Part VIII, Chapter 38

Discontinuous Galerkin

The goal of this chapter is to study the approximation of an elliptic model problem by the discontinuous Galerkin (dG) method. The distinctive feature of dG methods is that the trial and the test spaces are broken finite element spaces (see §18.1.2). Inspired by the boundary penalty method from Chapter 37, dG formulations are obtained by adding a consistency term at all the mesh interfaces and boundary faces, boundary conditions are weakly enforced à la Nitsche, and continuity across the mesh interfaces is weakly enforced by penalizing the jumps. The dG method we study here is called symmetric interior penalty (SIP) because the consistency term is symmetrized to maintain the symmetry of the discrete bilinear form. Incidentally, the symmetry property is important to derive optimal L^2-error estimates assuming full elliptic regularity pickup. We also discuss a useful reformulation of the dG method by lifting the jumps, leading to the important notion of discrete gradient reconstruction.

38.1 Model problem

For simplicity, we focus on the Poisson equation with homogeneous Dirichlet boundary conditions:

$$\begin{cases} \text{Find } u \in V := H_0^1(D) \text{ such that} \\ a(u,w) = \ell(w), \quad \forall w \in V, \end{cases} \tag{38.1}$$

with $a(v,w) := \int_D \nabla v{\cdot}\nabla w \,\mathrm{d}x$, $\ell(w) := \int_D f w \,\mathrm{d}x$, $f \in L^2(D)$, and D is a Lipschitz polyhedron in $\mathbb{R}^d$. This problem is well-posed owing to the Lax–Milgram lemma and the Poincaré–Steklov inequality in $H_0^1(D)$. We refer the reader to §41.4 for the more general PDE $-\nabla{\cdot}(\lambda\nabla u) = f$ with contrasted diffusivity λ.

38.2 Symmetric interior penalty

In this section, we derive the dG approximation of the model problem (38.1) using the SIP method and show that the discrete problem is well-posed.

38.2.1 Discrete problem

Although dG methods can be used on general meshes composed of polyhedral cells, we consider for simplicity a shape-regular sequence $(\mathcal{T}_h)_{h\in\mathcal{H}}$ of affine matching meshes so that each mesh covers D exactly. Let $W^{1,1}(\mathcal{T}_h;\mathbb{R}^q)$, $q \geq 1$, be the broken Sobolev space introduced in Definition 18.1. Recall that every interface $F := \partial K_l \cap \partial K_r \in \mathcal{F}_h^\circ$ is oriented by the fixed unit normal vector $\boldsymbol{n}_F$ pointing from K_l to K_r, i.e., $\boldsymbol{n}_F := \boldsymbol{n}_{K_l} = -\boldsymbol{n}_{K_r}$, and that the jump across F of a function $v \in W^{1,1}(\mathcal{T}_h;\mathbb{R}^q)$ is defined by setting $[\![v]\!]_F := v_{|K_l} - v_{|K_r}$ a.e. on F. We also need the following notion of face average.

Definition 38.1 (Average). *For all $F := \partial K_l \cap \partial K_r \in \mathcal{F}_h^\circ$, the* average *of a function $v \in W^{1,1}(\mathcal{T}_h;\mathbb{R}^q)$ on F is defined as*

$$\{v\}_F := \frac{1}{2}\big(v_{|K_l} + v_{|K_r}\big) \qquad \text{a.e. on } F. \tag{38.2}$$

As for jumps, the subscript F is dropped when the context is unambiguous.

To be more concise, it is customary in the dG literature dedicated to elliptic PDEs to define the jump and the average of a function at the boundary faces by setting $[\![v]\!]_F := \{v\}_F := v_{|K_l}$ a.e. on $F := \partial K_l \cap \partial D \in \mathcal{F}_h^\partial$ (i.e., K_l is the unique mesh cell having the boundary face F among its faces).

Let $(\widehat{K}, \widehat{P}, \widehat{\Sigma})$ be the reference finite element which we assume to be of degree $k \geq 1$. Let us consider the broken finite element space (see (18.4)) s.t.

$$V_h := P_k^{\mathrm{b}}(\mathcal{T}_h) := \{v_h \in L^\infty(D) \mid \psi_K(v_{h|K}) \in \widehat{P},\, \forall K \in \mathcal{T}_h\}, \tag{38.3}$$

where $\psi_K(v) := v \circ \boldsymbol{T}_K$ is the pullback by the geometric mapping $\boldsymbol{T}_K$. The approximation setting in dG methods is nonconforming since functions in V_h can jump across the mesh interfaces and can have nonzero boundary values, whereas membership in $V := H_0^1(D)$ requires continuity across the interfaces (see Theorem 18.8) and zero boundary values. Nonconformity implies that we cannot work with the bilinear form a. The construction of the discrete bilinear form a_h on $V_h \times V_h$ is a bit more involved than for the Crouzeix–Raviart finite element method from Chapter 36, where it was sufficient to replace the weak gradient ∇ by the broken gradient ∇_h (see Definition 36.3) to build a_h from a. Instead, the SIP method hinges on the following discrete bilinear form:

$$a_h(v_h, w_h) := \int_D \nabla_h v_h \cdot \nabla_h w_h \,\mathrm{d}x - \sum_{F\in\mathcal{F}_h} \int_F \{\nabla_h v_h\}\cdot \boldsymbol{n}_F [\![w_h]\!] \,\mathrm{d}s \\ - \sum_{F\in\mathcal{F}_h} \int_F [\![v_h]\!]\{\nabla_h w_h\}\cdot \boldsymbol{n}_F \,\mathrm{d}s + \sum_{F\in\mathcal{F}_h} \varpi(h_F) \int_F [\![v_h]\!][\![w_h]\!] \,\mathrm{d}s, \quad (38.4)$$

where the second and the fourth terms on the right-hand side are reminiscent of Nitsche's boundary penalty method. The second term is called *consistency term* since it is important to establish consistency/boundedness (see Lemma 38.9). The third term, which is called *adjoint consistency term*, makes the discrete bilinear form a_h symmetric and it is important to establish an improved L^2-error estimate (see Theorem 38.12). The fourth term is important to establish coercivity (see Lemma 38.6). It penalizes jumps across interfaces and values at boundary faces and is, therefore, called *penalty term*. Coercivity requires that the penalty parameter be s.t. $\varpi(h_F) := \varpi_0 h_F^{-1}$, where $\varpi_0 > 0$ has to be chosen large enough, and on shape-regular mesh sequences, the local length scale h_F can be taken to be the diameter of F.

We consider the following discrete problem:

$$\begin{cases} \text{Find } u_h \in V_h \text{ such that} \\ a_h(u_h, w_h) = \ell_h(w_h), \quad \forall w_h \in V_h, \end{cases} \quad (38.5)$$

where the discrete linear form is given by

$$\ell_h(w_h) := \int_D f w_h \,\mathrm{d}x, \qquad \forall w_h \in V_h. \quad (38.6)$$

This choice for ℓ_h is possible since the source term in the model problem (38.1) is assumed to be in $L^2(D)$. A more general setting, e.g., $f \in H^{-1}(D)$, is discussed in Remark 36.5. Furthermore, it is legitimate to extend a_h to $(H^{1+r}(D) + V_h)\times V_h$, $r > \frac{1}{2}$, since $\nabla u \in \boldsymbol{H}^r(D)$ implies that $(\nabla u)_{|F}$ is well defined as an integrable function for all $F \in \mathcal{F}_h$. To motivate the appearance of the consistency term in the definition of a_h, let us prove the following important result.

