full account is given in [2]. In order to simplify the sums over the element boundaries ∂K , we use identity (3.3) in [2], which states:

$$\sum_{K \in \mathcal{T}_h} \int_{\partial K} \varphi_K q_K \cdot n_K ds = \int_{\Gamma} [[\varphi]] \cdot \langle q \rangle ds + \int_{\Gamma_0} \langle \varphi \rangle [[q]] ds, \quad (17)$$

$\forall \varphi \in T(\Gamma)$ and $q \in [T(\Gamma)]^d$. If we introduce (17) into (15–16) we obtain:

$$\int_{\Omega} u_h \cdot \sigma dx = \int_{\Omega} \nabla_h \phi_h \cdot \sigma dx + \int_{\Gamma} [[[\hat{\phi} - \phi_h]]] \cdot \langle \sigma \rangle ds + \int_{\Gamma_0} \langle \hat{\phi} - \phi_h \rangle [[\sigma]] ds \quad (18)$$

$$\int_{\Omega} u_h \cdot \nabla_h v dx - \int_{\Gamma} [[v]] \cdot \langle \hat{u} \rangle ds - \int_{\Gamma_0} \langle v \rangle [[\hat{u}]] ds = 0. \quad (19)$$

The numerical fluxes in the DG formulation from Brezzi et al. [6] are defined as:

$$\hat{\phi} = \begin{cases} \langle \phi_h \rangle & \text{on } \Gamma_0, \\ \phi_h & \text{on } \Gamma_N \cup \Gamma_S, \end{cases} \quad (20)$$

$$\hat{u} = \langle u_h \rangle - \alpha_r ([[\phi_h]]) \quad \text{on } \Gamma_0, \quad (21)$$

$$\hat{u} \cdot n = g_N \quad \text{on } \Gamma_N, \quad (22)$$

$$\hat{u} \cdot n = -\frac{\partial^2 \phi}{\partial t^2} \quad \text{on } \Gamma_S, \quad (23)$$

where $\alpha_r(q) = \eta_{\mathcal{F}} \langle \mathcal{R}_{\mathcal{F}}(q) \rangle$ for $\mathcal{F} \in \mathcal{F}_h^I$, $\eta_{\mathcal{F}} \in \mathbb{R}^+$ a positive number, and $g_N \in H^{\frac{1}{2}}(\Gamma_N)$ the prescribed Neumann boundary data. If we introduce the relations for the lifting operators (8-9) into (18), we obtain the following equation for u_h , almost everywhere:

$$u_h = \nabla_h \phi_h + \mathcal{R} \left(\left[\left[\widehat{\phi} - \phi_h \right] \right] \right) + \mathcal{L} \left(\left\langle \widehat{\phi} - \phi_h \right\rangle \right). \quad (24)$$

The weak formulation for the potential ϕ_h is now obtained by introducing the relation (24) for u_h , the numerical fluxes (20–22), and the boundary conditions (1) and (4) into (19):

Find a $\phi_h \in V_h^p \times (C^2(t_0, t_n) \cap C^1[t_0, t_n])$, such that for all $v \in V_h^p$, the following relation is satisfied:

$$\begin{aligned} & \int_{\Omega} \nabla_h \phi_h \cdot \nabla_h v dx + \int_{\Gamma_0} \left[\left[\widehat{\phi} - \phi_h \right] \right] \cdot \langle \nabla_h v \rangle ds + \int_{\Gamma_0} \left\langle \widehat{\phi} - \phi_h \right\rangle \cdot \llbracket \nabla_h v \rrbracket ds - \\ & \int_{\Gamma_0} \llbracket v \rrbracket \cdot \langle \widehat{u} \rangle ds - \int_{\Gamma_0} \langle v \rangle \cdot \llbracket \widehat{u} \rrbracket ds + \int_{\Gamma_S} v \frac{\partial^2 \phi_h}{\partial t^2} ds - \int_{\Gamma_N} v g_N ds = 0. \end{aligned} \quad (25)$$

Here we used (8-9) to transform integrals over Ω with the lifting operators into boundary integrals over Γ . Using the definitions of the jump and average operators (5–7), the numerical fluxes (20–22), and the lifting operators (8-9), we can transform the following integrals along Γ_0 in (25) into:

$$\begin{aligned} & \int_{\Gamma_0} \left[\left[\widehat{\phi} - \phi_h \right] \right] \cdot \langle \nabla_h v \rangle ds = - \int_{\Gamma_0} \llbracket \phi_h \rrbracket \cdot \langle \nabla_h v \rangle ds \\ & \int_{\Gamma_0} \llbracket v \rrbracket \cdot \langle \widehat{u} \rangle ds = \int_{\Gamma_0} \llbracket v \rrbracket \cdot \langle \nabla_h \phi_h \rangle ds + \\ & \quad \sum_{\mathcal{F} \in \mathcal{F}^I} \int_{\Omega} \mathcal{R}_{\mathcal{F}}(\llbracket v \rrbracket) \cdot \mathcal{R} \left(\left[\left[\widehat{\phi} - \phi_h \right] \right] \right) dx - \\ & \quad \sum_{\mathcal{F} \in \mathcal{F}^I} \eta_{\mathcal{F}} \int_{\Omega} \mathcal{R}_{\mathcal{F}}(\llbracket v \rrbracket) \cdot \mathcal{R}_{\mathcal{F}}(\llbracket \phi_h \rrbracket) ds, \end{aligned} \quad (26)$$

while the other integrals along Γ_0 are zero. The integrand in the second integral at the righthand of (26), increases the stencil of the DG discretization for each element K beyond its nearest neighbors. Using (11), the fact that $\left[\left[\widehat{\phi} - \phi_h \right] \right] = 0$ at $\Gamma_N \cup \Gamma_S$, and the relation $\left[\left[\widehat{\phi} - \phi_h \right] \right] = - \llbracket \phi_h \rrbracket$ at Γ_0 , we can approximate this contribution as:

$$\begin{aligned} & \sum_{\mathcal{F} \in \mathcal{F}^I} \int_{\Omega} \mathcal{R}_{\mathcal{F}}(\llbracket v \rrbracket) \cdot \mathcal{R} \left(\left[\left[\widehat{\phi} - \phi_h \right] \right] \right) dx \cong \\ & \quad - n_f \sum_{\mathcal{F} \in \mathcal{F}^I} \int_{\Omega} \mathcal{R}_{\mathcal{F}}(\llbracket v \rrbracket) \cdot \mathcal{R}_{\mathcal{F}}(\llbracket \phi_h \rrbracket) dx, \end{aligned}$$

with n_f the number of faces of an element in the tessellation, e.g. $n_f = 6$ for hexahedral elements. This approximation has no effect on the accuracy and stability of the DG discretization, see Section 5.1, and the stencil now only contains contributions from the element itself and its nearest neighbors.

If we introduce the bilinear form $B_h^0 : V_h^p \times V_h^p \rightarrow \mathbb{R}$ and the linear form $L_h : V_h^p \rightarrow \mathbb{R}$, which are defined as:

$$\begin{aligned} B_h^0(\phi_h, \psi_h) &= \int_{\Omega} \nabla_h \phi_h \cdot \nabla_h \psi_h dx - \int_{\Gamma_0} ([[\phi_h]]) \cdot \langle \nabla_h \psi_h \rangle + [[\psi_h]]) \cdot \langle \nabla_h \phi_h \rangle ds \\ &\quad + \sum_{\mathcal{F} \in \mathcal{F}_h^I} (\eta_{\mathcal{F}} + n_f) \int_{\Omega} \mathcal{R}_{\mathcal{F}}([[\phi_h]]) \cdot \mathcal{R}_{\mathcal{F}}([[\psi_h]]) dx, \\ L_h(\psi_h) &= \int_{\Gamma_N} g_N \psi_h dx, \end{aligned} \quad (27)$$

then the semi-discrete weak formulation for linear free-surface gravity waves can be formulated as:

Find a $\phi_h \in V_h^p \times (C^2(t_0, t_n) \cap C^1[t_0, t_n])$, such that for all $v \in V_h^p$, the following relation is satisfied:

$$B_h(\phi_h, v) := \left(\frac{\partial^2 \phi_h}{\partial t^2}, v \right)_{\Gamma_S} + B_h^0(\phi_h, v) = L_h(v). \quad (28)$$

For the fully discrete weak formulation, we introduce the following backward difference approximations to $\partial \phi_h / \partial t$ at $t = t_n$:

$$\begin{aligned} \varsigma_h^1 &:= \frac{1}{\Delta t} (\phi_h^1 - \phi_h^0), \\ \varsigma_h^n &:= \frac{1}{\Delta t} \left(\frac{3}{2} \phi_h^n - 2 \phi_h^{n-1} + \frac{1}{2} \phi_h^{n-2} \right), \quad \text{for } n \geq 2, \end{aligned}$$

with $\phi_h^n = \phi_h(t_n)$. The weak formulation for the potential ϕ_h^n can now be formulated as:

Find a $\phi_h^n, \varsigma_h^n \in V_h^p$, such that for all $v_1, v_2 \in V_h^p$, the following relation is satisfied for $n \geq 2$:

$$\frac{3}{2\Delta t} \left(\varsigma_h^n, v_1 \right)_{\Gamma_S} + B_h^0(\phi_h^n, v_1) = L_h^n(v_1) \quad (29)$$

$$\frac{3}{2\Delta t} \left(\phi_h^n, v_2 \right)_{\Gamma_S} - \left(\varsigma_h^n, v_2 \right)_{\Gamma_S} = \frac{1}{\Delta t} \left(2\phi_h^{n-1} - \frac{1}{2}\phi_h^{n-2}, v_2 \right)_{\Gamma_S}, \quad (30)$$

and for $n = 1$:

$$\frac{1}{\Delta t} \left(\varsigma_h^1, v_1 \right)_{\Gamma_S} + B_h^0(\phi_h^1, v_1) = L_h^1(v_1) \quad (31)$$

$$\frac{1}{\Delta t} \left(\phi_h^1, v_2 \right)_{\Gamma_S} - \left(\varsigma_h^1, v_2 \right)_{\Gamma_S} = \frac{1}{\Delta t} \left(\phi_h^0, v_2 \right)_{\Gamma_S}. \quad (32)$$

Here $L_h^n(v_1)$ is defined as:

$$L_h^1(v_1) := L_h(v_1) + \frac{1}{\Delta t} (\zeta_h^0, v_1)_{\Gamma_S} \quad (33)$$

$$L_h^n(v_1) := L_h(v_1) + \frac{1}{\Delta t} \left(2\zeta_h^{n-1} - \frac{1}{2}\zeta_h^{n-2}, v_1 \right)_{\Gamma_S} \quad \text{for } n \geq 2. \quad (34)$$

Note, using (2) we have the relation $\zeta_h^n = -\zeta_h^n$, with ζ the wave height.

For the numerical simulations and part of the theoretical analysis it is beneficial to eliminate ζ_h^n , since this will result in a positive definite linear system for the expansion coefficients of the potential ϕ_h^n . We rewrite the weak formulations (29–32) therefore as:

Find a $\phi_h^n \in V_h^p$, such that for all $v \in V_h^p$ and $n \geq 2$, the following relation is satisfied:

$$\begin{aligned} B_h^n(\phi_h^n, v) &:= \frac{9}{4\Delta t^2} (\phi_h^n, v)_{\Gamma_S} + B_h^0(\phi_h^n, v) \\ &= L_h^n(v) + \frac{1}{\Delta t^2} (3\phi_h^{n-1} - \frac{3}{4}\phi_h^{n-2}, v)_{\Gamma_S}, \end{aligned} \quad (35)$$

and for $n = 1$:

$$B_h^1(\phi_h^1, v) := \frac{1}{\Delta t^2} (\phi_h^1, v)_{\Gamma_S} + B_h^0(\phi_h^1, v) = L_h^1(v) + \frac{1}{\Delta t^2} (\phi_h^0, v)_{\Gamma_S}. \quad (36)$$

5 Stability and error analysis

5.1 Accuracy and stability of spatial discretization

The proof of stability and the error analysis of the DG discretization for linear free-surface gravity waves strongly uses the results obtained in [6], and we refer for more details to [2].

We summarize the following theoretical results, which will be partly needed in the proof of stability and the error estimate for the fully discrete scheme discussed in Section 5.2:

Lemma 5.1 *Suppose that $\Omega \subset R^d$ is a bounded Lipschitz domain, then for any $\Delta t > 0$ the bilinear form B_h^n , given by (35-36), is bounded, coercive for any $v \in V(h)$ and $\eta = \min_{\mathcal{F} \in \mathcal{F}_h} \eta_{\mathcal{F}} > 0$. In addition, it is consistent, adjoint consistent, and there exists a positive constant C , independent of h and ϕ^n , such that for any $\phi_h^n \in V_h^p$ the following error estimate holds:*

$$\|\phi^n - \phi_h^n\|_{0,\Omega} \leq Ch^{p+1} |\phi^n|_{p+1,\Omega} \quad (37)$$

$$\| |\phi^n - \phi_h^n| \| \leq Ch^p |\phi^n|_{p+1,\Omega}. \quad (38)$$

Proof:

- *Boundedness* of $B_h^0(\phi_h, v)$ has been proven in [6] and [2]. Using the Schwarz inequality and applying the trace theorem on each element $K \in \mathcal{T}_h$ connected to Γ_S , we also obtain:

$$|(\phi_h^n, v)_{\Gamma_S}| \leq C_{\Gamma_S} \|v\|_{1,h} \|\phi_h^n\|_{1,h},$$

hence:

$$|B_h^n(\phi_h^n, v)| \leq C \|\phi_h^n\| \|v\|.$$

Here, we used the Poincaré-Friedrichs inequality to establish the equivalence of the $\|v\|_{1,h}$ norm with $|v|_{1,h}$, and the fact that for finite h the norms (12) and (13) are equivalent.

- *Coercivity* of $B_h^0(\phi_h, \phi_h)$ with homogeneous Dirichlet conditions in the $|v|_{1,h,*}$ norm was proven in [6] for triangles under the condition $\eta > 0$. A careful tracing of the proof in [6] shows that for general shape regular elements with n_f faces the same proof applies to B_h^0 defined in (27), we only need to use the relation:

$$\|\mathcal{R}(q)\|_{0,K}^2 \leq n_f \sum_{\mathcal{F} \in \partial K} \|\mathcal{R}_{\mathcal{F}}(q)\|_{0,K}^2, \quad (39)$$

instead of (43) in [6], which is obtained by applying the arithmetic inequality $(\sum_{i=1}^n a_i)^2 \leq n \sum_{i=1}^n (a_i)^2$ and (11). This implies that for finite $h > 0$ coercivity is ensured if $\eta > 0$, hence:

$$\alpha \|\phi_h^n\|^2 \leq B_h^0(\phi_h^n, \phi_h^n), \quad \forall \phi_h^n \in V_h^p, \quad (40)$$

with $\alpha = \min(1 - \epsilon, \eta + n_f - \frac{n_f}{\epsilon})$ and $\epsilon \in (\frac{n_f}{\eta + n_f}, 1)$ an arbitrary number. For $h = 0$ we need to apply the Poincaré-Friedrichs inequality, which ensures that in the space V the H^1 semi-norm is actually a norm, see also [5], Section 5.3. The coercivity of $B_h^n(\phi_h^n, \phi_h^n)$, for all $\phi_h^n \in V_h^p$ and $\Delta t > 0$ is then immediate:

$$\alpha \|\phi_h^n\|^2 \leq B_h^0(\phi_h^n, \phi_h^n) + \frac{9}{4\Delta t^2} \|\phi_h^n\|_{0,\Gamma_S}^2 = B_h^n(\phi_h^n, \phi_h^n). \quad (41)$$

- *Consistency* of $B_h^n(\phi_h^n, v)$, $\forall \phi_h^n, v \in V_h^p$, follows directly from the consistency of $B_h^0(\phi_h^n, v)$, see [2]. This also implies the orthogonality relation:

$$B_h^n(\phi^n - \phi_h^n, v) = 0, \quad \forall v \in V_h^p.$$

- For *Adjoint Consistency* we consider the adjoint problem:

$$\begin{aligned} -\Delta\psi &= 0, & \text{in } \Omega \\ n \cdot \nabla\psi &= 0, & \text{at } \Gamma_N, \quad n \cdot \nabla\psi + \frac{9}{4\Delta t^2}\psi = 0, & \text{at } \Gamma_S, \end{aligned}$$

which is obtained if we introduce $w, \psi \in V_h^p$ into (35) and integrate by parts:

$$B_h^n(w, \psi) = 0, \quad \forall w \in V_h^p,$$

hence the DG discretization is adjoint consistent.

- The *DG and L^2 approximation error* (37–38) are obtained analogously to [2] using the coercivity, boundedness, and orthogonality properties of $B_h^0(\phi_h^n, \psi_h^n)$. In addition we use the trace theorem for the free-surface contribution $(\phi_h^n, v)_{\Gamma_S}$. \square

Due to the coercivity, symmetry and boundedness of the bilinear form $B_h^0(\phi_h, \psi_h)$ we can also define the following inner product:

$$((\phi_h, \psi_h))_{\Omega, h} := B_h^0(\phi_h, \psi_h), \quad \forall \phi_h, \psi_h \in V_h^p, \quad (42)$$

with induced norm:

$$|||\phi_h|||_{DG}^2 = ((\phi_h, \phi_h))_{\Omega, h}. \quad (43)$$

Note, for $h = 0$ we have $((\phi, \psi))_{\Omega, h} = (\nabla\phi, \nabla\psi)_\Omega$, which is also an inner product for all $\phi, \psi \in V$ due to the Poincaré-Friedrichs inequality. The norm $|||\phi_h|||_{DG}$ is for finite h equivalent with the norm $|||\phi_h|||$.

We also define the Ritz-Galerkin projection $R_h : V \rightarrow V_h^p$, with respect to the inner product $((\cdot, \cdot))_{\Omega, h}$, so that

$$((R_h\phi, \psi))_{\Omega, h} = ((\phi, \psi))_{\Omega, h}, \quad \forall \psi \in V_h^p. \quad (44)$$

5.2 Stability and error estimate of fully discrete scheme

The following theorem proves stability and provides an error estimate for the fully discrete DG formulation. For the proof we follow the same strategy as outlined in [21], Chapter 1.

Theorem 5.2 *Let $\Omega \subset R^d$ be a bounded Lipschitz domain and ϕ_h^n, ζ_h^n , be solutions of (29–34), with $\zeta_h^0 = -\zeta^0$ and $|||R_h\zeta^0 - \zeta^0|||_{0, \Gamma_S} \leq Ch^{r+1} |\zeta|_{r+1, \Gamma_S}$ for any $\phi \in H^{r+1}(\mathcal{T}_h) \times (C^4(t_0, t_n) \cap C^3[t_0, t_n])$ with $r \geq 1$, and $\eta =$*

$\min_{\mathcal{F} \in \mathcal{F}_h} \eta_{\mathcal{F}} = \frac{1}{\Delta t}$, then the discontinuous Galerkin discretization is stable and we have the following error estimate for the wave height ζ_h^n and potential ϕ_h^n at time level t_n for $n \geq 0$, n integer:

$$\begin{aligned} \|\zeta^n - \zeta_h^n\|_{0,\Gamma_S} &\leq Ch^{r+1} \left(4 \int_{t_0}^{t_n} |\zeta,t|_{r+1,\Gamma_S} ds + \frac{5}{2} \int_{t_0}^{t_1} |\zeta,t|_{r+1,\Gamma_S} ds \right) \\ &\quad + \frac{5}{2} \Delta t^2 \int_{t_0}^{t_n} \|\zeta,ttt\|_{0,\Gamma_S} ds + \frac{5}{2} \Delta t \int_{t_0}^{t_1} \|\zeta,tt\|_{0,\Gamma_S} ds \\ &\quad + C_a \Delta t h^{r+1} \left(|\zeta^0|_{r+1,\Gamma_S} + \int_{t_0}^{t_n} |\zeta,t|_{r+1,\Gamma_S} ds \right) \\ \|\phi^n - \phi_h^n\| &\leq C_{\Gamma_S} \left(Ch^{r+1} \left(16 \int_{t_0}^{t_n} |\zeta,t|_{r+1,\Gamma_S} ds + 10 \int_{t_0}^{t_1} |\zeta,t|_{r+1,\Gamma_S} ds \right. \right. \\ &\quad \left. \left. + 2 \int_{t_{n-2}}^{t_n} |\zeta,t|_{r+1,\Gamma_S} ds \right) + 10 \Delta t^2 \int_{t_0}^{t_n} \|\zeta,ttt\|_{0,\Gamma_S} ds \right. \\ &\quad \left. + 10 \Delta t^2 \int_{t_0}^{t_1} \|\zeta,tt\|_{0,\Gamma_S} ds + \frac{5}{4} \Delta t^2 \int_{t_{n-2}}^{t_n} \|\zeta,ttt\|_{0,\Gamma_S} ds \right) \\ &\quad + C_a \Delta t h^r \left(|\phi^0|_{r+1,\Omega} + \int_{t_0}^{t_n} |\phi,t|_{r+1,\Omega} ds \right). \end{aligned}$$

Remark 5.3 If we assume that the velocity potential ϕ^n is sufficiently regular, e.g. $\phi \in H^2(\mathcal{T}_h)$, then using the adjoint consistency of B_h^n and a duality argument [2] it is straightforward to show that:

$$\|\phi^n - \phi_h^n\| \leq Ch \|\phi^n - \phi_h^n\|_{0,\Omega}.$$

Proof: Introduce:

$$\phi_h^n - \phi^n = (\phi_h^n - R_h \phi^n) + (R_h \phi^n - \phi^n) := \theta_1^n + \rho_1^n, \quad (45)$$

$$\zeta_h^n - \zeta^n = (\zeta_h^n - R_h \zeta^n) + (R_h \zeta^n - \zeta^n) := \theta_2^n + \rho_2^n, \quad (46)$$

with R_h the Ritz-Galerkin projection defined in (44). First, we will derive bounds for θ_1^n and θ_2^n . Consider the following relation for $n \geq 2$:

$$\begin{aligned} (\bar{D}\theta_2^n, v_1)_{\Gamma_S} + B_h^0(\theta_1^n, v_1) &= (\bar{D}\zeta_h^n, v_1)_{\Gamma_S} - (\bar{D}R_h\zeta^n, v_1)_{\Gamma_S} + B_h^0(\phi_h^n, v_1) - \\ &\quad B_h^0(R_h\phi^n, v_1) \\ &= L_h(v_1) - ((R_h\phi^n, v_1))_{\Omega,h} - (\bar{D}R_h\zeta^n, v_1)_{\Gamma_S} \quad (47) \end{aligned}$$

$$= L_h(v_1) - ((\phi^n, v_1))_{\Omega,h} - (R_h\bar{D}\zeta^n, v_1)_{\Gamma_S} \quad (48)$$

$$= \left(\frac{\partial \zeta^n}{\partial t}, v_1\right)_{\Gamma_S} - (R_h\bar{D}\zeta^n, v_1)_{\Gamma_S} \quad (49)$$

$$= -((R_h - I)\bar{D}\zeta^n, v_1)_{\Gamma_S} - \left(\bar{D}\zeta^n - \frac{\partial \zeta^n}{\partial t}, v_1\right)_{\Gamma_S}$$

$$:= -(\omega_1^n, v_1)_{\Gamma_S} - (\omega_2^n, v_1)_{\Gamma_S}, \quad \forall v \in V_h^p, \quad (50)$$

with $\bar{D}\theta_2^n = (\frac{3}{2}\theta_2^n - 2\theta_2^{n-1} + \frac{1}{2}\theta_2^{n-2})/\Delta t$, I the identity operator, and $\frac{\partial \varsigma^n}{\partial t} = \frac{\partial \varsigma}{\partial t}$ at $t = t_n$. Here we used in (47) equations (29) and (34), and the definition of the inner product (42). In (48) we used (44) and the linearity of R_h , and finally in (49), we used the consistency of B_h^0 ((3.13) in [2]), which implies $((\phi^n, v_1))_{\Omega, h} = (\nabla \phi^n, \nabla_h v_1)_{\Omega}$, $\forall v_1 \in V_h^p$, and (28) for $h = 0$ with $\varsigma = \frac{\partial \phi}{\partial t}$. Similarly, we obtain with $\bar{\partial}\theta^n = (\theta^n - \theta^{n-1})/\Delta t$ the relation:

$$(\bar{\partial}\theta_2^1, v_1)_{\Gamma_S} + B_h^0(\theta_1^1, v_1) = -(\omega_1^1, v_1)_{\Gamma_S} - (\omega_2^1, v_1)_{\Gamma_S}, \quad (51)$$

with $\omega_1^1 = (R_h - I)\bar{\partial}\varsigma^1$ and $\omega_2^1 = \bar{\partial}\varsigma^1 - \frac{\partial \varsigma^1}{\partial t}$. We use now the following relation ([21], (1.52), page 18):

$$\begin{aligned} \Delta t(\bar{\partial}\theta_2^1, \theta_2^1)_{\Gamma_S} + \Delta t \sum_{j=2}^n (\bar{D}\theta_2^j, \theta_2^j)_{\Gamma_S} \\ \geq \frac{3}{4} \|\theta_2^n\|_{0, \Gamma_S}^2 - \frac{1}{4} (\|\theta_2^{n-1}\|_{0, \Gamma_S}^2 + \|\theta_2^1\|_{0, \Gamma_S}^2 + \|\theta_2^0\|_{0, \Gamma_S}^2). \end{aligned} \quad (52)$$