Lemma 38.2 (Consistency term). *Assume that $u \in H^{1+r}(D)$, $r > \frac{1}{2}$. Then we have $a_h(u, w_h) = \ell_h(w_h)$ for all $w_h \in V_h$.*

Proof. We have $[\![u]\!]_F = 0$ a.e. on all $F \in \mathcal{F}_h$ (use Theorem 18.8 for $F \in \mathcal{F}_h^\circ$ and $\gamma^{\mathrm{g}}(u) = 0$ for $F \in \mathcal{F}_h^\partial$) and $\nabla_h u = \nabla u$ (see Lemma 18.9). Since $\nabla u \in \boldsymbol{H}^r(D)$, $r > \frac{1}{2}$, we also have $[\![\nabla u]\!]\cdot\boldsymbol{n}_F = 0$ a.e. on all $F \in \mathcal{F}_h^\circ$ (see Remark 18.4). We infer that

$$a_h(u, w_h) = \int_D \nabla u\cdot\nabla_h w_h \,\mathrm{d}x - \sum_{F\in\mathcal{F}_h} \int_F (\nabla u\cdot\boldsymbol{n}_F)[\![w_h]\!] \,\mathrm{d}s.$$

We conclude by performing elementwise integration by parts as follows:

$$\int_D \nabla u \cdot \nabla_h w_h \,\mathrm{d}x = \int_D -(\Delta u) w_h \,\mathrm{d}x + \sum_{K \in \mathcal{T}_h} \sum_{F \in \mathcal{F}_K} \int_F (\nabla u \cdot \boldsymbol{n}_K) w_{h|K} \,\mathrm{d}s$$
$$= \ell_h(w_h) + \sum_{F \in \mathcal{F}_h} \int_F (\nabla u \cdot \boldsymbol{n}_F) [\![w_h]\!] \,\mathrm{d}s. \qquad \square$$

Remark 38.3 (Literature). The SIP approximation has been analyzed in Arnold [15] (see also Baker [44], Wheeler [394]). □

Remark 38.4 (Nonmatching meshes). It is possible to consider nonmatching meshes if the diameter of each interface $F \in \mathcal{F}_h^\circ$ is uniformly equivalent to the diameter of the two cells sharing F. □

38.2.2 Coercivity and well-posedness

We equip the space V_h with the following norm:

$$\|v_h\|_{V_h}^2 := \|\nabla_h v_h\|_{\boldsymbol{L}^2(D)}^2 + |v_h|_{\mathrm{J}}^2, \qquad |v_h|_{\mathrm{J}}^2 := \sum_{F \in \mathcal{F}_h} \frac{1}{h_F} \|[\![v_h]\!]\|_{L^2(F)}^2. \tag{38.7}$$

That $\|\cdot\|_{V_h}$ is a norm on V_h (and not just a seminorm) can be verified directly: If $\|v_h\|_{V_h} = 0$, then v_h is piecewise constant and $[\![v_h]\!]_F = 0$ for all $F \in \mathcal{F}_h$. This means that v_h is constant on D and vanishes at ∂D, so that $v_h = 0$. Our first step in the analysis is to bound from above the consistency term. Recall that $\mathcal{T}_F := \{K \in \mathcal{T}_h \mid F \in \mathcal{F}_K\}$ is the collection of the mesh cells having F as face. Let $|\mathcal{T}_F|$ denote the cardinality of the set $\mathcal{T}_F$ ($|\mathcal{T}_F| = 2$ for all $F \in \mathcal{F}_h^\circ$ and $|\mathcal{T}_F| = 1$ for all $F \in \mathcal{F}_h^\partial$).

Lemma 38.5 (Consistency term). *Let us set for all $(v_h, w_h) \in V_h \times V_h$,*

$$n_h(v_h, w_h) := -\sum_{F \in \mathcal{F}_h} \int_F \{\nabla_h v_h\} \cdot \boldsymbol{n}_F [\![w_h]\!] \,\mathrm{d}s. \tag{38.8}$$

Then the following holds true for all $v_h \in V_h$:

$$\sup_{w_h \in V_h} \frac{|n_h(v_h, w_h)|}{|w_h|_{\mathrm{J}}} \le \left(\sum_{F \in \mathcal{F}_h} \frac{1}{|\mathcal{T}_F|} \sum_{K \in \mathcal{T}_F} h_F \|\boldsymbol{n}_F \cdot \nabla(v_{h|K})\|_{L^2(F)}^2 \right)^{\frac{1}{2}}. \tag{38.9}$$

Proof. The Cauchy–Schwarz inequality leads to

$$|n_h(v_h, w_h)| \le \sum_{F \in \mathcal{F}_h} h_F^{\frac{1}{2}} \|\boldsymbol{n}_F \cdot \{\nabla_h v_h\}\|_{L^2(F)} \times h_F^{-\frac{1}{2}} \|[\![w_h]\!]\|_{L^2(F)}$$
$$\le \left(\sum_{F \in \mathcal{F}_h} h_F \|\boldsymbol{n}_F \cdot \{\nabla_h v_h\}\|_{L^2(F)}^2 \right)^{\frac{1}{2}} |w_h|_{\mathrm{J}},$$

Letting $\boldsymbol{g}_h := \nabla_h v_h$, (38.9) follows from $\{\boldsymbol{g}_h\}_F = \frac{1}{|\mathcal{T}_F|}\sum_{K\in\mathcal{T}_F} \boldsymbol{g}_{h|K}$ and

$$\|\boldsymbol{n}_F{\cdot}\{\boldsymbol{g}_h\}\|^2_{L^2(F)} = \frac{1}{|\mathcal{T}_F|^2}\left\|\sum_{K\in\mathcal{T}_F} \boldsymbol{n}_F{\cdot}\boldsymbol{g}_{h|K}\right\|^2_{L^2(F)} \le \frac{1}{|\mathcal{T}_F|}\sum_{K\in\mathcal{T}_F}\|\boldsymbol{n}_F{\cdot}\boldsymbol{g}_{h|K}\|^2_{L^2(F)}.$$

□

We shall use the same discrete trace inequality as in Chapter 37 to prove a coercivity property. Let c_{dt} be the smallest constant such that

$$\|\boldsymbol{n}_F{\cdot}\nabla_h w_{h|K}\|_{L^2(F)} \le c_{\mathrm{dt}} h_F^{-\frac{1}{2}}\|\nabla_h w_h\|_{\boldsymbol{L}^2(K)}, \tag{38.10}$$

for all $w_h \in V_h$, all $K \in \mathcal{T}_h$, all $F \in \mathcal{F}_K$, and all $h \in \mathcal{H}$. Let $n_\partial := \max_{K\in\mathcal{T}_h}|\mathcal{F}_K|$ be the largest number of faces per mesh cell, i.e., $n_\partial \le d+1$ for simplicial meshes (the definition of n_∂ differs from that of Chapter 37).