Take $v_1 = \theta_2^n$ in (50–51), and sum over all time levels:

$$\begin{aligned} \Delta t(\bar{\partial}\theta_2^1, \theta_2^1)_{\Gamma_S} + \Delta t \sum_{j=2}^n (\bar{D}\theta_2^j, \theta_2^j)_{\Gamma_S} + \Delta t \sum_{j=1}^n B_h^0(\theta_1^j, \theta_2^j) \\ = -\Delta t \sum_{j=1}^n (\omega_1^j + \omega_2^j, \theta_2^j)_{\Gamma_S}, \end{aligned} \quad (53)$$

which, combined with inequality (52), results in the estimate:

$$\begin{aligned} \|\theta_2^n\|_{0, \Gamma_S}^2 + \frac{4}{3} \Delta t \sum_{j=1}^n B_h^0(\theta_1^j, \theta_2^j) \leq \frac{1}{3} (\|\theta_2^{n-1}\|_{0, \Gamma_S}^2 + \|\theta_2^1\|_{0, \Gamma_S}^2 + \|\theta_2^0\|_{0, \Gamma_S}^2) \\ - \frac{4}{3} \Delta t \sum_{j=1}^n (\omega_1^j + \omega_2^j, \theta_2^j)_{\Gamma_S}. \end{aligned} \quad (54)$$

Note, inequality (54) is valid for arbitrary values of $\theta_1^j \in V_h^p$, ($j = 1, \dots, n$), which follows directly from the fact that (52) does not depend on θ_1^j . We can take therefore the supremum of (54) for all values of θ_1^j and obtain the

following estimate for $\|\theta_2^n\|_{0,\Gamma_S}$:

$$\begin{aligned} \|\theta_2^n\|_{0,\Gamma_S}^2 + \frac{4}{3}\Delta t \sum_{j=1}^n \sup_{0 \neq \theta_1^j \in V_h^p} \frac{\left((\theta_1^j, \theta_2^j) \right)_{\Omega,h}}{\|\theta_1^j\|} \|\theta_1^j\| \\ \leq \frac{1}{3} (\|\theta_2^{n-1}\|_{0,\Gamma_S}^2 + \|\theta_2^1\|_{0,\Gamma_S}^2 + \|\theta_2^0\|_{0,\Gamma_S}^2) \\ + \frac{4}{3}\Delta t \sum_{j=1}^n \left(\|\omega_1^j\|_{0,\Gamma_S} + \|\omega_2^j\|_{0,\Gamma_S} \right) \|\theta_2^j\|_{0,\Gamma_S}. \end{aligned}$$

Suppose m is chosen so that $\|\theta_2^m\|_{0,\Gamma_S} = \max_{0 \leq j \leq n} \|\theta_2^j\|_{0,\Gamma_S}$, and using the Riesz representation theorem, which implies:

$$\sup_{0 \neq \theta_1^j \in V_h^p} \frac{\left((\theta_1^j, \theta_2^j) \right)_{\Omega,h}}{\|\theta_1^j\|} = \|\theta_2^j\|,$$

we obtain the following bound:

$$\begin{aligned} \|\theta_2^n\|_{0,\Gamma_S} &\leq \|\theta_2^m\|_{0,\Gamma_S} \\ &\leq \frac{1}{2} (\|\theta_2^1\|_{0,\Gamma_S} + \|\theta_2^0\|_{0,\Gamma_S}) + 2\Delta t \sum_{j=1}^n \left(\|\omega_1^j\|_{0,\Gamma_S} + \|\omega_2^j\|_{0,\Gamma_S} \right), \end{aligned}$$

since $\sum_{j=1}^n \|\theta_1^j\| \|\theta_2^j\| \geq 0$. Using a similar analysis, but now only restricted to (51), we obtain the inequality:

$$\|\theta_2^1\|_{0,\Gamma_S} \leq \|\theta_2^0\|_{0,\Gamma_S} + \Delta t (\|\omega_1^1\|_{0,\Gamma_S} + \|\omega_2^1\|_{0,\Gamma_S}).$$

If we choose $\zeta_h^0 = R_h \zeta^0$ as initial field, then $\|\theta_2^0\|_{0,\Gamma_S} = 0$, and the estimate for $\|\theta_2^n\|_{0,\Gamma_S}$ can be further simplified into:

$$\|\theta_2^n\|_{0,\Gamma_S} \leq 2\Delta t \sum_{j=1}^n (\|\omega_1^j\|_{0,\Gamma_S} + \|\omega_2^j\|_{0,\Gamma_S}) + \frac{1}{2}\Delta t (\|\omega_1^1\|_{0,\Gamma_S} + \|\omega_2^1\|_{0,\Gamma_S}). \quad (55)$$

For a bound on $\|\theta_1^n\|$ we use (50–51) with $v_1 = \theta_1^n$, the coercivity estimate (40) for $B_h^0(\theta_1^n, \theta_1^n)$, and the trace theorem with constant C_{Γ_S} :

$$\begin{aligned} \alpha \|\theta_1^n\|^2 &\leq |(\bar{D}\theta_2^n, \theta_1^n)_{\Gamma_S}| + |(\omega_1^n, \theta_1^n)_{\Gamma_S}| + |(\omega_2^n, \theta_1^n)_{\Gamma_S}| \\ &\leq C_{\Gamma_S} \left(\frac{1}{\Delta t} \left(\frac{3}{2} \|\theta_2^n\|_{0,\Gamma_S} + 2 \|\theta_2^{n-1}\|_{0,\Gamma_S} + \frac{1}{2} \|\theta_2^{n-2}\|_{0,\Gamma_S} \right) + \right. \\ &\quad \left. \|\omega_1^n\|_{0,\Gamma_S} + \|\omega_2^n\|_{0,\Gamma_S} \right) \|\theta_1^n\|. \end{aligned} \quad (56)$$

Combining (55) and (56) gives the estimate:

$$\begin{aligned} \|\theta_1^n\| &\leq \frac{C_{\Gamma_S}}{\alpha} \left(2(\|\omega_1^1\|_{0,\Gamma_S} + \|\omega_2^1\|_{0,\Gamma_S}) + \|\omega_1^n\|_{0,\Gamma_S} + \|\omega_2^n\|_{0,\Gamma_S} \right. \\ &\quad \left. + 8 \sum_{j=1}^n (\|\omega_1^j\|_{0,\Gamma_S} + \|\omega_2^j\|_{0,\Gamma_S}) \right) \quad \text{for } n \geq 1. \end{aligned} \quad (57)$$

Estimates for $\|\omega_1^n\|_{0,\Gamma_S}$, $n \geq 2$, can be obtained using the following relation:

$$\begin{aligned} \omega_1^n &= \frac{1}{\Delta t} (R_h - I) \left(\frac{3}{2} \varsigma^{n-1} - 2\varsigma^{n-2} + \frac{1}{2} \varsigma^{n-3} \right) \\ &= \frac{3}{2\Delta t} \int_{t_{n-2}}^{t_n} (R_h - I) \varsigma_{,t} ds - \frac{1}{2\Delta t} \int_{t_{n-2}}^{t_{n-1}} (R_h - I) \varsigma_{,t} ds. \end{aligned}$$

If we restrict $R_h \varsigma$ to Γ_S and use the standard local approximation property:

$$\|u - R_h u\|_{0,\Gamma_S} \leq Ch_K^{r+1} |u|_{r+1,\Gamma_S},$$

with $r \geq 1$ the highest index of the Sobolev space for which $\varsigma \in H^{r+1}(\mathcal{T}_h)$, then we obtain the following estimates:

$$\begin{aligned} \sum_{j=2}^n \|\omega_1^j\|_{0,\Gamma_S} &\leq \frac{2}{\Delta t} \int_{t_0}^{t_n} \|(R_h - I) \varsigma_{,t}\|_{0,\Gamma_S} ds \\ &\leq \frac{2Ch^{r+1}}{\Delta t} \int_{t_0}^{t_n} |\varsigma_{,t}|_{r+1,\Gamma_S} ds, \end{aligned} \quad (58)$$

$$\|\omega_1^1\|_{0,\Gamma_S} \leq \frac{Ch^{r+1}}{\Delta t} \int_{t_0}^{t_1} |\varsigma_{,t}|_{r+1,\Gamma_S} ds, \quad (59)$$

$$\|\omega_1^n\|_{0,\Gamma_S} \leq \frac{2Ch^{r+1}}{\Delta t} \int_{t_{n-2}}^{t_n} |\varsigma_{,t}|_{r+1,\Gamma_S} ds, \quad \text{for } n \geq 2. \quad (60)$$

Estimates for $\|\omega_2^n\|_{0,\Gamma_S}$ can be obtained using a Taylor series expansion with remainder, see also ([21], (1.52), pages 14–17), yielding:

$$\|\omega_2^1\|_{0,\Gamma_S} \leq \int_{t_0}^{t_1} \|\varsigma_{,tt}\|_{0,\Gamma_S} ds, \quad (61)$$

$$\|\omega_2^n\|_{0,\Gamma_S} \leq \frac{5}{4} \Delta t \int_{t_{n-2}}^{t_n} \|\varsigma_{,ttt}\|_{0,\Gamma_S} ds, \quad \text{for } n \geq 2. \quad (62)$$

$$\sum_{j=1}^n \|\omega_2^j\|_{0,\Gamma_S} \leq \frac{5}{4} \Delta t \int_{t_0}^{t_n} \|\varsigma_{,ttt}\|_{0,\Gamma_S} ds + \int_{t_0}^{t_1} \|\varsigma_{,tt}\|_{0,\Gamma_S} ds. \quad (63)$$

If we introduce (58-59) and (61-63) into (55) we obtain the final bound for $\|\theta_2^n\|_{\Gamma_S}$:

$$\begin{aligned} \|\theta_2^n\|_{0,\Gamma_S} &\leq Ch^{r+1} \left(4 \int_{t_0}^{t_n} |\varsigma,t|_{r+1,\Gamma_S} ds + \frac{5}{2} \int_{t_0}^{t_1} |\varsigma,t|_{r+1,\Gamma_S} ds \right) \\ &\quad + \frac{5}{2} \Delta t^2 \int_{t_0}^{t_n} \|\varsigma,ttt\|_{0,\Gamma_S} ds + \frac{5}{2} \Delta t \int_{t_0}^{t_1} \|\varsigma,tt\|_{0,\Gamma_S} ds. \end{aligned}$$

Similarly, introducing (58-63) into (57) and taking $\alpha = \frac{1}{\Delta t}$, which implies that the stabilization parameter η must satisfy $\eta = \frac{1}{\Delta t}$, results in the bound for $\|\theta_1^n\|$:

$$\begin{aligned} \|\theta_1^n\| &\leq C_{\Gamma_S} \left(Ch^{r+1} \left(16 \int_{t_0}^{t_n} |\varsigma,t|_{r+1,\Gamma_S} ds + 10 \int_{t_0}^{t_1} |\varsigma,t|_{r+1,\Gamma_S} ds \right. \right. \\ &\quad \left. \left. + 2 \int_{t_{n-2}}^{t_n} |\varsigma,t|_{r+1,\Gamma_S} ds \right) + 10 \Delta t^2 \int_{t_0}^{t_n} \|\varsigma,ttt\|_{0,\Gamma_S} ds \right. \\ &\quad \left. + 10 \Delta t^2 \int_{t_0}^{t_1} \|\varsigma,tt\|_{0,\Gamma_S} ds + \frac{5}{4} \Delta t^2 \int_{t_{n-2}}^{t_n} \|\varsigma,ttt\|_{0,\Gamma_S} ds \right) \end{aligned}$$

Estimates for ρ_1^n and ρ_2^n are obtained using the definition of the DG-norm (43), the coercivity and boundedness of $B_h^n(\phi_h^n, v)$, see Lemma 5.1, and interpolation estimate (4.22) in [2] (which can be straightforwardly extended to general convex elements with a Lipschitz boundary), with $v = u_I$ an elementwise discontinuous interpolant:

$$\begin{aligned} \|\rho_1^n\| &= \|R_h \phi^n - \phi^n\| \leq \frac{C_a}{\alpha} \min_{v \in V_h^p} \|\phi^n - v\| \leq \frac{C_a}{\alpha} h^r |\phi^n|_{r+1,\Omega} \\ &\leq \frac{C_a}{\alpha} h^r \left(|\phi^0|_{r+1,\Omega} + \int_{t_0}^{t_n} |\phi,t|_{r+1,\Omega} ds \right). \end{aligned}$$

Similarly, when we restrict $R_h \phi^n$ to the surface Γ_S , we obtain the estimate:

$$\|\rho_2^n\|_{0,\Gamma_S} \leq \frac{C_a}{\alpha} h^{r+1} \left(|\varsigma^0|_{r+1,\Gamma_S} + \int_{t_0}^{t_n} |\varsigma,t|_{r+1,\Gamma_S} ds \right).$$

Using the relation $\varsigma^n = -\zeta^n$ completes the proof. \square

6 Algebraic system for the discontinuous Galerkin discretization

The algebraic equations for the discontinuous Galerkin formulation (29-34) are obtained by representing the potential function $\phi_h^n \in V_h^p$ in each element

$K \in \mathcal{T}_h$ with the following polynomial representation:

$$\phi_h^n(x) = \sum_{j=1}^{N_p} \phi_{K,j} \mathcal{N}_{K,j}(x), \quad \forall x \in K, \quad (64)$$

with N_p the number of polynomial coefficients in the DG discretization. The polynomial order p is chosen equal to 1, 2 or 3. Similar expressions are used for the time-derivative of the potential function ζ_h^n and the test functions v_1, v_2 . As elements we use quadrilateral elements (hexahedral elements in three dimensions), but the procedure to derive the algebraic equations is essentially independent of the different type of elements. The basis functions $\mathcal{N}_{K,j}$ are defined as:

$$\mathcal{N}_{K,j}(x) = \widehat{\mathcal{N}}_j \circ F_K^{-1}(x),$$

with $\xi \in [-1, 1]^d$ and $d = \dim(\Omega)$, using the isoparametric mapping:

$$F_K : [-1, 1]^d \rightarrow K; x \mapsto \sum_{j=1}^{N_p} \widehat{x}_j \widehat{\mathcal{N}}_j(\xi), \quad (65)$$

with $\widehat{x}_j \in \mathbb{R}^d$ the coordinates of the N_p points used to define Lagrangian basis functions on the element K , see [11]. The basis functions $\widehat{\mathcal{N}}_j(\xi)$ are defined as the tensorproduct of p -th order polynomials in the local coordinates ξ_i , $i = 1, \dots, d$.