Lemma 38.6 (Coercivity, well-posedness). *Let the penalty parameter be s.t.* $\varpi(h_F) := \varpi_0 h_F^{-1}$ *with* $\varpi_0 > n_\partial c_{\mathrm{dt}}^2$. (i) *We have*

$$a_h(v_h, v_h) \ge \alpha\|v_h\|^2_{V_h}, \qquad \forall v_h \in V_h, \tag{38.11}$$

with $\alpha := \frac{\varpi_0 - n_\partial c_{\mathrm{dt}}^2}{1+\varpi_0} > 0$. (ii) *The discrete problem* (38.5) *is well-posed.*

Proof. Let $v_h \in V_h$. Our starting observation is that

$$a_h(v_h, v_h) = \|\nabla_h v_h\|^2_{\boldsymbol{L}^2(D)} + 2n_h(v_h, v_h) + \varpi_0|v_h|^2_{\mathrm{J}}.$$

Using (38.9) and (38.10), we infer that

$$|n_h(v_h, v_h)| \le \left(\sup_{w_h\in V_h}\frac{|n_h(v_h, w_h)|}{|w_h|_{\mathrm{J}}}\right)|v_h|_{\mathrm{J}} \le n_\partial^{\frac{1}{2}} c_{\mathrm{dt}}\|\nabla_h v_h\|_{\boldsymbol{L}^2(D)}|v_h|_{\mathrm{J}},$$

since $|\mathcal{T}_F| \ge 1$, $\sum_{F\in\mathcal{F}_h}\sum_{K\in\mathcal{T}_F}(\cdot) = \sum_{K\in\mathcal{T}_h}\sum_{F\in\mathcal{F}_K}(\cdot)$, and $|\mathcal{F}_K| \le n_\partial$, so that $\sum_{F\in\mathcal{F}_h}\sum_{K\in\mathcal{T}_F}\|\boldsymbol{g}_{h|K}\|^2_{\boldsymbol{L}^2(K)} \le n_\partial\sum_{K\in\mathcal{T}_h}\|\boldsymbol{g}_{h|K}\|^2_{\boldsymbol{L}^2(K)} = n_\partial\|\boldsymbol{g}_h\|^2_{\boldsymbol{L}^2(D)}$ with $\boldsymbol{g}_h := \nabla_h v_h$. This leads to the lower bound

$$a_h(v_h, v_h) \ge \|\nabla_h v_h\|^2_{\boldsymbol{L}^2(D)} - 2n_\partial^{\frac{1}{2}} c_{\mathrm{dt}}\|\nabla_h v_h\|_{\boldsymbol{L}^2(D)}|v_h|_{\mathrm{J}} + \varpi_0|v_h|^2_{\mathrm{J}},$$

whence we infer the coercivity property (38.11) by using the quadratic inequality from Exercise 37.2. Finally, the well-posedness of (38.5) follows from the Lax–Milgram lemma. □

Remark 38.7 (Penalty parameter). As in the boundary penalty method from Chapter 37, one needs a (reasonable) upper bound on the constant c_{dt} to choose a value of ϖ_0 that guarantees coercivity. The results of §12.2 show that c_{dt} scales essentially as k^2. An alternative penalty strategy allowing for an easy-to-compute value of ϖ_0 is discussed in Remark 38.17, but this technique requires local inversions of small mass matrices. □

Remark 38.8 (Discrete Sobolev inequality). Let ℓ_D be a length scale associated with D, e.g., $\ell_D := \mathrm{diam}(D)$. One can show that there is $C_{\mathrm{SOB}} > 0$ such that $C_{\mathrm{SOB}}\|v_h\|_{L^q(D)} \le \ell_D\|v_h\|_{V_h}$ for all $v_h \in V_h$, all $h \in \mathcal{H}$, and all $q \in [1,\infty)$ if $d = 2$ and $q \in [1, \frac{2d}{d-2}]$ if $d \ge 3$; see Buffa and Ortner [95], Di Pietro and Ern [164]. The reader is referred to Arnold [15], Brenner [86] for similar estimates in broken Hilbert Sobolev spaces ($q = 2$). □

38.2.3 Variations on boundary conditions

The non-homogeneous Dirichlet boundary condition $u = g$ on ∂D with $g \in H^{\frac{1}{2}}(\partial D)$ is discretized by modifying the right-hand side in (38.5) as follows:

$$\ell_h^{\mathrm{nD}}(w_h) := \ell(w_h) - \sum_{F\in\mathcal{F}_h^\partial} \int_F g(\boldsymbol{n}_F{\cdot}\nabla_h w_h - \varpi(h_F)w_h)\,\mathrm{d}s. \tag{38.12}$$

For the Robin boundary condition $\gamma u + \boldsymbol{n}{\cdot}\nabla u = g$ on ∂D with $g \in L^2(\partial D)$ and $\gamma \in L^\infty(\partial D)$ taking nonnegative values on ∂D ($\gamma := 0$ corresponds to the Neumann problem), the discrete bilinear form and the right-hand side become

$$a_h^{\mathrm{Rb}}(v_h, w_h) := \int_D \nabla_h v_h{\cdot}\nabla_h w_h\,\mathrm{d}x - \sum_{F\in\mathcal{F}_h^\circ}\int_F \{\nabla_h v_h\}{\cdot}\boldsymbol{n}_F[\![w_h]\!]\,\mathrm{d}s \tag{38.13a}$$

$$- \sum_{F\in\mathcal{F}_h^\circ}\int_F [\![v_h]\!]\{\nabla_h w_h\}{\cdot}\boldsymbol{n}_F\,\mathrm{d}s + \sum_{F\in\mathcal{F}_h^\circ}\varpi(h_F)\int_F [\![v_h]\!][\![w_h]\!]\,\mathrm{d}s + \sum_{F\in\mathcal{F}_h^\partial}\int_F \gamma v_h w_h\,\mathrm{d}s,$$

$$\ell_h^{\mathrm{Rb}}(w_h) := \ell(w_h) + \sum_{F\in\mathcal{F}_h^\partial}\int_F g w_h\,\mathrm{d}s. \tag{38.13b}$$

One can verify that Lemma 38.2 still holds true in both cases.

38.3 Error analysis

In this section, we derive an energy error estimate, that is, we bound the error by using the coercivity norm and the abstract error estimate from Lemma 27.5. We also derive an improved L^2-error estimate by means of a duality argument. We assume that $u \in V_{\mathrm{S}}$ with

$$V_{\mathrm{S}} := H^{1+r}(D) \cap H_0^1(D), \qquad r > \frac{1}{2}. \tag{38.14}$$

The assumption $u \in V_{\mathrm{S}}$ is reasonable in the setting of the Poisson equation with Dirichlet conditions in a Lipschitz polyhedron since it is consistent with the elliptic regularity theory (see Theorem 31.33). The important property that we use is that for any function $v \in V_{\mathrm{S}}$ the normal derivative $\boldsymbol{n}_K{\cdot}\nabla v$ is meaningful in $L^2(\partial K)$ for all $K \in \mathcal{T}_h$. Recall that the discrete space is $V_h := P_k^{\mathrm{b}}(\mathcal{T}_h)$ equipped with the $\|{\cdot}\|_{V_h}$-norm defined in (38.7). We set $V_\sharp := V_{\mathrm{S}} + V_h$ and we equip this space with the norm

$$\|v\|_{V_\sharp}^2 := \|\nabla_h v\|_{\boldsymbol{L}^2(D)}^2 + |v|_{\mathrm{J}}^2 + \sum_{K\in\mathcal{T}_h} h_K \|\boldsymbol{n}_K{\cdot}\nabla v_{|K}\|_{L^2(\partial K)}^2, \tag{38.15}$$

with $|v|_{\mathrm{J}}^2 := \sum_{F\in\mathcal{F}_h} \frac{1}{h_F}\|[\![v]\!]\|_{L^2(F)}^2$. A discrete trace inequality shows that there is $c_\sharp$ s.t. $\|v_h\|_{V_\sharp} \le c_\sharp \|v_h\|_{V_h}$ for all $v_h \in V_h$ and all $h \in \mathcal{H}$, i.e., (27.5) holds true. Using the discrete bilinear forms a_h and ℓ_h defined in (38.4) and (38.6), respectively, the consistency error is s.t. $\langle \delta_h(v_h), w_h\rangle_{V_h',V_h} := \ell_h(w_h) - a_h(v_h, w_h)$ for all $v_h, w_h \in V_h$.