The first step in deriving the algebraic equations for the discontinuous Galerkin discretization is to approximate the local lifting operator $R_{\mathcal{F}} \in \mathbb{R}^d$, defined in (10), which we briefly summarize. For more details see [20]. Since $R_{\mathcal{F}}$ is only non-zero in the two elements K_L and K_R which are connected to the face \mathcal{F} , we have:

$$\begin{aligned} & \int_{K_L} R_{\mathcal{F},L}([\phi_h]) \cdot v_L dx + \int_{K_R} R_{\mathcal{F},R}([\phi_h]) \cdot v_R dx \\ &= \frac{1}{2} \int_{\mathcal{F}} (\phi_{L,h} n_L + \phi_{R,h} n_R) \cdot (v_L + v_R) ds, \quad \forall v_L, v_R \in \Sigma_h^p. \end{aligned} \quad (66)$$

We approximate the k -th component of the lifting operator and test functions v_L, v_R as:

$$(R_{\mathcal{F},K}([\phi_h])(x))_k = \sum_{j=1}^{N_p} R_{jk}^{K;\mathcal{F}} \mathcal{N}_{K,j}(x), \quad \forall x \in K, \quad (67)$$

$$(v_K(x))_k = \sum_{j=1}^{N_p} v_{K,jk} \mathcal{N}_{K,j}(x), \quad \forall x \in K, \quad (68)$$

and introduce the following vectors $\mathcal{I}^{K;\mathcal{F}} \in \mathbb{R}^{N_p}$, matrices: $\mathcal{A}^K, \mathcal{B}^K, \mathcal{C}^{LR;\mathcal{F}}, \mathcal{S}^{K;\mathcal{F}} \in \mathbb{R}^{N_p \times N_p}$, and tensors $\mathcal{D}^{LR;\mathcal{F}}, E^{L;\mathcal{F}}, F^{LR;\mathcal{F}} \in \mathbb{R}^{N_p \times N_p \times d}$, which are defined as:

$$\begin{aligned}
\mathcal{A}_{ij}^K &:= \int_K \mathcal{N}_{K,i} \mathcal{N}_{K,j} dx, & \mathcal{B}_{ij}^K &:= \int_K \nabla_h \mathcal{N}_{K,i} \cdot \nabla_h \mathcal{N}_{K,j} dx, \\
\mathcal{C}_{ij}^{LR;\mathcal{F}} &:= \int_{\mathcal{F}} \mathcal{N}_{L,j} n_L \cdot \nabla_h \mathcal{N}_{R,i} ds, & \mathcal{D}_{ijk}^{LR;\mathcal{F}} &:= \int_{\mathcal{F}} n_{L,k} \mathcal{N}_{L,i} \mathcal{N}_{R,j} ds, \\
\mathcal{I}_i^{K;\mathcal{F}} &:= \int_{\mathcal{F}} g_N \mathcal{N}_{K,i} ds, & \mathcal{S}_{ij}^{K;\mathcal{F}} &:= \int_{\mathcal{F}} \mathcal{N}_{K,i} \mathcal{N}_{K,j} ds \quad (69) \\
E_{njk}^{K;\mathcal{F}} &:= \sum_{i=1}^{N_p} (\mathcal{A}^{-1})_{ni}^K \mathcal{D}_{ijk}^{KK;\mathcal{F}}, & F_{njk}^{LR;\mathcal{F}} &:= \sum_{i=1}^{N_p} (\mathcal{A}^{-1})_{ni}^R \mathcal{D}_{jik}^{LR;\mathcal{F}}. \quad (70)
\end{aligned}$$

Here, K is the index of the elements in the tessellation \mathcal{T}_h , which contains $N_{\mathcal{T}}$ elements, and L, R are the element indices of the two elements connected to each side of the face with index \mathcal{F} . The integrals are computed by transforming them to the reference element $[-1, 1]^d$ using the mapping (65) and approximating them with Gauss quadrature rules. Since most of the integrands are of polynomial type we choose the number of quadrature points such that the Gauss quadrature rule is exact.

If we introduce (67-68) into (66), and use the fact that this equation must be satisfied for arbitrary test functions v_L, v_R , then we obtain the following relations for the coefficients of the lifting operator in (67):

$$\begin{aligned}
R_{nk}^{L;\mathcal{F}} &= \frac{1}{2} \sum_{j=1}^{N_p} \left(E_{njk}^{L;\mathcal{F}} \phi_{L,j} + F_{njk}^{RL;\mathcal{F}} \phi_{R,j} \right), \\
R_{nk}^{R;\mathcal{F}} &= \frac{1}{2} \sum_{j=1}^{N_p} \left(F_{njk}^{LR;\mathcal{F}} \phi_{L,j} + E_{njk}^{R;\mathcal{F}} \phi_{R,j} \right).
\end{aligned}$$

The algebraic equations for the discontinuous Galerkin discretization are obtained by introducing the polynomial representation for ϕ_h^n and ζ_h^n and the test functions, defined in (64), into (35-36) and using the fact that these equations must be satisfied for arbitrary test functions. After some lengthy, but straightforward algebraic manipulations the algebraic equations for the discontinuous Galerkin discretization at the time level $t = t_n$ can be represented in matrix form as:

$$\mathcal{M} \Phi^n = F, \quad (71)$$

with $\mathcal{M} \in \mathbb{R}^{N_p N_{\mathcal{T}} \times N_p N_{\mathcal{T}}}$, $\Phi \in \mathbb{R}^{N_p N_{\mathcal{T}}}$ and $F \in \mathbb{R}^{N_p N_{\mathcal{T}}}$, where $N_{\mathcal{T}}$ denotes the number of elements in the mesh. The matrix \mathcal{M} at each time level $n \geq 2$

consists of $N_{\mathcal{T}}^2$ blocks $[M] \in \mathbb{R}^{N_p \times N_p}$ and is constructed as follows:

$$a) \quad \text{Initialize } \mathcal{M} = 0, \quad (72)$$

$$b) \quad [M_{ij}]^{KK} = \mathcal{B}_{ij}^K, \quad \forall i, j \in \{1, \dots, N_p\}, \forall K \in \{1, \dots, N_{\mathcal{T}}\}, \quad (73)$$

$$c) \quad [M_{ij}]^{LL} \leftarrow \sum_{\mathcal{F} \in \mathcal{F}_h^I} \left\{ [M_{ij}]^{LL} - \frac{1}{2} (\mathcal{C}_{ij}^{LL;\mathcal{F}} + \mathcal{C}_{ji}^{LL;\mathcal{F}}) \right. \\ \left. + \frac{1}{4} \tilde{\eta}_{\mathcal{F}} \sum_{k=1}^d \sum_{n=1}^{N_p} \left(E_{njk}^{L;\mathcal{F}} D_{ink}^{LL;\mathcal{F}} + F_{njk}^{LR;\mathcal{F}} D_{ink}^{LR;\mathcal{F}} \right) \right\}, \quad (74)$$

$$d) \quad [M_{ij}]^{LR} \leftarrow \sum_{\mathcal{F} \in \mathcal{F}_h^I} \left\{ [M_{ij}]^{LR} - \frac{1}{2} (\mathcal{C}_{ij}^{RL;\mathcal{F}} + \mathcal{C}_{ji}^{LR;\mathcal{F}}) \right. \\ \left. + \frac{1}{4} \tilde{\eta}_{\mathcal{F}} \sum_{k=1}^d \sum_{n=1}^{N_p} \left(F_{njk}^{RL;\mathcal{F}} D_{ink}^{LL;\mathcal{F}} + E_{njk}^{R;\mathcal{F}} D_{ink}^{LR;\mathcal{F}} \right) \right\}, \quad (75)$$

$$e) \quad [M_{ij}]^{RL} \leftarrow \sum_{\mathcal{F} \in \mathcal{F}_h^I} \left\{ [M_{ij}]^{RL} - \frac{1}{2} (\mathcal{C}_{ij}^{LR;\mathcal{F}} + \mathcal{C}_{ji}^{RL;\mathcal{F}}) \right. \\ \left. + \frac{1}{4} \tilde{\eta}_{\mathcal{F}} \sum_{k=1}^d \sum_{n=1}^{N_p} \left(E_{njk}^{L;\mathcal{F}} D_{ink}^{RL;\mathcal{F}} + F_{njk}^{LR;\mathcal{F}} D_{ink}^{RR;\mathcal{F}} \right) \right\}, \quad (76)$$

$$f) \quad [M_{ij}]^{RR} \leftarrow \sum_{\mathcal{F} \in \mathcal{F}_h^I} \left\{ [M_{ij}]^{RR} - \frac{1}{2} (\mathcal{C}_{ij}^{RR;\mathcal{F}} + \mathcal{C}_{ji}^{RR;\mathcal{F}}) \right. \\ \left. + \frac{1}{4} \tilde{\eta}_{\mathcal{F}} \sum_{k=1}^d \sum_{n=1}^{N_p} \left(F_{njk}^{RL;\mathcal{F}} D_{ink}^{RL;\mathcal{F}} + E_{njk}^{R;\mathcal{F}} D_{ink}^{RR;\mathcal{F}} \right) \right\}, \quad (77)$$

$$g) \quad [M_{ij}]^{K,K} \leftarrow \sum_{\mathcal{F} \in \mathcal{F}_h^S} \left\{ [M_{ij}]^{K,K} + \frac{9}{4\Delta t^2} \mathcal{S}_{ij}^{K;\mathcal{F}} \right\}, \\ \forall i, j \in \{1, \dots, N_p\}, \quad (78)$$

with $\tilde{\eta}_{\mathcal{F}} = \eta_{\mathcal{F}} + n_{\mathcal{F}}$. For $n = 1$ the matrix \mathcal{M} is identical, except that the coefficient $\frac{9}{4}$ is equal to 1 in (78). The righthand side vector F is constructed as follows:

$$h) \quad [F_i]^K = 0, \quad \forall i \in \{1, \dots, N_p\}, \forall K \in \{1, \dots, N_{\mathcal{T}}\},$$

$$i) \quad [F_i]^K \leftarrow \sum_{\mathcal{F} \in \mathcal{F}_h^N} \left\{ [F_i]^K + \mathcal{I}_i^{K;\mathcal{F}} \right\}, \quad \forall i \in \{1, \dots, N_p\},$$

$$j) \quad [F_i]^K \leftarrow \sum_{\mathcal{F} \in \mathcal{F}_h^S} \left\{ [F_i]^K + \sum_{j=1}^{N_p} \left(\frac{2}{\Delta t} s_{K,j}^{n-1} - \frac{1}{2\Delta t} s_{K,j}^{n-2} \right) \mathcal{S}_{ij}^{K;\mathcal{F}} \right\},$$

$$\forall i \in \{1, \dots, N_p\},$$

$$k) \quad [F_i]^K \leftarrow \sum_{\mathcal{F} \in \mathcal{F}_h^S} \left\{ [F_i]^K + \sum_{j=1}^{N_p} \left(\frac{3}{\Delta t^2} \phi_{K,j}^{n-1} - \frac{3}{4\Delta t^2} \phi_{K,j}^{n-2} \right) \mathcal{S}_{ij}^{K;\mathcal{F}} \right\},$$

$$\forall i \in \{1, \dots, N_p\},$$

with the obvious modification for $n = 1$, see (33) and (36). The construction of the matrix on a general unstructured mesh is now straightforward. First, we initialize all entries to zero, followed by a loop over all elements. Next under c) till g) we loop over all faces, compute the block matrix entries for the two elements with indices L and R which are connected to the face \mathcal{F} , and store these entries in the blocks $[M]^{LR}$. This whole procedure does not depend on the chosen type of elements and is suitable for any unstructured mesh.

The resulting linear system (71) is symmetric and positive definite, because the bilinear form (35) is coercive, bounded and symmetric, and can be solved efficiently, either directly with a Choleski decomposition or iteratively with a preconditioned conjugate gradient method. The linear system is non-singular due to the stabilization terms in the DG discretization, but has a large condition number because one of the eigenvalues is close to zero. This does not result in a serious loss of accuracy when the zero mean condition $\int_{\Omega} \phi_h dx = 0$ is imposed, which ensures that a unique solution is obtained in the space $H^1(\mathcal{T}_h) + V$. The zero mean condition can be imposed by correcting the numerical solution after each iteration step in the conjugate gradient method: $\phi_h \leftarrow \phi_h - \bar{\phi}_h$, with $\bar{\phi}_h = \frac{1}{|\Omega|} \int_{\Omega} \phi_h dx$, after which the new expansion coefficients of ϕ_h are obtained in each element K using an $L^2(K)$ projection onto the basis functions $\mathcal{N}_{K,j}$. In practice, it works however equally well if the zero mean condition is imposed after the solution of the linear system is obtained. Introducing the zero mean condition directly into the linear system is not beneficial since this results in a non-symmetric linear system and is computationally more expensive.

7 Analysis of dispersion and dissipation error of the DG discretization

A critical factor in the numerical simulation of free surface waves is the dissipation and dispersion of the waves due to numerical discretization errors. For this purpose we first conduct an analysis of the semi-discrete DG formulation (28) for time-harmonic waves. Next, we conduct a Fourier analysis of the fully discrete scheme. This will provide us with information on the dependence of the dissipation and dispersion properties of the numerical scheme on the time-step and spatial discretization.

7.1 Analysis for time-harmonic waves

In this section we assume that the potential function is time-harmonic:

$$\phi(x, y, z, t) = \widehat{\phi}(x, y, z) \exp(-i\omega t), \quad (79)$$

with frequency ω and $i = \sqrt{-1}$. If we introduce the time-harmonic ansatz (79) into (28) we obtain the following eigenvalue problem:

$$-\omega^2(\widehat{\phi}, v)_{\Gamma_S} + B_h^0(\widehat{\phi}, v) = 0, \quad \forall v \in V_h^p, \quad (80)$$

which can be transformed into a linear system using the results from the previous section. Without loss of generality we now assume that the elements are ordered such that the first N_S elements are connected to the free surface. The linear system for B_h^0 then is obtained by adding the contributions (72–77) to the matrix $\mathcal{B}^0 \in \mathbb{R}^{N_p N_T \times N_p N_T}$. The contribution $-\omega^2(\widehat{\phi}, v)_{\Gamma_S}$ results in the matrix $\mathcal{B}_{SS} \in \mathbb{R}^{N_p N_S \times N_p N_S}$, which is a block diagonal matrix $\mathcal{B}_{SS} = \text{diag}(\mathcal{S}^{1;1}, \dots, \mathcal{S}^{N_S;N_S})$, with blocks $\mathcal{S}^{K;\mathcal{F}}$ defined in (69). We partition the matrix \mathcal{B}^0 now as:

$$\mathcal{B}^0 = \begin{pmatrix} \mathcal{B}_{SS}^0 & (\mathcal{B}_{DS}^0)^T \\ \mathcal{B}_{DS}^0 & \mathcal{B}_{DD}^0 \end{pmatrix},$$

with $\mathcal{B}_{DD}^0 \in \mathbb{R}^{N_p(N_T-N_S) \times N_p(N_T-N_S)}$, $\mathcal{B}_{SS}^0 \in \mathbb{R}^{N_p N_S \times N_p N_S}$, and $\mathcal{B}_{DS}^0 \in \mathbb{R}^{N_p(N_T-N_S) \times N_p N_S}$, where the suffix \mathcal{S} refers to coefficients in elements connected to the free surface. The eigenvalue problem can now be expressed as:

$$-\omega^2 \begin{pmatrix} \mathcal{B}_{SS} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \widehat{\phi}_S \\ 0 \end{pmatrix} + \begin{pmatrix} \mathcal{B}_{SS}^0 & (\mathcal{B}_{DS}^0)^T \\ \mathcal{B}_{DS}^0 & \mathcal{B}_{DD}^0 \end{pmatrix} \begin{pmatrix} \widehat{\phi}_S \\ \widehat{\phi}_D \end{pmatrix} = 0, \quad (81)$$

with $\widehat{\phi}_S \in \mathbb{R}^{N_p N_S}$, $\widehat{\phi}_D \in \mathbb{R}^{N_p(N_T - N_S)}$. It is straightforward to eliminate $\widehat{\phi}_D$ from (81):

$$\widehat{\phi}_D = -(\mathcal{B}_{DD}^0)^{-1} \mathcal{B}_{DS}^0 \widehat{\phi}_S,$$

and we obtain the following eigenvalue problem for $\widehat{\phi}_S$:

$$L_S \widehat{\phi}_S = \mu \mathcal{B}_{SS} \widehat{\phi}_S, \quad (82)$$

with $\mu = \omega^2$ and $L_S \in \mathbb{R}^{N_S \times N_S}$ defined as:

$$L_S = \mathcal{B}_{SS}^0 - (\mathcal{B}_{DS}^0)^T (\mathcal{B}_{DD}^0)^{-1} \mathcal{B}_{DS}^0.$$

Due to the symmetry and coercivity of B_h^0 and B_h , both \mathcal{B}^0 and \mathcal{B}_{SS} are positive definite matrices, and it is straightforward to show that also L_S is positive definite. Using the matrix square root of \mathcal{B}_{SS} and the transformation $\widehat{y} = (\mathcal{B}_{SS})^{\frac{1}{2}} \widehat{\phi}_S$, we can transform (81) into the eigenvalue problem:

$$\widehat{L}_S \widehat{y}_S = \mu \widehat{y}_S, \quad (83)$$

with $\widehat{L}_S = (\mathcal{B}_{SS})^{-\frac{1}{2}} L_S (\mathcal{B}_{SS})^{-\frac{1}{2}}$ a positive definite matrix. All eigenvalues μ_i , $i = 1, \dots, N_S$ are therefore real and strictly positive, hence the eigenvalues $\omega_i = \pm \sqrt{\mu_i}$, $i = 1, \dots, N_S$ are also real. The L^2 -norm of the potential ϕ at the free surface, which is defined as:

$$\|\phi\|_{\Gamma_S} = (\phi, \overline{\phi})_{\Gamma_S}^{\frac{1}{2}} = \left\| \widehat{\phi}_S \right\|_{\Gamma_S},$$

where an overbar denotes the complex conjugate, is therefore constant, since ω_i in (79) is real. The spatial discretization therefore does not generate any growing or decaying modes in (79), is energy conserving and stable, and only generates a dispersion error. This also applies to the wave height ζ , since the wave height is defined as $\zeta = -\frac{\partial \phi}{\partial t}|_{\Gamma_S}$. Note, contrary to [19] or [22] the stability of the DG discretization does not depend on the mesh smoothness. Also, the numerical scheme does not require any extra damping terms in the free surface equation (2) to guarantee stability, as for instance in [19].

7.2 Analysis of fully discrete scheme

For the analysis of fully discrete scheme we assume, in addition to the time-harmonic potential, a two-dimensional domain with a uniform mesh with $N_x \times N_z$ coordinates $x_j = j\Delta x$, $z_m = m\Delta z$ and periodic boundary conditions in the x -direction. We use the Fourier ansatz for the coefficients $\widehat{\phi}_j$ in the DG discretization at time level $t_n = n\Delta t$:

$$\widehat{\phi}_j(z, t_n) = \lambda^n \exp(ikj\Delta x) \widehat{\phi}^F(z), \quad (84)$$

with $\lambda^n = \exp(-\omega n \Delta t)$ the amplification factor, k the wavenumber, and a similar relation for the coefficients $\widehat{\zeta}_j$.

If we introduce (84) into (71) we obtain the following generalized eigenvalue problem:

$$\begin{pmatrix} 0 & S \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \widehat{\phi}_S^F \\ \widehat{\phi}_D^F \end{pmatrix} = \mu(\lambda, \Delta t) M^F(k) \begin{pmatrix} \widehat{\phi}_S^F \\ \widehat{\phi}_D^F \end{pmatrix}, \quad (85)$$

with $\widehat{\phi}_S^F \in \mathbb{R}^{N_p}$ and $\widehat{\phi}_D^F \in \mathbb{R}^{N_p(N_z-1)}$, where the suffix \mathcal{S} refers to coefficients in elements connected to the free surface and D to elements not connected to the free surface. The matrix $M^F \in \mathbb{R}^{N_p N_z \times N_p N_z}$ is an Hermitian positive definite block-tridiagonal matrix with the following block-structure:

$$[M^F(k)]_m = \text{TriDiag}(A \exp(-\imath k \Delta x), B_m, A \exp(-\imath k \Delta x)),$$

with $A, B_m \in \mathbb{R}^{N_p \times N_p}$ for $m \in \{1, \dots, N_z\}$. The matrix $S \in \mathbb{R}^{N_p \times N_p}$ is given by (69). The eigenvalues μ_i are related to the amplification factor λ and time step Δt as:

$$\left(\frac{9}{4} + \frac{\Delta t^2}{\mu}\right) \lambda^4 - 6\lambda^3 + \frac{11}{2} \lambda^2 - 2\lambda + \frac{1}{4} = 0.$$

Hence for each eigenvalue μ_i we obtain four amplification factors $\lambda_{i,j}$, $j = 1, \dots, 4$. Since the matrix M^F is Hermitian and positive definite it is straightforward to solve the generalized eigenvalue problem, which will have only N_p non-zero real eigenvalues. The eigenvalues μ_i are computed with MATLAB for a wide range of Δt and $k \in [0, 2\pi)$ values, resulting for all cases in amplification factors $\lambda_{i,j}$ with a modulus less than or equal to one, hence the numerical discretization is unconditionally stable.

The Fourier analysis is also used to compute the dispersion and dissipation properties of the numerical scheme. If we solve the generalized eigenvalue problem (85) for a particular wave number k then the frequency and dissipation of the discrete modes are the real and imaginary part of $\frac{2}{\Delta t} \ln \lambda_{i,j}$, respectively, with one of the frequencies close to the frequency of the physical mode. In Section 8.3 we use this technique to compute the dispersion and dissipation of the numerical scheme as a function of the mesh size and time-step.

8 Numerical examples

In this section we discuss the numerical results of model problems which aim to verify the analysis and demonstrate the capabilities of the method

for the simulation of free surface water waves. After some initial test cases to validate the code and confirm the theoretical orders of accuracy for the Laplace equation, we consider two cases: (i) an harmonic wave in an unbounded domain, and (ii) the simulation of water waves generated by a wave maker.

For the free surface simulations in an unbounded domain (in the horizontal direction) we choose the domain as $[-1, 1] \times [-1, 0]$ with periodic boundary conditions given by (3) at both ends $x = \pm 1$ of the domain, and a flat bottom at $z = -1$ with a homogenous Neumann boundary condition representing a slip flow boundary. As initial free surface we use the projection of the analytic solution onto the finite element basis functions. The analytic solution of this problem is given by:

$$\phi = \phi_0 \cosh(k(z + 1)) \cos(\omega t - kx), \quad (86)$$

where ϕ_0 denotes the amplitude of the velocity potential, k the wave number, which is related to the wavelength λ as $k = 2\pi/\lambda$, and ω the frequency of the oscillations, which satisfies the dispersion relation:

$$\omega^2 = k \tanh(k),$$

All quantities in this section are in dimensionless form. For details see Section 2. The parameter η in the DG algorithm was set equal to $\eta = h^{-2p}$, with p the order of the polynomial basis functions, but the results do not significantly depend on this parameter.

8.1 Laplace equation

As a first step to verify the computer code we switch off the free-surface condition (2) and consider the Laplace equation with only Neumann boundary conditions. At the free surface we use:

$$\frac{\partial \phi}{\partial n} = \phi_0 k \sinh(k) \cos(kx).$$

At the other boundaries homogeneous Neumann boundary conditions are applied. As mentioned in the Section 3, the solution of this problem is indeterminate up to a constant. We can, however, find a unique solution in the space $H^1(\mathcal{T}_h) + V$. To keep the L^2 -errors within reasonable bounds we choose a wave length $\lambda = 1$ and amplitude $\phi_0 = 0.05$ in (86) for the basis functions with polynomial order $p = 1$ and 2, and $\lambda = 0.25$ and $\phi_0 = 0.1$ for $p = 3$, respectively. This makes it possible to easily compare cases with different order of polynomial basis functions. To illustrate the insensitivity

of the computed solution to the mesh smoothness, which is an important benefit of the DG scheme, we have chosen two meshes, a uniform and a randomly distorted mesh where the vertices have a random displacement of 30% of the edge length. The L^2 -error in the velocity potential as a function of the mesh size (on log – log scale) is shown in Figure 1, which confirms that the optimal order of convergence $O(h^{p+1})$ is achieved on both meshes, with the quality of the solution almost unaltered by the mesh smoothness.