Lemma 38.9 (Consistency/boundedness). *Assume* (38.14). *There is $\omega_\sharp$, uniform w.r.t. $u \in V_{\mathrm{S}}$, s.t. for all $v_h \in V_h$ and all $h \in \mathcal{H}$,*

$$\|\delta_h(v_h)\|_{V_h'} \le \omega_\sharp \, \|u - v_h\|_{V_\sharp}. \tag{38.16}$$

Proof. Let $v_h \in V_h$ and let us set $\eta := u - v_h$. Owing to Lemma 38.2 and since $[\![u]\!]_F = 0$ for all $F \in \mathcal{F}_h$, we infer that for all $w_h \in W_h$,

$$\begin{aligned}\langle \delta_h(v_h), w_h\rangle_{V_h',V_h} = {} & \int_D \nabla_h \eta{\cdot}\nabla_h w_h \,\mathrm{d}x + n_\sharp(\eta, w_h) \\ & - \sum_{F\in\mathcal{F}_h} \int_F [\![\eta]\!]\{\nabla_h w_h\}{\cdot}\boldsymbol{n}_F \,\mathrm{d}s + \sum_{F\in\mathcal{F}_h} \frac{\varpi_0}{h_F}\int_F [\![\eta]\!][\![w_h]\!]\,\mathrm{d}s,\end{aligned}$$

where $n_\sharp(v, w_h) := -\sum_{F\in\mathcal{F}_h}\int_F \{\nabla_h v\}{\cdot}\boldsymbol{n}_F [\![w_h]\!]\,\mathrm{d}s$ is understood as an extension to $V_\sharp \times V_h$ of the discrete bilinear form n_h originally defined on $V_h \times V_h$ by (38.8). (Note that the assumption $r > \frac{1}{2}$ in the definition of V_{S} is crucial for this extension to make sense.) The Cauchy–Schwarz inequality implies that

$$\begin{aligned}&\left|\int_D \nabla_h \eta{\cdot}\nabla_h w_h \,\mathrm{d}x + \sum_{F\in\mathcal{F}_h}\frac{\varpi_0}{h_F}\int_F [\![\eta]\!][\![w_h]\!]\,\mathrm{d}s\right| \\ &\quad \le \|\nabla_h\eta\|_{\boldsymbol{L}^2(D)}\|\nabla_h w_h\|_{\boldsymbol{L}^2(D)} + \varpi_0|\eta|_{\mathrm{J}}|w_h|_{\mathrm{J}} \le \max(1,\varpi_0)\|\eta\|_{V_\sharp}\|w_h\|_{V_h}.\end{aligned}$$

Since the bound (38.9) is still valid for $n_\sharp(\eta, w_h)$, we also have

$$\begin{aligned}|n_\sharp(\eta, w_h)| &\le \left(\sum_{F\in\mathcal{F}_h}\frac{1}{|\mathcal{T}_F|}\sum_{K\in\mathcal{T}_F} h_F\|\boldsymbol{n}_F{\cdot}\nabla(\eta_{|K})\|_{L^2(F)}^2\right)^{\frac{1}{2}} |w_h|_{\mathrm{J}} \\ &\le c\,\|\eta\|_{V_\sharp}|w_h|_{\mathrm{J}} \le c\,\|\eta\|_{V_\sharp}\|w_h\|_{V_h}.\end{aligned}$$

(This is where we use the contribution of the normal derivative to the $\|{\cdot}\|_{V_\sharp}$-norm.) Proceeding as in the proof of Lemma 38.5, we finally infer that

$$\left|\sum_{F\in\mathcal{F}_h}\int_F [\![\eta]\!]\{\nabla_h w_h\}\cdot\boldsymbol{n}_F\,\mathrm{d}s\right| \le |\eta|_{\mathrm{J}}\left(\sum_{F\in\mathcal{F}_h}\frac{1}{|\mathcal{T}_F|}\sum_{K\in\mathcal{T}_F} h_F\|\nabla(w_{h|K})\|^2_{\boldsymbol{L}^2(F)}\right)^{\frac{1}{2}}$$
$$\le n_\partial^{\frac{1}{2}} c_{\mathrm{dt}}|\eta|_{\mathrm{J}}\|\nabla_h w_h\|_{\boldsymbol{L}^2(D)} \le n_\partial^{\frac{1}{2}} c_{\mathrm{dt}}\|\eta\|_{V_\sharp}\|w_h\|_{V_h},$$

where we used the discrete trace inequality (38.10) as in the proof of Lemma 38.6. Collecting the above bounds shows that $|\langle\delta_h(v_h), w_h\rangle_{V_h',V_h}| \le c\|\eta\|_{V_\sharp}\|w_h\|_{V_h}$, i.e., (38.16) holds true. □

Theorem 38.10 (Convergence). *Let* u *solve* (38.1) *and let* u_h *solve* (38.5) *with the penalty parameter* $\varpi_0 > c_{\mathrm{dt}}^2 n_\partial$. *Assume* (38.14). (i) *There is* c *s.t. the following holds true for all* $h\in\mathcal{H}$*:*

$$\|u-u_h\|_{V_\sharp} \le c \inf_{v_h\in V_h}\|u-v_h\|_{V_\sharp}. \tag{38.17}$$

(ii) *Letting* $t := \min(k,r)$, *we have*

$$\|u-u_h\|_{V_\sharp} \le c\left(\sum_{K\in\mathcal{T}_h} h_K^{2t}|u|^2_{H^{1+t}(K)}\right)^{\frac{1}{2}}. \tag{38.18}$$

Proof. (i) The estimate (38.17) follows from Lemma 27.5 combined with stability (Lemma 38.6) and consistency/boundedness (Lemma 38.9).
(ii) We bound the infimum in (38.17) by taking $v_h := \mathcal{I}_h^\sharp(u)$, where $\mathcal{I}_h^\sharp : L^1(D)\to P_k^{\mathrm{b}}(\mathcal{T}_h)$ is the L^1-stable interpolation operator from §18.3. We need to bound $\|\nabla(\eta_{|K})\|_{\boldsymbol{L}^2(K)} + h_K^{\frac{1}{2}}\|\nabla(\eta_{|K})\|_{L^2(\partial K)}$ for all $K\in\mathcal{T}_h$ and $h_F^{-\frac{1}{2}}\|[\![\eta]\!]_F\|_{L^2(F)}$ for all $F\in\mathcal{F}_h$, with $\eta := u-\mathcal{I}_h^\sharp(u)$. Theorem 18.14 implies that $\|\nabla(\eta_{|K})\|_{\boldsymbol{L}^2(K)} \le c h_K^t|u|_{H^{1+t}(K)}$. Moreover, Corollary 18.15 implies that $h_K^{\frac{1}{2}}\|\nabla(\eta_{|K})\|_{L^2(\partial K)} \le c h_K^t|u|_{H^{1+t}(K)}$ and that $\|\eta_{|K}\|_{L^2(F)} \le c h_K^{t+\frac{1}{2}}|u|_{H^{1+t}(K)}$ for any face $F\in\mathcal{F}_K$. Since $[\![\eta]\!]_F := \eta_{|K_l}$ for all $F := \partial K_l\cap\partial D\in\mathcal{F}_h^\partial$ and $[\![\eta]\!]_F := \eta_{|K_l}-\eta_{|K_r}$ for all $F := \partial K_l\cap\partial K_r\in\mathcal{F}_h^\circ$, we can use the shape-regularity of the mesh sequence and the triangle inequality for the jump to infer that $h_F^{-\frac{1}{2}}\|[\![\eta]\!]_F\|_{L^2(F)} \le c\sum_{K\in\mathcal{T}_F} h_K^t|u|_{H^{1+t}(K)}$ for all $F\in\mathcal{F}_h$. This leads to (38.18). □