8.2 Time harmonic waves in an unbounded domain

As a next step we consider time harmonic waves in an unbounded domain with free-surface condition (2). The initial conditions for the free surface potential and wave height at $z = 0$ are:

$$\phi = \phi_0 \cosh(k) \cos(kx), \quad \zeta = \phi_0 \omega \cosh(k) \sin(kx).$$

For the free surface computations we restrict ourselves to $p = 1$ and 2. We choose the wave length $\lambda = 1$ and ϕ_0 such that the maximum amplitude of the free-surface height is 0.05, hence 5% of the water depth. First, we consider the $L^2(\Omega)$ -error in the velocity potential and the $L^2(\Gamma_S)$ -error in the wave height, which both converge with nearly optimal order, see Figures 2 and 3. This error virtually only consists of a frequency error, the dissipation is very small, see also Figures 5-8.

The free-surface height after 20 periods is shown in Figure 4. The mesh size is (approximately) $h = 0.125$ and 0.0625 for the basis functions with order $p = 1$ and $h = 0.125$ for $p = 2$, on a uniform as well as a randomly distorted mesh. The insensitivity of the computed wave solution to the mesh smoothness, already proved by the theoretical analysis, is confirmed and the computations did not require any artificial modification of the free surface boundary condition, such as for instance in [19]. The dissipation error in both cases, $p = 1$ and 2, is very small, but for $p = 1$ the phase error on the mesh with $h = 0.125$ is significantly larger than for $p = 2$. For $p = 1$ the computed wave is moving faster than the actual wave. After mesh refinement to $h = 0.0625$ the phase error for $p = 1$ is significantly reduced, see Figure 4. For $p = 2$ the numerical solution matches even closer with the analytic solution on the mesh with $h = 0.125$ and we can conclude that for simulating wave motion it is more efficient to increase the polynomial order of the basis functions than reducing the mesh size.

8.3 Analysis of dispersion and dissipation error

Since the time integration scheme is unconditionally stable the time step is only dictated by accuracy requirements. In order to obtain insight into the dependency of the accuracy on the time step and mesh size we performed the fully discrete Fourier analysis discussed in Section 7.2. The results for the frequency error and dissipation for the DG discretizations with linear and quadratic polynomial basis functions are shown in Figures 5-8. They show that the dissipation is very small and the error is dictated by the frequency error. Also, there is an optimal relation between the time step and mesh size. Figures 5 and 7 can be used to find optimal values for the time step, depending on the maximum wave frequency expected in the simulation and mesh size used. We also conducted a discrete Fourier analysis for a two-step Adams-Moulton time integration method. For this method the dissipation is on all meshes at the level of machine accuracy, but the simulation results did not differ substantial in comparison with the backward difference time integration method since the frequency error is the dominant factor.

8.4 Simulation of water waves generated by a wave maker

Next we present the numerical simulation of waves generated by a wave maker in a model basin. In this case the domain is $[0, 10] \times [-1, 0]$. We assume, homogenous Neumann boundary conditions at the bottom $z = -1$ and at the end of the domain at $x = 10$ opposite to the wave maker. The initial free surface height and velocity potential are zero, and a time periodic Neumann boundary condition is applied at the wave maker, which governs the normal velocity. The normal velocity profile is linear, starting with zero at the bottom and has an amplitude of 0.05. The frequency of the time harmonic motion is 1.8138. All simulations are done only on the random mesh. We choose quadrilateral elements with 3 cases of varying mesh size and polynomial degree, $h = 0.125$ with $p = 1$ and 2 and also for $h = 0.0625$ when $p = 1$. The wave profile in the domain at $T = 20$, when the wave starts approaching the wall opposite the wave maker, is presented in Figure 9. For a comparison see also [22]. Next, the profile at $T = 38.22$, when the wave gains full height against the wall, is shown in Figure 10, and finally the profile at $T = 75.32$, when the wave from the opposite wall is travelling back into the domain and affects the pattern of the generated wave, is given in Figure 11. Up to $T = 38.22$ there is hardly any difference visible between the results for $p = 1$ and 2 and the different meshes. At $T = 75.32$ the results for $p = 1$ on the coarse mesh are influenced by the

mesh size, but on the finer mesh with $h = 0.0625$ they are close to the $p = 2$ results on the mesh with size $h = 0.125$. Considering the computational expense it is therefore more efficient to use a polynomial order $p = 2$ for these wave simulations in order to obtain a higher accuracy than a finer mesh. Beyond $p = 2$ this requires, however, also a more accurate time integration technique. The extension of the backward difference or Adams-Moulton time integration methods to higher order is, however, not A-stable as a consequence of a famous theorem from Dahlquist. This causes serious stability problems in the simulations conducted with these methods and they will have to be replaced for instance by implicit Runge-Kutta methods to obtain an unconditionally stable discretization.

9 Conclusions

We have developed a discontinuous Galerkin finite element discretization for linear free surface gravity waves. The algorithm is unconditionally stable and does not suffer from a mesh dependent instability, which occurs when the Laplace equation and free surface boundary condition are only weakly coupled, as is commonly done in most algorithms. The discretization results in a symmetric positive definite linear system, which can be solved efficiently either with a Choleski decomposition or a preconditioned conjugate gradient method. A complete analysis of the stability and accuracy of the numerical discretization is conducted which shows that the method has an $L^2(\Gamma_S)$ -error for the wave height and an $L^2(\Omega)$ -error for the velocity potential both proportional to $O(\Delta t^2 + h^{p+1})$. The algorithm does not require a separate velocity reconstruction, for instance using finite differences or different order polynomials for velocity and potential, in order to preserve the accuracy in the wave height. Also, no smoothing or damping of unstable waves on a non-uniform mesh is necessary. An analysis of the DG discretization error for harmonic waves shows that the spatial discretization only produces a dispersion error and has zero dissipation. In addition, a fully discrete Fourier analysis is conducted to find optimal values of the time step as a function of the wave frequency, mesh size and wave length. This analysis shows that, in particular, the discretization using quadratic basis functions has a low numerical dissipation and dispersion error and is well suited for the simulation of wave motion. The theoretical results are confirmed by numerical simulations of water waves in an unbounded domain and waves generated by a wave maker. The DG discretization discussed in this paper will be used as the essential building block in the construction of a DG method for fully non-linear water waves which is a topic of current research.

Acknowledgement

This research was partly funded by the Maritime Research Institute Netherlands (MARIN), coordinator Dr. R.H.M. Huijsmans, and the Institute of Mechanics, Processes and Control Twente (IMPACT), whose support is gratefully acknowledged. The authors would also like to thank J.J. Sudirham, University of Twente, for her help during various stages of the project, in particular in computing the lifting operator, and Drs. R.M.J. van Damme, M.A. Botchev and O. Bokhove for their help and valuable comments on different aspects of the algorithm.

References

- [1] Arnold, D.N., (1982). An interior penalty finite element method with discontinuous elements, *SIAM J. Numer. Anal.* **19**(4), 742–760.
- [2] Arnold, D. N., Brezzi, F., Cockburn, B., and Marini, L.D., (2002). Unified analysis of discontinuous Galerkin Methods for Elliptic Problems, *SIAM J. Numer. Anal.* **39**(5), 1749–1779.
- [3] Bassi, F., Rebay, S., Mariotti, G., Pedinotti, S., and Savini, M., (1997). A high-order accurate discontinuous finite element method for inviscid and viscous turbomachinery flows. In Decupere, R., and Dibelius, G. (Eds.), *Proceedings of 2nd European Conference on Turbomachinery, Fluid Dynamics and Thermodynamics*, Antwerpen, Belgium, pp. 99–108.
- [4] Braess, H., and Wriggers, P., (2000). Arbitrary Lagrangian Eulerian finite element analysis of free surface flow. *Comput. Meth. Appl. Mech. Engrg*, **190**, 95–109.
- [5] Brenner, S.C. and Scott, L.R., (2002). *The mathematical theory of finite element methods*. Second edition. Springer-Verlag, New York.
- [6] Brezzi, F., Manzini, G., Marini, D., Pietra, P. and Russo, A., (1999). Discontinuous finite elements for diffusion problems. In *Atti Convegno in Onore di F. Brioschi*, Istituto Lombardo, Accademia di Scienze e Lettere, Milan, Italy, pp. 197–217.
- [7] Brezzi, F., Manzini, G., Marini, D., Pietra, P. and Russo, A., (2000): Discontinuous Galerkin Approximations for Elliptic Problems, *Numer. Meth. Part. Diff. Eq.*, 16(4), 365–378.
- [8] Cai, X., Langtangen, H.P., Nielsen, B.F., and Tiveto, A., (1998). A finite element method for fully nonlinear water waves, *J. Comput. Physics*, **143**, 544–568.

- [9] Castillo, P., (2002). Performance of discontinuous Galerkin methods for elliptic PDES, *SIAM J. Sci. Comput.*, **24**(2), 524–547.
- [10] Castillo, P., Cockburn, B., Perugia, I. and Schotzau, D., (2000). An a priori error analysis of the local discontinuous Galerkin method for elliptic problems, *SIAM. J. Numer. Anal.*, **38**(5), 1676–1706.
- [11] Ciarlet, P.G., (1991). Basic error estimates for elliptic problems. In Ciarlet, P.G., and Lions, J.L. (Eds), *Handbook of Numerical Analysis*, Vol. II (Part 1), North Holland, Amsterdam.
- [12] Cockburn, B., and Shu, C.-W., (1998). The local discontinuous Galerkin time-dependent method for convection-diffusion systems, *SIAM J. Numer. Anal.*, **35**, 2440–2463.
- [13] Dautray, R. and Lions, J.L., (1988). *Mathematical analysis and numerical methods for science and technology*, Vol. 2, Springer-Verlag.
- [14] Dawson, C., (2002). The $\mathcal{P}^{K+1} - \mathcal{S}^K$ local discontinuous Galerkin method for elliptic equations, *SIAM J. Numer. Anal.*, **40**(6), 2151–2170.
- [15] Dawson, C. and Proft, J. (2002). Discontinuous and coupled continuous/discontinuous Galerkin methods for the shallow water equations, *Comput. Meth. Appl. Mech. Engrg.*, **191**(41-42), 4721–4746.
- [16] Dawson, C. and Proft, J. (2003). Discontinuous/continuous Galerkin methods for coupling the primitive and wave continuity equations of shallow water, *Comput. Meth. Appl. Mech. Engrg.*, **192**(47-48), 5123–5145.
- [17] Lighthill, M.J., (1978). *Waves in Fluids*, Cambridge University Press, UK.
- [18] Newman, J.N., (1977). *Marine Hydrodynamics*, MIT Press, Cambridge, MA.
- [19] Robertson, I., and Sherwin, S.J., (1999). Free-surface flow simulation using *hp*/Spectral elements, *J. Comput, Phys.*, **155**, 26–53.
- [20] Sudirham, J.J., Van der Vegt, J.J.W. and Van Damme, R.M.J., (2003). A Study on Discontinuous Galerkin Finite Element Methods for Elliptic Problems, *Technical Memorandum Department of Applied Mathematics* 1690.
- [21] Thomée, V., (1997). *Galerkin finite element methods for parabolic problems*, Springer-Verlag.

- [22] Westhuis, J.H., (2001). *The numerical simulation of nonlinear waves in a hydrodynamic model test basin*, Ph.D. Thesis, University of Twente, Enschede, The Netherlands.
- [23] Whitham, G.B., (1974). *Linear and Nonlinear Waves*, Wiley, New York.
- [24] Wu, G.X., and Taylor, R.E., (1994). Finite element analysis of two-dimensional non-linear transient water waves, *Appl. Ocean Res.*, **16**, 363–372.

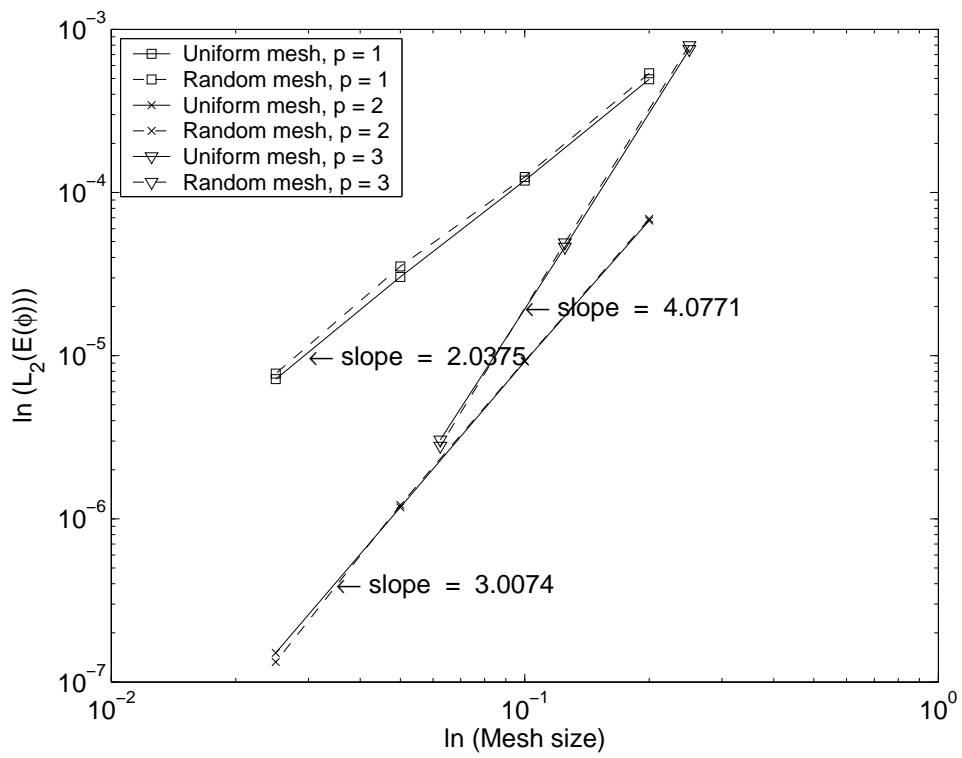


Figure 1: $L^2(\Omega)$ -error in the DG discretization of the Laplace equation with Neumann boundary conditions on a uniform and randomly disturbed mesh.