Remark 38.11 (L^2-orthogonal projection). Note that, as shown in Remark 18.18, $\mathcal{I}_h^\sharp$ is the L^2-orthogonal projection onto $P_k^{\mathrm{b}}(\mathcal{T}_h)$ since ψ_K is the pullback by the geometric mapping $\boldsymbol{T}_K$. □

We now derive an L^2-error estimate by invoking a *duality argument* as in §36.3.3. For all $g\in L^2(D)$, we consider the adjoint solution $\zeta_g\in V := H_0^1(D)$ s.t. $a(v,\zeta_g) = (v,g)_{L^2(D)}$ for all $v\in V$, i.e., $-\Delta\zeta_g = g$ in D and $\gamma^{\mathrm{g}}(\zeta_g) = 0$. Owing to the elliptic regularity theory (see §31.4), there is $s\in(0,1]$ and a constant c_{smo} such that $\|\zeta_g\|_{H^{1+s}(D)} \le c_{\mathrm{smo}}\ell_D^2\|g\|_{L^2(D)}$ for all $g\in L^2(D)$. In the present setting of the Poisson equation with Dirichlet conditions in a Lipschitz polyhedron, it is reasonable to assume that $s \in (\frac{1}{2}, 1]$.

Theorem 38.12 (L^2-estimate). *Under the assumptions of Theorem* 38.10 *and assuming that the elliptic regularity index satisfies $s \in (\frac{1}{2}, 1]$, there is c such that for all $h \in \mathcal{H}$,*

$$\|u - u_h\|_{L^2(D)} \leq c\, h^s \ell_D^{1-s} \|u - u_h\|_{V_\sharp}. \tag{38.19}$$

Proof. Apply Lemma 36.14 and use exact adjoint consistency; see Exercise 38.3. □

Remark 38.13 (Variations on symmetry). Let us set

$$\begin{aligned} a_h(v_h, w_h) := &\int_D \nabla_h v_h \cdot \nabla_h w_h \,\mathrm{d}x - \sum_{F \in \mathcal{F}_h} \int_F \{\nabla_h v_h\} \cdot \boldsymbol{n}_F [\![w_h]\!] \,\mathrm{d}s \\ &- \theta \sum_{F \in \mathcal{F}_h} \int_F \{\nabla_h w_h\} \cdot \boldsymbol{n}_F [\![v_h]\!] \,\mathrm{d}s + \sum_{F \in \mathcal{F}_h} \varpi(h_F) \int_F [\![v_h]\!] [\![w_h]\!] \,\mathrm{d}s, \end{aligned} \tag{38.20}$$

where θ is a real number ($\theta := 1$ corresponds to the SIP formulation). The choice $\theta := -1$ gives the method usually called nonsymmetric interior penalty (NIP). This choice is interesting since it simplifies the analysis of the coercivity in that the consistency term cancels with the added nonsymmetric term. The original idea can be traced back to the method in Oden et al. [318], where the nonsymmetric method is introduced without the penalty term. The convergence analysis when the penalty term is included can be found in Rivière et al. [335, 336], where it is shown that coercivity only requires $\varpi_0 > 0$; see also Larson and Niklasson [274] for the inf-sup stability analysis. The incomplete interior penalty (IIP) method corresponds to the choice $\theta := 0$. Similarly to SIP, a minimal threshold on the penalty parameter ϖ_0 is required for the coercivity; see Dawson et al. [157]. Whenever $\theta \neq 1$, the analysis of the L^2-error estimate proceeds as in §37.3.2 (accounting for an adjoint consistency error), and one only obtains $\|u - u_h\|_{L^2(D)} \leq ch^{\frac{1}{2}} \|u - u_h\|_{V_\sharp}$ even if full elliptic regularity holds true ($s = 1$). □

Remark 38.14 (L^∞-estimates). Pointwise dG error estimates are found in Kanschat and Rannacher [264], Chen and Chen [117], Guzmán [233]. □

38.4 Discrete gradient and fluxes

In this section, we introduce the notion of discrete gradient and use it to derive an alternative viewpoint on the SIP bilinear form. One interesting outcome is a reformulation of the discrete problem (38.5) in terms of local problems with numerical fluxes.

38.4.1 Liftings

Loosely speaking the discrete gradient consists of the broken gradient plus a correction associated with the jumps. This correction is formulated in terms of local liftings introduced in Bassi and Rebay [46] and analyzed in Brezzi et al. [93] (see also Perugia and Schötzau [323] for the *hp*-analysis). Let $F \in \mathcal{F}_h$ and an integer $l \geq 0$. Consider the local *lifting operator* $\boldsymbol{\mathcal{L}}_F^l : L^2(F) \to \boldsymbol{P}_l^{\mathrm{b}}(\mathcal{T}_h) := P_l^{\mathrm{b}}(\mathcal{T}_h;\mathbb{R}^d)$ s.t. for all $\varphi \in L^2(F)$, the discrete function $\boldsymbol{\mathcal{L}}_F^l(\varphi)$ is defined as

$$\int_D \boldsymbol{\mathcal{L}}_F^l(\varphi)\cdot\boldsymbol{\tau}_h \,\mathrm{d}x := \int_F \{\boldsymbol{\tau}_h\}\cdot\boldsymbol{n}_F \varphi \,\mathrm{d}s, \qquad \forall \boldsymbol{\tau}_h \in \boldsymbol{P}_l^{\mathrm{b}}(\mathcal{T}_h). \tag{38.21}$$

By localizing the support of $\boldsymbol{\tau}_h$ to a single mesh cell, we infer that $\boldsymbol{\mathcal{L}}_F^l(\varphi)$ is collinear to $\boldsymbol{n}_F$ and is supported in the set $D_F := \mathrm{int}(\bigcup_{K\in\mathcal{T}_F} K)$. In practice, the Cartesian components of the polynomial function $\boldsymbol{\mathcal{L}}_F^l(\varphi)$ can be computed in each $K \in \mathcal{T}_F$ by inverting the local mass matrix with entries $\mathcal{M}_{K,ij} := \int_K \theta_{K,i}\theta_{K,j}\,\mathrm{d}x$, where the functions $\theta_{K,i}$ are the local shape functions in K.

Consider now a function $v \in H^1(\mathcal{T}_h)$. We define the global lifting of the jumps of v as follows:

$$\boldsymbol{\mathcal{L}}_h^l([\![v]\!]) := \sum_{F\in\mathcal{F}_h} \boldsymbol{\mathcal{L}}_F^l([\![v]\!]).$$

This makes sense since $[\![v]\!]_F \in L^2(F)$ for all $F \in \mathcal{F}_h$. A consequence of $\mathrm{supp}(\boldsymbol{\mathcal{L}}_F^l([\![v]\!])) = D_F$ is that $\boldsymbol{\mathcal{L}}_h^l([\![v]\!])_{|K} := \sum_{F\in\mathcal{F}_K} \boldsymbol{\mathcal{L}}_F^l([\![v]\!])$ for all $K \in \mathcal{T}_h$, i.e., only the jumps across the faces of K contribute to the restriction to K of the global lifting $\boldsymbol{\mathcal{L}}_h^l([\![v]\!])$.

Lemma 38.15 (Stability). *The following holds true for all $l \geq 0$:*

$$\|\boldsymbol{\mathcal{L}}_F^l(\varphi)\|_{\boldsymbol{L}^2(D_F)} \leq c_{\mathrm{dt}} h_F^{-\frac{1}{2}} \|\varphi\|_{L^2(F)}, \qquad \forall \varphi \in L^2(F),\ \forall F \in \mathcal{F}_h, \tag{38.22a}$$

$$\|\boldsymbol{\mathcal{L}}_h^l([\![v]\!])\|_{\boldsymbol{L}^2(D)} \leq n_\partial^{\frac{1}{2}} c_{\mathrm{dt}} |v|_{\mathrm{J}}, \qquad \forall v \in H^1(\mathcal{T}_h), \tag{38.22b}$$

where c_{dt} is the constant from the discrete trace inequality (38.10)*.*

Proof. The proof of (38.22a) is proposed in Exercise 38.4. To prove (38.22b), we use the Cauchy–Schwarz inequality and the definition of n_∂ to infer that

$$\|\boldsymbol{\mathcal{L}}_h^l([\![v]\!])\|_{\boldsymbol{L}^2(K)}^2 = \int_K \left|\sum_{F\in\mathcal{F}_K} \boldsymbol{\mathcal{L}}_F^l([\![v]\!])\right|^2 \mathrm{d}x \leq n_\partial \sum_{F\in\mathcal{F}_K} \|\boldsymbol{\mathcal{L}}_F^l([\![v]\!])\|_{\boldsymbol{L}^2(K)}^2,$$

for all $K \in \mathcal{T}_h$. Summing over the mesh cells, recalling that the support of $\boldsymbol{\mathcal{L}}_F^l([\![v]\!])$ is D_F, and using (38.22a) yields (38.22b). □

Definition 38.16 (Discrete gradient). *Let $l \geq 0$. The* discrete gradient *operator $\mathfrak{G}_h^l : H^1(\mathcal{T}_h) \to \boldsymbol{L}^2(D)$ is defined as follows:*

$$\mathfrak{G}_h^l(v) := \nabla_h v - \boldsymbol{\mathcal{L}}_h^l([\![v]\!]), \qquad \forall v \in H^1(\mathcal{T}_h). \tag{38.23}$$

We can now use Definition 38.16 to derive alternative expressions for the SIP bilinear form a_h defined in (38.4). Recalling that $V_h := P_k^{\mathrm{b}}(\mathcal{T}_h)$, $k \geq 1$, we choose the polynomial degree of the liftings such that $l \in \{k-1, k\}$. Since $\nabla_h v_h, \nabla_h w_h \in \mathbb{P}_{k-1}^{\mathrm{b}}(\mathcal{T}_h) \subset \mathbb{P}_l^{\mathrm{b}}(\mathcal{T}_h)$ for all $v_h, w_h \in V_h$, we infer that

$$\begin{aligned}\int_D \nabla_h v_h \cdot \nabla_h w_h \,\mathrm{d}x - \sum_{F\in\mathcal{F}_h} \int_F \Big(\{\nabla_h v_h\}\cdot\boldsymbol{n}_F [\![w_h]\!] + [\![v_h]\!]\{\nabla_h w_h\}\cdot\boldsymbol{n}_F\Big)\,\mathrm{d}s \\ = \int_D \mathfrak{G}_h^l(v_h)\cdot\mathfrak{G}_h^l(w_h)\,\mathrm{d}x - \int_D \boldsymbol{\mathcal{L}}_h^l([\![v_h]\!])\cdot\boldsymbol{\mathcal{L}}_h^l([\![w_h]\!])\,\mathrm{d}x.\end{aligned} \tag{38.24}$$

Recalling the expression (38.4) of a_h, we obtain

$$a_h(v_h, w_h) := \int_D \mathfrak{G}_h^l(v_h)\cdot\mathfrak{G}_h^l(w_h)\,\mathrm{d}x + \tilde{s}_h(v_h, w_h), \tag{38.25}$$

with $\tilde{s}_h(v_h, w_h) := \sum_{F\in\mathcal{F}_h} \varpi(h_F) \int_F [\![v_h]\!][\![w_h]\!]\,\mathrm{d}s - \int_D \boldsymbol{\mathcal{L}}_h^l([\![v_h]\!])\cdot\boldsymbol{\mathcal{L}}_h^l([\![w_h]\!])\,\mathrm{d}x$. The estimate (38.22b) from Lemma 38.15 implies that

$$a_h(v_h, v_h) \geq \|\mathfrak{G}_h^l(v_h)\|_{L^2(D)}^2 + (\varpi_0 - n_\partial c_{\mathrm{dt}}^2)|v_h|_{\mathrm{J}}^2, \tag{38.26}$$

for all $v_h \in V_h$, showing again the relevance of the condition $\varpi_0 > n_\partial c_{\mathrm{dt}}^2$ for coercivity (see Lemma 38.6).

Remark 38.17 (Alternative penalty strategy). It is possible to penalize the liftings of the jumps instead of the jumps, leading to the following modification of the SIP bilinear form:

$$\begin{aligned}\check{a}_h(v_h, w_h) := \int_D \nabla_h v_h\cdot\nabla_h w_h\,\mathrm{d}x - \sum_{F\in\mathcal{F}_h}\int_F \{\nabla_h v_h\}\cdot\boldsymbol{n}_F[\![w_h]\!]\,\mathrm{d}s \\ - \sum_{F\in\mathcal{F}_h}\int_F [\![v_h]\!]\{\nabla_h w_h\}\cdot\boldsymbol{n}_F\,\mathrm{d}s + \sum_{F\in\mathcal{F}_h}\varpi_0\int_D \boldsymbol{\mathcal{L}}_F^l([\![v_h]\!])\cdot\boldsymbol{\mathcal{L}}_F^l([\![w_h]\!])\,\mathrm{d}x.\end{aligned}$$

The main advantage of this formulation is that coercivity holds true as soon as $\varpi_0 > n_\partial$, thereby avoiding the constant c_{dt} from (38.10). However, the discretization stencil is larger since the dofs in two cells $K, K' \in \mathcal{T}_h$ are coupled if there is $K'' \in \mathcal{T}_h$ s.t. $\partial K \cap \partial K'' \in \mathcal{F}_h^\circ$ and $\partial K' \cap \partial K'' \in \mathcal{F}_h^\circ$ (for the usual penalty strategy the coupling condition is $\partial K \cap \partial K' \in \mathcal{F}_h^\circ$). □