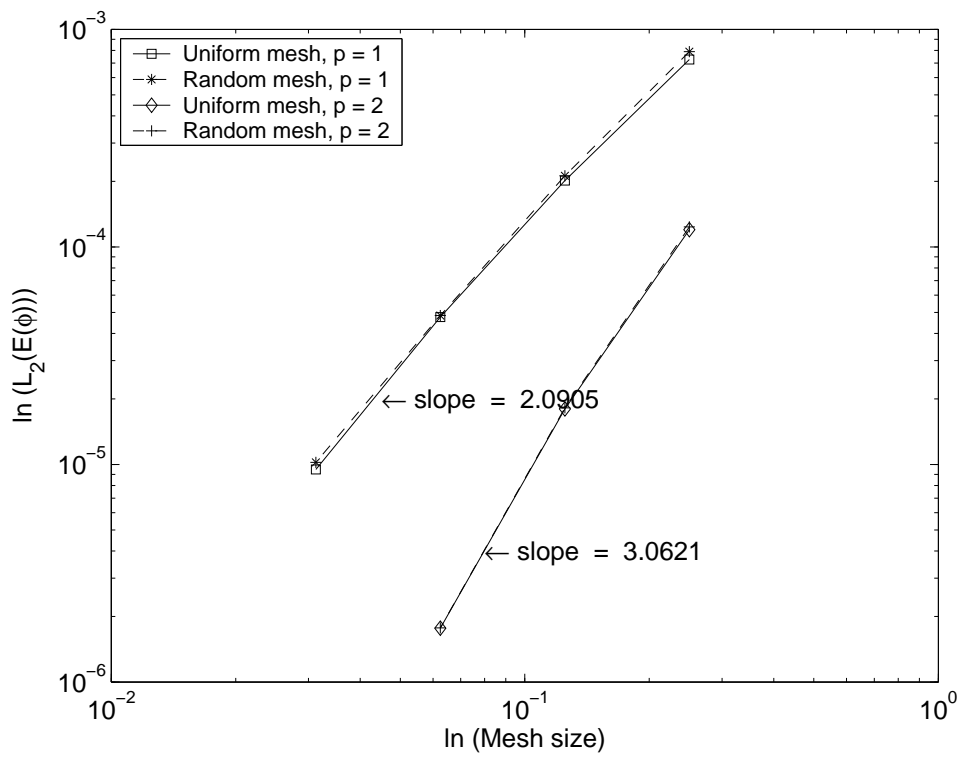


Figure 2: $L^2(\Omega)$ -error in the velocity potential as a function of mesh size for polynomial basis functions with order $p = 1$ and 2.

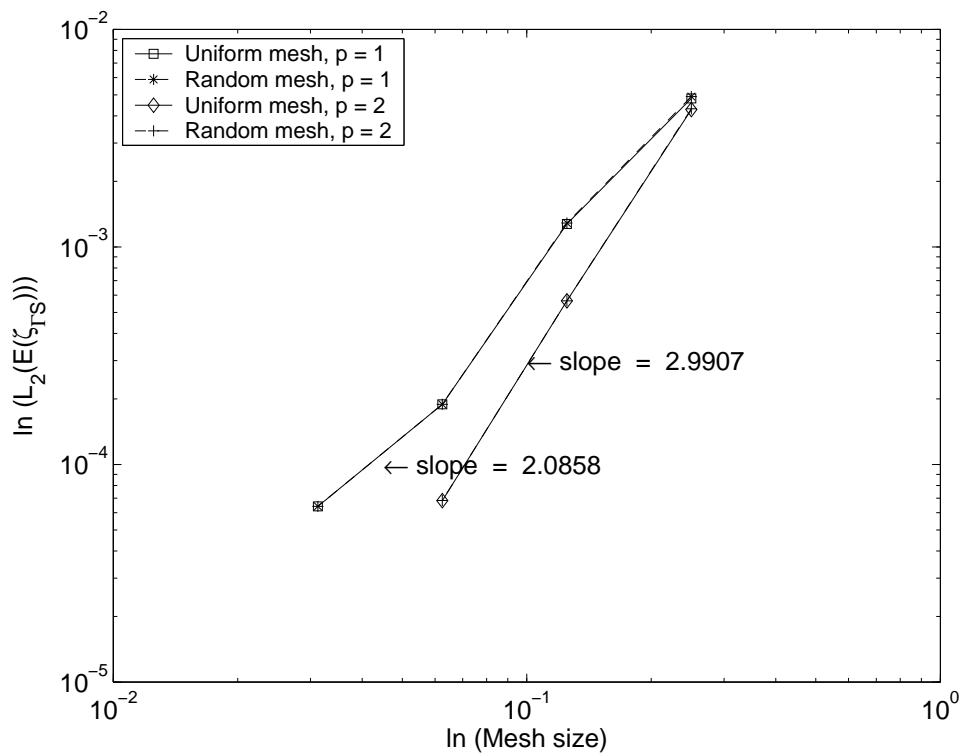


Figure 3: $L^2(\Gamma_S)$ -error in the wave height as a function of mesh size for polynomial basis functions with order $p = 1$ and 2.

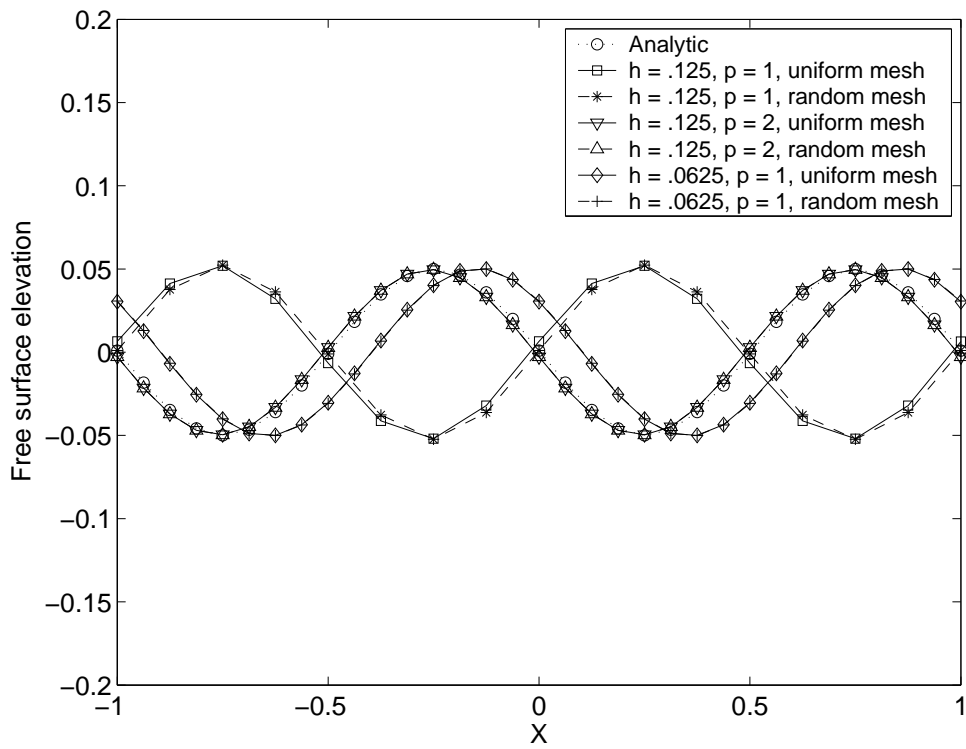


Figure 4: Wave profile on uniform and randomly disturbed mesh after 20 periods.

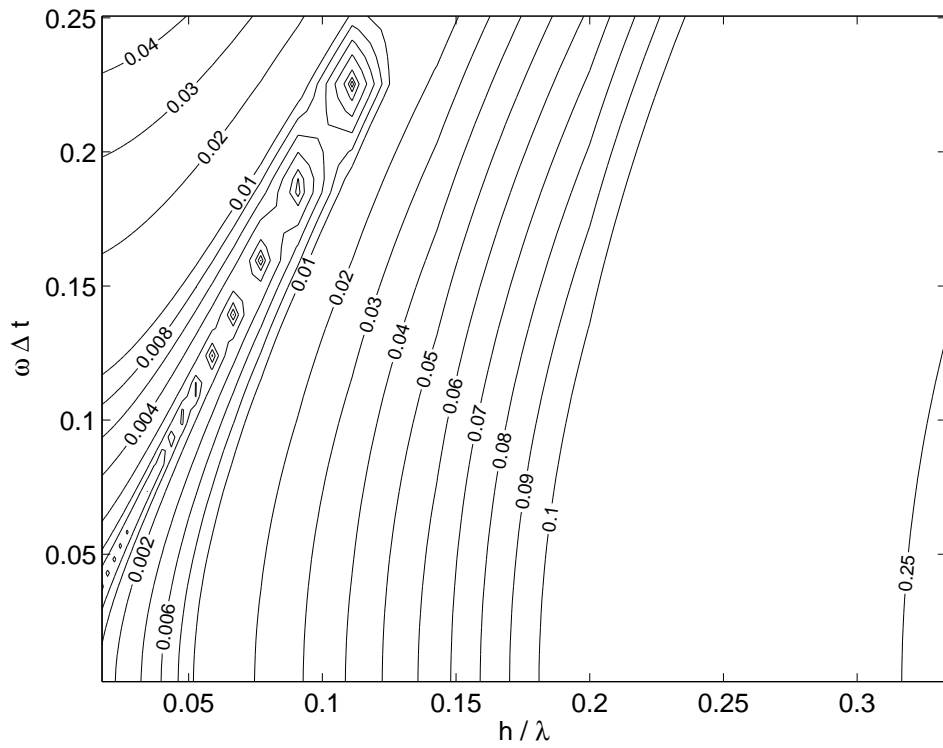


Figure 5: Absolute value of frequency error for time-harmonic waves as a function of mesh size/wave length (h/λ) and time step \times frequency ($\omega\Delta t$) for the DG discretization using linear basis functions.

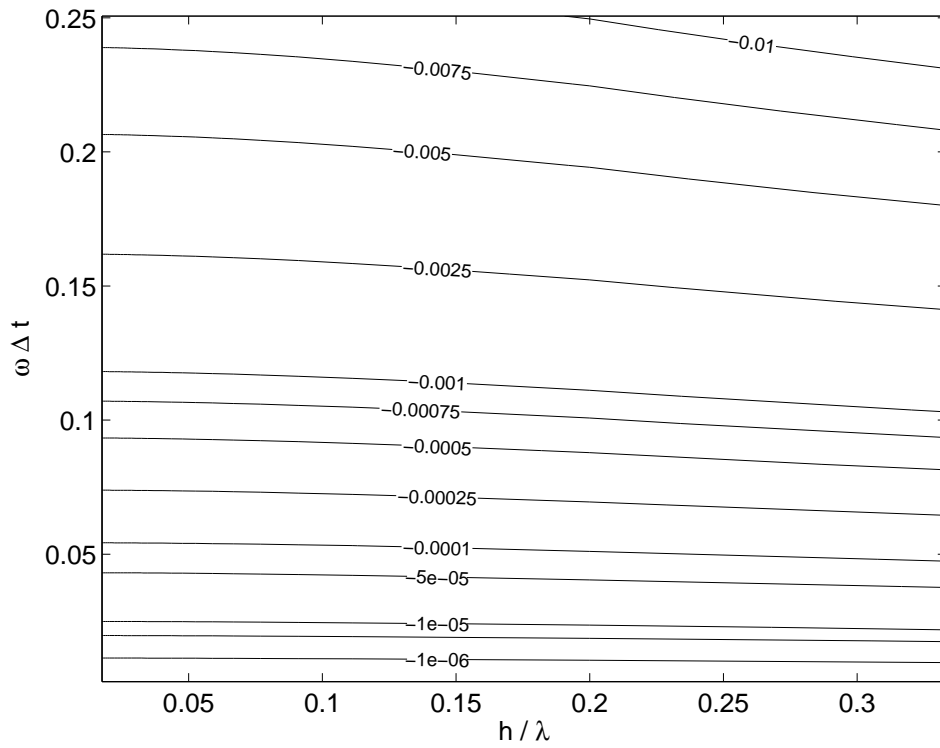


Figure 6: Dissipation for time-harmonic waves as a function of mesh size/wave length (h/λ) and time step \times frequency ($\omega\Delta t$) for the DG discretization using linear basis functions.