Remark 38.18 (Choosing l). The computation of the discrete gradient can be done with any $l \geq k-1$. The minimal choice is $l = k-1$, but choosing $l = k$ may be more interesting from the implementation point of view since it does not require the user to construct the finite element space $P_{k-1}^{\mathrm{b}}(\mathcal{T}_h)$. □

Remark 38.19 (Literature). The discrete gradient is an important notion in the design and analysis of dG methods for nonlinear problems. We refer the reader to Ten Eyck and Lew [364] for nonlinear mechanics, to Burman and Ern [100], Buffa and Ortner [95] for Leray–Lions operators, and to Di Pietro and Ern [164] for the incompressible Navier–Stokes equations. Moreover, an important stability result established in John et al. [260] is that if $l = k+1$, then there is c s.t. $\|v_h\|_{V_h} \le c\|\mathfrak{G}_h^l(v_h)\|_{L^2(D)}$ for all $v_h \in V_h$ and all $h \in \mathcal{H}$. Since the proof of this result invokes Raviart–Thomas functions, simplicial meshes are required, but hanging nodes are still allowed under some assumptions. An interesting consequence of this stability result is that for $l = k+1$, replacing a_h defined in (38.4) by $\tilde{a}_h(v_h, w_h) := \int_D \mathfrak{G}_h^l(v_h){\cdot}\mathfrak{G}_h^l(w_h)\,\mathrm{d}x$ gives a stable and optimally convergent dG discretization without any penalty parameters. Notice that $\tilde{a}_h$ does not deliver exact consistency because liftings are discrete objects; see Exercise 38.6. For the same reason, the bilinear form a_h defined in (38.4) coincides with the right-hand side of (38.25) on $V_h{\times}V_h$, but the two sides of the equality produce different results on $V_\sharp{\times}V_h$. □

38.4.2 Local formulation with fluxes

Let $K \in \mathcal{T}_h$ and consider a smooth function $\xi \in C^1(K)$. Integration by parts shows that the solution to (38.1), if it is smooth enough, satisfies

$$\int_K f\xi\,\mathrm{d}x = \int_K -(\Delta u)\xi\,\mathrm{d}x = \int_K \nabla u{\cdot}\nabla\xi\,\mathrm{d}x - \int_{\partial K} (\nabla u{\cdot}\boldsymbol{n}_K)\xi\,\mathrm{d}s.$$

Splitting the boundary integral over the faces $F \in \mathcal{F}_K$ yields

$$\int_K \nabla u{\cdot}\nabla\xi\,\mathrm{d}x + \sum_{F\in\mathcal{F}_K} \epsilon_{K,F} \int_F \Phi_F(u)\xi\,\mathrm{d}s = \int_K f\xi\,\mathrm{d}x, \tag{38.27}$$

where $\Phi_F(u) := -\nabla u{\cdot}\boldsymbol{n}_F$, $\epsilon_{K,F} = \boldsymbol{n}_K{\cdot}\boldsymbol{n}_F$, and $\boldsymbol{n}_K$ is the outward normal to K ($\boldsymbol{n}_K{\cdot}\boldsymbol{n}_F = \pm 1$, for all $F \in \mathcal{F}_K$, depending on the orientation of F). The function Φ_F is called exact flux since (38.27) expresses a balance between the source term in K, the diffusion processes in K, and the fluxes across all the faces in $\mathcal{F}_K$. An interesting feature of dG methods is that one obtains a discrete counterpart of (38.27) when the test function is supported only in the mesh cell K.

Lemma 38.20 (Local formulation). *Let u_h solve* (38.5). *Let the* numerical flux *on a mesh face $F \in \mathcal{F}_h$ be defined by*

$$\widehat{\Phi}_F(u_h) := -\{\nabla_h u_h\}{\cdot}\boldsymbol{n}_F + \varpi(h_F)[\![u_h]\!]. \tag{38.28}$$

Then the following holds true for all $q \in P_K$ and all $K \in \mathcal{T}_h$:

$$\int_K \mathfrak{G}_h^l(u_h){\cdot}\nabla q\,\mathrm{d}x + \sum_{F\in\mathcal{F}_K} \epsilon_{K,F} \int_F \widehat{\Phi}_F(u_h) q\,\mathrm{d}s = \int_K f q\,\mathrm{d}x. \tag{38.29}$$

Proof. Let $\mathbb{1}_K$ be the indicator function of K and let q be arbitrary in P_K. Using the test function $w_h := q\mathbb{1}_K$ in (38.5), we obtain $a_h(u_h, q\mathbb{1}_K) = \int_K fq\,\mathrm{d}x$. Then, (38.29) follows by invoking (38.24) and by making use of the identity $[\![q\mathbb{1}_K]\!]_F = \epsilon_{K,F}q$ if $F \in \mathcal{F}_K$ and $[\![q\mathbb{1}_K]\!]_F = 0$ otherwise. □

The numerical flux $\widehat{\Phi}_F(u_h)$ consists of a centered flux, $-\{\nabla_h u_h\}{\cdot}\boldsymbol{n}_F$, originating from the consistency term, plus a stabilization term, $\varpi(h_F)[\![u_h]\!]$, originating from the penalty term. A unified presentation of dG methods for the Poisson equation based on fluxes can be found in Arnold et al. [21].

38.4.3 Equilibrated $\boldsymbol{H}(\mathrm{div})$ flux recovery

The vector-valued function $\boldsymbol{\sigma} := -\nabla u$ is called *diffusive flux*. This function is important in many applications where the underlying PDE expresses a conservation principle in the form $\nabla{\cdot}\boldsymbol{\sigma} = f$ in D. Since $\boldsymbol{\sigma} \in \boldsymbol{H}(\mathrm{div}; D)$, Theorem 18.10 implies that $[\![\boldsymbol{\sigma}]\!]{\cdot}\boldsymbol{n}_F = 0$ for all $F \in \mathcal{F}_h^\circ$ (possibly in a weak sense if $\boldsymbol{\sigma}$ is not smooth enough). From a physical viewpoint, this zero-jump condition expresses the fact that what flows out of a mesh cell through one of its faces flows into the neighboring mesh cell.

The local formulation (38.29) provides a natural way of reconstructing a discrete diffusive flux $\boldsymbol{\sigma}_h$ in $\boldsymbol{H}(\mathrm{div}; D)$ that closely approximates $\boldsymbol{\sigma}$. Assuming that the mesh is matching and simplicial, we now describe a way to reconstruct $\boldsymbol{\sigma}_h$ in the Raviart–Thomas finite element space $\boldsymbol{P}_l^{\mathrm{d}}(\mathcal{T}_h)$ defined in (19.16) with $l \in \{k-1, k\}$. The reconstruction is explicit and amounts to prescribing the global degrees of freedom of $\boldsymbol{\sigma}_h$ in $\boldsymbol{P}_h^{\mathrm{d}}(\mathcal{T}_h)$; see Ern et al. [193], Kim [268].

Lemma 38.21 (Flux recovery). *Let $\boldsymbol{\sigma}_h \in \boldsymbol{P}_h^{\mathrm{d}}(\mathcal{T}_h)$ be such that*

$$\int_F (\boldsymbol{\sigma}_h{\cdot}\boldsymbol{n}_F)(q \circ \boldsymbol{T}_F^{-1})\,\mathrm{d}s = \int_F \widehat{\Phi}_F(u_h)(q \circ \boldsymbol{T}_F^{-1})\,\mathrm{d}s, \qquad \forall F \in \mathcal{F}_h, \forall q \in \mathbb{P}_{l,d-1},$$

$$\text{and if } l \geq 1, \quad \int_K \boldsymbol{\sigma}_h{\cdot}\boldsymbol{r}\,\mathrm{d}x = -\int_K \mathfrak{G}_h^l(u_h){\cdot}\boldsymbol{r}\,\mathrm{d}x, \qquad \forall K \in \mathcal{T}_h, \forall \boldsymbol{r} \in \mathbb{P}_{l-1,d},$$

where $\boldsymbol{T}_F$ is an affine bijective mapping from the unit simplex of $\mathbb{R}^{d-1}$ to F for all $F \in \mathcal{F}_h$. Let $\mathcal{I}_h^{\mathrm{b}}$ denote the L^2-orthogonal projection onto $P_l^{\mathrm{b}}(\mathcal{T}_h)$. Then we have

$$\nabla{\cdot}\boldsymbol{\sigma}_h = \mathcal{I}_h^{\mathrm{b}}(f). \tag{38.30}$$

Proof. Integrating by parts on a cell $K \in \mathcal{T}_h$ and using (38.29), we infer that

$$\begin{aligned}\int_K (\nabla{\cdot}\boldsymbol{\sigma}_h)q\,\mathrm{d}x &= -\int_K \boldsymbol{\sigma}_h{\cdot}\nabla q\,\mathrm{d}x + \sum_{F\in\mathcal{F}_K}\int_F (\boldsymbol{\sigma}_h{\cdot}\boldsymbol{n}_K)(q_{|F}{\circ}\boldsymbol{T}_F){\circ}\boldsymbol{T}_F^{-1}\,\mathrm{d}s \\ &= \int_K fq\,\mathrm{d}x,\end{aligned}$$

for all $q \in \mathbb{P}_{l,d}$ since $\nabla q \in \mathbf{P}_{l-1,d}$ and $q_{|F} \circ \boldsymbol{T}_F \in \mathbb{P}_{l,d-1}$ (see Lemma 7.10). Then (38.30) is a consequence of $\nabla{\cdot}\boldsymbol{\sigma}_h \in P_l^{\mathrm{b}}(\mathcal{T}_h)$. □

Equation (38.30) shows that $\nabla{\cdot}\boldsymbol{\sigma}_h$ optimally approximates the source term. By proceeding as in Di Pietro and Ern [165, §5.5.3], it is possible to show that $\|\boldsymbol{\sigma} - \boldsymbol{\sigma}_h\|_{\boldsymbol{L}^2(D)} \le c(\|u - u_h\|_{V_\sharp} + h\|f - \mathcal{I}_h^{\mathrm{b}}(f)\|_{L^2(D)})$.

Exercises

Exercise 38.1 (Elementary dG identities). (i) Let $F := \partial K_l \cap \partial K_r \in \mathcal{F}_h^\circ$. Prove that $2\{\boldsymbol{\sigma}{\cdot}\boldsymbol{n}_K q\} = (\{\boldsymbol{\sigma}\}[\![q]\!] + [\![\boldsymbol{\sigma}]\!]\{q\}){\cdot}\boldsymbol{n}_F$. (ii) Let $\theta_l, \theta_r \in [0,1]$ such that $\theta_l + \theta_r = 1$. Let $[\![a]\!]_\theta := 2(\theta_r a_l - \theta_l a_r)$ and $\{a\}_\theta := \theta_l a_l + \theta_r a_r$. Show that $\{ab\} = \{a\}\{b\}_\theta + \frac{1}{4}[\![a]\!]_\theta [\![b]\!]$.

Exercise 38.2 (Boundary conditions). (i) Assume that u solves the Poisson problem (38.1) with the non-homogeneous Dirichlet condition $u = g$ on ∂D. Let a_h^θ be defined in (38.20). Devise $\ell_h^{\theta,\mathrm{nD}}$ so that exact consistency holds for the following formulation: Find $u_h \in V_h$ such that $a_h^\theta(u_h, w_h) = \ell_h^{\theta,\mathrm{nD}}(w_h)$ for all $w_h \in V_h$. (ii) Assume that u solves the Poisson problem with the Robin condition $\gamma u + \boldsymbol{n}{\cdot}\nabla u = g$ on ∂D. Let ℓ_h^{Rb} be defined in (38.13b). Devise a_h^{Rb} so that exact consistency holds for the following formulation: Find $u_h \in V_h$ such that $a_h^{\theta,\mathrm{Rb}}(u_h, w_h) = \ell_h^{\mathrm{Rb}}(w_h)$ for all $w_h \in V_h$.

Exercise 38.3 (L^2-estimate). Prove Theorem 38.12. (*Hint*: see the proof of Theorem 37.8.)

Exercise 38.4 (Local lifting). Prove (38.22a). (*Hint*: use (38.10).)

Exercise 38.5 (Local formulation). Write the local formulation of the OBB, NIP, and IIP dG methods discussed in Remark 38.13.

Exercise 38.6 (Extending (38.25)). Let $\tilde{a}_h$ (resp., a_h) be defined by extending (38.25) (resp., (38.4)) to $V_\sharp \times V_h$. Show that $\tilde{a}_h(v, w_h) = a_h(v, w_h) + \sum_{F \in \mathcal{F}_h} \int_F \{\nabla_h v - \mathcal{I}_h^{\mathrm{b}}(\nabla_h v)\}{\cdot}\boldsymbol{n}_F [\![w_h]\!]\,\mathrm{d}s$ for all $(v, w_h) \in V_\sharp \times V_h$.

Exercise 38.7 (Discrete gradient). Let $(v_h)_{h\in\mathcal{H}}$ be a sequence in $(V_h)_{h\in\mathcal{H}}$ (meaning that $v_h \in V_h$ for all $h \in \mathcal{H}$). Assume that there is C s.t. $\|v_h\|_{V_h} \le C$ for all $h \in \mathcal{H}$. One can show that there is $v \in L^2(D)$ such that, up to a subsequence, $v_h \to v$ in $L^2(D)$ as $h \to 0$; see [165, Thm. 5.6]. (i) Show that, up to a subsequence, $\mathfrak{G}_h^l(v_h)$ weakly converges to some $\boldsymbol{G}$ in $\boldsymbol{L}^2(D)$ as $h \to 0$. (*Hint*: bound $\|\mathfrak{G}_h^l(v_h)\|_{\boldsymbol{L}^2(D)}$.) (ii) Show that $\boldsymbol{G} = \nabla v$ and that $v \in H_0^1(D)$. (*Hint*: extend functions by zero outside D and prove first that $\int_{\mathbb{R}^d} \mathfrak{G}_h^l(v_h){\cdot}\boldsymbol{\Phi}\,\mathrm{d}x = -\int_{\mathbb{R}^d} v_h \nabla{\cdot}\boldsymbol{\Phi}\,\mathrm{d}x + \sum_{F\in\mathcal{F}_h} \int_F \{\boldsymbol{\Phi} - \mathcal{I}_h^{\mathrm{b}}\boldsymbol{\Phi}\}{\cdot}\boldsymbol{n}_F [\![v_h]\!]\,\mathrm{d}s$ for all $\boldsymbol{\Phi} \in \boldsymbol{C}_0^\infty(\mathbb{R}^d)$.)