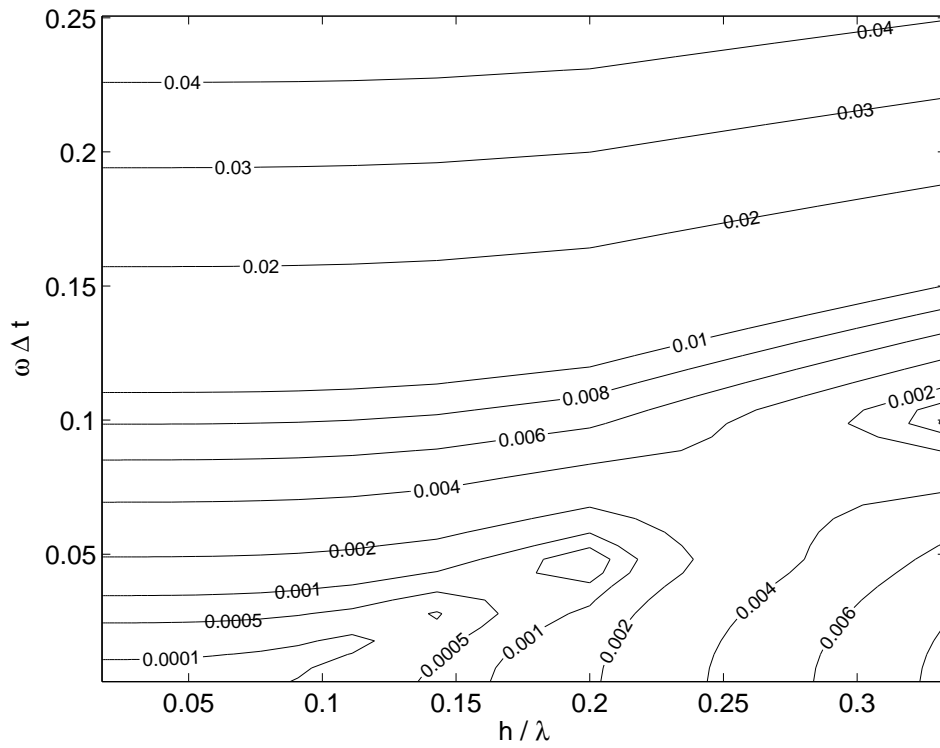


Figure 7: Absolute value of frequency error for time-harmonic waves as a function of mesh size/wave length (h/λ) and time step \times frequency ($\omega\Delta t$) for the DG discretization using quadratic basis functions.

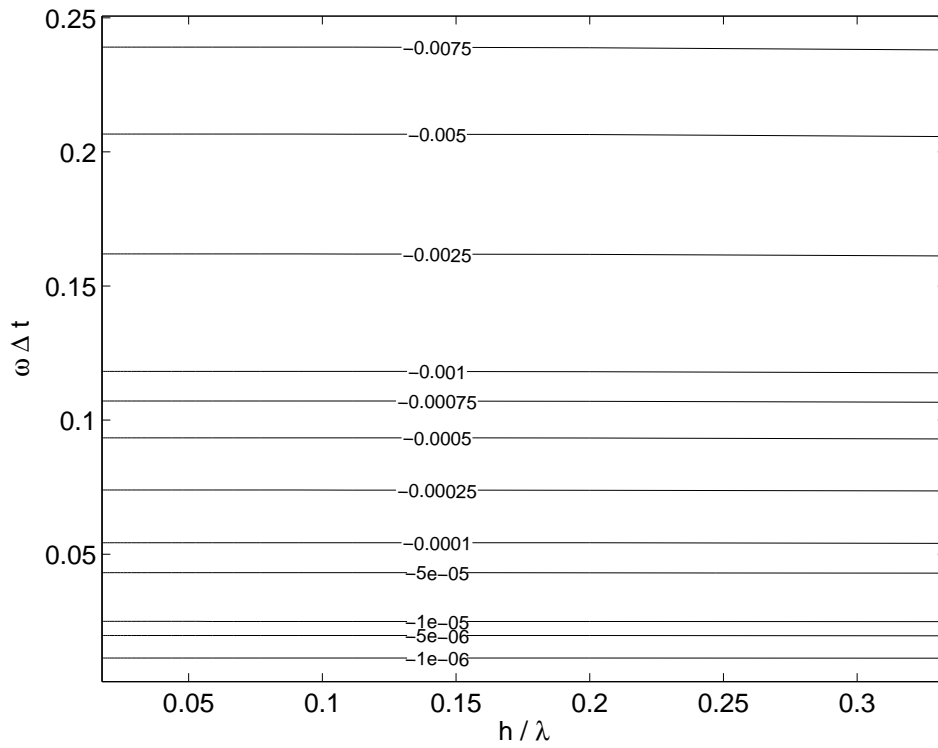


Figure 8: Dissipation for time-harmonic waves as a function of mesh size/wave length (h/λ) and time step \times frequency ($\omega\Delta t$) for the DG discretization using quadratic basis functions.

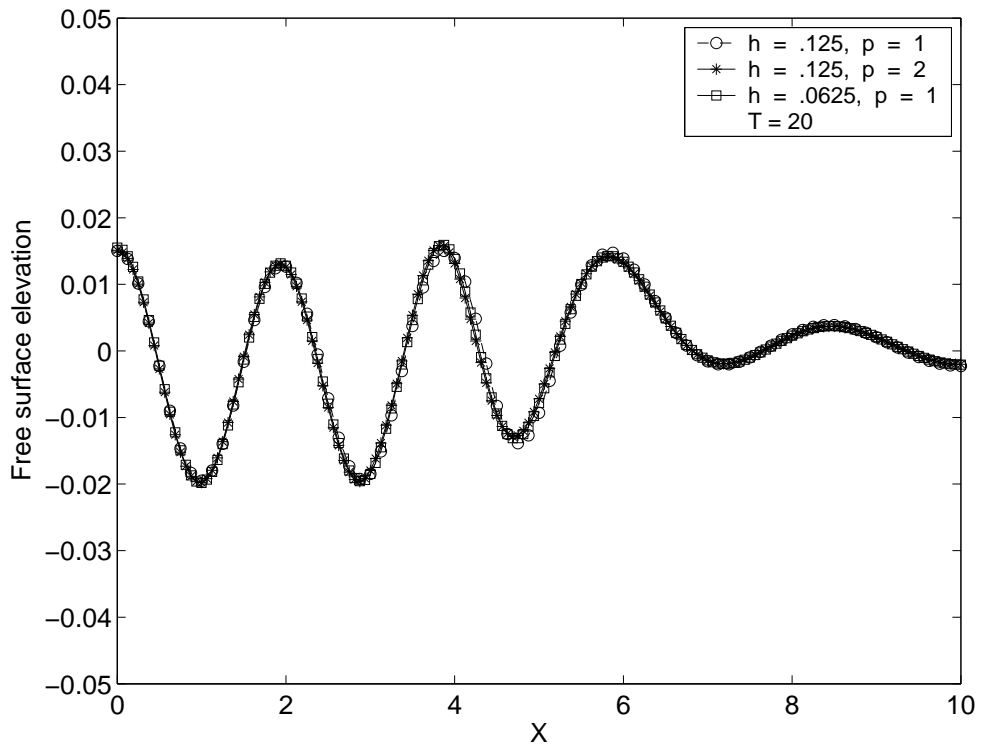


Figure 9: Wave profile at $T = 20$ generated by a wave maker at $x = 0$ (randomly disturbed mesh) for different mesh sizes h and order of the polynomial basis functions.

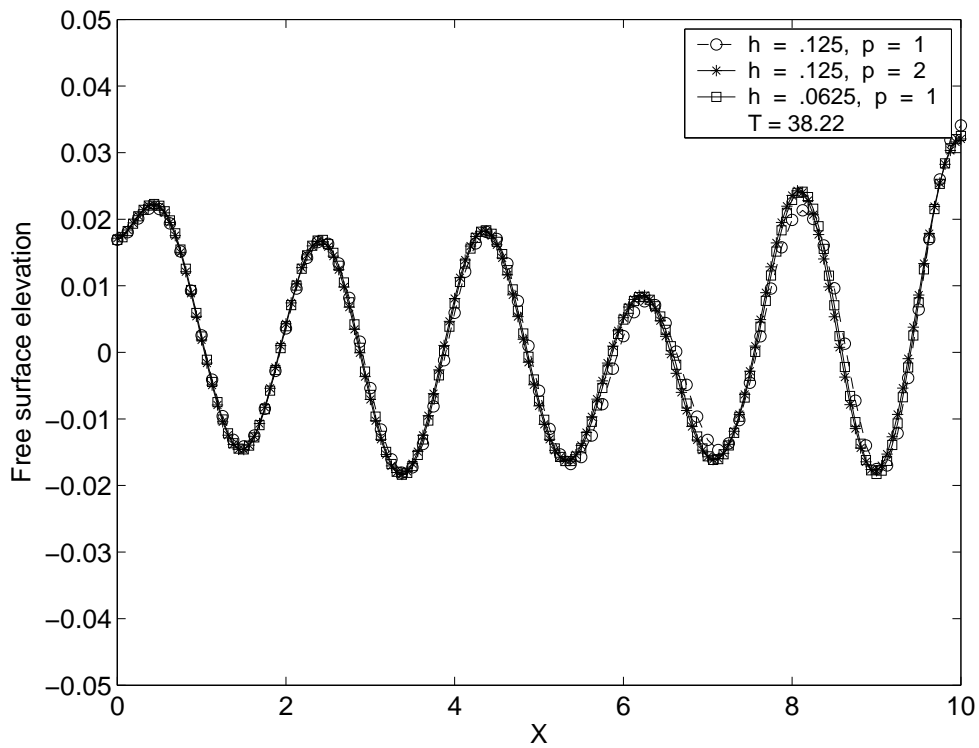


Figure 10: Wave profile at $T = 38.22$ generated by a wave maker at $x = 0$ (randomly disturbed mesh) for different mesh sizes h and order of the polynomial basis functions.

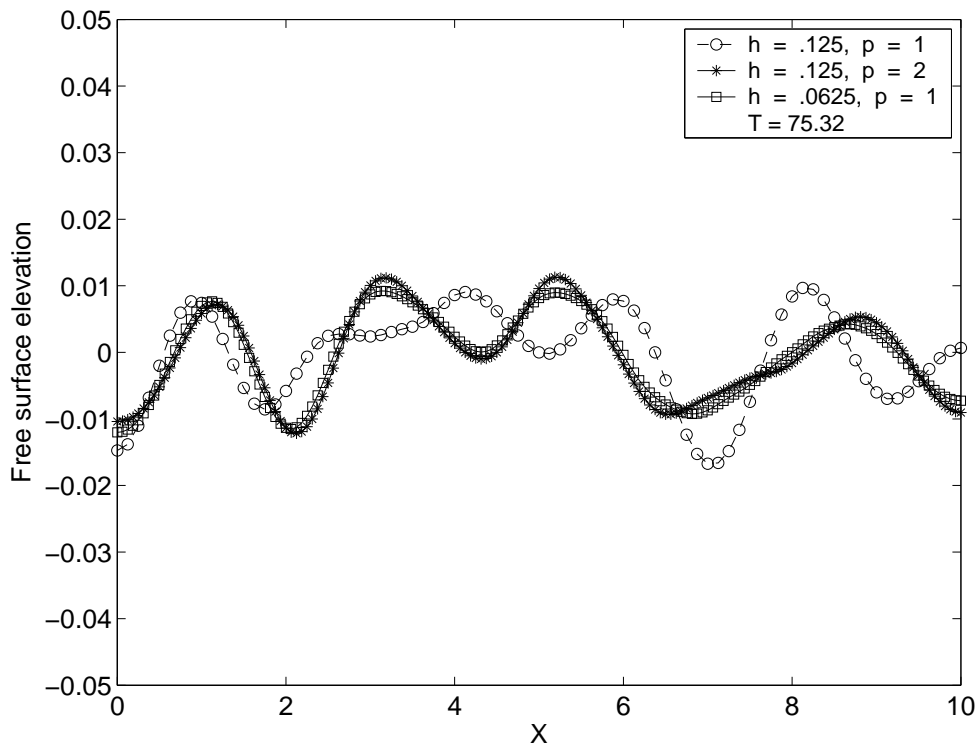


Figure 11: Wave profile at $T = 75.32$ generated by a wave maker at $x = 0$ (randomly disturbed mesh) for different mesh sizes h and order of the polynomial basis functions.